LOCAL HEATING
DURING ELECTRON BEAM PATTERNING OF LITHOGRAPHY MASKS

by
Lars Vicum

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
(Mechanical Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

1998

Abstract

The semiconductor industry is constantly seeking to improve the performance of integrated circuits by drastically increasing the density and reducing the pattern feature size. Accurate patterning of advanced lithography masks is a key issue in the production of integrated circuits with sub-0.13 μm feature size. Predicting and subsequently correcting for the errors produced in writing the pattern is essential.

One source of pattern error is the heating of the lithography mask during the electron beam patterning process. Mask heating during writing causes pattern errors through the resist stress relief, local over- or under-development of the resist due to the temperature dependent resist sensitivity, and by thermal distortions. These pattern errors depend upon the transient temperature distribution in the mask during the electron-beam-writing process.

In this thesis local mask temperature profiles are predicted for various writing conditions. The local mask heating of an optical reticle during writing of a single electron beam flash is analyzed. The contribution of multiple electron beam flashes on an X-ray mask to global mask heating during writing is determined by calculating the transient mask temperature profile during the patterning process using a finite element software package.

Two levels of model are compared. A very detailed small-scale model of pattern writing is developed with the element size that of a pattern shape. An averaging large-scale model in which the electron beam energy is distributed over many elements is also developed. The averaging technique is often used in thermal analyses to reduce computation time. The average temperature rise and local maximum of the two methods are compared. It is found that the averaging technique predicts the average temperature rise accurately, but significantly under-predicts the maximum local temperatures.

A procedure for developing thermal models with high accuracy, minimal number of elements, and reasonable calculation time is described. Design rules for finite element models used for the simulation of local mask heating are presented.

Acknowledgements

I would like to thank William A. Beckman and John W. Mitchell, my advisors, for their constant guidance, support and great attitude towards advising and dealing with students.

I am very grateful that I had the opportunity to complete my master degree here at the Solar Energy Laboratory. My time in the lab has been very intense and enjoyable. I am going to miss working in the lab, the seminars and the discussions with my lab mates. Thank you Bill, John, Sandy and all the students to make this laboratory not solely a place to work but a great place to be.

I would like to thank all those people who contributed to this thesis. Thanks to Professors Edward G. Lovell and Roxann L. Engelstad for their support. Thanks to Sergey Babin, who provided me with Monte Carlo simulation data. Thanks to Nate Blair and David Bradley for helping me with computer problems. Also thanks to SEMATECH for sponsoring my project work.

My stay in Madison was made possible by the German Academic Exchange Service (DAAD) and the Institut für Thermodynamik at the University of Hannover. I am grateful to

Stephan Kabelac and Dirk Labuhn for putting their personal effort in this program to make this great study abroad experience possible.

Finally very special thanks to my parents, my sisters Iris and Inga-Viola, and my grandfather for their financial and moral support.

Table of Contents

Chapter Three

Chapter Four

Chapter Five

Chapter Six

List of Figures

List of Tables

Chapter One

Introduction

1.1 History and Future of the Semiconductor Industry

The key invention for the beginning of the microelectronics era was the invention of the integrated monolithic circuit by Jack Kilby and Robert Noyce in 1958. Until 1958 electronic devices were produced with single transistors. The invention of the integrated monolithic circuit made internal interconnections between many solid state devices possible, which resulted in a device density increase of many thousand times. After this key invention the semiconductor market became the fastest developing segment of the electronic industry. The worldwide semiconductor production doubled about every two and a half years and reached a value of 208 billion dollars in 1997 which corresponds to 21.7 % of the total worldwide electronic equipment sales [1.1]. This very fast economic development was accompanied by a continuous technological revolution. The decrease in the minimum feature size and the increase in the chip area, the chip frequency and device density made the rapid increase of integrated circuit performance possible. From 1960 to 1996 the speed of integrated circuits, measured in floating point operations per second, increased every two years by the factor of a hundred [1.1].

Figure 1.1 illustrates the performance trend for the dynamic random access memory (DRAM). The reduction in feature size over time follows an exponential trend that will probably continue until a minimum feature size of 0.1 μm reached. Figure 1.1 also shows the devices corresponding to a specific integration level. Although the number of active devices per chip has continuously increased with decreasing feature size and increasing chip area the costs per chip stayed almost constant. This results in a very fast decrease in costs per bit from 1 cent per bit in 1970 to 10^{-4} cent per bit in 1996.

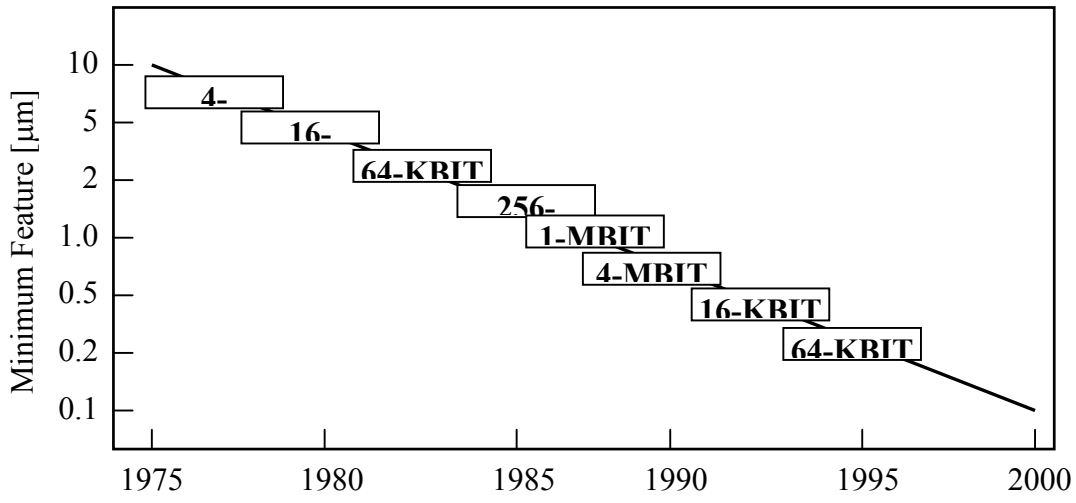


Figure 1.1. Minimum feature size on a semiconductor DRAM device as a function of the year the devices were first commercially available [1.2].

For the development of the semiconductor industry, lithography has been the driving technology. Since the invention of the IC the growth of the industry has been the direct result of improved lithography techniques. Lithography is also one of the major economic factors in the IC production process with currently 35% of the chip manufacturing costs [1.3]. The recent development of lithography techniques as well as the predicted future of

lithography for the next decade is presented in Fig. 1.2. So far optical lithography has been the mainstream technology for the industry. Currently optical lithography with dense ultraviolet light (DUV) is used for high volume production and leading-edge production with 0.25 μm minimum feature size.

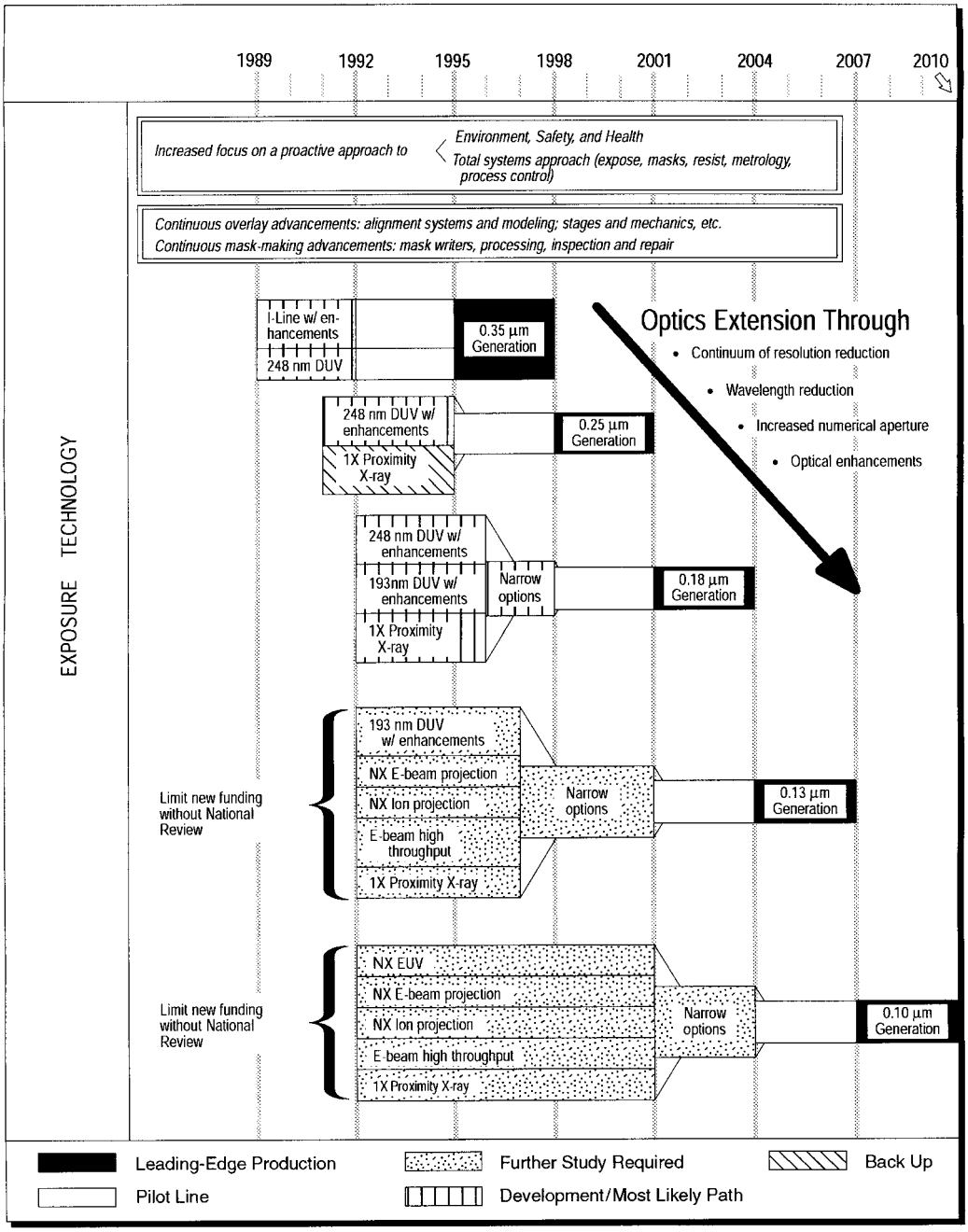


Figure 1.2. Exposure Technology Roadmap by the Semiconductor Industry Association [1.4].

The most important limit for optical lithography, with smaller minimum feature sizes, will be the depth of focus. For minimum feature sizes of 0.10 μm and smaller it will become necessary to use alternative lithography techniques as 1× proximity X-ray lithography, ion-beam lithography, EUV lithography or electron-beam projection lithography to achieve the necessary overlay accuracy and critical dimensions improvement [1.3]. Although it is still uncertain which technology will substitute optical lithography, X-ray lithography was often considered as the most likely option [1.5, 1.6].

Many areas of the lithography field have to be investigated further to obtain the production precision necessary for minimum feature sizes of 0.10 μm and smaller. This work will concentrate on investigating the sources of mask errors during the lithography mask electron-beam writing process, which is one step in the mask fabrication. The mask fabrication process is described in Chapter two for an X-ray mask. Chapter three characterizes in detail the electron-beam writing process. The sources of mask fabrication errors, caused by the electron-beam writing process, are pointed out. Mask heating caused by the electron-beam is analyzed in detail. In Chapter four an approach for the quantitative analysis of local mask heating is explained. The local mask heating when writing a spot on an optical mask is investigated in Chapter five. How the local mask heating contributes to global mask heating is examined in Chapter six. The mask heating during patterning of small mask regions is investigated and the accuracy of averaging techniques, used to approximate the contribution

of local to global mask heating, is analyzed. Chapter seven contains conclusions and the proposal for future research.

1.2 References

[1.1] ..Integrated Circuit Engineering Corporation, “Worldwide IC Industry Economic Update and Forecast,” pp. 1-5, 1996.

[1.2] L. F. Thompson, C. G. Willson, and M. J. Bowden, “Introduction to Microlithography,” ACS Professional Reference Book, 1994.

[1.3] Semiconductor Industry Association, "National Technology Roadmap for Semiconductors,” 1997.

[1.4] Semiconductor Industry Association, "National Technology Roadmap for Semiconductors,” 1994.

[1.5] W. Arden, "Optical versus X-ray Lithography for Future Device Fabrication,” *Microelectronic Engineering*, 13, 1991.

[1.6] J. P. Silverman, "X-ray Lithography: Status, Challenges, and Outlook for 0.13 μm,” *J. Vac. Sci. Technol. B*, 15, 1997.

Chapter Two

The X-ray Lithography Process

2.1 Introduction

Lithography is a process in which an image is transferred onto a receptive surface by shining light through the non-opaque portions of a pattern on a mask. In the X-ray lithography process a mask with a thin membrane and a pattern of absorber material on it is used to project the layout of an integrated circuit onto a silicon wafer that is covered with a photoresist material. Figure 2.1 shows a schematic of the exposure process with X-rays. The pattern on the mask is transferred repetitively onto the device wafer at different locations by moving the wafer relative to the mask. After the complete wafer surface is exposed the unexposed resist is washed away and the unprotected wafer surface is etched chemically. Then the hardened resist is chemically dissolved leaving the finished patterned layer on the wafer. The wafer is covered again with a photoresist and a new mask is aligned to the device wafer and another exposure process follows. This procedure is repeated for each layer of the integrated circuit, about 15 or more times per wafer [2.1].

The collimated X-rays used for X-ray lithography typically have a wavelength in the 7-12 Å range [2.2]. The source that is used in practice for these soft X-rays is an electron storage ring, which provides the light in the form of synchrotron radiation. Because there are no

suitable focusing lenses available for X-rays, the pattern on the mask has to be of the same size as the image. The fact that X-ray lithography is a 1:1 process creates the major difficulty for this technology. Because the X-ray mask has to be "1×", it is much more complicated to manufacture, and to inspect and repair the mask compared to optical masks. Optical masks use 4× optics so the size of the circuit pattern to be printed is obtained using optical lenses for demagnification.

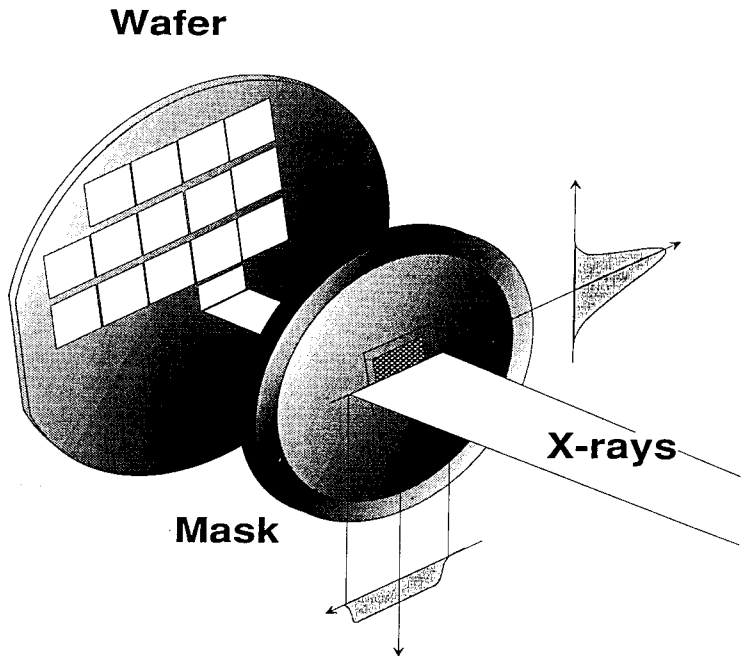


Figure 2.1. X-ray wafer exposure [2.3].

Although the mask making of X-ray lithography masks is more difficult for the current generation of integrated circuits, X-ray lithography could be the superior technology for future generations of integrated circuit generations beyond a minimum feature size of 0.13 μm. For the new generations, the use of optical lithography will be limited by the obtainable resolution due to a limitation in the depth of focus [2.1].

2.2. X-ray Lithography Masks

A typical X-ray lithography mask consists of a multi-layer structure of about 625 μm thickness with an approximately 2 μm thick membrane in the middle. The membrane is made of a material with a low atomic number for high transmission of X-rays. The membrane has to be transparent enough to have low X-ray exposure time and has to be thick enough to be mechanical stable. The absorber material that covers the top surface of the membrane at locations corresponding to the design of the integrated circuit is made of a material with high atomic number in order to have a large electron density to absorb the X-rays. Possible absorber materials are gold, tantalum, tungsten and platinum [2.4]. To make the membrane more resistant to distortions caused by the handling and the exposure it is bonded to a support ring. The geometry of an ARPA-NIST X-ray mask is shown in Fig. 2.2 [2.5].

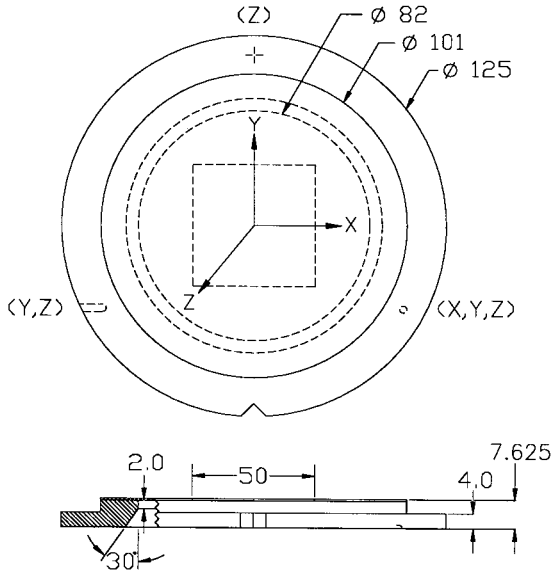


Figure 2.2. ARPA-NIST National X-ray Mask Standard format.
All dimensions are in millimeters [2.5].

Although the details of the mask fabrication process vary for each manufacturer the fundamental process steps are similar. Figure 2.3 shows a typical X-ray mask fabrication process. The fabrication starts with a silicon wafer of 625 μm. This wafer is covered with a

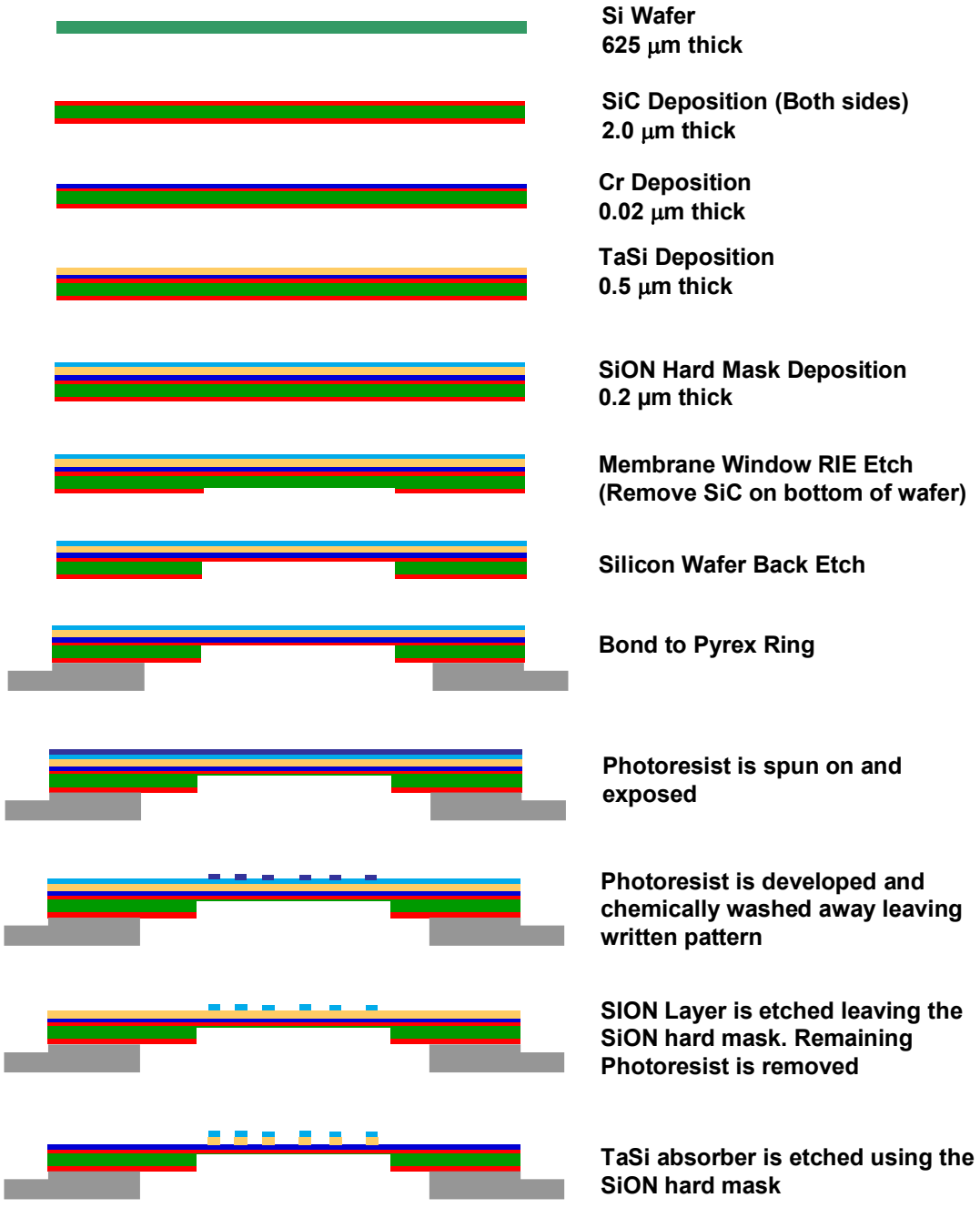


Figure 2.3. Process steps of typical X-ray lithography mask fabrication [2.6].

2 μm layer of silicon carbide. Silicon carbide is chosen as a membrane material, because, compared to other materials with low atomic number, it has a higher stiffness and radiation hardness [2.4]. In the next process step a chrome etch stop layer is sputtered onto the wafer. After this a 0.5 μm absorber layer made of tantalum silicate is deposited. On top of this layer a hard mask of SiON is deposited. The next process step is the creation of the membrane. The silicon carbide layer at the backside of the wafer is removed by reactive ion etching and the silicon wafer is etched away by wet etching to create the membrane in the middle of the mask. The last process step in the production of the mask blank is the bonding of the Pyrex™ support ring to the wafer. Once a layer of photoresist is spun on the top surface of the mask blank the integrated circuit pattern is written on the resist with an e-beam machine. After the developing of the photoresist the undeveloped material is removed by chemical washing. With etching processes the pattern is transferred from the photoresist to the hard mask and then to the absorber.

2.3. Sources of X-ray Mask Pattern Displacements

A significant portion of the pattern placement errors are due to mask distortions. The sources for distortion of X-ray masks can be divided into groups: the mask distortions that occur during the fabrication process of the mask and the distortions that are induced during the use of the mask for the exposure process. One reason for mask distortions in the fabrication process is the stress in the multi-layer structure of mask. The deposition processes create

significant stress in different material layers. Distortions can be caused by different stresses in layers stacked above each other, stress gradients across the mask area or by the partial removal of stressed layers. The amount of distortion depends on the non-uniformity of the pattern density. Another reason for pattern displacements in the mask production process is the accuracy of the e-beam writing machine. Improvement of the e-beam writing tools is considered to be one of the most critical issues in X-ray lithography [2.2]. In addition to pattern displacements due to the inaccuracy of the e-beam writing tool, displacements in the e-beam writing process are also caused by mask distortions connected with this process. These distortions can be divided in two groups: mechanical and thermal distortions. Mechanical distortions are caused by in plane vibrations of the membrane due to the rapid acceleration during mask stepping and in conjunction with the mounting or fixturing of the mask support ring. Additional mechanical displacements may occur if the stress in the resist is relieved during patterning. Thermal distortions occur due to the heating of the mask by the energy deposition of the electron-beam in the material. The heating effects during e-beam writing are described in detail in Chapter three.

When the X-ray mask is used in the exposure tool, other sources of distortions need to be considered. One source of distortion is the mounting of the mask in the exposure tool. Another important origin for pattern displacements is gravity. In the mask writing process the mask is in a horizontal position, but for exposure the mask is oriented in a vertical position. In addition, exposing the mask with X-rays heats up the mask. The amount of energy absorbed by the mask depends on the pattern density and the thickness of the mask membrane. The membrane typically absorbs thirty to fifty percent and the absorber material

almost a hundred percent of the incident radiation [2.4]. Reduction of the mask thickness to decrease the energy absorption is limited by a minimum thickness to maintain the necessary mechanical stability of the mask. Cooling of the mask with helium is a method used to decrease the thermal distortions due to the mask heating. Another source for mask distortions during the exposure process is radiation damage of the mask materials [2.7].

2.4 References

[2.1] ....M. C. Peckerar and J. R. Maldonado, "X-ray Lithography- an Overview," *Proceedings of the IEEE*, Vol. 81, No. 9, 1993.

[2.2] J. P. Silverman, "X-ray Lithography: Status, Challenges, and Outlook for 0.13 μm," *J. Vac. Sci. Technol. B*, 15, 1997.

[2.3] F. Cerrina, Center for X-ray Lithography, University of Wisconsin-Madison, private communications.

[2.4] ....P. Rai-Choudhury, "Microlithography, Micromachining, and Microfabrication," Vol. 1, 1997.

[2.5] M. F. Laudon, "Investigation of X-ray Lithography Mask Geometries and Mounting Techniques for a National X-ray Mask Standard," M.S. Thesis, University of Wisconsin, Madison, 1993.

[2.6] D. M. Puisto, IBM, private communication.

[2.7].... W. Arden, "Optical versus X-ray Lithography for Future Device Fabrication," *Microelectronic Engineering*, 13, 1991, p. 231-241.

Chapter Three

Electron-Beam Patterning of X-ray Lithography Masks

3.1 Introduction

A highly focused electron-beam is used to transfer the electric circuit pattern onto the polymeric photoresist layer of the X-ray mask membrane. By interaction with the molecules of the resist the incident electrons change the chemical structure of the polymeric material. The pattern transfer accuracy of the electron-beam writing process mainly depends on two factors: the spatial distribution of the electrons and the deposited heat within the membrane material. The splaying of the incident electrons in the matter results in an increase of the effective beam diameter with increasing depth in the material. For certain e-beam writing applications, as the writing of X-ray masks, the increase of the beam diameter with depth is the main physical limit for the resolution of the pattern transfer. Although it is possible to focus an electron-beam to a diameter of the order of 1 nm the resolution accuracy is limited to a diameter of the order of 10 nm [3.1].

3.2 Electron-Beam Writing Systems

Electron-beam writing systems can be divided into two basic categories: systems using a focused scanning e-beam and systems that project an entire pattern. The major advantage of the second method is a higher pattern writing speed compared to serial writing, where only small areas of pattern can be written at a time. Although e-beam writing with a scanning e-beam is comparably slow, it is the method used for mask fabrication. The requirements for the pattern transfer accuracy during e-beam writing of a mask are very high. For example making a 1× X-ray lithography mask requires a placement accuracy and linewidth control of 20 nm or better for typical minimum feature size from 0.25 μm to less than 0.1 μm. Only with the scanning electron-beam approach is the achievable resolution high enough to meet this accuracy requirements.

There are two different kinds of scanning e-beam systems as shown in Fig. 3.1.

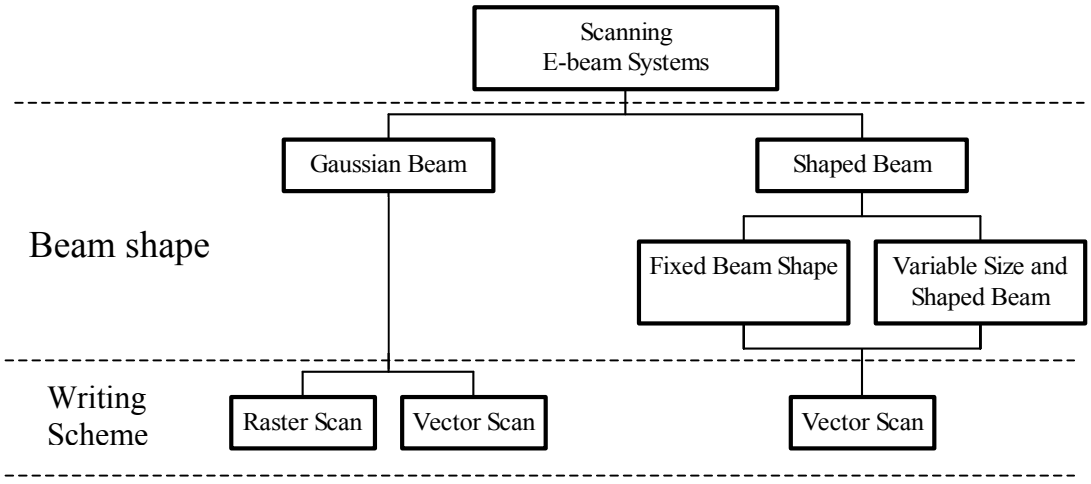
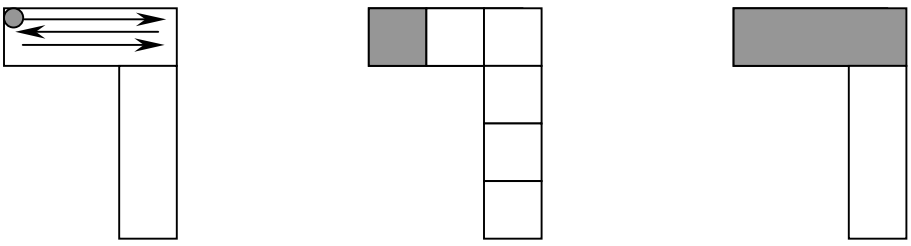


Figure 3.1. Classes of e-beam writing systems [3.2].

Gaussian systems use a round beam that scans the pattern pixel by pixel. The pixel diameter, equal to the e-beam diameter, determines the minimal feature size that can be written with the system. The second kind of e-beam systems are shaped beam writers with either fixed or variable shape. Figure 3.2 shows how a pattern is written with a Gaussian, a fixed shape and a variable shape electron-beam system. Shaped electron-beam systems allow the writing of a specific pattern with fewer flashes than with a Gaussian beam system. This results in a shorter pattern writing time for shaped electron-beam systems than for Gaussian systems.



Gaussian round beam Fixed Shaped Beam Variable Shaped Beam

Figure 3.2. Different pattern generation techniques. Adapted from [3.2].

The hardware of all electron-beam writing systems has the same basic components. Figure 3.3 depicts a schematic of a system. The basic system components are (a) the electron source, called the electron gun, (b) the electron optical column, (c) the mechanical table that carries the target and (d) the control system that drives and monitors the movement of the table and controls the shape of the e-beam. The electron gun is either a thermoionic or a field emission source. In thermoionic guns electrons are produced by heating a material to a high temperature at which it emits electrons at the surface. Field emission sources produce

electrons by extracting electrons from a sharp tip of tungsten with a very high electric field. One advantage of field emission electron guns is that a higher resolution can be obtained because of a smaller variations in the electron energy than for thermoionic guns. An even more important advantage of field emission electron guns is a much higher brightness of the source. Field emission electron guns can deliver up to a thousand times the current density produced by a thermoionic source. The brightness of the electron source is an important factor for the pattern writing time. The writing time for one spot of pattern, the so called minimum dwell time, is equal to the ratio of the current necessary to develop the photoresist and the current density that is provided by the source. Electron sources with a high brightness are needed especially for electron-beam writing systems with a shaped beam. Shaped beams expose, compared to Gaussian e-beams, a large area. For this reason a high current rate has to be provided by the electron source to avoid a long dwell time.

The electron gun is located in the column of an e-beam writing system. The optical hardware elements in the column can be divided into three groups: (α) the spot shaping aperture, (β) the system of lenses and (γ) the scanning deflectors. Not every e-beam writing system has the parts of the optical column configured like shown in Fig. 3.6. Most systems use a more complicated combination of different lenses and beam apertures to produce an e-beam with minimum possible aberration. The lenses in the optical column are used to focus and define the electron-beam. With the apertures the beam is limited to a certain size. Apertures are also used to define specific forms of the beam in shaped beam systems. The scanning deflectors move the beam to different locations on the target. Either electromagnetic or electrostatic coils are used as scanning deflectors. The deflectors can only be used to move the e-beam within a limited area of several millimeters. Only for small electron-beam deflections is the position accuracy high enough and the edge gradient and the distortion of the beam is below the limits given for a specific minimum feature size. To write the pattern on the complete mask membrane surface the positioning on the target is done by a combination of beam deflection and moving of the table that holds the mask.

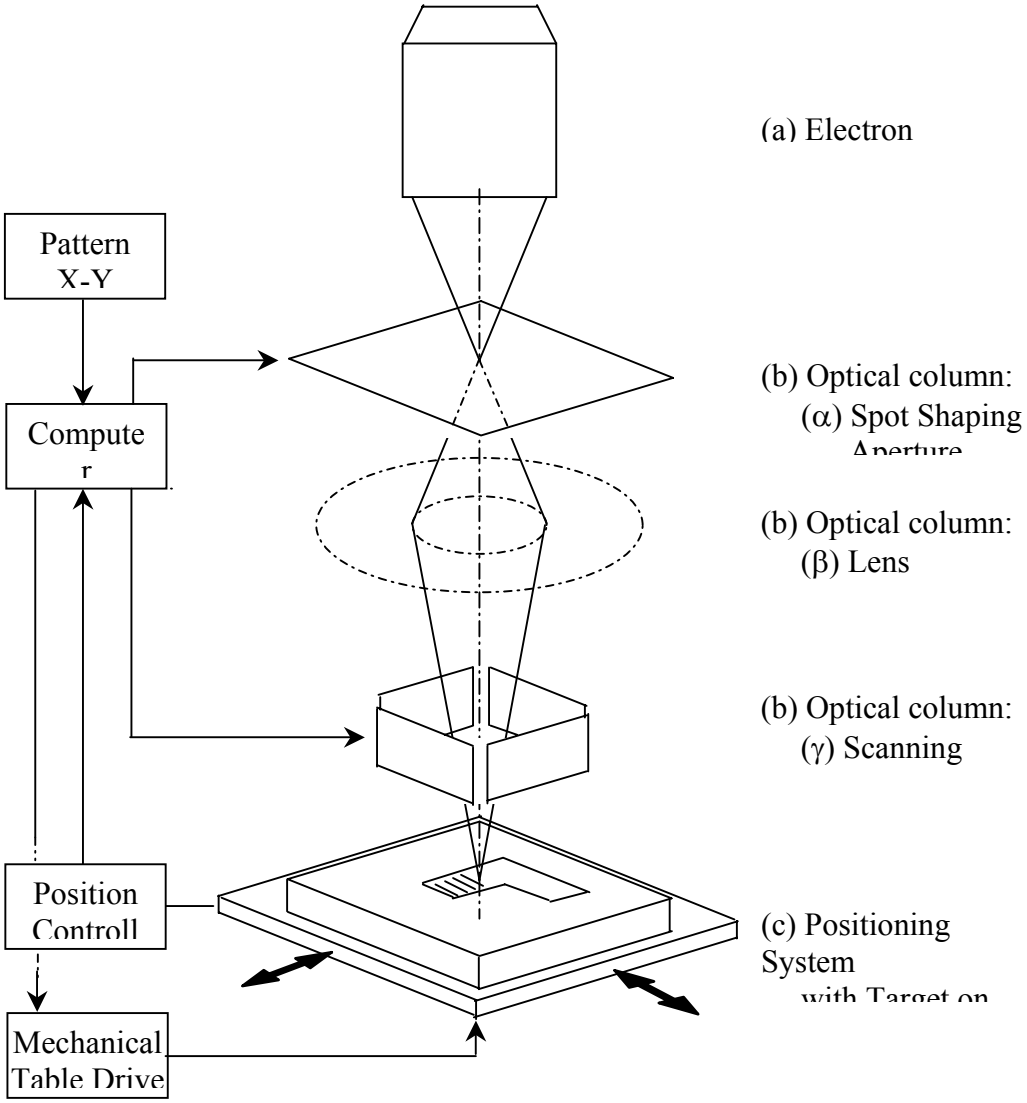


Figure 3.3. Schematic of an electron-beam writing system. Adapted from [3.2].

