PROCEEDINGS OF THE SYMPOSIUM ON

ELECTRON AND ION BEAM SCIENCE AND TECHNOLOGY

SEVENTH INTERNATIONAL CONFERENCE

Edited by

Robert Bakish
Bakish Materials Corporation
Englewood, New Jersey

ELECTRONICS AND ELECTROTHERMICS AND METALLURGY DIVISIONS

THE ELECTROCHEMICAL SOCIETY, INC., Post Office Box 2071, Princeton, New Jersey 08540

PROCEEDINGS OF THE SYMPOSIUM ON

ELECTRON AND ION BEAM SCIENCE AND TECHNOLOGY

SEVENTH INTERNATIONAL CONFERENCE

Edited by

Robert Bakish
Bakish Materials Corporation
Englewood, New Jersey

ELECTRONICS AND ELECTROTHERMICS AND METALLURGY DIVISIONS

THE ELECTROCHEMICAL SOCIETY, INC., Post Office Box 2071, Princeton, New Jersey 08540

Library of Congress Catalog Card Number: 71-120300
Printed in the United States of America

PREFACE

Proceedings preparation time is also time to reflect upon the meeting just passed. This meeting would not have taken place were it not for the efforts of the Conference committee, the session organizers, the support of the companies with which they are associated, and of course, you, the speakers and your employers. I wish to express to all of you my and The Electrochemical Society's thanks.

Thanks are also due to the Society Headquarter's staff for their help before and during the meeting, as well as in the printing of the proceedings. It has been a real pleasure to work with all of you, and I hope that the few inconveniences which some of you experienced, including the not entirely adequate meeting room for our session, have not upset you and will not prevent your participation in the sequel meetings of this conference. One thing I wish to assure you here is that at the 1976 meeting, the 8th International Conference on Electron and Ion Beam Science and Technology in Vancouver, B.C. Canada, you will have a 200 capacity room. I sincerely hope, and indications at hand are such, that it is highly probable that in our future meetings we will see even greater growth and returns for all those involved in this technology.

Now to the proceedings volume. There is considerable variety of subject matter in the papers presented at the 7th International Conference on Electron and Ion Beam in Science and Technology. This being the case it is always a problem to arrange the papers in the proceedings for a most orderly presentation of the material. After considerable thought and review of the material I have decided to group the papers in 3 main sections: Section I deals with all high power electron beam related subjects, and Section II will contain all papers dealing with low power electron beams. The second section itself will have 2 subsections, one handling miscellaneous low power beam topics, and the other x-ray and electron lithography and microfabrication systems. Section III will contain the papers on ion implantation.

I have used the key words provided by the author and those generated by me in reviewing the papers to provide a limited index to the proceedings. It is my belief that it will be quite adequate for the reader. Two papers, those of R.S. Payne and A. N. Broers, are not indexed. They arrived late and the writer preferred to include them unindexed rather than omit them from the proceedings. Both are review papers and important to our readership.

As you can see from the content of the papers here, the whole area of electron and ion beam applications continues to make important strides. The strongest development and progress is seen in the areas related to the manufacture of integrated circuits and solid state devices. With the danger of being accused of some bias and myopia I will hazard the guess that at our 1978 meeting in Vancouver, the

I must have them if you wish to be included in the proceedings volume, and now without further ado I will turn the floor over to my good friend of many years, Alex Broers, who will review for us the advanced lithography trends.

R. Bakish

Washington, May 4, 1976.

Acknowledgment

The speakers whose presentations are contained in these proceedings, members of the conference committee, and session chairmen and co-chairmen whose names are given below, have contributed to the success of this meeting. Their contribution is gratefully acknowledged. M. Allais, R. W. Allison, M. von Ardenne, K. Amboss, W. Barwicz, E. Bas, F. Benesowski, M. Boston, G. R. Brewer, A. El Kareh, L. Habraken, M. Hatzakis, C. Hayashi, D. R. Herriott, A. E. Jenkins, H. Koch, H. M. Luther, G. Molenstedt, S. Namba, H. I. Smith, W. C. Nixon, N. A. Olshanski, B. Paton, H. Pfeiffer, K. Pickar, J. G. Siekman, G. Slodzian, W. W. Smeltzer.

Special thanks are given to Sarah A. Kilfoyle of our ECS Headquarters, who is responsible for all arrangements in connection with the printing of this volume, and to my wife, Ellen, who patiently puts up with the never ending projects with which I get involved.

TABLE OF CONTENTS

Section I.

High Power Electron Beams.

INVESTIGATION ON BOLT-CATHODES WITH FLOATING ZONE MELTED POLYCRYSTALLINE AND MONOCRYSTALLINE LaB_6 EMITTERS

P.Hafner and E.B.Bas
Institut für Technische Physik/AFIF
ETH Hönggerberg
8093 Zürich,Switzerland

ABSTRACT

The LaB_6 samples were prepared by floating-zone melting in a special electron beam melting device at nitrogen gas pressure of about 20 Torr. In this way it is possible to grow LaB_6 single crystals of about 4 mm diameter and 25 mm length. From this material emitters of 1 and 2 mm diameter were fabricated for a bolt-cathode system. The thermionic characteristics of (100), (110) and (111) oriented single crystal emitters as well as of polycrystalline emitters were investigated at UHV conditions in an automated Richardson-plot measuring device. The highest thermionic emission is observed with (100)-oriented emitters (1 A/cm^2 at $\sim$ 1600 K and 10 A/cm^2 at $\sim$ 1790 K). The 1 mm diameter polycrystalline cathodes (zone melted material) were life tested in a 5 kV-electron gun at constant emission current, with the initial current densities of 10 A/cm^2 ($\sim$ 1880 K) and 5 A/cm^2 ($\sim$ 1810 K). Mass erosion rates of 2×10^{-8} $g/cm^2 \cdot s$ and 1.5×10^{-8} $g/cm^2 \cdot s$ were found. Based on these results and the investigation of oxygen-LaB_6 interaction at high temperatures with Auger electron spectroscopy we propose an erosion mechanism by oxidation.

1. INTRODUCTION

In 1951 Lafferty introduced for the first time LaB_6 as a thermionic emitter [1]. In the first applications LaB_6 emitters were prepared from LaB_6 powder sintered on to a suitable substrate such as carbonized tantalum or rhenium. In the fifties one of the authors tried to apply LaB_6 in the same way to the so called bolt-cathode developed in our laboratories for high intensity electron guns [2]. These experiments were unsuccessful due to strong erosion and unstable emission, especially under the influence of ion bombardment. In 1967 Broers was more successful in using similar bolt-cathode arrangement but with a hot pressed rod from Cerac-Hot-Pressing Inc. as bolt [3]. The hot pressed material is polycrystalline with LaB_4 as a binder between small LaB_6 crystallites as can be seen in Fig. 1a. The white phase in this micrograph is LaB_4. Such a material is not optimal for cathode technology and has many disadvantages relative to the life behaviour in an electron gun. The Russian scientists tried to improve the material by argon arc melting. Fig. 1b shows the micrographs of such a LaB_6 sample. The sample is polycrystalline and has no LaB_4 phase but many gas inclusions. In our laboratory we tried the improvement by zone melting, which ended eventually in LaB_6 single crystals. Fig. 1c shows a micrograph of such a sample of high homogeneity. Single crystals are very suitable for many physical investigations on LaB_6 and have many advantages as emitter material in electron guns. This paper is a first report on these investigations.

2. THE DESIGN OF BOLT-CATHODES WITH LAB_6 EMITTERS

The LaB_6 material used in our investigations was prepared by floating zone melting in an electron beam melting device. This 20 kW melting device has been described in an earlier publication [4]. The floating zone melting of presintered rods (6x6 mm in cross section) is accomplished in nitrogen gas at about 20 Torr pressure. In this way the evaporation of LaB_6 is greatly reduced. When melting in N_2 we were able to grow LaB_6 single crystals 4 to 5 mm in diameter and about 25 mm in length. The problems of floating zone melting of LaB_6 will be discussed elsewhere. Fig. 2 shows the appearance of a zone melted sample. The surface always has a fine crystalline or amorphous crust, which can be removed by spark erosion. Oriented single crystal emitters have been prepared from these rods using spark erosion machining combined with an x-ray diffractometer control. Fig. 3 shows optical micrographs, scanning electron micrographs and x-ray Laue patterns of three emitters with (100), (110) and (111) orientation of the emission surface (2 mm in diameter).

For the bolt-cathodes with LaB_6 emitters the original system for tungsten bolt-cathodes is used [2]. An essential difference exists in the design of the bolt. In the first case the bolt is either a tungsten rod clamped at the bottom end and polished on the top face as emission surface (up to 1 mm rod diameter) or a cylindrical emitter part (greater than 1 mm in diameter) electron beam welded to a 1 mm diameter tungsten rod as a holder. In the case of LaB_6 emitters the welding is impossible, therefore soldering with molybdenum disilicide is used. Fig. 4 shows drawings of three designs currently used. The type a is used for 1 mm diameter emitters. The LaB_6 emitter and the 1 mm diameter tungsten holder are simply coupled by a tantalum tube with $MoSi_2$-solder between them. Additionally a small hole is drilled through the tantalum tube and LaB_6 emitter and the hole is filled with a piece of tantalum wire. The soldering is accomplished by heating the bolt to the melting point of $MoSi_2$ (2000°C) in the cathode system. Type b is similar to type a with an additional screen made of tantalum. The function of this screen is twofold: Firstly it prevents the escape of the high energy electrons from the bombarding zone into the electron beam even at very low Wehnelt-cylinder potentials and secondly it screens the tungsten coil from the LaB_6 evaporation products. Type c is used for 2 mm diameter LaB_6 emitters. When cathodes with small emission surfaces are needed the 1 mm diameter emitters are shaped to a suitable form.

3. THERMIONIC EMISSION CHARACTERISTICS OF SINGLE CRYSTAL LAB_6 CATHODES

For the investigation of the electron emission characteristics an automated measuring system has been devised. Fig. 5 shows the principle scheme of this system. The diode consists of two bolt-cathode systems placed face to face. The bolt of one system has a single crystal LaB_6 emitter of 2 mm diameter. The bolt of the second system is made of single crystal tungsten 0.9 mm in diameter. It forms the measuring anode of the diode and is surrounded by a disc with a hole of 1.1 mm in diameter as screening anode. The screening anode is also made of zone melted tungsten. The measuring anode can be heated by electron bombardment for degasing. The length of the diode gap is about 0.5 mm. During the measurements the vacuum in the system was held at the 10^{-9} Torr level. The measuring electronics are automated. Square waves pulses of up to 5 kV amplitude and about 300 μs length are applied to the cathode, and the amplitude of the pulses is changed by a ramp generator. The diode voltage and anode current are measured simultaneously near the end of the pulse by sample-hold electronics. The stored signals are then digitalized and forwarded to a teletype. The true temperature of the cathode is measured in a small hole in the LaB_6 emitter by an optical micro pyrometer and typed manually on the teletype. All this data are punched on to the tape which is used for computer evaluation. The computer evaluates the Schottky plots and the Richardson plots by least square approximation. Fig. 6 shows such a Richardson plot for (111) face as plotted by the computer. In Fig. 7 the typical thermionic emission characteristics are drawn for the three faces (100) (110)

(111) as well as for a polycrystalline cathode made from zone melted material. The dashed line represents Lafferty's results for LaB_6 powder sintered on a carbonized tantalum substrate. The diagram shows that the differences in thermionic emission for different faces are much more pronounced in the pre-exponentional Richardson constants than in work functions. Only the (100) face delivers higher emission than the cathodes in Lafferty's work.

4. LIFE TEST OF BOLT-CATHODES WITH ZONE MELTED POLYCRYSTALLINE LAB_6 EMITTERS

It is a difficult task to define working lifetime of a cathode in an electron gun. Not only the vacuum level and the composition of the residual gas atmosphere has to be considered. The electron optics for electron beam generation is an essential factor in setting the point at which a cathode has to be replaced. The material lost by evaporation reduces the diameter of the emission surface and changes the axial position of it. In addition the surface topography and the emission distribution can be changed. All these changes influence the beam focussing. We tried to get some ideas on lifetime simply in measuring the diameter reduction of the emission surface under certain operation conditions. At present only the cathodes with polycrystalline emitters made from zone melted material have been investigated. The bolt-cathodes with 1 mm diameter LaB_6 emitters were operated in a 5 kV test electron gun with ion separation [5] at constant beam current during the whole run. The gun was evacuated by an oil diffusion pump without liquid nitrogen trap to 10^{-6} Torr. Two cathodes were tested. The first cathode was operated with a constant beam current of 79 mA corresponding to an initial emission current density of 10 A/cm^2 (initial temperature ca. 1880 K). The second cathode was operated at half of this current (i.e. 5 A/cm^2, 1810 K). The results of these tests are represented in Fig. 8. In the 10 A/cm^2 run a diameter reduction rate of 0.31 µm / h was measured. This represents a layer erosion rate of 0.155 µm/h $\widehat{=}$ 4.3×10^{-9} cm/s. Assuming the density of LaB_6 to about 4.6 g/cm^3 the mass erosion rate is 2×10^{-8} $g/cm^2 \cdot s$. The corresponding value for the 5 A/cm^2 run is 1.5×10^{-8} $g/cm^2 \cdot s$. From the evaporation measurements of Lafferty [1] one can calculate the evaporation rates for 1880 K and 1810 K to be 6.5×10^{-9} $g/cm^2 \cdot s$ and 1.4×10^{-9} $g/cm^2 \cdot s$ respectively. The rates in our test are higher by factors of 3 and 10. Lafferty measured the evaporation rate indirectly utilising the activation of a tungsten emitter by the lanthanum atoms avaporating from a graphite crucible filled with LaB_6. In our experiments the material loss has been measured directly. Furthermore it must be considered that our experiments were done at constant emission current and not at constant temperature. This means that with the shrinkage of emission area a continuous increase of cathode temperature occurs (automatic emission current control). At the start ($j=10\ A/cm^2$) the temperature was about 1880 K. Fig. 8 shows, however, that the experimental points stay quite well on a straight line. Although the temperature rises during the run, the material loss rate is not influenced appreciably. Therefore, the question arises about the nature of the material loss mechanism. The Auger electron spectrum of the emission surface demonstrates that at high temperatures some kind of surface oxidation occurs even in the 10^{-7} Torr range. The oxidation can proceed in principle at boron octahedron sites as well as at the lanthanum sites. But the removal of oxygen occurs most probably in the form of boron oxide, which is very volotile [6]. The excess of lanthanum or of oxidized lanthanum is then removed by evaporation of lanthanum atoms (after reduction by boron). The Auger electron spectra show that at the emitter temperatures higher than 1700 K the oxygen coverage is smaller than the sensitivity limit of the spectrometer (say about 5% of a monolayer). If at these high temperatures the oxidation still occurs the sticking time of boron oxide must be very short and the material removal rate is mainly controlled by the supply rate of oxygen, i.e. by the partial pressure of oxygen and probably H_2O. As a consequence the material loss from the cathode should not be strongly influenced by the cathode temperature. The second test at $j_{init.} = 5\ A/cm^2$ ($T_{init.} = 1810$ K) shows that this is really the case. Although the temperature is reduced by 70 K the effective material loss rate is only about 25% lower.

For comparison we have tested also a bolt-cathode with a (111) oriented tungsten single crystal emitter at the same conditions. As can be seen from Fig. 8 the diameter reduction rate is only a factor 2 higher for the same initial current density of $10\ A/cm^2$. Of course, the cathode temperature is 1000 K higher and the total heating power higher by more than factor 3.

The erosion of the emission surface during the life test of the LaB_6 emitter with $10\ A/cm^2$ is demonstrated by the micrographs in Fig. 9 which show the appearance of the surface at the start and after 111, 478 and 615 hours. Some selectivity of erosion for different crystallographic faces can be observed. The tests with single crystal emitters will bring more information about this.

We are investigating the oxygen-LaB_6 interaction in our high temperature Auger spectrometer which has been successfully used for the study of oxygen-tungsten interaction up to 3000 K [7][8]. As a preliminary result Fig. 10 shows the changes in Auger electron spectrum due to the high temperature oxygen interaction. The emitter investigated is similar to the emitter used in the life test. The spectrometer used has viton seals and is evacuated by a liquid nitrogen traped oil diffusion pump. The vacuum conditions are similar to those in an electron beam generating system. The diagram in Fig. 10 shows the successive Auger electron spectra during the cooling of the emitter from the operating temperature of about 1700 K to room temperature in about 20 min. Using a multiplex control the spectrum is scanned only around the three peaks boron, carbon and oxygen. At the beginning (0 min.), just before the heating power is turned down, no oxygen or carbon peak is detectable and a large boron peak at about 177 eV is registered. After 2 min. an appreciable oxygen peak appears and the boron peak becomes smaller. When the emitter advances in cooling the oxygen peak grows continuously and the height of the 177 eV boron peak diminishs. At the same time a new peak at about 170 eV grows until after 20 min. a stationary condition is established. The boron peak at 177 eV is now very small and the new peak at 170 eV has an appreciable amplitude. We interprete this new peak at 170 eV as being from oxidized boron. The energy shift due to this oxidation of boron has a value of about 7 eV. If the emitter is now slowly heated a reversed sequence of Auger electron spectra can be observed. Since the boron oxide is very volatile the oxygen is very probably removed as some kind of boron oxide. If so, every new heating will be accomplished by a loss of a certain surface layer and the surface becomes active for electron emission. The influence of oxygen reaction on the lanthanum Auger peak is much smaller. We hope that the experiments in progress with single crystal emitters will bring more light on the interaction of gases with LaB_6 at high temperatures.

5. REMARKS TO THE LIFETIME TESTS

We did not measure the brightness of electron guns with LaB_6 emitters as yet. It has been shown in an earlier publication [5] that the maximum experimental measured with tungsten bolt-cathodes agree very well with the maximum theoretical brightness defined by:

$$\beta_{max} = \lim_{\substack{r \to o \\ \alpha \to o}} \beta_{r,\alpha} = \frac{j_e}{\pi} \frac{eV_a}{kt}$$

For $j_e = 10\ A/cm^2$, $T = 1880\ K$ and $V_a = 20\ kV$ one can calculate:

$$\beta_{max} \approx 0.4 \times 10^6\ A/cm^2 \cdot sr$$

This brightness can be held at least for 500 hours as the micrograph 3 in Fig. 9 shows. If one makes a point cathode there will be some enhancement of the electrical field at the front of the emitter by a factor κ. Due to the Schottky effect the brightness raises:

$$\beta_{max_1} = \beta_{max_o} \exp[4.4\ E_o^{1/2}/T(\kappa^{1/2}-1)]$$

In our experiment E is approximately 20 kV/cm for V_a = 5 kV. If it is assumed κ = 10 for a point cathode at V_a = 20 kV the following expectable brightness results

$$\beta_{max_1} \approx 2\ \beta_{max_o} \approx 0.8 \times 10^6\ A/cm^2 \cdot sr$$

In a recent paper Verhoeven and Gubson [9] reported an estimated lifetime of 40 h for a point cathode with $\beta_{exp} \approx 1 \times 10^6 A/cm^2 \cdot sr$ (20 kV) and a tip radius of about 5 to 30 μ. They also pointed out the importance of the change in surface topography. Broers [10] reported recently lifetimes of 100 - 300 hours for $1 \times 10^{-6} A/cm^2 \cdot sr$ (20 kV), tip radii of 12 µm and 50 hours for $4 \times 10^{-6} A/cm^2 \cdot sr$, tip radii of 1 µm. However one must be cautious in adapting such statements without reference to special operation condition. Certainly one can expect that the point emitters made from oriented single crystals will bring improvements in lifetime for high brightness operation of electron guns.

REFERENCES

[1] J.M.Lafferty, J.Appl.Phys. 22 (1951) 299

[2] E.B.Bas, Z.angew.Phys. 7 (1955) 337

[3] A.N.Broers, J.Phys. E (Sci.Instrum.) 2 (1969) 273
J.Appl.Phys. 38 (1967) 1991
Rev.Sci.Instrum. 40 (1969) 1040

[4] E.B.Bas,P.Hafner,A.A.Latif and G.Wulff, Record of 11th Symposium on Electron Ion and Laser Beam Technology, (Ed.R.F.M.Thornley) IEEE, Inc. 1971, p. 643

[5] E.B.Bas and G.Cremosnik, First Int.Conf. on El.Ion Beam Sci.Technol. (Ed.R.Bakish) John Wiley & Sons Inc. (1965) p. 108

E.B.Bas, Optik 12 (1955) 377

[6] H.Rizzo, "Oxidation of boron at temperatures between 400 and 1300°C in "Boron Synthesis, Structures and Properties" (Edit. Kohn et al.) (1960) p. 175

[7] E.B.Bas and U.Bänninger, Surface Science 41 (1974) 1

[8] U.Bänninger and E.B.Bas, Surface Science 50 (1975) 279

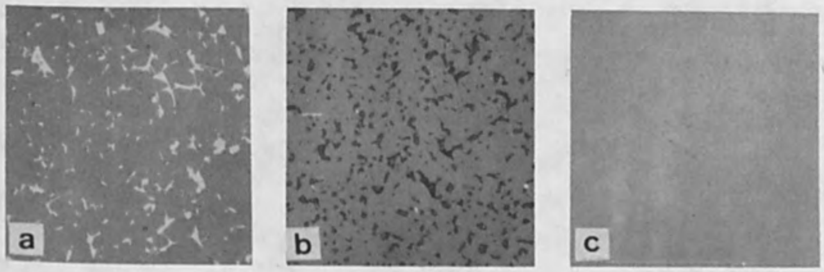


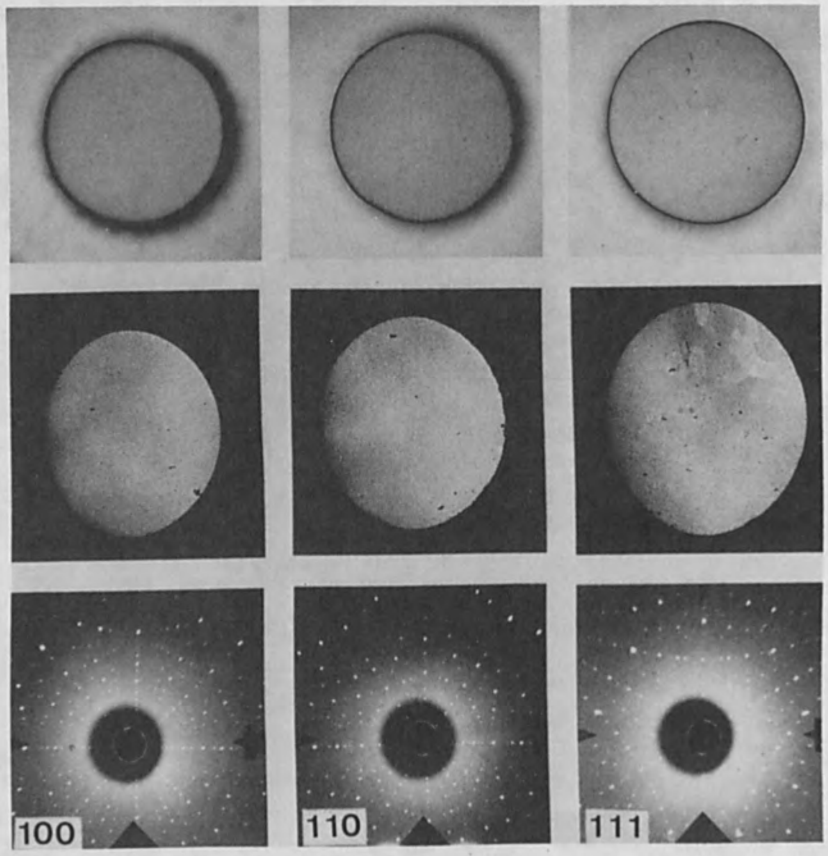
Fig. 1 Optical micrographs of LaB_6-samples from different origins:

a) Hot pressed polycrystalline sample, origin USA (white phase LaB_4)

b) Arc-melted polycrystalline sample, origin USSR (dark gas inclusions)

c) Electron beam zone melted monocrystalline sample origin ETH Zürich

Fig. 2 One-pass zone melted LaB_6 rod



a b c

Fig. 3 Oriented single crystal LaB_6 emitters

a) (100)-face
b) (110)-face
c) (111)-face

First row optical micrographs, second row scanning electron-micrographs, third row Laue-X-ray patterns.

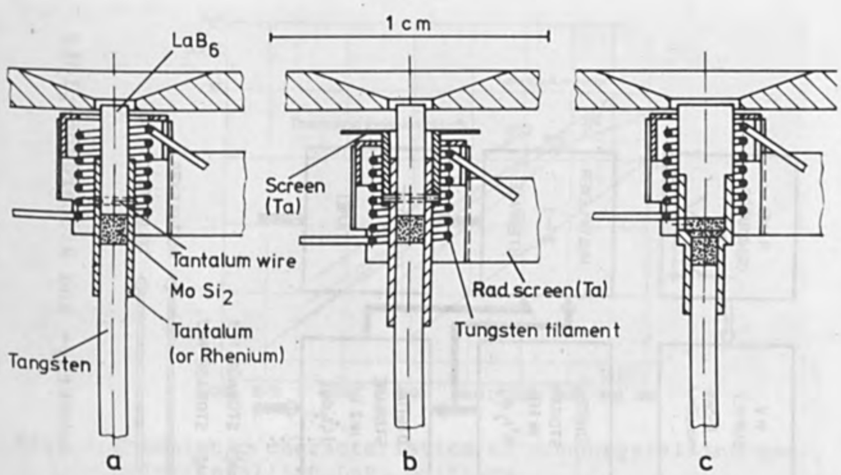


Fig. 4 Different types of LaB_6 emitter design for bolt-cathode systems

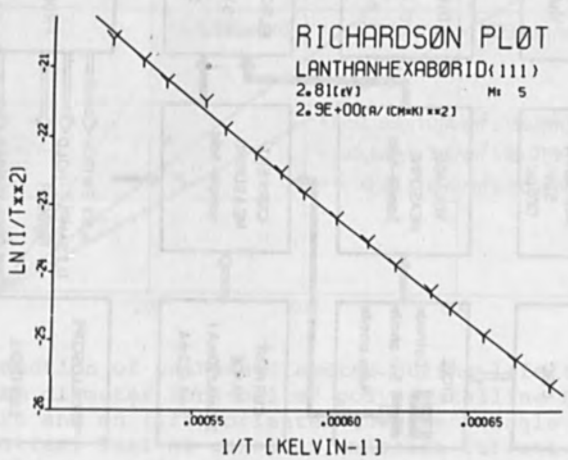


Fig. 6 Typical Richardson-plot for the (111) face of LaB_6

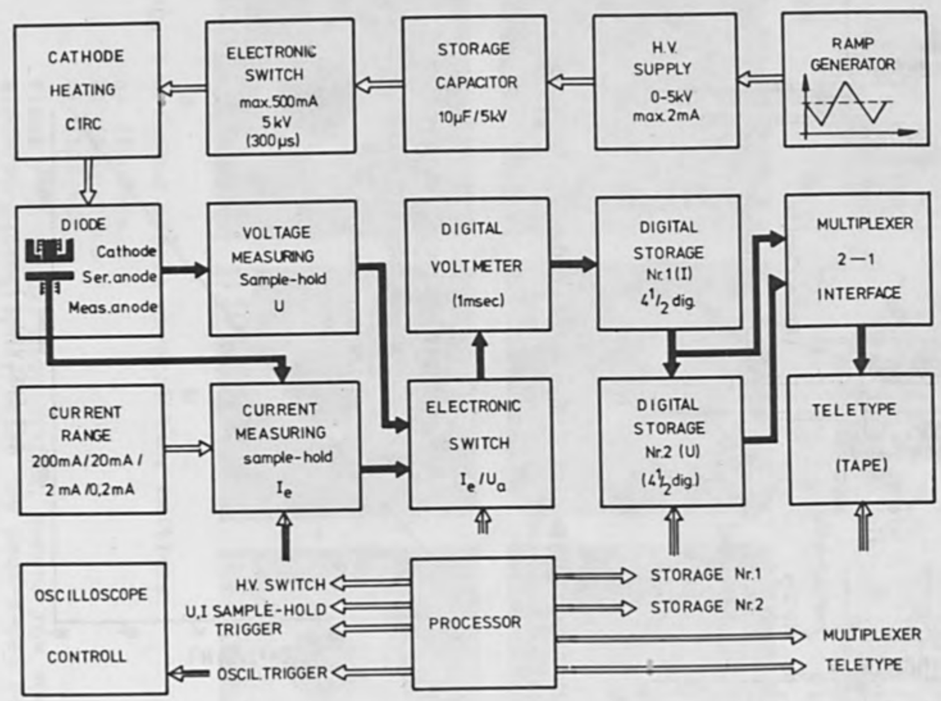


Fig. 5 Principal scheme of the measuring system for Schottky- and Richardson-plots

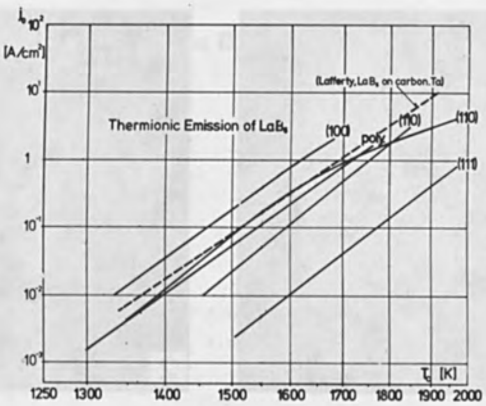


Fig. 7 Thermionic characteristics of monocrystalline and polycrystalline LaB_6 emitters

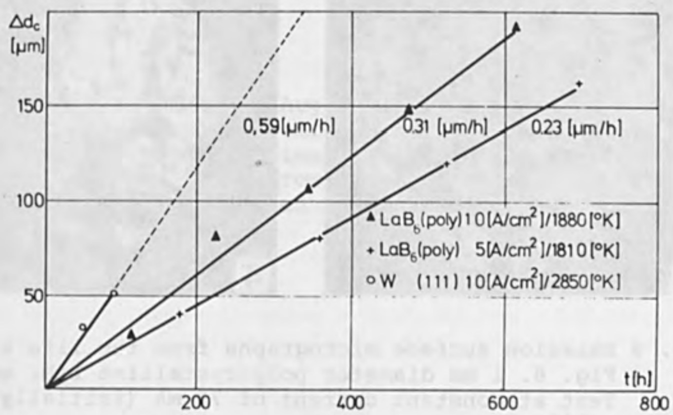


Fig. 8 Reduction of cathode diameter during life test of 1 mm diameter zone-melted polycrystalline LaB_6 emitters and an (111) oriented tungsten single crystal emitter. Test at constant emission current. The initial current densities and temperatures are indicated in the diagram.

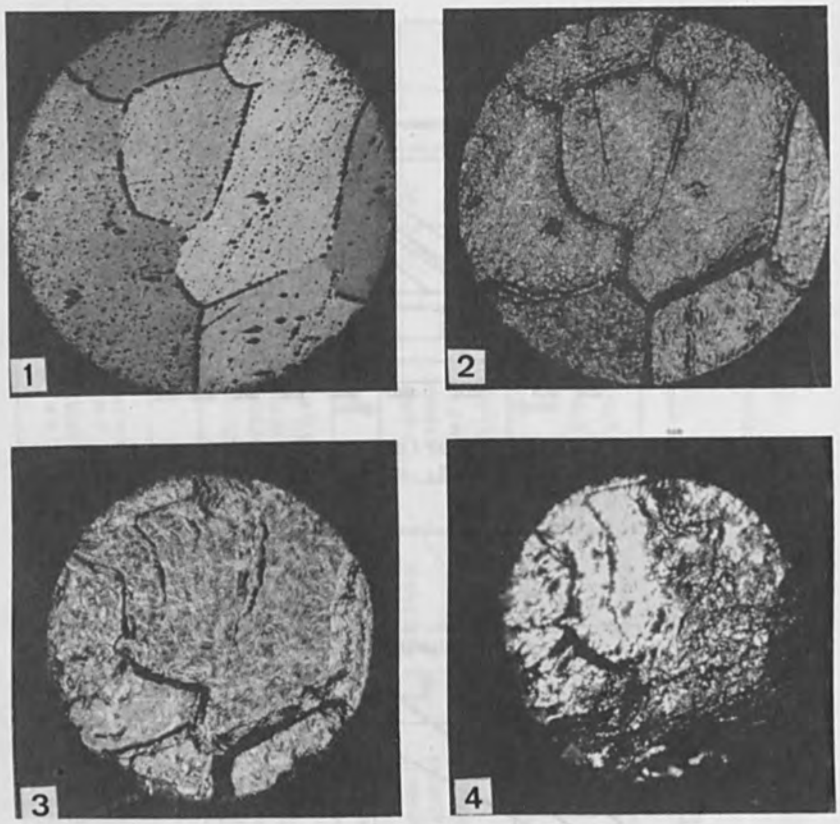


Fig. 9 Emission surface micrographs from the life test of Fig. 8. 1 mm diameter polycrystalline LaB_6 emitter. Test at constant current of 79 mA (initially 10 A/cm^2, $\sim$ 1880 K).

1) at the start of the test
2) after 111 h
3) after 478 h
4) after 615 h

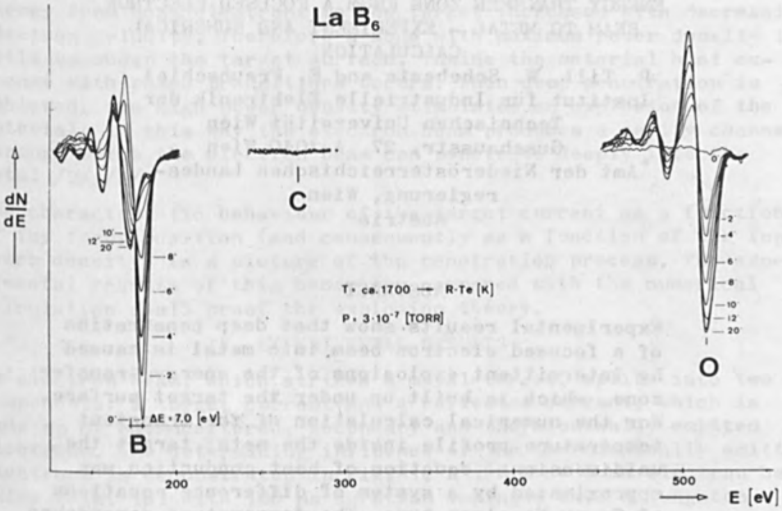


Fig. 10 High temperature Auger electron spectrum of polycrystalline LaB_6 emitter during the cooling from $\sim$ 1700 K to room temperature in a vacuum of $3x10^{-7}$ Torr. The times from the shut down of the heating power are indicated in minutes.

ENERGY DENSITY AND TEMPERATURE RISE IN THE ENERGY TRANSFER ZONE FROM A FOCUSED ELECTRON BEAM TO METAL - EXPERIMENT AND NUMERICAL CALCULATION

P. Till, W. Schebesta and E. Fraunschiel[+]
Institut für Industrielle Elektronik der
Technischen Universität Wien
Gusshausstr. 27, A-1040 Wien
[+]Amt der Niederösterreichischen Landes-
regierung, Wien
Austria

ABSTRACT

Experimental results show that deep penetration of a focused electron beam into metal is caused by intermittent explosions of the energy transfer zone, which is built up under the target surface. For the numerical calculation of the transient temperature profile inside the metal target the unidimensional equation of heat conduction was approximated by a system of difference equations of Crank-Nicolson type. The temperature dependence of heat conduction and the latent heat of fusion and evaporation was taken into consideration.

I. INTRODUCTION

For about 15 years the deep penetration of focused electron beams has been used for material machining. Nevertheless the processes of material removal and the formation of a cavity channel are still matter of investigations, because one aim is to find determining variables, which make possible adaptivè control during electron beam welding.

Single 100 keV electrons have a range of about 20 μm in steel, focused electron beam with 100 kV acceleration voltage and some mA beam current can penetrate several cm of steel. Therefore a cavity channel must be assumed which is free of molten or solid material during the welding process. The electron beam can remove 1 g steel with about 1000 J, whereas the evaporation of 1 g steel requires 8400 J; therefore evaporation alone cannot be the reason for the formation of the cavity channel as has been previously assumed /1/. There are evidences that the material is removed by intermittent explosions.

A 100 kV electron beam, which impinges on a metal, looses its energy by a large amount of collisions with electrons and with lattice atoms, if processes with little probability are neglected. The penetration depth of single electrons is proportional to the square of the acceleration voltage and inversely proportional to the density. The cross section of the collisions which transfer energy from beam electrons to the target increases with decreasing electron velocity. Therefore a zone with maximum power density is built up under the target surface. Inside the material heat exchange with phase transitions occurs. When deep penetration is achieved, the high power density stipulates an explosion of the material; in this way the electron beam produces a cavity channel, through which the electron beam can penetrate deeply into the metal /2/,/3/.

The characteristic behaviour of the target current as a function of the focus position (and consequently as a function of the input-power density) is a picture of the penetration process. The experimental results of this behaviour compared with the numerical calculation shall proof the explosion theory.

II. EXPERIMENTAL RESULTS

An electron beam, which strikes a metal target, splits into two components: a target current and a reflected current, which is made up of backscattered, secondary and thermionically emitted electrons. The determining influence of the thermionically emitted electrons is demonstrated in Fig. 1. A 100 kV - 5 mA electron beam pulse (Fig. 1a) with 100 μs duration impinges onto a tungsten target, while the current flowing through the target is displayed on a storage oscilloscope (Fig. 1b). After about 60 μs the target current decreases exponentially, leading to a negative target current according to the rise of temperature and the exponential increase of the thermionic emission.

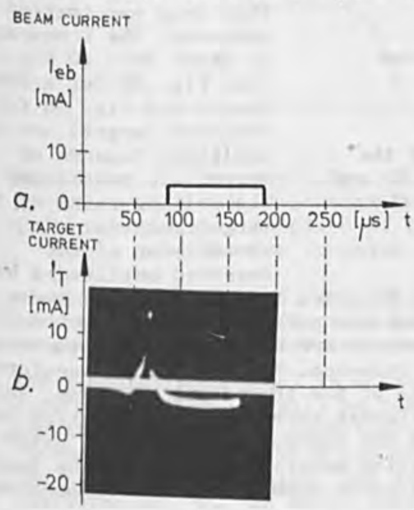


Fig. 1. Beam current (a) and target current (b) for a 100 μs electron beam pulse impinging a tungsten target (U = 100 kV, I = 5 mA).

The focus position influences the power density in the zone of impact. If the power density is not able to melt the surface, the splitting is only determined by the backscattered and the secondary electrons. If the target surface is molten, the thermionically emitted electron current reduces the target current (see Fig. 2a, which shows the dc-value of the target current, when a 100 kV - 1 mA electron beam continuously hits a steel target which moves with a velocity of 4 mm/s; Fig. 3a

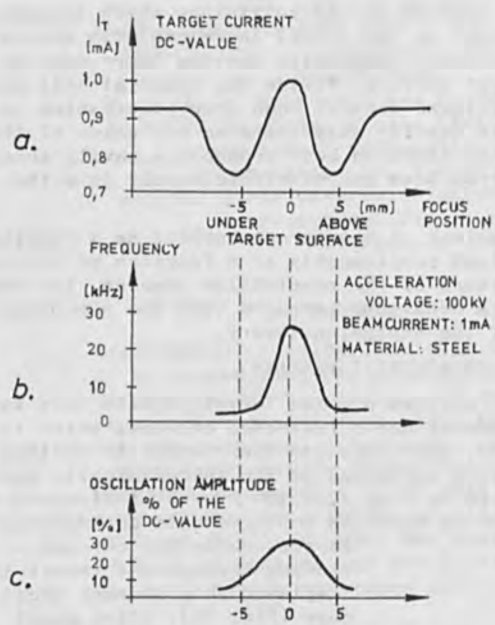


Fig. 2. Direct current value of the target current (a), frequency (b) and amplitude (c) as a function of the focus position for steel.

shows the dc-value, when a 100 kV - 2,5 mA electron beam hits a tungsten target). Between the two minima of the W-shaped curve the target current increases again. In the region of such focus position the electron beam is able to form a cavity channel; not all the electrons can leave the channel and therefore part of the reflected current is added to the target current. Only in this region deep penetration is observed /3/,/4/.

Both the target current or the reflected current show significant oscillations on an oscilloscope when deep penetration is achieved. The frequency is about 10 - 20 kHz (see Fig. 2b for a steel target and Fig. 3b for a tungsten target) and the amplitude depends on thermionic emissivity at the melting point of the target material, as a comparision of the measured amplitudes between steel target (Fig. 2c) and tungsten target (Fig. 3c) shows. The experimental results with various materials as a target confirm the assumption that the oscillations are due to periodic increase and decrease of the thermionic emission. The rise of temperature leads to an exponential increase of the thermionic emission, and therefore to a decrease of the target current. This reduction is interrupted by the explosion of the energy transfer zone, which removes the thermionically emitting material. Therefore the target current increases again. The electron beam raises the temperature and the cycle is repeated.

III. NUMERICAL CALCULATION

In order to pursue the processes in the microzone of impact, a computer program for the computation of the thermal behaviour was carried out.

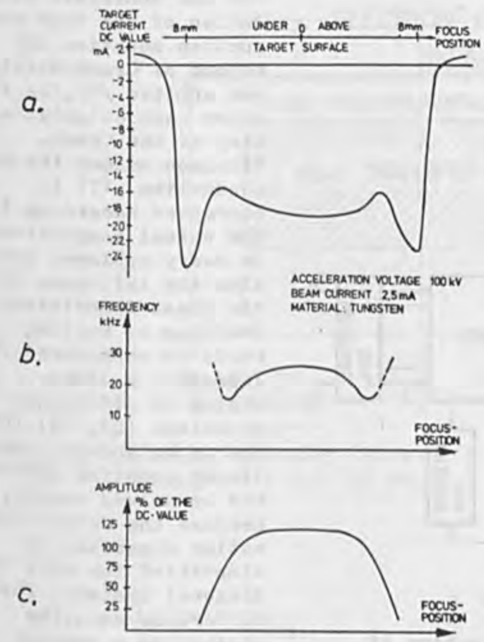


Fig. 3. Direct current value of the target current (a), frequency (b) and amplitude (c) as a function of the focus position for tungsten.

The physical function scheme of the thermal processes is shown in Fig. 4. In addition to the essential processes of energy absorption and heat exchange, the influence of the temperature dependent material coefficients, heat transfer at the boundaries, heat radiation and the phase transitions (melting and boiling) have to be taken into account.

Fig. 5 shows the idealized geometry. Under the target surface a zone with maximum power density will be built up. According to the physical energy transfer from highly accelerated electrons in metal a simplified cylindrical heat transfer zone or energy transfer zone is assumed; the diameter of it is equal to the beam diameter, the height of it is one half of the range of electrons; between surface and the energy transfer zone exists a layer transparent to electrons with one sixth of the range of electrons (see Fig. 5). The energy will be transfered completely into the energy transfer zone (ideal focus position). Steel and tungsten with 1 mm thickness are considered as target material.

The computational model is shown in Fig. 6. The Fourier heat conduction equation (1) will be considered as a thermal power balance equation. As boundary condition at the target surface there is thermal no load condition and at the bottom of the target there is thermal shunt transition. The temperature profile is calcualted unidimensionally. The metal target is subdivided by n-2 sublayers (index i) either equidistant (energy transfer zone and quasi transparent zone) or progressively rising (from the heat or energy transfer zone into the target). No losses due to heat radiation are assumed. However the heat energy radiated from the target surface was found to be small, compared with the heat input flow in the energy transfer zone.

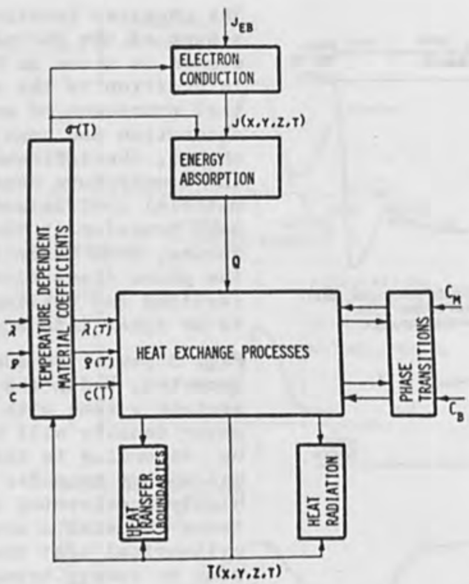


Fig. 4. Physical function scheme of the thermal processes

For the numerical calculation of the heat conduction equation the method of Crank-Nicolson was applied /5/,/6/,/7/. After each calculation step of the Crank-Nicolson method the heat conduction (T) is corrected according to the actual temperature in every sublayer and also the influence of the phase transitions (melting or boiling heat) is considered /7/. Therefore a linear system of difference equations (2), (3)-(6) has to be solved. This linear equation system can be solved easily, because the Gauss elimination algorithm is simplified for this tri-diagonal systems. This difference equation system has n unknown variables, but only n-2 equations. The two missing equations are delivered by the boundary conditions. This calculation has been made by the aid of the CDC CYBER 74 of the "Interfakultären Rechenzentrum der Technischen Universität Wien".

$$\iiint_V \left(c\cdot\rho\,\frac{\partial T}{\partial t} - \frac{\partial^2 T}{\partial z^2} - Q\right)dV = 0 \qquad (1)$$

$$e_i T_{i-1}^{j+1} + f_i T_i^{j+1} + g_i T_{i-1}^{j+1} = 1_i \qquad (2)$$

$$e_i = \frac{-k}{C_i\,(R_i + R_{i-1})} \qquad (3)$$

$$g_i = \frac{-k}{C_i\,(R_i + R_{i+1})} \qquad (4)$$

$$f_i = -\,e_i - g_i + 1 \qquad (5)$$

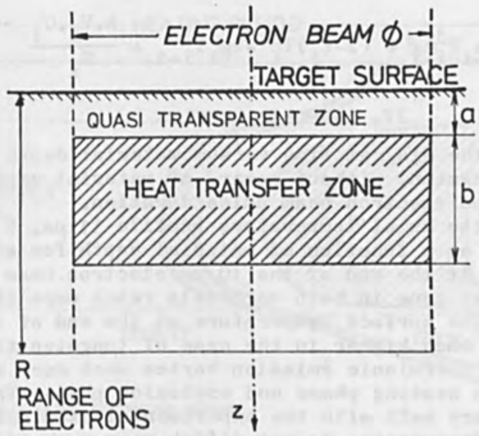


Fig. 5. Idealized geometry (steel : a = 4 µm, b = 12 µm, Ø = 35 µm, tungsten : a = 1,3 µm, b = 4,5 µm, Ø = 40 µm).

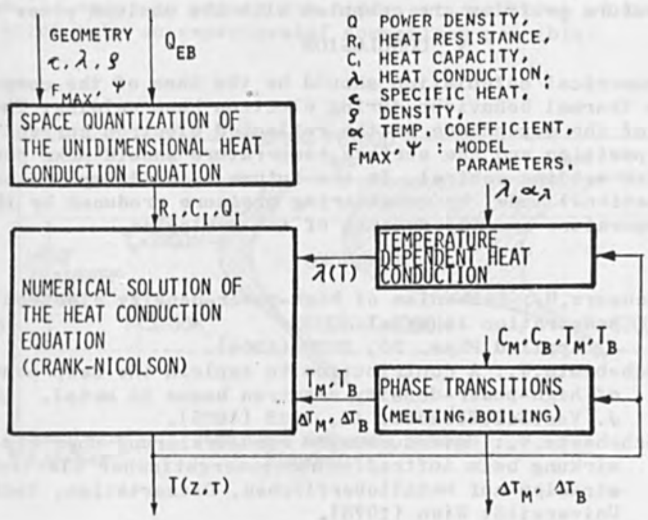


Fig. 6. Computational model

$$l_i = -e_i T^j_{i-1} + (2-f_i)T^j_i - g_i T^j_{i+1} + \frac{k.V_i.Q_i}{C_i} \tag{6}$$

IV. COMPUTED RESULTS

Fig. 7 shows the diagram time versus material depth for steel, and Fig. 8 for tungsten, without regard of material expansion within the time of the electron beam pulse duration.
Fig. 9 shows the axial temperature profile (1 µs, 5 µs and 10 µs heating time) as a function of material depth for steel and Fig.10 for tungsten. At the end of the 10 µs electron beam pulse the energy transfer zone in both materials reach more than boiling temperature. The surface temperature at the end of the electron beam pulse is much higher in the case of tungsten than in the case of steel, the thermionic emission varies much more with a tungsten target between heating phase and explosion phase. This result corresponds very well with the experimental examination.
The measured frequencies do not differ very much if different materials are applied. Therefore the comparison of the experiments with the computation shows that the time between energy absorption and material reaction does not depend much on the target material. It must be stated that in reality the penetration process will be slower than the computed results because the power density in the energy transfer zone depends very much on the focus position; the temperature profiles are computed with the maximum power density.

V. CONCLUSION

This numerical calculation should be the base of the computation of the thermal behaviour during electron beam welding. The knowledge of the dependence of the reflected electron current from the focus position and the surface temperature should make possible an adaptive welding control. In the future we will improve our computational model by considering pressure produced by the rise of temperature and the dynamic of the explosion.

REFERENCES

/1/ Schwarz,H.: Mechanism of high-power-density electron beam penetration in metal. J. Applied Phys. 35, 2020 (1964).

/2/ Schebesta,W.: A contribution to explain the deep penetration of high-power-density electron beams in metal. J. Vac.Sci.Technol. 12, 1218 (1975).

/3/ Schebesta,W.: Untersuchungen zur Aufklärung über die Tiefenwirkung beim Auftreffen höherenergetischer Elektronenstrahlen auf Metalloberflächen. Dissertation, Technische Universität Wien (1975).

/4/ Boncoeur,M.: Use of secondary, backscattered and thermal electrons in the electron beam welding process. 10.Symp. on Electron,Ion and Laser Beam Technology, Gaithersburg (1969).

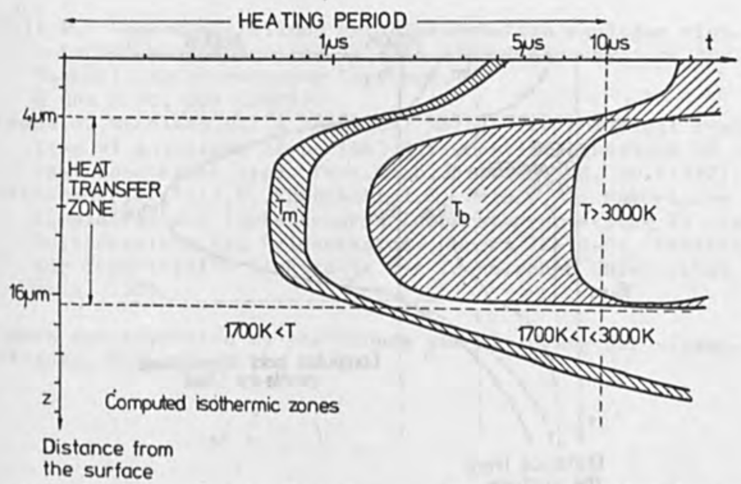


Fig. 7. Computed axial isothermic zones (melting, boiling) as a function of time and depth for steel (U = 100 kV, I = 1 mA, $Q = 0,832.10^{16}\ W/m^3$).

It was necessary to apply different beam currents for steel and tungsten, in order to regard the different densities of the materials, which make an experimental comparison possible.

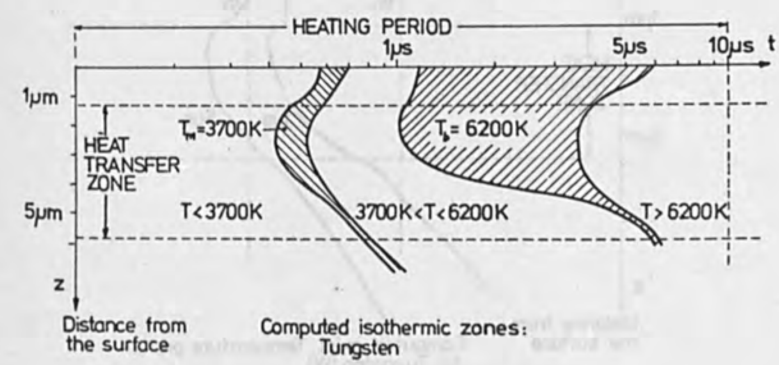


Fig. 8. Computed axial isothermic zones (melting, boiling) as a function of time and depth for tungsten (U = 100 kV, I = 2,5 mA, $Q = 0,442.10^{17}\ W/m^3$)

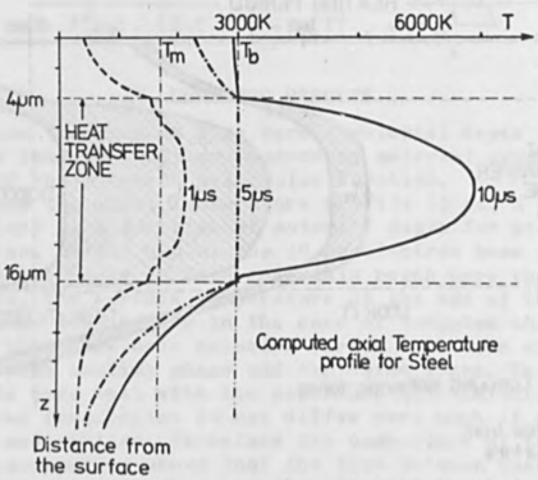


Fig. 9. Computed axial temperature profile as a function of temperature and depth for steel ($U = 100$ kV, $I = 1$ mA, $Q = 0{,}832.10^{16}$ W/m^3).

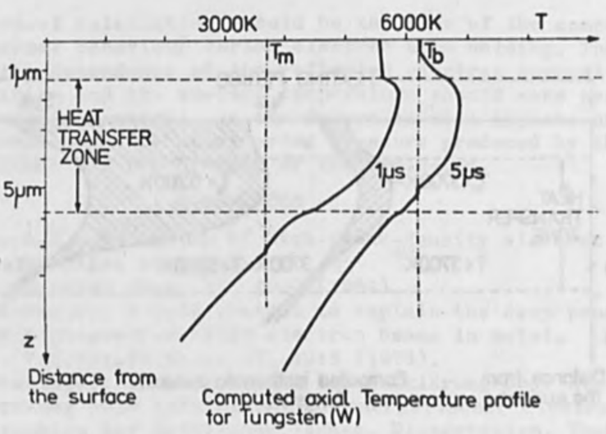


Fig.10. Computed axial temperature profile as a function of temperature and depth for tungsten ($U = 100$ kV, $I = 2{,}5$ mA, $Q = 0{,}442.10^{17}$ W/m^3).

/5/ Till,P.: Gegenüberstellung der Eigenschaften analoger elektrischer, digitaler und hybrider Simulationssysteme zur Nachbildung thermischer Vorgänge. E und M 92, 294 (1975).

/6/ Crank,J.,Nicolson,P.: A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type. Proc.Camb.Phil.Soc. 43, 50 (1947).

/7/ Mitterauer,J.,Till,P.,Fraunschiel,E.,Haider,M.: Numerische Simulation der instationären Temperaturverteilung in einer Emitterspitze bei Feldemission. Report IE-76-01, Institut für Industrielle Elektronik der Technischen Universität Wien, 1976.

This work was supported by the "Fonds zur Förderung der wissenschaftlichen Forschung".

A Study on Dynamic Behaviours of Electron Beam Welding (Report 1) -The Observation by a Fluoroscopic Method-

Yoshiaki ARATA,* Eiichi ABE** and Masanao FUJISAWA**

Abstract

An experimental technique utilizing a cine-fluoroscopy and a metal tracer was developed to continually monitor the electron beam welding process. This technique made it possible to observe the dynamic nature of the beam hole more directly and more clearly than some others.

Using aluminum alloys as the parent metal, some phenomena related to the beam-metal interactions such as the big wall-fluctuation at the back side of the beam hole ("wall-cave" and "wall-knob"), the formation of root porosity due to the arched curvature of the beam hole, the behaviours of spiking concerned with the hole shape, and in part the rapid movement of the molten metal were apparently revealed.

1. Introduction

It has been suggested that a vapor filled capillary is formed in the path of an electron beam[1,2] and is rapidly filled with molten metal as electron beam welding progresses. The term "beam hole" is here adopted to indicate such a capillary which has been called a (weld) cavity or (penetration) channel and so on.