The way of coupling beam deflection and table movement depends on the mask writing strategy. Figure 3.1 shows that there are two different kinds of writing schemes. The raster scanning method is only applied for e-beam systems with a Gaussian beam. In e-beam writing machines that use this writing style the beam changes its position relative to the mask by continuously moving the table. The complete mask area is written by scanning it with parallel lines either in serpentine or unidirectional style. Simultaneously with the scanning, the beam is switched on and off according to the position of the pattern. The beam deflection system of the optical column is only used to move the beam perpendicular to the table motion and to correct for positioning errors that are detected by monitoring the table movement with a laser interferometer. The raster scanning method is used by the commercially well established mask making tool, the MEBES system produced by Etec [3.3]. A disadvantage of the system is that the throughput decreases with increasing resolution because the density

of the pixels increases and the complete mask area has to be scanned. Systems that use the vector scanning method instead of the raster scanning do not have this disadvantage. The vector scanning machines deflect the e-beam sequentially to individual pattern elements. By limiting exposure only to mask areas that have to be patterned, scanning time is saved. Gaussian beam systems that use the vector scanning method move the beam to individual pattern elements that are written with the raster writing method as shown in Fig. 3.2. Vector scanning systems with a shaped beam have the advantage that they can develop the same pattern with less flashes than a Gausssian beam system and thereby can save exposure time. Writing systems with vector scanning and a variable shaped beam, as the EL-4 produced by IBM and the Excaliber produced by Etec, have the highest throughput of all e-beam writing systems [3.3].

Figure 3.4 shows a writing scheme typical of a variable shaped beam system with vector deflection. The mask membrane area that has to be patterned is subdivided into fields. The e-beam writing machine does the stepping from field to field in unidirectional writing style by moving the table [3.4].

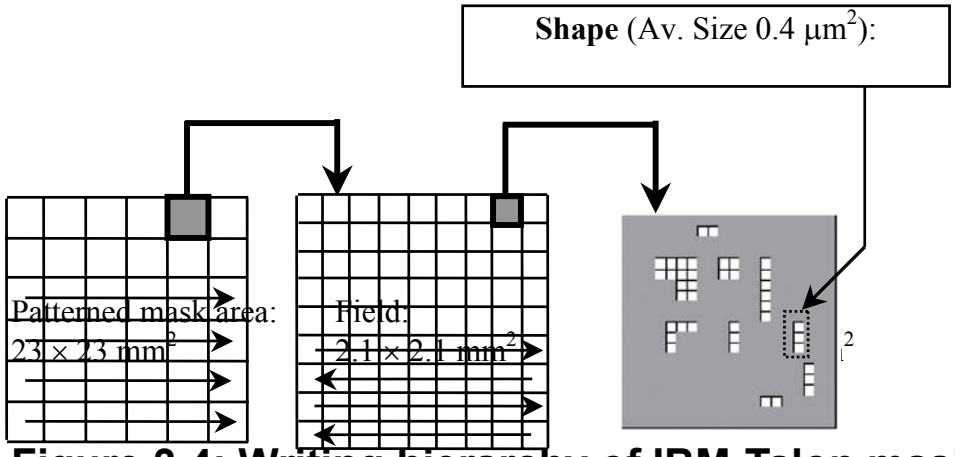
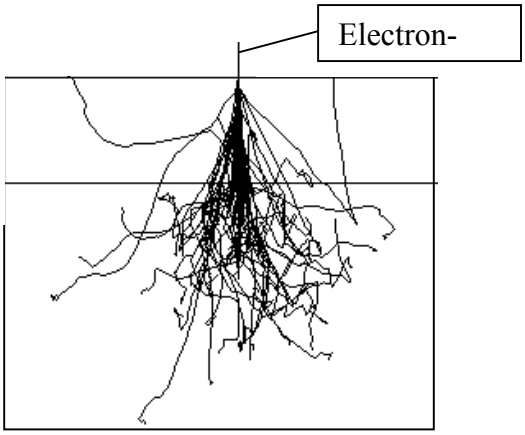


Figure 3.4: Writing hierarchy of IBM-Talon mask [3.5].

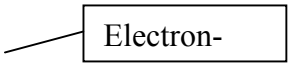
The mechanical stepping speed is smaller than the deflection stepping speed by a factor of five hundred. Beam deflection can be used for the stepping within the field from subfield to subfield without causing a large pattern error due to beam shape distortion. The field is written in serpentine style. Within the subfield a vector scanning fashion is used. Identical shapes of circuit pattern are written sequentially as shown in Fig. 3.2 for the variable shaped beam system.

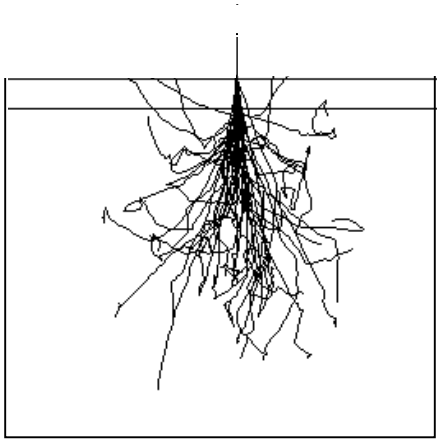
3.3 Electron Solid Interaction

When a finely focus electron-beam enters a material, as during e-beam writing where the e-beam enters the polymeric photoresist material from the surrounding vacuum, the electrons lose energy by elastic and inelastic collisions with the atomic particles of the material. Inelastic collisions occur when an incident electron hits an electron of the material. The light electrons of the material can take up a large portion of the incident electrons kinetic energy without deflecting it significantly. Because the nuclei of the material are much heavier than the incident electrons and have a much bigger charge, a collision with a nucleus results in a scattering of the electron but only a very small energy loss. Because of a very high density of atomic electrons in the material compared to the density of atomic nuclei, the probability for an electron-electron collision is much higher than for an electron-nucleus interaction.



Beam acceleration voltage 10 keV.





Beam acceleration voltage 20 keV.

Figure 3.5.
Fifty electron trajectories of a e-beam incident on a material of 0.5μm

PMMAon Si substrate calculated with the Monte Carlo program [3.6].

Therefore the motion of an incident electron in the material can be approximated as a continuously slowed down particle moving along a straight line changing its direction at points where a collision with a nucleus takes place. The scattering events caused by electron nucleus interaction can be divided into two classes: forward and backward electron scattering. Fast electrons entering at the material surface and penetrating the resist are only scattered by small angles. These scattering cases are called forward scattering events. Deeper in the material the scattering of slower electrons by large angles is called backscattering. The scattering of electrons entering a structure of silicon with a film of 0.5 μm photoresist, is depicted in Fig. 3.5. The three-dimensional electron trajectories projected onto a plane parallel to the e-beam and rectangular to the material surface are shown for electron-beams accelerated with a voltage of 10 keV and 20 keV. The forward scattering events of fast electrons occurring mainly in the resist layer are hard to identify, because of the small scattering angle. In Fig. 3.5 forward scattering can only be noticed as the broadening of the beam diameter in the resist with increasing depth. The increase in the beam diameter Δd_{beam} in nanometers is described by the following empirical equation with the beam acceleration voltage ΔV_{beam} in keV and the resist thickness δ_{resist} in nanometers [3.3]:

$$\Delta d_{beam} = 0.9 \cdot \left(\frac{\delta_{resist}}{\Delta V_{beam}} \right)^{1.5} \qquad [3.3.1]$$

For a beam acceleration voltage of 75 keV, typical for a modern e-beam writing system, and a resist thickness of 0.5 μm the beam would increase its size by 15.5 nm. This physical limit for the e-beam writing pattern transfer accuracy can be improved by using a higher acceleration voltage and developing resist materials that require a smaller thickness.

The large angle backscattering of slower electrons is clearly noticeable in Fig. 3.5. Since mainly low energy electrons are backscattered, increasing of the beam acceleration voltage moves the point where backscattering first occurs to a point deeper in the material. For an acceleration voltage of only 10 keV backscattering starts in the photo resist layer and causes a significant increase of the effective beam diameter in the lower region of the polymeric layer. Figure 3.5 shows that there are almost no backscattering events within the resist layer for a doubled acceleration voltage. Backscattered electrons cause additional resist exposure in regions with a significant distance from the e-beam center. The spreading of the beam by scattering and the resulting additional resist exposure is called the electron-beam proximity effect. This effect creates a limit to the electric circuit pattern density. For a pattern with a high density, the exposure of a certain pattern element causes an unwanted exposure of nearby pattern features. The impact of the proximity effect depends very much on the structure of the written target. In masks with thin material layers on a thick substrate, like optical masks, a significant number of electrons are backscattered to the resist layer. In X-ray masks that only have a very thin membrane, the proximity effect has much less importance. For a high acceleration voltage not many electrons are backscattered within the mask membrane, because most of the electrons go through the mask membrane and the distance they travel within the material is too short to decrease their kinetic energy enough to cause a backscattering event.

3.4 Calculation of the Energy Density Function with a Monte Carlo Simulation

The distribution of electrons deposited in a solid material by an e-beam and the distribution of their energy within the target material can be calculated most accurately by simulating electron trajectories in the target with a Monte Carlo method. Analytical solutions for these distribution functions have also been developed, but the analytical prediction has a small error only for a target with a single material [3.2].

In most Monte Carlo calculation approaches the electron trajectories are calculated with a single scattering model. In this kind of model electrons are assumed to move along straight lines until they get scattered by interaction with a target nucleus. An electron trajectory is calculated by successively calculating the path length between collisions with nuclei and the scattering angles until the electron stops moving within the material or exits it. By determining the trajectories for a large number of single electrons the electron distribution function in the target can be calculated. The angle that an electron is scattered from its original direction depends on how close it is to the force field of a target nucleus. The Thomas-Fermi equation describes the potential field of a nucleus with a high accuracy:

$$V(r) = -\frac{Z \cdot e^2}{r} \cdot \exp\left(\frac{-0.745 \cdot r \cdot \sqrt[3]{Z}}{a_0}\right) \qquad \text{with} \quad a_0 = 0.53 \times 10^{-8}\,\text{cm} \qquad [3.4.1]$$

Where Z is the atomic number, e the electronic charge and a_0 the Bohr radius. This potential field equation accounts for the decrease of the nucleus electrical field by the surrounding electrons. The interaction of the potential field of a nucleus described by the Thomas-Fermi equation and the field of an electron has been characterized first by Rutherford [3.7]. He developed a function that can be used to determine the elastic scattering cross-section of a nucleus σ_e, the area around a nucleus normal to a trajectory of a passing electron in that the electron is scattered. A scattering cross-section of a nucleus is depicted in Fig. 3.6.

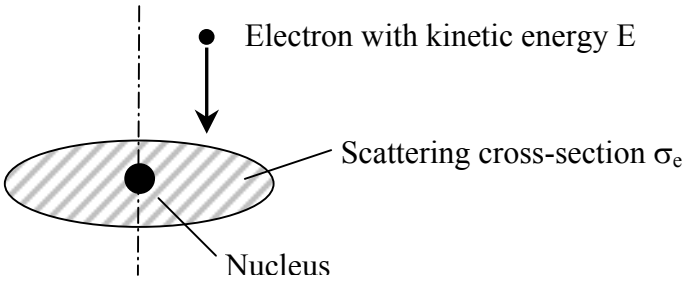


Figure 3.6. Scattering cross-section of a nucleus.

A slightly modified form of this equation developed by Browning is

$$\sigma_e = \frac{3\times 10^{-18}\cdot Z^{1.7}}{E + 0.005\cdot Z^{1.7}\sqrt{E} + \frac{0.0007\cdot Z^2}{\sqrt{E}}} \qquad [3.4.2]$$

with σ_e in cm^2, the atomic number Z and the kinetic energy E of the passing electron in kilo electron volts [3.8]. In the process of calculating the trajectory of an electron in a target material this equation is used to calculate the average distance Λ_e that the electron moves between scattering events. The distance Λ_e, the mean free path of an electron, is equal to

$$\Lambda_e = \frac{A}{N_A\cdot\rho\cdot\sigma_e} \qquad [3.4.3]$$

with A the atomic weight of the material, N_A the Avogadro number and ρ the density of the material. The electron energies E_i used to calculate the mean free path lengths are the kinetic energy values at the beginning of each straight line within a trajectory (Figure 3.6). The actual step length Λ_i of an electron between scattering events is approximated by multiplying the mean free path length Λ_e with the logarithm of a random number between zero and one. The scattering angles θ_i and φ_i for each deflection are also determined with the help of random numbers. Since the probability function of the azimuth angle has uniform distribution, the angles φ_i are calculated with

$$\varphi_i = 2\pi\cdot RND_1 \qquad \text{with } RND_1 = \text{random number in } [0\ldots 1]. \qquad [3.4.4]$$

The conical deflection angle θ_i does not have a uniform distribution. The random number used to calculate θ_i has to be weighted by eq. 3.4.5. Rutherford developed this equation by

combining eq. 3.4.1 with a differential electron mass balance for a unit solid angle around a nucleus known as the Rutherford differential scattering cross section equation [3.7]. The variable α in this equation is the atomic screening parameter accounting for the decrease of the nucleus potential by the surrounding electrons.

$$\cos(\theta_i) = 1 - \frac{2\alpha \cdot RND_2}{1 + \alpha - RND_2} \quad \text{with} \quad RND_2 = \text{random number in } [0 \ldots 1] \qquad [3.4.5]$$

The kinetic electron energies E_i, used to calculate the mean free electron path, are assumed to decrease continuously between elastic scattering events by the inelastic scattering with atomic electrons. This energy loss is described by the Bethe equation

$$\frac{dE}{dx} = -\frac{2\pi e^4 \rho_e}{E} \cdot \ln\left(\frac{1.166 \cdot \rho_e \cdot E}{I}\right) \qquad [3.4.6]$$

with ρ_e the density of atomic electrons and I the mean excitation energy of the solid material. Since this equation describes where a single electron deposits energy in the material a simulation of many electron trajectories can be used to determine the distribution function of the energy deposited by an e-beam in a solid. A computation of 50,000 to 100,000 electron trajectories is necessary is required to obtain an accurate approximation of this distribution function.

Figure 3.7 shows a schematic of an electron trajectory entering a solid material normal to the surface. The first scattering event is assumed to happen at the surface. The step length Λ_0

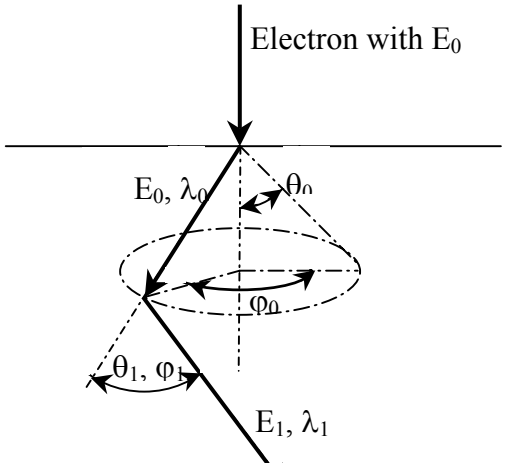


Figure 3.7. Trajectory of an electron. Adapted from [3.2].

and the first scattering angles θ_0 and φ_0 are calculated with the Monte Carlo technique as previously described. The next spatial electron position is calculated and the new electron energy E_1 obtained with eq. 3.4.5 is used to determine Λ_1, θ_1 and φ_1. This calculation procedure is repeated until the simulated electron leaves the material or the kinetic electron energy reaches a value of almost zero.

3.5. Heat Transfer Mechanisms in an X-ray Mask Membrane during E-Beam Writing

In this section of Chapter three the heating effects in a mask are characterized that occur due to the energy deposition during the e-beam writing process. The mechanisms of the energy deposition in the mask material by electron solid interaction were explained in Section 3.3 and 3.4.

In the following section the influence of different mask and e-beam tool parameters on the temperature characteristics of the mask is explained. The various effects of the mask temperature rise on the writing process, mask and pattern characteristics are illustrated in the second section of this chapter. The temperature response of the mask to the electron-beam

exposure depends mainly on three factors: the magnitude and spatial distribution of the energy generation rates, the exposure time and the heat diffusivities of the mask materials. The energy deposition rate and its distribution is determined by the system characteristics of the e-beam tool and the geometry of the pattern that has to be written. E-beam tool system characteristics determining the energy deposition rate are the current density of the beam and the beam acceleration voltage. High current densities result in a large amount of charge deposited per area and therefore high energy densities. The beam acceleration voltage determines the kinetic energy of the incident electrons and thereby the amount of deposited energy per electron. How the acceleration voltage corresponds with the shape of the energy distribution function is described in Section 3.3.

Pattern geometry parameters influencing the energy deposition rate and distribution are the pattern density and the e-beam flash size. High local pattern densities require that a high percentage of surface area is written with the e-beam and therefore result in high local energy densities. It is characteristic of the electron-beam writing process that energy is only deposited into a very small area. The size of the electron-beam flash is typically on the order of four to seven orders of magnitudes smaller than the total patterned area of the mask membrane. Although the area of energy deposition is very small, significant local temperature rises can occur because of the high density of deposited energy. The flash size determines how much surface area is exposed at the same time to the electron-beam. Since the current density is independent of the flash size, a large flash size results in a high energy deposition rate.

The necessary exposure time for a flash depends on the current density of the e-beam and the photoresist sensitivity. High current densities result in small exposure times to deposit the necessary charge. Because the amount of deposited energy per time is large, the local temperature rise is high. To minimize the total writing time of a mask the current density is maximized to the point where local temperature rises reach the allowable limit. Because a small photoresist sensitivity corresponds to a small amount of charge necessary to develop the resist, sensitive resists need smaller exposure times. For this reason an optimal resist is the one with maximal sensitivity.

To explain different effects of mask heating on pattern and mask characteristics it is useful to distinguish the heating effects of different geometry and time scales. A classification into three groups seems to be the best approach to describe the different heating mechanisms. Figure 3.8, showing schematically a section of a mask membrane written with squared flashes, is used as an example for an e-beam writing process to explain the different heating effects. The first contribution to heating is the direct heating. This is induced by the electrons, at the point and time of exposure of a specific flash. The second contribution to heating is the proximity heating. Proximity heating is the heating of a flash area due to heat diffusion from an adjacent flash area that was previously exposed. The last contribution to heating is the global heating. The category of global heating accounts for the slow uniform heating process of the total mask area.

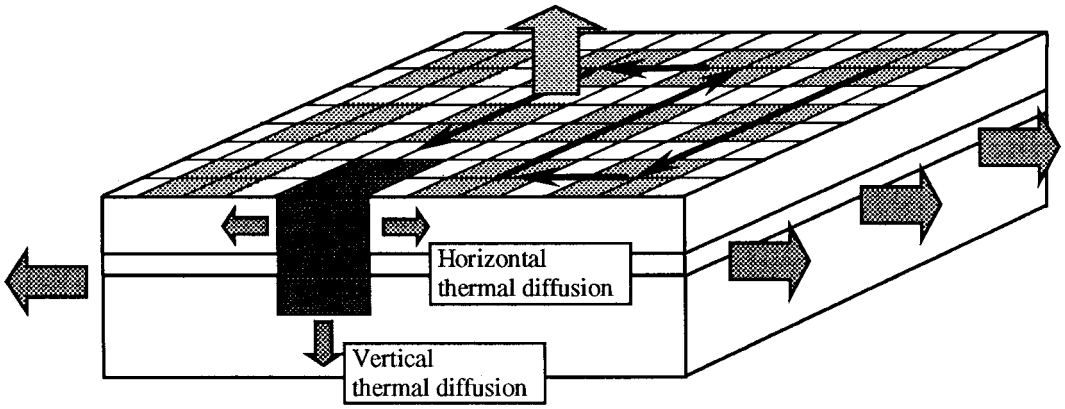


Figure 3.8. Electron-beam writing of squared flashes on a mask membrane.

The direct heating can cause significant temperature rises. Values between 15K to 800 K have been reported in the literature for different exposure conditions [3.9, 3.10, 3.11]. These significant local temperature rises are mainly important because of their influence on the temperature dependant resist properties. The high temperature dependence of the resist sensitivity has been determined experimentally for a variety of photoresist materials [3.9, 3.10, 3.12, 3.13, 3.14, 3.15]. For example the sensitivity of the photoresist material PMMA decreases by about 30 percent for a temperature rise of 50 K [3.14]. Because of non-uniform mask temperature rises over the area of a flash due to local heating, the resist sensitivity is non-uniform over this area. Since the current density is uniform over the flash area, direct heating can cause parts of the flash area to be either under or overexposed. The influence of this effect on the edge profile and the pattern shape accuracy was analyzed experimentally for many writing conditions [3.9, 3.10, 3.12, 3.13, 3.14, 3.15]. Another effect of local temperature rises due to local heating is the possible change of the mechanical stress in the resist material. For the resist material PMMA the stress can drop to a tenth of its initial

value. This stress relief can cause significant pattern distortions and displacements [3.16, 3.17, 3.18]. New resist materials, such as Shipley's negative resist SNR200, experience only a negligible stress relief [3.19].

Proximity resist heating, the heating of a flash area due to heat diffusion from an adjacent flash area that was previously exposed, has the same effects on patterning accuracy as direct heating by changing the resist sensitivity.

There are various methods used to avoid patterning inaccuracies by local heating. The reduction of the flash size makes the amount of deposited energy per flash and therefore the temperature rise smaller. Another method of decreasing the heat generation is the reduction of the current density. Writing the pattern in multiple paths also makes the heat generation per volume and time smaller, because with each e-beam flash only a fraction of the total dose is deposited. The local mask temperature rise can also be decreased by increasing the waiting time between flashes. The different methods of reducing the local heating were examined in [3.12, 3.13], and [3.14].

The influence of global heating on pattern inaccuracies due to change of resist sensitivity is negligible, because the uniform temperature rise over large areas of the membrane is too small to effect the sensitivity. Patterning errors connected with global heating are mainly caused by thermal distortions of the mask. Even a small mask temperature rise can cause significant thermal distortions, because the amount of distortions depends not only on the value of temperature rise, but also on the size of the area where the temperature rise occurs. The mask distortion due to heating is a reversible transient process. To estimate pattern

displacements for a specific pattern feature caused by this effect it is necessary to know the mask displacement at the writing location as a function of time for the pattern feature. To be able to predict local mask displacements accurately, it is necessary to have a precisely calculated transient mask temperature profile.

3.6 References

[3.1] O. C. Wells, "Scanning Electron Microscopy," McGraw-Hill, NewYork, 1974.

[3.2] L. F. Thompson, C. G. Willson, and M. J. Bowden, "Introduction to Microlithography," ACS Professional Reference Book, 1994.

[3.3] P. Rai-Choudhury, "Microlithography, Micromachining, and Microfabrication," Vol. 1, 1997.

[3.4] H. C. Pfeiffer, D. E. Davis, "EL-4 a new Generation Electron-Beam Lithography System," *J. Vac. Sci. Technol. B*, 11, 1993.

[3.5] D. M. Puisto, IBM, private communication.

[3.6] D. C. Joy, SS_MC: Turbo Pascal Program for Monte Carlo Simulation of single Electron Scattering in a Solid Material, 1991.

[3.7] J. D. Jackson, "Classical Electrodynamics," John Wiley and Son, 1962.

[3.8] J. R. Lowney, "User's Manual for the Program MONSEL-1: Monte Carlo Simulation of SEM Signals for Linewidth Metrology," National Institute of Standards and Technology, 1994.

[3.9] S. Babin, "Measurement of Resist Heating in Photomask Fabrication," *J. Vac. Sci. Technol. B*, 15, 1997.

[3.10] E. Kratschmer and T. R. Groves, "Resist Heating Effects in 25 and 50kV E-Beam Lithography on Glass Masks," *J. Vac. Sci. Technol. B*, 8, 1990.

[3.11] M. Karnezos and P. Weimar, "X-ray Mask Heating during Electron-Beam Patterning," *J. Vac. Sci. Technol. B*, 5, 1987.

[3.12] K. Nakajima, T. Honda, and H. Matsumoto, "New Compensation Methods for avoiding Proximity Resist Heating in Variable Shaped Electron-Beam Lithography," *J. Vac. Sci. Technol.*, 8, 1990.

[3.13] E. H. Mulder, K. D. van der Mast, and A. C. Enters, "Thermal Effects in Electron-Beam Lithography," *J. Vac. Sci. Technol. B*, 7, 1989.

[3.14] S. Babin, P. Hudek, and I. Kostec, "Quantitative Measurement of the Resist Heating in a Variable Shaped Electron Lithography," *J. Vac. Sci. Technol. B*, 15, 1997.

[3.15] N. K. Eib and R. J. Kvitek, "Thermal Distribution and the Effect on Resist Sensitivity in Electron-Beam Direct Write," *J. Vac. Sci. Technol. B*, 7, 1989.

[3.16] D. L. Laird, R. L. Engelstad, D. M. Puisto, R. E. Acosta, K. D. Cummings, and W. A. Johnson, "Predicting in-plane Distortion from Electron-Beam Lithography on X-ray Mask Membranes," *J. Vac. Sci. Technol. B*, 14, 1996.

[3.17] B. Shamoun, M. Sprague, F. Bedford, and R. Engelstad, "X-ray Mask Distortions during E-Beam Patterning," to appear in *Microelectronic Engineering*, 1998.

[3.18] F. Bedford, R. L. Engelstad, and F. Cerrina, "Thermal Analysis of an X-ray Mask Membrane during Electron-Beam Patterning," *Proceedings of TECHCON*, 1996.

[3.19] J. P. Silverman, "X-ray Lithography: Status, Challenges, and Outlook for 0.13 μm," *J. Vac. Sci. Technol. B*, 5, 1987.

Chapter Four

Calculation of the Transient Mask Temperature Profile during E-Beam Writing

4.1 Local Mask Heating

This chapter describes how the mask heating effects during e-beam patterning are analyzed. The different heating effects that occur, such as direct heating, proximity heating and global heating are qualitatively described in Section five of Chapter three. The primary goal of this research is to investigate local mask heating effects, the direct and the proximity heating.

Mask direct heating is the effect that is analyzed in Chapter five. The mask temperature profile is calculated for the e-beam writing of a single flash on a mask. Knowing the local mask temperature profile on the geometry level of a single flash is mainly important for determining photoresist sensitivity changes. The local temperature rise can lead to a non-uniformity in the resist sensitivity which causes patterning inaccuracies due to local under or over-development. In Chapter five the local heating is analyzed for writing a flash on an optical mask. The results obtained in Chapter five can only give a qualitative understanding of X-ray mask direct heating, because of the different mask geometry. The optical mask has a thin metal and polymer layer on top of bulk glass whereas an X-ray mask is only a thin membrane in the area of e-beam patterning.
Heating of an X-ray mask is examined in Chapter six. The main goal of Chapter six is to determine how local heating contributes to global mask heating and to estimate the accuracy of an averaging technique used for the calculation of global mask temperature profiles. Describing the global mask heating exactly is of interest, because a precise prediction of thermal mask distortions is only possible with an accurately calculated transient mask temperature profile. The calculation of the actual transient X-ray mask temperature profile is made difficult by the complicated heat loading conditions of the mask by the e-beam writing process. The energy is deposited in the mask in very many writing steps. An integrated circuit, as for example a Pentium Processor, consists of a few million transistors. A lithography mask for the production of this chip usually has more than one chip layout on its surface. Many steps are required to write each single feature with an electron-beam on the mask. The size of a single feature is much smaller than the total mask area that is written: the size of a feature is by the order of up to ten orders of magnitudes smaller than the total patterned mask membrane area. For this reason a numerical model of the full mask describing the heat deposition process on the real scale would need to be discretized with a

very fine mesh size. This large number of elements and load steps would result in a very complex numerical model that is not suitable to simulate the energy deposition process in a reasonable computational time. Consequently the real writing process has to be approximated with a simpler model.

In [4.1, 4.2], and [4.3] the actual writing process is simplified by describing the heat deposition during the beam writing of the features in a certain mask area as an averaged heat deposition in this area. The area used in [4.1, 4.2], and [4.3] for averaging the energy deposition is the mask area that a modern e-beam system writes by deflecting the beam without mechanically moving the mask. For a typical variable shaped beam system with vector deflection used to write an IBM-Talon mask this area has a size of 2.1 mm × 2.1 mm. This area is called a mask field.

In Fig. 4.1 it is shown that the patterned mask area is written by stepping fields in unidirectional style over the complete patterned area. Using a uniform heat generation over the writing time and area of a field to approximate the actual e-beam writing process makes it possible to simulate the complete mask writing process in a reasonable computational time. Because the ratio of the mask dimensions to the size of the field area, the area of constant energy deposition, is not very large, discretizing the numerical model does not require too many grid points. Also the number of simulation load steps, equal to the number of mask fields, is not very large for this kind of mask model.

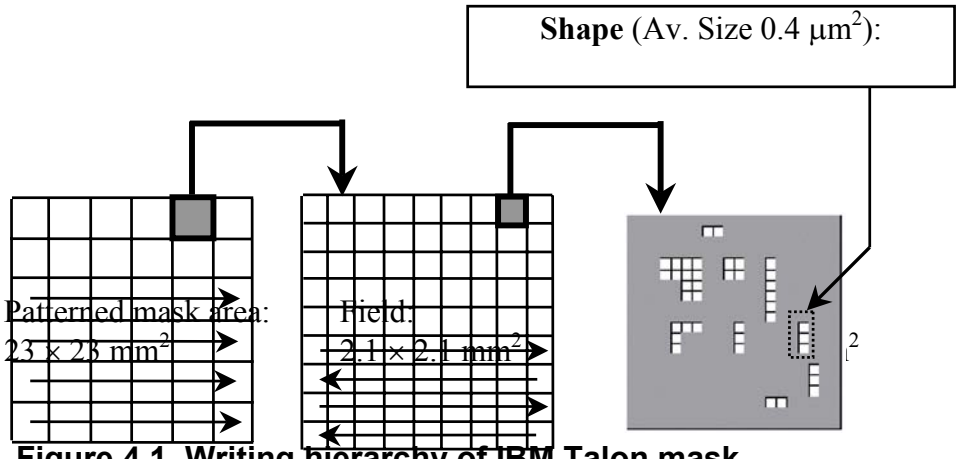


Figure 4.1. Writing hierarchy of IBM Talon mask.

To obtain accurate results for the transient mask temperature profile with the technique described in the previous paragraph, it is essential, that applying an averaged heat generation to the field area can describe the actual field writing process with reasonable precision. In this research the accuracy of the averaging is estimated. In Chapter six local mask temperature profiles are calculated by simulating the actual beam writing process and the results are compared to profiles calculated with the averaging technique. Temperature profiles for the actual writing process are calculated on the geometry level of a subfield, shown in Fig. 4.1, to limit the necessary computational time. The comparison of the two subfield temperature profiles is used to estimate the accuracy of the averaging technique.

4.1 References:

[4.1] F. Bedford, R. L. Engelstad, and F. Cerrina, "Thermal Analysis of an X-ray Mask Membrane during Electron-Beam Patterning," *Proceedings of TECHCON*, 1996.

[4.2] D. L. Laird, R. L. Engelstad, D. M. Puisto, R. E. Acosta, K. D. Cummings, and W. A. Johnson, "Predicting in-plane Distortion from Electron-Beam Lithography on X-ray Mask Membranes," *J. Vac. Sci. Technol. B*, 14, 1996.

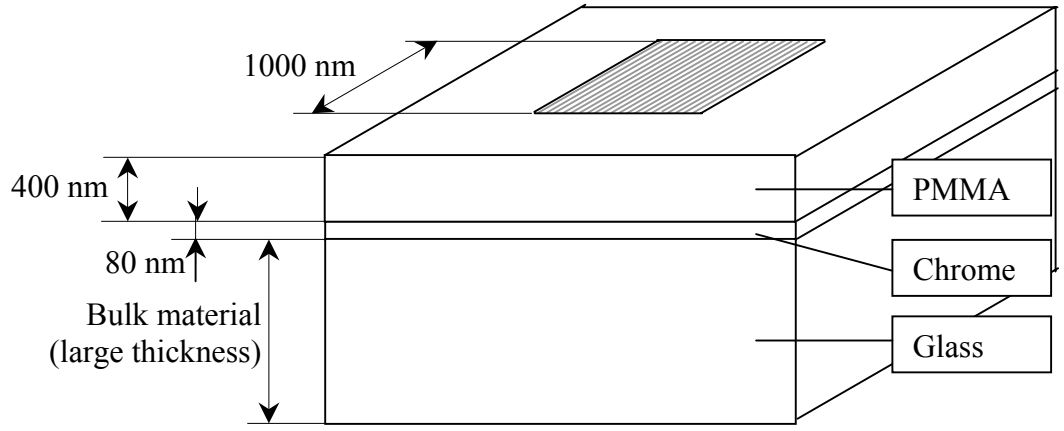
[4.3] B. Shamoun, M. Sprague, F. Bedford, and R. Engelstad, "X-ray Mask Distortions during E-Beam Patterning," 23rd Micro and Nano Enginering International Conference, 1997.

Chapter Five

Local Mask Heating

5.1 Introduction

In Section 3.5 the different mask heating mechanisms during e-beam mask writing were explained. In this Chapter one of these mechanisms, the mask direct heating, is examined in detail. As explained in Section 3.5 the effect described as direct heating is the local mask heating in the area of a flash during its writing time. Direct heating is mainly of interest because of its influence on the resist sensitivity. The non-uniform temperature rise in the flash area results in a non-uniform resist sensitivity change which causes local under or over development of the photoresist.



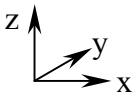


Figure 5.1. Geometry of optical mask and square flash.

Material	Density (kg/m^3)	Conductivity (W/m K)	Heat Capacity (J/kg K)
PMMA	1210	0.176	1170
Chrome	7200	62.9	465
SiN	2220	1.38	787

Table 5.1. Mask material properties.

In this chapter the exposure of a single flash on an optical mask is examined. The dimensions of the mask geometry and the flash are shown in Fig. 5.1. Material properties of the mask materials are listed in Table 5.1. The mask temperature profile is calculated for writing the squared flash with a typical electron-beam writing machine. The loading conditions used correspond to the energy input into the mask by an e-beam with 40 A/cm^2 current density and an acceleration voltage of 50 kV. Temperature profiles of the mask are calculated for different flash exposure times in a range related to typical sensitivity values for the photoresist material PMMA. The heat generation loading conditions for this exposure case are described in the following section.

5.2 Calculation of the volumetric heat generation rate function for square flash exposure of an optical mask

The volumetric distribution function of the energy deposited in the mask material by the electron-beam is calculated with a Monte Carlo simulation method as described in Chapter three. The result of the Monte Carlo simulation is the energy density function for the exposure of the mask with an electron-beam point source. The energy deposition function for a beam with squared shape is calculated with the result for the electron-beam point source.

The energy deposition function for the mask geometry shown in Fig. 5.1 and a point source with 50 KV was calculated by [5.1]. This deposition function is the result of simulating 100000 electron trajectories with a Monte Carlo technique. The form of the calculated function is a volumetric step function with values of constant energy densities for ring volume elements around the e-beam center as shown in Fig. 5.2.

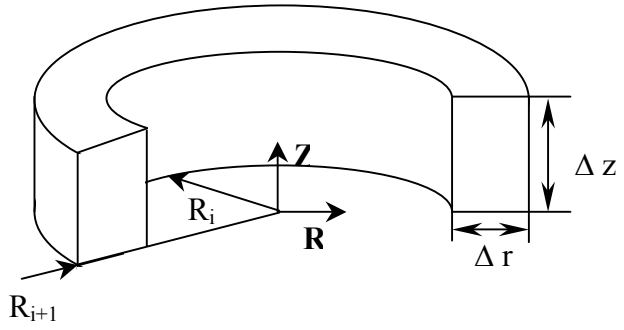


Figure 5.2. Volume element of constant energy density.

The function that describes the radial dimensions of the volume elements is given by

$$R_i = R_0 \cdot \exp(j \cdot \Delta rho) \qquad \text{for j=0.5, 1.5, 2.5, 3.5....} \qquad [5.1]$$

and i=j+0.5

The constants R_0 and Δrho of this exponential function and the Δz values are given in Table 5.2. for the three material layers of the optical mask.

Material	R_0 [nm]	Δrho	Δz [nm]	Number of sub-layers
PMMA	8	0.25	8	50
Chrome	8	0.25	8	10
Glass	125	0.25	500	50

Table 5.2. Constants for equation [5.1] and thicknesses of volume elements with constant energy density.

The number of sub-layers listed in Table 5.2 is the number volumes in depth with constant energy density and thickness Δz for each material layer. For the bulk material glass the number of sub-layers corresponds to the depth where the energy density reaches zero.

In Fig. 5.3 the energy density function is shown for one sub-layer of each material. The energy density values are normalized with the energy deposited by a single electron. Characteristic for this function is that the density of absorbed energy drops very fast with increasing distance from the e-beam center. The radial spread increases with depth in the material. The numerical values of the energy density function are listed in Appendix A.

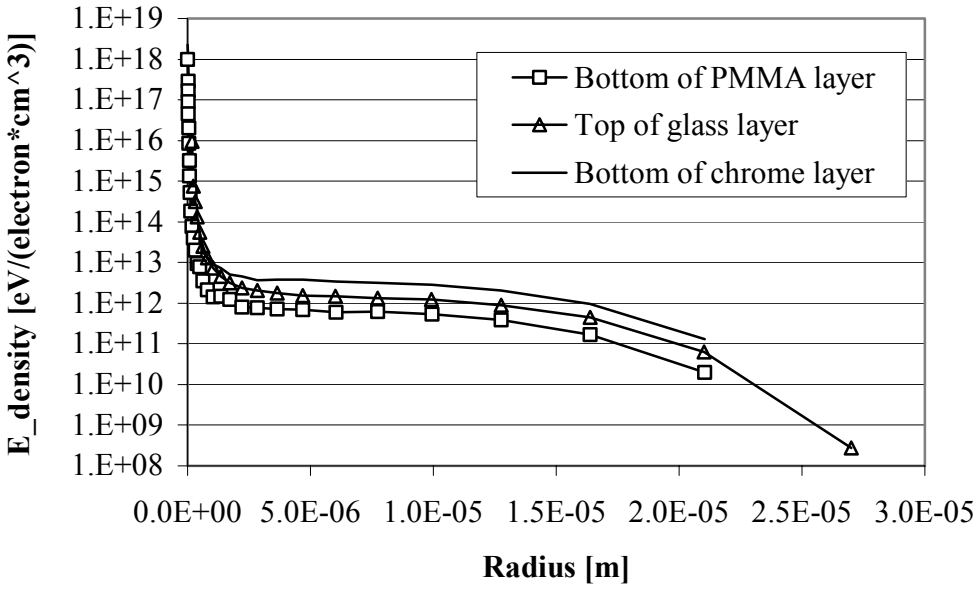


Figure 5.3. Energy density functions for sub-layers of PMMA, chrome and glass.

The trend of increasing radial spread with increasing depth in the material is clearly shown in the contour plot of the complete energy density function in Fig. 5.4. The shape of this function corresponds to the radial spread of the electron trajectories presented qualitatively in

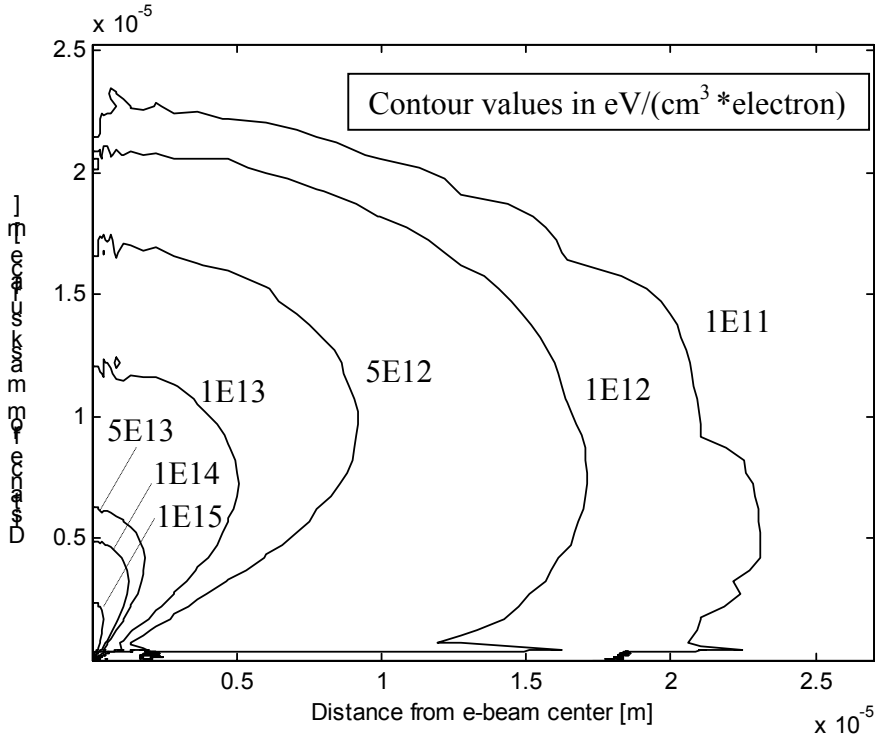


Figure 5.4. Contour plot of complete energy density function.

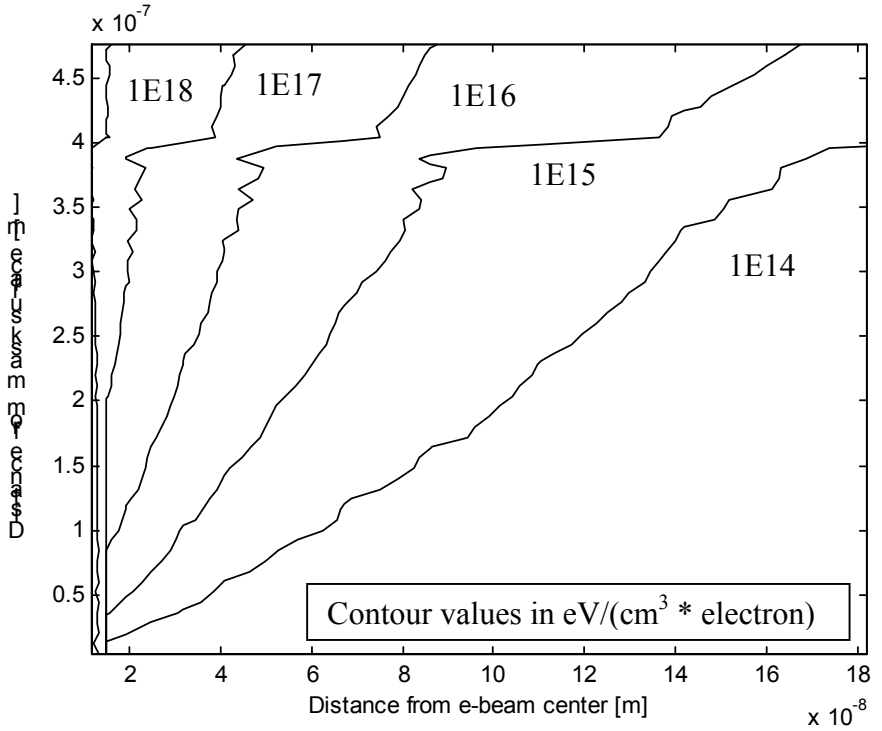


Figure 5.5. Contour plot of PMMA and chrome layer energy density function.

Fig. 3.1. Figure 5.5 shows a contour plot of the energy density function only for the PMMA and the chrome layer. In these layers most of the energy is absorbed very close to the e-beam center. Already within a distance of 40 nm from the center the energy density drops by the order of two magnitudes. With the function for an e-beam point source the mask loading conditions for a square flash have to be calculated. To determine the mask temperature profile with a finite element program the necessary input is the heat generation in the mask caused by the e-beam writing. The time dependence of the energy deposition process is given by the current density of the beam. With the value of the current density and the size of the square flash the charge deposited in the mask per unit time can be calculated. Together with the spatial energy density function that gives values of energy per volume and charge, a heat generation function is determined. The following section describes how the energy density distribution for a square is calculated with the function for a point source. The energy density distribution is later transformed into a heat generation distribution.

The energy density function for square flash writing has to be calculated in a form that is appropriate as an input for the finite element program. The finite element program ANSYS that is used for the calculations allows only one constant value of heat generation per element. For this reason the heat generation rate function for the square flash exposure has to be calculated as a spatial step function with step sizes equal to the size of the finite elements. Since the exposure area on the optical mask is a square, the most appropriate element shape is a cuboid with a top and bottom square side. The side length of the squares has to be small enough to get a sufficient resolution of the mask temperature profile. On the other hand the

number of finite elements of a model should be as small as possible to minimize the necessary computational time.

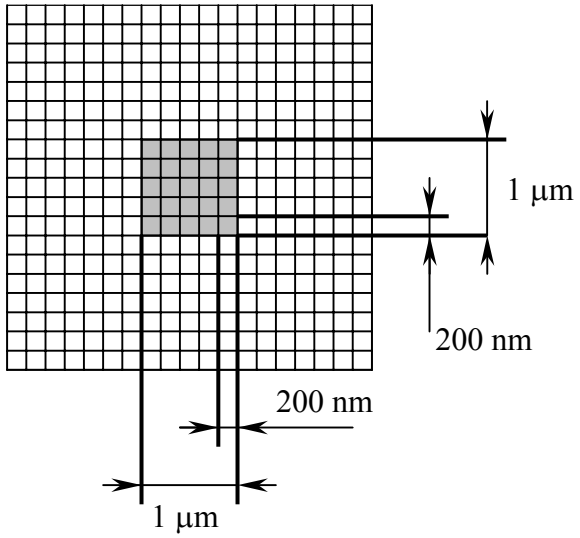


Figure 5.6. Element mesh in x and y-direction.

Figure 5.6 shows the finite element mesh in x and y-direction that is used for the calculation of the mask temperature profile. For an element side length of 200 nm the density of mesh grid points is high enough to get a sufficient resolution of the mask temperature profile in the area of the squared flash where the maximum temperature rise occurs. The lower limit for the thickness of the finite elements is the size of the Δz-increments in the energy density function. The upper limit is given by the fact that the shape of the energy density function has to be approximated with a sufficient accuracy. In Fig. 5.7 the energy density function of a point source is plotted along the center line of the e-beam. In the same figure the function averaged over five depth increments is shown. The averaged function approximates the shape of the original function. Because the highest gradient of the energy density function in z-direction is along the e-beam centerline, it can be assumed that the averaged function gives an accurate result for the complete energy density function in the PMMA and chrome layer. For this reason an element thickness of five times the sub-layer thickness listed in Table 5.2 is used in the PMMA and chrome layer. This gives a thickness of 40 nm for the elements in both materials.

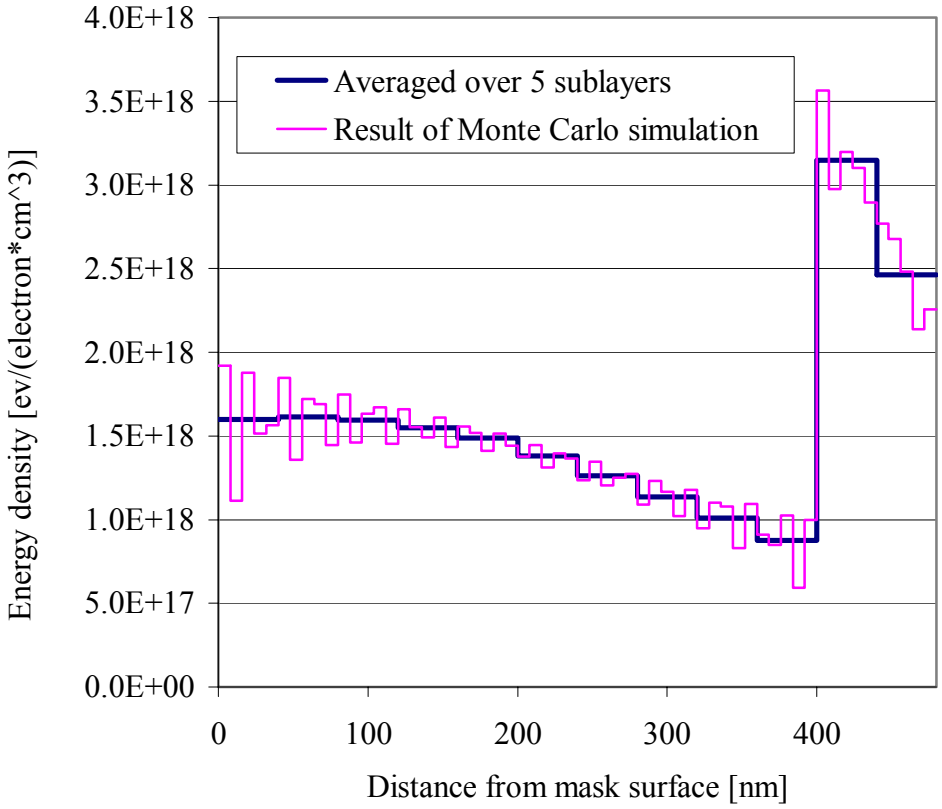


Figure 5.7. Energy density at e-beam center for PMMA and chrome layer.

For the glass layer no averaging of the energy density function over multiple sub-layers is used. The thickness of one sub-layer of 500 nm is sufficiently large. One element layer is used to model each sub-layer of the energy density function.

With the given dimensions of the finite elements the spatial energy density function for square flash exposure can be discretized. The energy density function of squared flash exposure is calculated separately for each layer of finite elements. The procedure of calculating the energy density function for one layer is pictured schematically in Fig. 5.8. The total area of a layer is subdivided into square cells of equal size. For the calculation procedure it is assumed for the cells in the flash exposure area that all the electrons are incident at the center of the cells although they uniformly distributed over the cell areas. Under this assumption the energy deposited in each cell of the layer by the exposure of a single cell of the flash area is calculated. Figure 5.8.a pictures schematically this distribution function. The darkest area in this figure represents one cell of the exposed flash area. The gray shading of the cells around this cell represents qualitatively the amount of energy that is deposited into these cells by the exposure of the dark gray cell in the center. The function describing the energy distribution for the exposure of only one cell of the flash has the same

shape for exposure of each single cell in the flash area. The only difference between the deposition functions of all the cells in the flash area is that the centers of the functions have a different location. For this reason the energy density function for the exposure of the complete flash area can be calculated by a simple superimposition of the functions describing each single cells within the flash area. This can be easily done by just moving the center of the function from cell center to cell center on a path as shown in Fig. 5.8.b. Adding the energy density values in every single cell for each superimposition step gives the energy density function of a square flash exposure. The accuracy of the method is limited by the assumption that within a cell all the electrons are incident at the center. With decreasing the cell size the error of the approximation is reduced. The error of the method can be estimated by calculating the energy density function for a layer with different cell sizes and comparing the results.

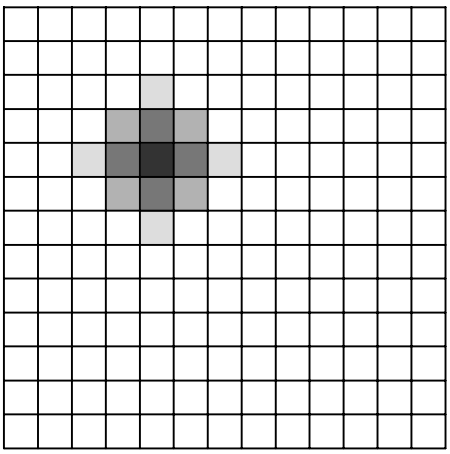


Figure 5.8.a.
Energy density function for exposure of one single cell.

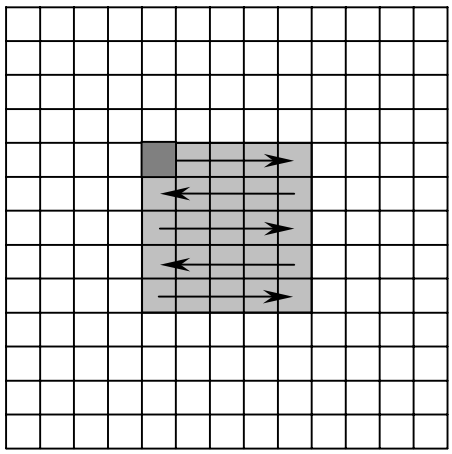


Figure 5.8.b.
Superimposing function for exposure of a single cell over complete flash area.

For a first approximation of the square flash energy density function the cell size is chosen to be equal to the element size of 200nm × 200nm. The procedure of calculating the energy density function for exposure of a single cell within the flash area is pictured in Fig. 5.9. This figure shows a quarter of the exposed cell in the top left corner. All the electrons are assumed to enter at the top left corner of the picture. The amount of energy that is deposited at a specific location in the layer depends on its distance from the location of the entering electrons. This distance, shown as ΔR in Fig. 5.9, gives with the corresponding point source energy density function of the layer the value for the location. To accurately calculate the energy deposited in each cell the cells are subdivided with a finer mesh as depicted in Fig. 5.9. Energy density values are calculated for each small element in a cell with the distance ΔR of its center. The energy density values of all the elements in one cell are averaged to obtain only one value per cell. The mesh within the cells has a small mesh size in all the cells close to the exposed cell shown in the top left corner of Fig. 5.9.

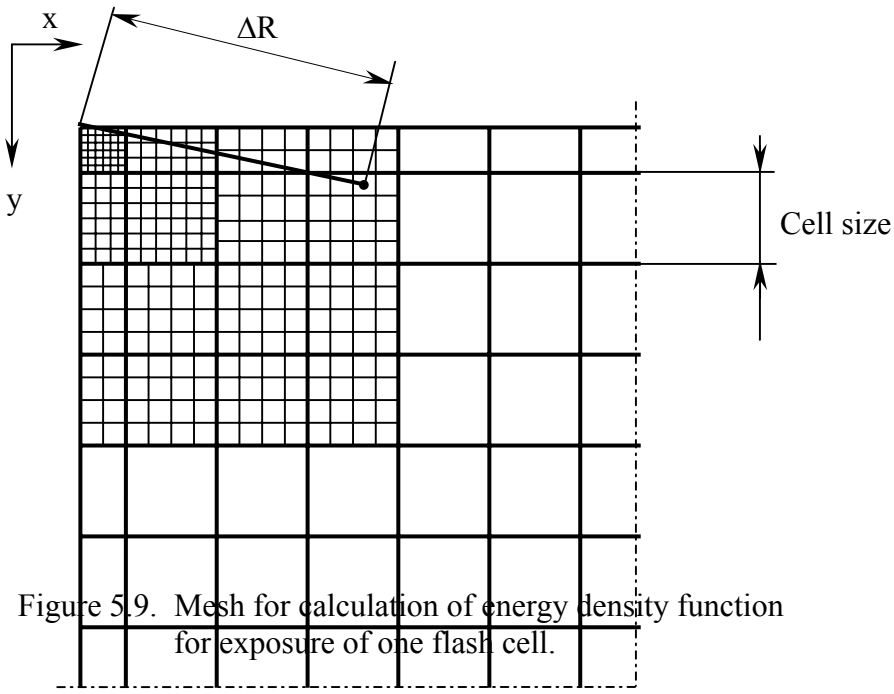


Figure 5.9. Mesh for calculation of energy density function for exposure of one flash cell.

The reason for using this small mesh is that the energy density function for point source exposure has a high slope close to the exposure point. Figure 5.3 shows that the energy density drops by the order of six magnitudes within a distance of 2μm form the beam center. The energy density function for square flash exposure calculated as described above is shown in Fig. 5.10 for the first PMMA layer. The figure shows the energy density function for a layer size of 20μm × 20μm. To estimate the error caused by the simplification that the electrons enter the optical mask only at the center points of the flash cells, the calculation of the energy density function is repeated for a smaller cell size.

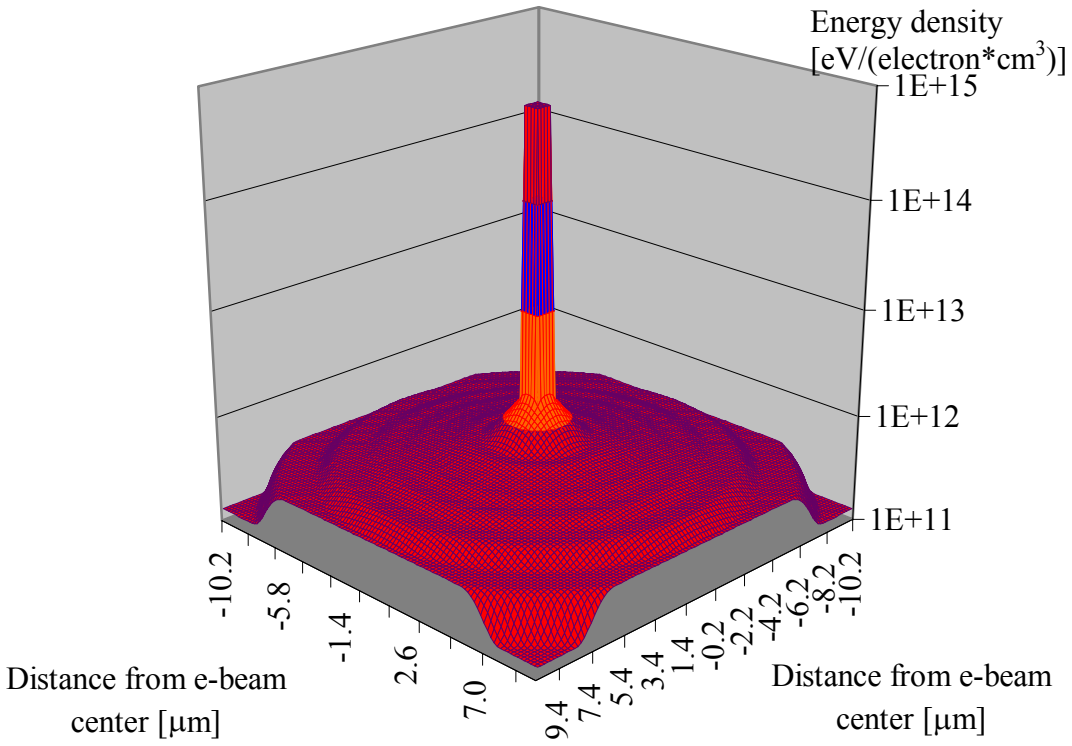
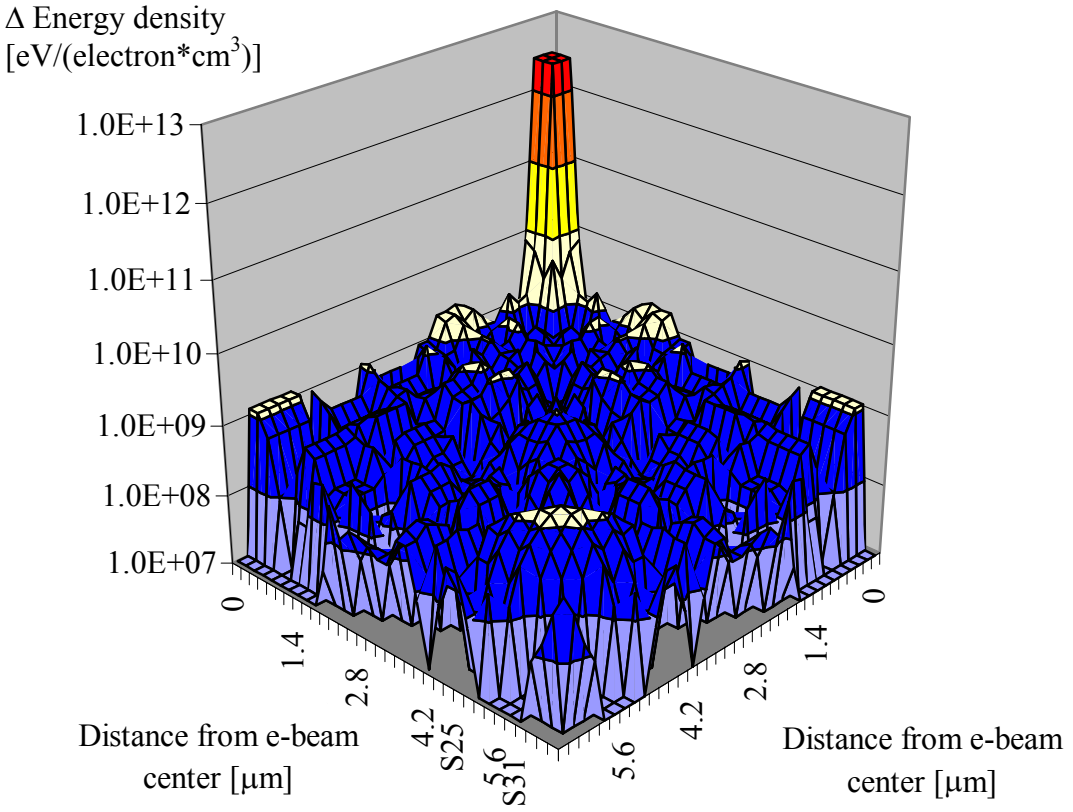


Figure 5.10. Energy density function of first PMMA layer for square flash exposure.

The chosen cell size is 100 nm, what is half of the previously used size. The absolute difference between the energy density functions calculated with a cell size of 100 nm and 200 nm is plotted for the first PMMA layer in Fig. 5.11. The average error is 0.1%, sufficiently small that it is not necessary to make the cell size smaller.

Another accuracy check of the calculated energy density function is the calculation of the energy density function volume integral for each element layer. The calculated value has to be the same as the energy density function volume integral for a point source, because both functions are normalized with the energy deposited by a single electron and the transformation from a point source function to a square flash function changes only the energy distribution in x-y direction.



..... Figure 5.11. Comparison of energy density functions calculated with cell size ..... 100 nm and 200 nm for first PMMA sublayer.

$$\int \text{Energy density of a point source}(r,z)\, dV_{layer} = \sum_{\text{elements in layer}} \text{Energy per element} \qquad [5.2]$$

To obtain an accurate value of the total energy per layer it is important to integrate the square flash energy density function over the complete part of the mask volume where energy is deposited by the electron-beam. In Fig. 5.12 the fraction of the total energy per layer is plotted versus the radius of the cylindrical volume in that the corresponding energy fraction is absorbed. The figure shows curves characteristic for the PMMA, chrome and glass layers.

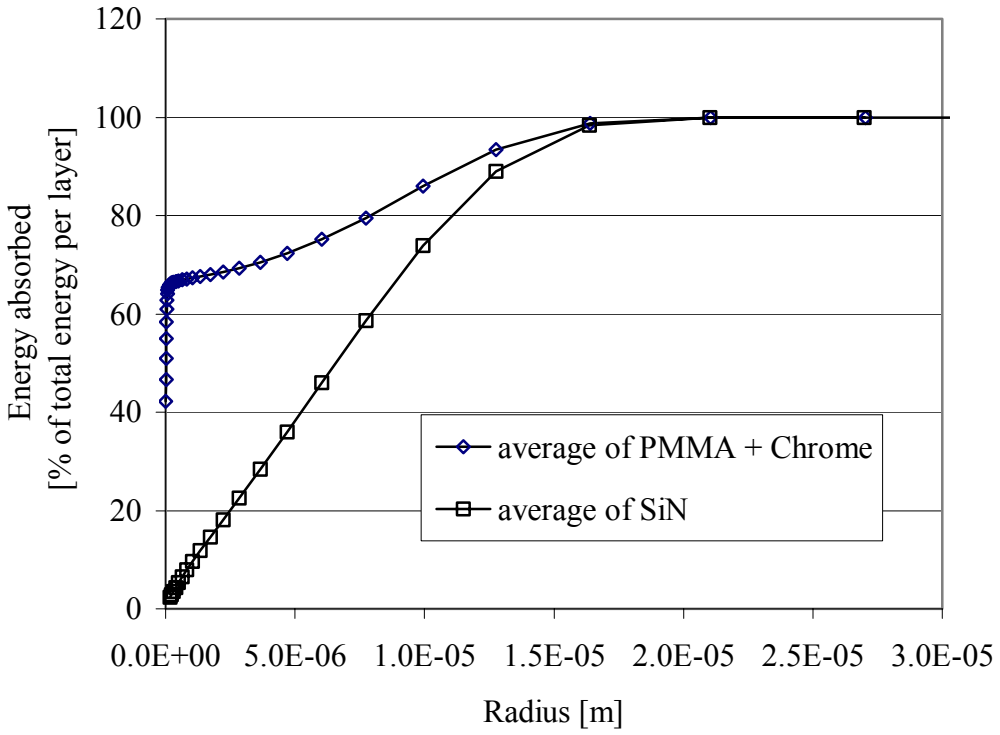


Figure 5.12. Fraction of total energy per layer vs. radius.

In each layer nearly a hundred percent of the energy is deposited by the point source within a volume of 22 μm radius. In the top layers of the optical mask the largest fraction of the energy is deposited very close to the e-beam center. With increasing depth in the material the energy is spread more in the radial direction.

The size of the volume in which one hundred percent of the e-beam point source energy is deposited determines the size of the minimum size of integration volume of the square flash energy density function. The integration volume of the square flash energy density function must be at least 22 μm from the outer edge of the flash to the edge of the integration volume. This means the layer area must have as size at least twice 22 μm plus 1 μm, which is equal to the flash side length. In the finite element model where the layers have the form of a square, this corresponds to an area of 43 μm × 43 μm with 225 by 225 cells for a cell size of 200 nm × 200 nm. The sum of the energies deposited in all of these cells is equal to the total deposited energy.

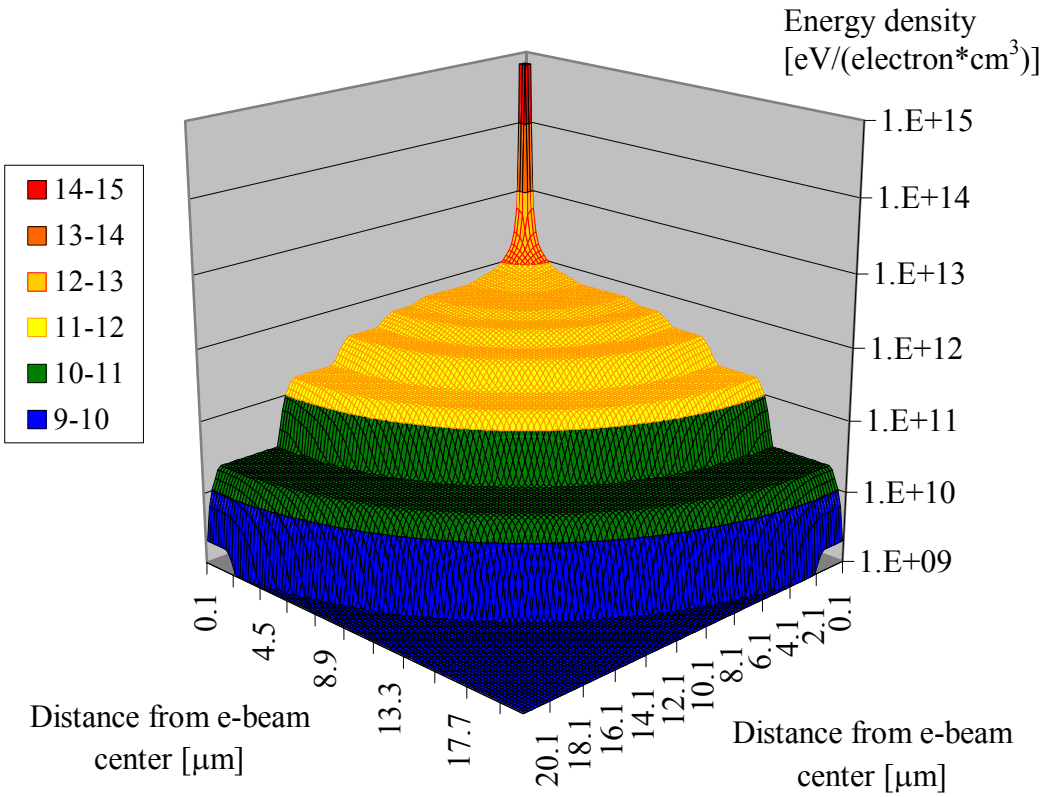


Figure 5.13. Quarter of square flash exposure energy density function for 1. PMMA sub-layer.

The energy density function for square flash exposure calculated for a layer size of 43 µm × 43 µm is plotted in Fig. 5.13 and Fig. 5.14. In the plot for the first PMMA layer the shape of the square exposure area can be clearly identified. The energy density function of the glass layer does not show the contours of the beam writing area asa peak, because of the radial spread of the electrons with increasing depth in the mask material. Characteristic for the energy density distribution functions of all layers is that there are areas of constant energy densities. These areas are visible as plateaus in Fig. 5.13 and Fig. 5.14.

Energy density
[eV/(electron*cm^3)]

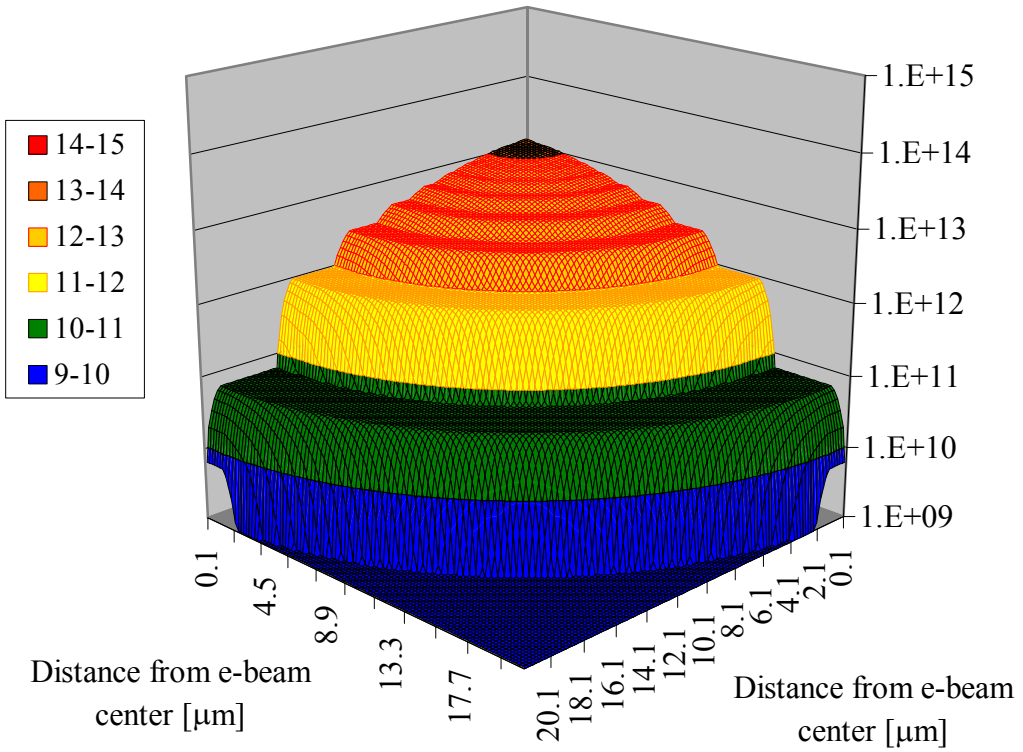


Figure 5.14. Quarter of square flash exposure energy density function for the 12. glass sub-layer.