Little is known, however, about the actual mechanism which determines the size and the shape of the beam hole and their time development during welding, because of many difficulties lying on the technique to observe directly those dynamic behaviours, while some works were attempted utilizing elavorate equipment. Tong and Giedt[1] first used radiography of pulsed X-ray for such investigation. However, since they could only take single pictures of which the definition and resolution were still unsatisfactory to reproduce, it was impossible to acquire continually an overall impression of the very rapid processes. Funk, McMaster et al.[2] employed a pinhole streak camera and recorded the X-ray intensity distribution which was emitted from the beam hole. While they could roughly reveal the time related position and intensity of electron beam impingement, dynamic behaviour of beam hole and molten pool could not be observed on their high speed pictures. For this reason, a trial was initiated in which a fluoroscopic technique was used for continuous observation of the phenomena.

2. Experimental Method

2.1 System for observation and recording

The experimental arrangement of the equipment is shown in Fig.1. The X-rays emitted from the X-ray source pass through a work piece during actual welding, and impinge onto an X-ray image intensifier in which they are collimated and amplified being converted into light-rays and electrons. Finally, the visible images are obtained on the second fluoresent screen and they are filmed with a movie camera.

---

* Professor, Welding Research Institute of Osaka University
** The Tokyo Metropolitan Research Institute of Industrial Technology

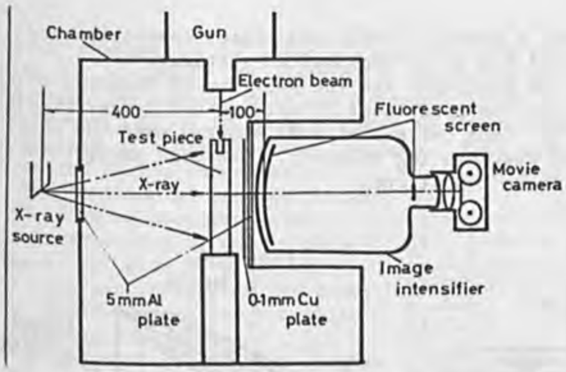


Fig.1. Schematic figure of experimental technique

In electron beam welding, the welder itself works as an X-ray generator and X-ray is emitted from the wall of the beam hole produced by welding electron beam. Such X-Ray is called "beam hole X-Ray."
The resolution and contrast of the image on the fluorescent screen are reduced by the incidence of the beam hole X-ray and the other scattering X-ray. Thus a mask of copper plate 0.2 mm thick was provided upon the input window of the image intensifier to absorb these disturbing X-ray. The decrease of discernability caused by the copper plate was only 0.5% in the penetrameter sensitivity (that is the indication of radiographic resolution). In order to obtain a good image quality, it is important to set up suitable distances between the X-ray sources, the welding zone to be detected and the image intensifier. The image intensifier should be set as close as possible to the welding zone for preventing the decrease of picture definition due to the enlargement. For the same reason, the smaller the focal spot of the X-Ray tube the better definition and resolution on the picture. So, in this experiment, a d.c. X-ray equipment Müller MG 150 was employed as the X-ray source with its focal spot of 0.7mm x 0.7mm and Philips image intensifier tube MB 13 (diameter 6 inches) was located at the bored vacuum chamber as illustrated. Being aligned with the center axis, a 16mm camera was placed behind the viewing screen of the tube. The films used were Fuji type RKS-FG for fluoroscopic use. The input and output side windows of the vacuum chamber were both constructed of 5mm thick aluminum plate instead of steel to decrease X-ray absorption.

For a given X-ray spectrum and intensity, the resolution and contrast of the image also depend to a great extent on the thickness ratio of X-ray penetration between the beam hole to be detected and the material to be welded. Hence a given void size can be most easily detected in thin specimens. But in thin specimens, the depth of weld penetration is severely limited since the specimens are apt to be over heated with a low power of electron beam, and show some behaviours different from the ordinary welding. Compromise specimen thickness of 12-20mm was chosen in this experiment.

2.2 Experimental Procedure

Experiments were composed of two series of welding runs. In the first series, aluminum alloy 5083 (4.6Mg, 0.01Zn, 0.02Ti, 0.15Cr, 0.04Cu, 0.21Fe, 0.64Mn, 0.14Si, Bal Al) was used as the parent metal and silver brazing alloy (75Ag-25Cu) was inserted as the tracer.

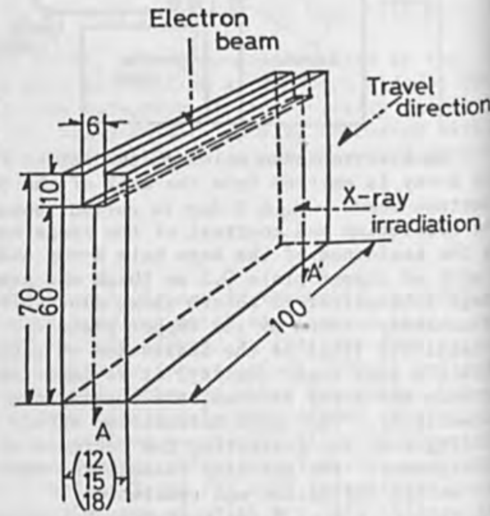
Without any tracers, the fluoroscopic images would yield no information about the extent and the flow motion of the molten metal since the density difference between melt and solid is very slight. To discern the movement of the melt at least in part, a material which had higher absorption coefficient than the parent metal had to be employed as a tracer.

The size and shape of the specimens and the position of inserted tracers were illustrated in Fig.2, and the conditions of welding and X-ray irradiation were as shown in Table 1. The lens current of 0.82 amperes made the focal point just on the surface of specimen (a_b value=1). The weld penetrations reached into about 50mm in these conditions.

Table 1 Experimental condition (series 1)

Welding Condition			Width of plate	X-ray irradiation (Focal spot 0.7 x 0.7 mm)		
Fixed factors	Lens current (A)	Process		Tube volage	–	Tube current
30kV	0.82	Bead-on -plate	12mm	60kV	–	8mA
250mA	0.84					
75mm (W.D.)	0.82	Butt weling	15mm	63kV	–	8mA
300mm/min	0.84					
	0.82	Bead-on -plate	18mm	65kV	–	8mA
2×10^{-5} Torr	0.84					

* Note W.D.= work distance



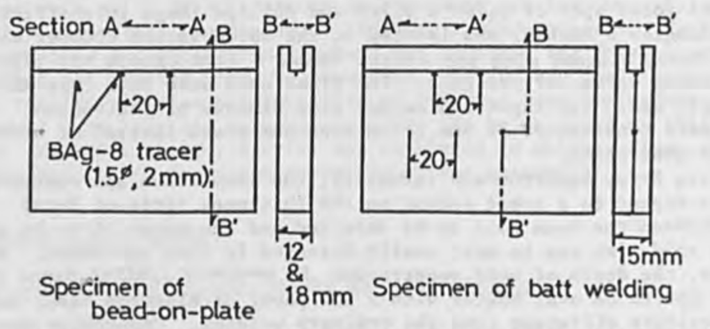


Fig.2. Test specimen and position of inserted tracers (Series. 1)

At the bottom part of the beam hole, the size of the void becomes smaller than at the upper part. The resolution of the film decreases to make it difficult to discern the shape of the hole.

In the second series, in order to improve the discernability, copper foils were inserted at the lower part of the specimen. It was considered to be important to observe the bottom part more clearly since the spiking, which is one of the inherent defects in partial penetration electron beam welding, appeared at the part. By changing the inserted metal from pure copper to some other alloys and also by changing its location, the effect of highly volatile element on the shape of beam hole and also on the formation of spiking was examined. In addition to employing the tracer foils, an attempt to increase the framing rate was successfully made in this series.

The movie films which were taken at 100 and 200 frames/sec had a fairly good quality of resolution and contrast. This contributed to the more precise observation of the phenomena.

The location of inserted metal and the size of specimens were illustrated in Fig.3, and in this series a grid made of stainless steel wire was located at the front of the first fluorescent screen to indicate the demensions. One spacing of this grid was about 8 millimeters on the obtained films. The welding conditions were as shown in Table 2.

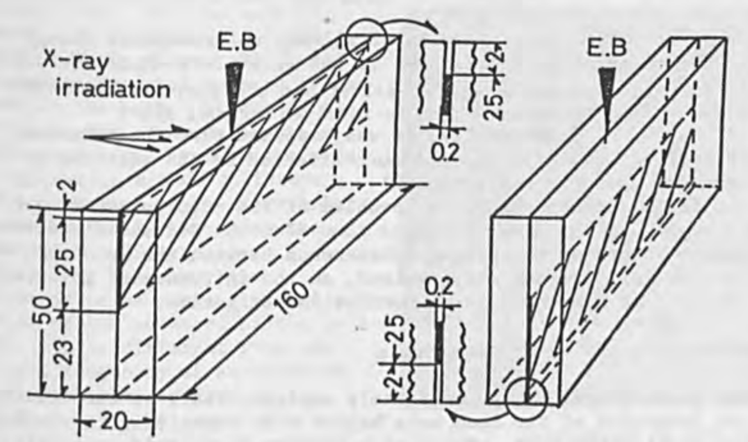
Experiments were conducted with a low voltage type welding unit which had the maximum capacity of 7.5 KW at 30 KV (and 15 KW at 60 KV).

The chamber was evacuated, welding parameters and X-ray irradiation parameters were set, and sequentially, the camera motor was started, and a little later, the welding operation was initiated. The film was moved at a framing rate of about 50 per second.

Table 2 Welding condition (Series 2)

Film No.	Material (width:20mm)	Welding speed cm/min	Work distance mm	Lens current Amp.	Inserted metal (0.2^t foil)	Framing rate frames/sec
6	Aluminum A1100	50	40	0.965	Copper	50
8					Brass	
9					Nickle silver	
10					Red brass	100
11					Copper nickle	50
12					Phosphor bronze	
4	5083 Alloy	30	65	0.84	Copper	
5		50	40	0.965	Brass	
13					Nickle silver	
14					Red brass	100
15					Copper nickle	50
16					Phosphor bronze	
7	Aluminum A1100				Brass*	
17						200
18	5083 Alloy					

note: Location of inserted metal tracer foils are bottom side of specimeus and mark (*) is at the upper side (butt welded)



(a) Inserted at the upper part (b) Inserted at the lower part

Fig.3 Dimension of specimens and position of foils (Series 2)

3. Experimental Results and Discussion

3.1 Quality of obtained films

The data was obtained in the form of exposed films which were several meters in length.

The contrast of the original film was rather low but it could be increased to a certain extent by means of reprinting the original on a reversal film. As for the resolution, wires in the penetrameter which had a diameter of more than 0.8 mm were clearly visible on the viewing screen of image intensifier, but on the frames of the film, the minimum diameter of discernable beam hole was to be believed about 1 mm. A resonable first approach to get a better resolution would be to select the material of the chamber windows such as beryllium which has a less X-ray absorption ability than aluminum employed in practice. It is not preferable to use thinner specimens for welding since the natural behaviour of the process may be disturbed in return for decreasing the total thickness of the materials through which the X-Ray must penetrate

In the second series of welding runs, the resolution and contrast of the film increased at the part where copper or brass foils was inserted.

3.2 Influence of inserted metals on the natural behaviour of beam hole and molten metal

As for the selection of added tracers and inserted foils, each of these materials should have a larger atomic number than aluminum, the parent metal, from the view point of X-ray absorption. On the other hand, such a material has a higher specific gravity, and there is the possibility that they settle at bottom of the molten pool

The tracers which have a better affinity for the parent metal are expected to flow together with the molten metal giving less effect upon its natural movement. Therefore, a silver brazing alloy was considered to be preferable because its melting point is not so far from that of aluminium, and silver has a good solubility for aluminum. It turns out that silver is readily availab wire form.

In spite of the high vapor pressure of silver, no remarkable change was recognized in the shape of beam hole even at the very place where the tracer was inserted. In turn of vapor pressures silver has a higher pressure than aluminum but lower than magnesium which is involved in 5083 alloy

Consequently, it was supposed that the behaviour of the molten pool or beam hole was not much affected by the addition of the small amount of such tracers.(Photo.2)

As copper foils are concerned, the problem of the vapor pressure can be neglected as copper foil has lower pressure than aluminum But the difference of the penetration depth or the spiking phenomenon between the weld with and without the copper foil was not yet examined, so the influence of inserted copper was remained as a problem for a further investigation.

3.3 Behaviour of the beam hole

Observing these films, it it immediately apparent that, as the welding is initiated, the formation of the beam hole begins with vaporizing the material along the path of electron beam. The beam hole increases its depth until an equilibrium is established, and then a nearly constant depth is continually produced for the duration of welding. While the beam hole varies its size and shape with the time, and shows a oscillatory nature.

In the first series, the cyclic wall-fluctuation of beam hole happened very rapidly and and developed upward. In this case the strong wall-fluctuation gave rise to the big cave and knob at the wall. The terms "wall-cave" and "wall-knob" are here adopted to indicate such caved part in and swelled part on the wall respectively. Such wall caves which appear as a consequence of local of gas bubbles due to the high pressure of the evaporating metal, were observed in the beam hole. These wall-caves grew up gradually moving upwards at the rate of 20 30 cm/sec. as shown in Photo.1, where some selected frames of the film were reproduced.

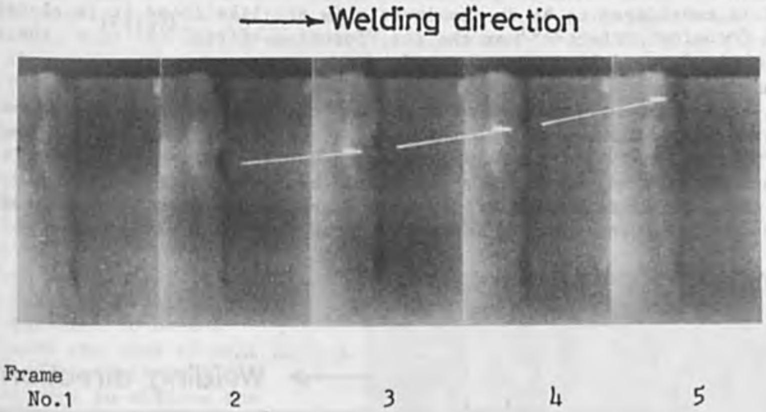


Photo.1 Upward movement of the wall-caves and wall-knobs in the beam hole

At the tip of the wall-knob, the beam hole was narrowed by the dinamic force of molten fluid on the back side. But the complete closure of the beam hole could not be recognized upon all these films excepting what occured at the bottom of the hole. It was difficult to estimate the wall-oscillation frequency, because the mode of fluctuation seemed different - at each part of the beam hole and the sharpened bottom of the hole was less discernable. However, the frequency was considered to be about 5 cycles/sec in view of the movement of the wall-caves. It is different from the "natural frequency of equilibrium state" mentioned by Funk et al. This frequency was comparable to that of the spiking which was revealed by means of etching the longitudinal section of specimen after welding as illustrated in Photo 2.

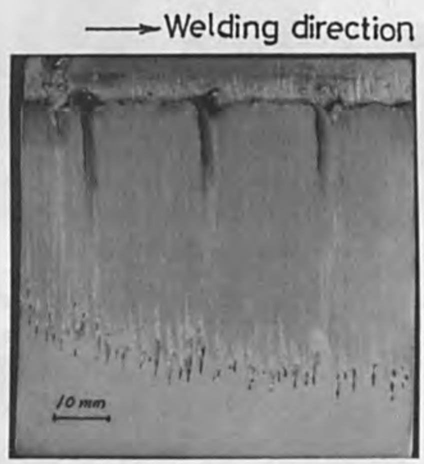


Photo.2 Macrograph of the spiking (Longitudinal section after welding)

So far as this experiment was concerned, the basic fluctuation mode of the beam hole was the same for both butt-and bead-on-plate welding.

3.4 Arched curvature of the beam hole and formation of root porosity

In most cases an arched curvature deflected in a direction opposite to that of moving was observed at the bottom side of the beam hole as shown in Photo.3. The reason for this event is yet unknown.
While it is considered to be due to inertia in all likelihood it is closely related to "wall focusing" effect[3)4)] or the gas "focusing effect"[5)6)12)] of the electron beam which leads to delay energy transfer

This event is also supposed to be concerned with the formation of root porosity. When the inflection point of the curvature is put upon a narrow part at the tip of the knob, a complete closure of the hole easily happens at the point. Near the bottom of beam, the cooling rate of molten metal is increased due to less input beam energy and becomes sufficiently rapid, and the molten metal may freeze before the trapped gas bubbles can be convected to the out side, thus giving rise to porosity in the root of weld

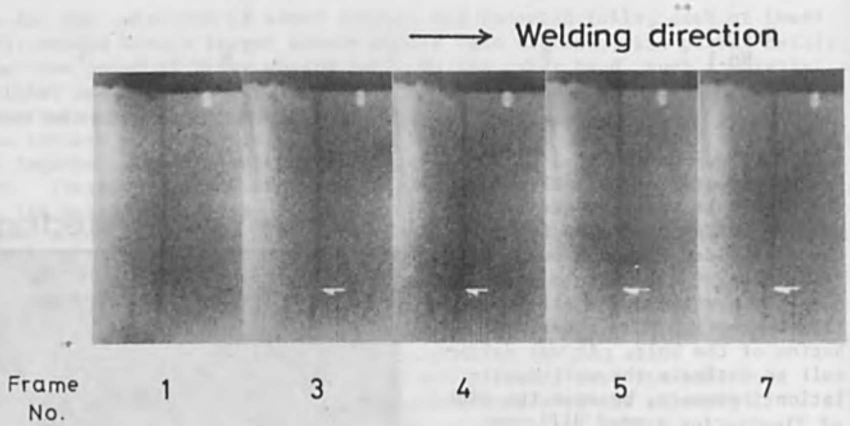


Photo. 3 Arched curvature of the beam hole and the formation of root porosity

3.5 Movement of the molten metal

The movement of molten fluid suggested by the inserted tracers of silver brazing alloy was as follows. The flow of molten metal was at a rapid rate and more than 9m/min in the vertical direction.

It was, of course recognized that the tracer was also spread to the rear wall through the horizontal direction which was perpendicular for the X-ray passage. The details of this movement could not be clarified because the diameter of the beam hole was very small

At the middle part of the beam hole, two ways of movement were observed One was that the molten metal tracer contained in the fluid at first rose up along the wall of the beam hole. at about 0.1 second later, it begun to fall down towards the bottom, and finally, it was distributed through out the beam hole in a symmetrical fashion. In this case, the beam hole had a sharp edge at the bottom as shown in Photo.4 series A. The other was that the tracer spread up and downwards at the beginning but the upward motion was less than the former and ceased after about 0.06 sec. while the downward motion continued 0.12 sec.. The final distribution was inclined towards the obtuse bottom of the hole as shown in Photo.4 series B..

Near the surface, the fluid moved most likely in a whirling fashion, and this event was expected to have a relation with the wave of melt reflected at the weld pool. But the attempt to observe the wave simultaneously with the hole shape was unsuccessful because the heat deposition at the surface could not make a normal bead as a result of the limitation by the employed shape of specimen head. (Photo. 5)

Although it should be noted that the difference of the specific gravity between the tracer and the base metal might disturb the natural behaviour, the turbulance in the beam hole would be very efficient for mixing the melt and resulted in a homogeneous fusion zone.

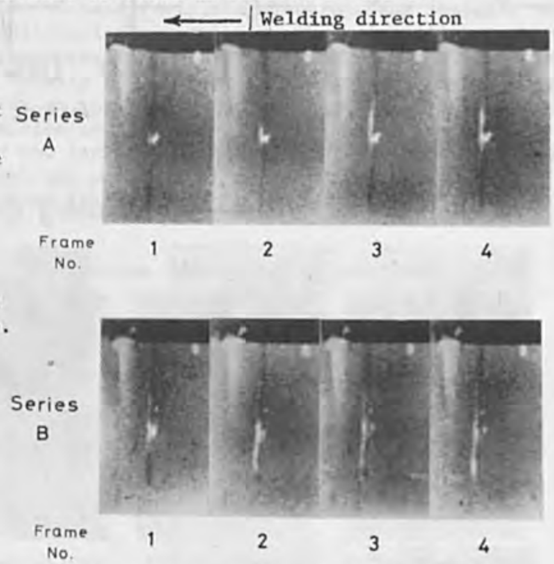


Photo.4 Flow of the molten tracer (1)
- middle part -

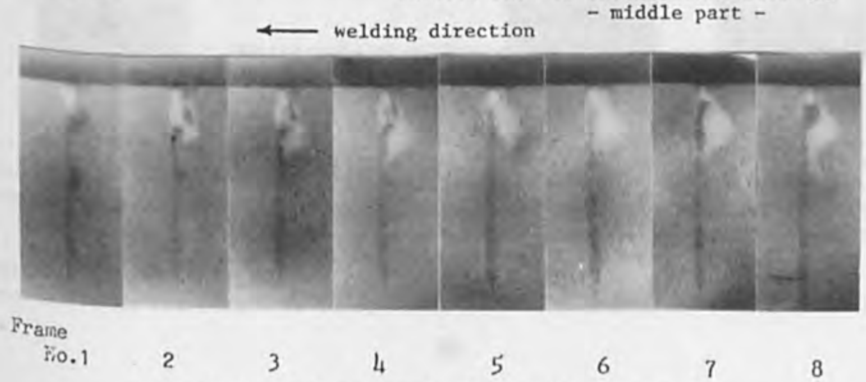


Photo.5 Flow of the molten tracer (2) -upper part-

3.6 Formation of spiking and the influence of elements of higher vaporization pressure

The spiking formation was revealed to have an intimate relation to the instantaneous shape of beam hole. The sequence of this phenomenon was as shown in Fig. 4, where the schematical illustrations were described. These descriptions were substituted for each frames of the movie film taken at 100 frames/sec.. The photograph of 50 frames/sec. was shown in Photo. 6, but it was difficult to represent all the sequence with one paticular series of frames.

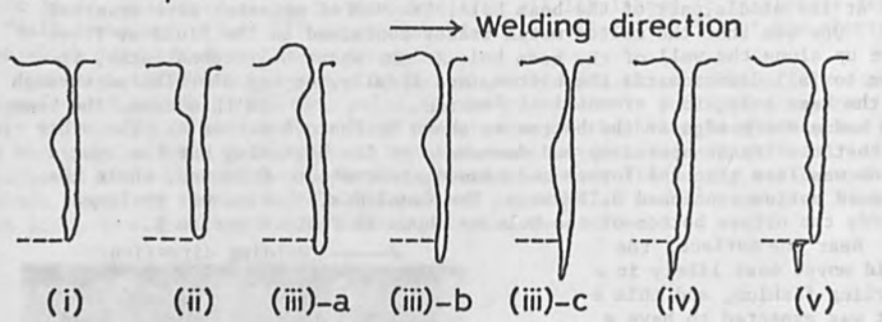


Fig. 4 Schematical illustration of spiking process

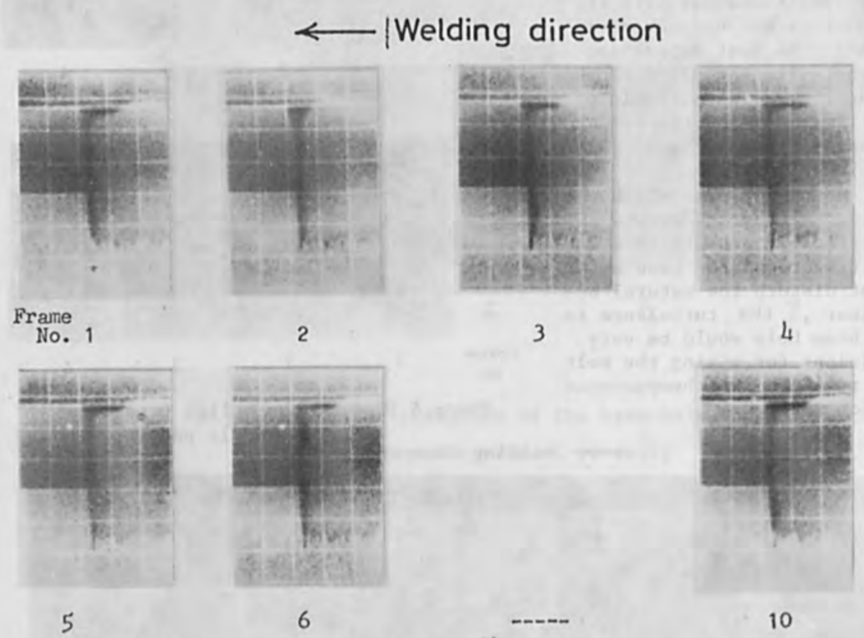


Photo. 6 Formation of spiking (from welding run No. 4)

note: Time interval of each frame is 0.02^{sec}

The spiking is formed through the following processes in the beam hole as shown in Fig. 4.

i) Growth of the wall-knob at the upper part and the constriction of beam hole at the upper openning

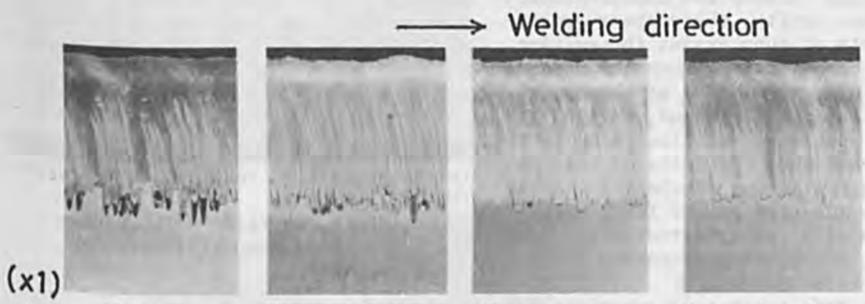
ii) Explosion in the beam hole

iii) Additional penetration (spike) occuring at the
- a) beginning
- b) middle
- c) finish of the process

iv) Axis discrepancy between spike and beam hole

v) Beam hole runs leaving the spike and returns to condition of i)

When the upper wall-knobs (about the first-knob) grow to constrict the upper openning, the constriction reaches to a deep place as shown in i) of Fig. 4. As a result of this, the knob comes upon the wall-cave and increases its pressure. Then at the next stage ii), an explosion occurs inside the beam hole and results in breaking the balance within the cavity. the pressure of the explosion exclude the melt arround the bottom of beam hole and there appears a rather round surface. Successively, the electron beam impinged on the dry bottom produces an additional penetration as shown in iii). As the result, spiking appears and increases in depth within the duration which is from 0.05 sec. to 0.03sec.. Finally the increase of hydrostatic head pressure makes the melt return downwards to the bottom and the beam hole runs leaving the spike and returns to condition of i). Metallurgical investigations of reference (10) which confirms the lack of molten metal at the root of spikes support this argument, although the movement of melt at the bottom of beam hole can not be observed in these films.



1)Brass (Zn:40%) | 2)German silver (Zn:17%) | 3)White copper (Zn:<1%) | 4)Phospher bronze (Zn:trace)

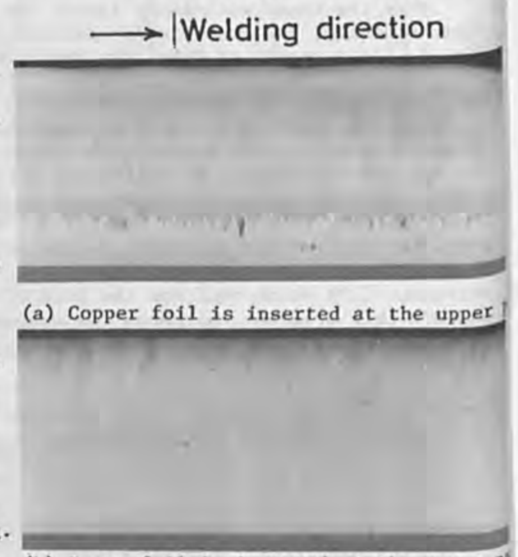
Photo. 7 Macrographs of aluminum specimens showing the profiles of spiking which vary with the kind of inserted metal foils

The profile of spike varies with the kind of inserted metal tracer foils. The inserted foils of brass, red bress and niskel silver make a little shrp and more long spikes, while those of nickel copper and phosphor bronze make a little roundish and short ones as shown in Photo. 7. The evaporation amount of the molten metal in the former is lager than that in the latter due to the more contents of the elements such as Zinc which occur high vaporization pressure. As the result of this, there is less molten metal arround the bottom in the former and stable beam hole with less knobs and caves is established . This is considered to profuce the difference of spiking above-mentioned.

Tong and Giedt[1] proposed a model of the penetration mechanism which is a step in this direction and asserted that the formation of spiking is due to th periodic closure of beam hole by the molten metal. On the other hand, some investigation[6),8),9)] hold that the fluctuation of the energy density of elec- tron beam caused by the collision of electrons with ionized metal vapor result in the irregular penetration which involves the spiking.

The results of our observation provide grounds for maintaining that the spiking can be formed without the existence of such complete closure events as proposed by Tong et al. In accordance with what has been stated above, the main factors which have great influence on the spiking are the metal vapor pressure which occurs due to the violent vaporization at the local zone and th overhanging of molten metal against the wall-caves. Namely occuring of the wall-caves and wall-knobs in the beam hole and their mutual inter- action originate the occurrence of the spiking. It can be easily con- sidered that the breaking of the delicate balance between these fac- tors is not necessery to have a precise periodicity because the fluctuation of energy density of electron beam is suspected to be not directly because the ripples of power input which shows general- ly a higher frequency[3),11)] than that of the spiking observed metallographically Consequently, if the beam hole has a sort of shape which is easy to keep the balance namely stable cavity, there appears possi- biloty of suppressing the spiking.

In practice, when the brass foils were inserted at the upper part of pure aluminum and 5083 al- loy specimens as illustrated in Fig. 3-b and when the foils of some al- loys containing a volatile element were inserted at the lower part of 5083 alloy as illustrated in Fig. 3-a, the spiking decreased as shown in Photo. 8.



(a) Copper foil is inserted at the upper

(b) Brass foil is inserted at the upper p

Photo. 8 Radiographs of aluminium specimen after welding

This event can be explained by the behavior of beam hole.

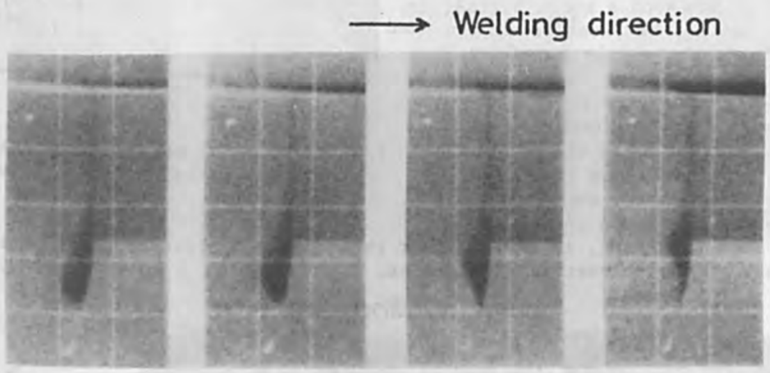
3.7 Behaviour of the beam hole when spiking is suppressed.

When the spiking is being suppressed with the manner described above,the beam hole keeps rather a stable shape because the variation of the wall-knobs is small as compared with the total volume of the beam hole, although the wall-caves and wall-knobs are in a strong fluctuation. The wall-knobs can not constrict strongly the beam hole due to the increased vapor pressure inside the wall-caves. Consequently, the abrupt change of vapor pressure is difficult and the beam hole indicates a stable state

That is the most likely cause of the decrease of spiking.

3.7.1 Behavior of spoon-like beam hole

In the case when the alloy foils are inserted at the lower part of 5083 alloy specimens, the beam holes swell out at the lower side and narrowed at the upper side bcause of heat dissipation at the upper zone due to evaporation of magnesium contained in the base metal. Their shapes are like a spoon as shown in Photo. 9.



1)Brass (Zn:40%) 2)German silver (Zn: 17%) 3)Copper nickel (Zn: 1%) 4)Phosphor bronze (P: 0.35%)

Photo. 9 Shape of spoon-like beam hole which differs with the kind of inserted foils

There is a little difference of their profile coming from the nature of the alloy foils.

This is mainly because of the inner pressure of beam hole which varies with the amounts of volatile elements involved such as Zn. But this gives almost no influence on the suppressiblity of spiking.

In a spoon-like beam hole, there are two ways of strong fluctuation. One is the total fluctuation of its volume as shown in Photo.10 which is represented by the variation of the maximum width of beam hole (d_c) in the longitudinal sections as illustrated in Fig.5.

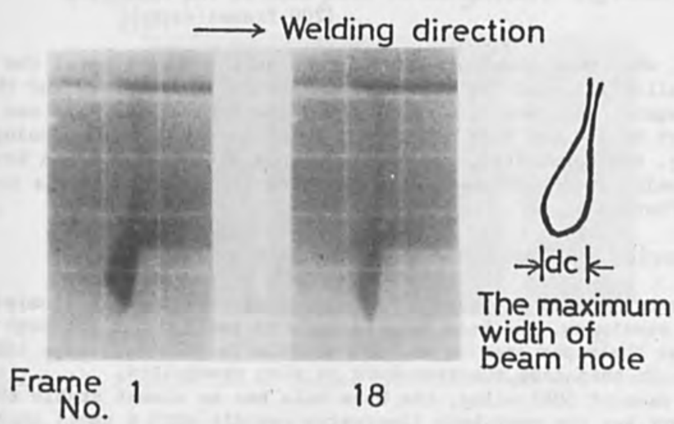


Photo. 10 Volumetric fluctuation of spoon-like beam hole (100 frames/sec)

No regular periodicity is observed in this fluctuation and the range of d_c-variation is more than 35% in the maximum case. It is supposed that the volumetric fluctuation is always irregular in time or it has such a short periodicity that it can not be revealed at the used rate of observation of 100 frames per second

The other is the fluctuation of "root-knob" which occurs characteristically at the rear side of the bottom as shown in Photo. 11 The root knob varies even in the cases where d_c-values are almost same. From this fact, it is seen that the root-knob fluctuation is shorter than that of the volumetric fluctuation.

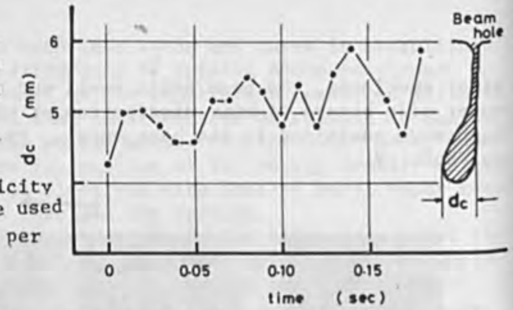


Fig. 5 Relation between the maximum width of beam hole (d_c) and time

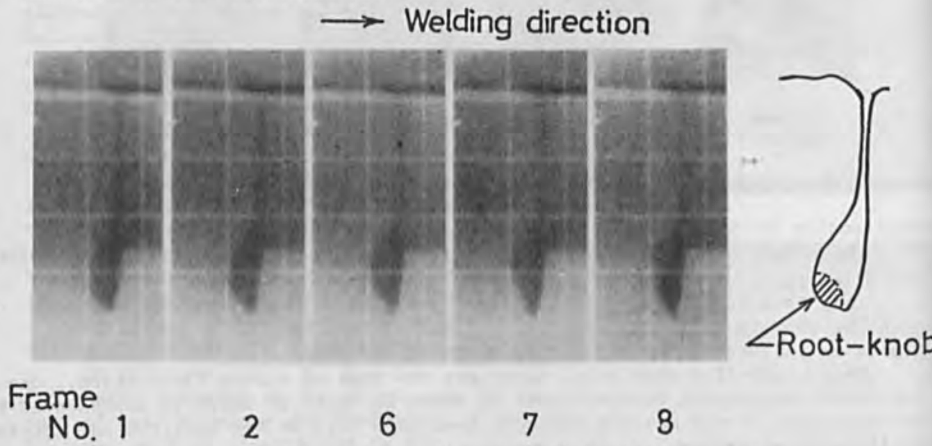


Photo. 11 Fluctuation of root-knob in the beam hole (200 frames/esc)

However, when pure aluminium is employed as the parent metal the effect of inserted alloy foils on the spiking suppression is decreased and the spiking appears again. In such case the surface opening of beam hole was wide. From this fact it is seen that though the metal vapor near the opening vents itself easily, the constriction of the beam hole at the middle and bottom side and the explosion inside the wall-caves at such region lead to the formation of spiking (Photo 23)

3.7.2 Behavior of bamboo leaf -like beam hole

When the brass foil is inserted at the upper part of pure aluminium and its alloying specimens, the beam hole becomes to swelled out although it is constricted at the upper opening and its profile is peculier shape like a bamboo lief. In this case the root-knob is also recognized.

In the case of 5083 alloy, the beam hole has an almost stable shape at the upper part but the root-knob fluctuates rapidly with a small amplitude as shown in Photo. 12.

In the case of pure aluminium the gentle wall-knob is observed at the middle of beam hole in most of frames and fluctuates with time as shown in

Photo. 13. In such both case, the spiking is suppressed.

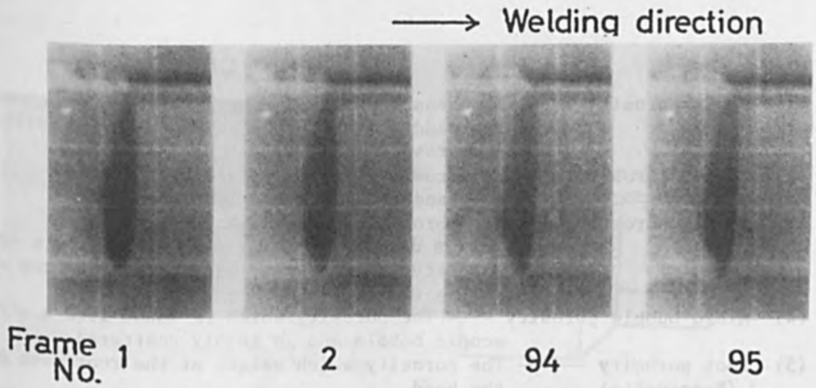


Photo. 12 Fluctuation of root-knob in the bamboo leaf -like beam hole (200 frames/sec)

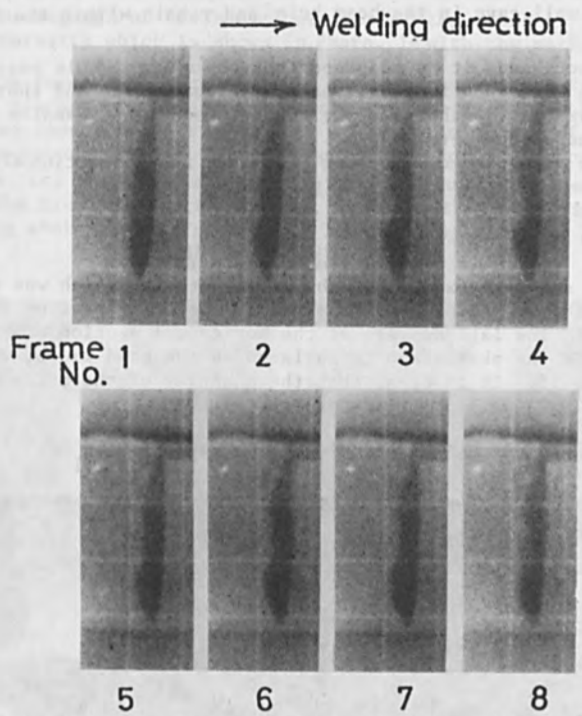


Photo. 13 Fluctuation of bamboo leaf-like beam hole in pure aluminium

3.8 Formation of porosities

There are five kinds of porosities observed in this experiment, and they are classified and termed here as follows.

(1) Cross porosity ----- The considerably big porosity which exists at the middle and the bottom side of the fusion zo[ne] and crosses solidification lines.

(2) Border porosity ----The porosity which lies along the solidification lines and has a rather slender shape.

(3) Bubble porosity ---- The porosity of round shape with a diameter of less than 3 mm. This type porosity is usually crowded at the bottom side of the fusion zone

(4) Micro bubble porosity ---- The porosity which is round like a microscopic bubble and is thinly scattered.

(5) Root porosity ------ The porosity which exists at the root zone of the bead.
(R-porosity)

The formation of porosity of (1),(2) and (3) is a result of the stron[g] fluctuation of wall-cave and wall knob, and is influenced by the cooling condition of their surrounding melt. It is considered that these porosities are left from the wall cave in the beam hole and remain within the fusion zone.

On the other hand, it is supposed that the micro bubble porosities have no direct relation with such fluctuations of the beam hole, and they are resulted from the vaporization of the volatile elements and the expansion of gases contained in the parent metal.

As for the formation of the root porosity, an explanation already presented in 3.4

3.8.1 Cross porosity

Photo. 14 shows an example of the cross porosity which was typically observed when the brass foil was inserted in the upper part of 5083 aluminum alloy specimen. The lateral size of the horizontal section of the porosity marked by (c) in the photograph is as large as the bead width, which is clearly shown in Photo. 15. It is clear that the quantity of the molte metal is quite little on both side walls of this porosity.

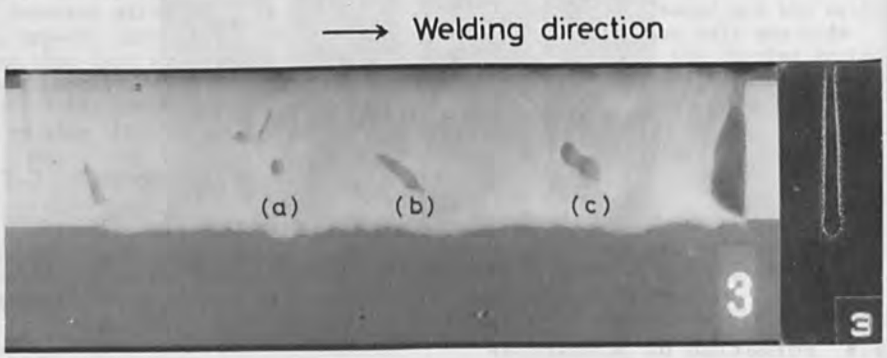


Photo. 14 Radiograph and trans verse section of the specimen which involves the cross porosities

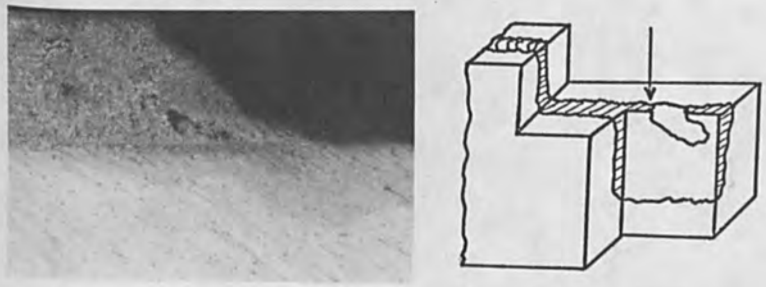


Photo. 15 Macrograph of horizontal section of the cross porosity which is shown in Photo. 14 with the mark (c)

To know the rormation process of the cross porosity, the framing photograph of this process of the porosity marked by (a) in Photo. 14 was taken with the framing speed of 200 frames/sec. The initial stage of the formation process of that porosity is shown in Photo. 16. As for the porosity marked by (c) in Photo.14, its initial stage (f_1, f_2, f_3) is sketched schematically in Photo.17. The final stage (f_4, f_5, f_6, f_7) of this process is also shown by these framing photographs (Frame No.1, 10, 12, 13).

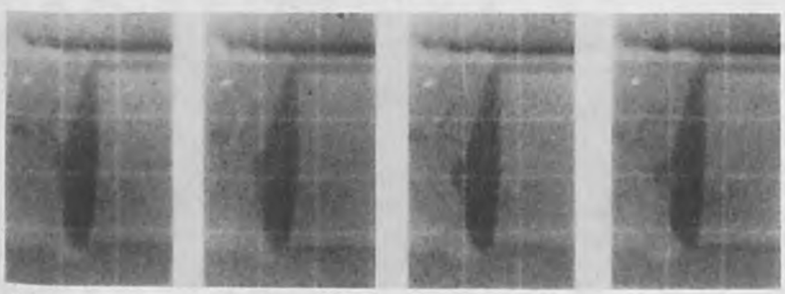


Photo. 16 Beginning of cross porosity formation (200 frames/sec)

From these results it can be explained that bacause of strong perturbation of rear side wall of the beam hole with high inner pressure, a deep wall-cave is formed and spreads out up to the width of the bead . It separates the molten pool into upper and lower parts, and the upper molten metal depresses this slender wall-cave towards the obliquely downwards direction as shown in Photo. 17.

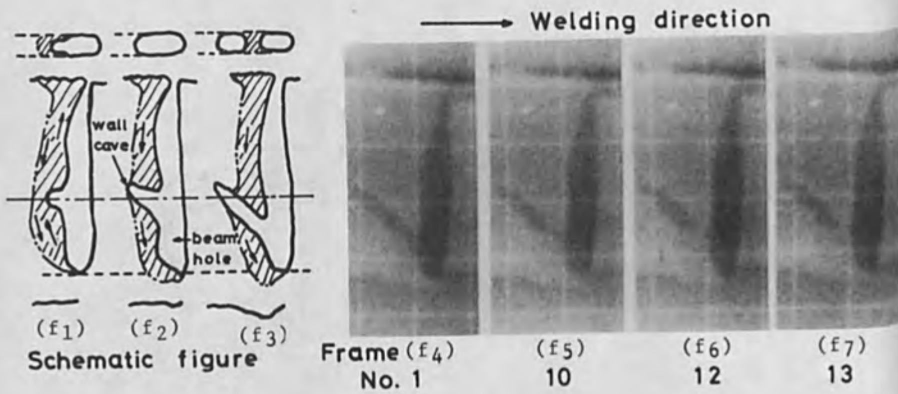


Photo. 17 Formation of cross porosity
(200frames/sec)

3.8.2 Border Porosity

This porosity is characteristically formed along the solidification line as shown in Photo. 18 and Photo. 19.

When the brass foil is inserted in the upper part of the pure alminum specimen, border porosities are frequently seen in the upper part of the beam hole.

In Photo. 19, the local expansion of the beam hole is not so large compared with that of the case of cross porosity, but the quantity of the molten metal is large and its motion is notably violent:. Due to such behaviors of the molten pool, the local buldge of the wall-cave is separated from the mother beam hole as a border porosity.

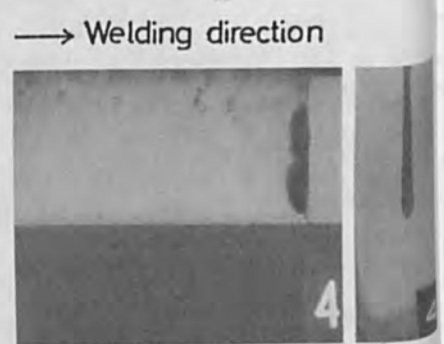


Photo. 18 Radiograph and transverse section of the specimen which involv the border porosities at the upper p

When several alloy foils are inserted in the lower part of the 5083-allo specimen, the border porosities appear at the upper part of the middle wall-cave, along its shoulder as shown in Photo. 20.

3.8.3 Bubble Porosity

This is particularly observed when Phosphor bronze is inserted in the lower part of the 5083-alloy specimen, and it is formed in the upper side of root-knob as shown in Photo. 21.

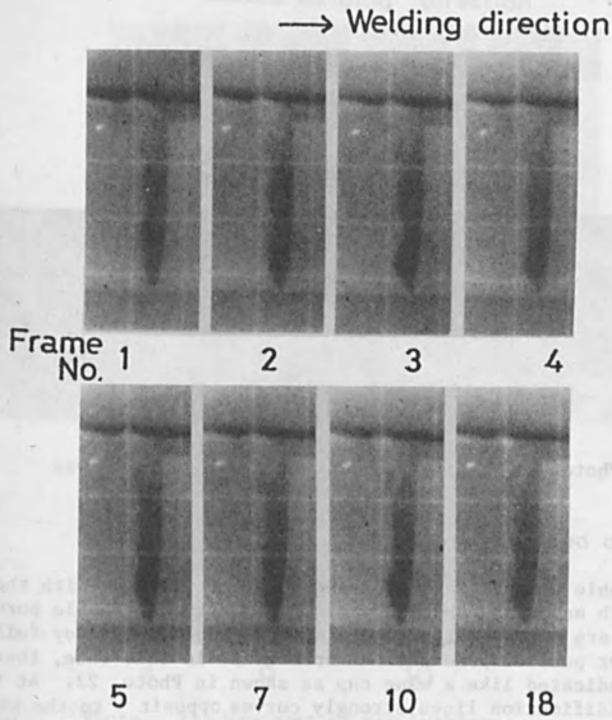
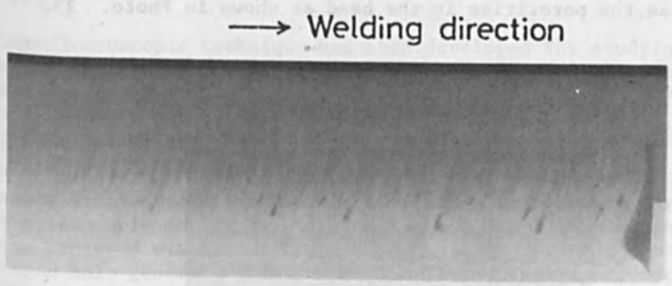
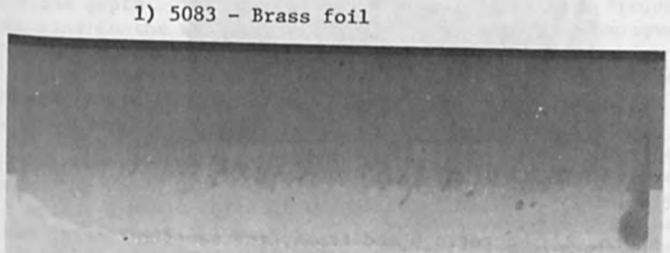


Photo. 19 Formation of border porosity at the upper part (200 frames/sec)



1) 5083 - Brass foil



2) 5083 - Nickel silver foil

Photo. 20 Radiographs of specimens width involve the border porosities

→ Welding direction

Photo. 21 Radiograph of specimen which involves the bubble porosities

3.8.4 Micro bubble porosity

Micro bubble porosity are observed to exist togather with the other porosities such as cross porosity, border porosity, or bubble porosity.

But they are also observed to exist alone when the alloy foils are inse ed at the lower part of pure alminum specimens In this case, there is large fusion zone indicated like a wine cup as shown in Photo. 22. At this upper part, the solidification lines strongly curves opposit to the welding dire tion and the dendritic crystals grow up toward the normal direction to the solidification line. Under such condition, big gas bubbles have enough time to rise up to the surface of the molten pool during the solidification and d not remain as the porosities in the bead as shown in Photo. 23.

→ Welding direction

magnitud (x6) (x1)

Photo. 22 Macrograph and transverse section of aluminum specimen after welding

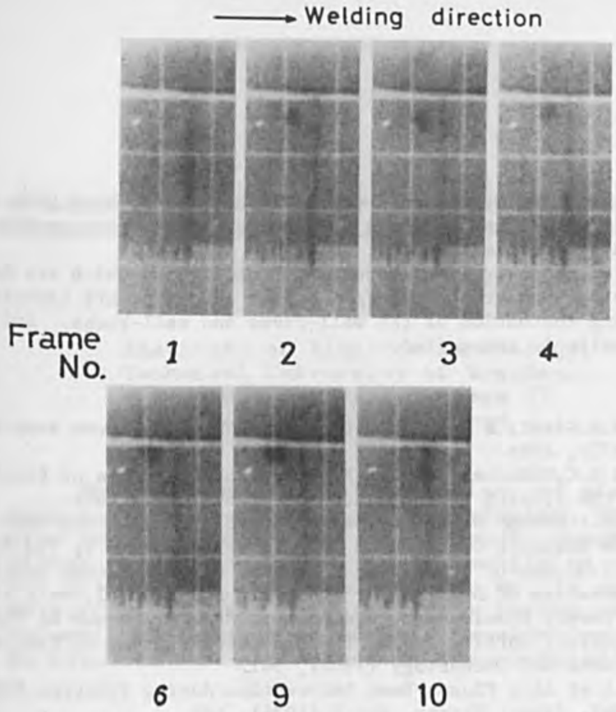


Photo. 23 Movement of gas bubbles
(50 frames/sec)

4 Summary

A cine-fluoroscopic technique has been developed for studying the dynamic mechanism of electron beam welding using aluminum alloys. Although further experiments are necessary to clarify the mechanism more precisely, because a different phenomenon may be observed in the specimens of steels, etc., the following conclusions are possible from the evidence obtained in this experiment.

(1) By means of the technique combined with metal tracers, a direct observation of the beam hole and in part the molten metal during electron beam welding can be achieved with a fairly excellent resolution on the movie film.

(2) The beam hole generally does not hold a steady shape and oscillates in width and depth. The strong wall-oscillation occurs frequently and it gives rise to the wall-caves and wall-knobs. They move upward at the rapid rate of 20-30cm/sec for example.

(3) The higher welding speed becomes the more the beam hole bends like an arch convexed toward the direction of welding progression. This event has a considerable effect on the formation of root porosity.

(4) The formation of spiking results from the breakdown of the balance between the vapor pressure and the hydrostatic head of the melt as well as the dynamic force of the molten fluid. It is revealed that it has intimate relation with the occuring wall caves and wall knobs in the beam hole. It can be surpressed when the stable beam hole is produced due to the swelling of itself in the profile like a bamboo leaf or a spoon by inserting alloy foils which contain a highly volatile element.

(5) The molten metal flows up and down along the wall of beam hole at quite high travel rate. That accounts for the favourable homogeneity of fusion zone caused by metal mixing.

(6) The formation of several characteristic porosities, which are here termed cross porosity, border porosity and bubble porosity, are revealed in relation to the strong fuctuation of the wall-caves and wall-knobs. And also micro bubble-porosity is recognized.

Reference

1) H.Tong and W.H.Giedt: A Dynamic Interpretation of Electron Beam Welding, W.J 49-6 (1970), 259s.

2) E.R.Funk and R.C.McMaster et al.: Penetration Mechanisms of Electron Beam Welding and the Spiking Phenomenon, W.J. 53-6 (1974), 264s.

3) Y.Arata et al.: Energy Distribution in the Hole of Electron Beam Welding Report of EBW Research Commitee of JWS (1975), EBW 123-75, (in Japanese)

4) Y.Arata and I.Miyamoto: Processing Mechanism of High Energy Density Beam (Rept.1) -Mechanism of Drilling- , Trans. of JWRI, Vol.2, No.2 (1973) 19

5) H.Schwarz: Present Knowledge of the Foundamental Processes of Electron Beam as a Material Working Tools, The 3rd Intern. Conf. on Electron and Ion Beam Science and Technology (1968), 301

6) V.P.Ledovskii et al.: Plasma-Beam Interaction during Electron Beam Welding IEE Conf. Publ. (Inst. Electr. Eng.) (1974), 563

7) N.A.Ol'shanskii et al.: Movement of Molten Metal during Electron Beam Welding, Svar. Proiz. 21-9, (1974), 12

8) R.E.Armstrong: Spiking in Partial Penetration Electron Beam Welds, The 4th Intern. Conf. on Electron and Ion Beam Science and Technology (1970),179

9) G.K.Hicken and W.G.Booco: Penetration Variations in Electron Beam Welding The 3rd Intern. Conf. on Electron and Ion Beam Science and Technology, (1968), 398

10) Y.Arata, S.Matsuda et al.: Study on Characteristics of Weld Defect and Its Prevention in Electron Beam Welding (Report III), Trans. of JWRI, 3-2, (1974), 81

11) V.I.Leskov et al.: Mechanism in Deep Weld Pools During Electron Beam Welding, Avt. Svarka. 28-1, (1975), 12

12) J.W.Meier: Electron Beam Welding Characteristics of Several Materials, Proc. on EB Tech., (1961), 145

INFLUENCE OF THE ELECTRON GUN PARAMETERS ON WELD PENETRATION DEPTH IN THE ELECTRON BEAM WELDING METHOD

K.P. Friedel

Institute of Electron Technology,
Technical University of Wrocław,
Wybrzeże S. Wyspiańskiego 27
50-370 Wrocław, Poland

The power density distribution of electron beam near cross-over image is evaluated. The region in which electron beam power density is higher than the critical power density is called the active zone. Influence of the cavity entrance orifice on the electron beam active zone is determined. Dependence of weld penetration depth on electron gun parameters is investigated.

INTRODUCTION

The effect of the deep weld formation is not thoroughly explained. However, it is known that a deep weld, in contradistinction to common hemispherical form of the weld is related with formation of the cavity [1,2]. Mechanism of the cavity formation may be explained by reaction of vapor stream [3], or by the micro-explosion phenomenon [4]. In each case formation of the cavity by the electron beam is possible after reaching a critical power density. The critical power density is a constant value and depends on the thermophysical properties of the solid material [5]. The cavity is an unstable creation [1,2,6]. The cause of instability is a liquid layer surrounding the cavity, especially a big quantity of melted material in the neighbourhood of the cavity entrance orifice [7].

The electron beam impacting on surface of the workpiece forms the cavity of a diameter dependent on power density distribution on surface of the workpiece /Fig. 1/. The cavity formation condition is fulfilled in central part of electron beam where power density exceeds the critical value. Peripheral part of electron beam contributes only to formation of a characteristic bowl at the surface, as a result of hemispherical type weld. The region in which electron beam power density is higher than the critical power density, is called the active zone [8,9]. The entrance orifice of the cavity acts as a diaphragm situated on the way of the electron beam, changing the former shape of the active zone of the electron beam penetrating into the workpiece. So, the cavity depth is closely dependent on the position of the electron beam active zone in relation to the workpiece surface, as well as on the diameter of the entrance orifice of the cavity. Instability of the entrance orifice diameter causes a fluctuation of the cavity depth.

Many authors have investigated the influence of electron beam parameters on the welding depth [8...20]. In this paper a trial was undertaken to connect the basic paramters of the electron beam with the welding depth, basing on a simplified model of the electron beam, taking into consideration the shape and position of the electron beam active zone in relation to the workpiece surface. It was assumed that the electron beam active zone is determined only by a kind of the electron source, by properties of ideal electron optical system, by Maxwellian emission velocity distribution by diameter of the cavity entrance orifice, and by a kind of the workpiece being welded. Thus, the influence of space charge, electron optical aberrations and dissipation of the electron beam by the vapours of the material being welded were not be taken into account. Moreover, theconsiderations concern the quasi-static conditions only, i.e. the welding speed is low and time of cavity formation is considerably shorter than the time required for displacement of electron beam by the distance equal to the cavity diameter. Such the assumptions are not fully adequate, but they allow to reveal qualitative dependencies between the electron beam parameters and the welding depth.

DETERMINATION OF THE ELECTRON BEAM ACTIVE ZONE

In order to calculate the configuration of the electron beam active zone it was assumed that the virtual electron source is a cross-over situated on the plane $z=z_0$/Fig. 2/. From the elementary area ds_0 of the cross-over plane the electrons go out with the intensity dependent on a direction and go into the region of the influence of axial symmetry magnetic field. In this region electrons change their motion direction but do not change their energy. The radiation characteristic of the cross-over is a known function $B(r_0,\omega_0,U)$. Let dN_0 denote the number of electrons going out from the elementary area ds_0 in the time τ, in solid angle $d\omega_0$. Assuming that the electron beam is monochromatic [21]:

$$\frac{dN_0}{d\tau} = B\, ds_0 d\omega_0 \qquad /1/$$

Thus, through the elementary area ds on the plane z=const., in a time unit pass:

$$\frac{dN}{d\tau} = dj_z\ ds \qquad /2/$$

electrons, where j_z denoting current density. From the condition of particles number conservation $dN_o = dN$ follows:

$$dj_z\ ds = B\ d\omega_o\ ds_o \qquad /3/$$

Then, power density in the plane z=const. is

$$dp = B\ U\ d\omega_o \frac{ds_o}{ds} \qquad /4/$$

where U - accelerating voltage.

Analytical determination of the elementary area ds requires knowledge of the electron trajectories equations, and would be very difficult if considerable simplifications would not be assumed.

As follows from Fig. 2, the current density in point Q(r,z), situated in the caustic region, will be a result of the super-position of all elementary electron beams deriving from the circular cross-over area of radius ρ with the centre in point r_o.

The elementary area on the plane z is

$$ds = 2\pi\xi d\xi \qquad /5/$$

where

$$\xi = a\alpha\frac{b - z}{b} \qquad /6/$$

$$d\xi = a\left(1 - \frac{z}{b}\right)d\alpha \qquad /7/$$

The elementary area on the plane of cross-over

$$ds_o = \rho d\rho\ d\psi \qquad /8/$$

is situated in the r_k distance from the axis:

$$r_k^2 = \rho^2 + r_o^2 + 2r_o\rho\cos\psi \qquad /9/$$

where $r_o = \frac{r}{R}$; $R = (M + 1)\frac{z}{b} - 1$, /10/

M - magnification.

The radiation characteristic B is essentially equivalent to electron brightness. For small beam semi-angle, when electron brightness is almost independent on α [22]:

$$B(r_k) = \frac{j_o(r_k)}{\pi\alpha^2} \qquad /11/$$

where $j_o(r_k)$ - current density in the point r_k of cross-ov

Assuming that current density distribution in the cro over is normal distribution [22]:

$$j_o(r_k) = j_o \exp\left(-\frac{r_k^2}{r_c^2}\right) \qquad /1$$

where j_o - peak-on-axis current density,

r_c - hypothetic radius of the cross-over.

Finally the radiation characteristic is

$$B = \frac{j_o}{\pi\alpha^2} \exp\left(-\frac{\rho^2}{r_c^2} - \frac{r^2}{r_c^2 R^2} - \frac{2\rho r}{r_c^2 R} \cos\psi\right) \qquad /1$$

Substituting the equations 5, 8 and 13 to the equation 4, taking into account that the element of the solid angle is

$$d\omega_o \approx 2\pi\alpha d\alpha \qquad /1$$

gives

$$dp = \frac{p_o \rho}{\pi\alpha^2 a^2 \left(1 - \frac{z}{b}\right)^2} \exp\left(-\frac{\rho^2}{r_c^2} - \frac{r^2}{r_c^2 R^2} - \frac{2\rho r}{r_c^2 R} \cos\psi\right) d\rho d\psi \qquad /1$$

As a result of integration within the whole circular area maximal radius ρ_m one gets

$$p(r,z) = \frac{p_o r_c^2}{\rho_m^2 R^2} \left\{1 + \frac{r^2}{R^2} + \frac{r^4}{2R^4} - \frac{\rho_m^2}{\exp(\rho_m^2) - 1}\left[\frac{r^2}{R^2} + \frac{r^4}{4R^4}\left(\rho^2 + 2\right)\right]\right\} x$$

$$x \left[1 - \exp\left(-\rho_m^2\right)\right] \exp\left(-\frac{r^2}{R^2}\right) \qquad /1$$

$$\text{where} \quad \rho_m = \frac{\xi}{R} \qquad /1$$

After substituting $p(r,z) = p_{cr}$ /where p_{cr} - critical power density/, this equation describes the surface of the electr beam active zone.

In the crossover image plane, for $z = b$ this distribut takes the form:

$$p(r,b) = \frac{p_o}{M^2} \exp\left(-\frac{r^2}{r_c^2 M^2}\right) \qquad /1$$

INFLUENCE OF THE CAVITY ENTRANCE ORIFICE ON THE ELECTRON BEAM ACTIVE ZONE

Determination of a modified power density distribution in the region of caustic with presence of a diaphragm - entrance orifice of the cavity makes hard difficulties. Then, determination was limited to finding of a modified power density on z axis only, which is enough for calculation of the maximum cavity depth.

The formula 15 for the points situated on z axis, i.e. for r=0, will take the form:

$$dp = \frac{p_o \rho}{\pi \rho_m^2 R^2} \exp\left(-\frac{\rho^2}{r_c^2}\right) d\rho \, d\psi \qquad /19/$$

Integration is carried out within circular area of maximal radius:

$$\rho_m' = \frac{r_A}{z - z_A} \cdot \frac{b - z}{M} \qquad /20/$$

where r_A - radius of cavity entrance orifice,
z_A - position of cavity entrance orifice on z axis.
The ρ_m' is different from ρ_m, because of the diaphragm presence /Fig. 3/.

As a result of integration one gets

$$p(0,z) = \frac{p_o r_c^2}{\rho_m^2 R^2} \left[1 - \exp\left(- \frac{r_A^2 (b - z)^2}{r_c^2 M^2 (z - z_A)^2} \right) \right] \qquad /21/$$

This formula describes the power density distribution along z axis, beginning from the entrance orifice of the cavity. The condition for formation of the cavity is that $p(0,z) > p_{cr}$.

DISCUSSION OF THE RESULTS

According to the assumptions, the electron beam will penetrate into the workpiece forming the cavity to the depth determined by the coordinate z, at which the power density on the axis falls down to the critical value. The cavity depth is than

$$H = z - z_A \qquad /22/$$

In order to calculate H, the orifice radius r_A must be found and substituted to the formula 21, and $p(0,z) = p_{cr}$ must be assumed. Radius of the cavity entrance orifice can be calculated from the formula 16, assuming $p = p_{cr}$ and $z = z_A$.

Dependence of so calculated depth of the cavity was investigated on:

- position of the active zone in relation to the workpiece surface;
- diameter of the cross-over;
- beam semi-angle in the cross-over;
- peak-on-axis power density in the cross-over;
- critical power density;
- position of the magnetic lens.

The calculations were carried out by means of digital computer.

Position of the electron beam active zone in relation to the workpiece surface can be changed by changing magnification M. As a measure of active zone position the distance h of the cross-over image from the workpiece surface was taken. In Fig. 4 dependence of the cavity depth on the active zone position is presented. The cavity of max. depth is obtained under the conditions when the cross-over image is under the workpiece surface. Thus, subjective focusing of electron beam on workpiece surface does not ensure obtaining of the maximum weld depth.

The larger the diameter of the cross-over /at constant overall power of the electron beam/, the smaller max. cavity depth /Fig. 5/. It is obvious, because power density in the cross-over is then decreased, and consequently, dimensions of the active zone of the electron beam are also decreased.

In Fig. 6 dependence of the maximum cavity depth on the beam semi-angle at established overall power and power density in the cross-over is presented. As is seen, for smaller angles, α, maximum depth of the cavity is considerably greater. The dotted line presents dependence of optimal position of the active zone h_{opt}, at which the maximum **weld** depth is obtained, on semi-angle, α. The smaller the beam semi-angle, the deeper the electron active zone should be shifted under the surface of the workpiece, in order to obtain the maximum weld depth.

On the ground of the cavity depth dependence on critical power density to peak-on-axis power density ratio p_{cr}/p_o /Fig. 7/ it may be found, what was to be expected, that for the workpieces of higher critical power density the cavity depth is smaller, and optimal position of the cross-over image is nearer the surface. Analogically changes the cavity depth and optimal position of the active zone when maximal power density in the cross-over decreases.

There is an optimal position of the magnetic lens at an established workpiece to cross-over distance. In Fig. 8 dependence of this position on the power density ratio p_{cr}/p_o is presented. As follows from this dependence, the more hardly weldable the workpiece is, i.e. the higher the critical power density, the nearer to the workpiece the magnetic lens must be shifted. It means that the gun-to-work distance must be shorten. It is characteristic that the magnetic lens position influences not only the maximum depth of the cavity but also considerably influences character of changes of the cavity depth as a function of the active zone position /Fig. 9/. With the increase in the gun-to-work distance in relation to the optimal distance, the smaller cavity depth is obtained but considerably more reproducible depth at eventual inaccuracy of electron beam focusing /fluctuation of the active zone configuration/.

Because the diameter of cavity entrance orifice may oscillate around the mean value, it was investigated how a ± 10% change in the diameter influenced the cavity depth /Fig.10/. It appeared that this influence took place, to the highest degree, in the neighbourhood of the active zone optimal position. The decrease in the fluctuations of the cavity depth can be obtained, among others, at the cost of decreasing the average depth of the cavity, either by incrising the gun-to-work distance or by the shift of the cross-over image towards the surface /subjective focusing on the workpiece surface/.

CONCLUSIONS

Basing on above considerations it was found that:

1. the cavity depth dependes on the position of the electron beam active zone in relation to the workpiece surface;
2. the diameter of the cavity entrance orifice considerably influences the shape of electron beam active zone;
3. maximum depth of the cavity is obtained for the image of the cross-over situated under the workpiece surface;
4. maximum cavity depth depends on a kind of the material, electron beam power density, electron beam convergence angle, and position of the magnetic lens;
5. the cavity depth highest fluctuations occur at approximately optimal position of the electron beam active zone in relation to the workpiece surface.

ACKNOWLEDGEMENT

Author would like to express his thanks to Mr.B. Kałat for his kind help in programming and carrying out the calculations.

REFERENCES

1. Tong H.,Giedt W.H., Weld.J.,Vol 49,No 6,1970,pp 259s-266
2. Arata Y.,Matsuda F.,Murakami T., Trans.of JWRI,Vol 2,No2 1973,pp 23-32;
3. Schwartz H., J.Appl.Phys.,35,7,1964,pp 2020-2029;

4. Zuev I.W.,Rykalin N.N.,Uglov A.A., Fiz.i khim.obrab.mat. No 5,1968,pp 7-15;
5. Zuev I.W.,Rykalin N.N.,Uglov A.A., Fiz.i khim.obrab.mat. No 3,1970,pp 3-7;
6. Leskov G.I.,Trunov E.N.,Zhivaga L.I., Avtomat.Svarka,No1 /262/,1975,pp 12-15;
7. Weber C.M., 5-th Int.Conf.on Electr.and Ion Beam Sci.and Technol.,1974,pp 13-16;
8. Jüptner W.,Bödecker V.,SepoldG., 6-th Int.Conf.on Electr. and Ion Beam Sci.and Technol.,1974,pp 366-376;
9. Arata Y.,Tomie M.,TeraiK.,Nagai H.,Hattori T., 6-th Int. Conf.on Electr.and Ion Beam Sci.and Technol.,1974, pp 377-398;
10. Bas E.B.,Cremosnik G., Vakuum-Technik,Vol 18,No 7,1958, p 181;
11. Hashimoto T.,Matsuda F., Trans.NRIM,Vol 7,No 3,1965,pp27 -40;
12. Passoja D.E., Brit.Weld.J., Vol 14,No 1,1967,pp 13-16;
13. Lubin B.T., Weld.J.,Vol 47,No 3,1968,Res.Suppl.,pp 140s -144s;
14. Sanderson A., Brit.Weld.J.,Vol 15,No 10,1968,pp 509-523;
15. Sandstrom D.J.,Buchen J.F.,Hanks G.S., Weld.J.,Vol 49, No 7,1970,Res.Suppl.,pp 293s-300s;
16. Konkol P.J.,Smith P.M.,Willebrand C.F.,Connor L.P., Weld J.,Vol 50,No 11,1971,pp 765-776;
17. Zuev I.W.,Rykalin N.N.,Uglov A.A., Fiz.i khim.obrab.mat. No 1,1972,pp 9-14;
18. Weber C.M.,Funk E.R., McMaster R.C., Weld.J.,Vol 51,No 2 1972,Res.Suppl.,pp 90s-94s;
19. Weidner C.W.,Shuler L.E., Weld.J.,Vol 52,No 3,1973,Res. Suppl.,pp 114s-119s;
20. Dumonte P.,Sayegh G., 6-th Int.Conf.on Electr.and IonBeam Sci.and Technol.,1974,pp 346-365;
21. Glaser W.,"Grundlagen der Elektronenoptik",1952,Springer Vienna;
22. Moss H.,"Narrow Angle Electron Guns and Cathode Ray Tubes",1968,Academic Press, New York, London.

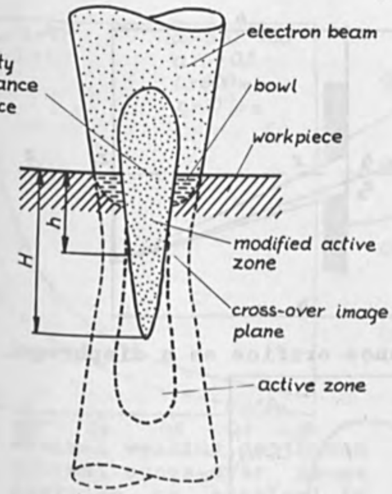


Fig.1 Effect of the cavity entrance orifice on electron beam active zone. H - cavity depth, h - position of cross-over image.

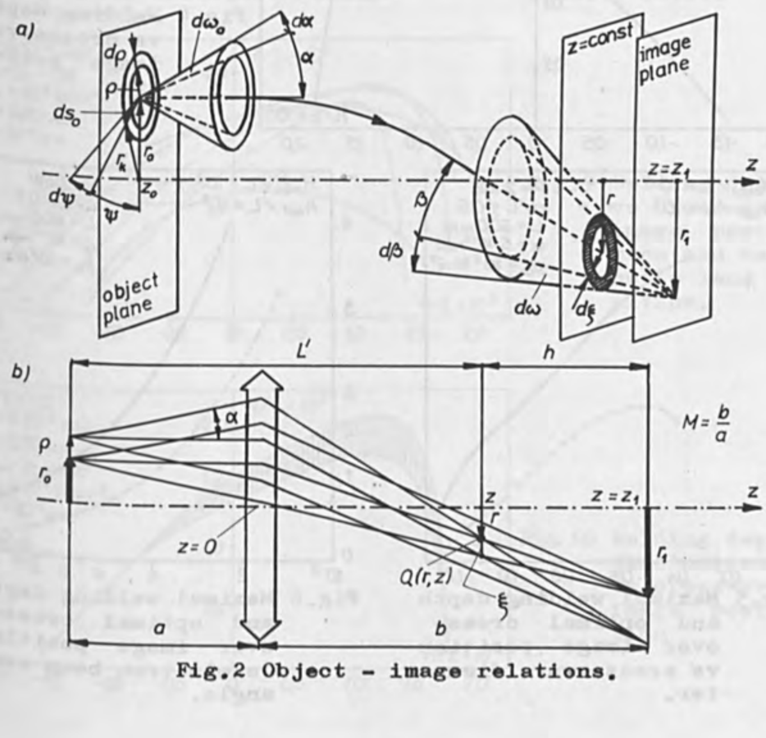


Fig.2 Object - image relations.

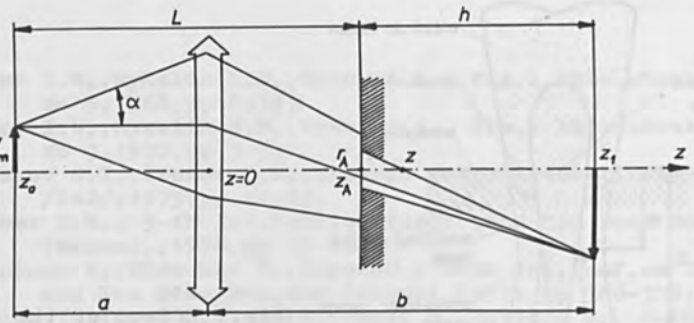


Fig.3 The cavity entrance orifice as a diaphragm.

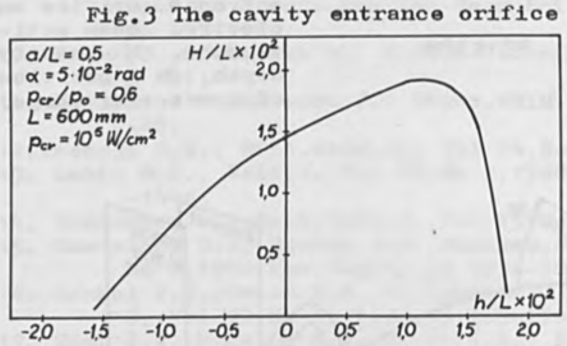


Fig.4 Welding depth vs cross-over image position.

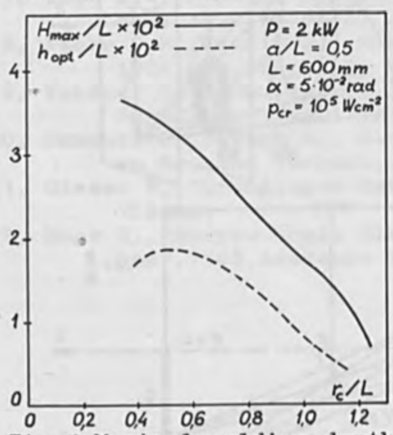


Fig.5 Maximal welding depth and optimal cross-over image position vs cross-over diameter.

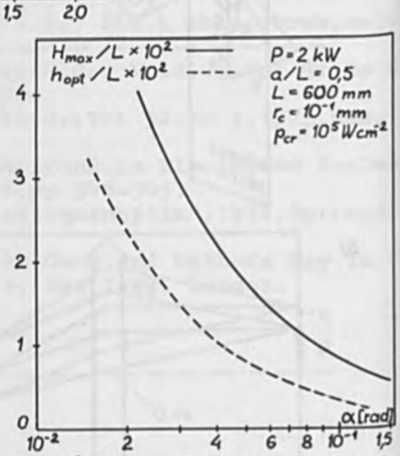


Fig.6 Maximal welding depth and optimal cross-over image position vs electron beam semi-angle.

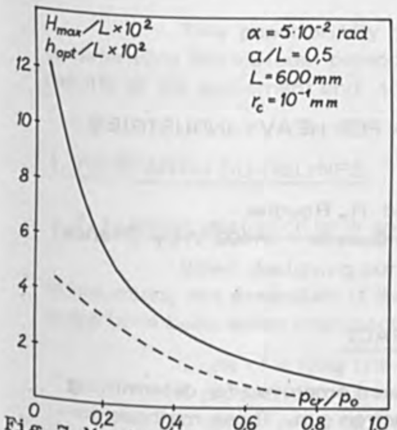


Fig.7 Maximal welding depth and optimal cross-over image position vs critical to peak-on-axis power density ratio.

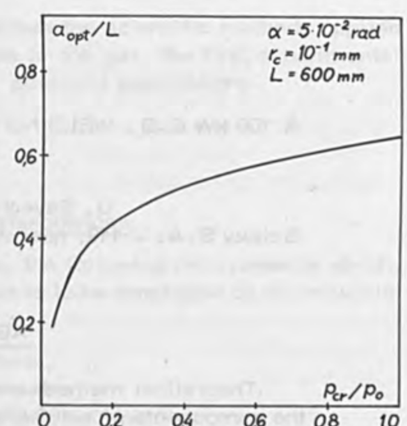


Fig.8 Magnetic lens position vs critical to peak-on-axis power density ratio.

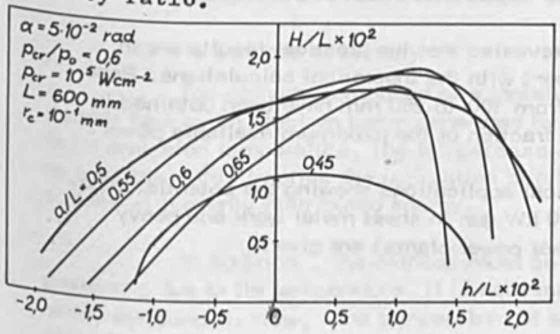


Fig.9 Welding depth vs cross-over image position and magnetic lens position.

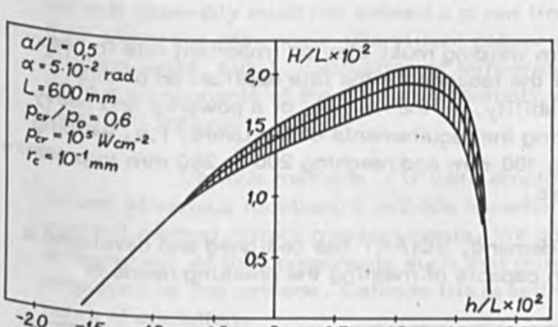


Fig.10 Welding depth fluctuations vs cross-over image position

A 100 kW E.B. WELDING GUN FOR HEAVY INDUSTRIES

G. Sayegh and R. Roudier
Sciaky S.A. - 119, quai Jules Guesde - 94400 Vitry (France)

ABSTRACT

Theoretical methods have been employed for determining the components of a 100 kW electron gun. These methods solve the various equations which define the electron beam in the presence of the electrical field in the gun, of the magnetic focusing field and under the influence of space charges with due account of the relativistic effect of the electrons.

Gun testing revealed that the practical results are in complete agreement with the theoretical calculations. Penetrations ranging from 100 to 250 mm have been obtained in steel with only a fraction of the maximum available power.

Finally, typical applications showing the potential capabilities of the 100 kW gun in sheet metal work and heavy machinery (nuclear power plants) are given.

INTRODUCTION

Electron beam welding must play an important role in heavy sheet metal work. One of the reasons for the late application of this process is the non-availability, on the market, of a powerful and heavy duty gun capable of meeting the requirements of the users, i.e. welding of steel gauges exceeding 100 mm and reaching 200 to 250 mm in pressure vessels or in turbine rotors.

To meet the demand, SCIAKY has designed and developed a 100 kW E.B. welding gun capable of meeting the pressing needs of manufacturers.*

* This project was financially supported by the "aide au développement" of the french Délégation Générale de la Recherche Scientifique et Technique.

This paper briefly describes the scientific methods applied to determine the various components of the gun, the first experimental results of the equipment and some potential applications.

1. RESEARCH GUIDELINES

1.1. Basic design criteria and consequences

When designing our gun, the following requirements which, in our mind, are essential if the gun is to be employed on an industrial scale have been taken into account :

- use of a long life cathode,
- production of a narrow beam,
- gun of reduced dimensions and mobile,
- use of a small diameter anode to avoid metal vapor release within the gap between the electrodes.
- long gun-to-work distance,
- oscillation of the beam onto the target.

1.1.1. Cathode assembly

The cathode composed of a tantalum pellet is bombarded at its rear face by an electron beam generated by an auxiliary gun and brought to its emission temperature. The temperature must be as homogeneous as possible : this implies the utilization of a beam striking the impact surface with evenly distributed energy.

In addition , the cathode must be mechanically held in position ; due to its temperature, it is necessary to heat the components which surround it. Now, the temperature of several components of the cathode assembly must not exceed a given limit or chances of damaging their properties may occur (Beryllium copper annealing temperature, silicone seals, etc.). Evaluation of temperature distribution in the various components of the cathode assembly was thus considered as an important feature.

Various methods /1/ were employed to estimate the temperatures at various locations of cathode assembly : analog method, finite element method, direct measurements. We could thus determine an arrangement of the components such that minimum heating power is employed for the cathode. Cathode life is estimated at more than 50 hours of operation.

Finally, it must be pointed out that the baseplate which supports the cathode assembly and through which the various voltages are conveyed includes, at its contact surface with atmospheric pressure, a cooling system via a compressed air jet. The cathode assembly and the cooling system have been patented.

1.1.2. Optimization of the electrodes

Electrode shapes (wehnelt, anode) have been optimized thanks to a computerized design program which makes it possible to calculate beam characteristics in a given geometric configuration of the gun.

This program gives the numerical solutions of the various integro-differential equations which characterize the electron gun : Poisson's equation, electron trajectories, cathode emission... Thus, for a given configuration of the electrodes, we can know beam characteristics (power, power distribution, envelope of the beam...).

For the 100 kW gun, we tried to find an arrangement between two contradictory trends as follows :

. use of a very high accelerating voltage, hence of a low current intensity. This will facilitate beam concentration on a narrow section - an interesting feature since the influence of space charge is less critical - (the space charge is proportional to current density/voltage square root) but has the disadvantage of necessitating a large, cumbersome gun ; this is a drawback because of the long insulation distances required to ensure adequate supply of high voltage.

. use of a low accelerating voltage, hence of large current quantities. The advantage is a gun of smaller dimensions since insulating distances will be shorter but the drawback is a rather wide beam.

For our gun, we selected / 2 / the following operating parameters : 100 kV - 1 Amp. The beam produced by the gun has a diameter of 0.6 mm at the cross-over located in the anode and a divergence of 35 milliradians at the same place.

Figure 1 shows electron-trajectories in the electrostatic portion of the gun for the various polarizations of the wehnelt. The shape of the hole drilled in the anode gives the beam free passage considering the manufacture and alignment tolerances of the different components.

1.1.3. Determination of beam characteristics in the magnetic portion

The magnetic focusing coil is essential in gun performances. For medium-power guns (15 kW - 100 kV for example) beam characteristics in the magnetic field can be determined using simplified methods featured in particular by Gauss optics assumptions / 3 /. It appeared to us that these simplified methods are not suitable because of the limiting assumptions admitted in the design of the 100 kW - 100 kV gun. Therefore, we undertook to develop a computerized calculation method capable of solving the equations of the electron trajectories in the presence of a magnetic field under general conditions and taking into account the influence of space charge and the relativistic influence of the electrons.

The method consists in calculating first, by the finite element technique (figure 2) the magnetic vector of an armored coil of a given number of ampere-turns. These magnetic vector potential values are introduced into a second computerized calculation program which operates on the finite difference method to give a digital solution for the various equations featuring the beam.

Figures 3 and 4 show the electronic trajectories drawn by the computer in the magnetic field between the anode and the target.

It must be noted first that the scales of axes r and z are markedly different.

The important influence of the space charge on beam shape is a significant feature. Beam envelope is augmented and the focusing point is completely modified by the space charge.

We also noted that beam characterisitcs completely differ from the ones which can be expected from the simplified methods employed in Gauss optics. We therefore recommand the use of the program we have developed to determine the performance of high power beams in narrow sections, such as in the 100 kW gun.

1.2. Description of the 100 kW gun

Figure 5 is a general view of the 100 kW gun in the horizontal position. In this paper, we shall give only a brief description of the setup employed for testing the gun.

The gun is located in a 5 m^3 vacuum chamber pumped by a suitable pump set. Gun overall length is 800 mm, the maximum diameter is 300 mm and the weight 220 daN.

The emission surface has a 5 mm diameter. The beam is concentrated by electric field components in a 0.6 mm spot near the anode. The diameter of the hole drilled in the anode is 2.5 mm and its length 45 mm, thereby reducing the section for metal-vapor passage towards the electrodes with the corresponding reduction in the chances of breakdown.

Joint locating in ensured by an optics system with a 90° prism designed to align the gun axis to the joint but this can also be done by a TV camera or a "Reflectron" (electronic seam tracker) / 4 /.

The gun is supplied by a source incorporating a transformer and a rectified bridge which generates a highly stable high voltage ; this source also includes a transformer for the supply of the auxiliary gun and a bias supply capable of completely cutting-off electronic emission at the highest accelerating voltage.

The compressed air (at 5 bars) is conveyed to cathode assembly supporting plate through a suitable pipe and, once heated, flows out via openings provided for that purpose.

Finally, the focus coil allows welding at gun-to-work distances ranging from 150 to 1000 mm while two pairs of deflection coils modify beam power distribution on the work.

2. ANALYSIS AND INTERPRETATION

2.1. Gun characteristics and experimental measurements

Figure 6 shows the comparison between the theoretical curve obtained by calculating beam current and the experimental results measured on the equipment, as a function of the accelerating voltage.

An excellent correspondance of the results can be noted and shows that :

. theoretical methods can be used to estimate beam characteristics.

. the gun remains stable from a dimensional point of view over the whole range of gun powers.

In addition, we used an experimental method / 5 / to measure the distribution of beam power at the focusing point. This method consists of deflecting very rapidly the beam on a Faraday cage covered with a very narrow (50 μ) straight slit and recording the collected signal. Figure 7 shows this distribution. From this curve, we can determine beam diameter which is 2.2 mm in the specific case under consideration (gun-to-work distance : 800 mm).

This diameter has an intermediate value between the two theoretical figures obtained by calculation (figure 8). This can be explained by the fact that at the focus coil, which is under soft vacuum, part of the electronic space charge is neutralized by the positive ions produced by the ionization of gas molecules hit by the beam. Consequently, the actual influence of the space charge is lower than indicated by theoretical calculation.

2.2. A few penetration results in A 533, grade B steel

A few typical examples of penetration using optimized welding parameters are given. Welding parameter optimization is achieved by trying to find the most suitable power, focusing point and beam vibration so as to obtain sound welds.

Figures 9, 10 and 11 are examples of penetration in 100, 150 and 200 mm steel workpieces with the corresponding welding parameters. The main features of the joint are : narrow and parallel sides and deep penetration.

It is to be noted that 200 mm penetration is obtained with a gun operating at two-thirds of maximum available power, thus leaving an important margin for further penetration.

We believe that 300 mm gauges can be welded at 10 cm per minute with the gun operating at full capacity.

3. INDUSTRIAL PROSPECTS AND EXAMPLES

This electron gun represents the basic tool we have been waiting for during several years to improve welding thick gauges in structural steelwork and heavy sheet metal work. Indeed, this process, now available for production units, has the following advantages :

- simpler joint preparation : machined single U or double Vee grooves are replaced by square edges ;

- much greater welding speed (example : butt-welding of dia. 3 m shells, thickness 200 mm, preparation 2 hours, welding 1 hour at 15 cm per minute) ;

- elimination or significant reduction of distorsion. Thus for certain configurations, it is possible to eliminate very expensive machining operations or shaping operations after welding ;

- reduction of the molten zone and consequently of the residual stresses ;

- the complete automatic control of the process ensures perfect reproducibility of the results ; the problems in connection with skilled welders become less critical. Parameter recording improves operation monitoring capability.

Below are listed several examples of potential 100 kW gun industrial applications for joining heavy structures. They will be classified according to their respective probability of realization, taking into account metallurgical checks and the solution to the pending problems related to manufactured codes.

The examples given involve the use either of local vacuum chambers / 6 / or of total vacuum large chambers. Chamber selection is done as a function of workpiece shape and of economic factors.

3.1. Welding nuclear reactor internal components

The tolerances imposed for the assembly on nuclear reactor internal components represent a heavy burden in these applications. Electron beam welding can be a leading advantage from the points of view of freedom from distorsion and elimination of after-weld machining. In the P.W.R. and B.W.R. reactors, the E.B. process can be employed for the manufacture of core barrels, heat shields and miscellaneous supports.

For H.W.R. (heavy water reactors), joining tubes to plates for end-shields is a typical application of the E.B. process. In this example, substantial savings can be made on machining : there is no need for reworking the bores of the channels and the pitch between tubes.

In F.B.R. (fast breeder reactors), the process is interesting for various structures of the core, especially for the fuel assembly support.

3.2. Steam turbine diaphragms

This process is already employed for the manufacture of these parts and for welding turbine wheels to the rim and the hub in 2 opposed passes with great speed .

The 100 kW gun will make it possible to improve the technique employed by allowing welding in one single pass of diaphragms with a maximum thickness of 200 mm and a reserve of power ensuring continuous mass production.

3.3. Manufacture and connection of pumps for reactor primary circuits

The pumps for reactor primary circuits are composed of heavy molded stainless steel elements whose thicknesses range between 200 and 300 mm. The use of the 100 kW gun with all its capabilities will permit important time reductions and less consumption of filler metals.

3.4. Welding steam-turbine rotor disks

In the next few years, if we take into account the probable increase of turbo-alternator power for the nuclear power industry, a "welded rotor" solution appears to be inevitable. In this case, the availability of a 100 kW gun and the possibility of welding 200 to 250 mm gauges in one single pass represent the first stage of an evolution which should end in the joining of the .rotors manufactured from forged discs.

The consequence of this manufacturing method will be a more flexible production cycle permitted by the elaboration and machining of several components separately.

3.5. Steam-generator shells

The thickness of steam generator shells ranges from 90 to 150 mm. These semi-circular casings can be welded with a local vacuum chamber along the edges (figure 12) and an inside E.B. equipment for butt-welding of the shells (figure 13).

To ensure a production of 18 steam-generators per year, we have noted that the use of the electron beam involves a significant saving (50%) with respect to conventional methods / 7 / and the overall investment for the EB welding equipment is amortized within 2 years.

3.6. Hydraulic turbine components

Several components of hydraulic turbines can be assembled in one pass with the 100 kW gun while reducing machining rework which is common practise for these parts. The problem of vacuum chamber configuration still remains to be solved as a function of workpiece geometry.

3.7. Pressure vessels

Many applications can be found in the field of heavy sheet metal work for pressure vessels thanks to the 100 kW gun. These applications will be preceeded by studies and metallurgical and mechanical checks.

REMARKS

Despite the significant advantages offered by electron beam welding in the field of heavy sheet metal work and mechanical engineering, the major users still hesitate to replace their conventional assembly methods (which they recognize as being expensive and time-consuming but well-proven) by a new welding technique.

There is thus an obstacle fabricators still hesitate to clear before obtaining complete confidence of practical applications on an industrial scale. The existing manufacturing codes have been elaborated for conventional welding processes. It is in this direction that important efforts shall have to be made with a view to adapting the manufacturing codes and procedures to the specific nature of electron beam welding.

CONCLUSION

The utilization of theoretical methods has made unnecessary the "cut and try" procedure employed formerly for defining gun components. The tests conducted on the gun show the limits of validity of the methods applied.

The first welding tests have revealed that 150 mm thick steel gauges can easily be welded using half the actual available power of this gun and that 200 to 300 mm steel gauges can be welded while keeping a sufficient power reserve so that the equipment is not operated at its maximum limit.

Finally, several potential applications of high-power electron beams in heavy industry are given.

BIBLIOGRAPHY

/ 1 / P. FRANQUIN and A. LANGLOIS : Mémoire de fin d'études ENSAM 1973/1974

/ 2 / G. SAYEGH, P. DUMONTE and NAKAMURA : 2nd Symposium on Advanced Welding Technology, Osaka - August 1975

/ 3 / P. GRIVET : Electron optics Pergamon Press - 1965

/ 4 / SCIAKY publication : Soudage par faisceau d'électrons - 1972

/ 5 / P. DUMONTE and G. SAYEGH : 6th electron and ion beam science and technique - 1974

/ 6 / G. SAYEGH and P. DUMONTE : Structural design and fabrication in shipbuilding London - nov. 1975

/ 7 / G. SAYEGH and P. DUMONTE : I.E.E. international conf. on metalworking London - nov. 1975

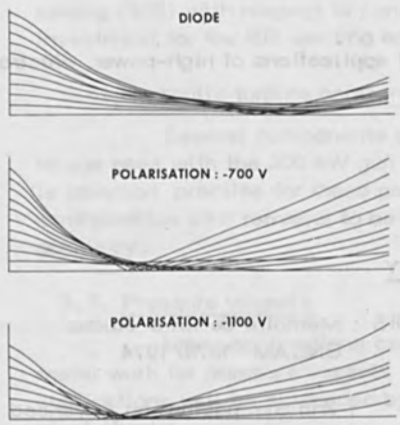


Fig. 1. Effect of bias on electron trajectories.

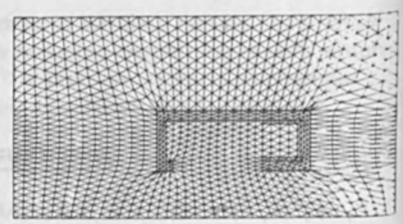


Fig. 2. Finite elements used in computing A.

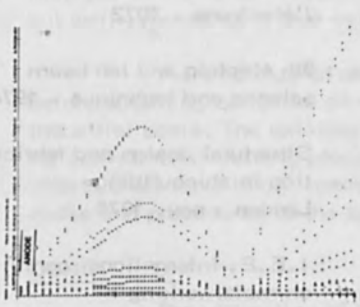


Fig. 3. Trajectories in the electromagnetic part of the gun (without space charge).

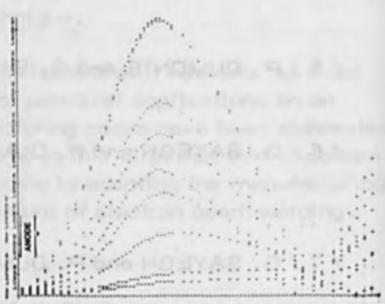


Fig. 4. Trajectories in the electromagnetic part of the gun (with space charge).

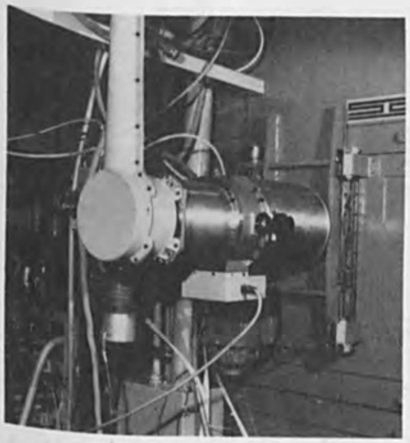


Fig. 5. 100 kW EB gun in horizontal position.

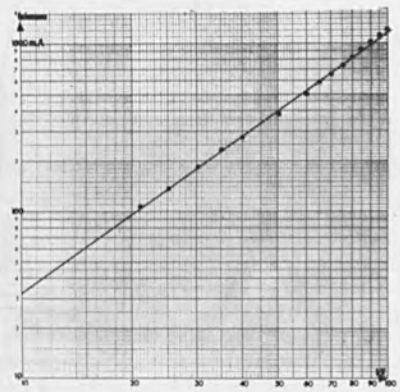


Fig. 6. Comparison of theory and experimentsl results. —— theory, • experimentsl results.

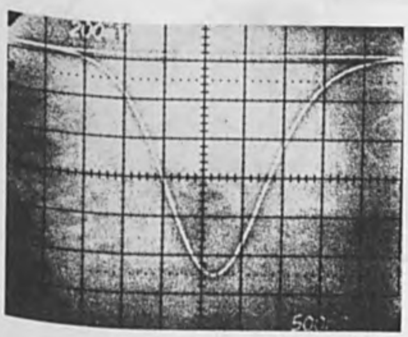


Fig. 7. Power distribution in the beam at 60 kV - 32 kW

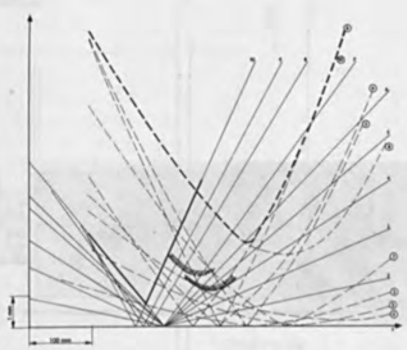


Fig. 8. Comparison of beam diameters at focus point. —— without space charge, ---- with space charge, \\\\\\ actual position of the beam, ///// partial neutralization of space charge.

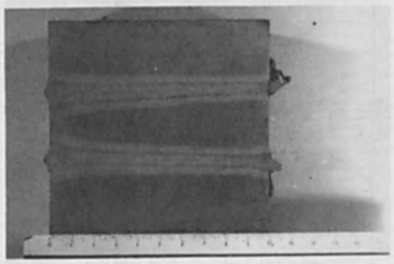


Fig. 9. Weld in 100 mm thick steel. V = 70 kV, I = 500 mA, v = 12 cm/mn.

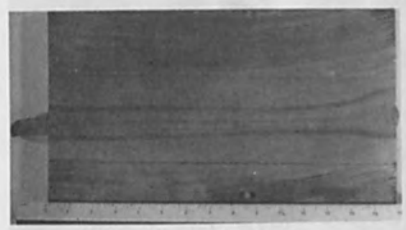


Fig. 10. Wel in 150 mm thick steel. V = 85 kV, I = 750 mA, v = 12 cm/mn.

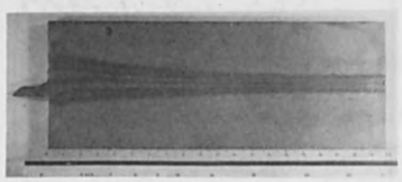


Fig. 11. Weld in 200 mm thick steel. V = 85 kV, I = 820 mA, v = 10 cm/mn.

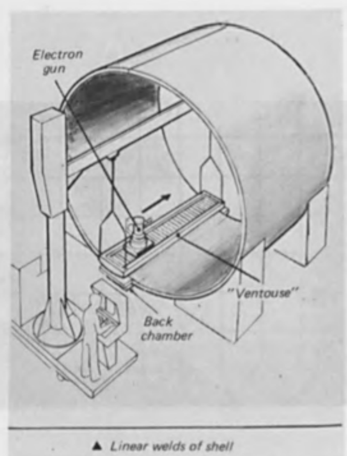
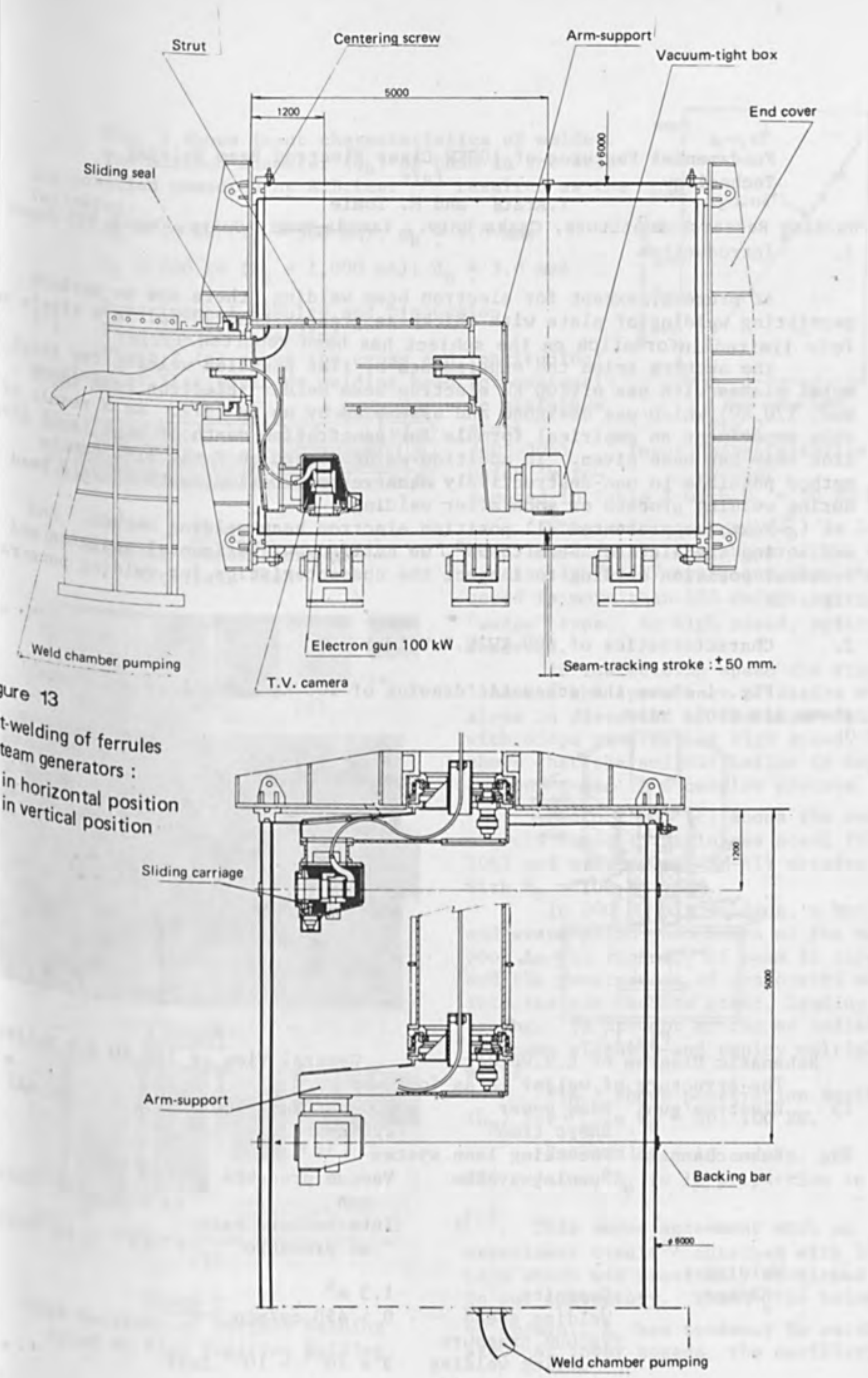


Fig. 12. "Ventouse equipment welding semi-circular compounds.



gure 13
tt-welding of ferrules
steam generators :
- in horizontal position
- in vertical position

Fundamental Features of 100KW Class Electron Beam Welding Technology

Y.Arata and M. Tomie
Welding Research Institute, Osaka Univ., Yamada-Kami, Suita, Osaka 565 Jap

1. Introduction

At present, except for electron beam welding, there are no methods permitting welding of plate with thickness greater than 10 cm. in a singl Only limited information on the subject has been reported. 1)2)6)

The authors tried the experiments of flat position welding for thic metal plates with use of 100 KW electron beam welder (electron beam input mas. 120 KW) which was designed and assembled by us in 1972. As a result this experiment an empirical formula for penetration depth of high input tron beam has been given. In addition we developed an X-ray slit camera method possible to non-destructively observe penetration depth of weld be during welding process or soon after welding.

Next, we presented all position electron beam welding method. And deflecting the electron beam to 90°, we carried out horizontal position a vertical position welding to inspect the characteristics for welding pene tion.

2. Characteristics of 100 KW E.B.Welder

Fig. 1 shows the schematic drawing of 100 KW E.B.welder and Photo shows its whole view.

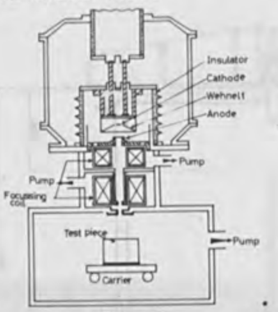


Fig. I
Schematic Diagram of E.B.Welder

Photo 1
General View of 100 KW E.B.Weld

The structure of welder is as follows:

1) Electron gun:	Beam power	100 KW (V_b = 100 KV, I_b = 1,000 m	
	Short time	120 KW	
2) Beam channel:	Focussing lens system		
	Pumping system	Vacuum pressure of-gun	: $10^{-6} \sim 10^{-4}$ T
		Intermediate vacuum pressure	: $10^{-4} \sim 10^{-3}$ T
3) Welding chamber :	Capacity	1.3 m^3	
	Welding speed	0 ∿ 450 cm/min.	
	Vacuum pressure during welding	$5 \times 10^{-4} \sim 10^{-2}$ Torr	

Fig. 2 shows input characteristics of welder.

The minimum diameter (d_b) of beam in the welding position measured by A.B.test[3)4)] resulted as the following:

W_b = 50 KW (I_b = 500 mA); $d_b \doteq 1.8$ mmϕ

W_b = 100 KW (I_b = 1,000 mA); $d_b \doteq 3.0$ mmϕ

3. Experimental Results and Discussion

3-1 Flat position Welding

Photo 2 (a) shows the cross and longitudinal sections of flat position welding bead for austenitic stainless steel (SUS-304) which is indicated in Table I with W_b = 50 KW beam.

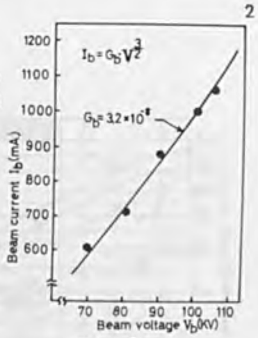


Fig.2

Input Characteristics of Welder

Wt(%) ← → PPm

	C	Si	Mn	Cu	Cr	Ni	Mo	P	S	Al	N	O
SM41	018	0.47	0.71	—	—	—	—	0015	0.010	0.050	68	36
SUS304	005	0.74	174	0.12	19.5	10.9	0.16	0030	0010	0015	365	70
HT80	011	0.32	092	026	046	1.03	0.47	0008	0006	—	34	10

Table I

Chemical Compositions of Materials Used

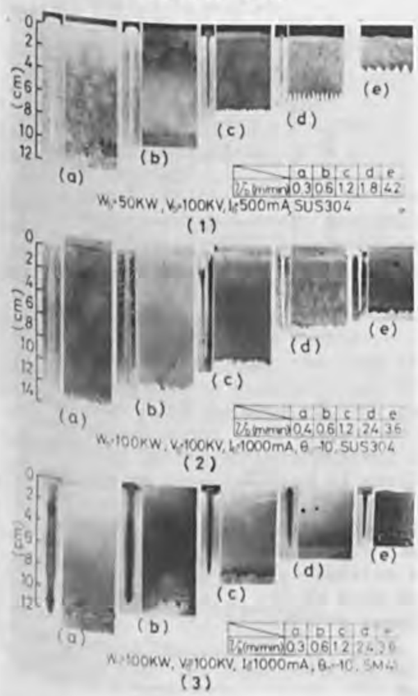


Photo 2

Bead Sections of Various Welding Speed by Flat Position Welding

When welding speed (v_b) is less than 120 cm/min., the shape of the cross section is "well" type, and when the speed is more than 120 cm/min. it is "wedge" type. At high speed, spiking is observed.

At low welding speed the ripple line has the appearance of stairs with slope in direction of the beam motion with slope reversal at high speed. This shows that the solidification is dependent on v_b and is a complex process.

Photo 2 (b) (c) shows the section of weld beads in stainless steel (SUS-304) and mild steel (SM-41) obtained with W_b = 100 KW beam.

In 100 KW E.B.welding, a boiling and evaporation phenomenon of the molten pool in the vicinity of beam is intensive and the penetration of evaporated metal into the gun becomes great, leading to arcing. To prevent arcing we deflect the beam slightly and employ multiple slits.

Fig.3 shows penetration depth (h_p) for v_b in W_b = 50, 100 KW.

Between v_b = 40 cm/min. and 200 cm/min., h_p is in proportion to $v^{-\frac{1}{2}}$. This shows agreement with an experiment result[5)] obtained with 30 KW beam which was previously developed in our laboratory. When v_b is below 40 cm/min., h_p has tendency to saturate. At lower speeds, the oscillation

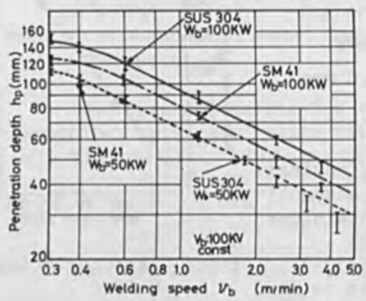


Fig.3

Effect of $_b$ and h_p

amplitude of the molten pool becomes larger, which interrupts the deep intrusion of the beam into the specimen. Under such conditions the beam energy is consumed to superheat the molten metal. Figs. 4, 5 and 6 show the relations between area of melting zone (S_B) of bead for $\sqrt{}_b$ of SUS-304, effective bead width ($\tilde{d}_B$) and effective penetro-parameters ($\tilde{P}_p$) = $h_p/\tilde{d}_B$. The maximum value of $\tilde{P}_p$ is 24. When the value is so high the features of beam welding with less thermal deformation appears.

Furthermore, small welding porosity and good welding results are apparent at high values and in the range of the curve with small rate variation.

As a result of examining the experiment data, $h_p \propto (\sqrt{}_b \times \tilde{d}_B)^{-0.83}$ is obtained.. With this formula we derived the following formula of penetration depth.

$$h_p = 0.32 \left(\frac{Q}{K\cdot Tm}\right)\left(\frac{k_D}{\sqrt{}_b\cdot \tilde{d}_B}\right)^{0.83} \ldots (1)$$

Note: Q : Heat input (Cal/s)
Tm: Melting point (^0C)
K : Thermal conductivity (Cal/s cm^{0}C)
K_D: Temperature conductivity (or thermal diffusivity)

Fig.7 shows the relation of the empirical formula and actual values.

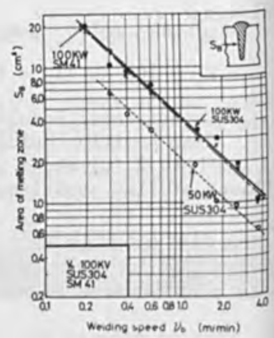


Fig.4

Effect of $\sqrt{}_b$ and S_B

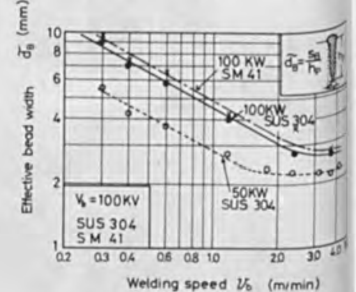


Fig. 5

Effect of $\sqrt{}_b$ and $\tilde{d}_B$

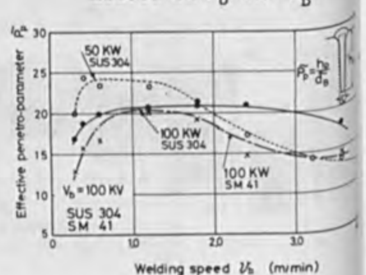


Fig.6

Effect of $\sqrt{}_b$ and $\tilde{P}_p$

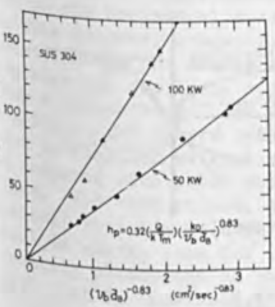


Fig.7
Relation between $(\mathcal{V}_b \cdot \tilde{d}_B)$ and h_p

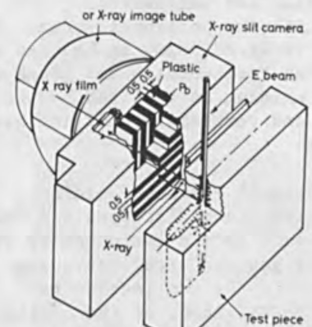


Fig.8
Schemtaic Diagram of
X-Ray Slit Camera

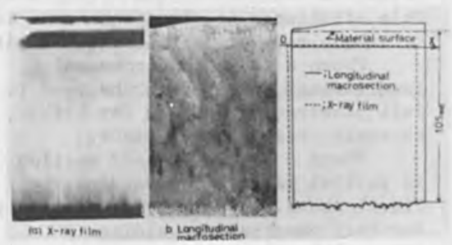
In usual measurements of penetration depth of more than 10 cm, it is necessary to cut the bead. This is associated with a number of problems.