The reason for this areas of constant energy density is that the square flash exposure energy density function is calculated with the point source exposure function that is given as a step function. The smooth transition from one level of energy density to the next lower or higher level occurs because of the superimposition of a step function as shown in Fig. 5.8.b. The error of the total energy per layer is determined as in equation 5.3 by calculating the difference between the volume integrals in equation 5.3.1 and 5.3.2.

$$\text{Error}_{\text{Energy density}} = \frac{(\text{Energy per layer}_{\text{Point source}} - \text{Energy per layer}_{\text{square}})}{\text{Energy per layer}_{\text{Point source}}} \cdot 100 \qquad [5.3]$$

$$\text{Energy per layer}_{\text{Point source}} = \int \text{Energy density of a point source}(r, z)\, dV_{\text{layer}} \qquad [5.3.1]$$

$$\text{Energy per layer}_{\text{Square}} = \sum_{\text{Elements in layer}} \text{Energy per element} \qquad [5.3.2]$$

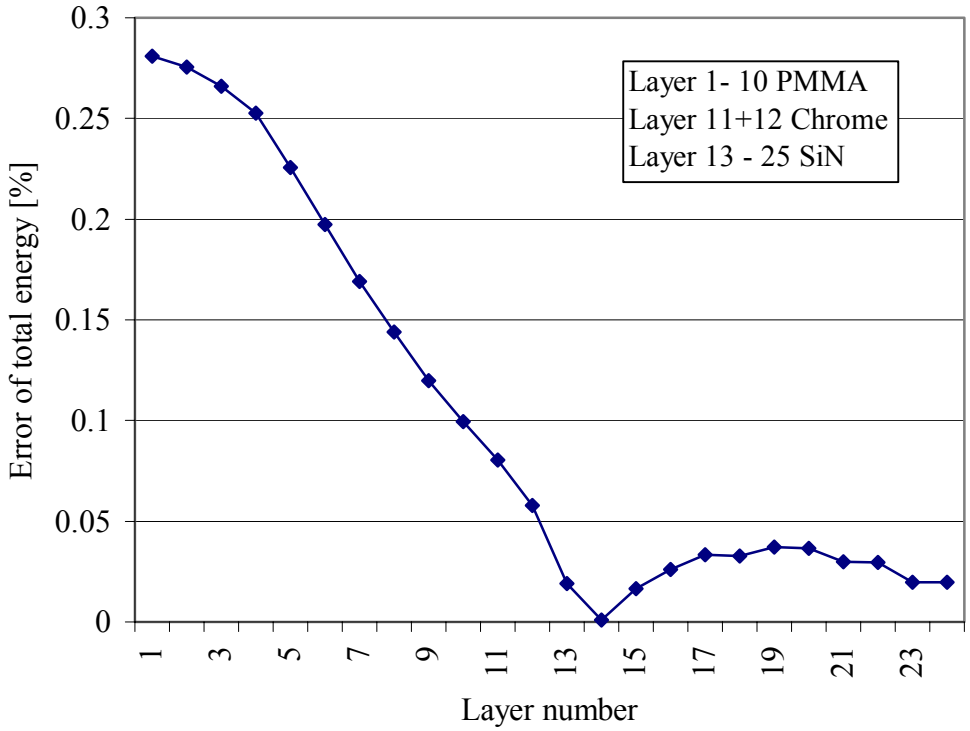


Figure 5.15. Error of total energy per layer for square flash exposure compared to point source exposure.

This percentage error in the total energy per layer is plotted in Fig. 5.15. With a maximum difference of 0.28 percent between the total energy per layer for square flash exposure and point source exposure, the energy error is sufficiently small. Thus, it can be assumed that the energy density function for square flash exposure is approximated with an acceptable accuracy.

5.3 The finite element model of the optical mask

In this section the finite element model is described that is used for the calculation of the optical mask temperature profile during electron-beam writing of a 1 μm × 1 μm square flash. The dimensions of the optical mask are shown in Fig. 5.1.
The optimal size of the finite elements used in the model is estimated in Section 5.2. The dimensions of the finite elements are summarized in Table 5.3.

Material	Δz [nm]	Δx [nm]	Δy [nm]	Number of Element layers
PMMA	40	200	200	10
Chrome	40	200	200	2
Glass	500	200	200	50

Table 5.3. Dimensions of cuboid finite elements and number of layers in each material.

The x-y dimensions of the mask volume where energy is deposited by the electron-beam are determined in Section 5.2. The maximum depth where e-beam energy is deposited is given by the heat generation function. To enclose the complete volume of energy deposition in a finite element model of cuboid shape the dimensions of this model must be at least 43 μm × 43 μm × 25.5 μm. Assuming this model dimensions and the element dimensions in Table 5.3 the estimated number of elements is 2,865,950. The transient mask temperature profile calculation with a finite element model of this size would not be possible in a reasonable computational time. For this reason the dimensions of the model have to be reduced if possible. In the following section of this chapter it is examined how a reduction of the model size effects the accuracy of the temperature profile calculation. Because the purpose of the calculations is to determine the mask temperature rise causing resist sensitivity changes due to local direct heating, only the temperature profile in the PMMA layer is of interest. The accuracy of the temperature profile calculations has to be optimized only for a small area around the written flash, where the maximum temperature rise occurs and therefore significant changes of resist sensitivity are expected. With this criteria for the temperature profile accuracy the dimensions Δx, Δy and Δz of the model volume have to be determined. Figure 5.16 shows the volume of e-beam energy deposition and enclosed in it the volume of dimensions Δx × Δy × Δz. The actual model volume dimensions Δx, Δy, Δz are

not necessarily smaller than the dimensions of the energy deposition volume as it is pictured in Fig. 5.16. The optimum of the dimensions Δx and Δy is estimated separately for each mask material layer as described below.

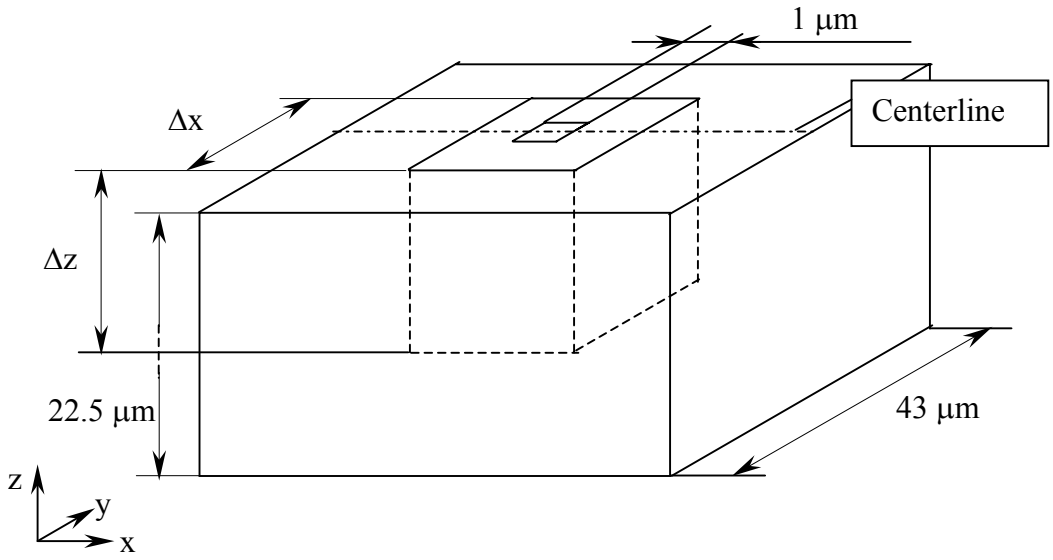


Figure 5.16. Finite element model dimensions.

The Δz dimension of the model is determined by repeating the mask temperature calculations with different values for Δz. The magnitude of the change in the surface temperature profile with increasing Δz gives the information to determine if it is necessary to increase Δz further.
First the optimum model dimensions of the PMMA layer are determined. If Δx and Δy are smaller than 43 μm, some of the e-beam energy deposited in the mask is not accounted for. To estimate the effect of cutting off a part of the energy deposition function the energy density values are related to the temperature rise that would occur in a single element under the assumption of zero conductivity.

$$C_{p,PMMA} \cdot m_{element} \cdot \Delta T = \text{Energy deposited} \qquad [5.4]$$

The temperature rise is calculated with equation 5.4 for the energy that is deposited in the typical flash writing time. The typical flash writing time is between 250 nsec and 1 μsec. The temperature rise is determined for the boundary values of this time interval. Figure 5.17 shows a plot of the calculated temperature rise for energy density values that

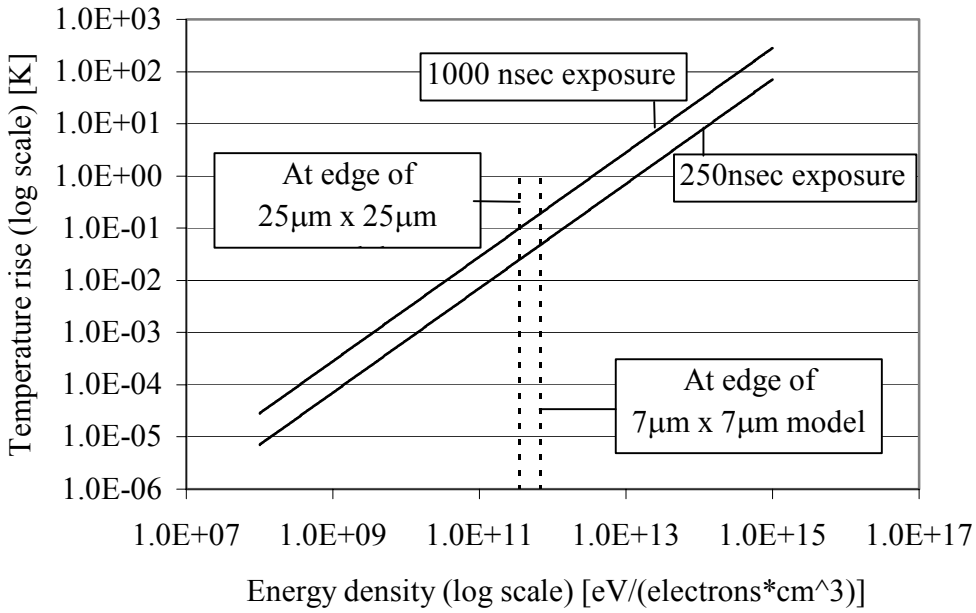


Figure 5.17. Temperature rise of a single element of PMMA vs. energy density of the element for writing times of 250 nsec and 1 µsec.

occur in the PMMA layer. The highest values of energy density in the PMMA layer correspond to a temperature rise of about 140 °C for a writing time of 1µsec. If the modeled mask area of the PMMA layer where only 25 µm × 25 µm the temperature rise in the elements at the edges of this model would be about 0.1 °C. Choosing the model size only 7 µm × 7 µm relates to a temperature rise at the edges of 0.3 °C. The temperature difference of 0.2 °C between the model edge for a 7 µm × 7 µm model size compared to a 25 µm × 25 µm model is small. It can be expected that there is no significant heat flux caused by a temperature gradient that small. The temperature rise of each element calculated with equation 5.4 neglects conduction in the PMMA layer. Because there are steep gradients in the energy density a the edge of the flash area as it is shown in Fig. 5.10 for the first PMMA layer, significant temperature gradients can be expected in this region of the mask. For this reason it is important to look at the energy transport by conduction in the PMMA layer. Since in the PMMA layer most of the energy is deposited in the flash area it is examined how fast the energy diffuses from this area to the edges of the modeled PMMA volume. Because the gradient of deposited energy in z-direction is small in the photoresist as it pictured in Fig. 5.7, it is enough to look at the energy diffusion in the PMMA sublayer with the highest energy deposition. To examine the transient conduction effects on the temperature

profile of this sublayer a heat generation load corresponding to the function in Fig. 5.13 is applied only to the flash area. The temperature profile is calculated for a model size of 25 μm × 25 μm and a writing time of 1 μsec. In Fig. 5.18 the temperature in the PMMA is plotted along the centerline marked in Fig. 5.16. The same temperature profile is plotted in Fig. 5.19 with a logarithmic temperature scale.

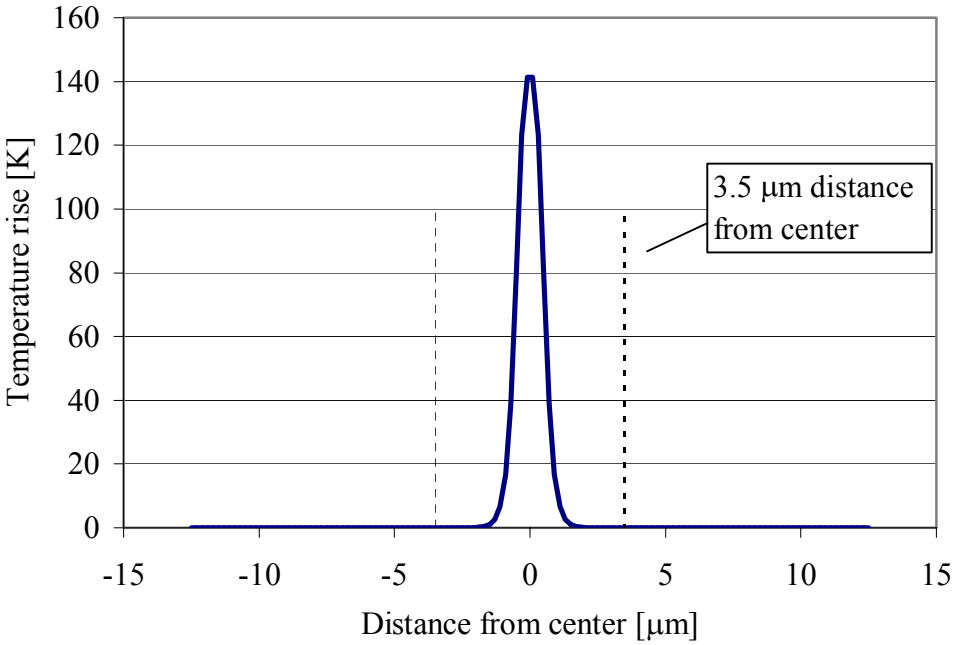


Figure 5.18. PMMA temperature rise in first sublayer for heat generation load applied only to flash area plotted along mask centerline. The writing time is 1μsec.

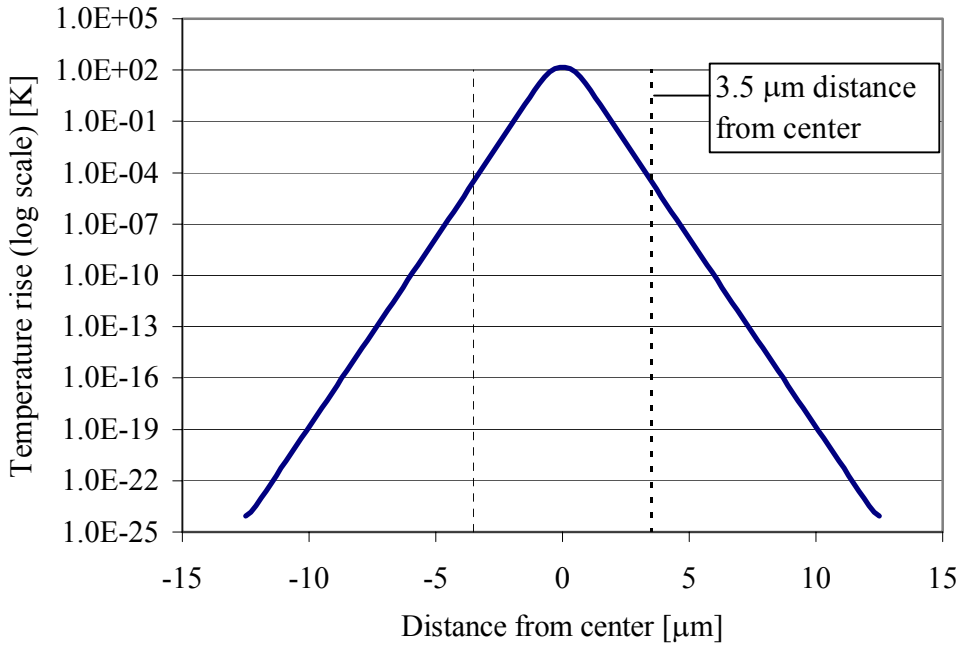


Figure 5.19. PMMA temperature rise in first sublayer for heat generation load applied only to flash area plotted along mask centerline. The writing time is 1μsec.

Even at a distance of 3.5 μm from the center, the temperature rise is only 1E-4°C after a writing time of 1 μsec. This means that for a model size of only 7 μm × 7 μm a temperature rise on the order of 0.0001 °C can be expected at the edges. It has to be further examined if this small model size is sufficient to accurately predict the temperature rise in the PMMA layer. A heat generation load is applied to the complete sublayer area as plotted in Fig 5.13 and not only to the 1 μm × 1 μm flash area as it is done to calculate the temperature profile in Fig. 5.17. The temperature profile is calculated for a model size of 7 μm × 7 μm and 25 μm × 25 μm. The difference between the two profiles along the centerline is plotted in Fig. 5.20.

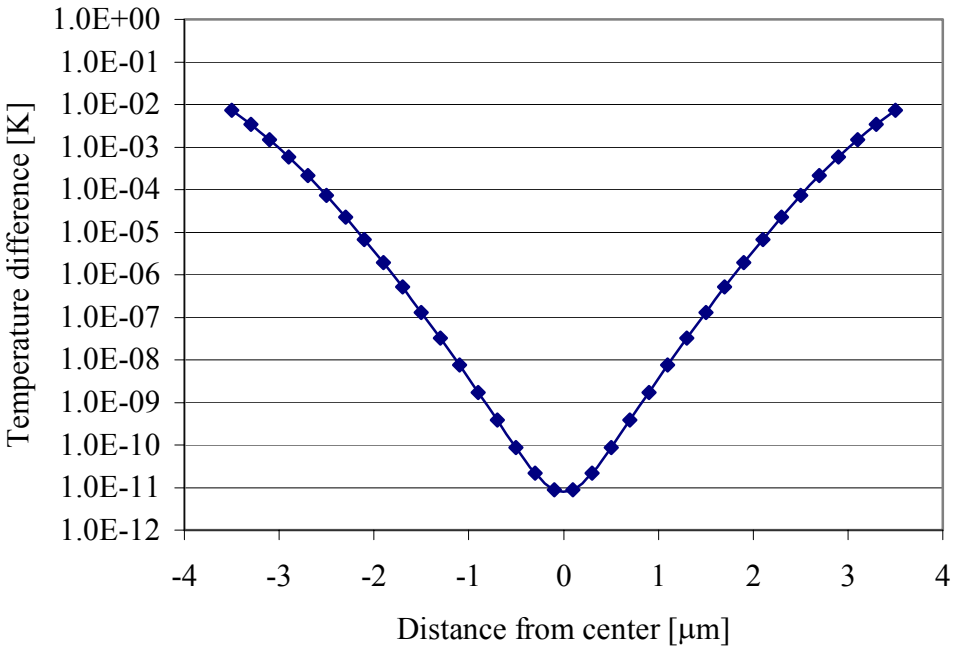


Figure 5.20. Difference in PMMA temperature rise in first sublayer between model size of 7 μm × 7 μm and 25 μm × 25 μm plotted along mask centerline. The heat generation loads are applied to complete model volume. The writing time is 1μsec.

A temperature difference of only 1E-2 °C at the edge of the model is small enough that for the PMMA layer Δx and Δy dimensions of 7 μm can be used for accurate modeling.

The temperature profiles plotted in Fig. 5.19, Fig. 5.21 and Fig. 5.21 are calculated with an ANSYS finite element model with a single layer of cuboid elements. To check the accuracy of the ANSYS calculations the results of ANSYS are compared to results calculated with a two dimensional Crank-Nicolson finite difference code for the case of a heat generation load applied only to the flash area.

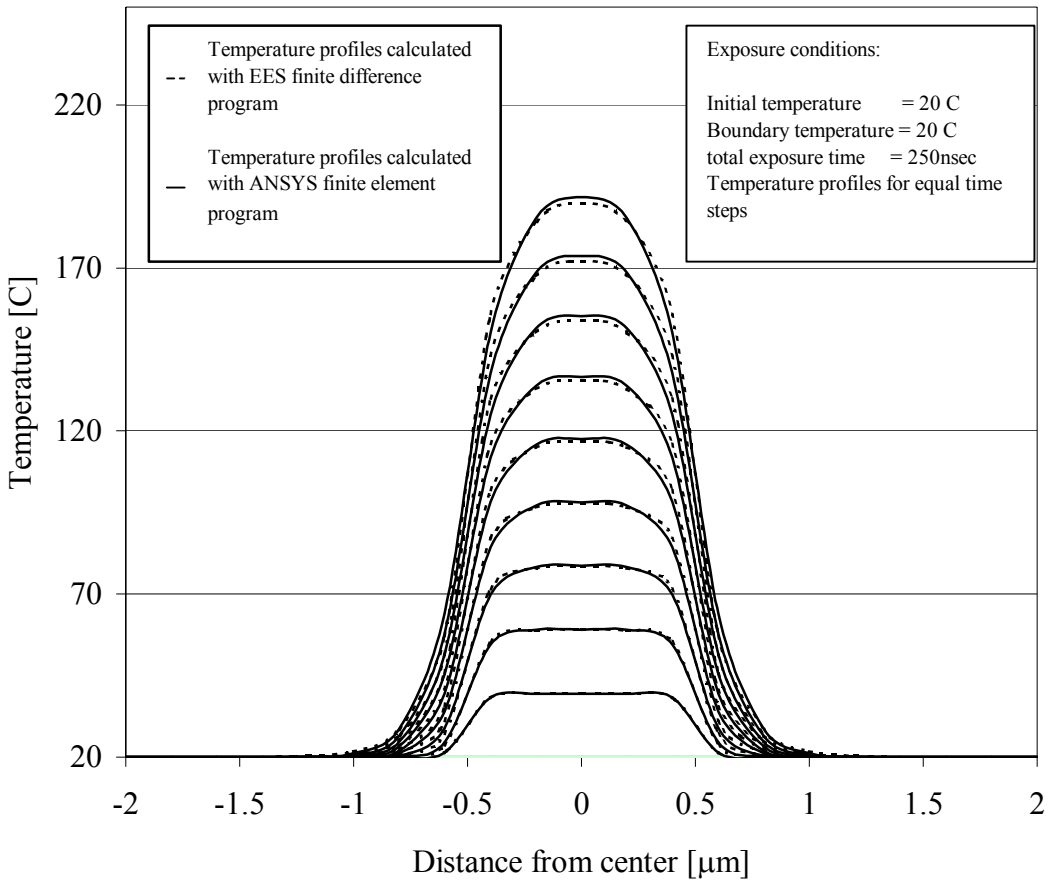


Figure 5.21. Temperature profile of a 4 μm × 4 μm PMMA sublayer for 250 nsec exposure of a 1 μm × 1 μm with 1E-3 W/μm^3. The temperature profile is plotted along the mask center line.

The heat generation of 10^{-3} W/μm^3 is uniform over the 1 μm × 1 μm flash area. The modeled sublayer area has a size of 4 μm × 4 μm and the temperature profile was calculated for a writing time of 250 nsec with isothermal boundary conditions at the edges of the model. In Fig. 5.21 the temperature profile calculated with ANSYS and the finite difference program is plotted along the mask centerline for nine equal time intervals from the beginning until the end of writing. The results of ANSYS and the finite difference code show good agreement.

For the chrome layer a similar procedure as for the PMMA layer is used to determine the optimal model dimensions Δx and Δy. Figure 5.22 is the equivalent of Fig.5.17 for the chrome layer. Because chrome has a high thermal conductivity compared to the polymer PMMA, the temperature rise prediction at certain locations in the layer with Fig. 5.22 is not very accurate.

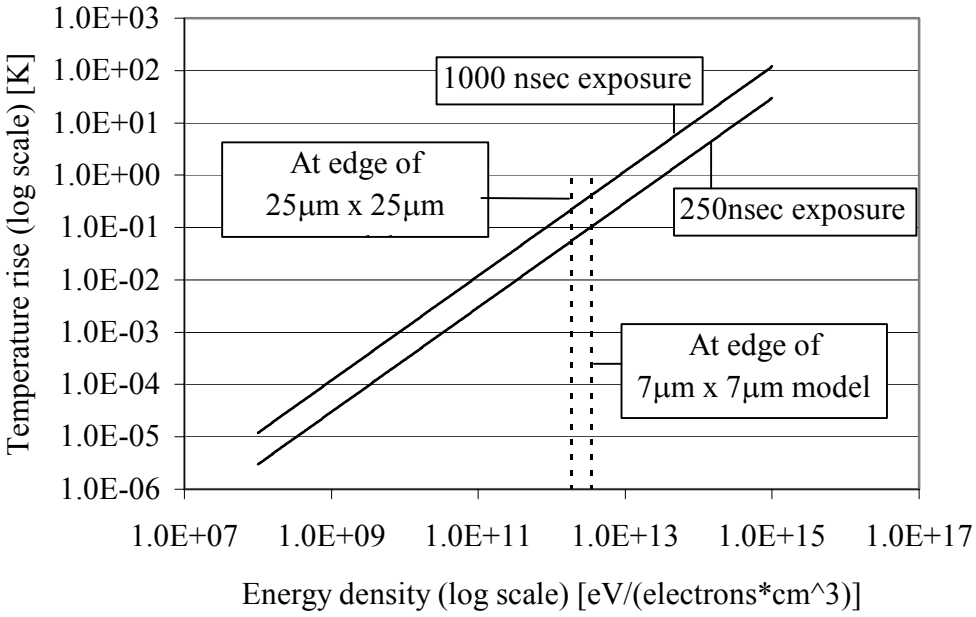


Figure 5.22. Temperature rise of a single element of chrome vs. energy density of the element for writing times of 250 nsec and 1 μsec.

The curves in Fig. 5.22 only relate the temperature rise of a single element with the heat generation that is applied to it for a specific time. Because of the high conductivity of chrome the limit for the dimensions Δx and Δy is controlled by the speed of heat diffusion from the flash area to the edge of the model. Figure 5.22 shows that cutting off parts of the heat generation function by choosing the dimensions Δx and Δy smaller than 43 μm does not appreciable change the temperature rise at the edge of the model. The temperature difference

of 0.1 °C between the model edge for a 7 μm × 7 μm model size compared to a 25 μm × 25 μm model is small.

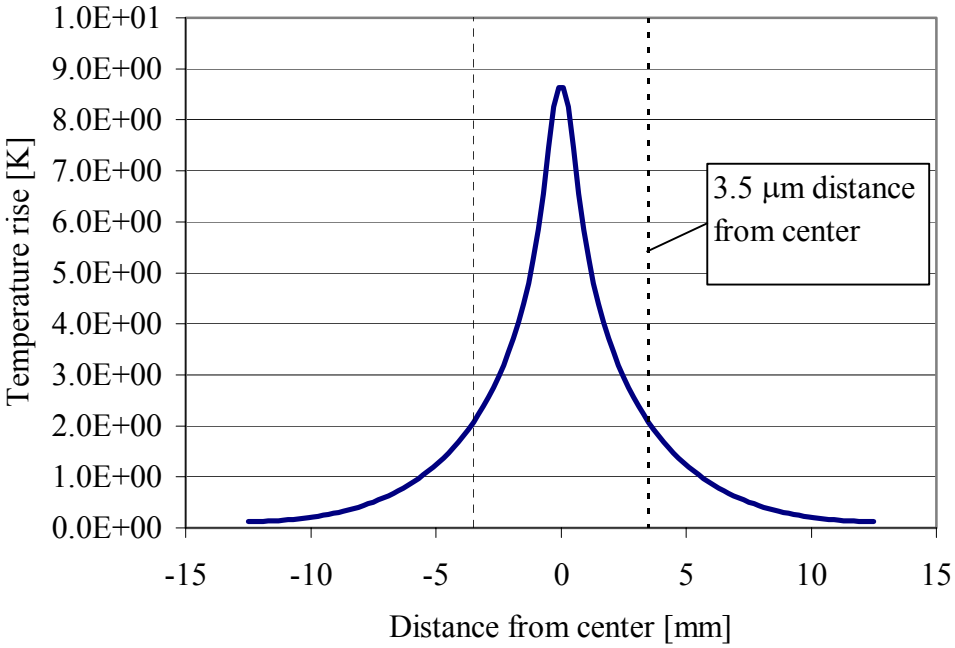


Figure 5.23. Chrome temperature rise in first sublayer for heat generation load applied only to flash area plotted along mask centerline. The writing time is 1μsec.

Figure 5.23 and Fig. 5.24 show the temperature rise in the chrome layer with maximum heat deposition plotted along the mask center line for a heat generation load applied only to the flash area. These figures correspond to Fig. 5.18 and Fig. 5.19 for the PMMA layer. The shape of the temperature rise function along the centerline for the chrome layer compared to the plot for the PMMA layer indicates the faster heat diffusion from the flash area to the edges of the modeled volume. The area where a significant temperature rise occurs due to the heat deposition in the flash area is much larger in the chrome layer. Figure 5.24 shows that within the writing time of 1 μsec the temperature rise at the edges of a 25 μm × 25 μm

model is 0.2 °C but the maximum rise is only 9 °C. This means that an amount of energy diffuses to the edges that can not be neglected. It is probable that the accuracy of the surface temperature profile would effected by choosing a size of only 25 μm × 25 μm for the chrome layer. In Fig. 5.25 the temperature profile for the same exposure case as for Fig. 5.23 and Fig. 5.24 is compared for a 7 μm × 7 μm and 25 μm × 25 μm model size.

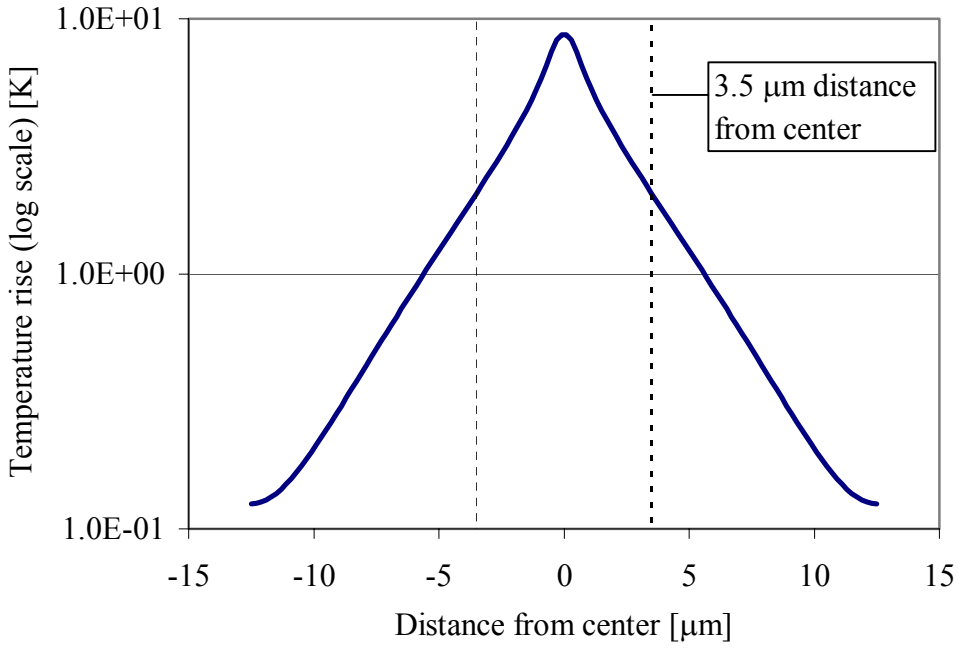


Figure 5.24. Chrome temperature rise in first sublayer plotted along mask centerline for heat generation load applied only to flash area. The writing time is 1μsec.

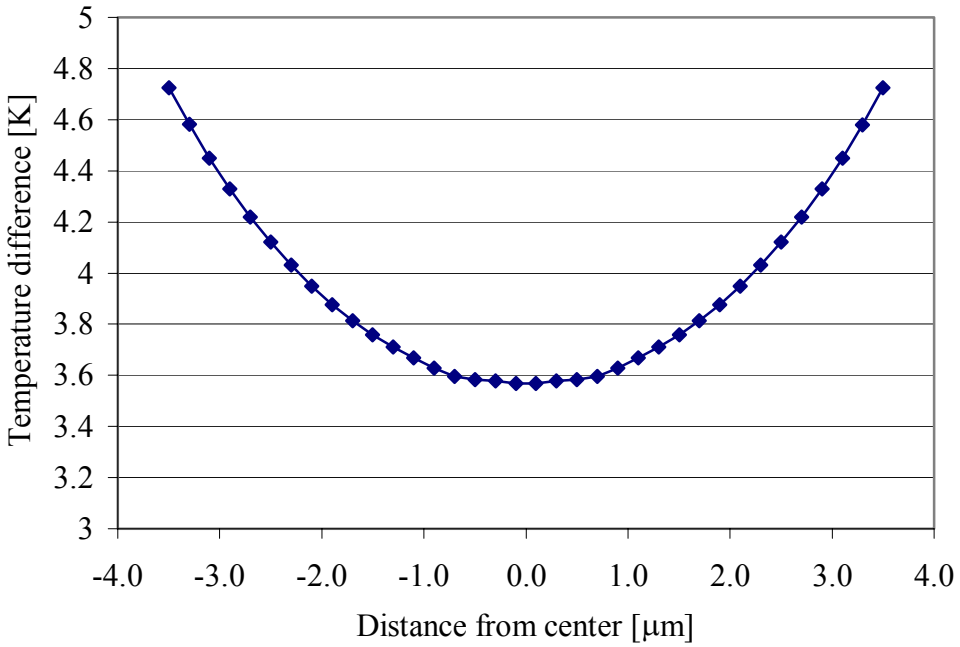


Figure 5.25. Difference in chrome temperature rise in first sublayer between model size of 7 μm × 7 μm and 25 μm × 25 μm plotted along mask centerline. The heat generation loads are applied to the complete model volume. The writing time is 1μsec.

The temperature disagreement of 3.559 °C is too big, so that a model size of 7 μm × 7 μm could not be used. For the chrome layer even a size of 25 μm × 25 μm is to small to model the chrome layer accurately. To get an accurate prediction of the temperature profile for the chrome layer the dimensions of this material layer have to be increased further. For the complete model the dimensions Δx and Δy should be the same for each material layer. For this reason two types of finite elements are used to model the chrome layer. In a 7 μm × 7 μm center region cuboid elements are used. With Δx and Δy dimensions that have the same values as for the PMMA layer. The center region of the cuboid elements in the chrome layer is connected to an area of shell elements. In ANSYS shell elements are defined as an area that is associated with a thickness parameter.