To cope with the problem, we developed an X-Ray slit camera system capable of non-destructively measuring of the penetrating depth of welding bead during the welding process or soon thereafter. Fig.8 shows a schematic drawing of the measuring method. Photo 3 shows the results of the measurement.

3-2 All Position Welding Method

As you know well, in the case of high output electron beam welding, arcing is liable to occur because of high rates of evaporation vapor from the molten pool. Molten metal is heated with violent stirring and can on occasion lead to shallow penetration.

As a solution for these problems, we have developed an all-position electron beam welding process. We also succeeded in completing a 100 KW beam deflector as shown in Fig.9, which is capable of deflecting the beam at any angle. We propose that the beam position welding include the following: first, the beam in the basic direction consist of "vertical beam" which is along the direction toward the earth and "horizontal beam" which meets at any angles to these beams.



W_b= 50 KW, I_b= 500 mA,
$\mathcal{V}_b$= 30 cm/min, SUS-304

Photo 3
Examples of Penetration
Photograph by X-Ray Slit Camera

Second, it is possible to perform flat position welding with the vertical position welding and moreover ro welding with the horizontal beam. Furthermore inclined beam can be used as well as the horizontal beam. Such all position welding is realized by the rotating and vertically shifting beam Deflector.

3-3 Horizontal Position Welding

In high input flat position welding, it leads to a great quantity of evaporated gas and metal intruding directly into the gun leading to extensive oscillation of the molten metal and a shallow h_p. Having solved these problems by developing all position welding with beam deflection of 90^o.

Fig. 10 shows a schematic drawing of the welder set with the new beam deflector which was produced experimentally.

Even if deflecting the beam to 90^o and making horizontal position welding, we could actually observe that the beam current more than 95% was deflectively collected and reached the work piece. This efficiency is almost the same as it is the case of flat position welding.

Photo 4 shows the cross and longitudinal sections of the bead in full penetration welding for high strength steel (HT-80) plate.

Photo 5 shows the bead section in partial penetration welding.
Fig.11 shows the beam power necessary for full penetration welding. (Uranami bead).

Partial penetration depth is nearly the same as that of flat position, but full penetration depth can be increased about 20% in W_b= 50 KW. This comes from the facts that in the case of full penetration welding the molten metal flows out of the both sides of the work which results in the decreases of molten metal reheating.

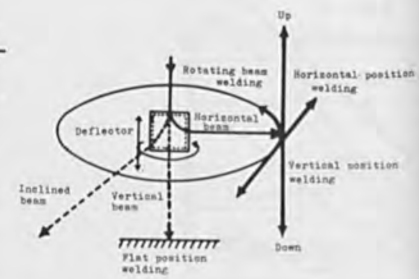


Fig.9

All Position Electron Beam Welding Method

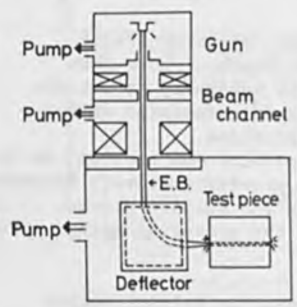


Fig.10
Schematic Diagram of Welder Set with Beam Deflector

3-4 Vertical Position Welding

The vertical position welding is made by use of a horizontal beam, similar to the horizontal position welding. In this respect, precautions must be taken against the arcing of an electron beam gun for large beam output. From a viewpoint of beam hole stability, the upward vertical position welding can be kept most stable, as compared with other welding positions. For that reason, the upward vertical position welding sees so small an energy loss due to the reheating phenomenon of the molten metal that h_p increases.

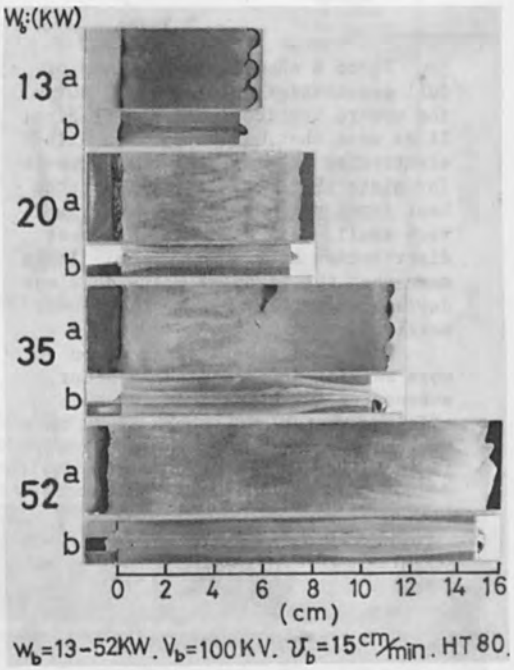


Photo 4
Full Penetrated Sections of Various Beam Power by Horizontal Position Welding

(a) Longitudinal Cross Section

(b) Lateral Cross Section

Photo 5
Bead Section of Partial Penetration Welding

Photo 6 shows bead sections of full penetration welding of HT 80 in the upward vertical position welding. It is seen that, in comparison with electroslag welding, the beam breadth for plate thickness is small and the heat input per unit weld length is very small. Fig. 12 shows hardness distribution at a welded zone. It is seen that the hardness value does not deviate much from that of the parent metal.

With increase in penetration more welding defects tend to occur, successive investigations for such thicker material in excess of 20 cm must be carried out if the problem is to be resolved.

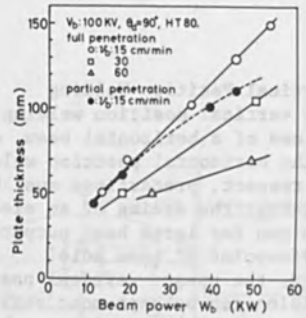
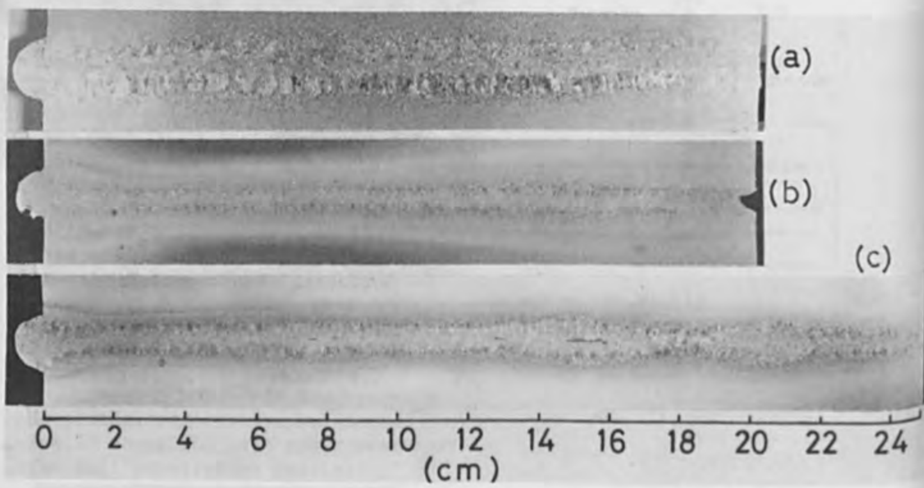


Fig.11

Beam Power Necessary to Full Penetration Welding



(a) W_b = 40 KW, v_b = 3.5 cm/min., HT 80

(b) W_b = 40 KW, v_b = 5.0 cm/min., HT 80

(c) W_b = 50 KW, v_b = 5.0 cm/min., HT 80

Photo 6

Full Penetrated Cross Sections of Upward Vertical Position Welding

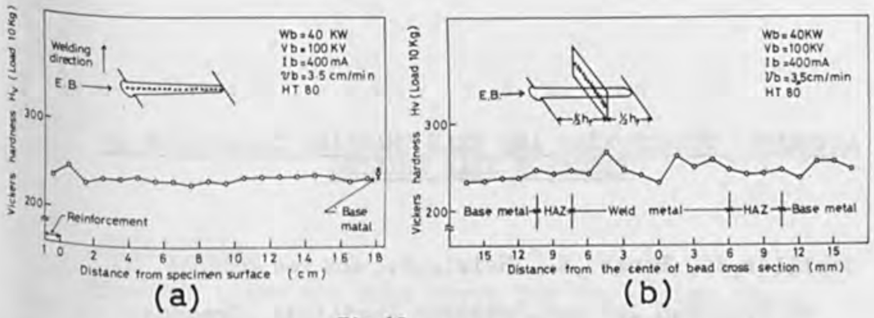


Fig.12
Hardness Curves at Upward Vertical Position Welding

4. Conclusion

1) Flat and horizontal position welding for steel plate thicker than 10 cm could be made in one pass by electron beam of W_b = 50~100 KW.
2) Empirical formula for the penetration depth for SUS-304 in 100 KW class electron beam welding was obtained as shown in Equation (1).
3) An X-Ray slit camera which is able to directly observe penetration depth during the welding process or soon after welding was developed.
4) High efficiency horizontal position welding is possible by deflecting the high energy density beam of the 100 KW class to 90° with the use of the newly developed deflector, without placing the electron beam gun to sideways. This result suggests that all position E.B.welding is possible by deflecting the beam of a fixed gun.
5) h_p in flat position welding attains to 11.5 cm with W_b = 50 KW, υ_b = 15 cm/min., and in the case of horizontal position welding it reaches 14 cm which is 20% higher than that of the flat position.
6) The upward vertical position welding at W_b = 50 KW gives h_p = 25 cm, which is about twice the value obtained in the case of flat position welding.

References

1) Y. Arata, M.Tomie: "100 KW Class Electron Beam Welding Technology (Report 1)", Trans JWRI Vol. 2, No. 1 (1973).
2) A. Sanderson: "A 75 KW Electron-Beam Installation for Thick-Section Welding," Metal Construction B.W.J., Vol. 6, No. 1 (1974).
3) Y. Arata: "Characteristics of Electron Beam Heat Source and View of Development on Electron Beam Welding Technology," J.Japan Welding Society, Vol. 41, No. 11 (1972).
4) Y. Arata, M. Tomie, K. Terai, H. Nagai and T. Hattori: "Shape Decision of High Energy Density Beam," Trans JWRI, Vol. 2, No. 1 (1973).
5) Y. Arata, M. Tomie, Y.Kato: "Some Properties of 30 KW Class Electron Beam for Welding," Trans JWRI, Vol. 2, No. 1 (1973).
6) G.Sayegh, P.Dumonte and T. Nakamura: "Design and Manufacture of a 100 KE Electron Gun," Advanced Welding Technology, The Second International Symposium of J.W.S., August (1975).

AUTOMATIC POSITIONING AND SEAM TRACKING CAPABILITY OF ELECTRON BEAM WELDING

Schiller, S., Panzer, S., Heisig, U., and von Ardenne, Th.

Manfred von Ardenne Research Institute, Dresden
German Democratic Republic

Abstract

It is especially electron beam welding with narrow and deep penetration where the beam has to be positioned relative to the weld-joint within tolerances of 0.1 mm or better. To this end, a system of complex process control was developed on the basis of backscattered electron emission. Signal generation is by means of duant electron collectors and special electronic signal processing facilities.

Signals obtained by EB scanning of the workpiece surface are used (a) to display an image to the weldjoint on the screen of a monitor, (b) for focus control and (c) for cold positioning; i.e. beam/joint alignment prior to welding.

EB seam tracking at welding power permits beam/joint alignment during welding. If positioning takes place by way of 'trend correction', a possible faulty measurement due to disturbances on the surface causes no misalignment. Constant positioning errors due to, say, vertical deviations at the weldjoint can be eliminated with the aid of additional signal evaluation.

The methods described in this paper turned out to be reliable in practice.

1. Introduction

Electron beams not only serve for the energy transfer in production processes but are also increasingly used for the simultaneous control of such processes. In particular, it is the utilization of backscattered electrons emitted at the point of beam incidence which plays a decisive role. This phenomenon is utilized especially for surface inspection and/or display as well as for beam action coordinate control. One outstanding example is the automatic matching of the beam deflection coordinate system to a given workpiece contour as found in thermal and nonthermal EB processing [1] [2]. Positioning accuracies of some µm are achieved in the case of thermal processing whilst those of nonthermal processes are in the order of some nm. In either case, however, primary contour detection by EB scanning takes place with a precision that corresponds to a fraction of the focal spot diameter.

It is known that EB welding calls for beam positioning relative to the weldjoint within tolerances down to approximately ± 0.1 mm. As a rule, attainable positional stability of the beam, work preparation and guidance during welding permit it to stay within the specified tolerances. Nevertheless, however, there are a great many cases where - in the interest of reduced production outlay for workpieces and less stringent demands on long-term beam position stability - it seems to be feasible to utilize the secondary actions of the beam also for process control and regulation. This applies particularly to cases where widely differing welding jobs have to be dealt with or where narrow but deep joints are to be made on expensive workpieces with a very high degree of reliability. But even if the mentioned tolerances are properly met, faults are likely to occur due to residual magnetism in either workpiece or fixture and, at least in part, through thermal distortion. So it may happen that the given tolerances for beam/joint alignment are exceeded. Shown in Figure 1 is the location of beam impingement on a workpiece along a seam length of 1 m. Even in cases where an obvious misalignment of the workpiece inclusive of a possible angular displacement between weldjoint center line and gun travel is avoided, there will be a residual positioning error between weldjoint and point of beam incidence around ± 0.2 mm.

In the following, the comprehensive utilization of electron scattering for displaying an image to the surface shall be explained for several variants of beam/joint positioning and focusing alignment.

2. Signal Generation

Based on known relationships, the emission of backscattered electrons - in particular with respect to their directional distribution - is directly correlated with the direction of the surface specimen block at the point of beam incidence relative to the incident direction. A very useful surface mark is the weldjoint which - to exert a direct effect on the emission of backscattered electrons - may be of special shape. Figure 2 shows some of the possible joint configurations.

When the direction of the joint is fixed in space - i.e. with either straight longitudinal seams, roundabout seams or circumferential edge welds - a joint configuration as per Figure 2b proves to be of advantage. An array of backscattered electron collectors adapted to such a joint geometry is depicted in Figure 3. The duants - that is the collectors 1 and 2 in Figure 3 - are arranged near the welding site symmetric to the non-deflected beam and to the direction of the weldjoint. They comprise one collector each and are shielded by a metallic screen. The part pointing toward the workpiece consists of a wire screen. Via variable load resistors the collectors are completed to ground and the screens have been put on a negative potential of approximately -50 V to keep thermo-electrons and secondary electrons away from the collectors. Adequate electronic coupling between collectors and associated screens noticeably reduces the capacitance between collectors and earthed environment. This allows a resolution of collector signals to frequency components of up to 5 MHz or better. The collector assembly is mounted on the exit end of the electron gun and directly connected to the load resistors and the follow-up impedance converters.

The beam has to be deflected vertically over the joint. In case it impinges on the workpiece completely outside of the joint, both collector currents are almost equal and the difference signal becomes zero. This holds true as long as the distance between point of beam incidence and plane of symmetry of the duant array is short against the distance between point of beam incidence and collector, provided that the angle between beam axis and surface specimen block on the point of incidence is small, too. But if the beam is going to hit the weldjoint either partially or in full, the collector difference currents will normally differ from zero. It is only precisely centered full admission of the joint where, for reasons of symmetry, the difference current signal also approaches zero. In Figure 3 the typical profile of the collector difference signal has been plotted versus beam position on the workpiece relative to the notch of the weldjoint. With given joint configuration, this signal profile always is of similar shape provided that the beam diameter on the workpiece surface is in the same order of magnitude as or smaller than the notch width b_k. With increasing values of $b_k/d_F > 1$ the signal amplitude reduces and, with $b_k/d_F \gg 1$, eventually becomes zero.

The characteristic zero passage of the difference signal obtained with a beam position in the joint center therefore is independent of beam current, beam diameter, notch width and working distance. On the other hand, however, these quantities effect the height of the signal amplitude. To avoid any further aggravation of signal evaluation by the said dependencies, the signal amplitude is normalized. This is accomplished automatically by electronic setting the load resistors of the collectors in dependence of the collector currents obtained with beam positions fully outside of the joint.

3. Using the Backscattered Electron Signals

3.1. Display of Surface Images

By EB scanning the workpiece surface, and transfer of the processed collector signals to the intensity modulation of a synchronized monitor, an image to the weldjoint can be displayed. Figure 4 gives the image of a notched joint. The distinct white-to-black transition in the weldjoint image corresponds to the joint position. Depicted in Figure 5 is the scanning electron image of a welding bead having a width of 3 mm. In addition, machining marks and weld splatter can also be seen.

Practical experience in welding engineering has shown that such a surface image gives all visual information needed for process control. In this respect it proves to be fully equivalent to light-optical viewing devices. As far as maintenance is concerned, scanning electron images will by far outdo light-optical or TV observation because even the thickest evaporated film on the collectors does not impair the definition of the image. Collectors need not be protected against evaporation during the welding phase even at higher welding powers. For many applications it proves to be very convenient that the site of observation is by choice. In addition, the image scale on the workpiece is also freely selectable via the dimensions of the scanning electron image.

The beam power to be transferred for scanning without fusion of the base material depends on (a) thermal properties of the material, (b) image size and (c) feed motion of the weldment. With steel the empirically determined power limit amounts to

$$P_{max} \approx 800 \cdot l^{3/2} \cdot v^{1/2}$$

for a square image, where

P_{max} = beam power in watts on which the workpiece surface is just prevented from fusing
l = lateral length of square scanning electron image in cm
v = feed motion of weldment in cm/s

The permissible beam power remains practically independent of the surface roughness. Customary beam power ratings for surface display with 5 to 20-fold image scale are in a span of 100 watts to some kilowatts.

In practice, a beam power around 1 kW will yield a resolution of less than 0.1 mm. Figure 6 gives the results of a resolution test conducted by displaying the image of wires with various diameters that have been stretched over a beam surface vertical to the direction of raster. Distinct images are obtained even with wire diameters that are as low as 0.06 mm.

3.2. Beam Focusing

The smaller the beam diameter at the point of incidence the better the resolution and the contrast of the image on the monitor. This gives a plain yet dependable possibility for obtaining a defined axial alignment of the focal spot on a level with the workpiece surface. This method turned out to be very useful for focusing on the workpiece as, referred to the image width, it yields an accuracy of ≤ |1 %| in industrial routine work and ≤|0.5 %| in case of very considerate alignment. These values correspond to almost 15 to 30 % of the half-length of a cylindrical high power density 'core' in the proximity of the focal spot.

In practice, this method proved to be a success particularly for welding small-lot workpieces of various shapes at different cycle times.

3.3. Beam/Joint Alignment

3.3.1. Cold Positioning

From the Figures 4 and 5 it is evident that, within the scanning electron image on the monitor, cross hairs are faded in for the deflection currents $I_{Ax} = 0$ and $I_{Ay} = 0$. Adequate design of the deflection unit ensures that eddy currents in the region of the deflection field are entirely avoided so that the zero passages of the deflection currents coincide with those of beam deflection. Hence the intersection point of the cross hairs defines that point on the workpiece surface that will be hit by the non-deflected beam during welding. With the aid of the image displayed on the monitor it is then possible to shift the workpiece in such a manner that the well defined joint center coincides with the intersection of the cross hairs. Apart from this mode of beam/joint alignment it is further possible to use a suitable d.c. deflection of the beam for positioning when the workpiece is fixed. It should be borne in mind, however, that this share in deflection must be eliminated during ascertainment of the cross hairs.

Instead of manual positioning via visual inspection of the image displayed on the monitor, the collector signal may also be used for automatic beam/joint alignment. This is best accomplished by generating one pulse each for the raster center line during the zero passages of both collector difference signal and deflection current, evaluating the said pulses in their timed sequence. Corresponding with the result of such an analysis, positional control then starts to takes place via a relative motion between beam and workpiece in the respective direction until both pulses coincide. In either case, however, the positioning accuracy is 0.1 mm or better.

After weldjoint center line and workpiece motion are well defined with respect to one point of the joint by means of the described beam alignment, welding can be started. But if such an unambiguous relationship does not exist, the weldjoint may also be scanned during positioning so as to store the obtained corrective values. During subsequent welding, relative motion then takes place according to stored information. Because of collector arrangement, signal generation for beam deflection and mode of positional alignment, however, the use of this method is restricted to continuous seams with only minor inclination angles between weldjoint center line and symmetry plane of the collectors. Hence, it is incapable of seam tracking with arbitrary course of the weld but is recommended for compensating the effects of manufacturing and positioning errors with longitudinal and circumferential seam welding.

With the discussed variants the beam power used for beam/joint alignment generally is much lower than welding power as such and therefore causes no thermal damage to weldments such as fusion of the workpiece surface etc. For this reason this method is termed 'cold positioning'.

Occasionally, surface scanning used for displaying the image to the weldjoint and for beam/joint alignment will also be employed to handle other tasks in production engineering. With weldments having joints that are not too thick it may be used to advantage for preheating the workpiece as well as for heat treatment after welding, just to cite some examples.

Within beam power limitations imposed by the thermal stressability of the workpiece surface, EB scanning can further be used for pulse weld processes where the gun operates in continuous mode. Starting from surface scanning, beam deflection then has to be interrupted during the weld pulses.

3.3.2. Hot Positioning

The point of beam incidence on the workpiece is independent of beam current and current flow through the focusing lens only with a properly centered beam path. Under manufacturing conditions, however, it is almost certain that, upon gating the beam current to welding power and during a change in focusing conditions, the focal spot on the workpiece may drift by some tenth of a millimeter. Moreover, it is likely that, during welding, the workpiece will be distorted by amounts that lie within the same order of magnitude. The result of both effects on beam/joint alignment can be determined and eliminated when positioning control takes place during the welding process at full beam power. Illustrated in Figure 7, this positioning principle was developed especially for seam tracking during longitudinal or circumferential seam welding. After a certain welding time the beam is guided cross-seam on a circular path in front of the point of weld. The direction of beam deviation from the joint is determined from the timed sequence of pulses that were also generated at the zero passages of both x-component of beam deflection and collector difference signal. Through workpiece displacement or beam deflection against the detected diversion by a fixed amount of, say 0.02 to 0.03 mm - i.e. an amount which is small against the permissible tolerance of beam/joint alignment - the positioning error is reduced. This is termed 'trend correction' and ensures that a possible faulty measurement causes no misalignment as normally found in case of 'true-to-amount' corrections by the determined full beam diversion. As a rule, faulty measurements are often due to disturbances on the surface; i.e. weld spatter in the scanned track and the like. Adequate selection of the scanning cycle repetition rate in a span of some Hz to some 10 Hz ensures that the realizable corrective amount per unit of seam length is always greater than the misalignment over the same unit of length.

Besides tracking of longitudinal and roundabout seams the scanning mode depicted in Figure 7 may further be used for circumferential edge welds of any desired diameter. In practice it has been shown that the previously described seam tracking facility permits edge welds of, say, 200 mm diameter without positioning error even if the untrueness is greater than ± 1 mm.

To avoid fusing of the workpiece surface along the scanning track the beam has to be moved over the workpiece fast enough. In coarse approximation the required scanning speed can be calculated via adiabatic heating [3]. For steel and accelerating voltages around 150 kV this yields a scanning speed of

$$v\ (m/s) \gtrsim 10 \cdot P\ (kW)$$

Here P denotes the beam power. For a welding power of up to 50 kW it is therefore necessary to use scanning speeds in the order of 500 m/s. With scanning circle diameters of, say, 15 mm this gives a scanning time of 100 µs. But the backscattered electron signal created during cross-seam motion amounts to only about 1 µs so that proper signal detection requires a frequency resolution of the collector assembly of $\gtrsim$ 5 MHz.

Because the described beam/joint alignment is made at full beam power during the welding process, it is termed 'hot positioning'.

3.4. Perturbing Effects and their Possible Elimination

The previously described positioning procedures may be affected by a variety of disturbances. Some of the most significant causes of trouble and their influence on the collector difference signal profile are depicted in Figure 8.

Jointing gaps proved to be non-critical with respect to the positioning process for as long as they remain small against the beam diameter on the plane of the workpiece surface. This also applies to hot positioning provided that welding takes place with a focal spot position well below the workpiece surface, and that the beam diameter on surface level is not much greater than the width of the joint marks. Larger cross-sectional areas to be welded and heavy subfocusing, however, require somewhat wider joint marks.

The effect of vertical deviations at the weldjoint depends on (a) 'visual angle' of both collectors, (b) point of beam incidence and (c) elevation difference. The apparent joint position indicated by the zero passage of the collector difference signals deviates from the actual position by an amount Δx. Time integral over the signal and resulting misalignment are related as shown in Figure 9; a capability that may serve as proximity control for continuous correction of elevation.

When pronounced weld spatter is likely to impair hot positioning, misalignments may be detected by evaluating the timed sequence of signal polarities. Should such an analysis reveal that the signal was triggered by a splash, then the resulting positional correction is suppressed. To restrict the frequency of such detrimental effects, signal detection can be confined to a narrow 'fixed error range' in the immediate vicinity of the joint. However, difficulties are liable to occur if workpiece geometry restricts the angle of divergence covered by the collectors. Such an example is given in Figure 10 which illustrates that signal evaluation, joint configuration and workpiece geometry have to be adequately matched.

4. Conclusions

Process control by backscattered electrons may extensively be used for objectivation and automation of EB welding processes. Scanning electron images of the workpiece surface are a profitable alternative to light-optical or TV observation. They are characterized by sturdiness and only minor maintenance requirements. In addition, they offer a definite criterion for focal setting on the workpiece surface. Further use of backscattered electron signals for cold positioning (i.e. beam/joint alignment prior to welding) as well as hot positioning (beam/joint alignment during welding) renders the demands with respect to work preparation and guidance during the welding process less stringent. This also applies to the long-term stability of EB generation. Possible workpiece and process dependent perturbing effects on positioning reliability caused, for instance, by improper jointing and weld spatter, can be eliminated to quite an extent by way of suitable electronic evaluation of the backscattered electron signals. The achieved high standards of development fully meet the requirements of up-to-date welding engineering. The described methods proved to be of advantage especially for welding extensive small-lot workpieces in the heavy engineering industry where high depth-to-width ratios of the seam are demanded and where reproducible welding conditions are a must. Figure 11 shows an actual plant where the described procedures have been put into being.

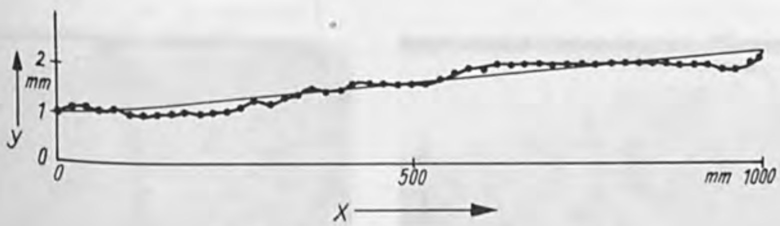


Fig. 1. Beam incidence deviations over a travelling path of 1000 mm relative to weldjoint center line as caused by workpiece misalignment and stray magnetic fields. Accelerating voltage U_B = 100 kV; Working distance z_A = 500 mn; Required joint position: x-axis.

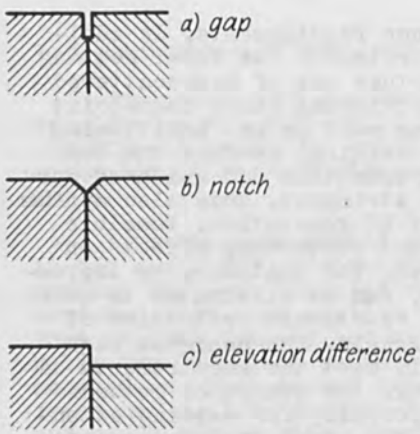


Fig. 2. Joint configurations suitable for weldjoint positioning by backscattered electrons illustrated by way of butt-weld joints.

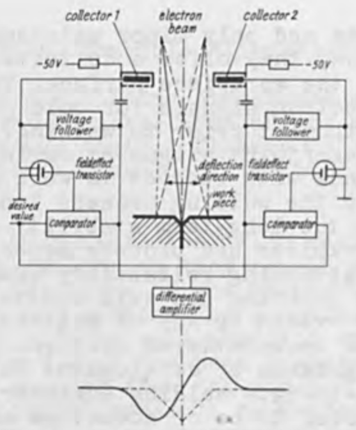


Fig. 3. Backscattered electron collectors in duant array relative to the workpiece; circuitry for signal generation and profile of the obtained difference signal (schematic); focal spot diameter about half the notch width.

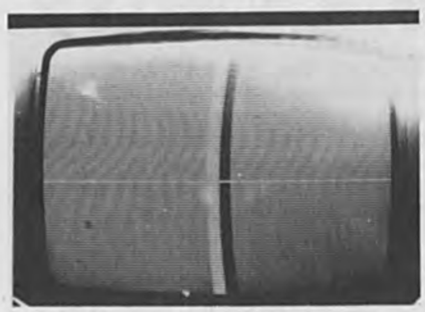


Fig. 4. Backscattered electron image to workpiece surface with positioned circular joint. The x-y cross hairs correspond to deflection conditions $I_{Ax} = 0$ and $I_{Ay} = 0$, respectively. Raster parameters: $U_B = 125$ kV; $I_B = 8$ mA; Scanning electron image: 16 mm x 12 mm. Notch width: 1.5 mm; Material: St 52.

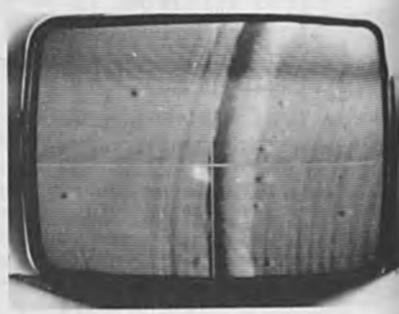


Fig. 5. Backscattered electron image to workpiece surface with welding bead, weld splatter and machining marks. Raster parameters: $U_B = 125$ kV; $I_B = 8$ mA; Scanning electron image: 16 mm x 12 mm. Bead width: 3 mm; Material: St 52.

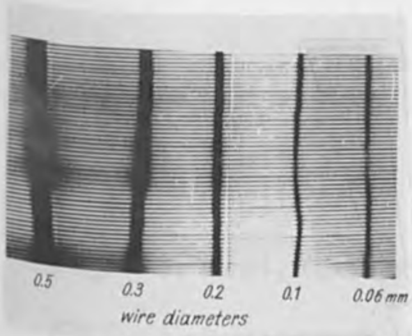


Fig. 6. Backscattered electron image to tungsten wires placed on a steel surface (resolution test). Raster parameters: U_B = 100 kV; I_B = 10 mA; Scanning electron image: 24 mm x 18 mm. Working distance: 600 mm.

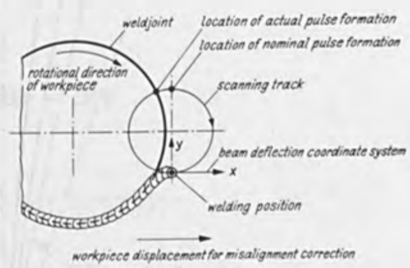


Fig. 7. Hot positioning with circular beam path for seam tracking during cicumferential seam welding (schematic).

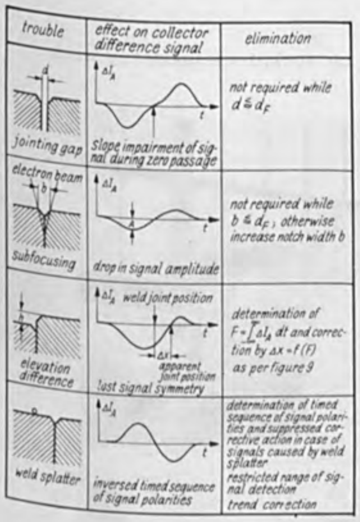


Fig. 8. Possible disturbances, their effects on collector signals and trouble elimination.

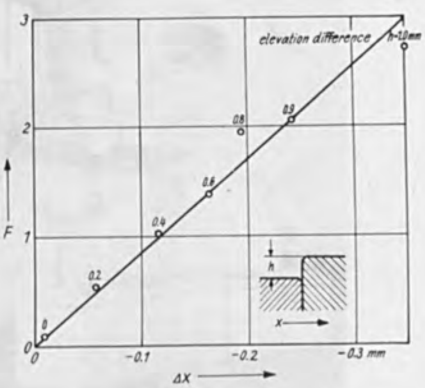


Fig. 9. Time integral of collector difference signal versus distance between apparent and actual joint position.

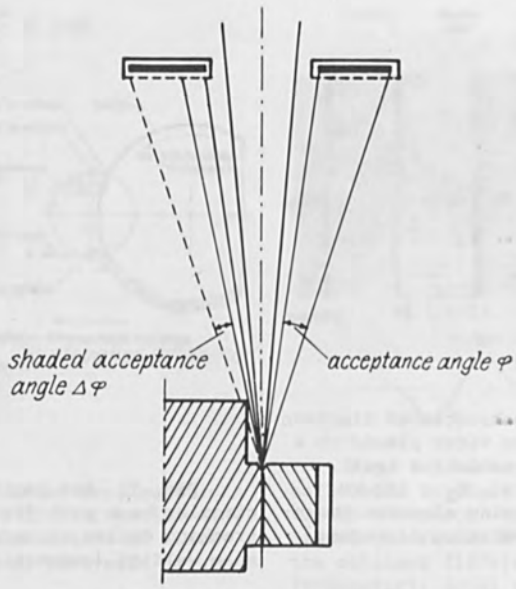


Fig. 10. Signal generation impaired by workpiece geometry with unmatched collector array.

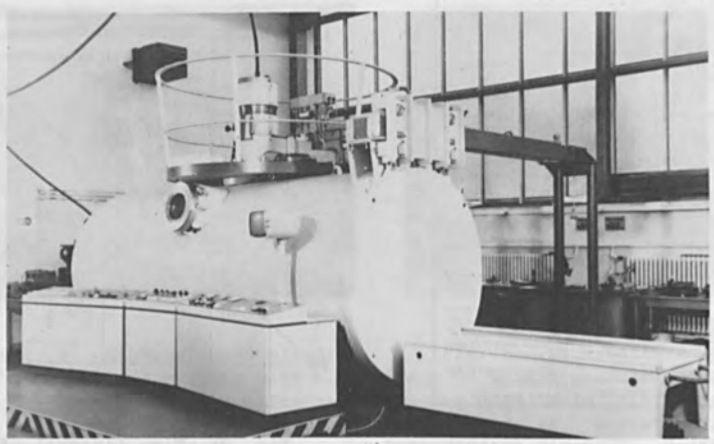


Fig. 11. EB welding plant for small-lot workpiece of the heavy industry with monitoring device, designed for cold and hot positioning. Acceleration voltage: 125 kV; Maximum beam power: 30 kW.

References

[1] S. Panzer, et al, DDR - DWP 87 874 (1971)

[2] B. Mundt, Optik 38, 351 (1973)

[3] S. Schiller, et al, Strahltechnik 7, DVS-Berichte, Düsseldorf 34, 41 (1975)

A NEW CONTROL METHOD OF ELECTRON BEAM FOCUSING IN THE EBW MACHINE

Cz. Kirczuk

INSTITUTE OF ELECTRON TECHNOLOGY,
TECHNICAL UNIVERSITY OF WROCŁAW,
Wybrzeże St. Wyspiańskiego 27
50-370 Wrocław, Poland

A new control method of electron beam focusing in EBW machines is presented. The idea of this method lies in making use of the alternate component of the current particle collector which is placed above the workpiece.

The fluctuations of the current of the particle collector are caused by the phenomena of dynamic character, which accompany the process of electron beam welding. It was found that there was a relation between the sound spectrum of the fluctuations and the conditions of the electron beam focusing.

A block diagram of the focusing control system in the low voltage EBW machine, as well as some results which confirm the usability of this method in practice are given. Attention was turned to a possibility of automatization of focusing by using the extreme regulation system.

I. INTRODUCTION

In order to use fully the possibilities of electron beam welding process, the optimal focusing conditions must be found. Under optimal conditions the deepest welds are obtained. There is a series of control methods for beam focusing in EBW machines. Some of them, such as visual estimation of the diameter of non-welding /low power/ beam, or estimation of the luminous intensity of ionized metal vapors are subjective methods. Their effectiveness depends on the skill and experience of the operator

The up to now known methods of objective control of the beam focusing have a series of disadvantages which make them difficult to use in practice, e.g.

- the control method based on phenomenon of mechanical vibrations of the sample [1] requires vibration damping of the workpiece manipulator;
- the method of the workpiece current measurement [2] requires insulation of the workpiece from the workpiece manipulator;

- the measuring method of the infra-red radiation [3] requires the use of special measuring indicator.

It seems that the most useful in practice may be methods based on measurements of current particles /electrons and ions/ emited by the workpiece [4], [5].

The method of focusing control presented below may be included to these methods. This method is in substance different from the ones described under [4], [5], because the alternate component of the current of the particle collector is used. The advantage of this method is a lack of sources polarizing the particle collector, and simplicity of the collector construction [6].

During electron beam welding the phenomenon of current fluctuations of the particles emited by the workpiece occur due to the interaction of the electron beam with plasma and other effects of dynamic character accompanying the mechanism of partial penetration electron beam welding [7], [8], [9], [10], [11].

Observations and measurements of current fluctuation of particles emitted by the workpiece indicated a possibility of using them for the purpose of focusing control.

The principle of the focusing control method in the low voltage EBW machine is illustrated in Fig.1. Max. output voltage from the f-U converter /frequency-voltage/ proves that the optimal focusing of the electron beam is obtained.

II. EXPERIMENT

Fluctuation of the collector current were observed by means of the oscilloscope on two low voltage EBW machines /P =2 kW, U =30 kV/ under different welding conditions for stainless steel 1H18N9T, copper M1G, aluminium PA1 and steel St5.

The observations indicated strong dependence of signal character on the I_L lens current at constant welding conditions. Under optimal focusing conditions max. amplitude and max. average signal frequency were observed. The observation revealed random character of the signal.

It was decided to make the measurements of the average signal frequency, since the signal amplitude depended on the beam current at constant focusing current. A diagram of the measuring system is shown in Fig.2.

A typical curve ofthe output voltage from the converter f-U U_{out} vs lens current I_L has a characteristic maximum /Fig.5/. The metalographic investigations confirmed the suggestion that maximum U_{out} corresponded to the deepest weld. The measurements of U_{out} vs I_L under various welding conditions indicated that with the increase in the beam power the lens current interval of optimal focusing extended /Figs. 6,7/. Similar shapes of the characteristic were observed when, instead of the converter f-U, the converter AC-DC was used.

The measurements of sound spectrum in the acoustic band were caried out during welding of stainless steel 1H18N9T copper M1G, aluminium PA1N and tungsten. The sound spectrum was measured in the system as in Fig.4 for three currents $I_L=I_{opt}$, $I_L=I_{opt}-$ 12.5 mA, $I_L=I_{opt}+$ 12.5mA. Relative deviations of the lens current from optimal lens current in all measurements did not exceed 5%. Time of sound spectrum measurement was 60 sec. One of obtained distributions is illustrated in Fig.3. As can be seen in Fig.3a the spectral distribution under optimal focusing conditions becomes near the continous distribution and get features of band noise. Under optimal conditions the increase in spectral width of the signal is observed, and effective voltage reaches the highest value. The fluctuation power of the collector current reaches maximum under these conditions. It may be supposed, that then intensity of the phenomenons occuring in the workpiece under electron beam bombardement is also maximum. Confirmation of this fact may be that at optimal focusing spiking is maximum. Under the conditions different than optimal the spectral distribution becomes more discreet. Distinct maxima of the distribution at the frequencies of 50, 100, 150, 4000, 8000 Hz are then observed. The frequencies of 1000, 4000, 8000 Hz prove the occurrence of pulsation of the beam power and are related to the work of tyrystor inverters in the high voltage supply unit, but the frequency of 50 Hz and harmonics come from the mains supplying the device.

III. CONCEPTION OF THE CONTROL METHOD OF ELECTRON BEAM FOCUSING

Because fluctuations of the collector current are of random character, they cannot be used directly for focusing control. For this purpose some determined functions should be used , dependent on focusing conditions, obtained by statistical treatment of the signal coming from the collector of the particles emitted by the workpiece $E\{\psi(u)\}$ where $\psi(u) \geqslant 0$ and $\psi(u) \neq$ const., such as:

- $E\{u^2(t)\}$ expected value of the signal voltage u^2;
- $E\{|u(t)|\}$ expected value of the voltage after full-rectification of the signal;
- $E\left\{\frac{|u(t)|+u(t)}{2}\right\}$ expected value of the voltage after half-wave rectification of the signal:
- $E\{\lambda(u_0)\}$ expected value of number of excedings by the signal the established level u_0 in a time unit.

For the presented focusing control system /Fig. 1./ the function $E\{\lambda(u_0)\}$ was used.

Because operation of the statistical averaging cannot be physically realized for estimation of the expected value of stochastic process $\{\psi[u(t)]\}$, averaging in the time must be assuming that it is stationary. A block diagram of the focusing control system, based on the above consideration, is shown in Fig.8. Of course, this diagram does not illustrate all possibilities of the conception building the focusing control system in which fluctuations of the current of the particles emitted by the workpiece, are used.

CONCLUSIONS

1. Alternate component of the collector current of the particles emitted by the workpiece may be used for control of beam focusing in EBW machines, because it includes the information about the weld depth.
2. Application of the converter f-U in the focusing control system allows to decrease the influence of the beam current on the value of the output signal.
3. Under optimal focusing conditions the fluctuations of the collector current of the particles get the features of band noise.
4. Because the maximum of the output voltage from the converter f-U corresponds to the optimal focusing conditions, the use of extreme regulator should automatically ensure these conditions during welding process.
5. It seems that other solutions of the focusing control system /general conception presented under III/ may be attractive considering their practical applications.

REFERENCES

1. Patent USA No 3542995
2. Patent England No 1259146
3. Patent France No 2095821
4. Patent West Germany No 2013950
5. Patent France No 2024057
6. A. Hałas, J. Felba, Cz. Kirczuk:"Kontrola ogniskowania wiązki w spawarce elektronowiązkowej", Raport Instytutu Technologii Elektronowej Politechniki Wrocławskiej, No 29, Wrocław 1974.
7. H. Tong, W.H. Giedth:"A Dynamic Interpretation of Electron Beam Welding", Welding Journal 49 p.259s-266s, June 1970.
8. Y. Arata, F. Matsuda and T. Murakami:"Some Dynamic Aspect of Weld Molten Metal in Electron Beam Welding", Transaction of JWRI 2 p.23-31, No 2, 1973.
9. T.Miyazaki, N. Taniguchi:" An Experimental Analysis of Penetration Mechanism of High Density Electron Beam", 5th International Conference on Electron and Ion Beam Science and Technology ed. by R. Bakish, 1972.

10. G.L.Mara,E.R.Funk, R.C. McMaster and P.E. Pence: "Penetration Mechanism of Electron Beam Welding and the Spiking Phenomenon",Welding Journal, 53 p.246s-251s, No 6, 1974.
11. C.M. Weber:"Direct Observations of the Penetration Mechanism of Partial Penetration Electron Beam Welding", 5th International Conference on Electron and Ion Beam Science and Technology ed. by R. Bakis 1972.

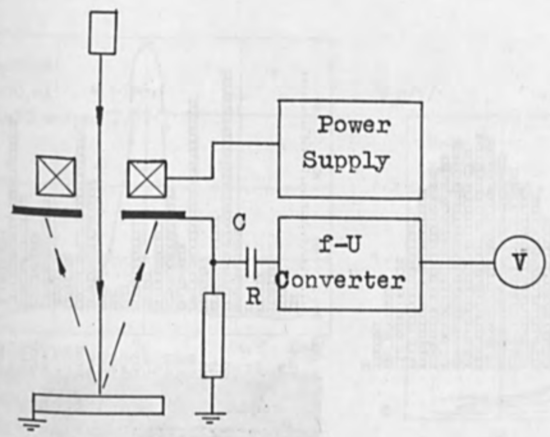


Fig.1.Block diagram of the focusing control system.

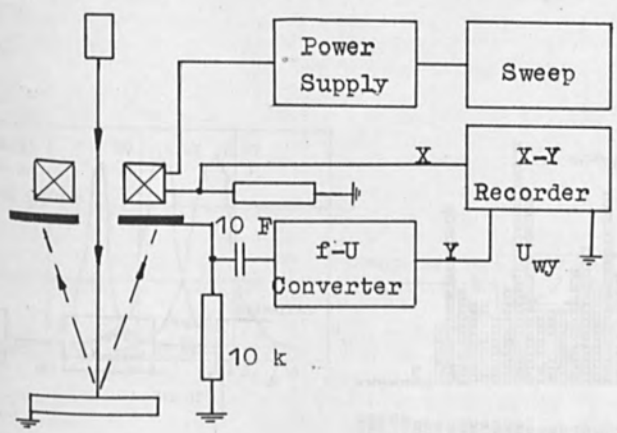
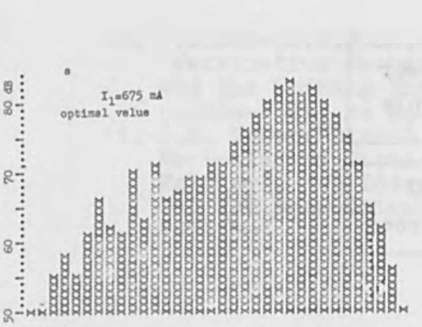


Fig.2 Diagram of the system for measurement of average frequency of the collector current fluctuations.



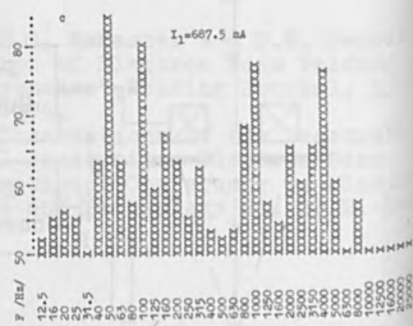
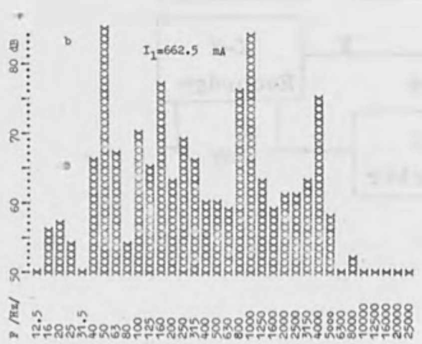


Fig. 3. Sound spectrum of the collector current during welding of stainless 1H18N9T. Welding parameters U = 15 kV, I = 30 mA, V = 170 cm/min. a - optimal focusing, b - $I_L = I_{opt}$ - 12.5 mA, c - I_L = I_{opt} = 12.5 mA.



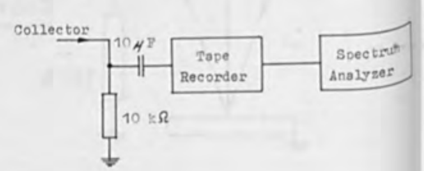


Fig. 4. Block diagram of the system measurement of sound spectrum of the collector current.

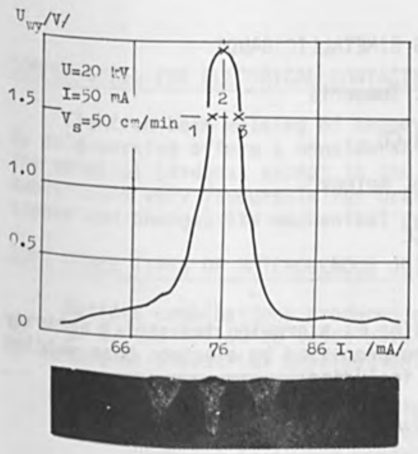


Fig. 5. Typical characteristic U_{out} vs I_L and cross-sections of the stainless steel 1H18N9T welds.

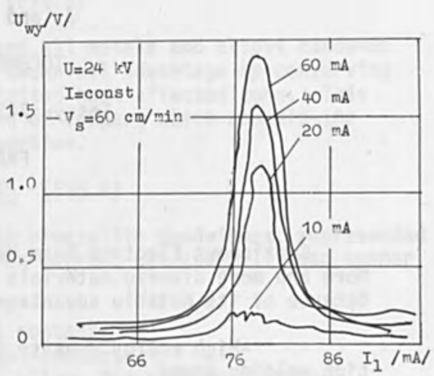


Fig. 6. Influence of beam current on the characteristic U_{out} vs I_L. workpiece - stainless steel 1H18N9T.

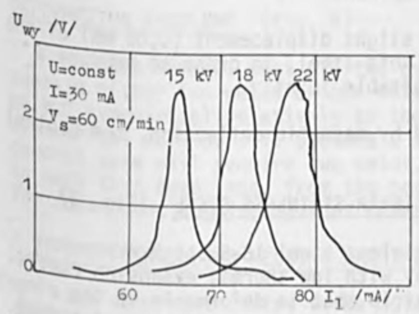


Fig. 7. Influence of the accelerating voltage on the characteristic U_{out} vs I_L. Workpiece - stainless steel 1H18N9T.

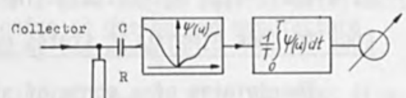


Fig. 8. Block diagram of the focusing control system.

WELDING OF CONTINUOUS BIMETALLIC BANDS

J. and M. Sommeria

TECHMETA CO.

Tessy- 74370- PRINGY

FRANCE

Continuous Electron Beam Welding is a growing industrial activity. More and more diverse materials are processed by electron beam welding because of its notable advantages including:

-High energy density (20 000 KW/cm^2) and consequential high welding speed,
-Great reduction of the width of the fused and heat affected zone,
-Low cost, due to the great speed.

<u>BAND SAW BLADES</u> (Fig. 1)

A well known application of continuous electron beam welding is the manufacture of bimetallic saws and only the electron beam can produce a sufficiently narrow heat affected zone to avoid destruction of the structure of the cutting component.

It is possible to maintain a slight displacement (0,05 mm) of the weld line in the firection of the soft-steel, in order to produce a ductile fused zone and obtain a suitable joint.

This displacement is obtained by magnetic deflection, the control is visual, with optical sighting.

<u>EDGE TO EDGE BIMETALLIC STRIPS OF INVAR-STAINLESS STEEL</u> (Fig. 2)

The joining of a strip of stainless steel 18-8, high expansion coefficient, to a strip of an alloy with low thermal expansion coefficiency such as Invar, produces a strip that is deformable in the direction of the plane when the temperature varies (thermostatic bimetal strip) able to transmit a considerable force due to its great moment of inertia.

Bimetallic pieces of suitable form are cut out from the continuous strip and employed as a semi-finished product.

Hence too, the electron beam permits a unique solution, due to the joint that it can produce between two alloys.

Effectively, it is this joint that will determine the quality of the bimetallic strip as factor of thermic deformation.

COPPER-STEEL FOR ELECTRICAL CONTACTS (Fig.3)

Electron beam welding of copper and all metals and alloys hardened by cold-working offers a considerable technical advantage by conserving the original hardness except in the limited heat affected zone. This substitutes very favourable for brazing of copper, which anneals the copper and changes its mechanical properties.

SOME OTHER TYPES OF HETEROGENOUS JOINTS (Fig.4)

Besides combinations produced with bimetallic sandwiches, represented by Fig. 4, the following tests have been carried out and will lead sooner or later, to industrial applications.

- -Steel-Silver, for electrical contact.
- -Copper-Silver, for electrical contact.
- -Magnetic alloy-Non magnetic alloy, for electrical relays.

MONO-METALLIC PROFILES (Fig.5)

The techniques for manufacturing of a great variety of profiles might be expensive or difficult, and particularly in the case of sections with different thicknesses. In many cases, the profiles can be broken down into two or more simple elements that can be fitted together inexpensively by electron beam welding. Fig. 5 shows the example of manufacturing ski edges out of light and hard alloy.

DEFORMATION PROBLEMS (Fig. 6)

We must remark, that all the examples of welding presented here contain either two welds or one weld and a melt called "equilibrating melt," symmetrical relatively to the medium line. In fact, all dissymmetrical welding will produce a sword shaped deformation. The most general case will require two welds, or a weld and an equilibrating melt. We have thus been lead, from the beginning to design and manufacture for continuous welding devices powered by two guns.

A TYPE OF CONTINUOUS WELDING MACHINE AIR-VACUUM-AIR, WITH TWO GUNS (Fig.7)

The component strips to be welded are stocked outside of the welding chamber. After cleaning they pass through a series of dynamic seals with intermediate pressure between the atmospherical pressure and the welding vacuum (5.10^{-2} Torrs). Thus, two medium levels between these two pressures are continuously pumped, by rotary pumps. The welded band will emerge into the atmosphere through a similar system of seals.

The welding chamber contains the device which guides and moves the strip (friction rolls).

The two electron guns are each provided with a sighting telescope and a magnetic deflection. The operator observes regularly the position of the points of impact on the joints and corrects manually the electrical deflection currents.

CONCLUSION

For a long time vacuum has been considered as an important obstacle for the development of continuous industrial and economic welding by electron beam.

Thus the much more expensive non-vacuum electron beam process equipment has been developed to overcome the limitations but this has not been without additional penalty in beam performance.

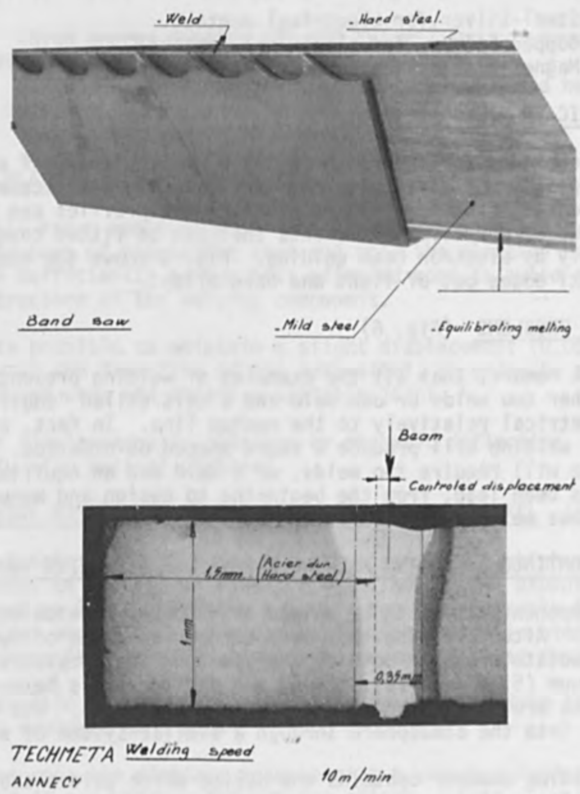
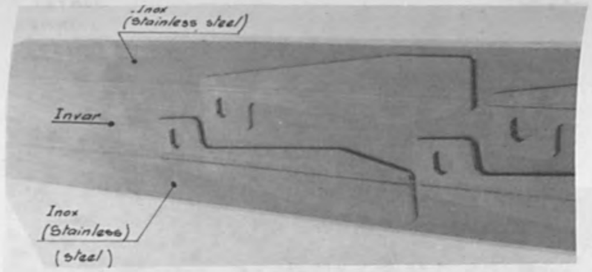
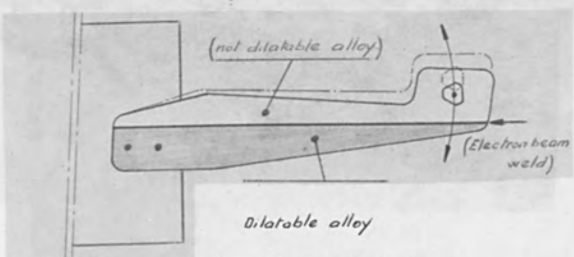


Fig 1



(Edge to edge bimetal strip)

Invar . Inox.

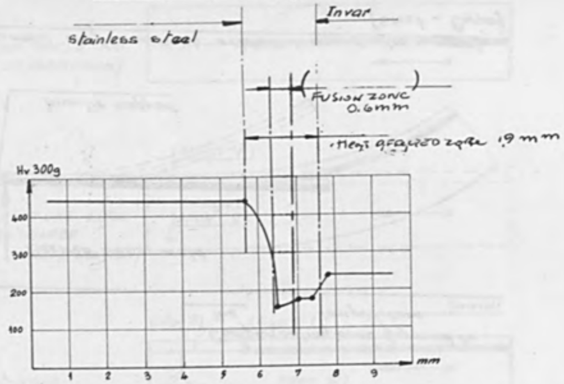


(Bimetal strip)

TECHMETA
ANNECY
FRANCE

-Fig 2/1-

-(Discontinuity of the welded zone determinated by of micro-hardness.)



Thikness - : 1,2 mm

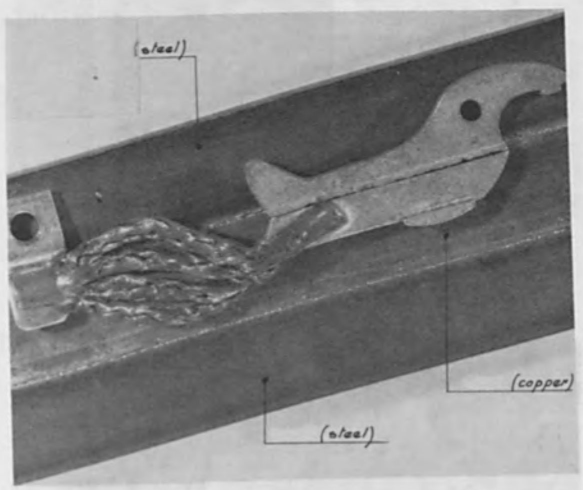
Welding speed : 7 m/min

TECHMETA
ANNECY
FRANCE

Fig 2/2-

- Bande cuivre - acier pour le découpage de pièces de contact électrique épaisseur : 3 mm.

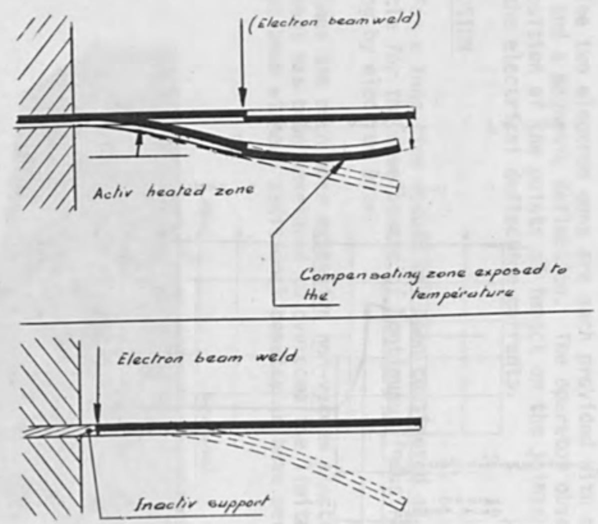
- (Copper steel strip for cutting of electrical contact components)



TECHMETA
ANNECY
FRANCE

- Fig 3 -

Bimétal strip sandwich with inverted components)

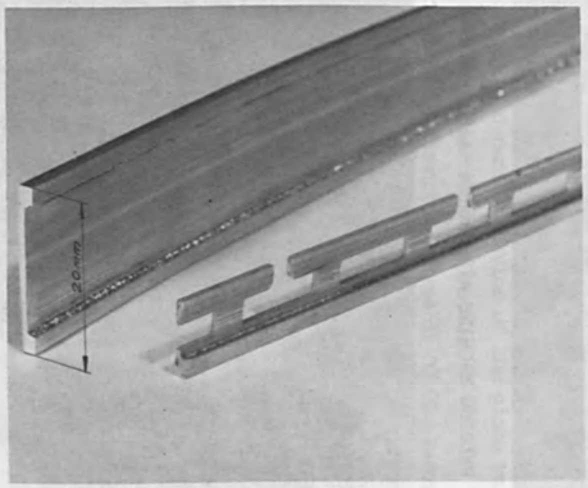


Bimétal strip sandwich welded to an inactiv support (Steel - Copper)

TECHMETA
ANNECY
FRANCE

- Fig 4 -

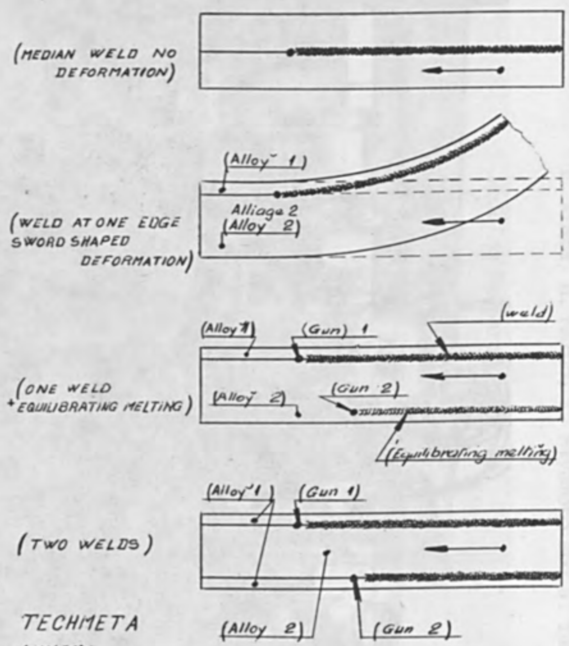
- (Profile in welded Zicral "light, hard alloy" for cutting of ski. edges)



TECHMETA
ANNECY
FRANCE

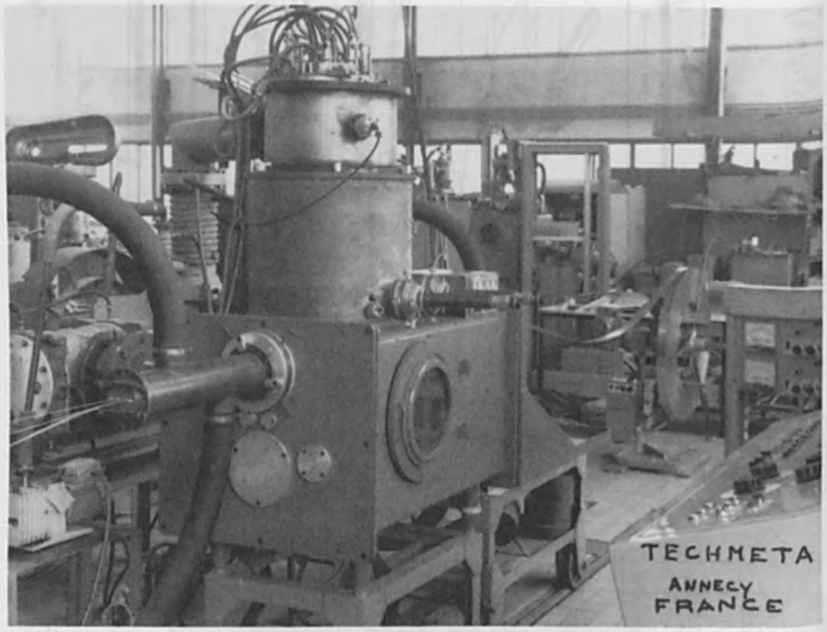
- Fig 5 -

(COMPENSATION OF DEFORMATIONS)

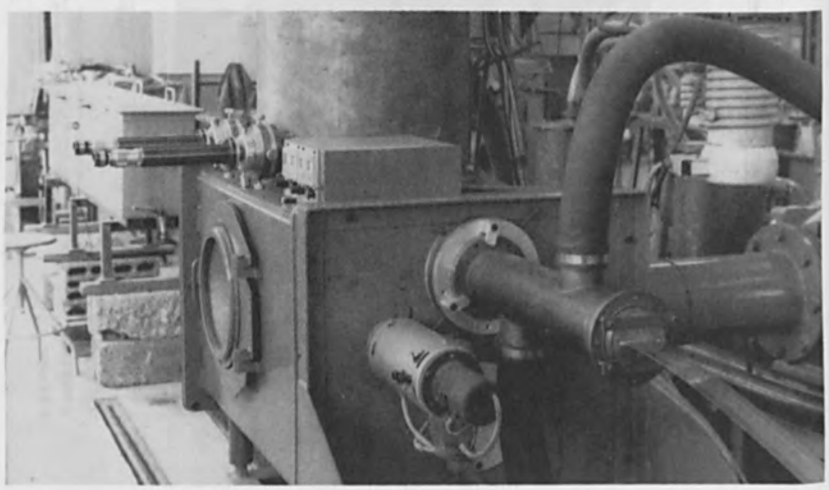


TECHMETA
ANNECY
FRANCE

- Fig 6 -



CONTINUOUS WELDING MACHINE AIR-VACUUM -AIR,WITH TWO GUNS.



DYNAMIC BEHAVIOR OF LASER WELDING AND CUTTING

Y. Arata, H. Maruo, I. Miyamoto and S. Takeuchi

Welding Research Insitute

Osaka University

Yamada-Kami, Suita-city, Osaka

Japan

1. Introduction

In recent years, the processing of materials by CO_2 lasers has rapidly advanced. Little is known, however, about the actual mechanism that occurs during laser welding or cutting, although some work has been done in this area. This is due to the difficulties of observation of the interaction between the laser beam and materials.

Accordingly, this study is directed mainly to the direct ovservation of the laser-material interaction. This paper consists of four sections: Section 2 gives welding mechanism obtained by using a transparent material. Section 3 demonstrates the dynamic behavior of the glass welding and section 4 shows the cutting mechanism of metals.

In this experiment, a conventional type of DC CO_2 laser with a total length of 6 meters and a 15 mm bore was used. The beam from the laser cavity was deflected downwards by a plane mirror to a focusing lens with a focal length 100 mm, and the beam was then focused to a spot roughly 0.3 mm in diameter. In laser gas cutting, reactive or non-reactive gas was directed to the workpiece through a convergent nozzle placed concentric to the laser beam.

2. Laser welding

It is well-known that the cavity exists in the path of the laser or electron beam resulting in the large depth-to-width ratio of the fusion zone. The possible unsteady nature of this cavity had been suggested[1], and some work has been done by using the X-rays technique[2] in order to observe the events occuring in the weld cavity. But the contrast of the radiographs was too low to understand the deep penetration mechanism.

To overcome this problem a new exprimental technique using transparent material has been developed to directly observe the dynamic behavior occuring during laser or electron beam welding. This technique and some of the results obtained in laser welding will be described here.

2.1 Shape of beam hole (cavity shape)

Various kinds of transparent material were tested, and it was found that glass is the best because it has a high heat diffusivity, and the appropriate amount of molten fluid is formed without excessive evaporation loss by a high power laser irradiation. The bead or cavity shape of the soda-lime glass was much similar to that of metals as shown in Fig.1.

(a) (b)

Fig.1 Examples of EB-weld bead and laser weld cavity (soda-lime glass).
(a) EB-weld bead (150KV-0.5mA, 4cm/min)
(b) Front view of laser cavity (150W, 4cm/min).

In this figure examples of the laser and electron beam welds are illustrated. No preheating was made and cracking could be suppressed during welding in spite of the fracture characteristics of the soda-lime glass. Here the influence of welding parameters on the shape of the cavity formed during laser welding is illustrated using the soda-lime glass.

The change in cavity shape produced by changes in laser power is shown in Fig. 2. Only a molten pool similar to the arc welding is observed when low laser power is used, but a typical cavity grows in the front part of the pool as the power increases. In the welds of the soda-lime glass, a lot of bubbles were left in the weld bead.

As shown in Fig. 3 the cavity is short and rather straight in shape when the welding speed is high, but as the speed decreases a constriction near the surface and a few swellings become obvious. In Fig. 4 the change in cavity shape produced by changes in the focal position is illustrated. The maximum penetration depth was obtained with the focal point of the beam 2.0 mm below the top surface.

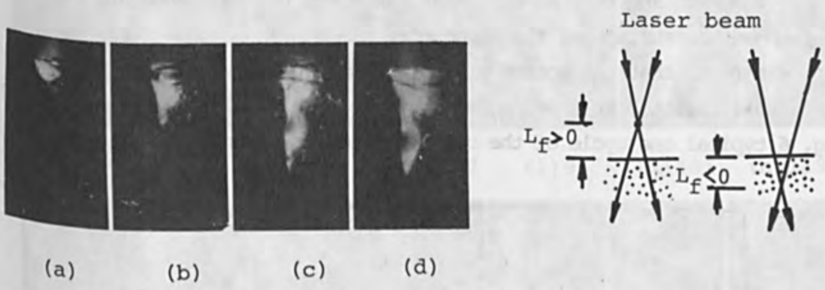


Fig.2 Effect of laser power on the cavity shape (7cm/min, Lf=+2mm).
Laser power: (a) 50W (b) 70W (c) 100W (d) 120W

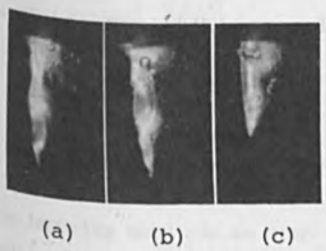


Fig.3 Effect of welding speed (150W, Lf=0). Velocity:
(a) 4cm/min (b) 6cm/min
(c) 10cm/min

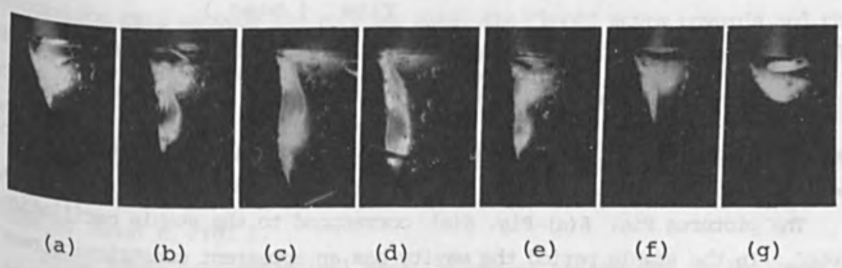


Fig.4 Effect of focal position (150W, 7cm/min).
L_f: (a) -6mm, (b) -4mm, (c) -2mm, (d) 0mm,
(e) +2mm, (f) +3mm, (g) +4mm.

2.2 Motion of beam hole (cavity motion)

Here dynamic behavior of laser welding is discussed based on the high-sp motion pictures. The pictures were taken for the bead-on-plate weld pass in soda-lime glass at about 1000 frames per second with no illumination. The wel speed was 4 cm/min and the laser beam of 150 W was focused through a Ge lens (f=100mm) on the surface of the workpiece.

It was found that the motion of the cavity was very rapid and periodic and that this resulted in a periodic change in the weld depth as shown in Fig. In Fig. 6 typical one cycle of the cavity motion is shown.

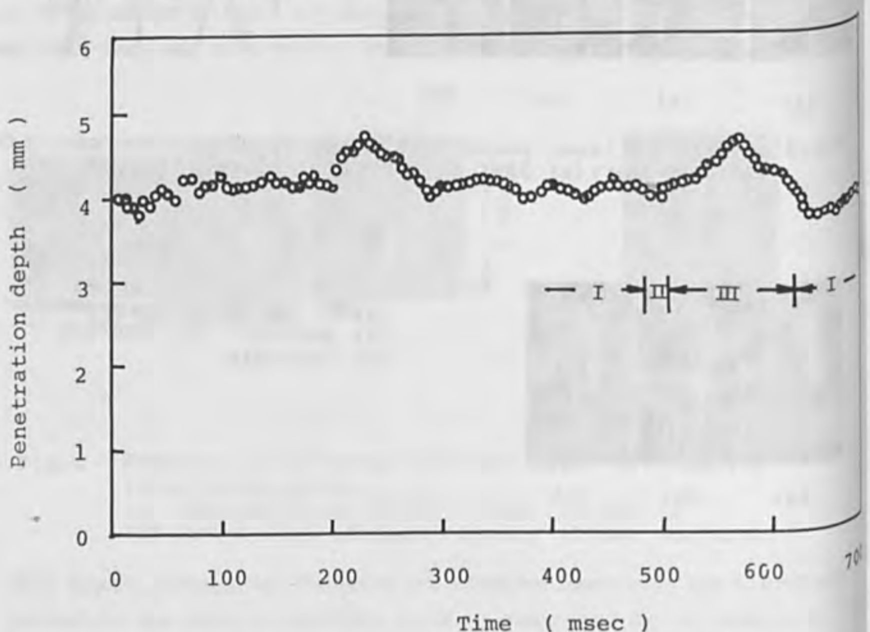


Fig.5 Variation of weld depth with time (150 W,4cm/min, L_f=1 mm). I: Stable oscillation period, II :Instability formation period, III:Recovary period)

Every cycle can be divided into three periods: the stable oscillation period, the instability formation period and the recovery period.

The pictures Fig. 6(a)-Fig. 6(e) correspond to the stable oscillation period. In the stable period the cavity has an appearent constriction near the surface with a projection, which the authors have named "knob" and swell on the back side, which the authors refered to as "cave", with smaller swell. The molten fluid on the back side of the cavity is not obvious in these pictures because the films were developed so that the brighter portion would clear.

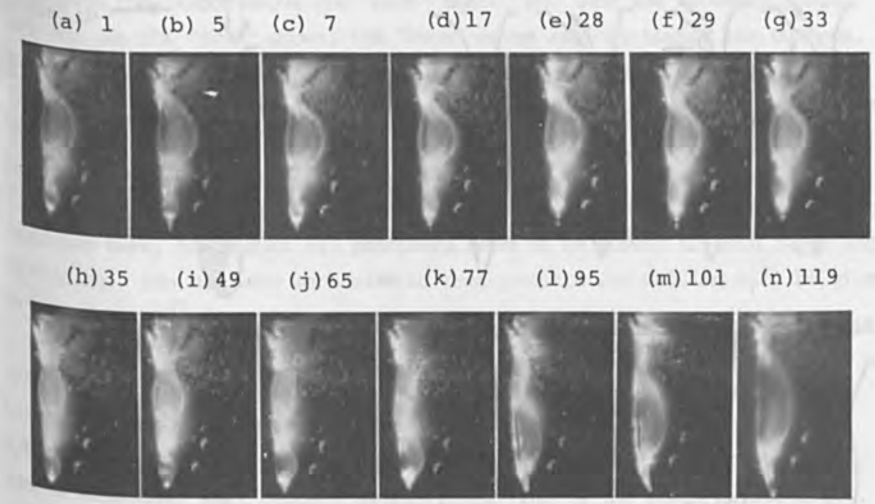


Fig.6 Typical one cycle of the cavity oscillation (1000fps).

In order to obtain the contour of the molten fluid, metallic powder was put on the glass surface and the melt run was made. Thus the outline of the impurity band can be seen after welding and it represents the back boundary, which is indicated by a dotted line in Fig. 7. In this period the evaporation rate in the cavity can be considered to be approximately in balance with the exhaust rate of vapor through the constriction. But if the evaporation rate exceeds the exhaust rate, the "knob" moves upwards and the hatched part of molten fluid flows rapidly out of the cavity through the constriction as shown in Fig. 7. Thereafter the "knob" moves downwards because the exhaust rate grows larger.

The oscillation of the lower "cave" and the sharp point at the bottom were also observed, but the variation of the penetration depth was not so large as shown in Fig. 5.

After these near-periodical oscillations of the cavity, it followed that the penetration depth suddenly increased. This phenomenon occured periodeically with the frequency of approximatly 4 Hz as shown in Fig. 5. This was observed to have started with a backward motion of the "knob" when the molten fluid flowed violently through the constriction. In the stable

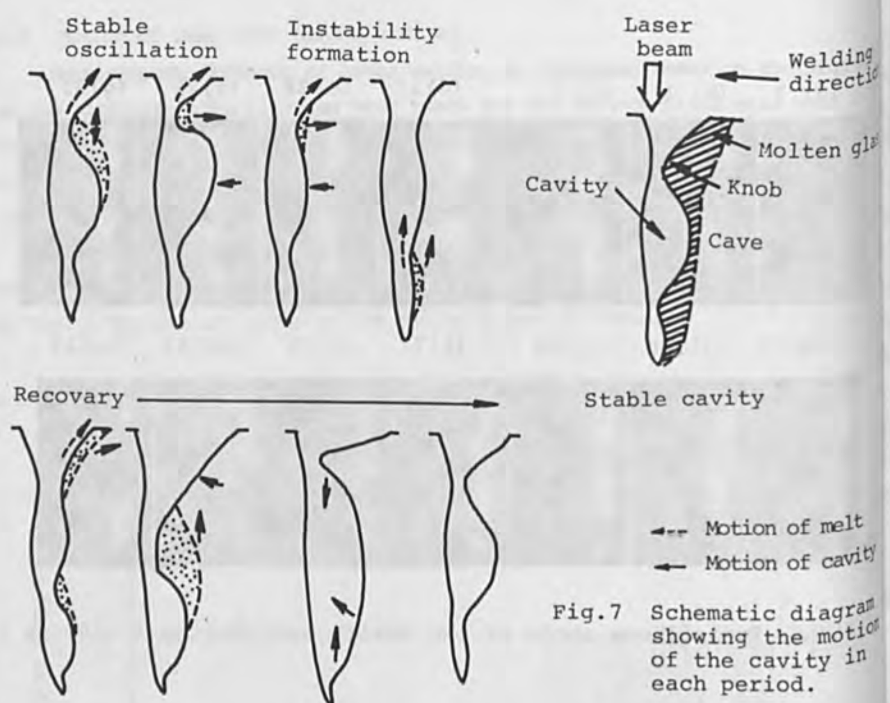


Fig.7 Schematic diagram showing the motion of the cavity in each period.

period the diameter of the cavity at the constriction also varied but it did not increase beyond a certain threshold under which the cavity remained stable.

When the cavity at the constriction increased in diameter beyond the threshold in very short time, about 20 msec, it could not shrink any more. This is the instability formation period and corresponds to Fig. 6(f)~Fig. 6 (i). This period results in a long and slender cavity having symmetry with respect to the vertical axis, but such a cavity can not be maintained, because the balance between the evaporation rate and the exhaust rate is suddenly broken and the temperature distribution cannot be symmetrical from the point of view of heat conduction theory. Because the constriction near the surface dissapears, the laser beam is not interrupted. The lower constriction opens and the penetration depth begins to increase. This is probably enhanced by the "wall-focusing effect"[3]. The penetration depth continues to increase until the middle of the next period.

In this situation the temperature of the bottom is considered to be so high that the lowest "cave" grows due to the high evaporation rate and a certain amount of molten fluid goes upwards. This molten fluid is observed to produce the "knob" above the "cave" again, and then the recovery period sets in. As the "cave" grows, the "knob" moves rapidly toward the surface, and one cycle is completed to produce the stable cavity as shown in Fig. 6 (j)~Fig. 6(n). The period of the recovery takes about 100 msec in this case. The shape of cavity in the recovery period varies with each cycle, but the recovery is considered to take place essentially in the same manner.

Typical examples of the dynamic behavior of the laser welding have been described here, but almost all phenomena seem to be common to both laser and EB welding. The complete mechanism of deep penetration welding will be given in near future.[4)]

As shown in Fig. 8 the cavity obtained for bead-on-plate on the 96% silicate glass is wedge-like for the low weld speed , and is bent near bottom. As the welding speed increases, the front wall of the cavity is apt to incline to the vertical line so that the incident angle of the laser beam to the front wall becomes smaller resulting in the lower reflectivity.

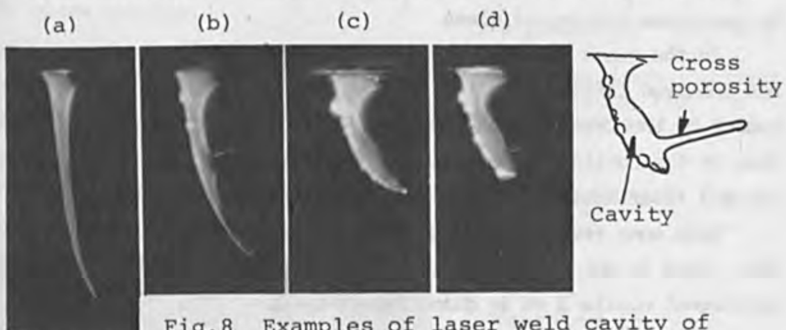


Fig.8 Examples of laser weld cavity of 96% silicate glass (90W, L_f=0). (a) 1cm/min, (b) 2cm/min, (c) 3cm/min, (d) 4cm/min.

The bottom of the bent cavity is heated by the wall-focusing effect although the direct laser beam cannot reach it. A violent motion of the cavity like that of soda-lime glass was not observed due to its high viscocity, but interesting phenomena were observed, for example, the bottom of the cavity swung backwards and forwards corresponding to the location of the wall heating as shown in Fig. 9. When the welding speed is higher, the "cave" was formed on the back side of the cavity, and it then grew in the direction perpendicular to the cavity wall, resulting in the long porosity inclined away from the horizontal line. The authors have named this "cross porosity". Occasionally two or three cross porosities were produced simultaneously.

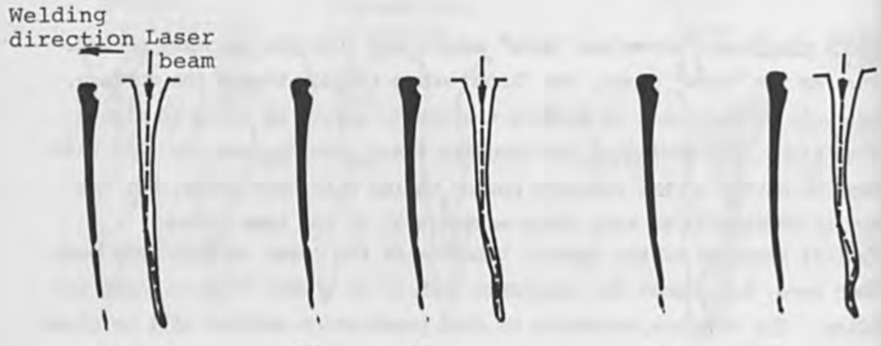


Fig.9 Motion of the cavity in 96% silicate glass.
(90 W, L_f=0, 1000fps, negative pictures)

3. Laser-gas cutting of glass

In the glass industry, the use of lasers for processing such as cutting has received little attention inspite of its usefulness. This seems to be caused by the fracture characteristics of glass, and some work[5)] has been done on the cutting of hot glass. But it is possible to cut cold glass by using a laser assisted by gas jet without producing any cracks.

Here some results obtained by observing the dynamic behavior which occur when glass is cut by the CO_2 laser are illustrated. In this experiment, a convergent nozzle 3 mm in diameter was used.

Very high power is necessary to allow separation of the glass or ceramic means of laser beam alone, because the molten fluid buries the space behind the cavity with the progression of the laser beam as described in the last section. By using gas jet assist the molten fluid is removed so that the cutting speed increases. In this experiment, borosilicate glass was cut with N_2 gas jet assist and no preheating, and no crack developed as long as the complete separation took place.

As shown in Fig. 10, the maximum cutting speed increases rapidly as the N_2 gas pressure increases as long as the gas pressure is low, but it reaches an almost constant value at the pressure around 1 kg/cm^2. The kerf, however, becomes more straight and narrower as the gas pressure increases as is shown in Fig. 11. The inclination of the cutting front is apt to decrease. On the other hand, the removed volume per unit of laser energy was found to decrease as the gas pressure increased.

High speed pictures were taken at about 1000 frames per second with the amera perpendicular to the cut. Figure 12 shows the pictures of the cutting ront of the glass of 3.5 mm hickness.

hen the pressure was 0.5 kg/cm^2 he workpiece could not be eparated and so only the half f the thickness was removed y evaporation and welding as carried out to a certain epth.

An obvious spiking phenom-non was observed to occur eriodically. The resultant roove surface was smooth and o drag line could be seen. n this case the glass specimen racked.

When the pressure was .0 kg/cm^2 separation occured esulting in downward flow of olten glass in droplet along the cutting front, and the cutting front in the ottom part inclined away from the vertical line.

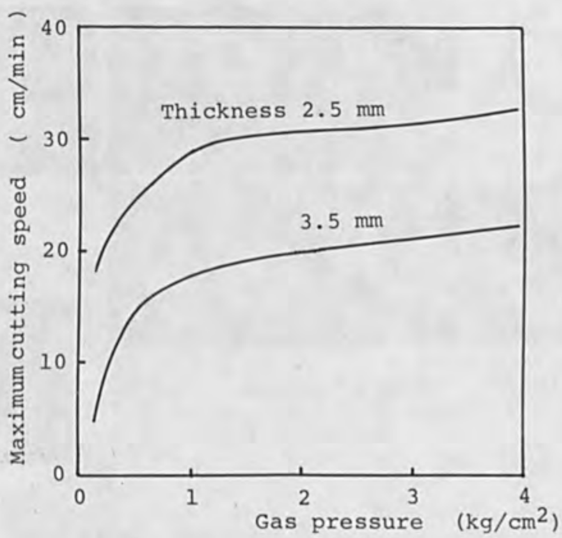


Fig.10 Relation between maximum cutting speed and gas pressure for the borosilicate glass (180W, L_f=0).

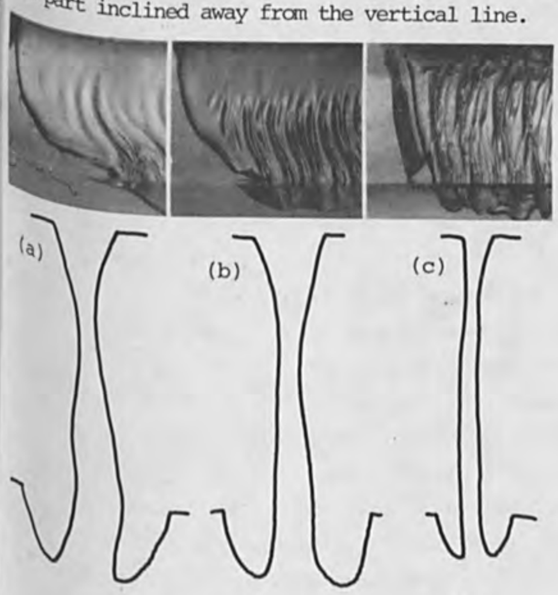
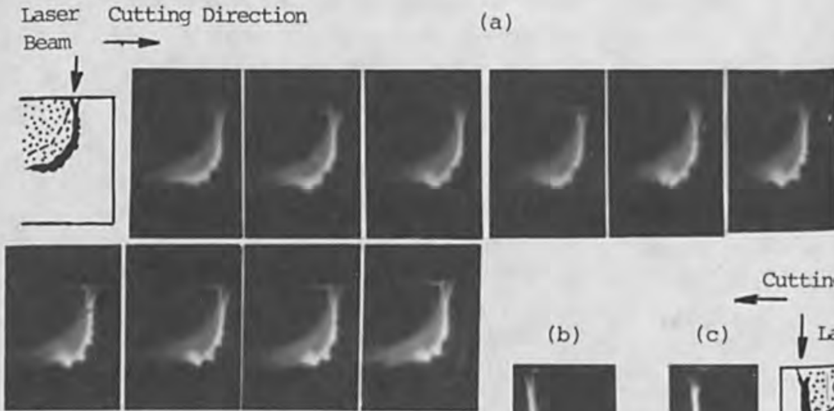


Fig.11 Examples of laser-gas cuts of the borosilicate glass (180W, L_f=0, 2.5 mmt, 23cm/min).
(a) 0.3 kg/cm^2
(b) 0.8
(c) 4.0



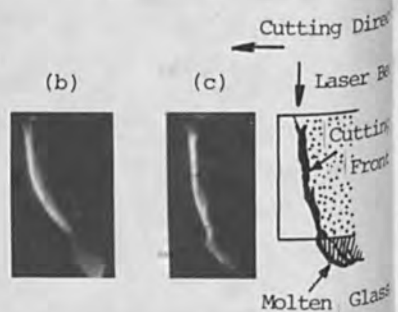


Fig.12 Effect of the gas pressure (5 mmt, 15 cm/min)
(a) $0.5 kg/cm^2$
(b) 1.0
(c) 4.0

There was no remarkable flow near the top surface creating a rather smooth cutting edge there. In this zone the material is removed mainly by evaporation and the melting rate of a smaller part of the glass is accelerated. The flow velocity was 15 cm/sec ~25 cm/sec.

In the lower zone where the fine and clear drag lines are visible, the removal by the flow of molten glass becomes remarkable. This molten glass is believed to heat the lower part and to be evaporated by the laser beam when it flows. Thus the small part of molten glass attaches itself to the rear surface, and is not blown off. When the gas pressure is lower, the cutting mechanism is believed not to be homogeneous, and this fact is comfirmed by the analysis based on the heat conduction theory [6].

When the pressure is 4.0 kg/cm^2, glass is observed to melt from near the top surface, and the velocity of the molten droplet was considerably higher than that of 1.0 kg/cm^2. The flow velocity along the cutting front which is almost vertical is found to have minimum value, about 30 cm/sec, near the center of the thickness. The flow velocity vear the top and bottom surface was about 50 cm/sec. From the rather straight kerf obtained when the pressure is 4.0 kg/cm^2 the cutting is considered to be more homogeneous from the top to the bottom of the thickness. The amount of molten glass attached to the rear surface was much smaller than the volume of the kerf, and almost no molten glass flowed out of the plate. These facts show that the gas jet makes the cutting phenomenon homogeneous throughout thickness by heating the lower part due to the downwards flow, and it makes the kerf straight. But if the gas pressure is too high, the cutting speed will decrease due to the cooling effect of the gas.

4. Laser-gas cutting of metal

A laser-gas cutting in which CO_2 laser is combined with a reactive gas jet for cutting metallics results in a very narrow kerf width and heat affected zone. In this area some papers have been published[7], but little is known about the cutting mechanism. Here a cutting mechanism based on the observed dynamic behavior and some experimental results obtained uaing mild steel and austenitic stainless steel will be discussed. In this experiment a convergent nozzle of 1.0 mm in diameter was used, and oxygen pressure was 3Kg/cm^2.

4.1 Classification of cutting edge

Fig. 13 shows the relation between the thickness of stainless steel and mild steel and their respective maximum cutting speeds. In Figs. 14 and 15 the variation of the edge condition with cutting speed for mild steel are shown. The kerf width at the surface is about 0.3 mm which corresponds to the beam diameter, and is almost independent of oxygen pressure.

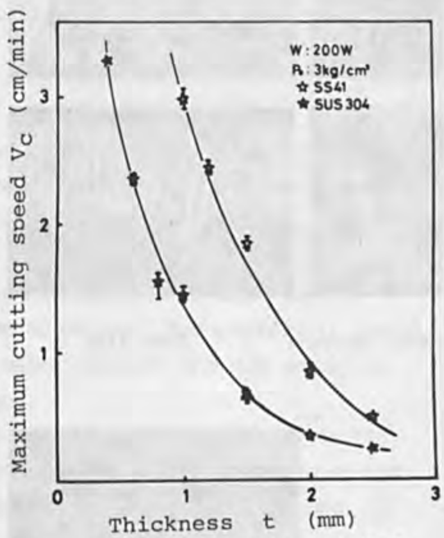


Fig.13 Relationship between thickness and maximum cutting speed (SS41: mild steel, SUS304:18/8 stainless steel; 200W, 3kg/cm^2, L_f=0).

In general it is observed that the drag lines inherent to the laser-gas cutting are discontinuous at point a and b, and that the kerf width varies correspondingly as is shown in Fig. 16. Then the cutting edge can be classified into three zones; A, B and C zone as Fig. 16: the A zone has a taper kerf, the B zone has a straight kerf and the C zone has a very rough kerf. But these all zones do not necessarily exist: when the thickness of the workpiece is very thin, the C zone does not exist. (See Fig.22) From these facts the phenomenon in each zone is considered to be different.

The cutting edge of stainless steel also has three zones, although its drag lines are generally not as clear as those of mild steel. (Fig. 17)

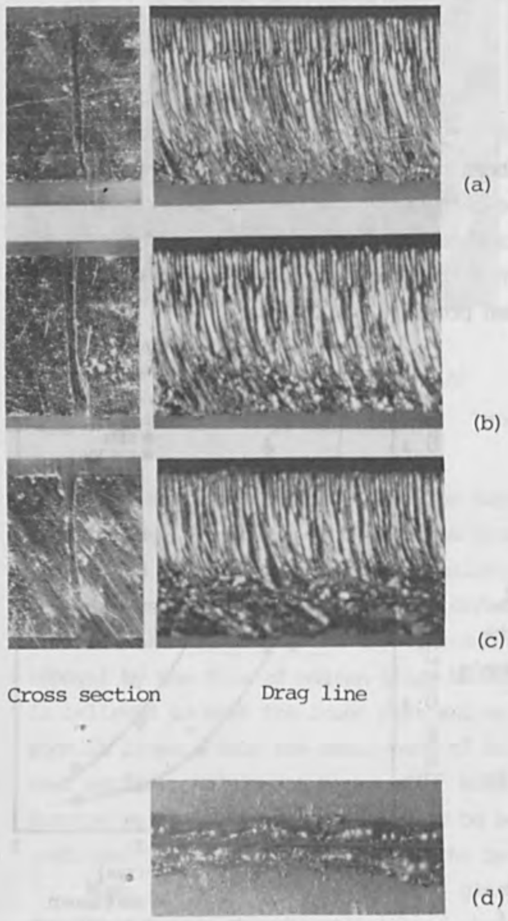


Fig.14 Cut edges of mild steel (SS41 2mmt, 200W, $3kg/cm^2$, L_f=0).
(a) 0.2m/min
(b) 0.4
(c) 0.8
(d) 0.4

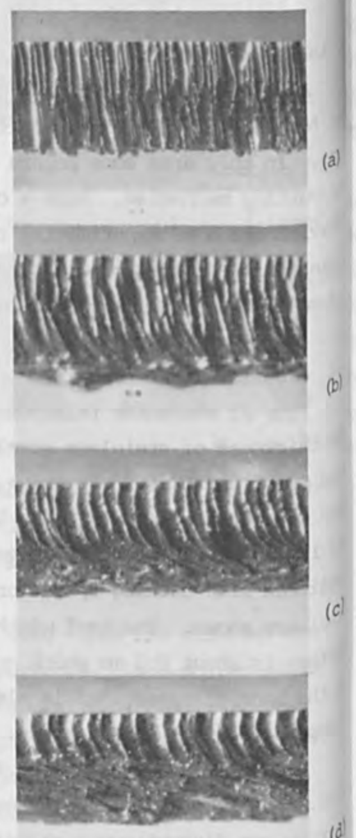
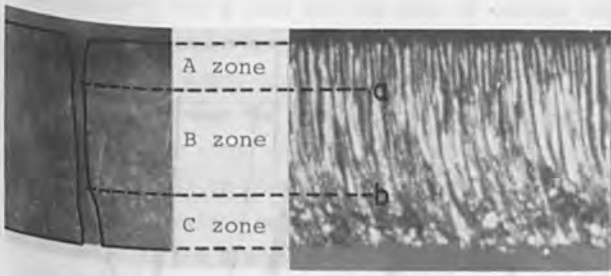


Fig.15 Cut edges of mild steel (SS41 1mmt, 200W, 3kg/min, L_f=0).
(a) 0.065m/min
(b) 1.8
(c) 2.4
(d) 2.9



← Cutting direction

Cross section   Drag line

Fig.16 Classified three zones of cut edge (SS41 2mmt, 200W, 3kg/cm^2, 0.8m/min, L_f=0).

Drag line

Fig.17 Drag line of austenitic stainless steel (SUS304 1mmt; 200W, 3kg/cm^2, 0.2m/min, L_f=0).

4.2 High speed filming of edge cut

The laser-gas cutting of mild steel with an edge cutting was filmed at about 1000 frames per second. Extreme care was exercised to assure that the focal point was just on the edge of the workpiece surface so as to machine a thin slice from the workpiece edge.

Viewing the film prints, the cutting phenomenon was observed to vary periodically. In case of the cutting speed 60 cm/min , for example, it varied with 70~80 Hz, which agreed well with the value obtained from the resulting drag lines whether they were edge cut or not.

One cycle (14~16 msec) of pictures is illustrated schematically in Fig. 18, where the sequence is arbitrarily started with the picture having no visible radiation in the A zone and narrow molten oxide flow in the B zone, ① ②, this part of the sequence continues for 10 msec. Then oxidation reaction starts to grow into the A zone and the molten layer in the B zone becomes gradually wider ③. Subsequently, droplets of oxide begin to flow down along the oxidation interface ④. On reaching to the top surface, the oxidation interface suddenly moves forwards very rapidly (≃0.8 m/sec) towards the cutting direction in the upper part inclining away from a vertcal lime in the A zone in 4~5msec, and dissapears. In this period the shape of the oxidation front at the B zone almost does not change, but the flow of the melt becomes much more obvious leaving a drag line at the back of melt layer due to the solidification ⑤ ⑥. Thus one cycle is completed.

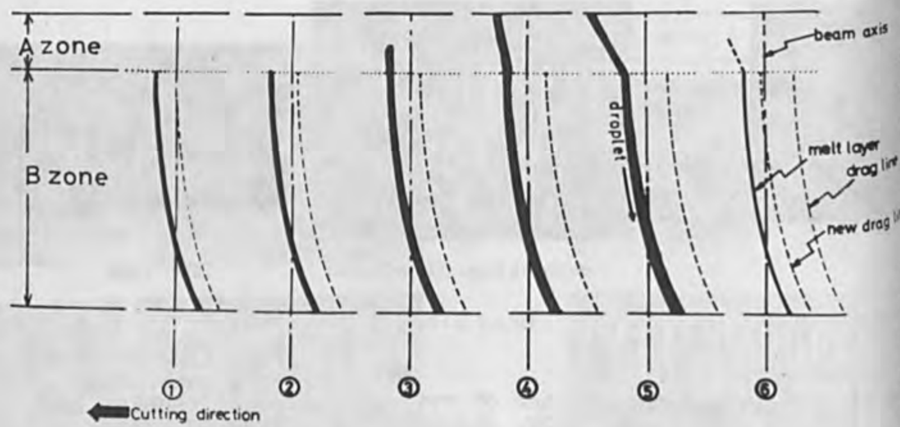


Fig.18 Periodical motion of cutting front.

Films of edge cut show that in laser-gas cutting the reaction proceeds periodically leaving a drag line for each cycle, and that the cutting phenomenon differs with each zone. The next section proposes a dynamic model for laser-gas cutting and shows how this model accounts for the phenomenon in each zone.

4.3 Discussion

(1) A zone

The mechanism of the laser-gas cutting is explained by using the dynamic model shown in Fig. 19 on the basis of the facts mentioned above.

In the A zone the laser beam does not seem to come in contact with the cut edge for a while, because the cutting front goes ahead of the beam and the edge is out of range of the laser beam. This corresponds to ① and ② in Fig. 18. As the laser beam progresses at constant speed, it comes in contact with the edge

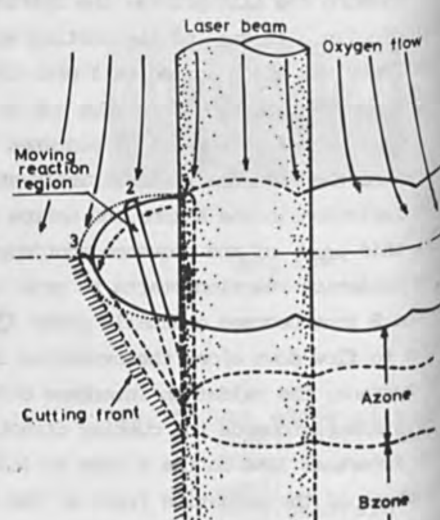


Fig.19 Movement of cutting front in the A zone.

from the bottom of the A zone and the area of cantact expands up to the top surface due to the taper kerf shown in Fig. 19. Thus the linear reactior interface 1 is produced, which corresponds to ③ and ④ in Fig. 18. Because the oxygen is very pure near the top surface, the resultant interface 1 moves forward very rapidly along the conical surface of the edge with the trigger of the laser beam, until it reaches line 2 then line 3. On reaching line 3, the linear reaction interface suddenly disappears then there is obvious flow of molten material in the B zone.

Thus the rapid oxidation reaction removes a thin slice from the taper cutting front in the A zone, and the cutting front goes ahead of the beam. This periodical oxidation reaction produces the cut edge with periodical variation in the kerf width as shown in Fig. 20. No reaction occures until the laser beam catches up with the cutting front. Increasing the cutting speed increased the pitch of drag lines and reduced the lengh of the A zone as shown in Figs. 21 and 22. Similar results were found for stainless steel, but the drag lines were more closely spaced than for mild steel resulting smooth cut face.

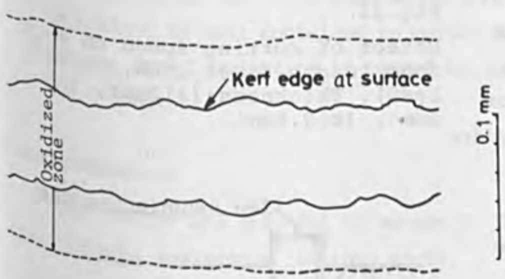


Fig.20
Typical kerf appearance of mild steel (SS41, 200W, $3kg/cm^2$, 0.065m/min, L_f=0).

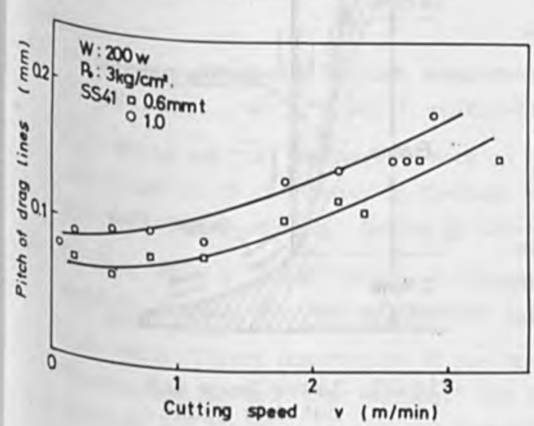


Fig.21
Effect of cutting speed on the pitch of drag lines for mild steel (SS41, 200W, $3kg/cm^2$, L_f=0).

(2) B zone

In contrast with the intermittent motion of the reaction interface in the A zone, the reaction interface in the B zone moves at a constant speed. As the cutting speed increases, the drag lines incline diagonally especially near the bottom, and their lengths decrease in the B zone as shown in Figs. 14 and 15 and Fig.22. The drag, however, does not exceed 0.3 mm, that is to say it is the same as the diameter of the focused laser beam. This means that the laser power is indispensable for maintaining the oxidation in the B zone as shown in Fig. 23. The energy required for cutting in the B zone is supplied by the laser beam and the exothermic reaction, and is adjusted by the inclination angle of the cutting front.

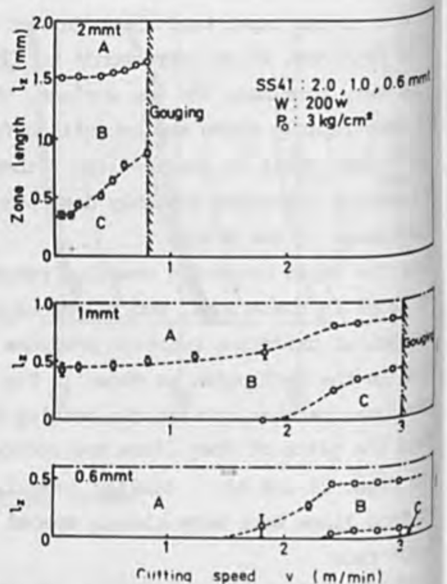


Fig.22

Effect of cutting speed on each zone length (SS41,200W,3kg/cm^2, L_f=0). Thickness:(a)2mmt, (b) 1mmt, (c)0.6mmt.

(3) C zone

The C zone is considered to be the zone which is out of the beam path, and thus the laser beam contributes little to the cutting mechanism. In this zone heat is supplied by the high temperature molten oxide flowing down from the upper zones, oxidizing away the edge in the oxygen flow. Therefore the inclination angle of the cutting front to the vertical line increses greatly, in some cases even up to almost 90°. As the inclination angle increases, the dross is apt

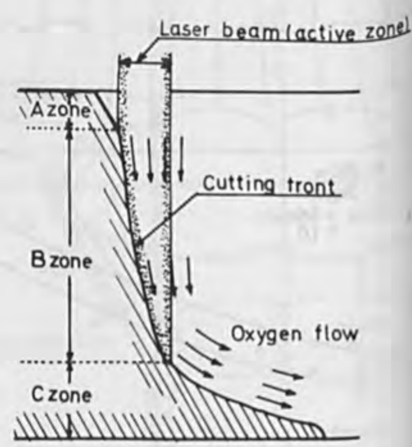


Fig.23 Laser beam and cutting front.

to attach itself to the cut edge or to the rear surface by solidification. Then the oxygen flow passing through the plate is interrupted and the cutting stops.

5. Conclusion

Conclusions obtained from this study are summarized as follows:

1) The welding phenomena of the glass are very similar to those of metal. The motion of the cavity maintained during welding is periodic and each cycle is divided into three periods: the stable oscillation period, the instability formation period and the recovery period. Some examples of cavity motion and weld defect were shown.
2) The laser-gas cutting of the cold borosilicate glass could be done without any crack. The gas jet assisted was found to make the cutting homogeneous throughout the thickness.
3) The mechanism of laser-gas cutting in metal is periodcal, and the cutting front is classified into three zones: a zone where an intermittent oxidation triggered off by the laser beam takes place, a zone where the reaction area is confined by the laser beam and a zone where the oxidation reaction is maintained by heat contained in molten oxide flowing down from the upper zones. Speed limitation is caused by the interruption of the oxidation in the C zone.

Acknowledgements

The authors are grateful to Messrs S. Ishida, N. Enomoto, T. Tokuhisa for their helpful assistance in this study.

Reference

1) H. schwartz,"Mechanism of High Power-Density Electron Beam Penetration", J. Appl. Phys., Vol. 35, No. 7, pp2020-2029 (1964).

O.C. Weils and T.E. Everhart,"A Note on the Physical Principles Underlying the Formation of the Cavity in Electron Beam Welding", 4th Annual Meeting Electron Beam Technology, Bakish, pp 105-122 (1962).

2) H. Tong and W.H. Giedt,"Dynamic Interpretation of Electron Beam Welding", Welding Journal, Vol. 47, pp259s-266s (1970).

C.M. Weber,"Direct Observation of the Penetration Mechanism of Partial Penetration Electron Beam Welding," 5th International Conference on Electron and Ion Beam Science and Technology, pp 307-320 (1972).

Y. Arata, E.Abe and M. Fujisawa,"Study on Dynamic Behaviours of Electron Beam Welding by a Fluoroscopic Observation," 7th International Conference on Electron and Ion Beam Science and Technology (1976)

3) Y. Arata and I. Miyamoto,"Wall-Focusing Effect of Laser Beam", The second International Symposium of the Japan Welding Society, Advanced Weld. Tech., 1, pp125-130 (1975).

4) To be published.

5) G.K. Chui,"Laser Cutting of Hot Glass", Amer. Ceram. Soc. Bull.,Vol. 48, pp514-518 (1975).

W. Haefner,"Method of Working [Cutting] Glass with Absorbent by a Laser Beam", U.S. Pat.3, 453. 097, July 1 (1969).

6) To be published.

7) A.B.J. Sullivan and P.T. Houldcraft,"Gas-jet laser cutting", Brit. Weld.J., pp443-445, August (1967).

F.W. Lunau and F.W. Paine,"CO_2 Laser Cutting", Weld. and Met. Fab. pp9-14, January (1969).

Simon L. Engel,"Laser Cutting of thin materials", SME Technical Paper, MR 74-960 (1974).

W.W. Duley and J.N. Gonsalves,"CO_2 laser cutting of thin metal sheets with gas jet assist", Optics and Laser Technology, pp78-81, April (1974).

M.J. Adams,"Gas jet laser cutting", Proccedings of the Conference on Advances in Welding Processes, The Weld. Inst., pp140-146, April (1970).

HIGH-RATE SPUTTERING WITH A TORUS PLASMATRON

Heisig, U., Goedicke, K., and Schiller, S.
Manfred von Ardenne Research Institute,
Dresden, GDR

Abstract

High-rate sputter sources for cylindrically symmetric arrays of target and substrate can be put into being by using the principle of ring gap discharges. Electrostatic and magnetic focusing of a gas discharge allows to generate a toroidal plasma of high density on either the inner or outer surface of a tubular target. Against cylindrically symmetric diode arrays with homogeneous discharge such a setup is characterized by an increase in sputter rates of about one order of magnitude. Such high rates may also be attained with substrate arrangements extending toward the cylinder axis by joining several toroidal discharges one after the other. With copper and sputtering toward the outside, for instance, the deposition rates exceed 1 µm/min (tube distance about 50 mm). Depending on the selected array, sputtering toward the inside gives deposition rates that are several times as high.

Further advantages of Torus Plasmatron Sputter Sources are low working pressure, low specific energy demand, high degree of material utilization and almost universal coating possibilities; i.e., also of profiles.

1. Introduction

With respect to its geometry, a sputtering plant is characterized by the shape of target and substrate and their positions relative to each other. Customary arrangements are diode arrays with plane substrate and target surfaces as well as cylindrical electrode arrays with coaxial tubular target and substrate surfaces.

In order to solve coating problems in a profitable manner it has been attempted time and again to enhance the sputter rates; among other things also by way of magnetic fields. From all the envisaged magnetic field configurations it is especially the closed ring gap field that substantially increases the ion current density on the target and thus both sputter rate and specific energy utilization during sputtering. By using these ring gap discharges [1] [2] for sputtering with plane targets, sputter rates were increased against those of conventional electrode arrays with plain longitudinal magnetic fields by one order of magnitude. Referred to copper and a discharge power of 5 kW, for example, sputter rates of about 20,000 Å/min can be attained when using such a planar high-rate sputter source with diode arrays [3][4][5].

Magnetic fields were further used in cylindrical sputtering arrangements to increase the discharge current density and thus the sputter rate. With a homogeneous magnetic field toward the cylinder axis - that is, in a magnetron arrangement - sputter rates of 1,000 to 4,000 Å/min were measured with copper and a discharge power of 2.5 kW, just to cite an example [6].

The presented paper deals with cylindrically symmetric high-rate sputter sources with ring gap discharges; i.e. with socalled Torus Plasmatrons.

2. Performance

In the case of a ring gap discharge the target - which serves as cathode of the discharge - is arranged in the closed gap field of a magnet. With tubular targets and a toroidal ring gap, a toroidal region will also be formed on the cathode within which the magnetic field strength has the characteristic pattern plotted in Figure 1. Arranged parallel to the cathode surface, the component B_Z of the magnetic induction reduces not only toward the anode but also from the gap center toward both gap edges. In the pressure range of 10^{-1} Pa (10^{-3} Torr) a potential of about 500 V between cathode and anode results in a heavy-current discharge provided that the magnetic induction, B_Z, on the cathode surface is greater than $3 \cdot 10^{-2}$ Vs/m^2. Characteristic of such a discharge is the formation of a toroidal region of high plasma density above the ring gap in the immediate vicinity of the cathode.

Charge carrier production and parameters of the ring gap discharge are best explained by way of the motion of electrons that are liberated at the cathode by ion action. To this end, let's consider their motion in the crossed electric and magnetic ring gap field. Figure 2 shows the plain model of a potential distribution between cathode and anode together with the schematic trajectory of an electron which emerges from the cathode. The orbit of this electron lies within the ring gap near the tubular cathode. In the said illustration, this trajectory has been rolled out on a straight line, $R_g \cdot \varphi$, where R_g is the radius of the directrix of electron motion. First the electron is accelerated in the cathode fall region I which, in the case of heavy-current discharges, covers only some tenths of a millimeter. In the adjacent ionizing zone II - the extension of which is also small against the distance between cathode and anode - the electron is returned into the cathode fall on a circular orbit. There it is moderated until its direction of motion reverses and then accelerated anew. On its travel along the said directrix the electron repeatedly collides with molecules of the discharge gas. In this way it generates ions, slow electrons and excited atoms, releasing its energy until it has been moderated to the quasi-thermal energy level. Unlike to glow discharges (where a Townsend avalanche is released) the cathode electron generates plasma electrons and ions the number of which depends on the multiplication constant M. Based on some simplifying assumptions, this multiplication constant then is

$$M \approx \frac{U_C - \sum_{\nu} eU_{\nu}}{eU_I}$$

where

U_C = potential difference in the cathode fall

$\sum_{\nu} eU_{\nu}$ = total energy released by exciting collisions

eU_I = ionizing energy of the discharge gas.

Still another characteristic of ring gap discharges is the fact that the electron emerging from the cathode is guided into one of the directrices. The position of this master directrix relative to the ring gap is fixed by the condition $B_r = 0$. This guiding action results from the inhomogeneity of the magnetic field which causes direct magnetic focusing owing to the force components e $(w \times \mathcal{B}_r)$ being directed toward the said master directrix. (Here $\mathcal{B}$ is the radial component of magnetic induction: i.e. $\mathcal{B}_r$ = $\mathcal{B}_{H,\mathcal{M}_r}$; cp. Figure 3a). Still another focusing effect results from the curvature of the cathode fall transverse to the ring gap. This curvature causes that, upon entering the ionizing zone, the velocity vector v of the cathode electrons encloses an angle toward the radial direction; with the plane of the master directrix being the sole exception. Such a curvature of the cathode fall is caused by the inhomogeneity of the current density transverse to the ring gap and the erosion profile of the cathode; i.e. the equipotential surface $U = U_T$. Plotted in Figure 3 are the mentioned focusing actions on electrons emerging from the cathode.