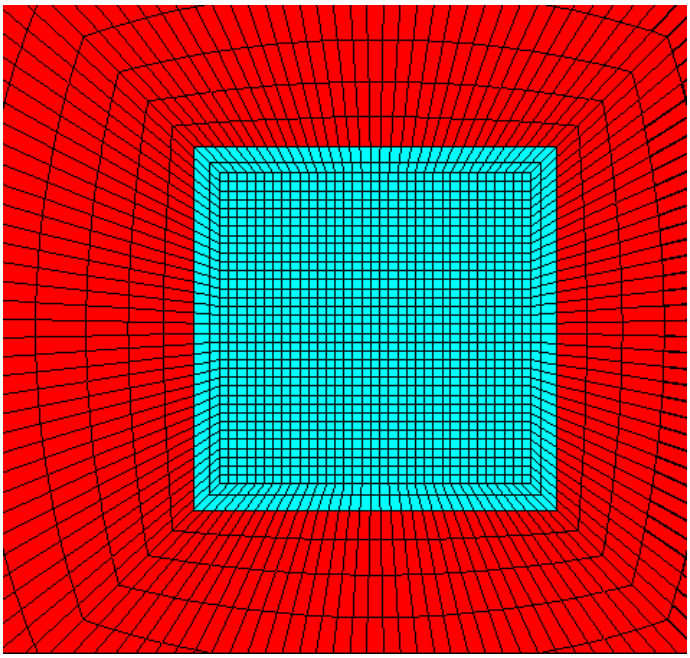


Figure 5.26. Finite elements in the chrome layer. The turquoise elements are solid elements with cuboid shape and the red elements are shell elements.

With these kind of elements two dimensional energy transport can be simulated. The advantage of using shell elements compared to cuboid elements is that the results calculated with shell elements are more accurate for a model with elements that have a large aspect ratio, the ratio of the dimensions of the side length of a single element. Figure 5.26 shows a plot of the elements in the chrome layer. The turquoise colored elements are the cuboid elements of the 7 μm × 7 μm center region and the red elements are shell elements in a circular area around the center region. It is of advantage to use shell elements in the red area, because a lower element density is necessary with this type of element. Shell elements permit the use of large element dimensions in x-y direction compared to the thickness of the chrome layer of only 80 nm. Another advantage of using shell elements is that only four nodes are required to define one element of this type. To define an element of cuboid shape

a minimum of eight nodes is necessary. For these reasons modeling a part of the chrome layer with shell elements decreases the number of necessary nodes compared to the necessary number of cuboid elements. It is sufficiently accurate to model the chrome layer outside the center region as a two dimensional problem, because big temperature gradients in depth direction occur only in the center region of the mask. In Fig. 5.27 the temperature profile of the chrome layer is plotted along the mask centerline for a heat generation load applied only to the flash area. Because the temperature rise at the edge of the model is only 1E-18°C, the model of the chrome layer can be considered to have sufficient size to give accurate temperature profile results.

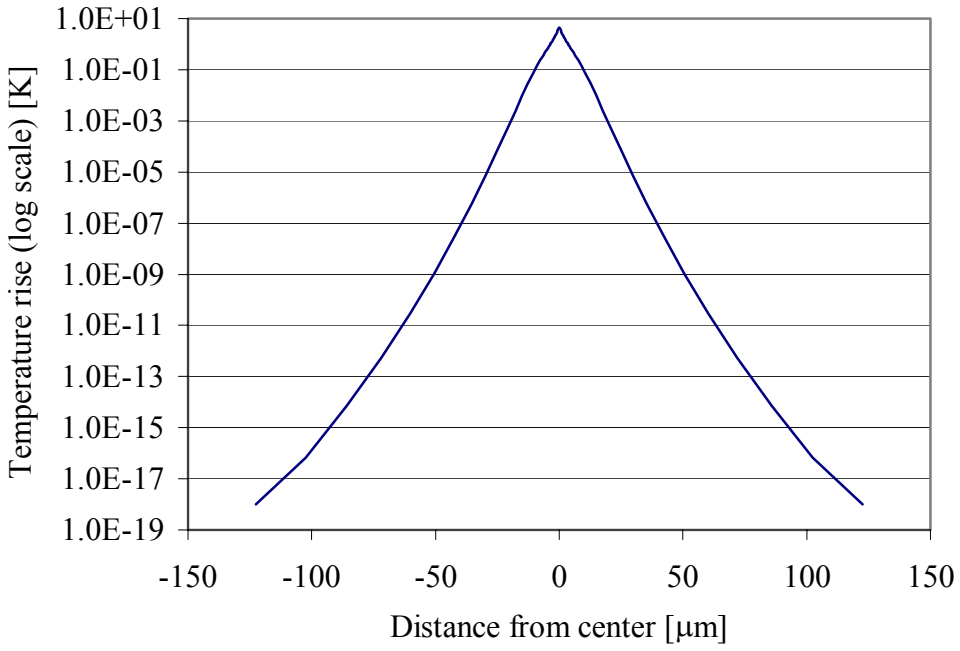


Figure 5.27. Chrome temperature rise in first sublayer for heat generation load applied only to flash area plotted along mask centerline The writing time is 1μsec. The chrome layer is modeled with cuboid elements in a 7 μm × 7 μm center region and a circular area of 125 μm with shell elements.

For the glass sublayer with maximum heat generation in the flash area the temperature rise is plotted along the mask centerline in Fig. 5.28. Heat

generation loads are only applied to the flash area of the 25 μm × 25 μm modeled glass sublayer. For a writing time of 1μsec the temperature rise at the edge of the 25 μm × 25 μm model is only 1E-14 °C. Because the Δx and Δy dimensions of the PMMA and chrome solid element layers are chosen to be 7 μm, these dimensions should be the same for the glass layer. The temperature rise at the edges of a 7 μm × 7 μm glass area predicted with Fig. 5.28 is around 0.03 °C for a maximum temperature rise of 48 °C in the glass layer. This indicates that the amount of energy diffusing to the edges of the 7 μm × 7 μm model within the writing time is small.

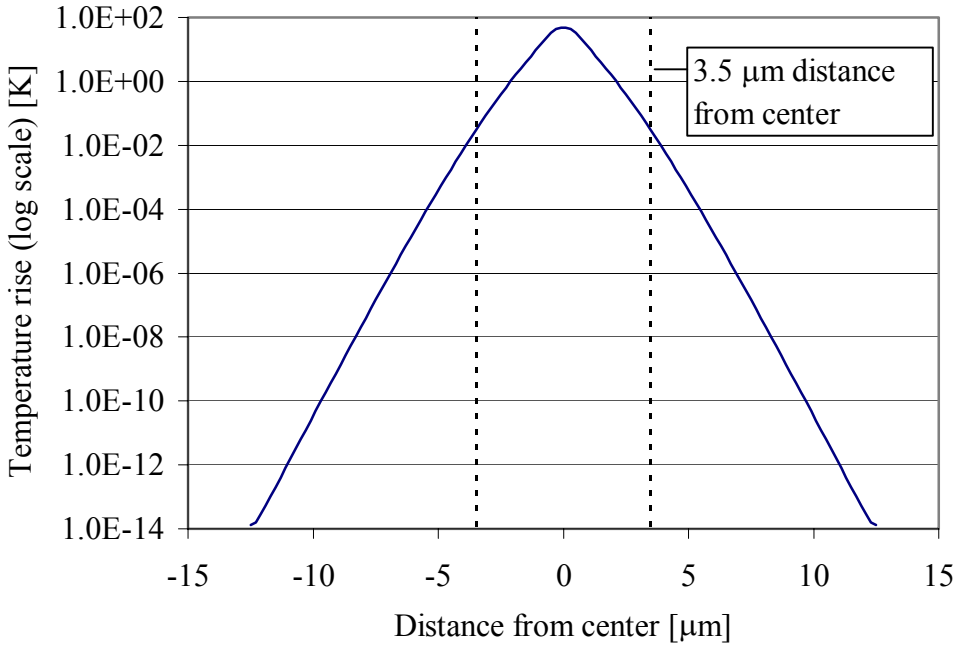


Figure 5.28. Glass temperature rise in first sublayer plotted along mask centerline for heat generation load applied only to flash area. The writing time is 1μsec.

In Fig. 5.29 the temperature profile for a heat generation load applied to the complete sublayer area is compared for a 7 μm × 7 μm and 25 μm × 25 μm model size. The temperature difference at a distance of 3.5 μm from the flash center seems to be big considering that Fig. 5.28 predicts that almost no energy diffuses from the flash area to this location within the writing time. The reason for the high temperature difference is that the

heat generation is distributed more evenly in the glass sublayer than in the PMMA and chrome layer. Therefore estimating the amount of energy reaching the edge of the 7 μm × 7 μm area with Fig. 5.28 does not give a good approximation to the actual value for a heat generation applied to the total volume. By only taking into account the energy deposited in the flash area a significant amount of energy is neglected.

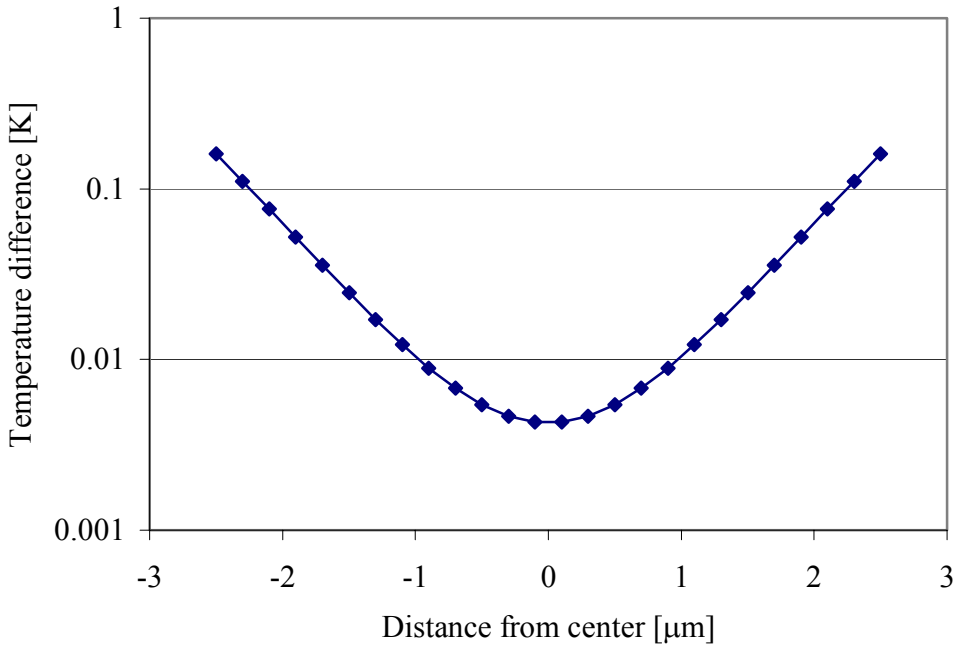


Figure 5.29. Difference in glass temperature rise in first sublayer between model size of 7 μm × 7 μm and 25 μm × 25 μm plotted along mask centerline. The heat generation loads are applied to the complete model volume. The writing time is 1μsec.

It can be expected that using Δx and Δy dimensions of only 7μm for the glass layer is the limiting factor influencing the accuracy of the optical mask model. Figure 5.20 for the PMMA layer and Fig. 5.28 for the chrome layer predict that the error caused by limiting the size of these materials solid element layers to 7 μm × 7 μm is smaller than for the glass layer.
Taking into account that as a final result only the temperature profile at the top surface of the photoresist in a small region around the flash area is of interest, the accuracy of the temperature profile calculation in the glass is considered as good enough.

After determining the dimensions Δx and Δy of the finite element model the only model dimension that needs to be set is the thickness Δz of the complete modeled volume as shown in Fig 5.16. To determine this model parameter the surface temperature profile of the mask is calculated for different values of Δz with a heat generation load applied to the complete model volume. The heat generation load corresponds to a square flash exposure and it is applied to the mask for a writing time of 1 μsec. The surface temperature profile is calculated for Δz values of 11.48 μm, 16.48 μm and 20.48 μm. The results for the different thicknesses are compared in Fig 5.30. Increasing the total thickness from 11.48 μm to 16.48 μm results in a maximum change of 0.0075 °C in the surface temperature profile. Increasing the thickness further to 20.48μm causes a maximum change of only 0.001614 °C in the surface temperature profile. These very small changes in the surface temperature profiles with varying thickness Δz show that cutting off parts of the glass material in the depth direction does not effect the accuracy of the surface temperature calculations for the used range of Δz values. For the final calculations of the mask temperature profile a total thickness of 16.48 μm is used.

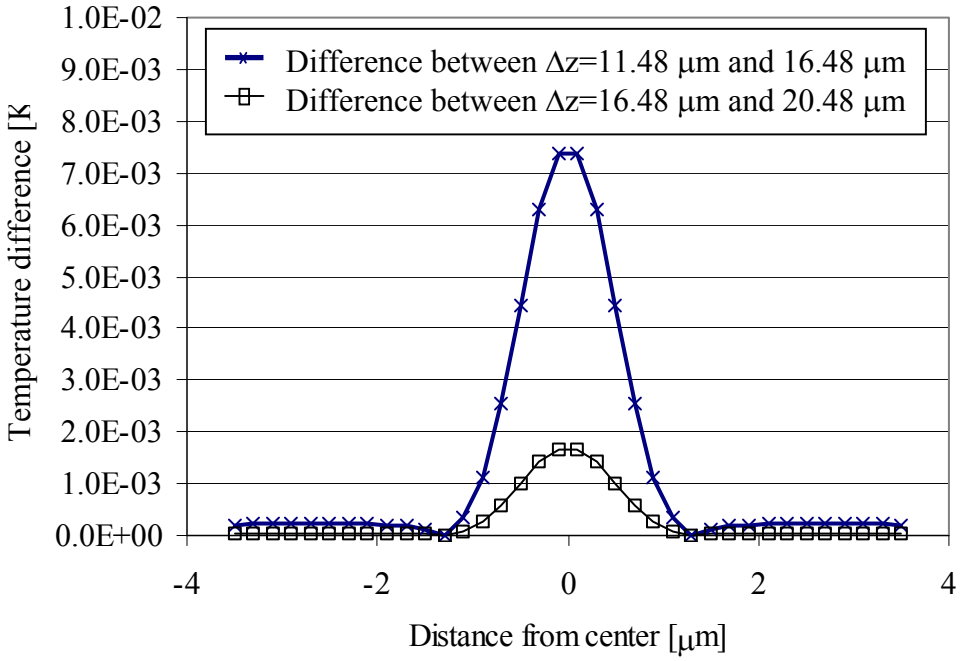


Figure 5.30. Difference in mask surface temperature profile between models with different total thickness Δz plotted along mask centerline.

The heat generation loads are applied to the complete model volume. The writing time is 1μsec.

5.4 The optical mask temperature profile for e-beam writing of a 1 μm × 1 μm square flash

In this section the calculated results of the optical mask temperature profile calculations for square flash are presented. Mask temperature profiles are shown for various e-beam writing times in the time interval from 250 nsec to 1 μsec, the typical range for writing times of a 1 μm × 1 μm square flash. The ANSYS finite model that is used for the calculations is described in detail in Section 5.2. The total number of nodes in this model is 51656 and the number of elements is 47950. A heat generation load, corresponding to the energy distribution function for square flash writing, is applied to each of these elements. The

computational time for the temperature profile calculation at one specific writing time is about three hours using a 200 MHz Pentium four processor system with 1Gigabyte of RAM.

The surface temperature profile of the optical mask that is of main interest for resist heating effects is plotted in Fig. 5.31.a to Fig. 5.31.d for different square flash writing times. The maximum temperature rise is 44.2 °C for a 250 nsec writing time, 72.2 °C for 500 nsec writing time, 88.0 °C for 750 nsec writing time and 97.6 °C for 1000 nsec writing time. In Fig. 5.32.a and Fig. 5.32.b the mask temperature profile in depth is plotted for a writing time of 250 nsec and 1000 nsec. The cut shown in these two figures is along the mask centerline. The square flash area is located in the middle of the figures and has a size of five elements in the horizontal direction. The PMMA material is modeled by ten element layers in vertical direction shown at top of the plots. The following two element layers in the vertical direction are chrome layers. All the following element layers represent the bulk glass material. In both plots in can be clearly seen that that maximum temperature rise occurs in the PMMA layer due to a high local heat generation in the material and the low diffusivity of the polymer. Compared to the PMMA and glass layer there is only a negligible temperature rise in the chrome. Because of the high conductivity of chrome the energy is conducted rapidly out of the flash center region where the maximum heat generation occurs. Although the local heat generation is much smaller in the glass than in the chrome and PMMA, there is a significant temperature rise of 55 °C in the glass material. The reason for this high temperature is the relatively low diffusivity of glass and the fact that the beam energy is distributed more uniform in the glass layer than in the upper material layers, which causes smaller local temperature gradients.

5.5 The optical mask temperature profile for e-beam writing of a 1 μm × 1 μm square flash

In this section the calculated results of the optical mask temperature profile calculations for square flash are presented. Mask temperature profiles are shown for various e-beam writing times in the time interval from 250 nsec to 1 μsec, the typical range for writing times of a 1 μm × 1 μm square flash. The ANSYS finite model that is used for the calculations is described in detail in Section 5.2. The total number of nodes in this model is 51656 and the number of elements is 47950. A heat generation load, corresponding to the energy distribution function for square flash writing, is applied to each of these elements. The computational time for the temperature profile calculation at one specific writing time is about three hours using a 200 MHz Pentium four processor system with 1Gigabyte of RAM.

The surface temperature profile of the optical mask that is of main interest for resist heating effects is plotted in Fig. 5.31.a to Fig. 5.31.d for different square flash writing times. The maximum temperature rise is 44.2 °C for a 250 nsec writing time, 72.2 °C for 500 nsec writing time, 88.0 °C for 750 nsec writing time and 97.6 °C for 1000 nsec writing time. In Fig. 5.32.a and Fig. 5.32.b the mask temperature profile in depth is plotted for a writing time of 250 nsec and 1000 nsec. The cut shown in these two figures is along the mask centerline. The square flash area is located in the middle of the figures and has a size of five elements in the horizontal direction. The PMMA material is modeled by ten element layers in vertical direction shown at top of the plots. The following two element layers in the vertical direction are chrome layers. All the following element layers represent the bulk glass material. In both plots in can be clearly seen that that maximum temperature rise occurs in the PMMA layer due to a high local heat generation in the material and the low diffusivity of the polymer. Compared to the PMMA and glass layer there is only a negligible temperature rise in the chrome. Because of the high conductivity of chrome the energy is conducted rapidly out of the flash center region where the maximum heat generation occurs. Although the local heat generation is much smaller in the glass than in the chrome and PMMA, there is a significant temperature rise of 55 °C in the glass material. The reason for this high temperature is the relatively low diffusivity of glass and the fact that the beam energy is distributed more uniform in the glass layer than in the upper material layers, which causes smaller local temperature gradients.

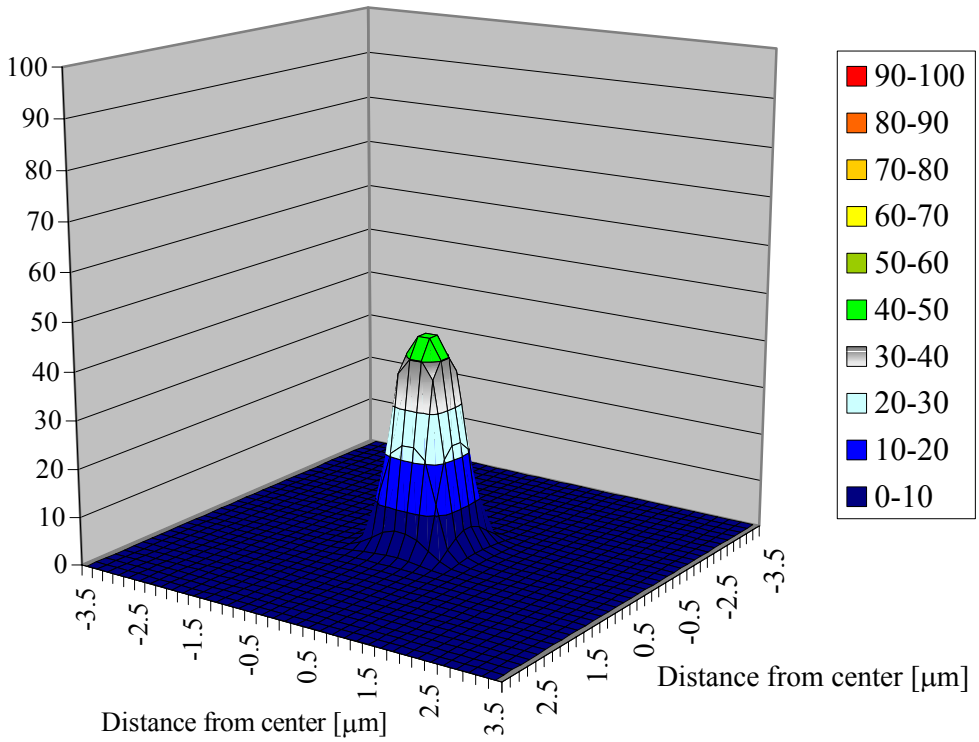


Figure 5.31.a. Mask surface temperature profile for e-beam writing of 1 μm × 1 μm square flash. The writing time is 250 nsec. Both horizontal axis are described by the same scale.

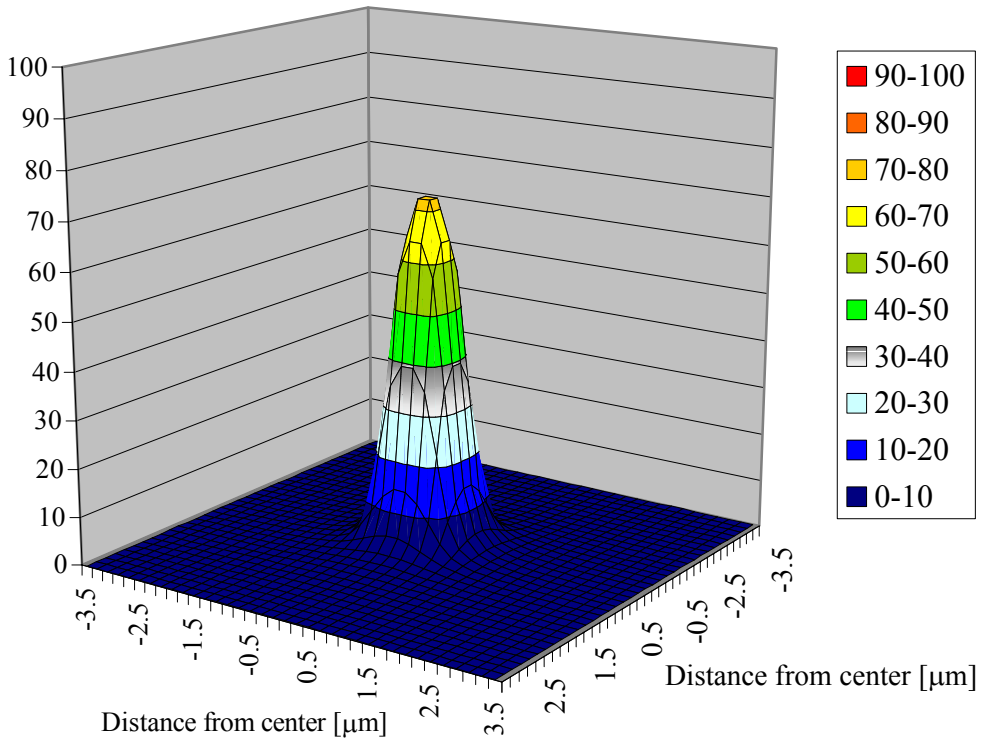


Figure 5.31.b. Same as Fig.5.31.a , but for writing time of 500 nsec.

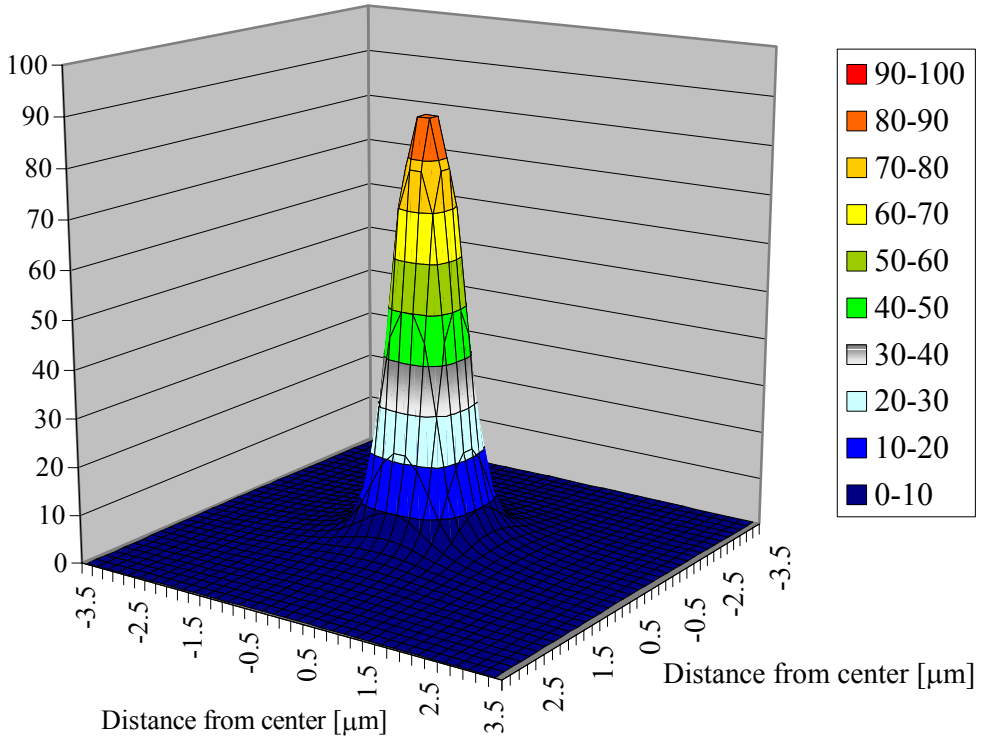


Figure 5.31.c. Same as Fig.5.31.a , but for writing time of 750 nsec.

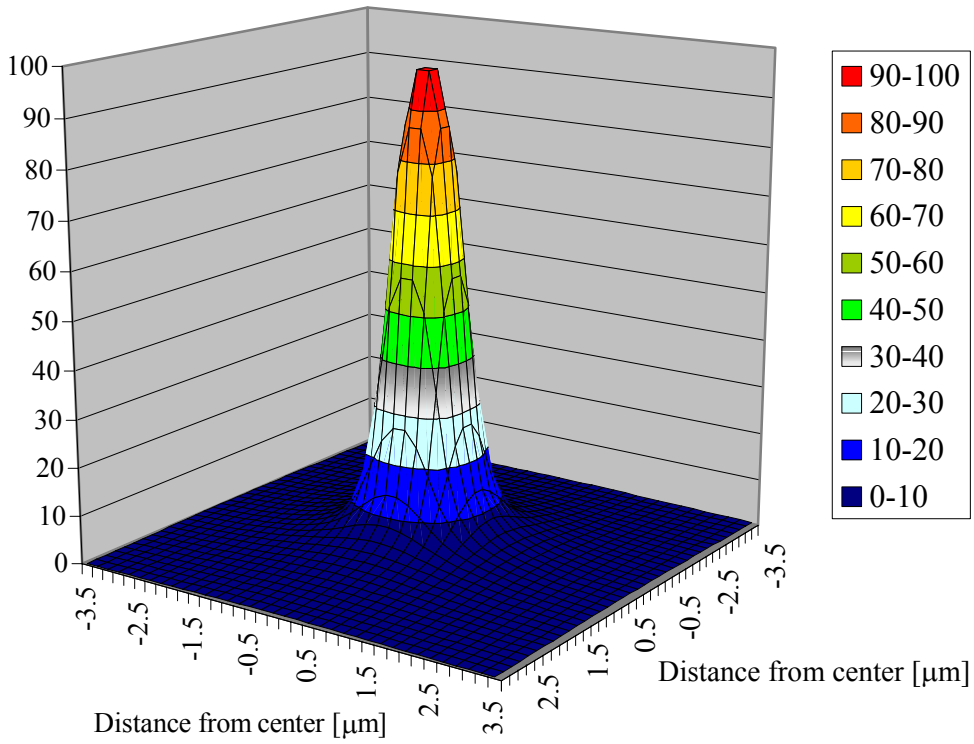


Figure 5.31.d. Same as Fig.5.31.a , but for writing time of 1000 nsec.

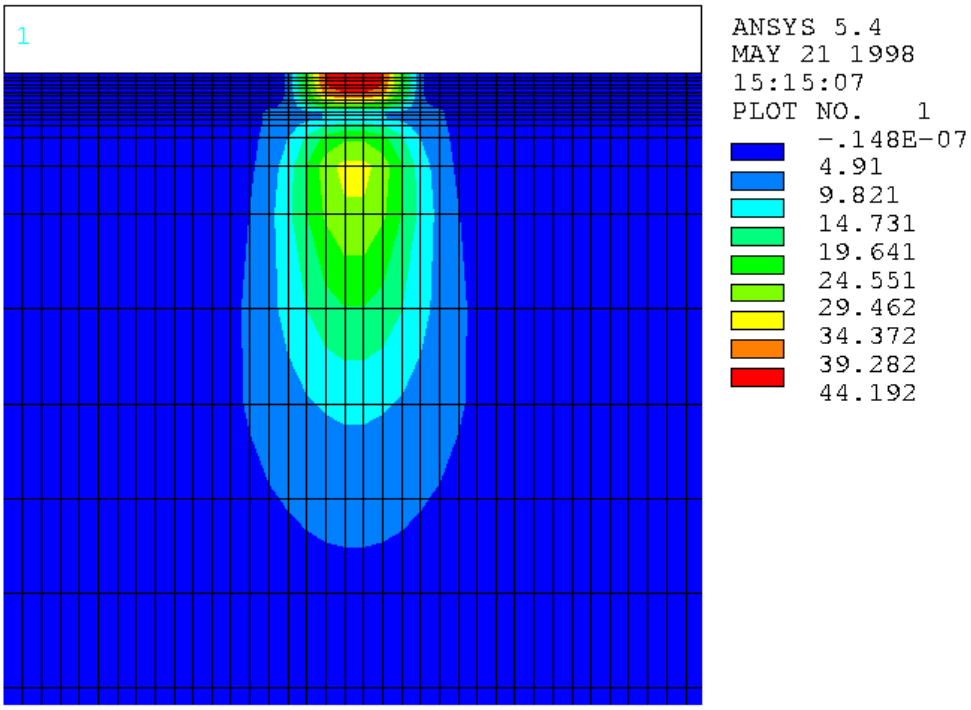


Figure 5.32.a. Optical mask temperature profile in depth. The figure shows a cut along themask center line. The writing time is 250 nsec.

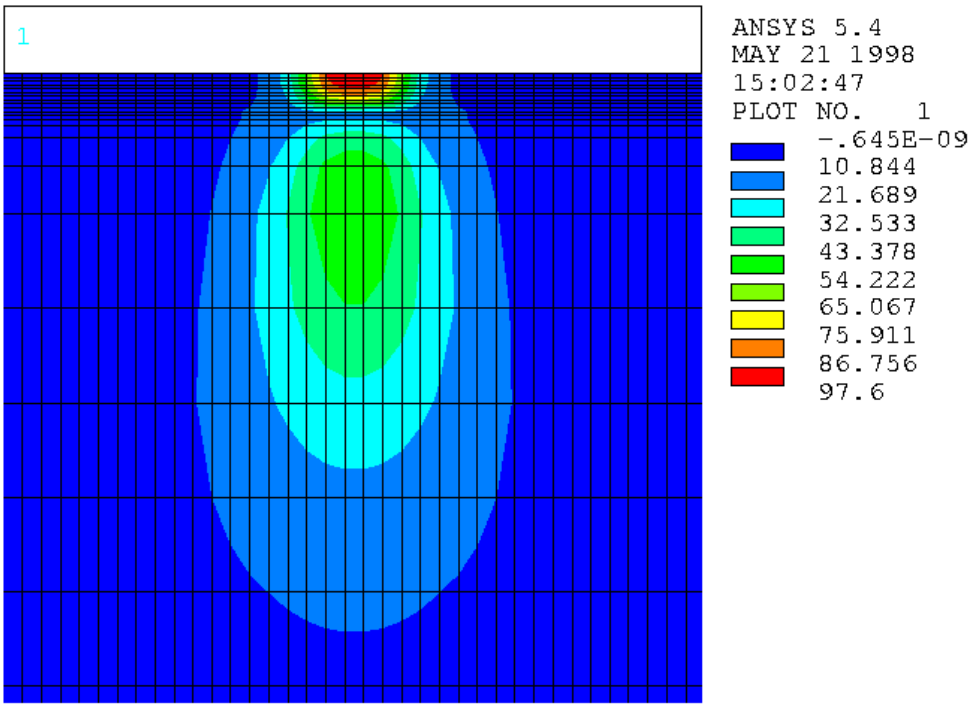


Figure 5.32.b. Optical mask temperature profile in depth. The figure shows a cut along themask center line. The writing time is 1 μsec.

5.5 References

[5.1] S. Babin, Etec, private communications.

Chapter Six

Local heating during e-beam writing of X-ray masks

6.1. Introduction

In this chapter the local heating of an X-ray mask during electron-beam writing is examined. The goal of this chapter is to determine how local heating contributes to global mask heating. The answer is important as an area thermal averaging technique was used in [6.1], [6.2], and [6.3] for the calculation of global mask temperature profiles. It is of interest to verify that the global mask heating can be simulated accurately with a averaging technique, because, as described in Chapter four, it is the only numerical method than can simulate the complete mask writing process in a reasonable computational time. An accurate transient mask temperature profile is needed to obtain a precise prediction of thermal mask distortions.

In this chapter local mask temperature profiles of subfields are calculated with simulations that approximate the actual e-beam writing process with a high accuracy. These results are compared to temperature profiles calculated with an averaging technique. In Section two the e-beam writing process of a subfield is analyzed for writing a simple test pattern in the subfield area. The heat loss mechanisms of the mask are examined and compared to the heat loss during subfield writing with an arithmetic averaging method. In Section three the e-beam writing process is examined for writing a subfield on the geometry level of pattern shapes. The heat loss mechanisms and the mask temperature profile are compared to results for subfield writing with averaged heat generation.

6.2 E-beam writing of a test pattern on an X-ray mask subfield

6.2.1 Introduction

In this section of Chapter six local mask heating during beam writing of a simple test pattern on a subfield is investigated. The goal is to obtain a qualitative estimate of the difference in the mask heat loss mechanisms between simulating the subfield writing with a detailed and an averaging technique.

The pattern that is used for the detailed simulation of the beam writing process is pictured in Fig. 6.1. The size of the subfield is with 40 μm × 40 μm, which is similar to the size of a 37.5 μm × 37.5 μm Talon mask subfield shown in Fig. 3.4 and Fig. 4.1.