3. Torus Plasmatron with One Ring Gap Unit

A ring gap discharge as described in the previous chapter has been used to realize a torus sputter source for coating the outer surface of the substrates. This source consists of a water-cooled tubular target having an inner diameter of 62 mm and a disk-shaped permanent ring magnet which sorrounds the target so as to generate a magnetic gap field on its inner wall. During operation of the Torus Plasmatron the toroidal zone of highest plasma density is formed within the region of the inner target wall. Figure 4 gives the geometry of a sputtering arrangement designed for coating the outer wall of tubular substrates having a diameter of 12 mm. This figure further shows the film thickness distribution along the axis of a fixed substrate. With a discharge power of 5 kW and a pressure of 5 m Torr the presented geometrical arrangement yields very high deposition rates of up to 50,000 Å/min. Owing to the pronounced inhomogeneity of the source distribution of sputtered particles, however, the thickness distribution on the substrate will differ accordingly. But averaged over a substrate length of 5 cm the deposition rate still amounts to 37,000 Å/min. Comparative measurements with respect to a planar plasmatron sputter source have shown that - based on equal magnetic induction, ring gap length and pressure as well as an adequate distance between target and anode - discharge power, voltage and sputter rate at the cathode are almost the same. With 5 kW discharge power and an argon pressure of 0.6 Pa (5 m Torr), the geometrical arrangement plotted in Figure 4 permits, for instance, to sputter 0.7 g/min of copper. The deposition rate on the substrate then follows this sputter rate in compliance with the geometrical parameters of both target and substrate (cp. Section 5).

To achieve a uniform thickness of the condensed layer with torus sputtering of long substrates irrespective of the inhomogeneous source distribution on the target, it is a customary practice to move target and substrate toward the common axis relative to each other. This is the more so as the high sputter rates, obtained within reach of the ring gap assembly, allow the use of a comparatively high relative velocity.

The inhomogeneity of ion current density on the cathode must, of course, cause a dissimilar erosion of the target material within the region of the ring gap. Plotted in Figure 5 is the erosion profile of a tubular copper target obtained after 5 hours of operation at a discharge power of 5 kW.

Generally, the formation of such a profile during ring gap discharges results in the drawback that only 30 to 40 % of the target material is actually utilized for coating. When this portion of the target material is used up, the mechanical strength of the target becomes questionable. Shifting the erosion profile of the target by altering the position of the magnet relative to the latter yields only minor improvements as far as plane targets are concerned. But with the Torus Plasmatron, the tubular shape of the target enables a more regular erosion of target material simply by shifting the magnet relative to the target. So the utilization of target material can be increased to about 70 %. Still another advantage of target shifting is the gain in continuous working hours obtained for a sputter source with given target.

4. Torus Plasmatron Sputter Source with Several Ring Gap Units

An even film thickness distribution on long rod-shaped or tubular targets is readily obtained without relative motion between sputter source and substrate by arranging several ring gap units on one and the same target. Shown in Figure 6 is a source consisting of five ring gap units as well as the attained thickness distribution along the axis of a fixed substrate. With given diameters of substrate and target, the tolerance of film thickness depends on number, width and distance of the individual ring gap units. In the above example these parameters were selected in a way that, in practice, two neighbouring thickness maxima cannot affect each other any more. Although this fixes the tolerance for the given geometrical conditions to ± 12 %, these limits can only be maintained over the maximum tube length by using the given number of ring gap units. For comparison, this illustration not only gives the measured film thickness distribution but also a distribution as obtained by geometrical superposition of five ring gap units of similar design.

With a sputter source of five ring gap units as illustrated in Figure 6 a discharge power of 25 kW actually brought about a copper deposition rate of 0.56 g/min measured on a tubular substrate of 12 mm diameter. Based on a plasmatron sputter rate of 3.5 g/min at the cathode and the mentioned parameters the portion of deposited material on the substrate amounts to 16 percent of the sputtered material.

Figure 7 gives the discharge voltage plotted versus discharge power for a torus plasmatron discharge in argon with copper target.

It is mainly the heavy-duty torus plasmatron with several ring gap units where a relative motion between substrate and target is preferably used. This not only improves the tolerance of film thickness as compared to the mentioned values but also permits to coat substrates having a length that is a multiple of the gap width. At the specified discharge ratings it will be possible, for instance, to coat a tube with a copper film of 1 µm thickness at an advance of more than 1 m/min.

Apart from the described mode of operation, where the flow of sputtered particles is directed toward the axis, torus plasmatron sputter sources with one or several ring gap units may further be used in quite a different manner. When the ring gap fields are produced by suitable magnets arranged within a tubular target, a flow of sputtered particles can be obtained on its outer surface that is directed toward the outside. Some examples of practicable substrate arrangements for both versions of torus plasmatron sputter sources are depicted in Figure 8.

5. Condensation Rate and Share of Deposited Material on the Substrate

The mean rate of condensation, a_c, and the portion α of the material being deposited on the substrate can also be obtained theoretically for a Torus Plasmatron with cylindrical substrate and target. This is best accomplished by introducing some simplifying assumptions: - First of all, a source in form of a closed circle line with cosine distribution [7] shall be used for describing the directional distribution of the particle flow caused by an actual ring gap discharge with parameters as mentioned in the previous chapters. This feature has been verified experimentally for the planar plasmatron [5][8]. Let's further assume that, for reasons of symmetry, it shall be sufficient for a circular cylinder array to consider only one of its elements - i.e., a small-area source - instead of the whole circular ring gap source. Moreover, losses due to dispersion of sputtered particles into the front faces of the sputtering arrangement shall not be taken into account.

Based on these suppositions, the portion α then is

$$\alpha = \frac{D_S}{D_T}$$

where

D_S = substrate diameter and

D_T = target diameter.

Whilst practically all the sputtered material condensates on the substrate when D_S is almost equal to D_T, only a certain part of the sputtered material will be deposited on the substrate the portion of which increasingly diminishes with reducing substrate diameter.

With geometrical data as specified for the described experimental setup a value of $\alpha = 0.19$ is to be expected. The difference to the measured value ($\alpha = 0.16$) can be explained by the dispersion of sputtered particles into the front faces of the target.

By definition we now obtain

$$\alpha = \frac{m_S}{m_T} = \frac{D_S}{D_T}$$

where

m_S = mass deposited on the substrate per unit of time

m_T = sputtered target mass per unit of time.

From this, it follows that within the target the relation

$$a_c = a_T$$

holds good for any substrate diameter; i.e. the mean film thickness a_c deposited per unit of time is equal to the

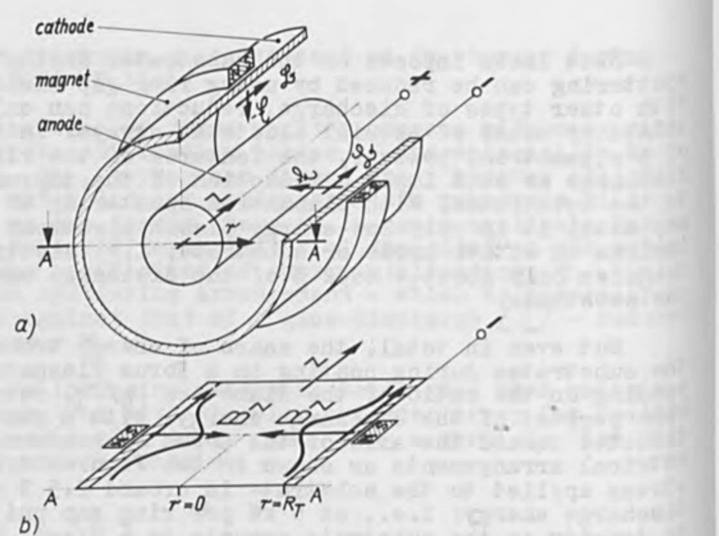


Fig. 1. Torus-Plasmatron ring gap discharge. a) electrode arrangement and directions of the electric field $\mathfrak{E}$ and magnetic field $\mathfrak{B}$ in the gap region for a hollow cathode. b) distribution of the magnetic field component B_z (r, z) [B_z (R_T, 0) > B_z (r, 0) for r < R_T]

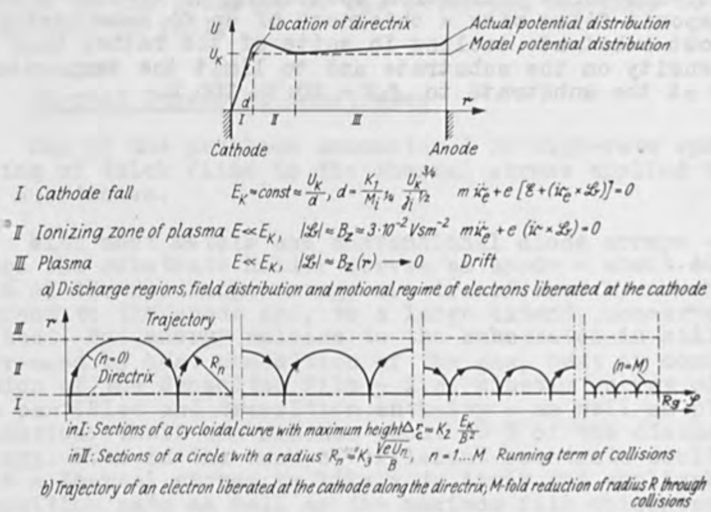


Fig. 2. On the mechanism of ring gap discharges

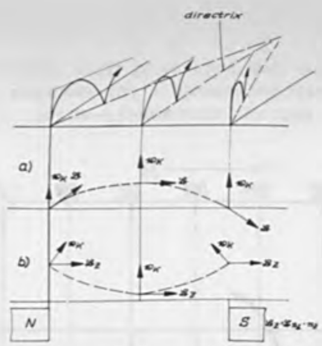


Fig. 3. Ring gap focusing of electron emerging from the cathode. a) Influence of an inhomogeneous magnetic field $\mathfrak{B}$ alone. b) Influence of an inhomogeneous electric field caused by cathode curvature.

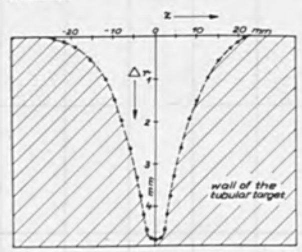


Fig. 5. Torus-Plasmatron source with one ring gap unit, profile of erosion Δr in the tubular target obtained by sputtering. Parameters: target material copper, P = 5 kW, p = 5 m Torr ($\approx$0,6 Pa) Ar, $t_D \gtrsim$ 300 min.

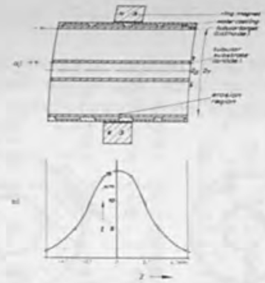


Fig. 4. Torus-Plasatron with one ring gap unit. a) Geometrical arrangement. b) Thickness distribution t on the substrate along axis z. Parameters: target material copper, P = 5 kW, p = 5 m Torr ($\approx$0,6 Pa) Ar, t_D = 3 min, D_r = 62mm D_S = 12 mm.

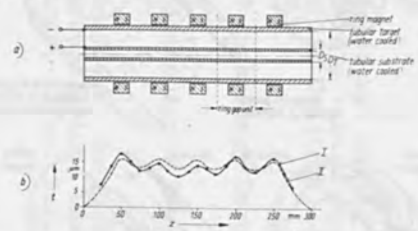


Fig. 6. Torus Plasmatron with five ring gap units. a) Geometrical arrangement. b) Thickness distribution t on the substrate along the axis z (I measured values ——, II geometrical superposition of five curves according to a single ring gap unit ---). Parameters: target material copper, P = 25 kW, p = 5 m Torr ($\approx$0,6 Pa) Ar, t_D = 3 min, D_T = 62 mm, D_S = 12 mm.

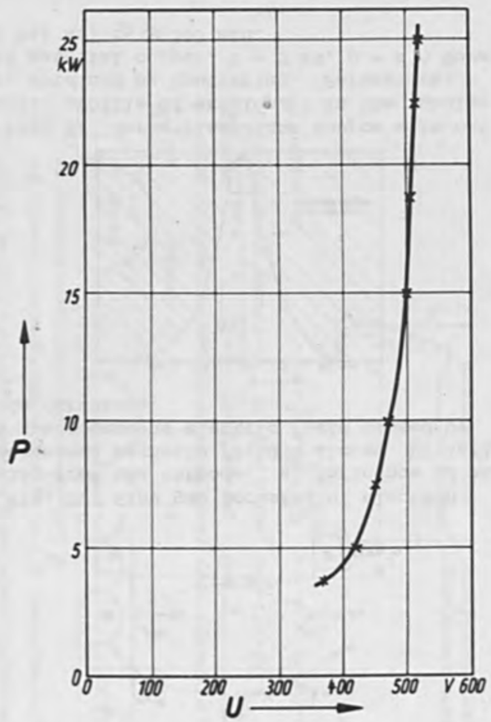


Fig. 7. Torus-Plasmatron sputter source. Power (P) - voltage (U) - characteristic of a Torus-Plasmatron with five ring gap units.

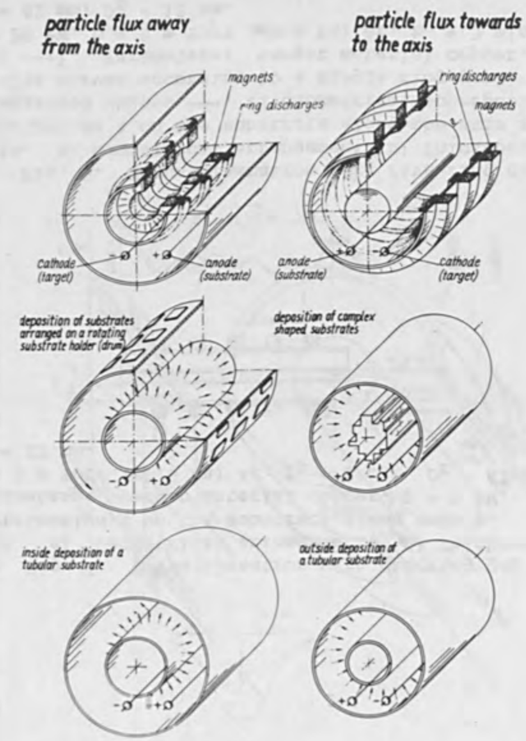


Fig. 8. Torus-Plasmatron arrangements with several ring gap units for sputtering with high local and high average deposition rate.

References

1. I.G. Kesajev, V.V. Paškov, Jurn. Techn. Fiz. 29 (1959) 3, p. 287 - 298

2. S. Schiller, U. Heisig, K. Goedicke: On the use of ring gap discharges for high-rate vacuum coating. Proc. 3rd Int. Conf. on Metallic Coatings, J.Vac. Sci. Technol. 13 (1976) Sept.

3. Anonym: Documentation, SLOAN Technology Corp., Santa Barbara (Ca.) 1972

4. Anonym: Physical Vapor Deposition. Airco Temescal Publ. 1976

5. U. Heisig, S. Schiller, K. Goedicke: Using the ring gap Plasmatron for high-rate sputtering. Proc. 3rd Int. Conf. on Metallic Coatings, J.Vac. Sci. Technol. 13 (1976) Sept.

6. J.A. Thornton, V.L. Hedgroth: Tubular hollow cathode sputtering onto substrates of complex shape. J.Vac. Sci. Technol. 12 (1975) 1, 93 - 97

7. L. Holland: Vacuum deposition of thin films. Chapman and Hall Ltd. London 1961, 154

8. J. Chapin: Documentation, Vacuum Technology Assoc. Inc. (VTA), Boulder (Cal.) 1974

Section II A

Low Power Electron Beams

Miscelaneous Topics

HIGH RESOLUTION ELECTRON GUN WITH NOVEL FOCUS AND DEFLECTION SYSTEM

Holger M. Luther
Carson-Alexiou, Inc.
20 Oak Street
Westport, Ct. 06880

ABSTRACT

An electron gun for high density mass memory systems as well as electronic camera systems is described. The gun consists of an electron source with a brightness of $5 \times 10^4 A/cm^2/sr$ and an electrostatic focussing and deflection structure capable of deflecting the beam up to 10^4 spot diameters with off-axis resolution losses primarily determined by the type of deflector used. The electron optical column utilizes superimposed focussing and deflecting electrodes arranged such that the electron beam crosses the image plane at normal angle of incidence and at very low energies ($20 < E < 150eV$). Computer solutions to the paraxial ray equations are compared with experimental results. Estimates of $C_s(\infty)$ are given. Two deflection systems are compared, and experimental results are shown to demonstrate the claim that the off-axis performance of the gun is limited by aberrations characteristic of the particular deflection system used in the E-O column.

Introduction

An electron gun has been developed at CBS Laboratories which represents a significant departure from electron guns used in most electronic camera and electron beam addressed mass memory systems in use today. Similar to the Return Beam Vidicon (RBV) column, this gun extracts a return beam of secondary electrons from the target; but unlike the RBV column, it needs no retarding-accelerating mesh in front of the target. The absence of the mesh allows much finer beam probes than possible from RBV columns which are resolution-limited by the mesh aperture. Another departure from the conventional electron guns is the superposition of the electrostatic focussing and deflection fields. This superposition eliminates the dominant anistropic lens aberrations introduced by a pre-lens deflection system. Thus the limit of the probe resolution is determined by the spherical aberration of the lens and the deflection aberration characteristic of the deflecting system.

Figure 1 shows a schematic of the gun. Electrons from a source aperture are aimed into a magnetic field which deflects the beam into the center of the first lens. This lens is an equidiameter two-tube lens formed by the end of the drift tube and the end of the four-fold deflector. The function of this lens is to collimate the incident electron beam. The second lens parameters are adjusted such that the deflection center of the deflector is coincident with the object-space focal point of this lens. The effect of this adjustment is a focussed beam of electrons incident normally upon the target plane, with an energy determined by the image space potential.

The optics is complicated by the fact that there is a third lens field near the target which is formed by the second lens electrode and the target plane. This field creates a departure from normal incidence near the edge of the target. There the electron beam is bent sharply away from the center of the target. This departure from normality may be corrected with the aid of a ring electrode to which is applied a small negative potential with respect to the target. Secondary electrons ejected from the target plane are accelerated by the second lens electrode into the lens field whence they retrace the trajectory of the probe until they reach the magnetic field. From there they are relayed into the entrance aperture of the return beam energy analyzer.

Analysis of Gun

Initial analytical efforts in the development of this electron gun had been limited to the prediction of the location of the image plane from the published properties of two-tube equidiameter electron lenses. Experimentally it soon became apparent that this analysis was incomplete, since it did not insure the coincidence of the deflection center with the upstream focal point of the second lens. Neither did this analysis compute the effect of the third lens field on the location of the image plane. As a result some electron optical columns were constructed which could not simultaneously satisfy the conditions of normal incidence and sharp focus.

This situation was improved with the computer generated solutions of the generalized, paraxial ray equation shown in Figure 2. The axial potential assumed for the lens field is composed of a superposition of lens fields of two-tube equidiameter lenses plus a mirror field at the target given by the last two terms of the expression for $\phi(Z)$ in Figure 2. The deflection field $E(Z)$ was broken up into three separate regions as shown in Figure 3. Only first order terms of the fields were considered.

A ray trace program based on the Runge-Kutta solution of the paraxial ray equation generated the gaussian trajectories of selected electron rays. These are shown in Figure 4. The solutions to the ray equation clearly show that both normality of incidence and sharp focus may be obtained simultaneously. A deflection sensitivity of about 400 V/cm predicted by the computer solution was verified experimentally within 10 per cent. The ray trace analysis also yielded the gaussian magnification which was shown to be 0.28. Edge gradient analysis of the focussed beam indicated a beam diameter at the 2σ point of 1 μm, while the divergence angle of the unfocussed beam was determined to be $1\text{x}10^{-3}$ rads at the 2σ point. The coefficient of spherical aberration $C_S(M)$ computed from these parameters was found to be about $8.8\text{x}10^4$ cm, which agrees well with the experimentally determined magnitude of $8.72\text{x}10^4$ cm obtained from a Hartman test.

Off-Axis Aberrations

Previous experiments on electron guns with lens deflectors had shown that the off-axis probe resolution was limited only by the deflection aberration characteristic of the deflector. This was demonstrated by installing first a standard fourfold parallel plate deflector into the column and measuring its off-axis resolution. Then an eightfold deflector shown in Figure 5 was installed. The eightfold electrostatic deflector was developed by John Kelley at the Stanford Research Institute. He established analytically as well as experimentally that this deflecting geometry eliminates deflection coma and reduces astigmatism. Consequently any coma in the beam from the column using the eightfold deflector would have to be due to the focussing system.

On- as well as off-axis resolution (MTF) measurements were made with the aid of edge gradient analysis. The results of these measurements are shown in Figure 6. The data for each set of measurements was normalized with respect to the on-axis beam radius taken to be the 1σ width of the gaussian current distribution. A best fit analysis of these sets of data showed that the relative increase in the beam radius $\Delta\sigma/\sigma_{OA}$ for the fourfold deflector is given by:

$$\Delta\sigma/\sigma_{OA} = A\theta + B\theta^2 + O\theta^3$$

in which θ is the deflection angle and the coefficient $O \ll A$ or B over the range of θ considered. The data for the eightfold deflector, however, suggest that $\Delta\sigma/\sigma_{OA}$ is related to θ by the equation:

$$\Delta\sigma/\sigma_{OA} = (B-A)\theta^2 + O'\theta^3$$

in which as before $O' \ll (B-A)$. Since coma is related to the first power of θ, it is clear that the off-axis probe is only affected by the deflector aberration, and not those of the lens.

The Electron Source

The electron source used in this gun is shown in Figure 7. It is essentially the design of Bruck and Bricka[1], but with an important modification. Since the immersion optics is very sensitive to mechanical misalignment, the Wehnelt was quartered and a differential deflection voltage was applied to opposing quarters. This permits the operator to steer the beam over the source aperture located at the end of the anode. The gun was determined to have a brightness of about 6×10^4 A/cm^2/sr at 12KV. Recently a directly heated LaB_6 cathode installed in the source increased that brightness to about 2×10^5 A/cm^2/sr. With a limiting source aperture of about 3.5 μm diameter, the source divergence angle was found to be about 1×10^{-3} radians depending upon the beam current drawn. Figure 8 shows the beam current obtained from the LaB_6 emitter.

The Energy Analyzer

The return beam is focussed close to the entrance aperture of the electron energy analyzer. This analyzer is a modified version of a filter lens described by Simpson and Marton.[2] The energy resolution $\Delta E/E$ is estimated to be about 2×10^{-3}. Higher resolution is obtained by limiting the entrance aperture angle to below 10 mradians. Figure 9 shows the collector current I_c given by:

$$I_C = \int_{eV_T}^{\infty} e\left(\frac{dN}{dE}\right) dE$$

as a function of the target potential V_T. N is the number of secondary electrons ejected from the target per unit energy per second. Differentiation of the curve in Figure 9 yields the usual energy distribution of secondary electrons emitted from a metallic surface.

Visual Data

Figure 10 is a photograph of the TV monitor display of the entire target area. Tool marks on the metallic surface are readily discernible. The 0.9 x 1.75 cm silicon chip in the center of the target delineates the quality area of the field of view. The chip is covered with a 200 Å layer of SiO_2 into which three L-bar resolution wedges have been etched. These are not resolved in the picture, as that level of detail is beyond the resolution capability of the monitor.

Figure 11 is a view of a portion of a different resolution wedge of photo-etched nickel. The smallest vertical and horizontal lines near the center of the target are 2 μm wide. The sun figure in the center is used to test the electron beam for astigmatism and/or coma. The slight vertical contrast gradient seen in the picture is due to a minor misadjustment of the magnetic separator coil current.

Applications

There are many possible applications for the CBS return beam electron gun ranging from high resolution storage camera systems to surface analysis using very low energy electron probes.

The gun has the further virtue that a simple change of the lens potentials V_2, V_3, and V_4 allows the focussed beam to land on the target with an energy of 5 KeV. It was found that in this mode of operation the gaussian magnification was smaller, and the deflection sensitivity was lower by about 10%. Both the return beam information and the normality of incidence on the target were lost however.

The above feature of the gun is being used to advantage in a fast access beam addressed memory tube. Other applications envisioned include Auger spectroscopy of selected target sites previously identified visually using the gun in the low energy mode of operation.

A field emission source has recently been designed by the author which will increase the resolution capability of the return beam gun into the submicron range. This should further extend the range of applications of this device.

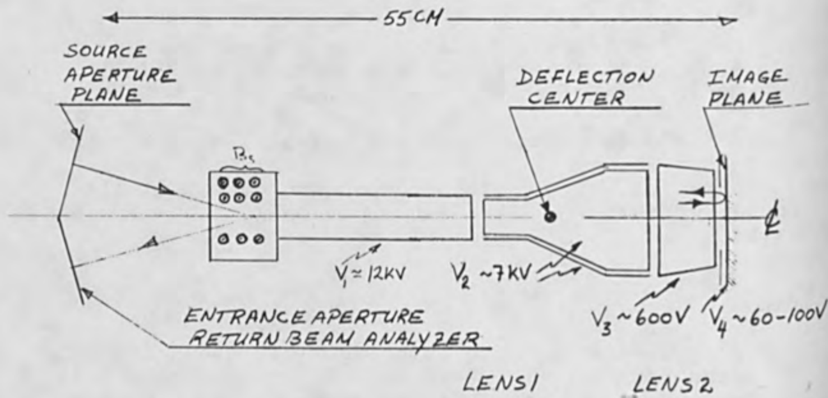
Acknowledgement

I am pleased to cite both Dr. W. E. Glenn and Dr. Hsing Yao Chen for their significant contributions to the development of this electron gun.

References

1. M. Bricka, H. Bruck, Ann. de Radioelectricite, 3, 339, 1948.

2. J. A. Simpson, L. Marton, Rev. Sci. Inst., 32, 802, 1961.



CBS ELECTROSTATIC RETURN BEAM ELECTRON GUN

FIGURE 1

$$1.\quad Y_{zz} + \frac{\phi}{2\phi} Y_z + \frac{\phi_{zz}}{4\phi} Y + \frac{E(z)}{2\phi} = 0$$

$$2.\quad \phi(z) = \frac{V_1}{2}\Big\{ 1 - \xi_{10} + \frac{v_1}{2}(1+\xi_{10})(1-\xi_{23}) + \frac{v_1 v_2}{2}(1+\xi_{23})(1-\xi_{34}) + v_1 v_2 v_3 (1-\xi_{34}^2) + v_1 v_2 (1-v_3)(1+\xi_{34}) \Big\}$$

with $\xi_{ij} = \tanh \frac{z - z_j}{\alpha_i R_i}$ $\qquad j \in \{0,3,4\}$

and $v_i = V_{i+1}/V_i$ $\qquad i \in \{1,2,3\}$

1. GENERALIZED PARAXIAL RAY EQUATION
2. AXIAL POTENTIAL DISTRIBUTION

FIGURE 2

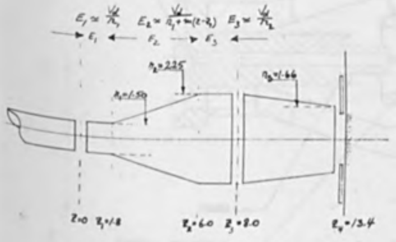
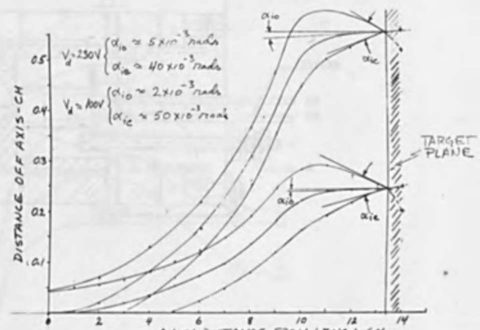
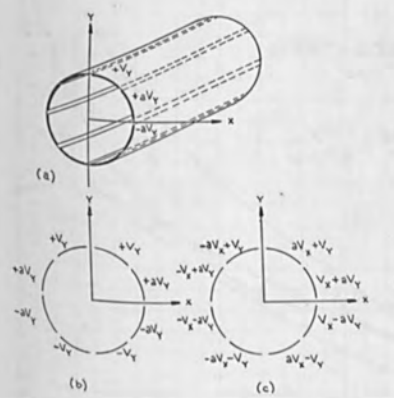


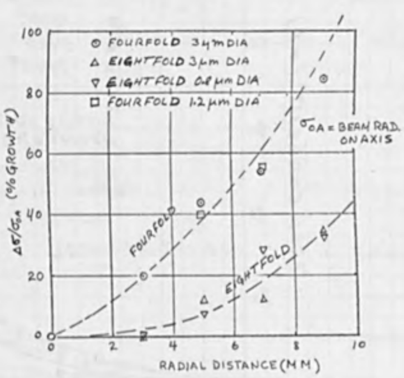
FIGURE 3



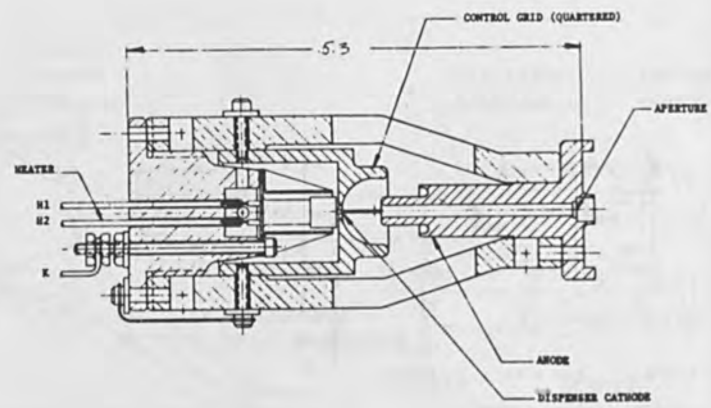
ELECTRON TRAJECTORIES (GAUSSIAN)
FIGURE 4



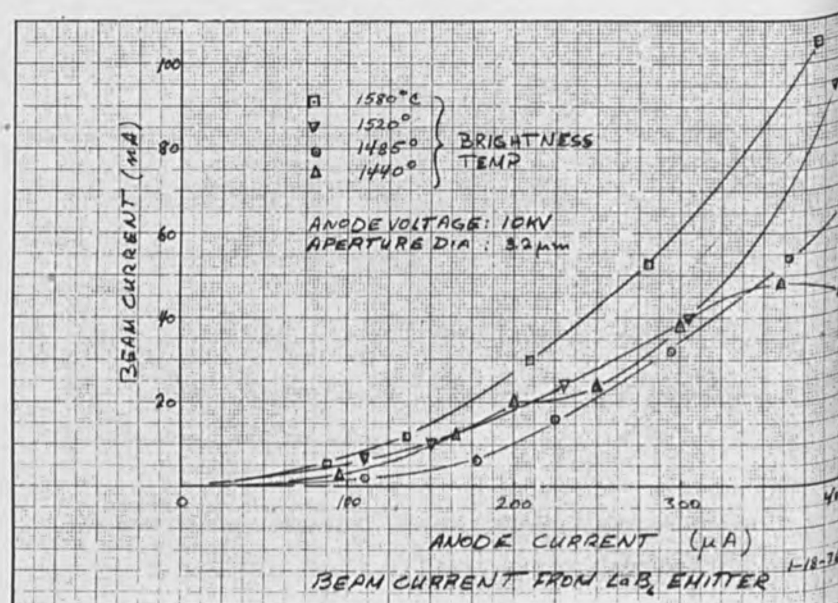
OCTUPOLE DEFLECTOR CONFIGURATION AND DEFLECTION VOLTAGE DISTRIBUTION
a) Octupole b) Vertical Scan only
c) Vertical and Horizontal Scan
FIGURE 5



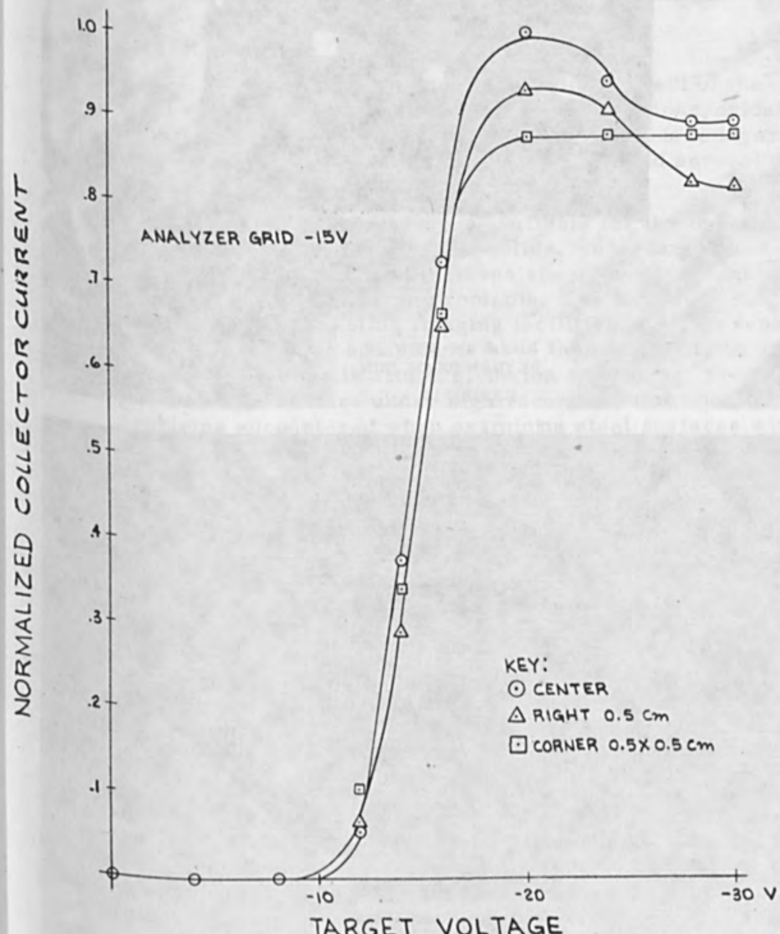
OFF-AXIS RESOLUTION
FOURFOLD VS EIGHTFOLD
FIGURE 6



ELECTRON SOURCE FOR CBS
RETURN BEAM ELECTRON GUN
FIGURE 7



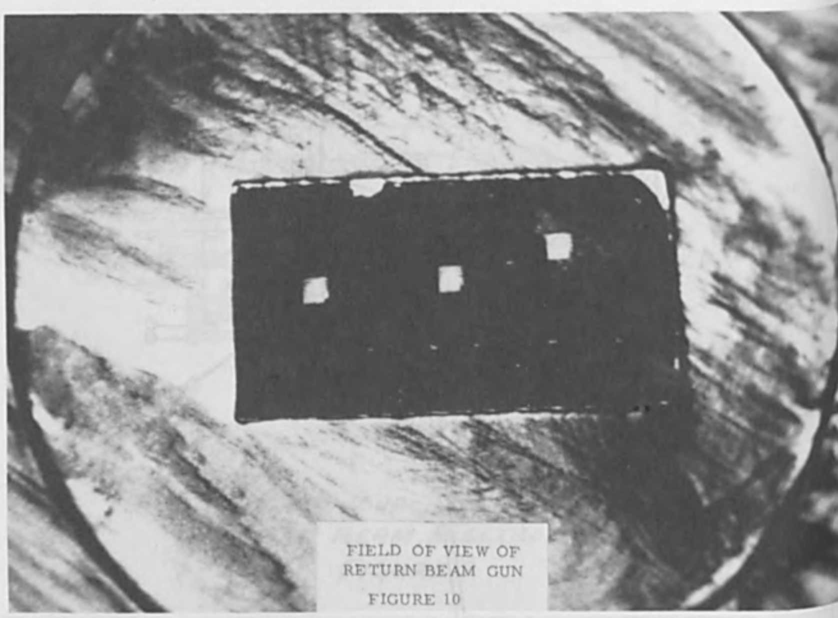
BEAM CURRENT FROM LaB_6 EMITTER
FIGURE 8



ENERGY ANALYZER RESPONSE

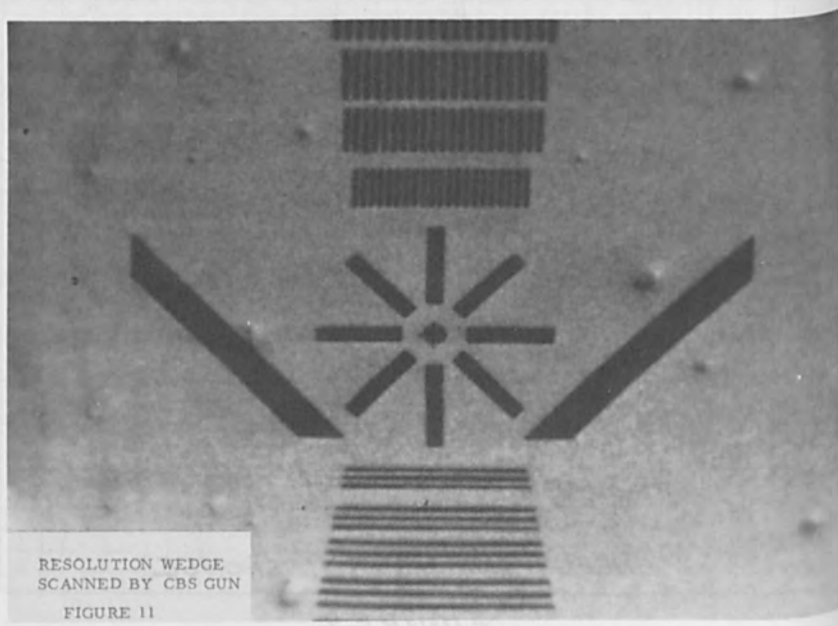
FIGURE 9

11-27-73



FIELD OF VIEW OF
RETURN BEAM GUN

FIGURE 10



RESOLUTION WEDGE
SCANNED BY CBS GUN

FIGURE 11

SURFACE ANALYSIS OF STEEL BY SECONDARY IONS AND AUGER ELECTRONS SPECTROSCOPIES.

J.P. SERVAIS, C. EMOND, J. RICHELMI, V. LEROY
Centre de Recherches Métallurgiques
Abbaye du Val Benoît, 4000 Liège, Belgium

ABSTRACT.

Many problems encountered in the field of the steel surface, e.g. thermal induced segregations, oxidation adherence of protective coatings, concern surface layers whose depths are included between one ten and several thousands Angströms.

Among the techniques available for the investigation of the outermost layers of solids, secondary ions mass spectroscopy and Auger electrons spectroscopy seem to be very useful to solve these problems. The ion probe microanalyser has interesting imaging facilities and high sensitivity. Most Auger equipments have the possibility to prepare the specimen in situ e.g. by ion sputtering, thermal processing, fracture under high vacuum. Some analytical problems encountered when examining steel surfaces with these techniques and examples of applications to the study of the modification in surface composition that are induced by thermal processing are presented in this paper.

ION PROBE MICROANALYSER.

The ion probe microanalyser, issued from Castaing and Slodzian's works (1, 2) combines a mass spectrometer with an ion microprobe (fig. 1). The mass spectrometer has special stigmatism properties in both radial and transverse directions which provide a direct imaging facility. With this instrument it is possible to observe the distribution map of the various ions emitted from the specimen surface or to record the secondary ions intensities versus the time thus as a function of the penetration in the specimen.

Use of the ion microprobe mass analyser for the study of surfaces is dependent on the control of the two basic phenomena involved in the technique, viz that of cathodic sputtering which successively exposes deeper and deeper atomic layers, and that of secondary ion emission which makes it possible to analyse the elements contained in the target by means of the mass spectrometer. These phenomena are very complex, and there is actually no adequate general theory of the interactions between ion beams and solids (3).

However, the many studies carried out up to now have led to identifying the fundamental features of these interactions :

- these phenomena are anisotropic (4). The sputtering rate and the secondary ion emission yield (number of secondary ions of a given species emitted as a result of the impact of one primary ion) vary with the orientation of the crystal lattice with respect to the primary ion beam. For instance, the scanning electron microscope image of figure 2.a shows differences in level between the grains of a poly-crystalline Fe - 35 Ni - 8 Cr alloy after bombardment with an argon ion beam. Variations of the secondary ion yield from grain to grain show up on the $^{52}Cr^+$ ion image (fig. 2.b) supplied by the ion probe analyser
- two types of secondary ion emission are generally distinguished (5). Some atom or groups of atoms which are sputtered away from a metallic target bombarded by ions of an inert gas are in so highly excited electronic states that their de-excitation outside the target leads to the formation of secondary ions. The yield of this mode of ion emission, designated kinetic emission, is generally low because the emission is related to the probability of finding sufficiently excited atoms in the immediate vicinity of the target.

Chemical emission occurs when the atomic bonds within the target are of a strongly ionic nature. In this case, the break of the ionic bonds by the incident particles gives rise to the direct emission of positive and negative ions. The ionization yield is much larger in the process of chemical emission than in that of kinetic emission. Chemical emission is experienced when oxides, sulphides or silicates are bombarded.

The so-called chemical flash is also observed when the first few layers of a freshly prepared metallic target (which is covered with a layer of chemisorbed gases) are sputtered (6). For instance, the presence of small chromium rich oxide inclusions in the Fe - 35 Ni - 8 Cr alloy of figure 2.b gives rise to very bright spots on the $^{52}Cr^+$ ion image. On the other hand, the micrograph (fig. 2.a) shows that these inclusions have been loosened from the metallic matrix, which indicates a difference between the sputtering rates of the oxide and the metallic matrix.

Analysis of the metallurgical samples with extensive oxide scales is highly unreliable due to the influence of the type of chemical bond on the mechanisms of secondary ion emission and of cathodic sputtering.

However, metallic targets can be caused to produce chemical emissions by bombarding them with ions of reactive gases (O_2^+, O^-) (6), or else by admitting in the vicinity of the target, a reactive gas such as oxygen at a pressure comprised between 10^{-6} and 10^{-3} torr (7, 8).

In the latter situation, a layer of chemisorbed gas with a strongly ionic character forms continuously on the surface; it results from an equilibrium between chemisorption and sputtering due to the primary ion beam, its thickness being a direct function of gas pressure and an inverse function of sputtering rate. The secondary ion emission then assumes the features of chemical emission.

The conditions of penetration of the primary ion beam into the target are modified to a large extent by the chemisorbed oxygen layer, which exerts a screening effect on the primary ions. This considerably reduces the anisotropy of secondary ion emission and cathodic sputtering (fig. 2. c and d).

When the oxygen pressure near the target reaches 10^{-5} torr, the sputtering rates for a metal and its oxide become practically equal (fig. 3); furthermore, the ion yields no longer depend on the nature of the chemical bonds in the sample. The same calibration relationships can then be used for the analytical determination of metals in both alloys and oxides (fig. 4). By suitably adjusting the oxygen pressure in the target chamber and the conditions of primary bombardment both oxide layers and metallic alloys can be analysed quantitatively (9).

As an example, table I lists the results of chemical and ion microprobe mass analyser determination of the composition of a series of homogeneized steels. Good agreement is observed between these results.

AUGER ELECTRON SPECTROSCOPY.

Auger electron spectroscopy is a technique intended for analysing solid surfaces, which rests on the following principles. The surface to be analysed is placed in an ultra-high-vacuum chamber and bombarded with an electron beam with an energy of about 2 keV. Under the impact of the electrons, the atoms of which the surface is made are excited. They can de-excite according to two different but complementary processes either by emitting photons whose energy is characterized by the chemical nature of the emitting species, or by emitting so-called "Auger electrons" whose kinetic energy also depends on the nature of the emitter. An electron spectrometer analyses the energy distribution N(E) of the electrons emitted by the surface. The results, presented as the variation of

$$\frac{d\,N(E)}{dE}$$

(the derivative of the distribution N(E) with respect to the energy) with E give a spectrum of lines which characterizes the nature of the elements present at the surface. As a first approximation, the peak to peak amplitude of a line is proportional to the atomic concentration at the surface.

Table I. Composition of steels as supplied by ion microprobe mass analyser (in wt %)
$I_p = 100$ nA, $S_o = 500 \times 500 \mu$ m, $P_{O_2} = 4.10^{-5}$ torr.

		Mn	Si	Cr	Ni	Mo	Fe
AISI 430	chemical	0.46	0.48	16.2	-	-	82.4
	IMMA	0.48	0.32	15.9	-	-	83.3
AISI 304	chemical	1.55	0.48	18.48	9.92	-	69.56
	IMMA	2.12	0.46	18.35	8.6	-	70.65
AISI 316	chemical	0.95	0.56	17.4	13.4	2.96	64.8
	IMMA	0.91	0.58	16.6	12.0	2.6	62.8
I 810	chemical	1.8	0.87	22.3	13.1	-	61.9
	IMMA	2.4	0.88	23.0	11.7	-	61.9
I 812	chemical	0.38	0.39	13.2	0.21	-	85.8
	IMMA	0.39	0.41	16.1	0.27	-	82.8

Within the range of energies involved in this technique, the yield of Auger emission i. e., the number of atoms which de-excite by emitting Auger electrons per 100 excited atoms, is distinctly larger than the yield of X rays emission (10). Therefore, the sensitivity of the method is high (0.01 mono-atomic layer in practice); moreover the method is specific i. e. the informations supplied concern only the first five atomic layers of the surface; as a matter of fact, the mean free path of electron with energies ranging from 500 to 1000 eV is of the order of 5 atom layers (11). As a result, Auger emission spectrometry calls for ultra-high vacuum, so as to avoid all contamination through adsorption of residual active gases. U.H.V. technology makes possible to reach vacuums of the order of 10^{-10} torr. In the equipment used in the present work, the pumping system comprised a turbomolecular pump, a liquid nitrogen cryopump and a titanium sublimation trap.

The impact of the primary electrons does not cause any detectable modification of the surface chemistry; therefore, the technique is not destructive. With a view to studying the variation of the chemical composition as a function of penetration into the first few surface layers, the analytical system was fitted with an ion gun which makes it possible to sputter the sample away at a controlled rate. Sputtering rates ranging from one to 30 angströms per minute can be achieved.

The U.H.V. chamber (fig. 5) including a cylindrical mirror analyser electron spectrometer, an ion gun, an electron gun for carrying out heat treatments in a vacuum or in a gas reduced pressure, and a quadripolar mass spectrometer for analysing residual gases present in the chamber or evolved during heat treatment of the sample. As an example, table II shows the composition of the residual atmosphere in the chamber during pumping down. In the case of the best vacuum obtained, the residual atmosphere consists mainly of carbon monoxide, carbon dioxide and hydrogen. The samples are mounted on a revolving holder which can be positioned under vacuum by remote control. The sample temperature is measured by means of Pt-Pt Rh thermocouples welded to the rear face of the samples.

A quantitative analysis of the chemical composition at the surface of a specimen can be obtained as a first approximation using a procedure developped by Palmberg (12).

Spectra of the Auger electrons emitted by pure elements recorded under identical experimental conditions, are first collected. The atomic concentration C_i of the element i at the surface of the sample that is being analysed is given by the relationship

$$C_i = k \frac{I^e_{i,j}}{I^t_{i,j}}$$

where $I^e_{i,j}$ and $I^t_{i,j}$ are the peak to peak amplitudes of the Auger line j for element i in the sample and in the blank.

The fact that the secondary electron spectra from the sample and the blank can be recorded using different sensitivities is taken into account by means of the factor k.

When several elements i are present at the surface of the sample, one can write :

$$C_i = \frac{I^e_{i,j} / I^t_{i,j}}{\sum_i I^e_{i,j} / I^t_{i,j}}$$

since $\sum_i C_i = 1$.

The validity of this method is illustrated in table III, which lists the analyses, obtained by chemical means and by Auger electron spectrometry of a series of samples which electron microprobe examination had demonstrated to be homogeneous.

APPLICATIONS TO THE STUDY OF SURFACE SEGREGATIONS INDUCED IN COLD ROLLED STEEL SHEETS BY THERMAL PROCESSING.

A given structural condition and hence suitable properties can be conferred upon cold-rolled steel strip by means of appropriate heat treatments. However, it is currently recognized that the influence of such treatments on the strip surface is far from negligible.

As a matter of fact, on heat treating, the material is subjected to conditions of temperature and chemical environment at which chemical reactivity and atomic diffusivity are very large, so that a certain number of physico-chemical reactions can take place at the surface of the steel strip. These reactions are the consequence of the general tendency of all systems (here the steel-environment system) to lower its internal energy by means of physico-chemical reactions at the interfaces.

Selective oxidation reactions can occur even when protective atmosphere with very low dew points are used. Similarly, secondary reactions can take place between the steel and external agents such as rolling-mill oils and coolants or degreasing solutions used at different cold transformation stages. Finally, even when such reactions do not occur, there is a possibility that certain alloying elements or impurities will migrate to the surface where their concentration will lower the surface energy of the ferritic matrix. The simplest case of such surface segregations is that of mild steels previously cleaned by ion etching and submitted to a heat treatment in the U.H.V. environment of the Auger spectrometer system. This made it possible to analyse the sample surfaces without exposing them to air after the annealing treatment.

Table II. Total and partial pressures (in torr) and contents (in %) of the major components of the atmosphere present at the successive stages of pumping down, preheating of the chamber 1 1/2 hour at 400° C.

		after preheating	after introducing liquid N_2	after titanium sublimation
CO_2	torr	$1.0\ 10^{-10}$	$2.4\ 10^{-11}$	$2.4\ 10^{-11}$
	%	3	1.2	10.3
CO	torr	$4.6\ 10^{-10}$	$1.3\ 10^{-10}$	$1.1\ 10^{-10}$
	%	12.7	6.3	47.1
N_2	torr	$6.9\ 10^{-10}$	$2.1\ 10^{-11}$	$9.7\ 10^{-12}$
	%	19	1	4.1
O_2	torr	$1.8\ 10^{-12}$		
	%	0.05		
H_2	torr	$2.3\ 10^{-9}$	$1.8\ 10^{-9}$	$8.9\ 10^{-11}$
	%	64.1	91	37.6
Ar	torr	$1.3\ 10^{-11}$	$5\ \ 10^{-12}$	-
	%	0.4	0.3	-
H_2O	torr	$3.3\ 10^{-11}$	$3.1\ 10^{-12}$	$2.2\ 10^{-12}$
	%	0.9	0.2	0.9
total pressure calculated		$3.6\ 10^{-9}$	$2\ \ 10^{-9}$	$2.4\ 10^{-10}$
measured		$2.5\ 10^{-9}$	$1.1\ 10^{-9}$	$3.2\ 10^{-10}$

Table III. Compositions as supplied by chemical analysis and Auger spectrometry (in at. %)

sample n°	method	Fe	C	N	P	S	Cr	Ni	V	Nb
1005	chemical	77.9						22.1		
	AES	77.5						22.5		
1006	chemical	69.4					8.62	21.9		
	AES	72					10	18.5		
1007	chemical	62					16.1	21.9		
	AES	61.5					18.5	20		
1008	chemical	56.8					21.4	21.8		
	AES	57.5					24.5	18		
1009	chemical	52.7					25.6	21.7		
	AES	51					21	18		
VN	chemical			50					50	
	AES			51					49	
FeS_2	chemical	33				67				
	AES	41				59				
Fe_3P	chemical	75			25					
	AES	79			21					
VC	chemical		50						50	
	AES		50						50	
NbC	chemical		50							50
	AES		45							55

Figure 6 shows a series of electron spectra recorded after ion bombardment cleaning and after heat treatment for 30 minutes at 550° C in vacuum; these results apply to an aluminium killed steel and a rimmed steel. After ion-bombardment cleaning (fig. 6. a), the samples surfaces are still slightly contaminated with oxygen as well as with some argon picked up during the cleaning operation.

After heat treating, traces of chromium and manganese are detected as well as considerable enrichment of the surface with elements present at very low concentrations or as impurities in the bulk of the steel samples. In the case of the rimmed steel (fig. 6. c), the surface is enriched in nitrogen, phosphorus and sulphur, in that of the killed steel (fig. 6. b), phosphorus, carbon and sulphur are detected; the shape of the Auger line for carbon shows that this element is present as graphite.

The nature and the amount of the elements which have segregated to the surface depends not only on the type of steel but also on the temperature and the heating rate of the annealing treatments (13, 14) It should be noted that these surface segregations induced by heat treatment in an ultra-high-vacuum effect extremely thin layers (of the order of a few atom layers) except in the case of carbon.

Several mechanisms have been proposed to explain these surface segregations (15, 16, 17, 18). According to one of these known as the equilibrium segregation mechanism or Gibbs mechanism (15, 16), elements with low solubility in the ferrite decrease the free energy of the system significantly by migration to the interfaces and more particularly to the metal free surface. According to this mechanism, the segregated layers do not exceed a few atom layers.

In the case of carbon segregation, contrary to what occurs within the bulk of the metal where supersaturated ferrite decomposes into ferrite and cementite according to the metastable iron cementite diagram, at the surface steel behaves in accordance with the iron-graphite equilibrium diagram. In actual fact, the absence of neighbouring atoms at the free surface favours nucleation and growth of graphite which has a higher atomic volume than cementite.

Most of the industrial heat treatments applied on the cold rolled mild steels are performed in protective atmospheres, e. g. $H_2 + N_2$ mixture, with very low dew point.

As an example of surface segregations induced by such annealing the figure 7 shows distribution profiles for the secondary ion intensities vs depth of penetration beneath the surface. These distribution curves obtained by the ion microanalyser indicate that there is in fact two sub-layers viz :

- an outer sub-layer with a thickness of the order of 200 Å; This sub-layer is enriched in elements such as sodium, carbon, phosphorus, sulphur, silicon, oxygen, manganese and chromium. The apparent manganese and chromium concentration in it can reach 20 and 5% respectively
- an inner sub-layer about 1200 Å thick, which is enriched essentially in manganese, chromium and oxygen.

Fig. 8 shows some positive and negative ion images of the surface of this steel. From this figure, it appears that the segregated elements are present at the surface but mainly in the grain boundaries where they form some complex oxides. In a deeper layer (1200 Å) manganese and chromium are present as oxide in the grain boundaries. The nature, extent and lateral distribution of the surface segregations in both sub-layers mentioned above depent mainly on three factors : the annealing procedure, the bulk chemical composition of the steel and the surface preparation prior to annealing (13).

In the case of stainless steel which can be annealed in oxidizing atmospheres, the ion microanalyser is able to distinguish some different sub-layers in the thin oxide film formed during the heat treatment and to define their compositions (fig. 9).

An external sub-layer of about 200 Å in thickness contains iron, chromium and manganese oxide. Behind this external sub-layer, an important silicon enrichment occurs in a sublayers of 500Å in thickness. Due to the chromium migration into the external oxide layer, a chromium depleted zone is apparent in the metallic material. The particular silicon distribution may qualitatively explain the favourable effect of this element on the oxidation resistance of stainless steels. The silicon enrichment in the inner sublayer has a screening effect on the diffusion of the other ion species toward the external oxide sublayer (19).

CONCLUSIONS.

Quantitative analysis of steels surfaces (i. e. polycristalline materials usually covered by more or less thick oxide layers) by the ion microanalyser is made possible by introducing oxygen at a pressure of about 10^{-5} torr in the specimen chamber.

In the case of Auger electron spectrometry, we have used the method of Palmberg for quantitative analysis. The possibilities of these techniques have been illustrated by some examples issued from our studies on the thermal induced segregations on steels.

BIBLIOGRAPHY.

1. R. CASTAING, G. SLODZIAN
J. Microscopie, 1, 395, 1962.
2. J.M. ROUBEROLL, J. GUERME, P. DESCAMPS, J.P. DAGNOT
J.M. GUYON
Proc. 16th Annual Conf. Mass Spect. Pittsburgh, 216, 1968.
3. H.W. WERNER
Surface Science, 47, 301, 1975.
4. M. BERNHEIM
Thèse Paris 1972.
5. G. SLODZIAN, J.F. BERNHEIM
C.R. Acad. Sc. Paris, 263, B 1246, 1966.
6. C.A. ANDERSEN
J. Mass Spectrometry and Ion Physics, 2, 61, 1969.
7. M. BERNHEIM, G. SLODZIAN
Surface Science, 40, 169, 1973.
8. G. BLAISE, M. BERNHEIM
Surface Science, 47, 324, 1975.
9. V. LEROY, J.P. SERVAIS, L. HABRAKEN
Metallurgical Reports CRM, 35, 69, 1973.
10. H.E. BISHOP, J.C. RIVIERE
J. Appl. Phys. 40, 1740, 1969.
11. R.E. WEBER, A.L. JOHNSON
J. Appl. Phys., 40, 314, 1969.
12. P.W. PALMBERG, G.E. RIACH, R.E. WEBER, N.C. MAC DONALD
Handbook of Auger Electron Spectroscopy, Electronics Industries Inc. 1972.
13. J.P. SERVAIS, H. GRAAS, V. LEROY
Metallurgical Reports CRM, 44, 29, 1975.
14. V. LEROY, H. GRAAS, C. EMOND, L. HABRAKEN
Communication presented at "Les Journées d'Automne"Paris
29 Sept. to 2 Oct. 1974.
15. D. MC LEAN
Grain boundaries in metals, London 1957.
16. H.P. SEAH, E.D. HONDROS
Proc. Roy. Soc. Lond., A 335, 191, 1973.
17. J.W. CAHN, J.E. HILLIARD
Acta Met., 7, 219, 1959.
18. R.E. HANNEMAN, T.R. ANTHONY
Acta Met., 17, 1133, 1969.
19. C.D. SHUNGU
Thèse Université de Liège (Belgium) to be published.

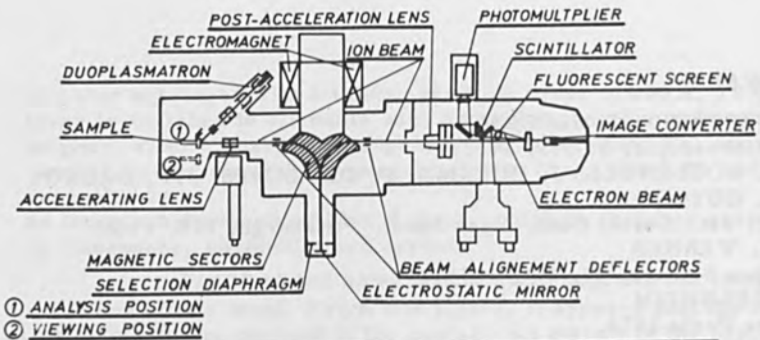


Fig. 1. Schematic diagram of the ion microanalyser.

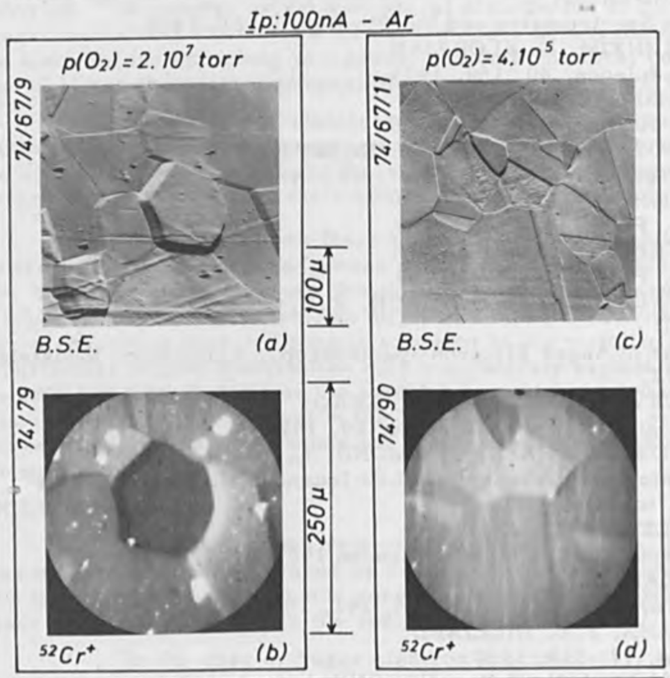


Fig. 2. Craters on an Fe - 35 Ni - 8 Cr alloy
a) Scanning electron microscope examination
b) Ion image of the chromium distribution
c) Scanning electron microscope examination
d) Ion image of the chromium distribution.

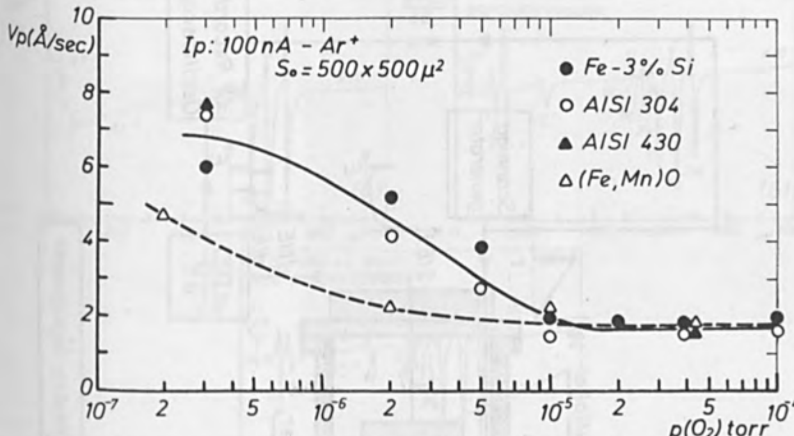


Fig. 3. Dependence of cathodic sputtering rate on the residual pressure in the case of ferrous alloys or oxides.

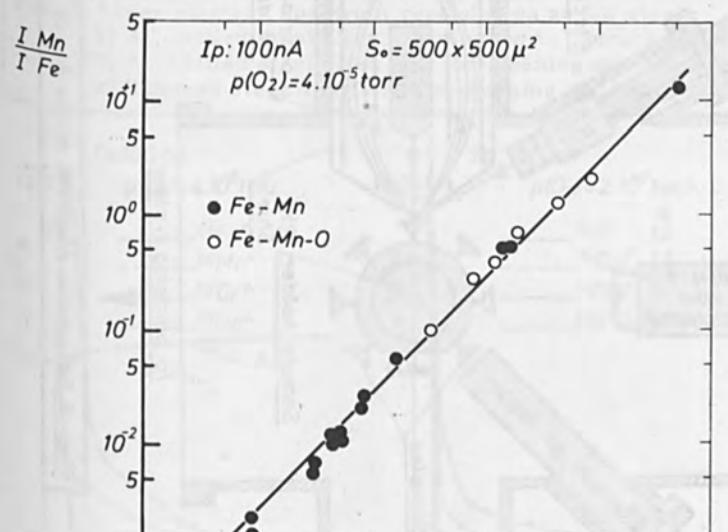


Fig. 4. Calibration curve for Fe - Mn and (Fe, Mn)O systems.

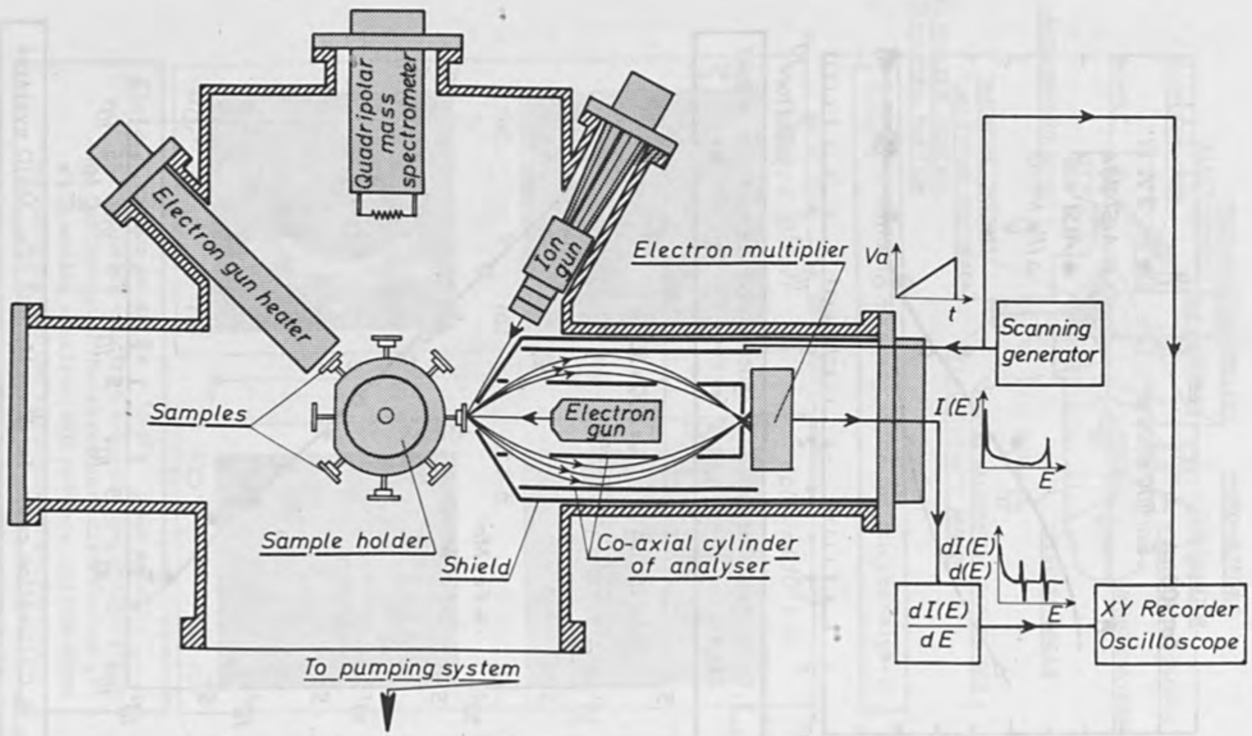


Fig. 5. Schematic diagram of the Auger spectrometry equipment

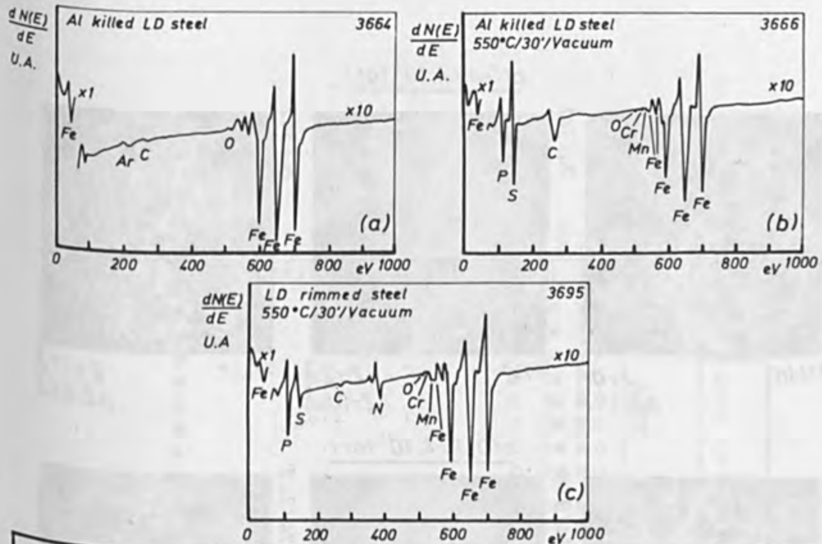


Fig. 6. Auger electron spectrum recorded on an LD steel
a) Al - killed steel after cleaning by Ar^+ bombardment
b) Al - killed steel after Ar^+ ion cleaning and heat treating
c) Rimmed steel after Ar^+ ion d eaning and heat treating.

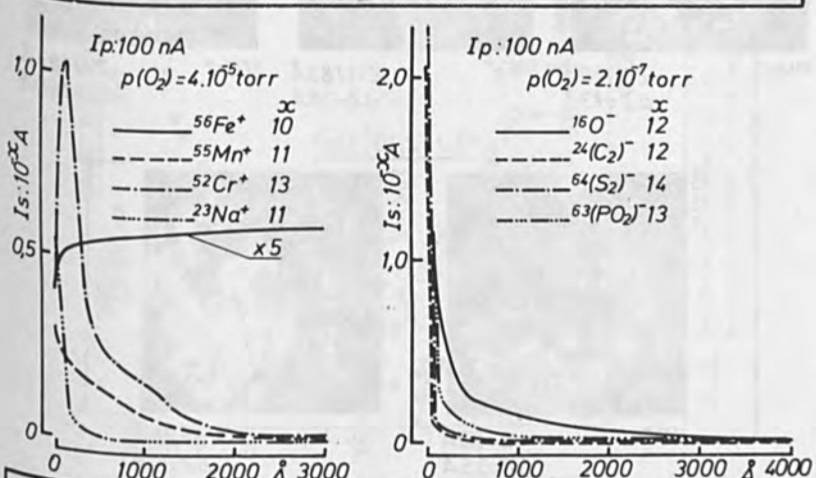


Fig. 7. Distribution profiles for the secondary ion intensities on a function of depth of penetration beneath the surface. Rimmed LD steel after open-coil annealing.

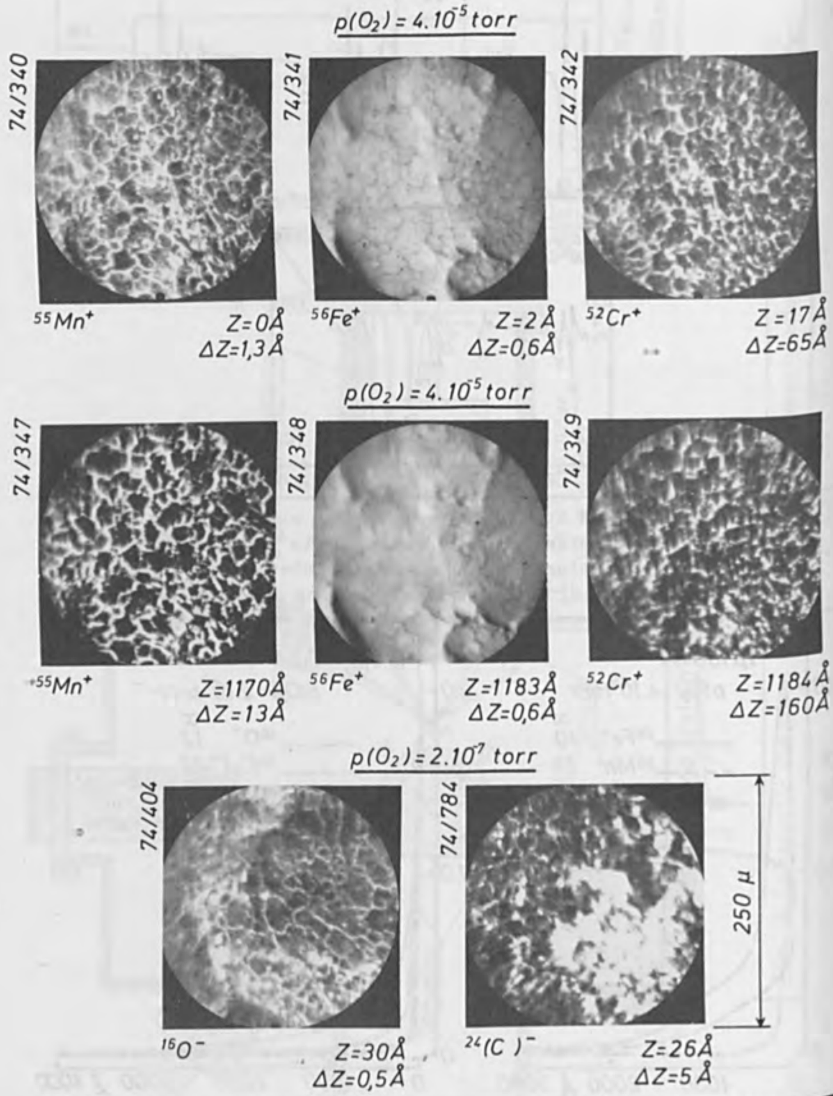


Fig. 8. Ion images showing the distribution of Fe, Mn, Cr, O and C in a LD rimmed steel, open-coil annealing.

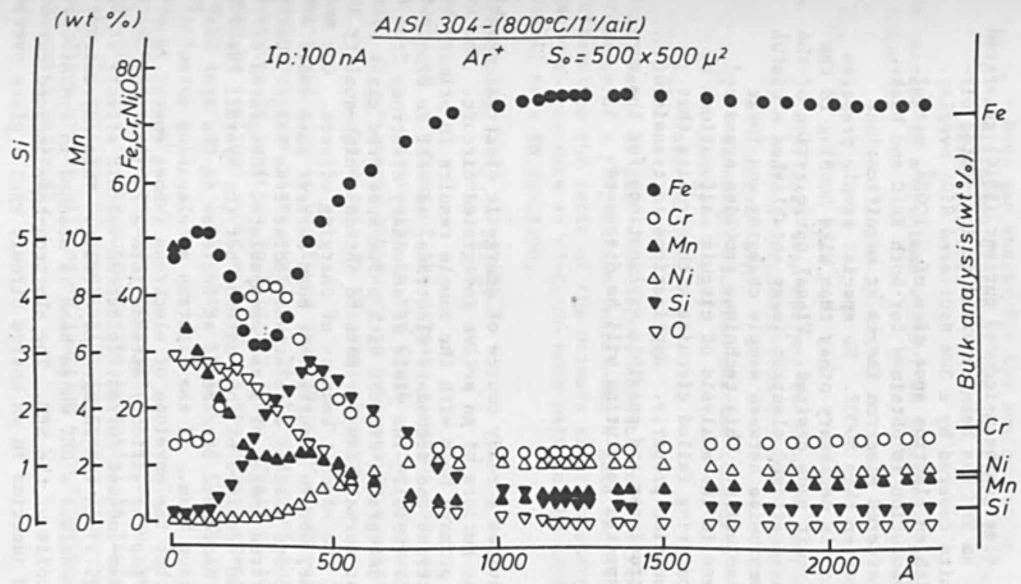


Fig. 9. In-depth analysis of an AISI 340 stainless steel, heat treated at 800° C during 1' in air.

SEM ELECTRON-BEAM-INDUCED CURRENT IN INTEGRATED CIRCUITS WITH A 3μm SPUTTERED SiO_2 OVERLAY

L. A. Kasprzak
IBM System Products Division, East Fishkill
Hopewell Junction, New York 12533

Electron-beam-induced current (EBIC) generated by an SEM has been observed in integrated circuits covered by a 3μm sputtered SiO_2 overlay. With an electron spot size of ≈ 1000Å, useful resolution is obtained for both EBIC and backscattered electron images at magnifications greater than 5000X. No special sample preparation is necessary other than wire bonding to the circuit pads desired. Visual optimization of the backscattered electron image establishes a useful compromise between sample charging and image resolution. This technique provides speed and ease in the analysis of circuit malfunctions by comparing failed circuits with circuits that function properly. Applications to transistor structures and possible explanations for the unexpected resolution will be discussed.

Introduction

The SEM provides a ready source of energetic electrons capable of penetrating the surface of an active integrated circuit. The interaction of the primary beam with the sample results in primary backscattered electrons and secondary electrons, as well as electron-beam-induced current (EBIC). The yield of secondary electrons from the surface of an integrated circuit with a 3μm sputtered glass overlay is insufficient at normal beam currents to obtain a high-quality image of the surface of the circuit because of charging effects. One must, by trial and error, obtain a sufficient beam current in a small enough spot size that one obtains a relatively noise-free image. Useful backscattered electron images of glass-encapsulated integrated circuits at 5000X have been obtained by this method, but the overall resolution of the system is decreased because of an increase in the spot size of the scanning electron beam. If the electron accelerating potential is increased such that the envelope of electrons (whose energy is greater than the band gap for silicon) extends into a PN junction, one can measure the beam-induced current (electron-voltaic effect) at the junctions (Fig. 1) in the integrated circuit. This current was used to intensity-modulate a CRT whose beam is scanned in synchronism with the scanning coils of the SEM. The electron-beam-induced current (EBIC) at a PN junction in an active circuit with a glass overlay is directly proportional to V_0 and i and inversely proportional to t, z, E_B, and d,

where
- V_0 = electron accelerating potential
- i = beam current
- t = thickness of overlay and thermal SiO_2
- z = atomic number of material above the junction
- E_B = band gap energy for the semiconductor
- d = diffusion depths of the junctions in the semiconductor.

The current in a specimen without PN junctions (say, for example, a bulk alloy) is

$$I_{SX} = I_B - I_{BS} - I_S \tag{1}$$

where
- I_B = primary beam current incident on sample
- I_{BS} = backscattered electron current (electron energy = eV_0)
- I_S = secondary electron current (electron energy < 1 keV)
- I_{SX} = specimen current.

A typical value of EBIC under these conditions is a few nanoamperes. The presence of a PN junction introduces a gain factor that is directly proportional to the ratio of the primary electron energy to the energy necessary to generate an electron hole pair in the semiconductor. Therefore, at a PN junction,

$$\text{EBIC} \propto (I_B - I_{BS} - I_S)\,\frac{eV}{E'_B} \tag{2}$$

where
- eV = energy of the primary electron when it penetrates to the depth of the PN junction
- E'_B = energy necessary to generate an electron hole pair in the semiconductor.

Experimentally, it was found that 20-30 kV electrons can penetrate 2-3μm of sputtered glass overlay. If the energy of the primary electrons at the PN junction is 1-10 keV, then a gain of 10^3 to 10^4 in EBIC can be realized. For example, if the beam current of energy eV at the PN junction is 1 nA, then the EBIC at the PN junction can be 1-10μA.

The effect of electron accelerating potential on the EBIC can be seen in Fig. 2. As the accelerating potential is increased from 20 to 30 kV the envelope of penetration increases from the silicon surface to the bottom of the silicon PN junction (4μm from the sputtered glass overlay outer surface).

EBIC in Devices with a Glass Overlay

A simple test circuit (Fig. 3) was examined in the EBIC mode (Fig. 4). In the center of Fig. 4 is a 50X backscattered electron image of the circuit under investigation. The detection circuit for the EBIC images can be seen in Fig. 5.

The following junctions are clearly delineated when EBIC is collected from the indicated pads.

Pad #	Junctions Delineated
1	CB of Tx A
2	EB & CB of Tx A
3	NP^+ & NR (880)
4	NP^+ & NR (880)
5	3NR (880) & EB all Tx's
6	EB & CB of group Tx's
7	P^+N
8	CB of group Tx's & NR on 7 resistors
9	3NR (880) & EB all Tx's
10	CB of Tx B & NR (77)
11	CB of Tx B & NR (77)
12	CB & EB of Tx B

N = N bed
C = collector
B = base
E = emitter
P^+ = P^+ isolation
R = resistor
Tx = transistor

A circuit with an N bed to P^+ short is shown in Fig. 6. A circuit with an EB short on Tx B, as well as an NP^+ short, is shown in Fig. 7. Experimentally, excellent EBIC image quality was obtained at magnifications up to 5000X (Fig. 8).

Applications and Future Work

The applications of this technique are rather extensive. First, however, high-quality circuits must be studied to fully understand the contrast effect caused by EBIC. Some of the applications are as follows:

- Circuit diagnostics--locating the cause of a circuit malfunction.
- Real-time circuit functional analysis by use of chopped-beam stroboscopy.
- Study of thin metal films--use of a shallow PN junction as a transmitted electron detector.

Caution must be exercised to prevent the introduction of noise, signal distortion, or artifacts peculiar to preparation or observation technique. For example, contacts must be reliable lest spurious signals be interpreted as circuit malfunction. For proper perspective one must incline the specimen surface at 90° to the electron beam axis. Otherwise the delineated junctions are displaced from the expected positions relative to surface topography. Finally, amplification of the EBIC prior to intensity modulation of the CRT is most critical with the glass overlay. Further study of EBIC is necessary, especially in the depth of penetration and envelope shape of primary electrons, and in damage to the specimen by the electron beam. These questions are complex because of the multilayer nature of the structure through which the electron beam must penetrate. However, as demonstrated, the technique can be utilized on an empirical data base.

Acknowledgment

James Greco deserves special thanks for his help with sample preparation.

Reference

1. P. R. Thorton, Scanning Electron Microscopy (Chapman and Hall, Ltd., London, 1968).

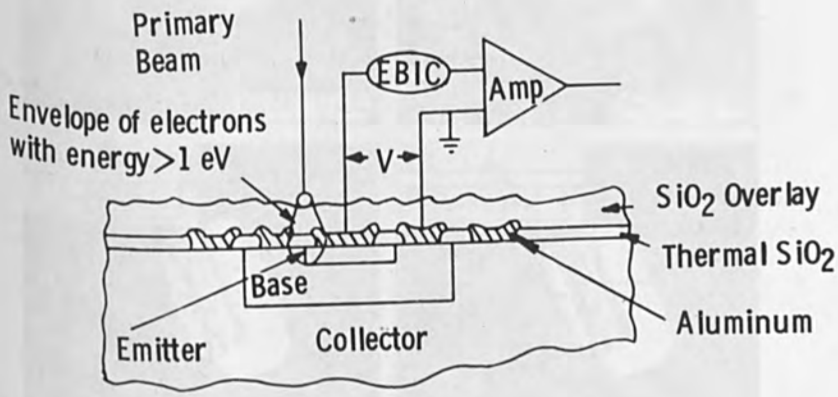


Fig. 1. Electron-beam-induced current at PN junction.

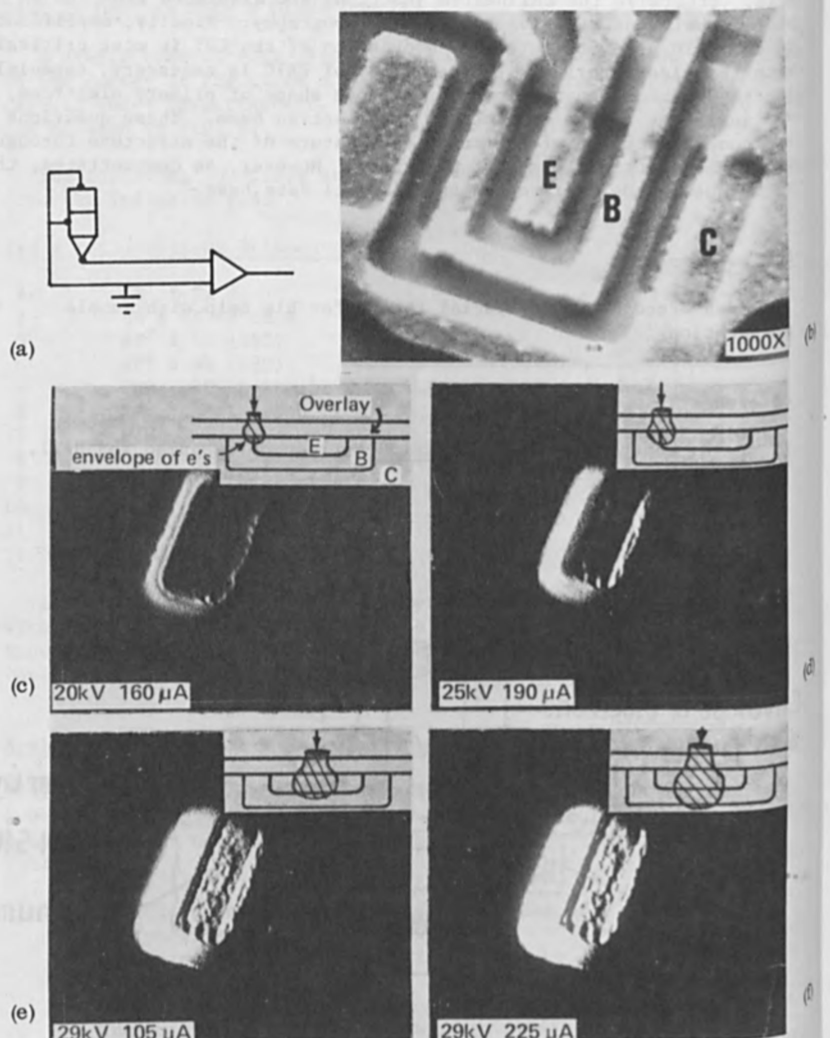


Fig. 2. EBIC detected (a) in a transistor (b) as beam potential and current are increased (c-f)

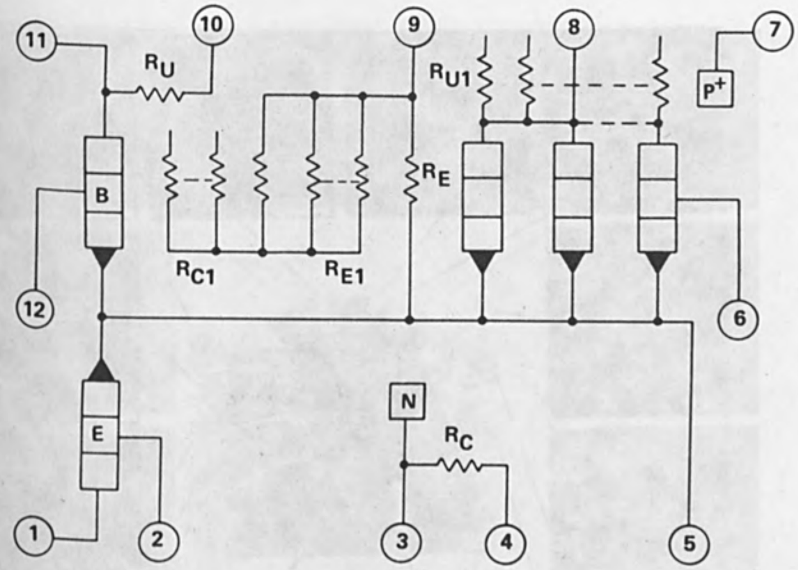


Fig. 3. Test circuit for EBIC study.

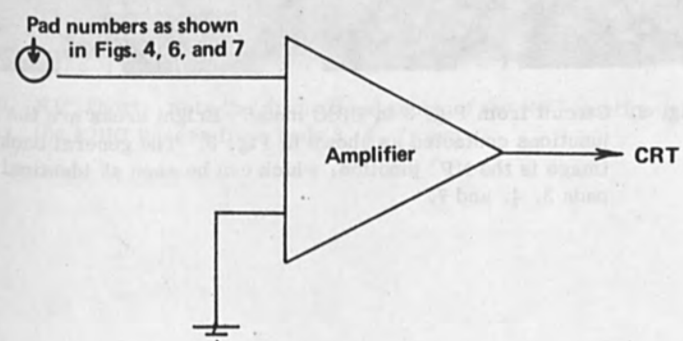


Fig. 5. Detection circuit for EBIC images in Figs. 4, 6, and 7.

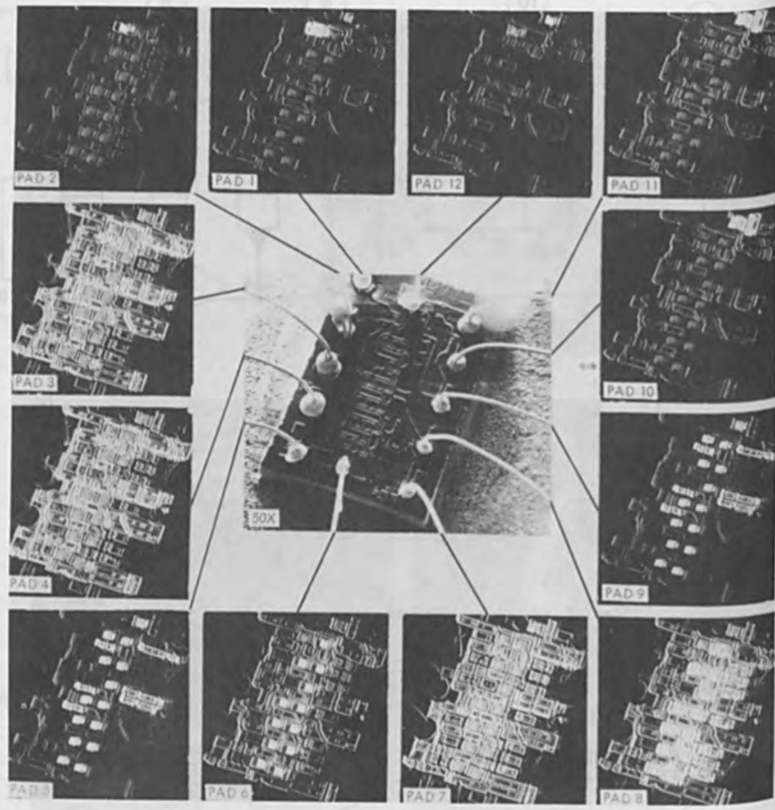


Fig. 4. Circuit from Fig. 3 in EBIC mode. Bright areas are the junctions contacted as shown in Fig. 5. The general background image is the NP^+ junction, which can be seen as identical for pads 3, 4, and 7.

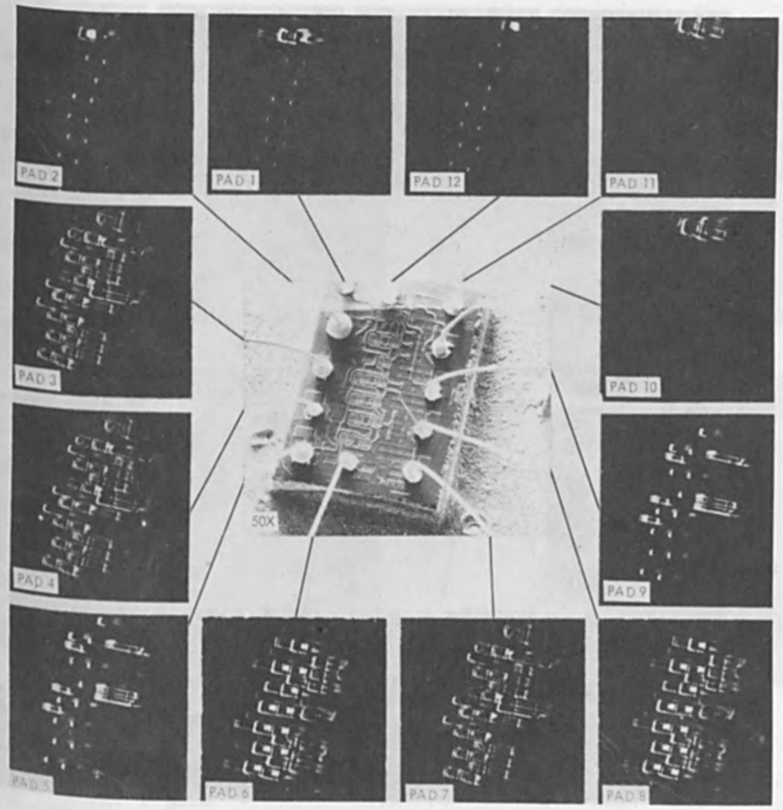


Fig. 6. NP$^+$ short. Note the dramatic absence of the NP$^+$ junction in the EBIC images from pads 3, 4, 7, and 8.

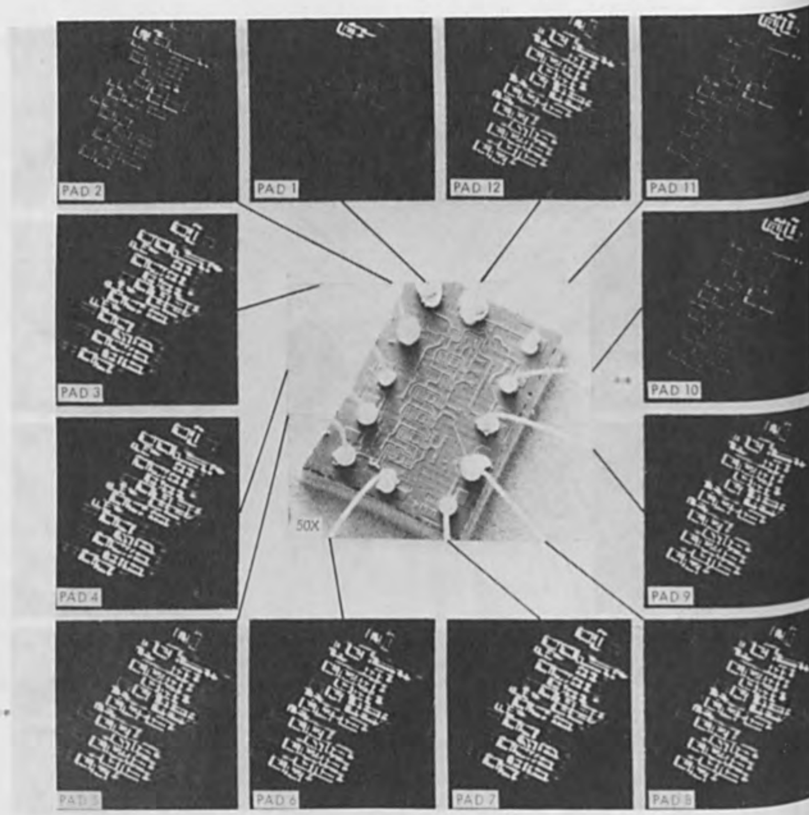


Fig. 7. An EB short on Tx B, and an NP^+ short. Again, note the absence of the NP^+ junction in the EBIC images from pads 3, 4, 7, and 8, as in Fig. 6. Pad 5 is the common emitter of all transistors and the base of Tx B; thus the fact that the images from pads 12 and 5 are identical indicates that there is an EB short on Tx B.

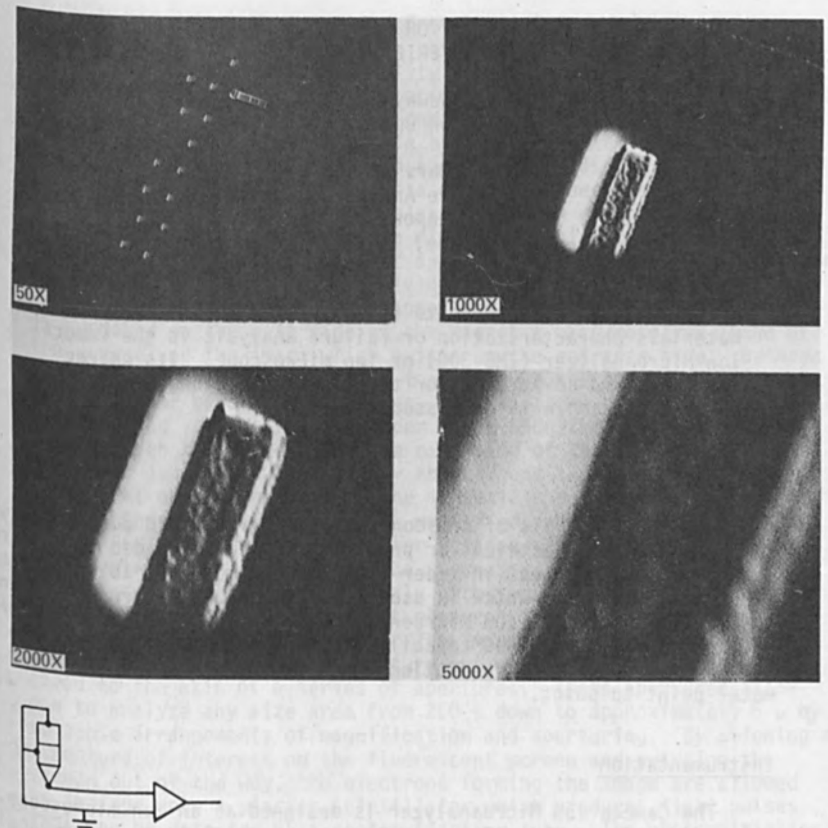


Fig. 8. EBIC at pad 5 isolation-suppressed.

THE APPLICATION OF ION MICROSCOPY TO SEMICONDUCTOR MATERIALS CHARACTERIZATION AND FAILURE ANALYSIS

Bradway F. Phillips
Electron Optics Laboratory
and
Marvin Jackson
Failure Analysis Section
Naval Weapons Support Center
Crane, Indiana 47522

One of the newer instruments to be applied to semiconductor materials characterization or failure analysis is the Cameca Ion Microanalyzer (IMS 300) or ion microscope. Its characteristics and advantages for these analyses along with a pertinent example is discussed.

Introduction

Accurate analysis of semiconductor materials both during fabrication and after electrical or physical failure is needed to help redesign processing steps in order to produce a more reliable product. One type of analysis which is useful at all stages of production and testing is the Cameca Ion Microanalyzer with its rapid picture production capability. This capability can help to sort out contamination problems with a resolution in the imaging mode of 1 micrometer point to point.

Instrumentation

The Cameca Ion Microanalyzer is designed as an ion microscope, not an ion microprobe. This distinction is vital to an understanding of its capabilities. The ion microscope is analogous to the familiar optical or light microscope in that the spatial (x-y) origin of a bit of information from the sample is transmitted unchanged (except for mass analysis) by the instrument and presented in its proper relative orientation at the readout point. Thus, both optical and ion microscopes examine a discrete field, and the details within this field are transmitted with magnification to the viewing or output end of the instrument. The major advantage of this method is that the images formed are composed of separate masses from the periodic chart; that is, 23 Na, 27 Al, 28 Si, etc.

A basic description of the instrument is as follows: It consists of a high intensity duoplasmatron ion gun inclined at an angle of 45^{o} to the specimen surface, with the gasses normally used to produce the ion beam being oxygen, argon, and nitrogen.

The beam diameter can be varied from 10 to 300 μ and primary-ion deflection plates allow it to be rastered to form any rectangular shape up to 2 mm^2. This capability to raster the beam becomes particularly important when a specimen must be analyzed in depth with very high Z resolution. The secondary ions are extracted from the area of the surface of the sample by an electrostatic immersion lens in which the surface of the specimen acts as one element of the lens. This arrangement produces a unresolved secondary-ion image at the level of the deflection plates near the entrance of the electromagnetic mass spectrometer. The electromagnetic spectrometer is a dual pass system with a plane of symmetry vertical through the center of the magnet and the mass-resolving slits. Also, the entrance and exit angles of the electromagnet are calculated such that there is produced on the exit side of the magnet a mass-resolved image of the unresolved ion image which exists on the entrance side. Between the two passes of the secondary ions through the magnet, an electrostatic mirror held at a ΔV (range 0-100 V) above the secondary-ion accelerating voltage sets an upper limit upon the energies that can pass through the magnet. On the exit side of the instrument, a stigmator is used to correct any aberrations in the secondary-ion image that are not removed in the magnetic spectrometer. A projection lens is used to vary the magnified secondary-ion image from 100 to 250X enlargement, and after passing through this lens, a postacceleration lens focuses the secondary-ion image on a specially constructed concave spherical Daly detector. This device changes the ion image into an equivalent electron image which is projected back onto a fluorescent screen. The image-viewing binoculars are used to select a particular feature one wishes to analyze and it is moved to the axis of a series of apertures. These apertures allow one to analyze any size area from 250 μ down to approximately 5 μ by suitable arrangements of magnification and aperturing. By aligning a structure of interest on the fluorescent screen and swinging the screen out of the way, the electrons forming the image are allowed to impinge upon a plastic scintillator which produces light pulses that can be detected by a photomultiplier tube. The photomultiplier tube has both digital and analog amplifiers so that the advantages of both large dynamic range and precise counting at low count rates can be utilized.

By rotating a deflecting magnet, the secondary-electron image can be moved from the fluorescent screen to a position where the image implies upon an extremely thin photographic emulsion within the vacuum chamber. This allows up to 40 photographs to be taken in series without having to open up the vacuum chamber. Also, with the film receiving the electron image directly, better than 1-μ point-to-point resolution can be obtained in the images.

A more complete discussion of the instrument and technique can be found in ref. 1-4.

Applications to Semiconductor Materials

The two basic modes of ion microscopy consist of mass pictorial distributions as a function of depth in a specimen and depth distributions of an element assumed to exist uniformly distributed in planes parallel to the surface of the specimen. Since this uniformly distributed material and the techniques to analyze it are being discussed in papers 90, 100, 131, 133, and 206 of this conference, it will not be mentioned further.

However, the application of ion microscopy to the problems of contaminant distribution in integrated circuits is a newly emerging technique in which the specimen preparation is at least as important as the analytical technique itself. In many integrated circuits the removal of the package to expose the silicon die is extremely difficult and a number of specialized procedures have been devised to accomplish this without destroying the spatial relationship of the impurity elements which must be analyzed. As Figure 1 shows schematically, Ni-Cr resistors are usually deposited upon a thick oxide and connected to other circuit elements by a thick deposited Al lead. Any attempts to sputter away this heavy Al lead in order to examine its interface with the Ni-Cr is doomed to failure because the Al lead is not a flat monocrystalline or microcrystalline surface and therefore selective sputtering occurs and the Al to Ni-Cr interface is approached unevenly. However, if the Si die is turned over and sealed to a mounting block for careful mechanical polishing all of the Si and some of the SiO_2 can be carefully ground away. Then the surface can be lapped smooth so that the ion microscope beam only has to remove less than 1 micrometer of amorphous SiO_2 which it can do easily.

By monitoring the mass peaks at m/e 27, 52, 58 expected from the Ni-Cr resistor and Al leads, the material removal rate can be slowed greatly as the last few hundred angstroms of SiO_2 are removed and the mass spectra shown in Figure 2 can be produced of small areas related to the Al-Ni-Cr interface and the Ni-Cr itself by the use of mass image limiting apertures. These spectra were produced from good integrated circuits and provide a baseline for the comparison of good and bad units. After mass scans are produced the picture mode is used to locate all major and minor constituents. As an indication of this we see the pictures in Figure 3. The elements Ni, Cr, Al, Si all appear to be normally distributed and the Na image demonstrates that the SiO_2 insulator contains appreciable Na. The CrO image shows that Cr has gettered O from the Al during the diffusion anneal used to lower the resistance of the Al-Ni-Cr interface. Another way to visualize the Ni-Cr resistor elements is presented in Figure 4. Here we see the elements and complexes as a function of depth and can show that the diffusion anneal causes the Cr to migrate deeply into the Al. Overannealing can lead to problems at the end of the Ni-Cr resistor where it goes under the Al lead if too much Cr migrates into the Al and leaves a thin area next to the Al lead. This leads to a resistor burnout and is one of the identified failure modes of the Ni-Cr resistors.

Many other examples could be presented but the above does illustrate the technique and a much more detailed account of the instrumental and sample factors affecting ion microscopy can be found in Ref. 5.

Conclusion

Ion microscopy has been shown to be a powerful technique for the analysis of semiconductor materials and has capabilities which are only being developed at this time.

References

1. Bradway F. Phillips, J. Var. Sci. Technol, 11, 1093-99 (1974)
2. H. W. Werner, Vacuum, 24, 10, 493-504
3. H. W. Werner, Appl. Phys., 7, 65-69 (1975)
4. Charles A. Evans, Jr., Analyt. Chem. 47, 9, 818A-826A, 856A-866A
5. George H. Morrison and Georges Slodzian, Analyt. Chem., 47, 11, 932A-943A (1975)

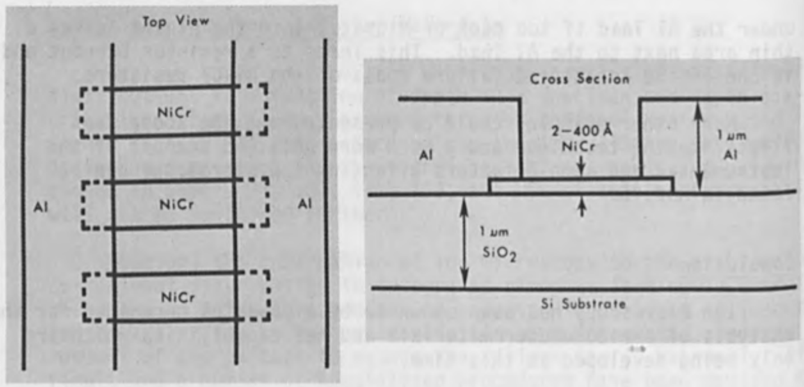


Figure 1. Diagram of nichrome resistors on integrated circuit

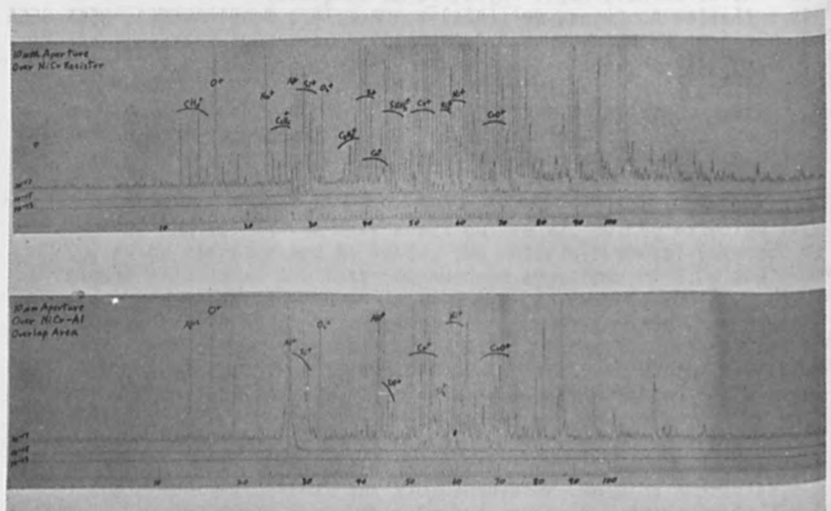


Figure 2. Mass scans of 10 micrometer diameter area on nichrome resistors

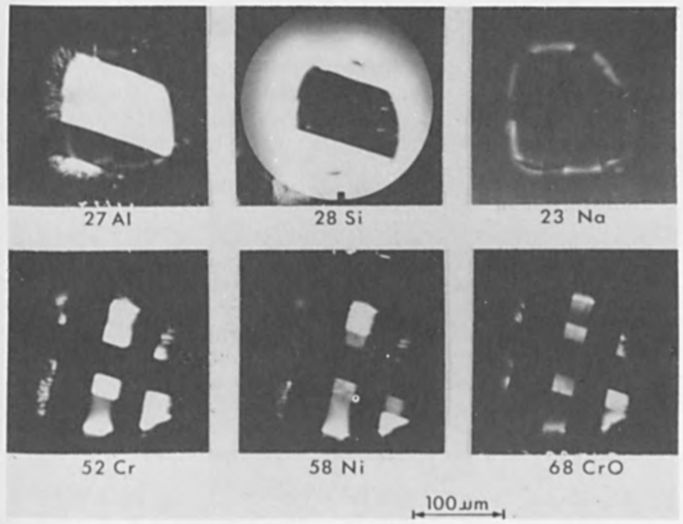


Figure 3. Major element photographs of nichrome resistor structure

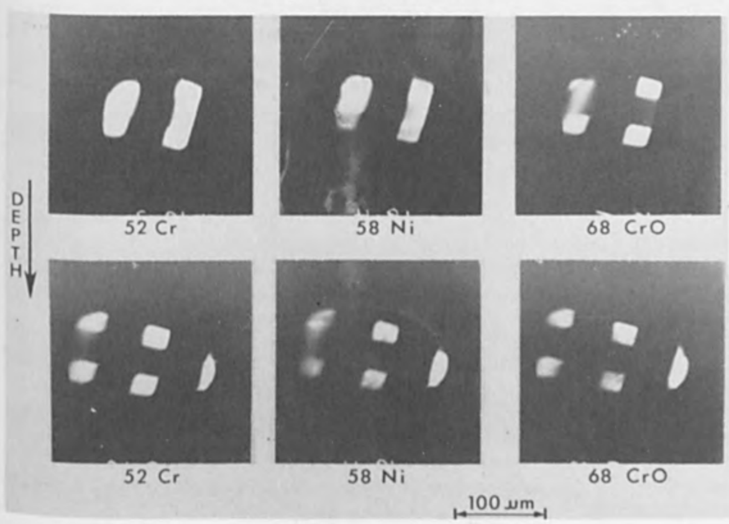


Figure 4. Major element photographs as a function of depth through nichrome resistors

Section II B

X-Ray Lithography

Electron Lithography

Micro Fabrication System

MODELING VALIDATION, TECHNIQUES AND APPLICATIONS FOR X-RAY LITHOGRAPHY

P. I. Hagouel and A. R. Neureuther
Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

ABSTRACT

The experimental development of soft X-ray lithography in conjunction with fabrication of blazed diffraction gratings is discussed. Interferometrically produced gold masks on thin silicon windows and a technique for protecting the mask during etching of the window are described. Images resulting from a multiple source angle exposure technique are shown in support of this technique as a workable concept. Prints of Ni grids and submicron period Au gratings on thin silicon windows are also shown. The constructive use of thermal effects during both exposure and development are also considered.

I. INTRODUCTION

In an earlier publication [1], the feasibility of producing blazed diffraction gratings with X-ray lithography was explored. This work as motivated by the desire for better grating groove shapes and the possibility of using a single mesh to replicate gratings with a variety of blaze angles. It was shown [1] that soft X-ray contact printing with a columnated beam appears quite promising for producing deep blazed gratings with up to 10,000 lines/mm. Initial experimental technique were based on the X-ray lithography techniques described by D. L. Spears et al. [2]. The process was modified and new techniques have been developed in order to accomplish our goals within the framework of our laboratory research facilities.

II. MASK FABRICATION

X-ray masks have been fabricated by placing gold patterns on thin silicon windows as described by Mccoy and Sullivan [3], and by Cohen et al. [4]. Here we describe the process used in our laboratory and in particular those steps that differ from those described elsewhere [3] [4].

The first step in mask fabrication is the diffusion of Boron into the front surface of <100> n-type 3-cm diameter silicon wafer. This is done at a temperature of 1100°C in a flowing gas atmosphere of B_2H_6, O_2 and N_2 for 6 hours. The membrane thickness obtained after preferential etching is around 3 μm. Drive-in at a temperature of 950°C in a wet atmosphere for 3 hours with subsequent etching in concentrated HF will remove the boron glass. A 12000 Å film of SiO_2 is sputter deposited on back side of the wafer and a 1500 Å film, again of SiO_2, is deposited on the front side. A thin layer of Nichrome (300 Å) is then evaporated on the front surface followed by the evaporation of a 2000 Å gold film.

The wafer is then coated with 3000 Å of Shipley AZ1350J positive photoresist and exposed interferometrically using a Helium-Cadium laser (3250 Å wavelength). After exposure, the pattern unit is developed in a 1:1 solution of AZ1350 developer and water. For a 14 minute exposure, at a total laser power output of .4 mW, the development time is 30 seconds. The subsequent postbaking of the developed resist profile is critical for two reasons: The resist has to hold and not flow during r.f. sputter etching. This requires a rather elevated postbaking temperature. Also the preservation of the developed resist profile (a submicron period grating structure) during postbaking imposes the opposite requirement, i.e. a low postbaking temperature. We chose a two step process where we postbake for a while at a low temperature (90°C) and immediately afterwards at a higher temperature (110°C).

Both r.f. sputter etching and wet (chemical) etching have been used to produce the gold pattern on the front side of the wafer; however, due to edge resolution and undercutting of chemical etching r.f. sputter etching is preferred over chemical etching. Figures 1 and 2 show a typical mask after r.f. sputter etching. The roughness of the lines is apparently due to roughness in the developed photoresist surface.

A major fabricational problem was encountered due to the fact that the preferential etchant of silicon (ethylenediamine pyrochatechol-water) would attack the gold surface. The gold would acquire a frosty appearance after several hours in the etching solution and the grating lines would be literally ripped off. Two alternative procedures were examined.

The first one, proposed by Bernacki et al. [4] was to use a teflon fixture so that only one side of the wafer is in contact with the etchant. However our fixture could only accomodate a certain size circular wafer, and it was difficult to handle a variety of wafer sizes and pieces of various shapes.

Another procedure was to open the window first and then deposit the gold mask. The difficulty of this approach was that the handling of the wafer through the remaining process steps had to be extremely careful.

A variety of different resist and wax films were tried for protection during etching, but they were almost immediately lifted off upon immersion in the etching solution. One known material that etches very slowly is SiO_2; therefore, it was decided to use SiO_2 as a protective film. To this end, the remaining mask fabrication steps are as follows:

1) Evaporate 300 Å of Nichrome on the gold relief structure. This thin film is used as a convenient adhesive interface between gold and SiO_2.

2) Sputter deposit a 1.2 μm film of SiO_2 on top of the Nichrome. This film will protect the gold pattern during the preferential etching of silicon.

3) Coat both sides of the wafer with AZ1350J and then open a SiO_2 window on the back side of the wafer, directly beneath the gold pattern. The SiO_2 window is 7.5 mm square.

4) The wafer is then immersed in the preferential etchant and upon etching of the silicon up to the boron doped layer, the wafer is removed from the solution, placed in warm methanol, then rinsed in DI H_2O and blown dry.

5) The SiO_2 films on both sides of the wafer are then removed by immersion of the wafer in buffered HF.

6) The thin Nichrome film may be left on top of the gold pattern or etched with dilute HCl.

A sketch of the topology for this process is shown in Figure 3.

III. EXPOSURE DEPENDENT RESIST PROPERTIES

The X-ray exposure system consists of a horizontal evaporation gun and a water cooled target at angle of 45° with interchangeable target material. The electron beam diameter size is 4 mm and the target is capable of dissipating the heat produced by a 200 mA current beam. Typical exposure parameters are a beam current of 100 mA, a source to substrate distance of 8 cm and one hour exposure time. Due to equipment availability the above exposure system together with the mask substrate holder had to be incorporated as a unit in one vacuum system. This imposes a severe limitation on exposure time because of the amount of heat generated by the gun cathode filament and the unavailability of suitable conduction or convection paths for efficient cooling of the mask substrate system. Heat is generally detrimental but in some particular instance it can be beneficial. In particular we noted that the pattern was visible in some cases immediately after exposure. The explanation is related to the type of scission events taking place during exposure.

It is well known that PMMA exhibits predominantly random chain scission upon X-ray exposure with a subsequent decrease in the initial average molecular weight. The final average molecular weight after exposure is inversely proportional to the radiochemical yield G for main chain fracture. It has been shown that G increases with temperature. This can be utilized so that exposure time may be decreased in typical X-ray lithographic applications by increasing the ambient temperature of the X-ray exposure system or by incorporating small amounts of acrylonitrile in the PMMA resist [5].