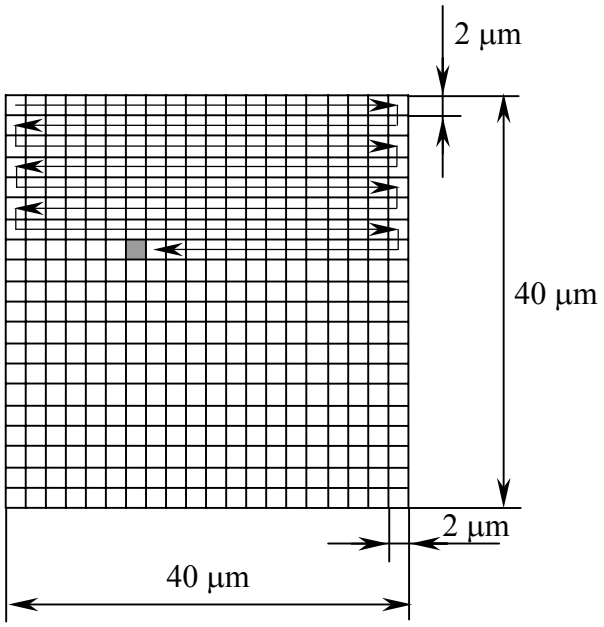


Figure 6.1. Subfield test pattern.

The size of the beam is 2 μm × 2 μm, which is equal to the maximum pixel size that a modern variable shaped e-beam can write [6.4]. Since the current density of the beam does not depend on the beam size, the 2 μm × 2 μm size corresponds to the maximum energy that can be deposited in a mask by a beam. A maximum local energy deposition is a worst scenario for the comparison of the detailed writing to writing with an averaged heat generation rate. The higher the local heat generation is, the higher are the local temperature rises and gradients. For this reason the difference between detailed and averaged writing can be expected to be the highest for maximum local energy deposition. For the same reason a subfield pattern coverage of a hundred percent, single pass writing and zero stepping time

between two spots is assumed. The writing style that is chosen for simulating the detailed subfield writing is a serpentine style as pictured in Fig. 6.1. Many e-beam machines write the pattern with a raster scanning technique, as described in Chapter three. This writing fashion is chosen for the simulation of the detailed subfield writing because of the simplicity of the method and because proximity heating increases local temperature rises. With a raster scanning technique all the stripes of a pattern can be written either in uniform direction or in serpentine style. For serpentine style writing the proximity heating contribution to local heating is bigger than for uniform style writing, because after finishing the writing of one line, the new line is written close to the previous line. This is the reason for choosing serpentine style writing for the detailed simulation of subfield writing.

Figure 6.2 shows the dimensions of the X-ray mask membrane for the mask used for the simulation of writing a test pattern. The material properties for the different layers are listed in Table 6.1. It is assumed that the pattern is written with a variable shaped electron-beam system with the typical system parameters of 50 KV acceleration voltage and a current density of 20 A/cm^2. For the resist sensitivity of PMMA a value of 55 μC/cm^2 is used. For these e-beam and resist characteristics, the writing time Δt_{Spot} of a single spot is, calculated with equation 6.2.1.1, equals 2.75 μsec.

$$\Delta t_{Spot} = \frac{\text{Sensitivity of Resist } [C/\mu m^2]}{\text{Current Density of Electron Beam } [A/\mu m^2]} \qquad [6.2.1.1]$$

In Chapter five the calculation of the heat deposition function for writing a square flash on an optical mask has shown that close to the mask surface most of the beam energy is deposited

in the beam writing area. After the beam has penetrates a few micrometers of optical mask material in depth, a significant radial spread of deposited energy is observed.

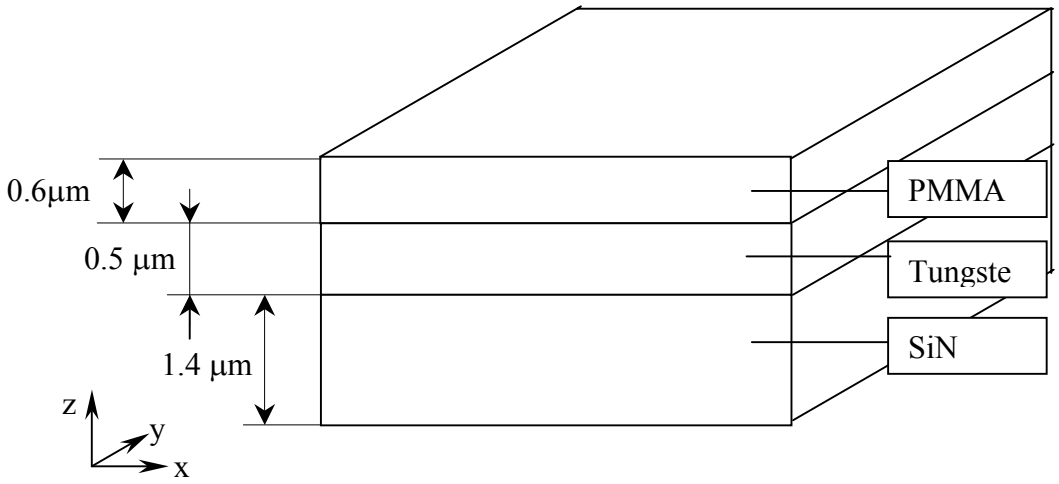


Figure 6.2. X-ray mask membrane with dimensions in depth.

Material	Density (kg/m^3)	Conductivity (W/m K)	Heat Capacity (J/kg K)
PMMA	1210	0.176	1170
Tungsten	19300	17.8	130
SiN	2220	1.38	787

Table 6.1. Mask material properties.

In Section three of Chapter three it is explained that the radial spread of the deposited beam energy occurs mainly because of electron backscattering. Since an X-ray mask is very thin in the patterning area, very few of the incident electrons are slowed down enough in the material that a backscattering event could occur. Therefore, electron scattering in the X-ray mask is dominated by forward scattering. Even so, the radial spread of electrons by forward scattering is almost negligible. Calculating the increase of the beam dimensions of the resist layer with the empirical equation 3.3.1 for forward scattering gives a value of 112.5 nm for the PMMA layer thickness of 0.5 μm and the acceleration voltage 20 KV. The value of 112.5 nm describes how much the beam is spread in the PMMA layer. Assuming a similar beam diameter increase in the chrome

and silicon nitride layer, the total increase of the beam dimensions is about half a micrometer. This minimal increase of the beam dimensions makes it possible to approximate the energy that is deposited by a square 2 μm × 2 μm beam in an X-ray mask with an energy deposition volume of the beam size times the mask thickness. Assuming a uniform energy distribution in depth the heat generation in the mask can be calculated with equation 6.2.2. The energy that is deposited in the mask during the writing of a single spot is equal to the product of the beam acceleration voltage, the resist sensitivity and the absorption factor. For the simulation of test pattern writing a value of 0.1 is used for the energy absorption factor of the X-ray mask. This value is typical for an X-ray mask membrane. With a mask membrane thickness of 2.5 μm and the spot writing time of 2.75 μsec, that is calculated with equation 6.2.1.1, the volumetric heat generation q'''_{spot} is 0.4E-3 W/μm^3, determined with equation 6.2.1.2.

$$q'''_{spot} = \frac{\text{Absorbed Energy } [J/\mu m^2]}{\Delta t_{Spot}\ [s] * \text{Thickness}_{Membrane}\ [\mu m]} \qquad [6.2.1.2]$$

The conditions that correspond to the writing of the test pattern, described in the previous paragraph, are summarized in Table 6.2.

Spot writing time Δt_{Spot}	2.75 μsec
Waiting time between spots	0 μsec
Size of square spot	2 μm × 2 μm
Volume of uniform heat generation $V_{Energy\,deposition,spot}$	2 μm × 2 μm × 2.5 μm
Volumetric heat generation	0.4E-3 W/μm^3
Writing style	Serpentine, single pass

Table 6.2. Summary of mask conditions for writing of a serpentine test pattern on an X-ray mask.

In the following paragraph the heat loading conditions of the X-ray mask membrane are determined for the simulation of subfield writing with the averaging technique used by [6.1],

[6.2], and [6.3]. The averaging technique describes the subfield writing process by applying a uniform heat generation to the total written subfield area during the complete writing time. The total subfield writing time $\Delta t_{Subfield}$ is equal to the sum of the spot writing times Δt_{Spot} for all the N_{Spots} spots within the subfield, since there is no waiting time between writing the spots. The subfield writing time calculated with eq. 6.2.1.3 is 0.0011 sec for N_{Spots} equal 400. The number of spots N_{Spots} is calculated with the area ratio of subfield to spot area.

$$\Delta t_{Subfield} = \Delta t_{spot} * N_{Spots} \qquad [6.2.1.3]$$

The uniform heat generation that is applied to the subfield area for the simulation with the averaging technique is calculated with equation 6.2.1.4 and equation 6.2.1.5.

$$q'''_{Subfield} \cdot V_{Energy\,deposition,subfield} \cdot \Delta t_{Subfield} = q'''_{spot} \cdot V_{Energy\,deposition,spot} \cdot \Delta t_{spot} \cdot N_{Spots} \qquad [6.2.1.4]$$

$$V_{Energy\,deposition,subfield} = V_{Energy\,deposition,spot} \cdot N_{Spots} \qquad [6.2.1.5]$$

Equation 6.2.1.4 expresses that the energy deposited by the electron-beam in the mask during writing the subfield in serpentine style is equal to the energy deposited in the subfield for approximating the writing process by applying an average heat generation $q'''_{Subfield}$ to the complete subfield. With equation 6.2.1.5 and equation 6.2.1.3 the form of equation 6.2.1.4 can be changed to equation 6.2.1.6. The equivalent heat generation rate $q'''_{Subfield}$, calculated with equation 6.2.1.6, is 0.1E-5 W/μm^3.

$$q'''_{Subfield} = \frac{q'''_{spot}}{N_{spots}} \qquad [6.2.1.6]$$

The conditions that correspond to the approximation of the test pattern writing by an averaging technique are summarized in Table 6.3.

Subfield writing time $\Delta t_{Subfield}$	0.0011 sec
Size of subfield	40 μm × 40 μm
Volume of uniform heat generation, V	40 μm × 40 μm × 2.5 μm

generation $V_{Energy\,deposition,spot}$	
Volumetric heat generation	0.1E-5 W/μm^3
Writing style	Uniform applied heat generation

Table 6.3. Summary of mask loading conditions for simulating the writing of a serpentine test pattern on an X-ray mask with an averaging technique.

6.2.2 The finite element model of the X-ray mask subfield

The dimensions of a typical X-ray mask are shown in Fig. 2.2.1. Because the dimensions of the complete mask patterning area are about three orders of magnitudes larger than the subfield dimensions, it is impossible to model the complete mask for the simulation of the subfield writing. By only modeling a part of the mask patterning area, the computational time of the writing simulation can be significantly reduced. Figure 6.3 shows schematically the idea for a finite element model that is used to model the writing of the subfield area. The dark gray colored square in Fig 6.3 represents a section of the X-ray mask membrane area.

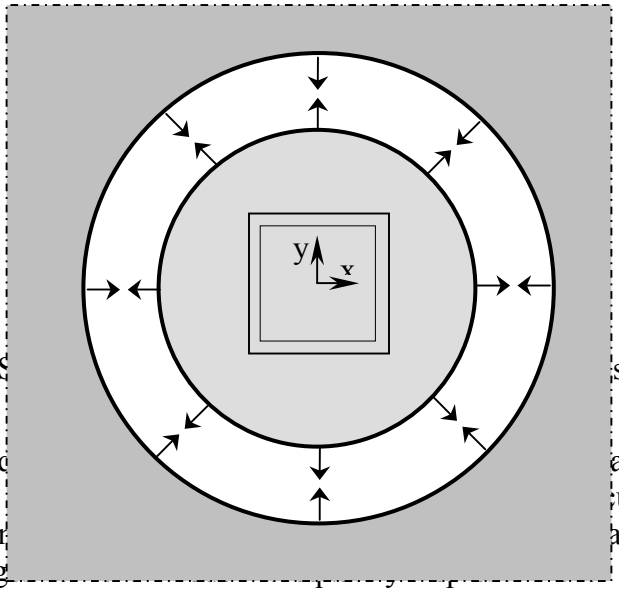


Figure 6.3. S sk subfield.

The light gray colore a in the center is the membrane area that is urate results with this model for a simulatio ant, that the boundary conditions at the edg interactions with the surrounding mask membrane area. To avoid the definition of complicated boundary conditions at the edges of the finite element model, the dimensions of the circular model area can be chosen sufficiently large so that big that there is almost no interaction between the edges of the model an the surrounding membrane area. The modeled membrane area has to be large enough that within the e-beam writing time of the complete subfield almost no thermal energy diffuses to the edges of the model. Under these conditions the boundary conditions at the edges of the finite element model do not affect the results of the e-beam writing simulation. A good way to minimize the necessary computational time of the subfield writing simulation is to minimize the number of elements in the model by modeling

only a part of the circular area, shown in Fig. 6.3, with a three dimensional model. It is sufficient to model that part of the membrane with a three dimensional model where temperature gradients in membrane depth are expected to occur. It is probable that significant gradients in depth can only be found in that part of the membrane where energy is deposited by the electron-beam. In Fig. 6.4 the heat flow in an X-ray mask membrane is shown for writing a single spot with the writing conditions listed in Table 6.1. The finite element model that is used for the simulation of writing a single spot is described in detail in the following paragraphs. Figure 6.4 verifies the assumption that gradients in membrane depth direction occur only in a region close to the point where the electron-beam deposits energy in the mask. In Fig. 6.4 heat flux vectors are plotted for a membrane cross-section through the center of the written spot. The dimension of the spot writing area is indicated in the figure.

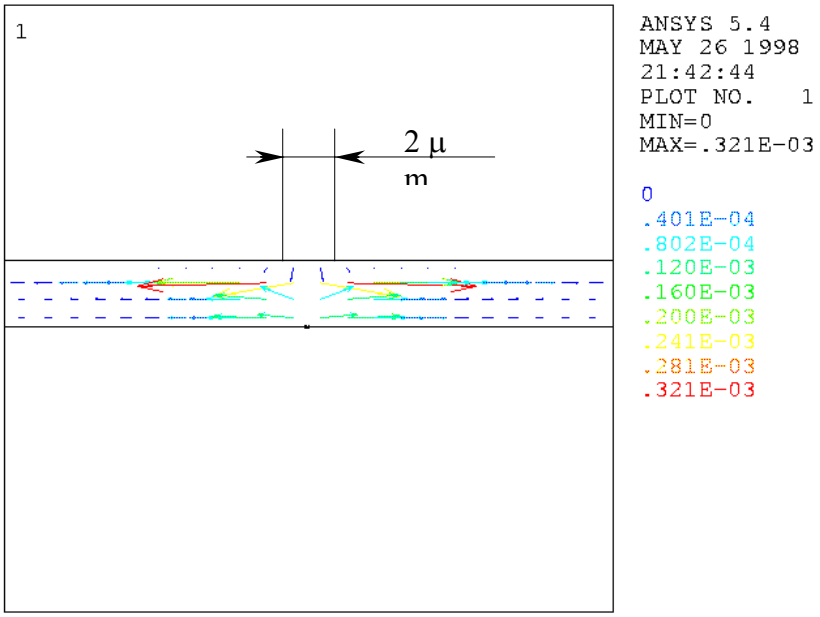


Figure 6.4. Heat flux vectors in the mask membrane for writing a single spot on the mask membrane with writing conditions listed in Table 6.1.

The directions of the heat flux vectors show qualitatively that temperature gradients in the depth direction only occur within a distance of about twice the spot dimensions from the spot center. Outside of this region the heat flux vectors are almost parallel which indicates that the in-depth temperature gradient in this region is negligible. For this reason it is sufficient to model only the subfield, the part of the mask where energy is deposited by the beam, and a small mask region around the subfield with a three dimensional model of finite elements. The size used in the model is 53 μm × 53 μm × 2.5 μm. In Fig 6.3 the boundaries of the three dimensional part of the model are shown as

the big square in the center. For the circular mask region around the square the heat transfer is assumed to be two-dimensional. Shell elements are used to model this part of the mask membrane. The advantage of using shell elements in a finite element model for the simulation of a two dimensional heat transfer problem is described in Section two of Chapter five. Because shell elements represent properties of a single material, average material properties need to be assigned to the finite elements in the circular mask region. The heat transfer in the material with averaged properties has to be equivalent to the heat transfer in the stack of three materials as shown in Fig. 6.2. The average heat capacity, density and thermal conductivity is calculated with equation 6.2.2.1, equation 6.2.2.2, and equation 6.2.2.3.

$$C_{P\,average} = \frac{C_{P\,PMMA} \cdot \Delta_{PMMA} + C_{P\,Tungsten} \cdot \Delta_{Tungsten} + C_{P\,SiN} \cdot \Delta_{SiN}}{\Delta_{Mask\,Membrane}} \qquad [6.2.2.1]$$

$$\rho_{average} = \frac{\rho_{PMMA} \cdot \Delta_{PMMA} + \rho_{Tungsten} \cdot \Delta_{Tungsten} + \rho_{SiN} \cdot \Delta_{SiN}}{\Delta_{Mask\,Membrane}} \qquad [6.2.2.2]$$

$$k_{average} = \frac{k_{PMMA} \cdot \Delta_{PMMA} + k_{Tungsten} \cdot \Delta_{Tungsten} + k_{SiN} \cdot \Delta_{SiN}}{\Delta_{Mask\,Membrane}} \qquad [6.2.2.3]$$

In these equations Δ_{PMMA}, $\Delta_{Tungsten}$, and Δ_{SiN} are the thicknesses of the mask material layers. The equations are derived, as shown in [6.6], for heat transfer in x-y direction of the materials pictured in Fig. 6.2.

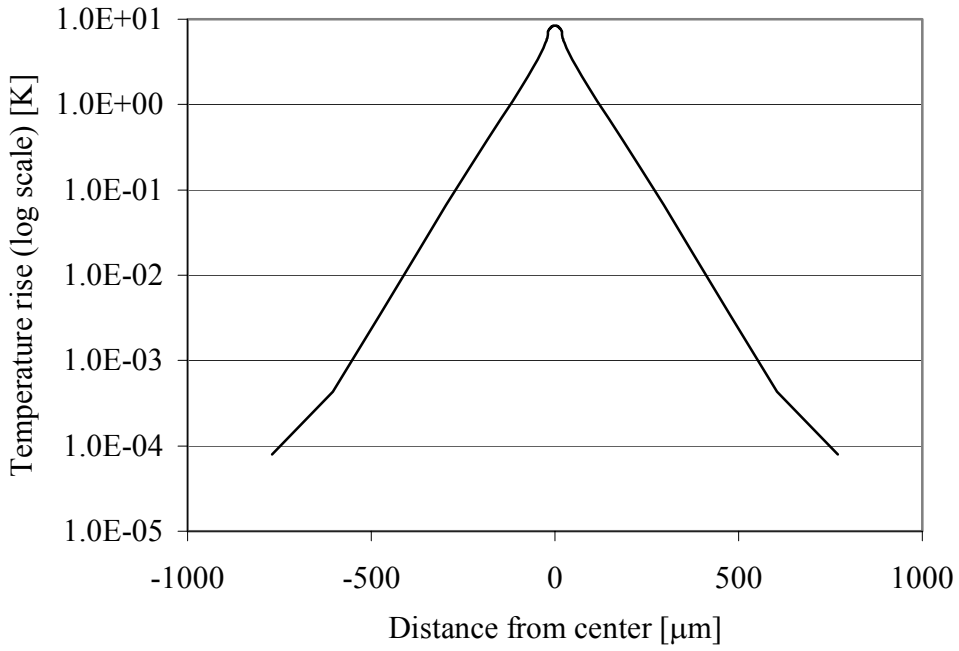


Figure 6.5. Mask surface temperature plotted along subfield centerline for subfield writing as described in Table 6.3.

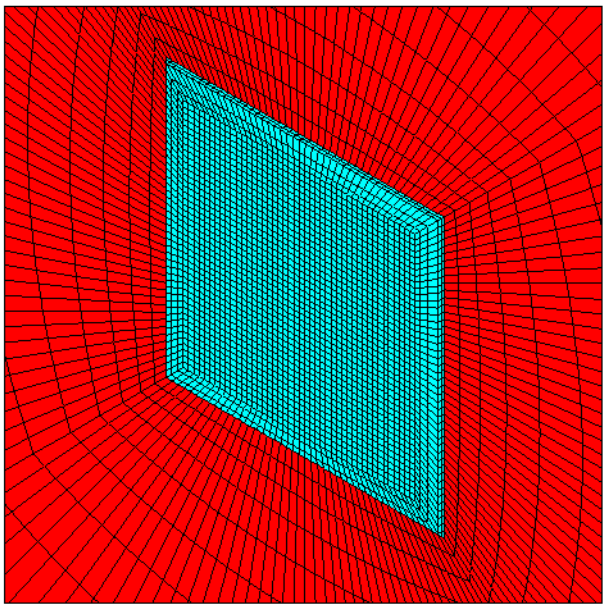


Figure 6.6. Finite element model used for simulation of subfield writing process.

As explained in the previous paragraph the diameter of the finite element model, schematically shown in Fig. 6.2, has to be large enough, that an insignificant amount of thermal energy diffuses to the edge of the model in the writing time of the mask. The model diameter size is checked by simulating the e-beam writing of the subfield with the averaging method as listed in Table 6.3.

Size of subfield	40 μm × 40 μm
Size of the three dimensional finite element region	53 μm × 53 μm× 2.5 μm
Radius of the two dimensional finite element region	770 μm

Table 6.4. Summary of finite element model dimensions used for simulating the writing of a subfield test pattern on an X-ray mask.

The surface mask temperature in Fig. 6.5, plotted along a cut through the subfield center, shows that the temperature rise of 0.0001°C at the model edges

is very small at the end of the subfield writing time. This temperature rise indicates that only a negligible amount of thermal energy diffuses to the edges of the model in the subfield writing time. Thus, the used model diameter is sufficiently large. The dimensions of the finite element model that is used for simulating the subfield writing are listed in Table 6.4. A schematic of the model is shown in Fig. 6.6. The square turquoise area in the center is the three-dimensional finite element region modeled with cuboid elements. The rectangular elements shown in red are shell elements with equivalent material properties. Not shown in Fig. 6.6 is a layer of surface effect elements covering the shell and cuboid element region. The surface effect elements are needed to include radiation effects in the finite element model. Radiation is the only heat transfer interaction between the mask membrane and the surrounding, because the electron-beam writing of masks is done in a vacuum. The emissivity of the mask surface is assumed to be 0.19, an experimentally determined value for a similar X-ray mask membrane [6.7]. Although it can be expected that the actual emissivity of the mask geometry shown in Fig. 6.2 is not exactly 0.19, it is shown later in this chapter that using this emissivity value does not significantly affect the accuracy of the temperature calculations.

After determining the dimensions of the finite element model an adequate model mesh must be defined. It has to be kept in mind that the main goal of this chapter is to estimate the difference between a detailed simulation of a patterning process and a simulation with an averaging technique. For this reason the mesh width of the model is not optimized with the intention of obtaining the highest possible simulation accuracy. The mesh is chosen as large as possible to limit the necessary computational time and small enough to guarantee a sufficient accuracy. In the three-dimensional model region with cuboid elements the PMMA and the tungsten layers are modeled with only one layer of elements in depth. The silicon nitride material is modeled with two layers of cuboid elements in depth. For the used mesh width of 1 μm in x-y direction the maximum ratio of a single finite element's side length is two, which much lower than the recommended limit [6.7]. The mesh width of 1 μm in x-y direction is chosen to allow the calculation of a temperature profile over the 2 μm × 2 μm area of a single electron-beam spot.

To verify that the finite element model is programmed correctly with the finite element software ANSYS® a simple test case is simulated. A uniform volumetric heat generation of 2.5E-5 W/μm^3 is assigned to the complete

volume of the mask membrane and the membrane surface temperature is calculated for the equilibrium of the mask with the surroundings by radiation. For equilibrium conditions the rate of energy deposition in the mask $\dot{E}_{in}$ is equal to the radiation loss rate $\dot{E}_{out}$. With equation 6.2.2.4 and eq. 6.2.2.5 the mask surface temperature $T_{surface}$ can be calculated and compared to the results obtained with the finite element model.

$$\dot{E}_{in} = q''' \cdot V_{mask\,model} \qquad [6.2.2.4]$$

$$\dot{E}_{out} = \sigma \cdot \varepsilon \cdot A_{model\,surface} \cdot \left(T^4_{surface} - T^4_{surroundings}\right) \qquad [6.2.2.5]$$

In Fig. 6.7 the mask surface temperature, calculated with the finite element program is plotted versus time. After about 0.0018 seconds the equilibrium temperature of the mask surface is reached. The difference between the surface temperature calculated with the finite element program and the equations 6.2.2.4 and eq. 6.2.2.5 is plotted in Fig. 6.8 for different mesh width in x-y direction.

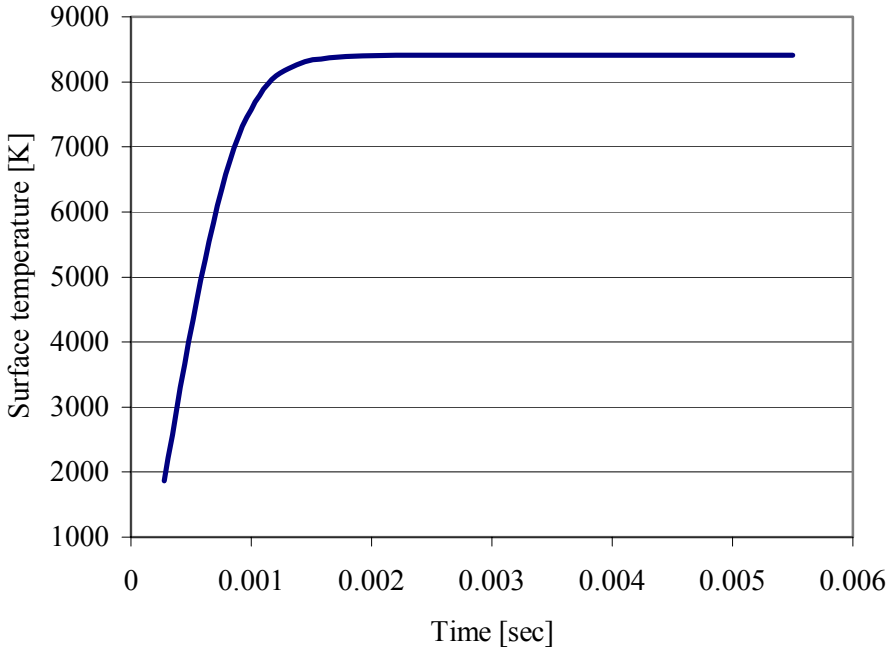


Figure 6.7. Mask surface temperature vs. time for uniform heat generation of 2.5E-5 W/μm^3 in the mask membrane.

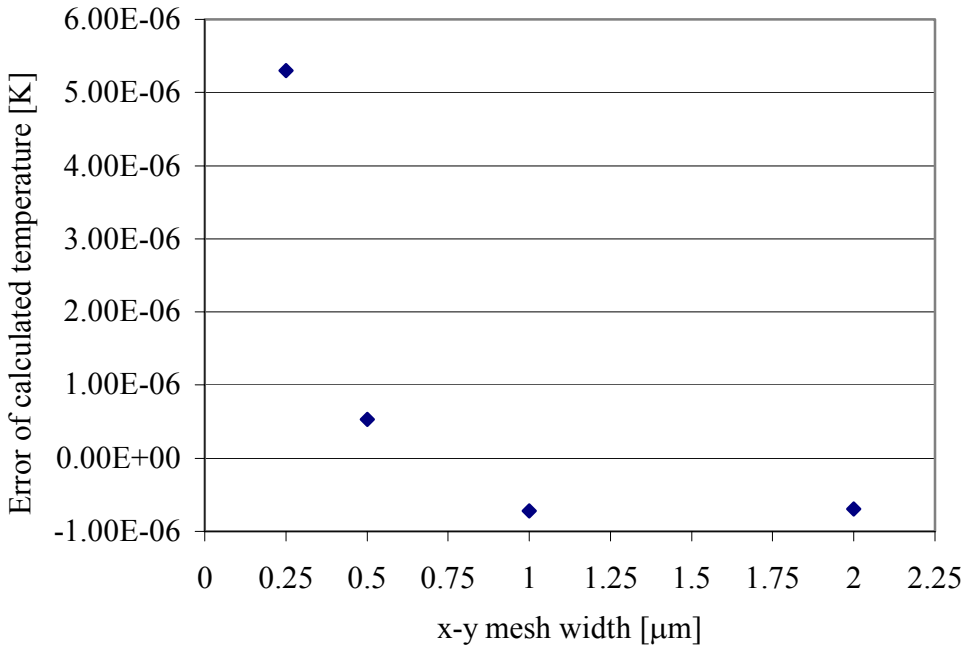


Figure 6.8. Finite element calculation error of the exact radiation equilibrium temperature.

For all of the finite element calculations the temperature error is below the calculation accuracy of 1E-5°C, specified in the finite element program for this simulation. By this simple test case simulation it is verified that all the elements of the model are connected correctly, because the energy deposited in the energy generated in the complete mask volume has to be conducted to the surface where it is radiated away to the surrounding.

6.2.3 Comparison of subfield electron-beam writing simulation with detailed and with averaging method

The complete finite element model, described in detail in the previous section, has 20064 elements and 15406 nodes. Simulating of the serpentine style electron-beam writing process, as specified in Table 6.2 and Fig 6.1, needs a computational time of 84 hours on an HP 9000/780 with 512 Mbyte RAM.

The surface temperature profile of the subfield is plotted in Fig. 6.9 after the writing of 135 spots. The surroundings temperature for the writing process is 294°K. Characteristic for the temperature profile is a significant temperature rise to 587°K at the center of the last written spot and almost no temperature rise in the rest of the subfield area. This indicates that the energy, deposited during the writing of a single spot, is rapidly conducted away from the deposition location.

In the following paragraph the heat transfer mechanisms in the mask are compared for the subfield electron-beam writing process simulated with the detailed and the averaging technique. To compare the the radiation loss mechanisms for the two simulation techniques the total energy that is lost to the surroundings by radiation in a given writing time is calculated. With the output of finite element program ANSYS® and eq. 6.2.3.1 the energy $E_{\text{radiation loss}}$ that is lost by radiation can be determined.

$$E_{\text{radiation loss}} = \int_{\text{Writing time}} \int_{\text{Model surface}} \dot{q}''_{\text{radiation}} \, d(\text{Area}) \, d(\text{time}) \qquad [6.2.3.1]$$

In Fig. 6.10 the energy $E_{\text{radiation loss}}$ is plotted versus the fraction of the total surface area that is already written. With values of fraction of the total surface area written the corresponding writing time can be calculated by multiplying the fraction value with the total subfield writing time. For example, a fraction of fifty percent is equivalent to writing half of the complete subfield in the writing time of half the total subfield writing time $\Delta t_{\text{Subfield}}$. In Fig. 6.10 energy values $E_{\text{radiation loss}}$ are plotted for writing the subfield with a spot in serpentine style and with an equivalent averaged heat generation. The uniform average heat generation is calculated with equation 6.2.1.6.

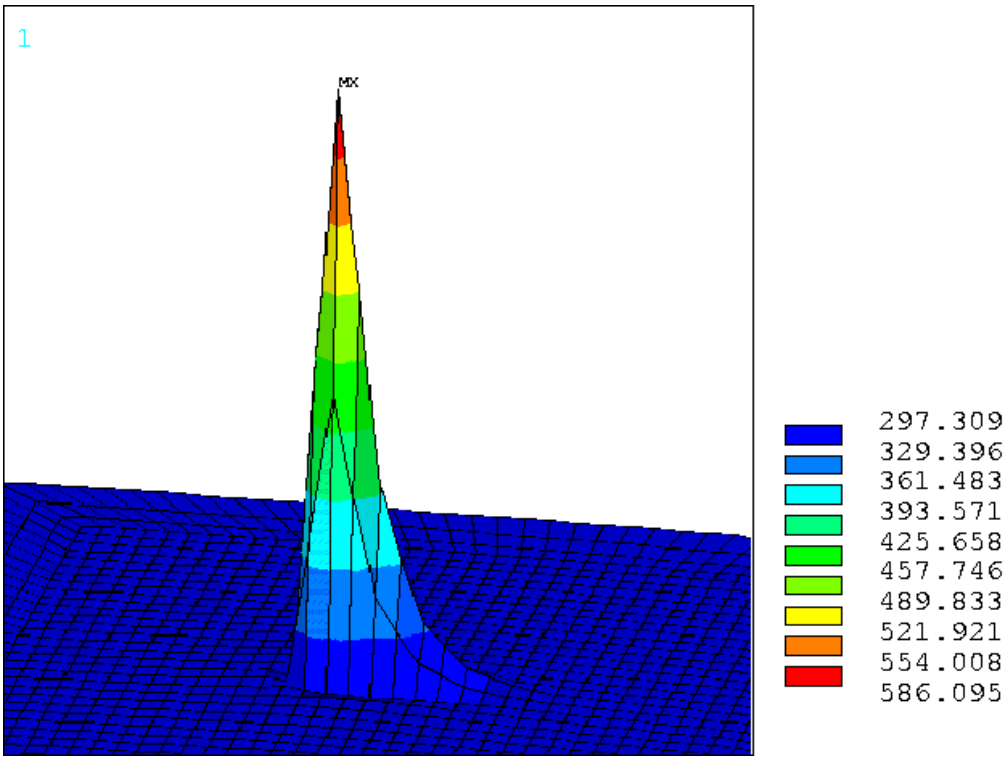


Figure 6.9. Temperature profile of the subfield for electron-beam writing from the left to the right of the picture. Temperatures in °K.

The number of spots N_{spots}, in this equation is determined with the fraction of the written subfield area and the total number of spots in a subfield. The values of $E_{radiation\ loss}$ in Fig. 6.10 for the serpentine style electron-beam writing are calculated with the finite element program ANSYS® only for writing a third of the total subfield area. The $E_{radiation\ loss}$ values for writing the last two thirds of the subfield are extrapolated. The reason for not determining the radiation loss of the complete writing process with ANSYS® is a limit by the finite element software. To compute the radiation loss ANSYS® requires saving a large amount of information in a file. Because the limit for a maximum possible file size set in ANSYS® to 2.1 Gbyte is reached after simulating a third of the complete writing process, it is not easily possible to simulate the complete writing process.

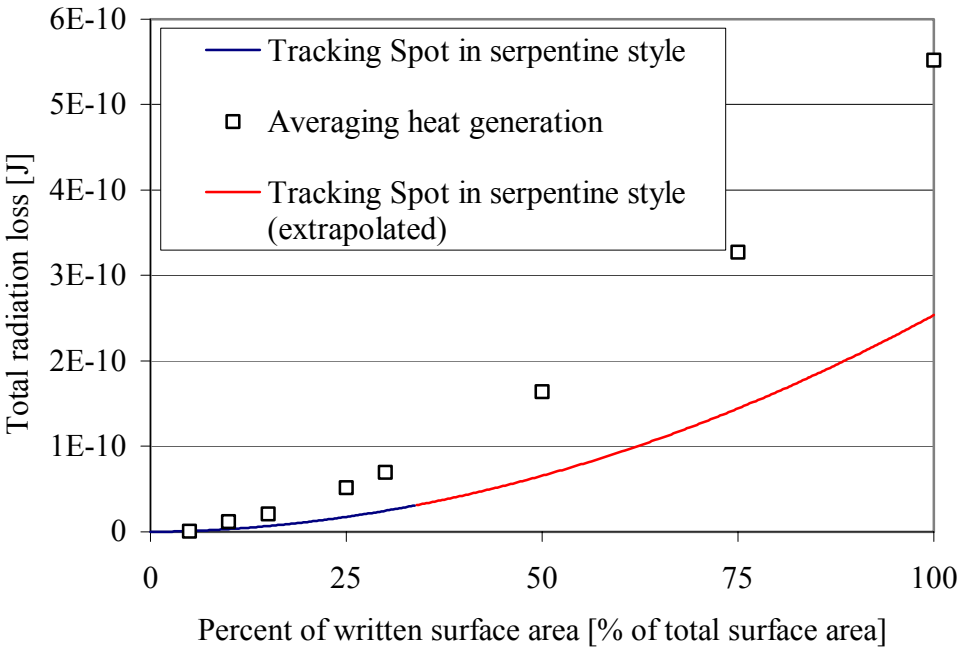


Figure 6.10. Total energy lost by radiation during the electron-beam writing process vs. the fraction of surface area that is already written .

The comparison in Fig. 6.10 of the energy lost by radiation for the serpentine and averaged subfield writing shows that using an averaging technique the radiation loss of the actual writing process is underpredicted. The radiation loss, determined with the averaging technique, is about double that for the actual writing process. However , as will be shown, the radiation loss is a negligible amount of the energy input.

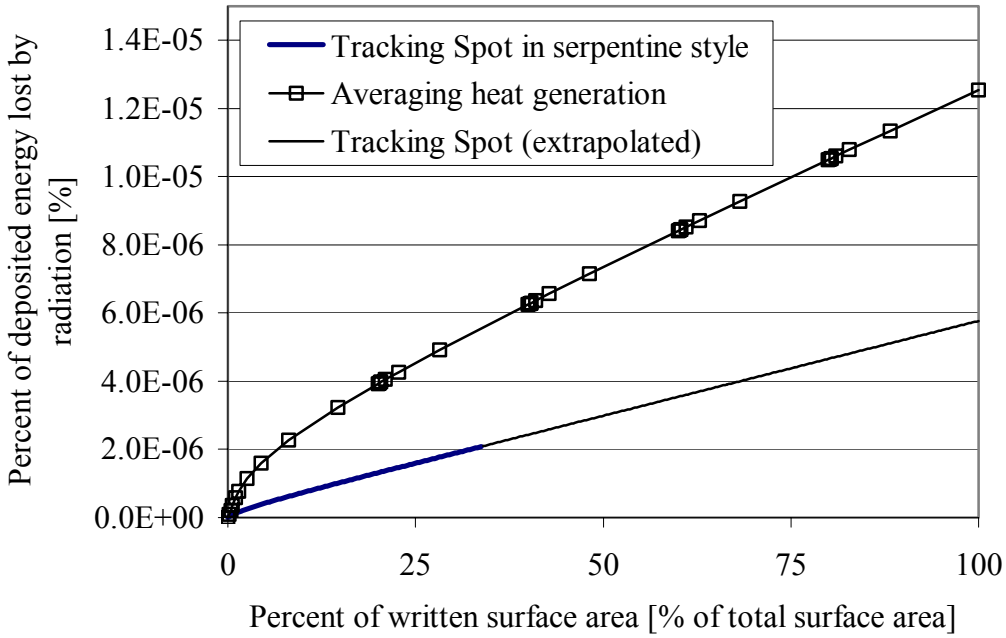


Figure 6.11. Ratio of total energy lost by radiation and deposited energy during the electron-beam writing process vs. the fraction of surface area that is already written.

The total energy lost by radiation is compared to the total energy deposited by the beam in the subfield in Fig. 6.11. The plot indicates that an insignificant fraction of the energy deposited in the subfield is lost to the surroundings by radiation. During the writing time of the complete subfield only 1.3E-5 percent of the energy deposited by the beam is lost by radiation for simulation the writing with the averaging technique. Because the radiation loss is negligibly small, it does not matter that the radiation loss for the actual beam writing process does not match with the radiation loss determined with the averaging technique. The conclusion is that the dominating heat transfer mechanism during the beam writing process is conduction.

To compare the conduction heat transfer for the averaging and serpentine style simulation techniques, the average surface temperature of the subfield area is calculated for both cases at the end of writing the complete subfield area. For an initial temperature of 294.26 °K the average surface temperature rise for simulating the writing of the subfield with the averaging technique is 7.50°K. For simulating the writing with the serpentine writing method the average subfield surface temperature rise is 5.27°K.

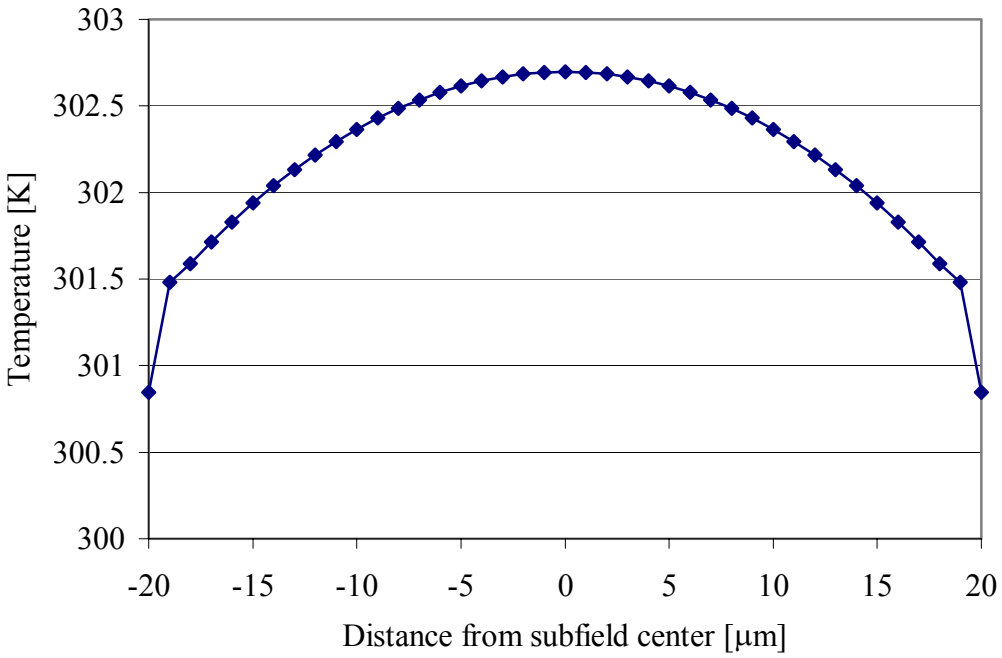


Figure 6.12.a. Subfield surface temperature profile plotted along centerline of subfield. The temperature profile for the end of the subfield writing time is plotted for simulation the energy deposition with the averaging technique.

The predicted average temperature rise is forty percent higher using the averaging technique than for modeling the serpentine style writing. The reason for this significant difference is the different temperature profiles generated in the subfield area during the time of energy deposition in the mask. In Fig. 6.12.a the average surface temperature profile of the subfield is plotted along the subfield centerline at the end of writing the complete subfield.

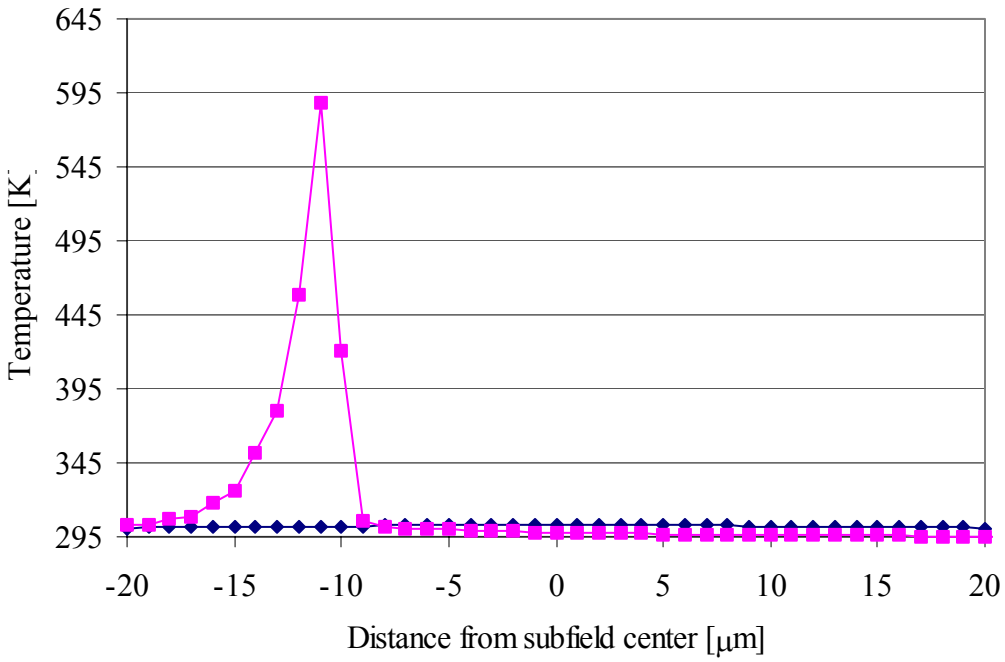


Figure 6.12.b. The temperature profile for the end of the subfield writing time is plotted for simulation the energy deposition with the averaging technique and the serpentine style writing. For the averaging technique the subfield surface temperature is plotted along the centerline of subfield. For the serpentine style writing the temperature is plotted along a line through the center of the moving beam parallel to the moving direction.

A comparison of the temperature profiles at the end of writing the subfield calculated with the averaging and with the serpentine style writing techniques is shown in Fig. 6.12.b. This figure illustrates the significant differences in the temperature profiles. The temperature is almost uniform and the rise of the subfield calculated with the averaging technique is very small. In contrast, the temperature profile for the serpentine writing shows a very large temperature rise of nearly 300 °C at the position that was last written. Even though the average temperature is lower, the maximum temperatures are much higher when using the detailed model.

Comparison of the local temperature gradients in the subfield in Fig. 6.12.b, illustrates another significant difference between simulating the beam writing with the averaging and the detailed technique. During writing of the subfield

very large local temperature gradients occur close to the written spot. The deposited beam energy is thus rapidly conducted away from the location of energy deposition. For the simulation of the writing process with the averaging method the gradients are lower and the deposited beam energy leaves the subfield slower by conduction. Therefore the average subfield temperature is higher for the e-beam writing simulation with the averaging technique.

The significant difference between simulating the subfield writing process with an averaging and a detailed technique suggest that a difference may also exist when comparing the averaging technique to the detailed simulation of a real pattern. It can be expected that the difference between the two techniques is smaller when using a real pattern instead of a test pattern. In a real pattern the beam writing spots are usually much smaller than used in the test pattern in Fig. 6.1. Therefore, the local temperature rises are smaller for the writing process of a real pattern. Because of this, the prediction error of the average subfield temperature with the averaging technique is expected to be smaller than for writing the test pattern. Although the peak temperatures of the subfield are expected to be smaller for a pattern with smaller features, a significant difference between the maximum subfield temperatures calculated with the two different simulation techniques is to be expected.

6.3 E-beam writing of a pattern on an X-ray mask subfield

6.3.1 Introduction

The goal of this section is to examine the e-beam writing process of a subfield on the geometry level of pattern shapes. The heat loss mechanisms and the mask temperature profile are compared to simulation results of subfield writing with averaged heat generation. The difference with the analysis of the second section is that in Section two the writing of a test pattern is examined. In this section a subfield writing process is investigated that represents closely the writing of a real integrated circuit pattern. For the analysis of the subfield writing process it is assumed that the patterning is done with a variable shaped beam system that writes in vector scanning fashion. Section two of Chapter three describes in detail the mask writing strategy of this kind of e-beam machine. In Fig. 3.4 this writing strategy is pictured for the writing of an IBM Talon mask [6.8]. This kind of mask is also used for the analysis in this chapter.

As described in Chapter three, the pattern of a subfield is divided into shapes of identical circuit pattern for writing a subfield with a variable shaped vector scan machine. A shape usually contains only a few basic integrated circuit elements. Identical shapes of circuit pattern are written sequentially as shown in Fig. 3.2. After the writing of one kind of shape is complete, the writing of the next shape is started. With this technique the complete pattern in a subfield is written. To approximate the shape writing process of a subfield without knowing a real circuit pattern it is assumed that there is only one kind of pattern shape in the subfield area. This kind of shape is assumed to have average characteristics of a typical pattern. For a heat transfer analysis of a writing process the most important shape characteristic is the shape size, since the size determines the amount of energy deposited at a shape location. For this reason it is assumed that in a subfield only shapes of average size are written. A vector scanning e-beam machine writes identical shapes in the subfield in arbitrary order. Therefore the writing process is sometimes described as random access writing [6.8]. To model the electron-beam writing of a subfield, a subfield area is divided into cells with the area of average shape size. The writing process is simulated by writing shapes in the subfield area in random order and at random locations. The total number of written shapes per subfield corresponds to the percentage of subfield pattern coverage. The percentage of pattern coverage describes, which fraction of the surface area is covered with pattern features. To correlate the number of written shapes per subfield to the percentage of pattern coverage it is assumed that within a shape the coverage is a hundred percent. The total number of written shapes N_{shapes} per subfield is calculated with eq. 6.3.1.1 with the total area $A_{subfield}$ of a subfield and the average area of a shape $A_{average\ shape.}$

$$N_{shapes} = \frac{A_{subfield}}{A_{average\,shape}} \cdot \text{fraction of surface area covered with pattern} \qquad [6.3.1.1]$$

For a Talon mask the size of a subfield is 37.5 μm × 37.5 μm. For the Talon mask with an average shape size of 0.4 μm^2, the number of shapes per subfield N_{shapes} is equal to 3516 times the fraction of pattern coverage. For a typical value of pattern coverage, as for

example twenty percent, the number N_{shapes} is equal 703. It is shown in the following section, that for a full Talon mask subfield model and the corresponding numbers N_{shapes} of written shapes, the simulation of the electron-beam writing process requires a unreasonable long computational time. For this reason the subfield writing process is analyzed for a reduced subfield size. The reduced size is chosen to be 18 μm × 18 μm, which is about a quarter of the actual subfield size. In Fig. 6.13 the pattern of a 18 μm × 18 μm subfield is shown for a pattern coverage of 17.5%. In this picture a grid subdivides the subfield area into cells of average shape size. The light blue spots in fig 6.13 represent the shapes that are written by the electron-beam. Only one of the spots is written at a time. The spots are written in random order.

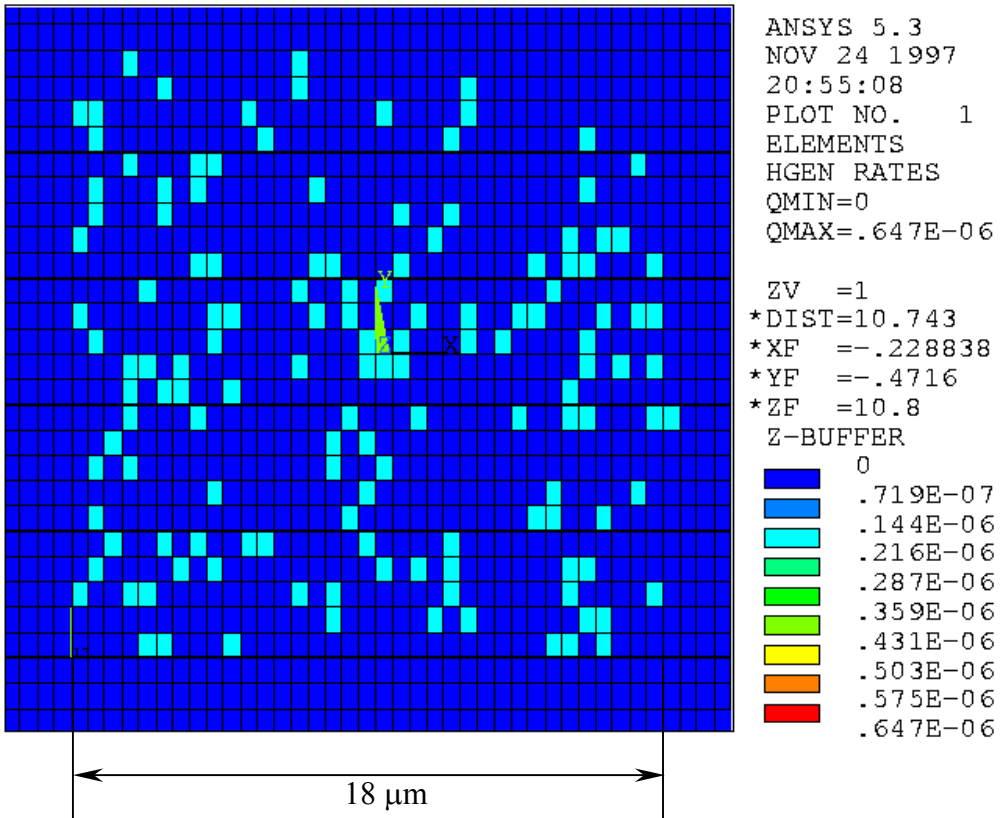


Figure 6.13. Pattern of a subfield for percentage coverage of 17.5%.

The geometry of the Talon mask is different from the geometry of the mask used in Section one of this chapter for the analysis of writing a beam test pattern. The dimensions of the talon mask, that are of importance for the analysis of the subfield e-beam writing process, are shown in Fig. 6.14. The Talon mask material properties are listed in Table 6.5. Because the material properties of SNR200, SiON and TaSi are not accurately known, the listed values are only estimated property values.

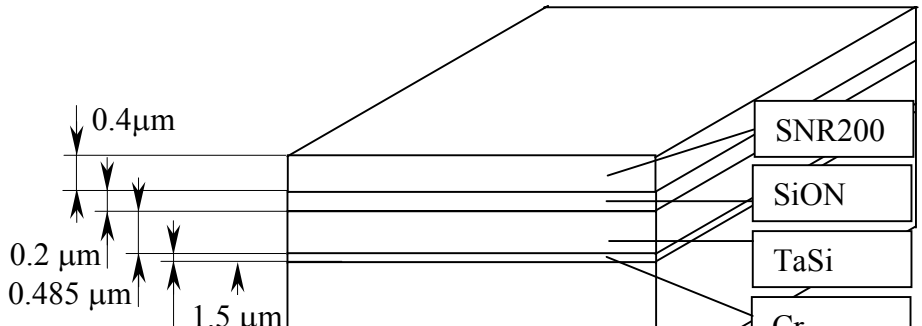


Figure 6.14. IBM Talon mask geometry.

Material	ρ [kg/m^3]	k [W/m - K]	c_p [J/kg - K]
SNR200	1200	0.2	1500
SiON	3000	30	750
TaSi	3000	100	700
Cr	7190	91	4605
SiC	3500	41	670

Table 6.5. Properties of Talon mask materials.

It is assumed that the subfield pattern is written with a variable shaped electron-beam system with an acceleration voltage of 75 kV and a current density of 30 A/cm^2. In Section two of Chapter six it is explained that the increase of the beam size with increasing depth in the mask material, caused by electron scattering, can be neglected for the writing of an X-ray mask. For the simulation of the Talon mask subfield writing, the variation of the deposited beam energy with depth in the material is taken into account. The results of a Monte Carlo simulation are used to determine the energy distribution in the depth direction of the mask [6.9]. The fraction f_i of the total deposited energy that is deposited in each mask material is listed in Table 6.6. Because the X-ray mask has the small thickness of 2.6 μm, only seven percent of the incident electrons are absorbed in the mask material. With the overall mask absorption coefficient α of 0.07 and the absorption coefficients f_i the volumetric heat generation $q'''_{Shape}(Material_i)$ for each mask material layer, caused by writing a single pattern shape, can be calculated with eq. 6.3.1.2.

$$q'''_{Shape}(Material_i) = \frac{\alpha \cdot f_i \cdot (V_{beam} \cdot \text{Sensitivity of Resist})\ [J/\mu m^2]}{\Delta t_{Shape}\ [s] * \text{Thickness}_{Material_i}\ [\mu m]} \qquad [6.3.1.2]$$

In eq. 6.3.1.2 V_{beam} is the beam acceleration voltage. The resist sensitivity of SNR200 is 20 μC/cm^2. The writing time of a pattern shape Δt_{Shape} of 0.667μsec can be calculated with eq. 6.2.1.1.

Material	f_i Fraction of total energy deposited in this material
SNR200	0.042

SiON	0.036
TaSi	0.314
Cr	0.013
SiC	0.595

Table 6.6. Fraction of the energy deposited in each material layer [6.9].

The process parameters that describe the writing of an X-ray mask subfield with pattern shapes in random order are summarized in Table 6.7. As listed in this table the stepping time necessary to move the beam from shape to shape is neglected. The stepping time of a variable shaped beam vector scan machine is, as written in [6.4], about 10 nsec, which is very small compared to the writing time of a single pattern shape.

Shape writing time Δt_{Shape}	0.6667 μsec
Waiting time between shapes	0 μsec
Size of average shape	0.5 μm × 0.75 μm
Volumetric heat generation in Material $_i$	f_i * 0.7875E-2 W/μm^3
Writing style	Random access writing

Table 6.7. Summary of mask loading conditions for writing an X-ray mask subfield with random order pattern shape writing.

The heat loading conditions of the mask for the simulation of the subfield e-beam writing with an averaging technique are determined as in Section one of this chapter. The writing time for the complete subfield area is calculated with eq. 6.2.1.3. Instead of the number of written spots N_{Spots} and the spot writing time Δt_{Spot}, the corresponding values N_{Shapes} and Δt_{Shape} for the writing of a pattern shape are used in equation 6.2.1.3. The number of written shapes in equation 6.2.1.3 is calculated with eq. 6.3.1.3, the ratio of subfield and shape area, and the percent pattern coverage of the subfield.

$$N_{Shapes} = \frac{\text{percent pattern coverage}}{100} \cdot \frac{A_{Subfield}}{A_{Shape}} \qquad [6.3.1.3]$$

The uniform volumetric heat generation in the mask subfield is determined as in Section two with eq. 6.2.1.6. In this equation the variable N_{Spots} is replaced by N_{Shapes} and for q'''_{spot} the variable $q'''_{Shape,av}$, calculated with eq. 6.3.1.4, is used.

$$q'''_{Shape,av} = \frac{\alpha \cdot \left(V_{beam} \cdot \text{Sensitivity of Resist}\right)\ [J/\mu m^2]}{\Delta t_{Shape}\ [s] * \text{Thickness}_{membrane}\ [\mu m]} \qquad [6.3.1.4]$$

The loading conditions that correspond to the approximation of the test pattern writing by an averaging technique are summarized in Table 6.8.

Subfield writing time $\Delta t_{Subfield}$	Depends on pattern coverage
Size of subfield	18 μm × 18 μm
Volumetric heat generation	Pattern coverage*0.6058E-3 W/μm^3
Writing style	Uniform applied heat generation

Table 6.8. Summary of mask loading conditions for simulating the writing of a serpentine test pattern on an X-ray mask with an averaging technique.

6.3.2 The finite element model of the X-ray mask subfield

For the simulation of the Talon mask subfield writing process a finite element model is used that is similar to the one described in Section one of Chapter six. As explained in Section one, it is sufficient for the analysis of the subfield writing process to model only a part of the complete patterned mask region. An appropriate size of this region, corresponding to the subfield writing conditions listed in Table 6.7 and Table 6.8, has to be determined. As written in Section one, the heat transfer in the mask membrane is described as a three dimensional problem only for a square center region of the model. The size of this region, modeled with cuboid finite elements, needs to be set corresponding to the subfield size and mask loading conditions listed in Table 6.7 and Table 6.8. As in Section one, the appropriate size for the three-dimensional finite element model region is estimated by examining the heat flux pattern in the mask membrane for writing a single pattern shape. The writing conditions are listed in Table 6.7. The heat flux pattern in the mask membrane, shown in Fig. 6.15, is calculated with the finite element model, which is described further in the following paragraph of this section. The vector plot in fig 6.15 shows that significant temperature gradients in membrane depth direction are limited to a small region around the written shape. Therefore it is sufficient to model only a membrane region of 28.8 μm × 28.8μm, what is a little bit larger than the 18 μm × 18μm subfield, with cuboid elements.

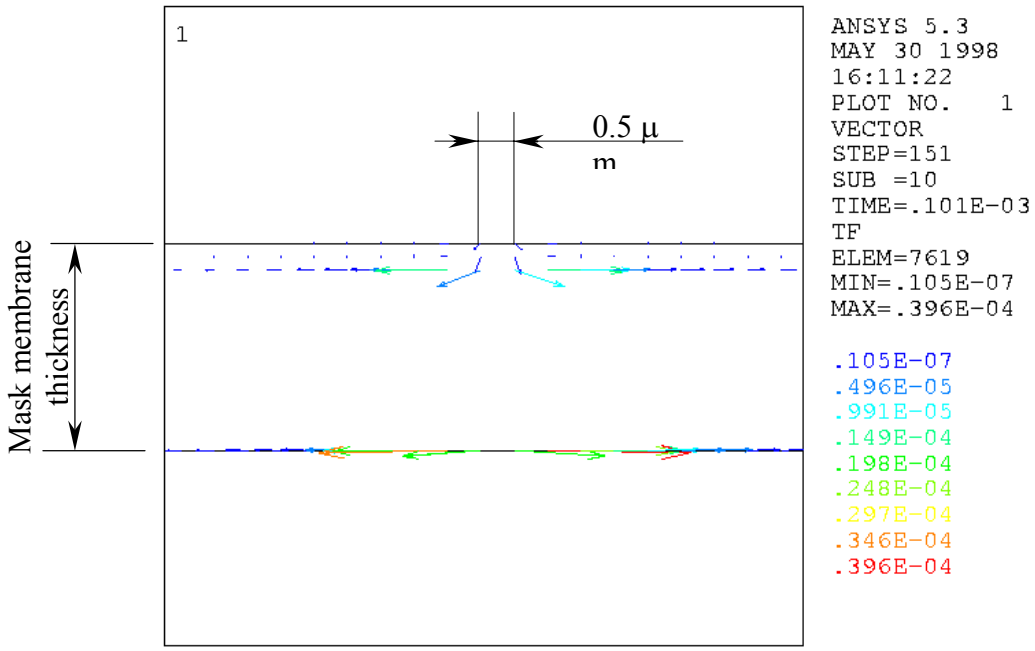


Figure 6.15. Heat flux vectors in the mask membrane for writing a single shape on the mask membrane with writing conditions listed in Table 6.7.

The size of the circular membrane region, shown in Fig. 6.3, modeling the heat transfer in the mask as a two-dimensional problem, is determined as in Section one. The mask temperature profile is calculated by simulating the complete subfield writing process. Then it is estimated how much energy diffuses to the edges of the finite element model during the subfield writing time by checking the mask temperature rise at the edges of the model. Because adiabatic boundary conditions are used for the edges of the model, a temperature profile as shown Fig. 6.5 describes how much energy diffuses to the model edges. For the simulation of the complete subfield writing process a pattern coverage of thirty percent is assumed. The subfield writing process is simulated with the averaging technique using the parameters in Table 6.8. In fig 6.16 the temperature rise at the model edge is plotted versus the corresponding model radius.

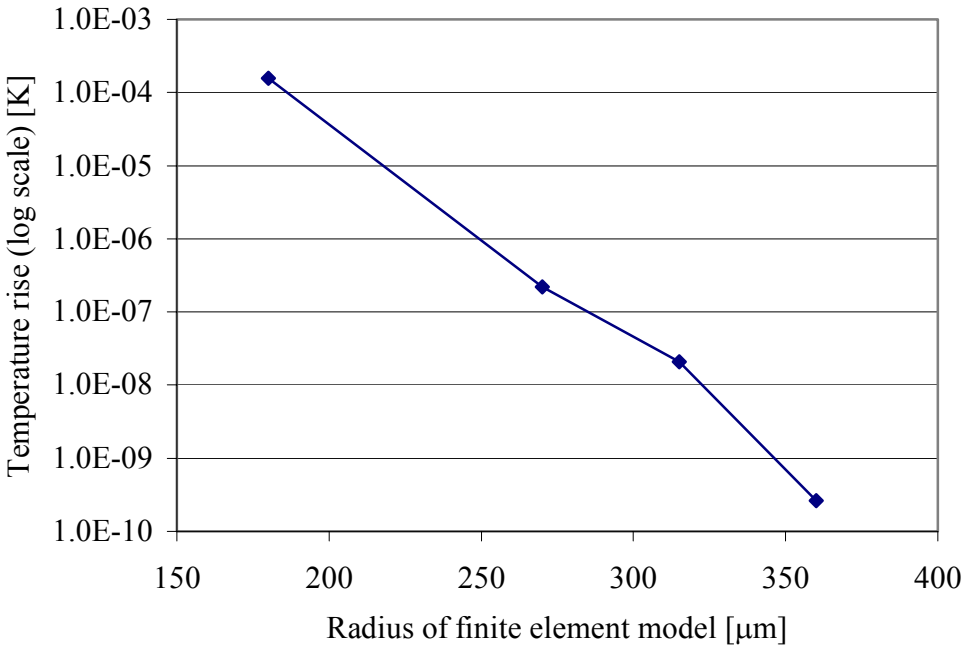


Figure 6.16. Temperature at the edge of the finite element model for simulation of subfield writing process with parameters in Table 6.8 and pattern coverage of 30%.

For the final model a radius of 315 μm, corresponding to a temperature rise of 2.1E-8 °C at the model edge, is used. A summary of the model dimensions is listed in Table 6.9.

Size of subfield	18 μm × 18 μm
Size of the three dimensional finite element region	28.8 μm × 28.8 μm× 2.6 μm
Radius of the two dimensional finite element region	315 μm

Table 6.9. Summary of finite element model dimensions used for simulating the writing of a Talon mask subfield.

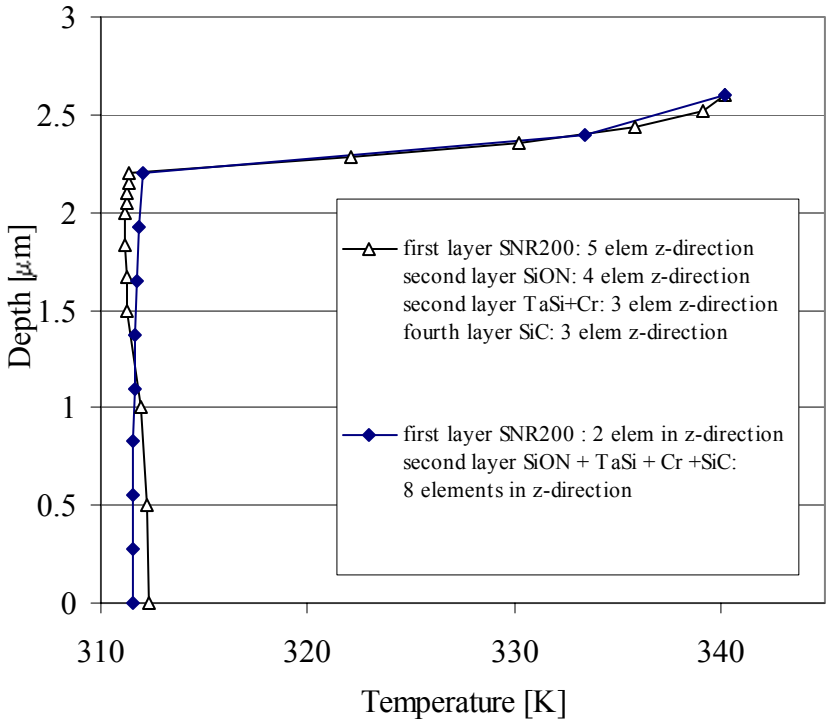


Figure 6.17. Temperature profile in mask depth direction at subfield center for writing and area of 4×4 shapes at subfield center. The writing conditions are listed in Table 6.7.

After determining the dimensions of the finite element model, an adequate mesh width of the model is determined. To optimize the number of elements in mask depth direction, the temperature profile in depth is calculated for varying numbers of elements per material layer and for substituting different combination of the five mask materials with a material of equivalent properties. The depth temperature profiles are determined at the center of the subfield area for writing 4×4 shapes at the subfield center with the writing conditions listed in Table 6.7. In Fig. 6.17 temperature profiles are plotted for modeling the mask with different combinations of materials substituted by a new material with averaged material properties. In one case only the tantalum silicate layer is combined with the chrome layer. The chrome layer is considered to have only a negligible influence on the heat flux pattern in the mask, because of its small thickness. Averaged material properties for this layer are calculated with eq. 6.2.2.1, eq. 6.2.2.2 and eq. 6.2.2.3. This temperature profile, representing the exact profile of the mask shown in Fig. 6.14, is compared to a profile of a model with combined tantalum silicate, chrome, siliconoxinitride and silicon carbide layer. The comparison in fig 6.17 shows that the small variation of the temperature in the four bottom layers can be approximated with sufficient accuracy by the profile calculated with the model that combines the bottom material layers. In Fig. 6.18 temperature depth profiles are plotted calculated with finite element models with a different number of elements per material layer. All profiles are calculated with models combining the four bottom material layers in a layer with equivalent properties. The comparison shows, that a variation of the number of

elements in the bottom material layer does not change the shape of the profile. For this reason only one element is used to model the four bottom material layers combined in one material with equivalent properties. The comparison in Fig. 6.18 also shows that the modeling the photo resist layer with only one element layer in depth is not sufficient to approximate the exact temperature profile. Therefore two element layers are used to model the SNR200 material. Figure 6.17 comparing profiles for two and five element layers for the photo resist shows that it is not necessary to use more than two element layers for the SNR200 material.

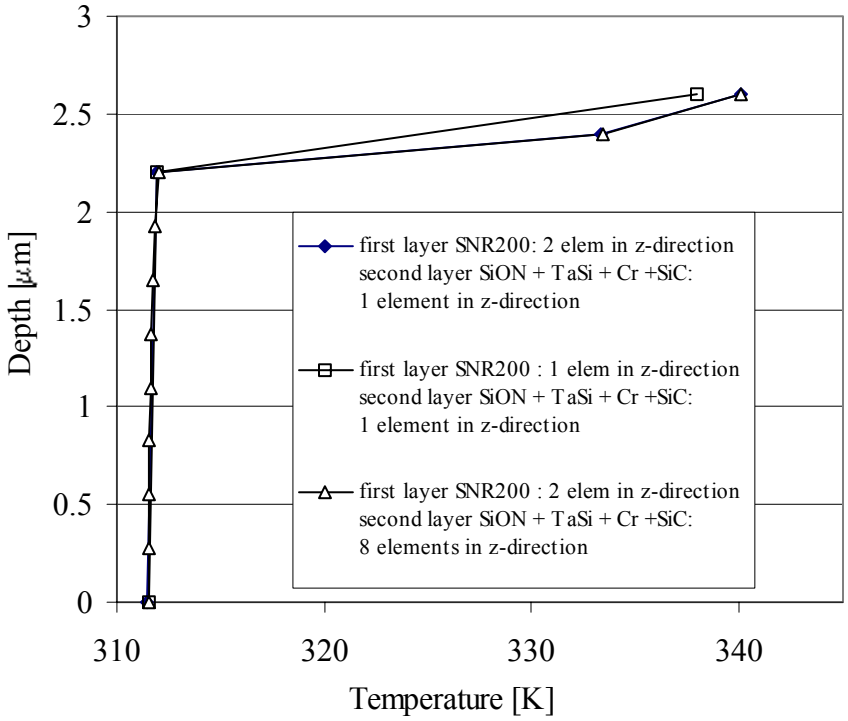


Figure 6.18. Temperature profile in mask depth direction at subfield center for writing and area of 4×4 shapes at subfield center. The writing conditions are listed in Table 6.7.

To limit the number of necessary elements and nodes of the complete model the grid width in x-y direction is chosen equal to the dimensions of a single shape. As shown in the following section, the temperature gradients in the subfield are not as high as for the writing of a subfield test pattern that is described in Section one of Chapter six. For this reason a mesh width of 0.5 in x-direction and 0.75 in y-direction is assumed to be small enough.

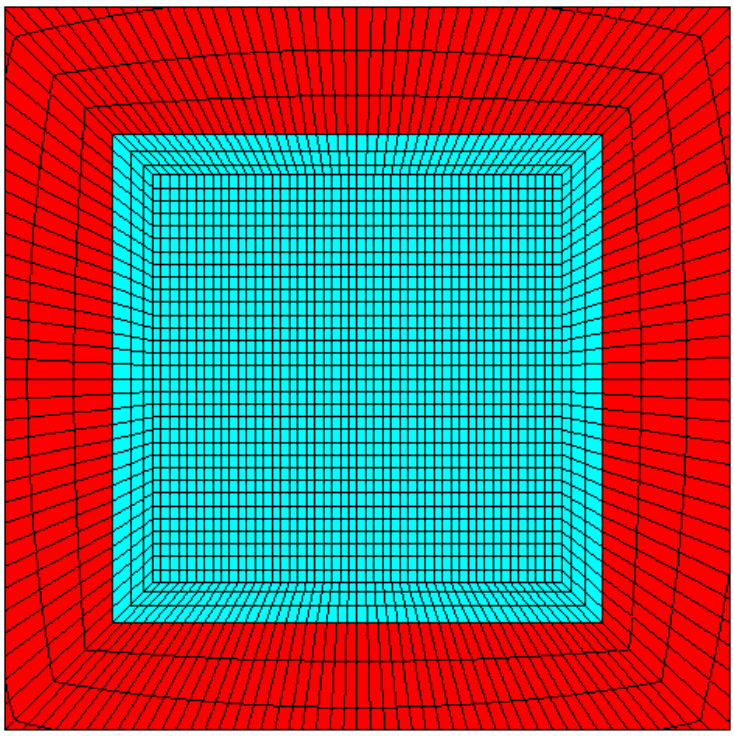


Figure 6.19. Finite element model used for simulation of subfield writing process.

In Fig. 6.19 the finite element model is shown with the dimensions and mesh width determined in the previous paragraphs. The square turquoise area in the center is the three-dimensional finite element region modeled with cuboid elements. The elements shown in red color are shell elements with equivalent material properties. The equivalent material properties are calculated as shown in Section one of Chapter six. Not shown in Fig. 6.19 is a layer of surface effect elements covering the shell and cuboid element region. The surface effect elements model the are needed to model the radiation loss at the top and bottom surface of the mask. For the emissivity a value of 0.19 is used.

6.3.3 Comparison of subfield electron-beam writing simulation with detailed and with averaging method

This section of Chapter six describes the results of subfield electron-beam writing simulations with the detailed and averaged technique. Both techniques are described in Section 6.3.1 and the writing conditions are summarized in Table 6.7 and Table 6.8. The ANSYS® finite element model used for the simulations is described in Section 6.3.2. The model dimensions are summarized in Table 6.9. The complete finite element model has 16480 elements and 11590 nodes. Simulating a writing process with the detailed method for a subfield with thirty percent pattern coverage requires computational time of 65 hours on an HP 9000/780 with 512 Mbyte RAM.

The surface temperature profile of the subfield with fifteen percent pattern coverage is shown in Fig. 6.20.a for the end of the writing process. The maximum temperature rise that occurs at the location of the last written shape is 12 °C for a surroundings temperature of 294.26 °C. The figure shows temperature peaks at five different shape locations. These are the locations of the last five written shapes. The height of the peaks depends on how much time has passed since the writing of the corresponding shape. It is characteristic for the temperature profile, that significant temperature rises occur only at the shape writing location. For the rest of the subfield area Fig. 6.20.a does not show a noticeable temperature rise. This indicates that the energy, deposited during the writing of a single shape, is rapidly conducted away from the deposition location. This is also proven by the fact that only five temperature peaks are visible in Fig.6.20.a. The subfield temperature rise resulting from writing the shapes before the last five shapes is not noticeable, because of the rapid conduction of deposited energy away from shape location. The high conduction speed is visualized also in Fig. 6.20.b to Fig.6.20.g. These figures show a time sequence of cooling of a subfield after a pattern with fifteen percent pattern coverage is written. The figures show the cooling process with time intervals of one shape writing time Δt_{shape}. Figure 6.20.f shows, that all

temperature peaks have disappeared after five writing times Δt_{shape}, 3.3 μsec, after the end of writing.

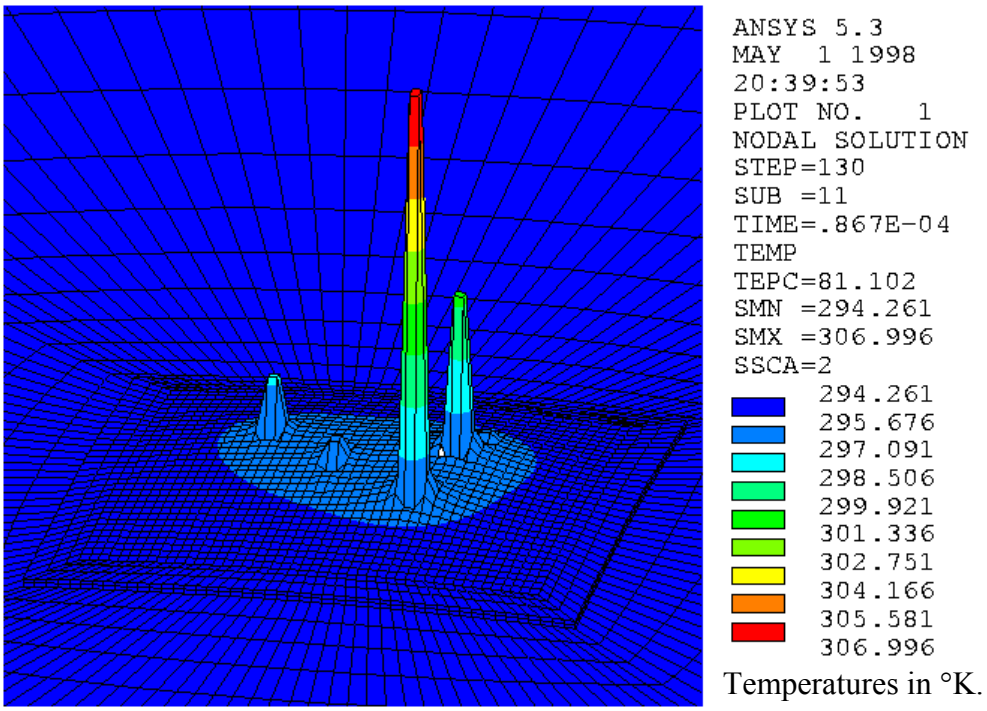


Figure 6.20.a. Temperature profile of subfield with 15 % pattern coverage at end of writing.

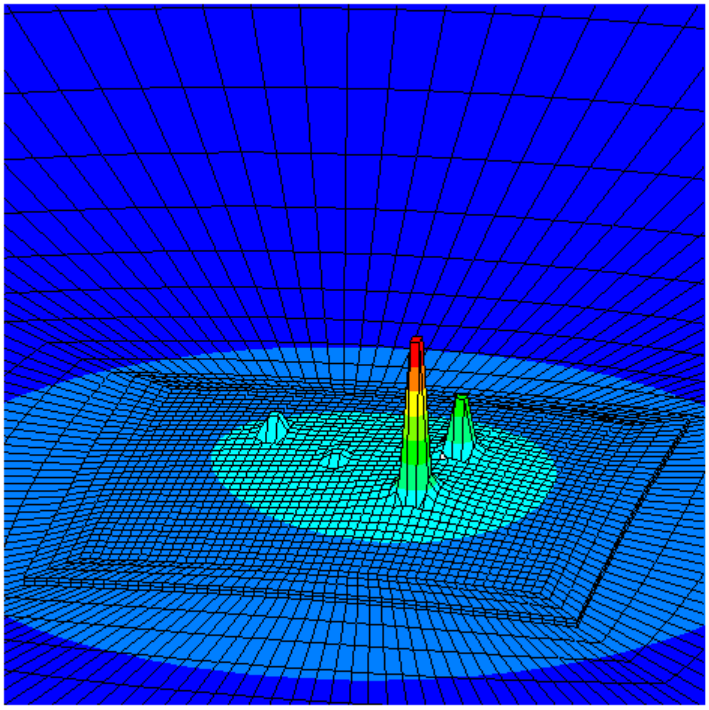


Figure 6.20.b. Temperature profile of subfield with 15 % pattern coverage at $1 \times \Delta t_{shape}$ after end of writing.

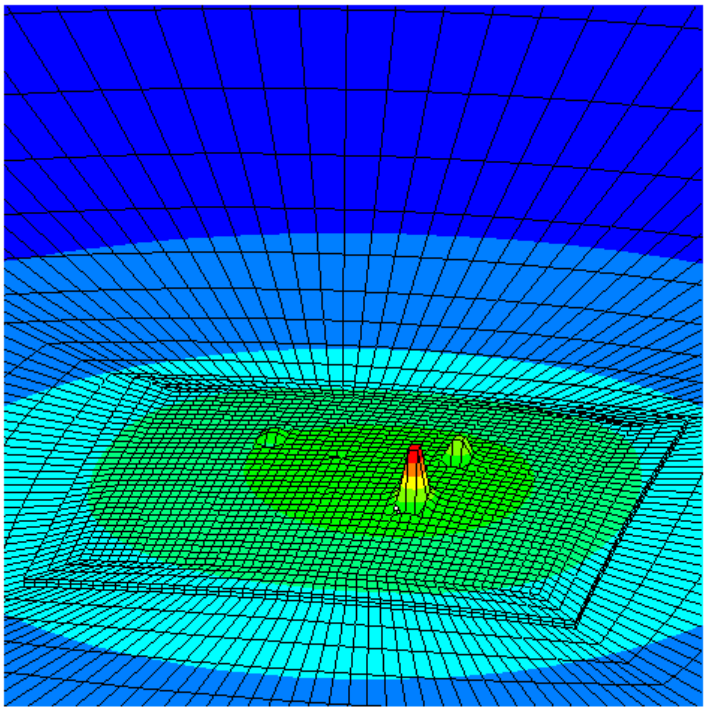


Figure 6.20.c. Temperature profile of subfield with 15 % pattern coverage at $2 \times \Delta t_{shape}$ after end of writing.

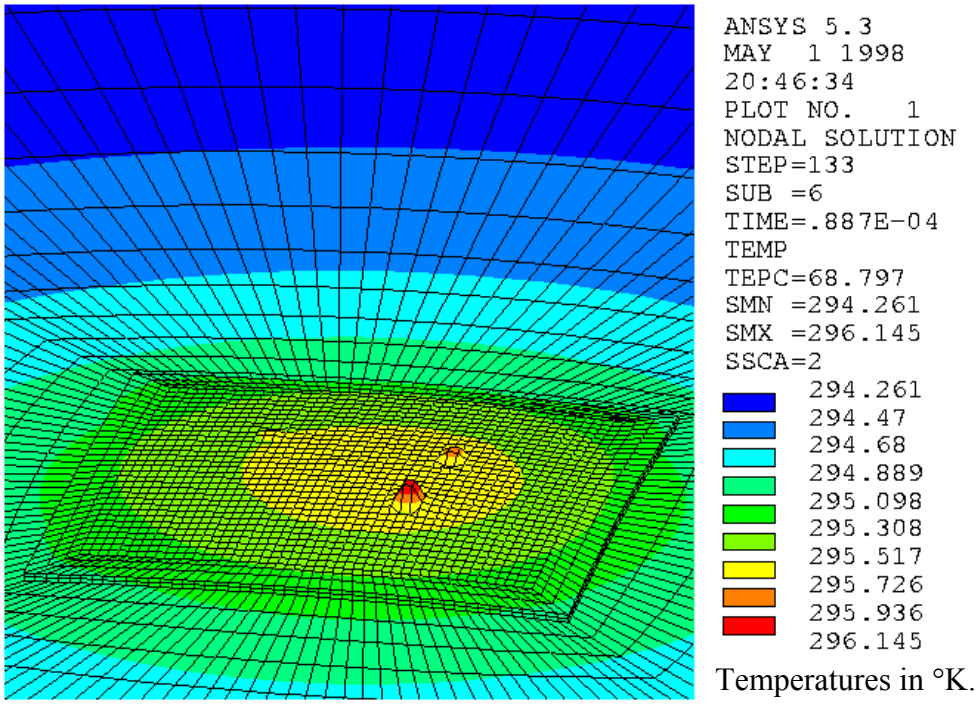


Figure 6.20.d. Temperature profile of subfield with 15 % pattern coverage at $3 \times \Delta t_{shape}$ after end of writing.

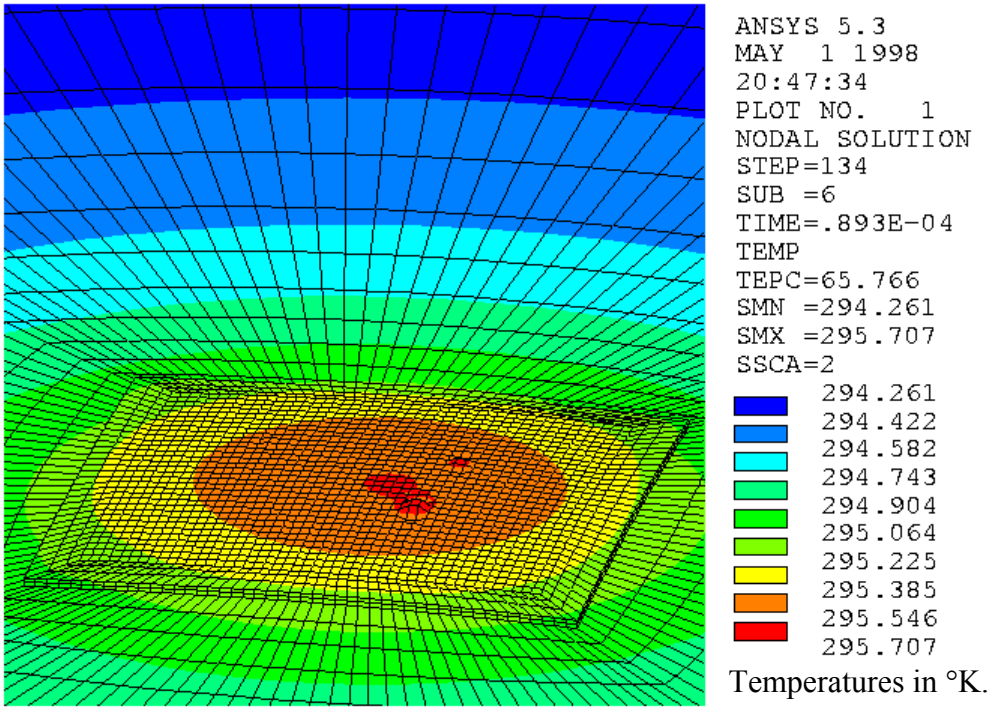


Figure 6.20.e. Temperature profile of subfield with 15 % pattern coverage at $4 \times \Delta t_{shape}$ after end of writing.

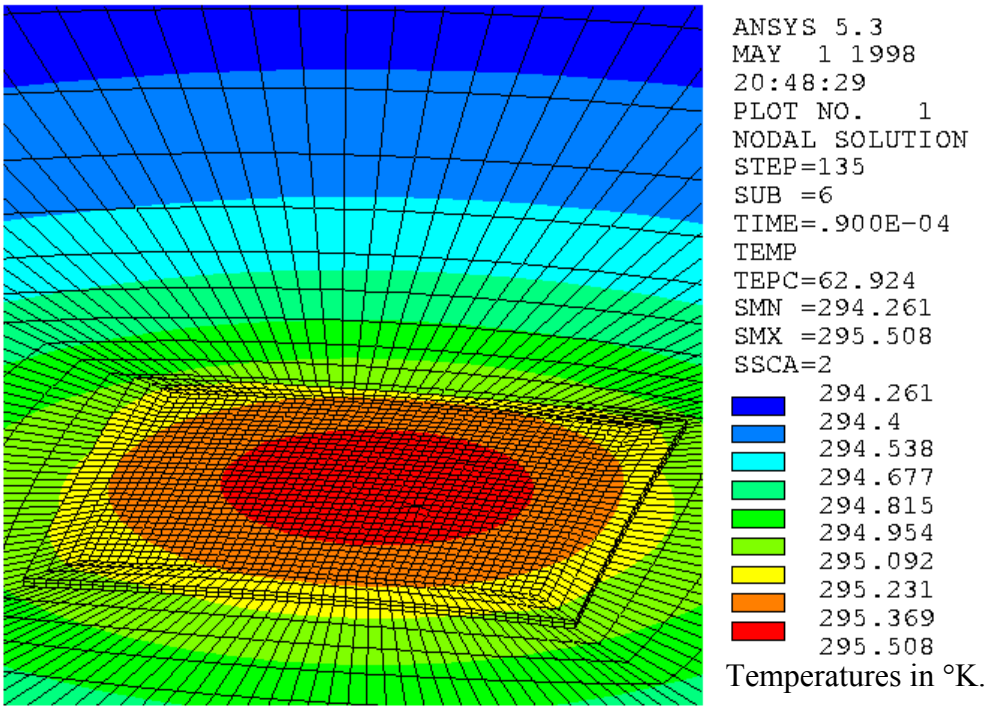


Figure 6.20.f. Temperature profile of subfield with 15 % pattern coverage at $5 \times \Delta t_{shape}$ after end of writing.

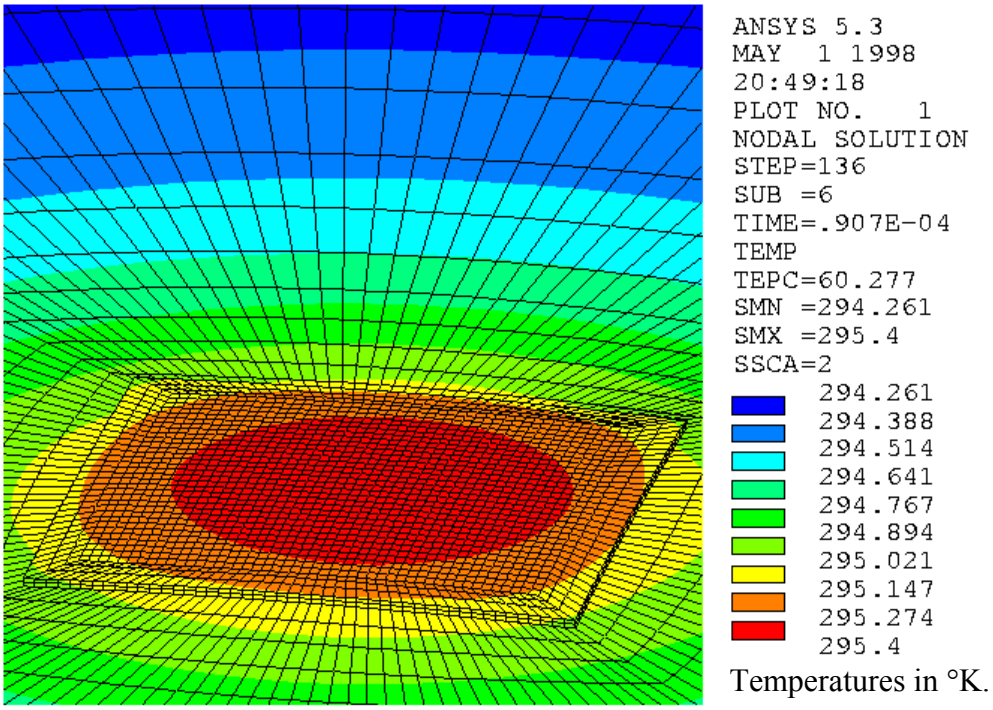


Figure 6.20.g. Temperature profile of subfield with 15 % pattern coverage at $6 \times \Delta t_{shape}$ after end of writing.

In the following paragraph the heat transfer mechanisms in the mask are compared for the subfield electron-beam writing process simulated with the detailed and the averaging technique. To compare the the radiation loss mechanisms for the two simulation techniques the total energy that is lost to the surroundings by radiation in a certain writing time is calculated with eq. 6.2.3.1, as it is done in Section 6.2.3 for the subfield test pattern writing.

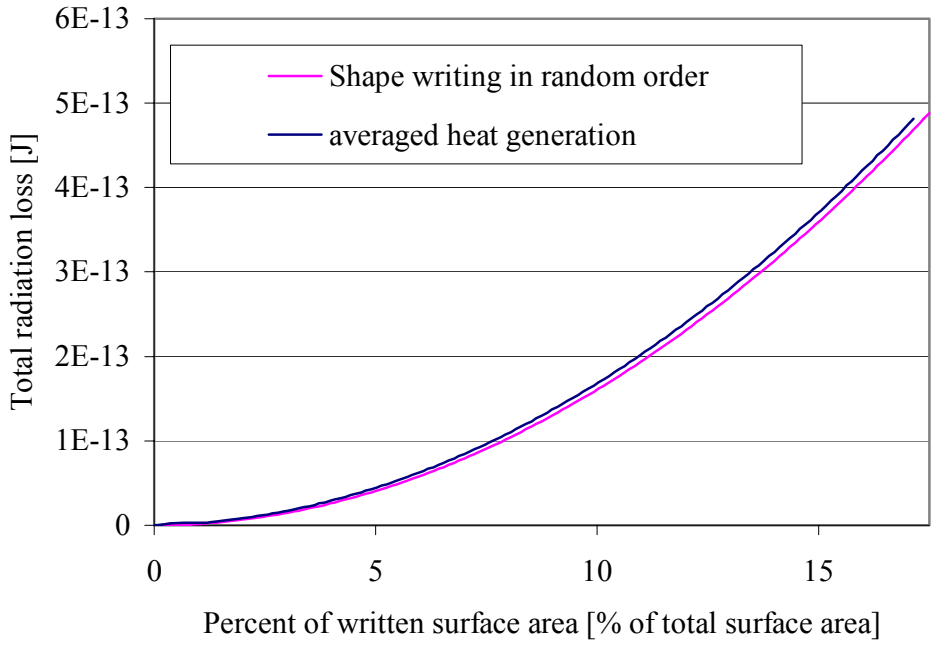


Figure 6.21. Total energy lost by radiation during the electron-beam writing process vs. the fraction of surface area that is already written.

In Fig. 6.21 the energy $E_{\text{radiation loss}}$ is plotted versus the fraction of the total surface area that is already written. The values of $E_{\text{radiation loss}}$ are plotted for writing the subfield with pattern shapes in random order and with an equivalent averaged heat generation. Comparing the two total radiation loss curves, calculated with the different simulation techniques, shows the good agreement between both cases. The reason for a better agreement than shown in Section 6.2 is, that the peak temperatures that occur during the shape writing simulation are much smaller than those that occur during the serpentine style spot writing. Another reason for the better agreement might be that the writing of shapes at random location in the subfield creates a more uniform energy distribution in the subfield than the serpentine style subfield writing. In Section 6.2 it is shown that only a negligible amount of the deposited beam energy is lost by radiation. Because the peak temperatures, occurring in the subfield for the shape writing at random locations, are smaller than for the serpentine spot writing, it can be assumed that radiation is also negligible for the subfield shape writing.

The comparison of the conduction heat transfer in the mask for the two simulation techniques is done as in Section 6.2 by comparing average surface

temperature of the subfield area at the end of writing. In Fig. 6.22 the average subfield surface temperature is plotted versus the percentage of subfield pattern coverage for the subfield writing simulated with the shape writing in random order and an applied average heat generation. The comparison shows that a simulation of the subfield writing process with the averaging method underpredicts the actual average subfield temperature. The difference between the average subfield temperature rises calculated with the two techniques for the surroundings temperature of 294.26 °C is fifteen percent for the pattern coverage interval form zero to thirty percent. In fig 6.22 the small fluctuation in the average subfield temperature for the simulation of the writing process with shape writing occur because the plotted average temperature is calculated only for the subfield area. The calculated average surface temperature depends upon how close to the edge of the subfield the last shapes are written. When the last shapes are written close to the edge of the subfield, the corresponding average subfield temperature is lower than for a case when the last shapes are written close to at the subfield center. When writing occurs near the center less energy diffuses out of the subfield region before the time for which the average subfield temperature is calculated.

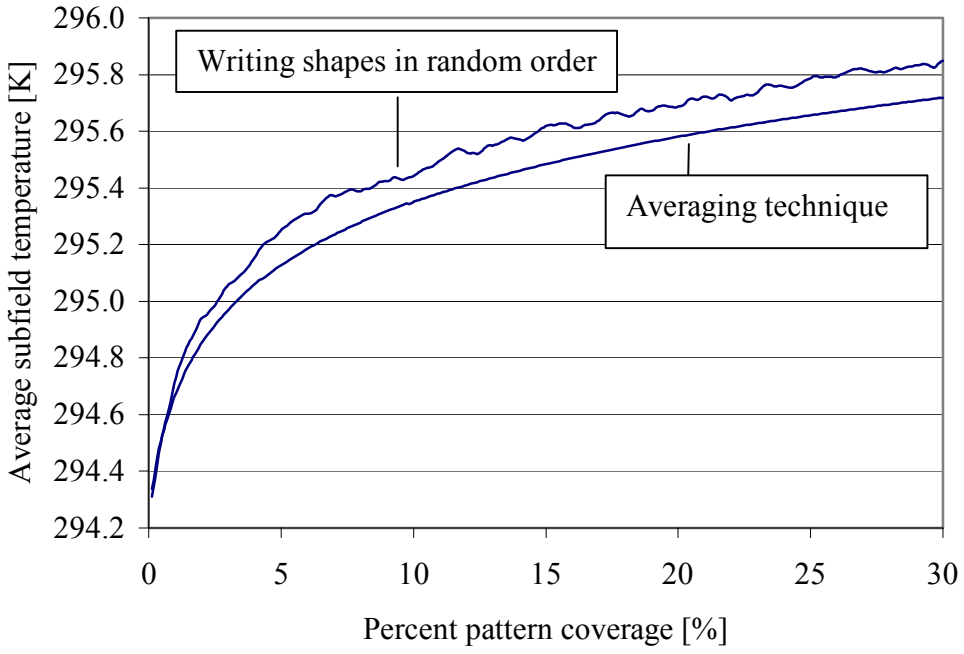


Figure 6.22. Comparison of average subfield temperature vs. percent pattern coverage for simulation of subfield writing process with shape writing in random order and averaged heat generation.

In Fig. 6.23 the highest subfield peak temperature is plotted versus the percent pattern coverage. The maximum peak temperature depends on how close consecutive shapes are written. When shapes are written close to each other, the writing of a previous shape heats up the subfield region where the following shapes are written. For example, the two temperature peaks in fig 6.23, at 6 and 28% coverage show a significantly higher temperature rise because two adjacent shapes were written consecutively. The maximum and average subfield temperatures for simulation of subfield writing process with shape writing in random order, given in Fig. 6.24, show the significant temperature difference between maximum and

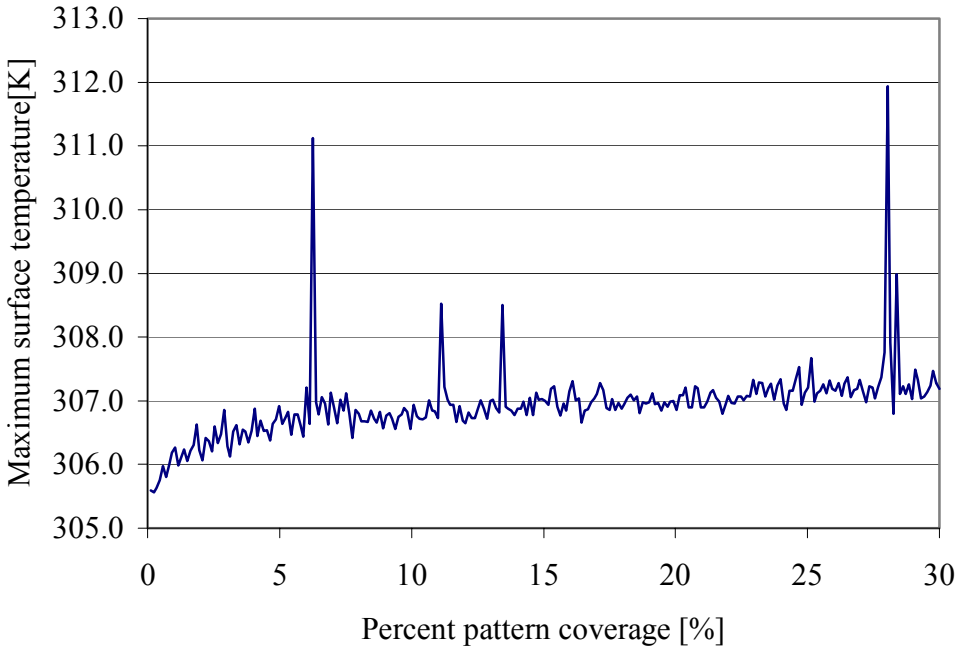


Figure 6.23. Maximum subfield temperature vs. percent pattern coverage for simulation of subfield writing process with shape writing in random order.

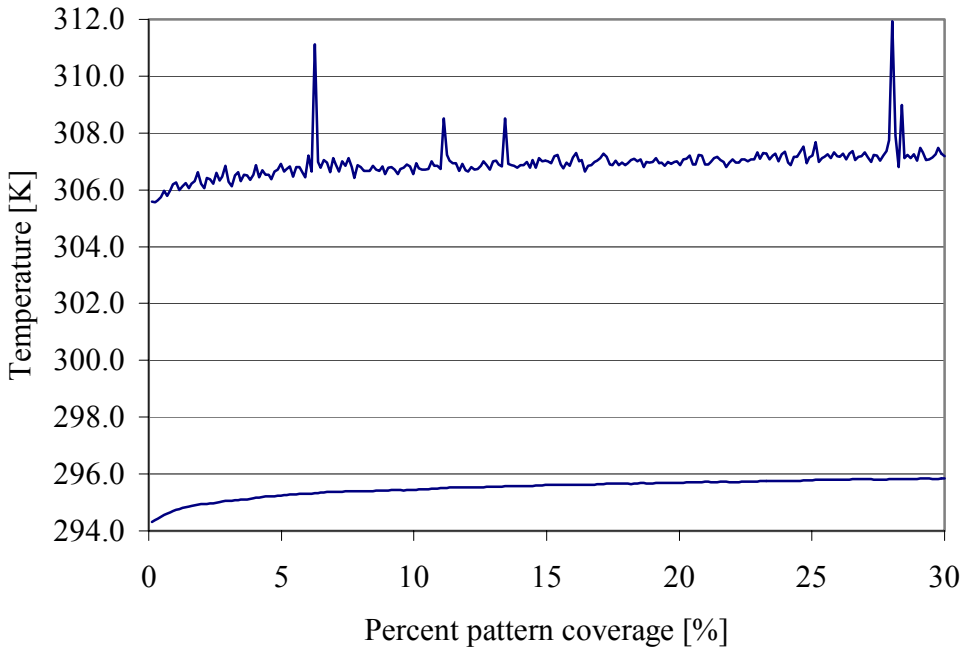


Figure 6.24. Maximum and average subfield temperature vs. percent pattern coverage for simulation of subfield writing process with shape writing in random order.

average temperature. The average temperature is also the maximum temperature predicted by the averaging technique. It is seen that the averaging technique significantly under predicts the maximum temperature, but it predicts the average subfield temperature with a reasonable accuracy.

6.4 References

[6.1] F. Bedford, R. L. Engelstad, and F. Cerrina, “Thermal Analysis of an X-Ray Mask Membrane during Electron-Beam Patterning,” *Proceedings of TECHCON*, 1996.

[6.2] D. L. Laird, R. L. Engelstad, D. M. Puisto, R. E. Acosta, K. D. Cummings, and W. A. Johnson, "Predicting in-plane Distortion from Electron-Beam Lithography on X-ray Mask Membranes,” *J. Vac. Sci. Technol. B*, 14, 1996.

[6.3] B. Shamoun, M. Sprague, F. Bedford, and R. Engelstad, “X-ray Mask Distortions during E-Beam Patterning," to appear in *Microelectronic Engineering*, 1998.

[6.4] H. C. Pfeiffer, D. E. Davis, "EL-4 a new Generation Electron-Beam Lithography System,” *J. Vac. Sci. Technol. B*, 11, 1993.

[6.5] P. Rai-Choudhury, "Microlithography, Micromachining, and Microfabrication," Vol. 1, 1997.

[6.6] F. P. Incropera, D. P. DeWitt, "Introduction to Heat Transfer," Wiley, 3rd edition, 1934.

[6.7] ANSYS® Help System, Release 5.4, 1997.

[6.8] D. M. Puisto, IBM, private communication.

[6.9] K. D. Cummings, private communication.

Chapter Seven

Conclusions and Recommendations

7.1 Conclusions

The analysis in Chapter five of optical mask direct heating by writing a square electron-beam flash of 1 μm× 1 μm has shown that a significant temperature rise occurs in the photoresist layer. The calculated mask surface temperature profile shows that the temperature rise is not uniform in the writing area. Due to the temperature dependence of the resist sensitivity, local under- or over-exposure of the resist may occur when writing a spot with a beam of uniform electron distribution. This effect is expected to be a source of patterning inaccuracies. The in-depth temperature distribution of the optical mask shows that the thermal properties of the thin chrome layer between the resist and the bulk glass material significantly affect the local temperature distribution. The chrome layer has a high thermal conductivity and a significant amount of electron-beam energy is rapidly transported away from the deposition location. A detailed model is necessary to evaluate the local temperature distribution.

The study of local mask heating during X-ray mask electron-beam writing contained in Chapter six provides an estimate of the accuracy of an averaging technique for heating that has been used in other studies to approximate the global mask heating. It is shown that for a small region of the mask that is patterned, the averaging technique approximates the average mask temperature rise of the actual writing process with an accuracy of about fifteen percent.

The study also shows that using the averaging technique to predict local heating effects, as for example the temperature rise in the photoresist layer, significantly underpredicts the maximum temperature rise. The prediction of stress relief in the photoresist layer may be affected.

In Chapters five and six a procedure for modeling the local mask heating with a finite element analysis is presented. In Chapter five it is shown that to analyze local heating in an optical mask it is necessary to take into account the radial spread of deposited beam energy with increasing depth in the material. Because of the large thickness of an optical mask, electron scattering in the mask is important so that it necessary to know the spatial energy deposition function in order to obtain a prediction of the mask temperature profile.

In Chapters five and six it is shown that to predict local mask temperature profiles it is sufficient to model only a small part of the lithography mask with a finite element model. A method for estimating the necessary model size to obtain a good approximation of the mask temperature profile has been developed. Methods for estimating the optimal mesh size of a finite element model are presented.

In Chapter six the modeling of a lithography mask with a combination of three-dimensional and two-dimensional finite elements for the calculation of local mask temperature profiles is described. It is shown that approximating the actual mask structure with two-dimensional finite elements and averaged material properties can significantly reduce the size of the finite element model and the computational time without reducing the accuracy of the calculated temperature profile.

7.2 Recommendations and Future Work

In the analysis in Chapter six the accuracy of the averaging technique for approximating the global mask temperature profile is estimated for a small pattern area of 18 μm × 18 μm size. Because the averaging technique is often used to approximate the global mask temperature over areas much larger than this size, the accuracy of the averaging method needs to be evaluated as a function of pattern size. The accuracy of the energy averaging technique is evaluated in Chapter six only for one mask geometry and construction. The averaging technique accuracy should be evaluated for other specific mask configurations and writing styles. The procedure developed in this thesis will aid the evaluation of the averaging method.