Polymers used for X-ray resists are also degrated by other types of processes such as thermal degradation (Pyrolysis) [6]. Thermal degradation can be divided roughly into two general categories: random chain scission and depolymerization. The latter is essentially a depolymerization process in which monomer units release from the chain ends. In this type of degradation the molecular weight of the residual polymer remains almost equal to the initial average molecular weight.

For low average molecular weight PMMA thermal degradation consists almost exclusively of depolymerization. Volatile monomer (methyl-methacrylate) is almost the exclusive product [7]. The simultaneous radiation degradation and thermal depolymerization can be useful in X-ray lithography since X-ray-thermal exposure is effectively a simultaneous exposure and development process.

IV. EXPOSURE RESULTS

Commercially available nickel grids with 10 μm and 6 μm linewidth were contact printed in PMMA on flexible glass. As it was mentioned earlier, we noted that profiles were visible in some cases immediately after exposure; however the contrast was low. Subsequent development in a mixture of 60% isopropyl alcohol and 40% isobutyle keytone for 50 minutes increased the contrast substantially by deepening the pattern in the exposed regions. This validated the simultaneous radiation degradation and thermal depolymerization effect.

The 10 μm grid was also contact printed in PMMA using a two source exposure as follows: first the mask substrate system was positioned at an angle of 35° with respect to the horizontal and exposed for 20 minutes at a source to substrate distance of 8 cm and a beam current of 100 mA. Then the mask substrate system was rotated around its axis of rotation so that the angle for the second case was -55°. It was again exposed for the same time.

This exposure scheme produced three distinct regions of exposure.

A. A region where the exposure was maximum (effectively a 40 minute exposure).

B. A region where one exposure was partially attenuated to the mask.

C. A region where both exposures were partially attenuated.

The results after development in Figure 4 show a distinct 3 step structure due to the different etch rates in the different regions of exposure. The visible edge ripple of the replicated grid bars is due to the exact reproduction of the origin of Nickle mask which had rippled edges.

Next a submicron gold grating on top of a thin silicon window was replicated at an angle of 10°. The diameter of the grating area was

6 mm. Due to the penumbra the grooves of the replicated grating in PMMA shown in Figure 5 are wider than the originals of the gold grating. Also, limitation in exposure time (One hour total exposure time at 100 mA beam current at source to substrate distance of 6 cm) due to excessive heating of the mask substrate system, resulted in poor contrast after development. Total development time was 50 min in a 60:40 mixture of IPA:MIBK. The eventual goal is the control of replicated grating groove shape through the use of a multiple angle exposure technique as depicted in Figure 6.

V. CONCLUSIONS

The feasibility of X-ray lithographic fabrication of blazed diffraction gratings has been explored experimentally through development of interferometrically produced masks and multiple source angle exposures. Gold masks with a .6 μ period on a 7.5 mm square 3 μ thick silicon window were produced using the 3250 Å line of a He-Cd laser, positive photoresist (AZ1350J), and r.f. sputter etching. A nichrome and SiO_2 coating technique was developed for protecting the gold mask during the etching of the thin silicon window.

Replication was accomplished with a $Al_{K\alpha}$ line (8.3 Å) exposure system. During exposure additional thermal effects were noted and explored theoretically. These effects were generally detrimental, but sometimes useful as in the case of replicating nickel grids. A two angle exposure produced images which established the multiple source exposure technique as a workable concept. The .6 μ period gratings were successfully replicated; however, due to our exposure system limitations the extent to which the replicated grating groove shape can be controlled is a bit uncertain.

ACKNOWLEDGEMENT

Research sponsored by the Joint Services Electronics Program Contract F44620-76-C-0100.

REFERENCES

[1] A. R. Neureuther, P. I. Hagouel, "X-Ray Lithographic Fabrication of Blazed Diffraction Gratings," Proceedings of the Sixth International Conference on Electron and Ion Beam Science and Technology, edited by R. Bakish (The Electrochemical Society, Princeton, 1974) pp. 23-33.

[2] D. L. Spears, Henry I. Smith, and E. Stern, "X-Ray Replication of Scanning Electron Microscope Generated Patterns, Proceedings of the Fifth International Conference on Electron and Ion Beam Science Science and Technology, edited by R. Bakish (The Electrochemical Society, Princeton, 1972) pp. 80-91.

[3] J. H. McCoy and P. A. Sullivan, "Progress in X-Ray Lithography," Op. cit. pp. 3-18.

[4] R. A. Cohen, R. W. Mountain, D. L. Spears, H. I. Smith, M. A. Lemma, and S. E. Bernacki, "Fabrication Procedure for Silicon Membrane X-Ray Lithography masks," Technical Note 1973-38, 20 September 1973, Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, Massachusetts 02173.

[5] Y. Hatano, H. Morishita, S. Nonogaki, "A Modified Methyl Methocrylate Positive Resist," ACS Organic Coating and Plastics Preprints, vol. 35, no. 2, pp. 258-265, August 1975.

[6] R. T. Conley, THERMAL STABILITY OF POLYMERS Marcel Dekker, New York, 1970.

[7] H. H. G. Jellinek, "Some Aspects of Polymer Degradation Process," The Effects of Polymer Degradation of Flow Properties of Fluids and Lubricants Containing Polymers, American Society for Testing and Materials, 1967.

[8] P. I. Hagouel and A. R. Neureuther, "Modeling of X-Ray Resists for High Resolution Lithography," Op. cit. vol. 35, no. 2, pp. 298-305, August 1975.

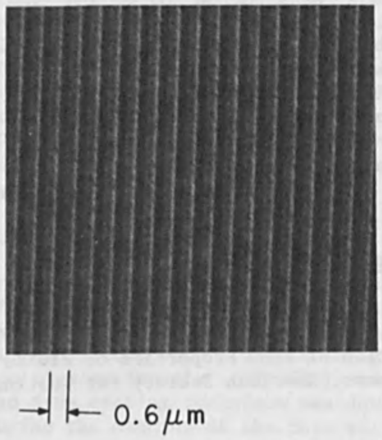


Fig. 1 Gold mask on thin silicon window produced by r.f. sputter etching interferometrically exposed AZ1350J at 3250Å.

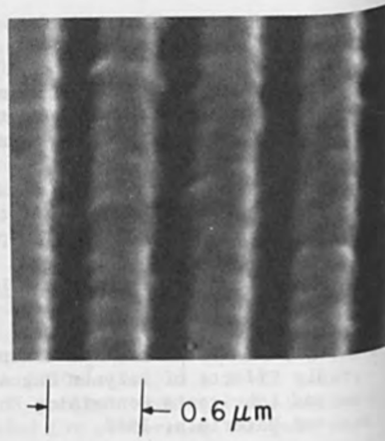


Fig. 2 Close up of gold mask showing roughness apparently related to developed photoresist roughness.

Fig. 4 Replication of 10 μ nickel grid in PMMA showing superpositions of two images made through the use of exposures at two source angles.

Fig. 5 Replication-.6 μ period grating in PMMA.

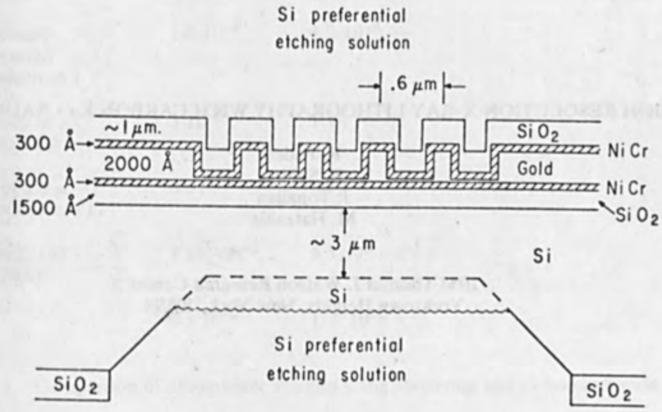


Fig. 3 Wafer topology for nichrome-SiO_2 protection of Au mask during window etching.

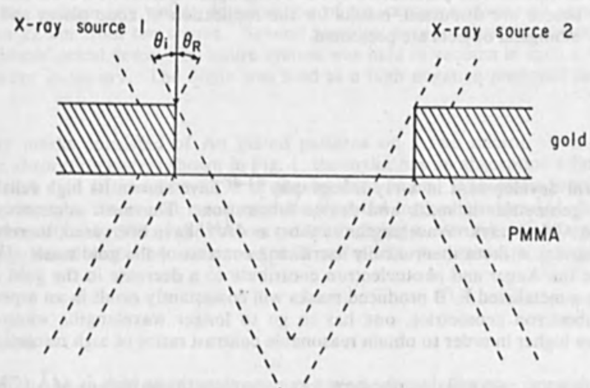


Fig. 6 Multiple angle exposure technique for groove shape control.

HIGH RESOLUTION X-RAY LITHOGRAPHY WITH CARBON Kα - RADIATION

R. Feder
E. Spiller
J. Topalian
M. Hatzakis

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

ABSTRACT: Electron beam generated high resolution patterns (linewidth less than 0.1μ) are replicated by x-ray lithography using Carbon Kα radiation with $\lambda = 44.8$ Å. Carbon radiation offers several advantages over the shorter wavelengths used previously; these are high contrast of mask and higher resolution of the resist with an acceptable transmission of mask substrates made of organic materials. The parameters of the lithographic system with a carbon source are discussed, results for the replication of zone plates and biological objects are presented.

Recent development in x-ray lithography [1-4] have shown its high reliability in copying sub-micron geometries in mask and device fabrication. The most commonly used radiation source is the AlKα x-rays. Wavelengths as short as 4Å[4] have been used, thereby allowing much thicker substrates without theoretically sacrificing contrast of the gold mask. However spurious effects from the Auger and photoelectron contribute to a decrease in the gold contrast.[5] As a general rule a metallized E/B produced masks will consistantly result in an aspect ratio less than 1:l. For submicron geometries, one has to go to longer wavelengths where the absorption constants are higher in order to obtain reasonable contrast ratios of high resolution patterns.

In this paper, we will describe how x-ray wavelengths as high as 44Å (CKα) were used to copy very thin Au masks. The copies will be seen to be of high resolution with high aspect ratios. The quantum efficiency of carbon x-ray generation is reported[6] to be more than twice that of Al. Table I compares the two targets showing that the power absorbed by PMMA per unit power of the electron gun is approximately the same. The fact that the resist is about four times more absorbing for the CKα than for AlKα also means the need to prevent dust, especially non-organic dust, from contaminating the mask or substrate surfaces.

	AlKα(λ=8.3 Å	CKα (λ=44.8 Å)
Operating voltage	8KV	10KV
quantum efficiency of x-ray generation (photons/sr electron)	$1.6\ 10^{-4}$	$4.2\ 10^{-4}$
power efficiency of x-ray generation (η)	$3\ 10^{-5}$	$1.16\ 10^{-5}$
transmission of 2.5 μ thick Mylar substrate (T)	0.67	0.32
absorption coeff. (α) of resist (PMMA)	$1\ 10^{3}\ cm^{-1}$	$4.5\ 10^{3}\ cm^{-1}$
$\eta\, T\alpha$	2×10^{-2}	$1.7\ 10^{-2}$

Table 1. Comparison of lithographic systems using aluminum and carbon radiation

Experimental System

The C-x-rays were obtained using the vacuum generator electron gun evaporation system. The hearth used had a 3mm undercut blind hole drilled in the copper cap. A rod of high prity carbon was hydraulically pressed into the hole to increase the carbon. density and be sure of a tight fit in the copper. With a higher melting point, we were able to use l0 kV (power supply limited) and 50 ma without any visible deterioration of the carbon surface. All exposures were made on samples 22 cm from the source. Several samples were exposed in parallel and were positioned on a hemisherical dome. The entire system was held in vacuum in such a way that no x-ray windows were necessary. The dome was held at a high negative potential to repel stray electrons.

The x-ray masks consisted of Au plated patterns on a two micron thick mylar film. According to the absorption curves shown in Fig. 1, the mylar has an absorption edge just below the carbon line such that it transmits 67% of the incident beam. The curve representing the absorption of Au includes the recent measurements reported in the literature.(7) It shows that 0.1μ thick gold transmits only 5% of the carbon radiation giving a contrast ratio of approximately 20:1. With a gap less than 6μ between mask and wafer, the penumbral blurring is less than 300Å.

Experimental Results

The generally accepted mechanism of resist exposure by x-rays is by the ejection of electrons from various shells which in turn causes the breaking of polymer bonds. The maximum electron energy available is equal to the energy of the incident photon. Although it is possible to estimate the distance traversed by the electron to determine the ultimate resolution, a series of experimental measurements were made to obtain this value. The substrate consisted of a 50 Å thick evaporated layer of Erbium on top of the PMMA resist. This layered structure was then exposed to carbon Kα radiation using a fine copper grid as the mask. After the exposure the Erbium film was etched away and the undeveloped resist coated wafer was broken into many pieces. Each piece was then developed in MIBK for different times ranging from 5 seconds to several minutes. The surfaces of all developed samples were coated with silver and the difference in resist thickness between exposed and unexposed areas was measured with a multi-beam Tolansky interferometer and plotted in Fig. 2. The curve for the Al radiation was reported earlier[8] and is included here for comparison puposes. The intercept on the Y axis represents the amount of resist that was instantaneously developed away. A series of MIBK concentrations were used resulting in different slopes of the development curve. However, all curves gave the same intercept to within ± 20Å. Only one development curve is shown.

An x-ray zone plate was fabricated by an electron beam system and was used as a resolution target. The Au plated lines on a mylar substrate ranged from 1μ to 500Å. The Au was approximately 1000 Å thick. This mask was replicated by the carbon source on a PMMA coated silicon wafer. Figure 3 shows a series of SEM pictures of a) the Au plated E/B zone plate, b) the developed resist pattern after x-ray exposure and c) The plated zone plate. Not all areas of the resist pattern in (b) plated due to a partial residue which was difficult to discern during an optical examination prior to Au plating. Figure 4 shows a similar series of SEM pictures of a bubble pattern. In this case, an optical examination of the resist pattern was sufficient to determine when the residue was removed.

Carbon radiation is strongly attenuated by nearly all elements except carbon. In organic objects therefore all sections containing other elements can be replicated with high contrast. Making use of this effect and the high dissolution rate of PMMA with exposure a series of biological specimens were used as masks to replicate their structure on resist after exposure to C radiation. Since biological specimens vary in thickness and density, we felt that the exposure on the resist should result in a variable dissolution of resist. This concept is shown in the schematic in figure 5. The first biological specimen to be photographed in this manner was a smear of salamander blood cells. These cells are of the primitive type containing a nucleus and to further exaggerate the differences in density a uranium stain was used which is known to migrate to the nucleus. Fig. 6 is a series of optical and SEM photographs where a) is an 800X optical micrograph of the blood cells, b) is an 800X SEM photograph of a similar set of cells and C) is an SEM photograph of an x-ray replica of one of the cells taken at 2000X.

The high resolution capabilities of this technique are shown by x-ray replicas which were made of a number of different diatoms one of which is shown in 7a, b and c. The three SEM photographs represent photographs taken at progressively higher magnifications.

Discussion

With the use of a relatively transparent mask substrate such as a polymer or carbon film, the use of a long wavelengths, e. g., $CK\alpha$ (44Å) can be very useful in copying low contrast E/B masks. Furthermore the high melting point of carbon and the greater efficiency of carbon x-ray generation will allow higher input power thereby decreasing the exposure time. However, the high absorption of most materials to C radiation means that greater care be taken in preventing dust particles acting as unwanted absorbers.

The ultimate resolution capability of C radiation can be seen from the intercepts in Fig. 2 to be approximately 50 ± 20Å (average values of a series of curves) whereas Al radiation is found to be in the order of 400 Å. Fig. 3 is an example of $CK\mu$ being used to copy a thin high resolution mask. The wall height was increased from 1000 Å to approximately 1 μ with a wall thickness approximately 0.1μ.

Although no practical system exists today for focusing x-rays, it is possible to use the x-ray lithographic technique with C radiation as a simple x-ray density microscope. The x-ray micrographs are of unity magnification. However, the high resolution capabilities are useful in studying the replicas by electron microscopy. The three dimensional pictures can be of value in determining the internal structure of biological specimens.

References

1. D. L. Spears and H. I. Smith, Electron Letters, *8*, pp 102 (1972).
2. D. L. Spears and H. I. Smith, Solid State Technol. *15*, pp. 21, 1972
3. S. E. Bernacki and H. I. Smith, IEEE Trans. on Electron Devices *22*, pp. 421, 1975.
4. D. Maydan, G. A. Coquin, J. R. Maldonado, S. Samekh, D. Y. Lou and G. N. Taylor. IEEE Trans. Electron Devices, *22*, pp. 429, (1975).
5. J. R. Maldonado, G. A. Coguin, D. Maydan and S. Somekh, J. Vac. Sci., *12*, pp. 1329, Nov/Dec. 1975
6. R. M. Dolby, Br. J. Appl. Phys., *11*, pp. 64, 1960
7. H. J. Hagemann, W. Gudat and C. Kunz, J. Opt. Soc. *65*, p. 742, (1975)
8. R. Feder, E. Spiller and J. Topalian, J. Vac. Sci. Tech. *12*, pp. 1332 (1975).

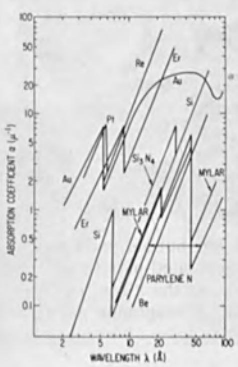


Fig. 1.

Absorption coefficient of some of the most absorbing and most transparent materials for soft x-rays. Data are obtained from standard references except for the recently reported measurements on Au.[7] An important feature is the relative high transmission and absorption of selected material for $\lambda = 44.8$ Å.

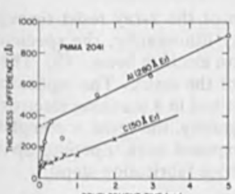


Fig. 2.

Thickness of a film of PMMA coated with erbium (280 Å for AlKα curve and 50 Å for CKα curve) and exposed to x-rays as a function of development time in MIBK. The erbium film is removed oafter the exposures in diluted HCl. The exposed wafer is cut into several pieces which are developed for different times. The thickness of the silver coated resist is measured with mutiple-beam interferometry. the variations in exposed and unexposed resist is defined by a 200 mesh grid in front of the water during exposure.

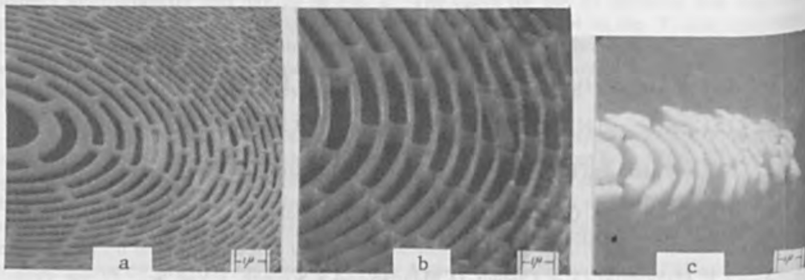


Fig. 3. Scanning electron micrographs of a zone plate. (a) is a gold plated E/B generated mask 1000 Å thick with line widths varying from 1μ to 500 Å. (b) is an x-ray image of (a) in resist approximately 1μ thick. (c) is the plated resist pattern of (b). Due to residual residue not all areas plated. All micrographs were taken at 60° angle of incidence.

Fig. 4. Representative area of 1μ I-T bar bubble pattern. (A) represents the plated E/B generated pattern approximately .1μ thick. (B) represents the x-ray resist image approximately .1μ thick. Micrographs were taken at a 60° angle of incidence

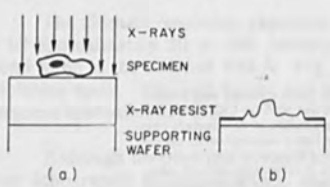


Fig. 5.

(a) Irradiation of the x-ray resist through the specimen. In x-ray lithography, the specimen is a mask generated by an electron beam. (b) The replica after development of the resist. The replica can be metallized and examined in a scanning electron microscope. In x-ray lithography, the resist is completely removed in the fully exposed area, opening up this area for subsequent device fabrication steps.

Fig. 6. Optical and electron micrographs of salamander blood cells stained with uranyl acetate. (a) is an optical photograph taken at normal incidence. (b) is an SEM micrograph of similar cells coated with a thin layer of aluminum and taken at a 60° angle of incidence. (c) is an aluminum coated x-ray resist replica of the blood cells taken at a 60° angle of incidence.

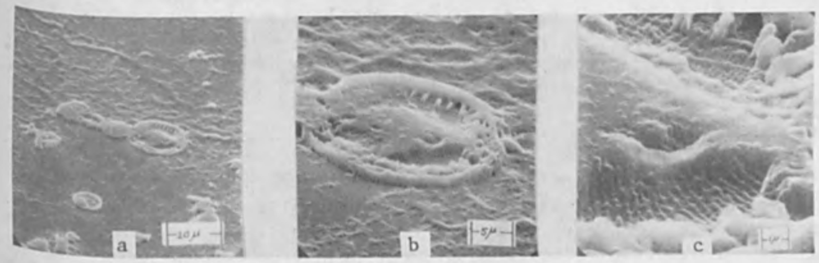


Fig. 7. SEM micrographs of an aluminum coated x-ray resist replica of a diatom. A, B, and C represents progressively higher magnifications of the same area. The smallest discernible structure is less than 0.1μ.

DIRECT WAFER EXPOSURE USING ELECTRON BEAM LITHOGRAPHY VS MASK REPLICATION: A COST COMPARISON MODEL

R. C. HENDERSON
HUGHES RESEARCH LABORATORIES
MALIBU, CALIFORNIA 90265

ABSTRACT

The model developed quantifies the principle that it is less expensive to expose directly the initial wafers rather than to incur the expenses of making masks. There is a crossover point in number of wafers processed where it becomes cheaper to replicate. This crossover point depends strongly upon the area of the chip and the mask defect density. Considering a given electron beam machine, if the chip area times mask defect density is large enough the crossover point is infinite, and it is cheaper to make the circuit with direct writing no matter how many wafers must be processed. For some of the machines in use today this infinite crossover point is reached with chip area > 0.1 cm^2 and defect densities $\gtrsim 2\ cm^{-2}$. The model also analyzes scaling a chip showing that the high resolution capability of direct wafer writing can be used to make circuits with chip area > 0.01 cm^2 more economically as well. The model is general, permitting comparisons with emerging replication techniques. These techniques must use masks with less than 10 defects/cm^2 in order to compete with direct electron beam writing.

INTRODUCTION

One potential use of electron beam lithography is direct exposure on semiconductor substrates to make LSI circuits. This technique is more expensive per wafer than conventional photolithography. Consequently, although a number of electron beam machines have been built,[1-8] they have been used to make the high resolution devices that photolithography cannot easily supply.[9-12]

This situation is changing in two respects. First, the trend to very large circuits creates the problem that typical values for mask defect densities decrease chip yield significantly. If the yields with maskless direct writing were sufficiently better, then the costs incurred processing extra replication wafers would outweigh the higher per wafer processing cost with direct writing. Alternatively, smaller features are obtainable with the electron beam method. This might permit demagnifying large circuits, thereby keeping cm^2/device low enough that typical values of defects/cm^2 do not make manufacture prohibitive.

A second factor is the interest in developing new replication techniques such as x-ray lithography,[13] projection electron lithography[14] or deep UV lithography.[15] These high resolution techniques require a mask in contrast to the (serial) electron beam machines where a computer controls deflection coils, blanking systems and workpiece motion. The new replication methods have shown nearly the same fine feature capability as the electron beam direct writing. However, we need to model the throughput, yield and capital costs of the new methods in order to compare with direct writing.

To analyze these developments a cost comparison model has been derived. This model quantifies the decision to process wafers either by direct writing (DW) or by exposing masks (EM). It is sufficiently general that it is possible to compare with current as well as future replication methods.

The model reflects several qualitative trade-offs between DW and EM. For example, by eliminating a mask step there is one less opportunity for random fluctuations to occur. This factor should mean lower defects/cm^2 as well as superior linewidth control with DW techniques. Also, the pattern written with a serial electron beam machine is adjustable. This should improve alignment by correcting level-to-level runout or compensating for process-induced wafer distortions. However, these potentially economic advantages of DW are weighed against its lower throughput and greater capital costs compared to EM techniques.

In the following section the basic model is derived that compares DW and EM when accounting for mask defects. The results show that chip size is critical to determine when DW is economical. Several illustrations using available performance data are presented. Later, modifications to the model are made that (1) account for demagnifying a chip designed to be made optically, (2) using combinations of DW and EM, or (3) accounting for yield improvement with DW attributable to better linewidth control or alignment. Finally, a number of ramifications are presented reflecting on future development for both the electron beam direct writing as well as the new replication methods.

BASIC MODEL DERIVATION

It simplifies the comparison and does not lose generality to assume that the serial-type electron beam machine generates the masks used by a replication system.* We further assume that the basic wafer fabrication methods are in hand, so that the question addressed is, "If I need a certain number of good chips of a particular circuit, is it cheaper to use DW or EM?" Obviously, for the first wafer or so the answer is to use DW rather than incur the additional expense making the masks. As more wafers are processed, EM eventually becomes more economical. We will refer to this quantity of wafers as the crossover number.

---

*Comparing with electron beam generated masks makes the estimate of DW usage conservative, because such masks should be less expensive than those made by conventional means. See reference 3.

To compute the crossover number, we partition the wafer fabrication cost into three parts: 1) exposure, 2) handling, and 3) materials. For EM, the exposure costs arise from the time spent with the electron beam machine making the masks, as well as from aligning the masks to the wafers and exposing the wafers. For DW, exposure costs arise solely from the electron beam machine time required to pattern each wafer. Handling costs occur from operations common to both types of pattern generation. These operations are wafer diffusions, oxidations, cleans, etc., as well as coating with resists, baking, etching and inspections. It could include chip testing. Finally, the price of materials for wafers and masks must be counted. By partitioning the costs in this way, the handling and wafer materials costs are the same for EM and DW. The wafer exposure costs are different.

Writing down the total as the sum of these three categories gives

$$C = \text{handling} + \text{materials} + \text{exposure}$$

$$C_{EM} = n_{EM}\,(H+W) + n_{\ell}(t_m E + M + n_{EM} R) \tag{1}$$

$$C_{DW} = n_{DW}\,(H+W) + n_{\ell} n_{DW} t_w E \tag{2}$$

where

C_{EM}, C_{DW}	refers to total cost for EM or DW
n_{EM}, n_{DW}	refers to the number of wafers processed by EM or DW
n_{ℓ}	the number of lithographies
E	electron beam machine exposure cost in \$/hr including labor and equipment depreciation
t_m, t_w	time to expose a mask or wafer in hr/exposure. $t_m = t_w$ in general.
M	mask fabrication cost in \$/mask including labor, material and equipment depreciation
W	wafer material cost in \$/wafer
R	replication exposure cost in \$/exposure including labor and equipment depreciation
H	all other processing costs handling wafers in \$/wafer.

If both DW and EM were defect free methods then $n_{DW} = n_{EM}$ and we could plot C_x versus n_x directly as in Fig. 1 (x refers to DW or EM). This shows the costs with replication are higher initially because masks must be made. However, the processing cost per wafer (the slope in Fig. 1) is higher with DW because $t_W E$ is always greater than R. This means there is a crossover number n_c, where C_{DW} becomes greater than C_{EM}. In the defect free case this occurs when n_x is greater than $(t_m E+M)/(t_W E-R)$. Clearly, the greater the mask fabrication cost or the closer $t_W E$ approaches R, the greater the number of devices that must be made to justify replication.

In general, however, n_{DW} will not equal n_{EM} because yields are finite and different for the two processes. A relationship between n_{DW} and n_{EM} can be derived since both must produce the same number of good chips, Q. If there are N chips/wafer then

$$n_x = \frac{Q}{NY_x} \tag{3}$$

where Y_x refers to the yield of either DW or EM. Therefore,

$$n_{EM} = n_{DW} \frac{Y_{DW}}{Y_{EM}} \tag{4}$$

Knowing a relationship between Y_{DW} and Y_{EM} we could replace n_{EM} in Eq. (1) and find cost per wafer on an equivalent basis.

Y_{EM} will differ from Y_{DW} because mask manufacture inevitably introduces defects. Assuming these are random spot defects and that defects on one mask level are unrelated to those on other levels, we can write an approximate expression for the yield Y_M[16]

$$Y_M = \left(\frac{1}{1+AD_1}\right)\left(\frac{1}{1+AD_2}\right)\cdots\left(\frac{1}{1+AD_{n\ell}}\right) \tag{5}$$

where A is the active area of the chip and D_i $(i=1...n_\ell)$ is the number of defects in the i^{th} mask.* In addition, chip yield is reduced due to spot and line defects on the wafers. Denoting yield attributed to these defects by Y_R and since Y_{EM} is the product of Y_M and Y_R we obtain

$$Y_{EM} = \left(\frac{1}{1+AD_m}\right)^{n_\ell} Y_R \tag{6}$$

*The function for yield here, is a form of the more general expression $Y = (1 + A\sigma/\sqrt{D})^{-D/\sigma}$ proposed by Ref. 16 with $\sqrt{D}$ ($\equiv$ mean defect density) = $\sqrt{\sigma}$ ($\equiv$ distribution of defects). Yet $D \approx \sigma$ is often experimentally observed, Ref. 16, 17, and 18. In addition, Eq.(5) predicts yield more optimistically than an exponentially decaying function so that the estimate of the crossover point is conservative.

where D_m is the mean value of mask defects/cm^2. For a first approximation we take $Y_{DW} = Y_R$ because chip loss due to spot defects on wafers should be the same whether they are processed by the resists used with EM or those with DW. Using $Y_{DW} = Y_R$ and substituting Eq.(6) in Eq.(4) we have

$$n_{EM} = (1 + AD_m)^{n_\ell} n_{DW} \tag{7}$$

relating n_{EM} to n_{DW} accounting for mask defects. Yield differences due to pattern misalignment or feature size variation will be treated in a later section.

We can now obtain values for cost per wafer for both EM and DW as a function of the number of equivalent DW wafers. The equations are

$$C_{DW}/n_{DW} = k_1 = n_\ell t_W E + H + W \tag{8}$$

and

$$C_{EM}/n_{DW} = k_2 + k_3/n_{DW} \tag{9}$$

where $$k_2 = (1 + AD_m)^{n_\ell}(n_\ell R + H + W)$$

and $$k_3 = n_\ell(t_m E + M)$$

The general relationships from Eq. (8) and (9) are shown in Fig. 2.

Figures 1 and 2 depict similar relationships except that cost per wafer is used in Fig. 2. This plotting emphasizes the decrease in cost/wafer as more wafers are processed with EM where the mask costs become a smaller fraction of the total cost. With DW, on the other hand, more wafers require more electron beam exposure, so that cost/wafer remains the same. As before, there is a crossover number n_c where if the quantity of wafers that must be processed is less than n_c then it is more economical to use DW. Solving for n_c by setting Eqs. (8) and (9) equal, we have

$$n_c = \frac{k_3}{k_1 - k_2} \tag{10}$$

which reduces to $(t_m E + M)/(t_W E - R)$ in the defect free case.

In contrast to Fig. 1, however, where the pair of curves were determined by wafer and machine parameters, now different chip designs and mask defect densities create a family of curves. As we consider chips with larger A or D_m, k_2 increases, and the crossover number in Fig. 2 moves to the right. Eventually $k_2 = k_1$, and DW is always less expensive than EM no matter how many chips must be made - the crossover number is infinite. This situation arises when, to get a given number of good chips, the extra costs incurred processing more EM wafers outweighs the higher per unit processing costs with DW. This conclusion follows independent of throughput, because machine depreciation and labor costs are included. For $k_2 > k_1$ it is more economical to have additional serial electron beam machines than to process more low-yield wafers with a higher throughput replication system.

Notice if n_c is finite we need to know Q/NY_{DW} in order to determine whether a given project is best processed with DW or EM. This does not apply for infinite n_c which, consequently, is the easier case to use.

To compute whether a given device has an infinite n_c, rearrange Eq. (10) after substituting from Eq. (8) and (9). As $n_\ell R + H + W$ is the total processing cost to make one wafer via the replication technique we will refer to this as C_W. If we expand $(1 + AD_m)^{n_\ell}$ and keep the first three terms we get ($AD_m < 1$)

$$AD_m\left(1 + \left(\frac{n_\ell - 1}{2}\right) AD_m + \ldots\right) = \frac{t_wE-R}{C_W}\left(1 - \frac{t_mE+M}{t_WE-R}\cdot\frac{1}{n_c}\right) \quad (11)$$

Except for n_c the terms on the right-hand side of Eq.(11) are insensitive to larger chips. Thus, passing to the limit $A \to A_{max}$, then $n_c \to \infty$. A_{max} is the largest chip that can be replicated using a given defect density before DW becomes the more economical route regardless of the number of wafers required. Eq. (11) becomes

$$A_{max}D_m\left(1 + \left(\frac{n_\ell - 1}{2}\right) A_{max}D_m + \ldots\right) = \frac{t_wE-R}{C_W} \quad (12)$$

GRAPHS AND EXAMPLES OF THE BASIC MODEL

Plotting Eq. (12) effectively portrays the relationships between DW and EM. This is shown in Fig. 3 with A_{max} plotted versus t_WE-R divided by the total wafer fabrication costs, C_W. The curves are parameterized by D_m and n_ℓ. Figure 3 is a universally applicable curve to evaluate any serial electron beam machine with respect to any replication technique accounting for mask defects; its use is straightforward. First determine values for the abscissa. Next find the intersection of the abscissa with the curve representing the appropriate D_m and read A_{max}. If the chip to be fabricated is larger than A_{max}, DW is the more economical route. On the other hand, if the intersection falls below the corresponding D_m curve, the crossover point is finite. In this case if the demand for chips is large enough, EM will be more economical.

Equation (11) which includes mask costs, provides the ability to calculate a finite n_c once values for E, R, M and t_W are fixed. A useful plot similar to Fig. 3 may be constructed, this time using n_c as a parameter and fixing D_m. To do this requires knowing t_m which should equal t_W. One case where $t_m > t_W$ would be constructing masks for projection printers where the area of the mask may be four times that of the wafer. Conversely, if electron beam registration on the wafers takes excessive time, then $t_W > t_m$. We will consider the case where $t_m = t_W$ and where the costs of processing the masks - labor, materials, plus mask inspection - are approximately five times the exposure cost ($M \approx 5t_W E$). For small R then

$$\frac{t_m E + M}{t_W E - R} \approx 6 \tag{13}$$

Using Eq. (13), Eq. (11) has been plotted in Fig. 4 for $D_m = 2\ cm^{-2}$. The curves rise initially very rapidly as n_{DW} changes. If $n_{DW} = 6$ or less, all chips for the case graphed in Fig. 4 are made more cheaply with DW. The $(n_{DW}) = \infty$ curve is the same as the $D_m = 2\ cm^{-2}$ curve in Fig. 3.

Previously published data show that some of the electron beam machines already developed have $t_W E/C_W$ values as low as 0.1. For example, the cost of the system developed at Bell Laboratories has been placed[19] at 600 K$. With a ten-year amortization, using a two-shift operation, and allowing 10% for the cost of money, produces a machine cost of 17$/hr or E = 42$/hr allowing 25$/hr labor cost. This machine exposes a two-inch wafer, including pattern registration, in $\lesssim$ 1/2 hour typically.[20] Allowing $100-$200 for processing 2" wafers, then $t_W E/C_W$ ranges from 0.1 to 0.2 for this electron beam system. The machine developed at Texas Instruments, although taking twice as long to expose a wafer, costs $\sim$ 350K$.[9] Using the same assumptions for machine depreciation and labor, then $t_W E/C_W$ here ranges 0.17 to 0.34

Most previous work [9-12] used DW to make devices with the very small dimensions that photolithography cannot do easily. Surprisingly, using the machines discussed above, DW compares favorably with replication by contact printing to make LSI chips with conventional dimensions. For example, with 5-10 μm design rules good chromium masters are being made[21] with D_m values $\sim 0.7\ cm^{-2}$. After making copies D_m increases to $\sim 1\ cm^{-2}$ and after printing a few wafers D_m rises to 2 to 3 cm^{-2}.[22] The costs of a contact printer and the working copy masks are low compared to $t_W E$, so we can ignore the R in the abscissa of Fig. 3. Consequently, from that figure we see that direct writing is more economical than contact printing for chips > 0.1 cm^2 in area when using machines that have $t_W E/C_W$ values 0.1-0.2.

MODIFICATIONS OF THE MODEL

The economic model presented above can be maintained, with minor modifications, to analyze additional situations: (1) making the electron beam chip smaller than the optical one, (2) accounting for increased yield with DW attributable to better linewidth control or alignment, and (3) using DW to replace only some of the lithographic steps.

Case (1). Presumably the chip under consideration has an optimum size with respect to yield for a particular replication technique. If this technique is conventional photolithographic processing, however, the chip could be made smaller by utilizing the ability of the electron beam techniques to pattern finer features. (This approach has been taken in the past,[23,24] although a general application requires the vertical structures of the devices to also change. In the case of MOS devices these scaling rules have already been presented.[25]) Different chip areas can be accounted for in the model used here by letting A denote the area of photolithographic chip so that $A_{DW} = A/S^2$. Here S is the amount by which the lateral dimensions are shrunk. With smaller dimensions the designs are more susceptible to defects so that $Y_{DW} = \alpha Y_R$ with, in this case $\alpha < 1$. Equation (7) becomes

$$n_{EM} = (1 + AD_m)^{n_\ell}(\alpha S^2)n_{DW} \qquad (14)$$

which gives a new k_2

$$k_2 = (1 + AD_m)^{n_\ell}(\alpha S^2)\,(n_\ell R + H + W) \qquad (15)$$

Thus we retain Figs. 3 and 4, but with an abscissa that is $(t_w E-R)/\alpha S^2 C_W$. For a given electron beam machine (fixing $t_w E$), for a given device technology (fixing $n_\ell R + H + W$), and for a given replication process (fixing R and D_m), it is straightforward to determine the amount of scaling required to make the serial electron beam exposure of a given chip economical.* For example, with present electron beam machines $t_w E/C_W$ can range 0.1 to 0.3. Estimating Y_{DW} to be 70-90% of Y_R, and with a value of S = 3 this range is 0.013-0.05. Referring to Fig. 3, and assuming again that R is negligible and D_m is 2-3 cm^{-2}, the range of chips that are best made with DW expands from > 0.1 cm^2 with no scaling to > 0.01 cm^2 with scaling.

---

* With the regime being considered here, t_w is independent of S. For very large values of S t_w will increase because the address size must decrease.

Case (2). The new replication techniques[13-15] share with serial electron beam DW methods the ability to make fine geometry devices. As evident from inspecting Fig. 2, the mask-making problem must be solved for these new techniques so that $D_m < 10\ cm^{-2}$. Moreover, compared to DW, variations in linewidth should be greater with any mask replication method. This is due to random processes making the mask which arise in addition to printing the wafers. Also, random variations in chip placement occur due to instabilities making the masks and/or wafer distortions that arise with processing. Since the serial electron beam machines can adjust the pattern position, gain and orientation, alignment yield for DW should be better. We can use the modification developed in Case (1) to predict what the magnitude of improvement must be in order to impact economically on the question DW or EM. Here, $\alpha > 1$ and $S^2 = 1$. As Figs. 3 and 4 apply to any replication technology, for $t_W E/C_W$ values $\gtrsim .2$, α must be $\gtrsim 10$ in order that most chips will be made more economically with DW.

Case (3). To simplify the above presentation we have assumed a uniform value for D_m. A better strategy would be to use DW for some steps, and retain EM for those steps that have low D_m. In principle, this could be accounted for by assigning D_m values to each of the n_ℓ levels and optimizing the \$EM/wafer and \$DW/wafer formulas. In light of the approximate nature of these relations it works as well to compute the cost per wafer as a fraction of the total wafer cost. That is, for processes using eight lithographic steps, if four of them are to be made with DW, then the value of C_W to be used in Figs. 3 and 4 is only 1/2 the total cost. This increases the value of $t_W E/C_W$ but at the same time the average D_m increases. Whether this leads to more or less chips applicable to DW depends very much on the values of the mask defect densities.

DISCUSSION

The worthiness of using DW for design checkout has been realized since the early days of electron beam lithography.[26] With the model developed in this paper it is now possible to quantify this use. It has also been clear that DW should be the process method of choice for fabrication of the highest performance (low volume) specialty devices and circuits.[26] It has not been generally appreciated, however, that when using the high speed electron beam machines currently being developed, DW is more economical for making large area chips when compared to conventional lithography.

Moreover, as Eqs. (11) and (12) are general, we can use them to analyze a wide range of applications. The graph in Fig. 3 makes clear the benefits that arise from improved electron beam machines with lower products of exposure time and operating costs ($t_W E$). Different wafer processing can be analyzed by changing the values of H and n_ℓ. New replication techniques can be evaluated if R and typical values of D_m are determined. GaAs processing can be analyzed by changing t_m and t_W, due to the smaller wafer sizes as well as changing H and W values. One could analyze the worthiness of remaking a mask a number of times (increasing $t_mE + M$) to get one with a low D_m. Clearly, the factors favoring DW are wafer processing that is expensive, difficult chip designs that give higher D_m, larger chips, or increased replication costs.

The implications of the model for serial electron beam machines are: (1) From Fig. 3, t_wE/C_W values should be < 0.2 in order to compete with replication. (2) Large-area devices must be manufacturable. This means either a large size for the electron beam scan field or else field butting. (3) For checking new designs DW is the more economical route in practically all cases of chip area, because the number of desired chips is small. Finally, (4) The difference between $n_C = 30$ and $n_C = \infty$ in Fig. 4 is insignificant. With the exception of design checking there are few applications where thirty wafers provide the needed supply of chips. Consequently, for production runs Eq. (12) or Fig. 3 are the appropriate bases for deciding DW or EM.

The implications of these two curves for the new replication techniques are also clear. (1) The new techniques should be studied with the view of determining typical D_m's. If the masks cannot be made with a $D_m < 10$ cm^{-2}, the process will not have many LSI applications that will be economical compared to DW. This requirement will become increasingly stringent as the dimensions of features within the chip areas get small. (2) For chips with $A \lesssim 0.01$ cm^2 mask defects are not dominant and replication techniques are promising. (3) The improvement in yield with DW attributed to better pattern alignment or linewidth control (an unknown factor) must be large (~ 10) when $t_wE/C_W \sim .2$ (a value achieved with present electron beam machines). Consequently, this factor of itself does not preclude replication. However, (4) if all factors combine so that values of $(t_wE - R)/\alpha C_W$ fall below .02, replication even with very low defect masks will not be competitive with DW for chips with area greater than 0.01 cm^2.

What values for t_w can a serial electron beam machine achieve in the future? The answer is a complicated interplay between gun brightness, resist sensitivity, small signal frequency, address size and exposure strategy. For example, speed/area is increased by exposure schemes that use (1) vector scanning[5] which requires only 20-80% area exposure, (2) combinations of large beam spots, for defining gross features, with small beam spots for fine features,[9] or (3) digital signals combined with analog ramps[27] to obtain the equivalent of a high small frequency. On the other hand, various overheads such as stage settling and registration increase time/cm^2 values. For a given system, these factors may be readily analyzed using a model developed by Ozdemir, et al.[27] Nonetheless, an approximate estimate of time/cm^2 achievable can be obtained from assuming 100% area coverage. Then $t_w = 1/fa^2$ where f is the small signal frequency and a is the address size. Frequency cannot be increased indefinitely however, because of the limited current available or limited resist sensitivities. Consequently, most systems under consideration today envision 1 to 20 MHz operation. Hence writing speeds ~ 1 min/cm^2 will be practical for address sizes 1/2 to 1/4 μm. This is $t_W \sim 1/3$ hr for a two-inch wafer. To further reduce t_W, or to maintain this t_W but use finer address sizes, requires special developments. In this regard, the multibeam approach being investigated by the ETEC Corporation might prove beneficial.[6]

SUMMARY

A cost comparison model has been developed to permit quantifying the decision whether to process wafers by directly writing the patterns using a serial electron beam machine or to expose masks with some type of replication technique. The model shows it is more economical to use DW for the first few wafers up to a value termed the crossover number. A relationship to predict the crossover number was developed accounting for differences in yield between EM and DW due to mask defects. The number predicted is conservative, because the function chosen to describe yield deterioration does not fall off as rapidly (with area) as the exponentially decaying functions expected from a random distribution of defects. If the product of mask defects per cm^2 times chip area is large enough, the crossover is infinite, meaning for these chips it is more economical to use DW no matter how many wafers must be processed.

This latter concept was graphed using mask defects per cm^2 as a parameter and with chip area plotted versus $t_W E/C_W$ - a term that is the ratio of the DW exposure costs to the wafer processing costs. Another graph showed the crossover number as a function of chip area when mask defects are 2 cm^{-2} and mask making costs approximately five times that exposing a wafer. In this case all chips are more economically made with DW if only six wafers are required.

The economic model was also applied to scaling an optically made chip, using combinations of DW and EM techniques and accounting for improved yield with DW due to better alignment or linewidth control.

The graphs clarified the machine cost and throughput that serial electron-beam machines must achieve in order to make DW a production technique. They also showed the electron-beam machines must be capable of making large-area chips. For design checking (making a few chips) DW is economically justified. New replication techniques, even though capable of making fine features, must be able to obtain masks with defect densities less than 10 cm^{-2} in order to compete with DW for making LSI chips.

Previously published data of some advanced electron-beam machines establish that achievable values of $t_W E/C_W$ are already 0.1 to 0.3. With current replication techniques mask defect densities are 1-3 cm^{-2} and DW is the more economical approach for devices with $A > 0.1$ cm^2. By applying device scaling, however, DW could be used to make devices 0.1 $cm^2 < A < 0.1$ cm^2 as well.

In the future, replication techniques such as x-ray lithography, projection photolithography, deep UV lithography, and projection electron lithography will compete with DW for chips with $0.1 < A < 0.1$ cm^2. However the smaller feature sizes within these chips will place additional pressure on the values of mask defect density achievable, particularly for those techniques that use x1 masks. Moreover, should very high frequency or multibeam machines develop, values of t_w will decrease from the best value of 1/2 hr for a 2-inch wafer reported to date. Probably a lesser role will be played by the demands of linewidth control and alignment accuracy, even though these demands increase R as well as the ratio Y_{DW}/Y_R. The combination of all these factors will probably preclude EM with the new replication techniques for $A > 0.01$ cm^2.

ACKNOWLEDGEMENT

I would like to thank Dr. G. R. Brewer and Dr. E. D. Wolf for their suggestions regarding this manuscript.

REFERENCES

1. F.S.Ozdemir, E. D. Wolf, and C. R. Buckey, IEEE Trans. Elect. Dev. ED-19, 624 (1972).
2. G. L. Varnell, D. F. Spicer, A. C. Rodger, and R. D. Holland, "Electron and Ion Beam Technology, 6th Int. Conf." R. Bakish Ed., (The Electrochem. Soc., Princeton, NJ, 1974). p.97.
3. D. R. Herriott, R. J. Collier, D. S. Alles and J. W. Stafford, IEEE Trans. Elec. Dev. ED-22, 385 (1975).
4. T. H. P. Chang, A. D. Wilson, A. J. Speth, and A. Kern, "Electron and Ion Beam Technology, 6th Int. Conf." R. Bakish Ed., (The Electrochem. Soc. Princeton, NJ, 1974), p.580.
5. T. H. P. Chang, IEEE Trans. on Magnetics 10, 883(1974).
6. N. C. Yew, "Electron and Ion Beam Technology, 6th Int. Conf." R. Bakish Ed., (The Electrochem. Soc. Princeton, NJ, 1974), p.111.
7. B. P. Piwczyk and K. G. McQuhae, J. Vac. Sci. Technol. 10, 1008 (1973).
8. O. Cahen, R. Sigelle, and J. Trotel, "Electron and Ion Beam Technology, 5th Int. Conf." R. Bakish Ed (The Electrochem. Soc. Princeton, NJ, 1972), p. 92.
9. Final Report, Contract No. DAAB-05-71-C-3715, September 1973. Texas Instruments, Inc., Dallas, Texas (Prepared for U. S. Army Electronics Command, Industrial Engineering Division, Philadelphia, PA).
10. R. F. W. Pease, J. P. Ballantyne, R. C. Henderson, A. M. Voshchenkov, and L. D. Yau, IEEE Trans. Elect. Dev. ED-22, 393 (1975).
11. F. Fang, M. Hatzakis and C. H. Ting, J. Vac. Sci. Technol. 10, 1082 (1973).
12. H. N. Yu, R. H. Dennard, T. H. P. Chang, C. Osburn, V. DiLonardo, H. E. Luhn, J. Vac. Sci. Technol. 12, 1297 (1975).
13. H. I. Smith, D. L. Spears, and S. E. Bernacki, J. Vac. Sci. Technol. 10, 913 (1973).

14. W. R. Livesay and R. B. Fritz, IEEE Trans. on Elect. Dev. ED-19, 647 (1972).
15. B. J. Lin, J. Vac. Sci. Technol. 12, 1317 (1975).
16. C. H. Stapper, IEEE Trans. Elec. Dev. (Corresp.) ED-20, 655 (1973).
17. A. P. Turley and D. S. Herman, IEEE Trans. Parts Hybrids and Pack. PHP-10, 230 (1974).
18. R. B. Seeds, IEEE Int. Conv. Rec. Part 6, 60 (1967).
19. D. R. Heriott, "13th Symposium on Electron and Photon Beam Technology", Colorado Spring, 1974.
20. R. C. Henderson, A. M. Voshchenkov and G. E. Mahoney, "13th Symposium on Electron, Ion, and Photon Beam Technology," Colorado Springs, 1975.
21. R. D. Wasson, Proc. Kodak Microelectronics Seminar Interface 1973, p. 66.
22. J. G. Skinner, Proc. Kodak Microelectronics Seminar Interface 1973, p. 53.
23. R. F. W. Pease, R. C. Henderson, and J. V. Dalton, J. Vac. Sci. and Technol. 10, 1078 (1973).
24. R. C. Henderson, R. F. W. Pease, A. M. Voshchenkov, R. F. Helm and R. L. Wadsack, IEEE J. Solid State Cir. SC-10 92 (1975).
25. R. H. Dennard, F. H. Gaenssler, L. Kahn, H. N. Yu, IEDM Tech. Digest, Washington, D.C., p. 168 (1972).
26. G. R. Brewer, Solid State Technology (August 1972), p. 43.
27. F. S. Ozdemir, C. R. Buckey, and E. D. Wolf, J. Vac. Sci. and Technol. 12, 1246 (1975).

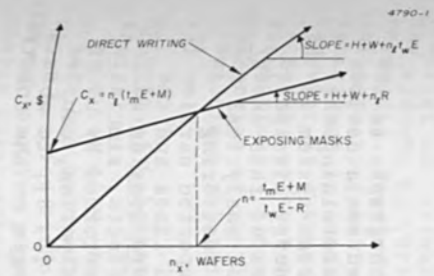


Fig. 1. Schematic Behavior of Total Costs vs Number of Wafers Processed for EM and DW Methods of Wafer Fabrication.

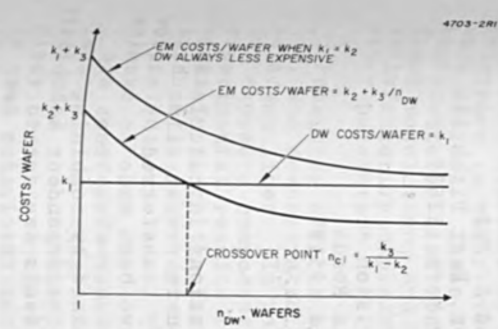


Fig. 2. Schematic Behavior of Costs Per Wafer vs Number of Equivalent DW Wafers for Both EM and DW Methods of Wafer Fabrication.

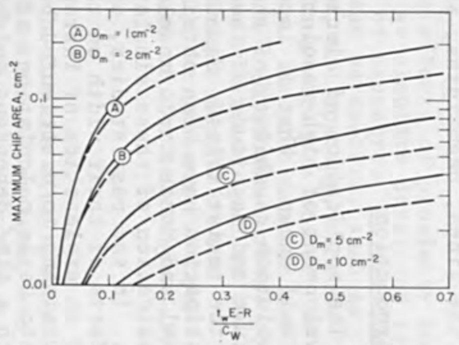


Fig. 3. Chip Area Versus $(t_w E-R)/C_W$ for Infinite Crossover Number, Parameterized by Mask Defect Density and Number of Lithographies. Heavy Line n_ℓ = 4, Dashed Line n_ℓ = 8.

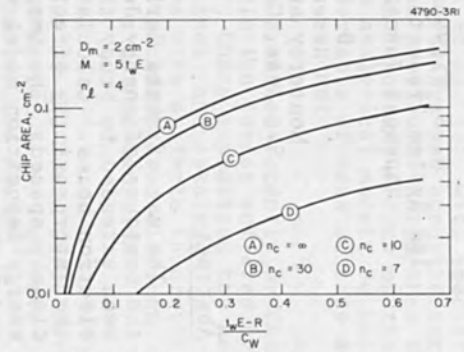


Fig. 4. Chip Area Versus $(t_w E-R)/C_W$ for Various Finite Crossover Numbers, Parameterized by Mask Defect Density, Mask Costs, and Number of Lithographies.

SIMULATION OF LITHOGRAPHIC IMAGES IN ELECTRON-BEAM TECHNOLOGY

N. S. Viswanathan and R. Pyle
IBM System Products Division, East Fishkill
Hopewell Junction, New York 12533

D. F. Kyser
IBM Research Division
Monterey and Cottle Roads
San Jose, California 95193

Abstract

The use of Monte Carlo simulations in calculating the contours of energy deposition in resist films has been extended to study the energy transferred by narrow electron beams. Calculations have been made to examine the interproximity effects that occur when lines are closely spaced. The images obtained by combining the energy-deposition model with a time-dependent solubility model show that the proximity effects are due to the backscattered electrons. Hence the resolution loss reaches its maximum when the acceleration voltage is low and the atomic number of the substrate is high. The effects of other parameters, such as spacing between the lines, beamwidths, and dose, are also examined.

Introduction

The application of electron-beam lithography to the development of high-resolution microelectronic devices is well known. What is important in determining the resolution capability of the electron-beam system is the nature and magnitude of the electron energy dissipated in the resist films. Simulated profiles of this energy dissipation have been obtained by use of analytical models [1-4] and Monte Carlo models [5, 6] of scattering. An application of these simulations was reported earlier [2, 7]. In the past studies, only isolated lines and the effects of their width on the system were considered. For an explanation of intraproximity effects, simulations pertaining to distributed beams were performed by using Monte Carlo calculations of energy dissipation, coupled with a time-dependent solubility model [2, 8], to obtain the contours of the final developed images on the substrate.

Electron-beam lithography, however, has other applications than the isolated lines of identical width that are usually dealt with in simulations of images. The practical effects of using arrays of lines, for example, have been examined by Hawryluk et al. [3], with special reference to their utility in gratings and acoustic surface-wave devices. These authors in particular noted that the backscattered electrons have an important effect on the final resolution, and that the analytical models do not satisfactorily explain the observable effects of size and other parameters.

Proximity effects have been examined recently by Chang [9], who demonstrated the resolution losses suffered by adjacent lines. In particular, Chang emphasized the need for exposure compensation of lines, and obtained experimental results that agree with a simple theoretical model in which the forward and backscattered electron intensities are represented as a set of Gaussian distributions.

In view of the successful use of the Monte Carlo method in obtaining energy profiles of resist films, we have extended the calculations to examine the energy distributions obtained in exposing multiple lines adjacent to each other. Since the final linewidths are of practical importance, we have simulated the images obtainable after the development step, using the solubility model reported earlier for isolated lines.

In this paper we examine the effects of proximity in greater detail. Specifically, we simulate developed resist profiles at various conditions of exposure. We examine the effects of variations in beam parameters, such as beamwidth, acceleration voltage, and spacing, as they influence proximity effects, along with the effects of variations in dose and thickness. We also compare the behavior of a substrate having a high atomic number, gold ($Z = 79$), to that of the silicon substrate used in the simulations. Since the final images depend on the properties of the developer [10], we present further simulations related to the nature of the developer solvents.

Methods of Calculation

The use of Monte Carlo methods for the single scattering models has been well documented in the literature, and their use for lithographic applications has been examined in our earlier papers [5, 7]. In the present work, the energy dissipation for line scan conditions was obtained in the manner we described there. The programs have been modified somewhat to generalize the output to conform to any desired beam shape (an ideal delta function, a Gaussian function, and a square beam of a specified edge slope) with a preselected beamwidth. Also, the energy deposition contours of an array of lines have been computed by superpositioning the energy matrices in two dimensions with prespecified offset distances. All the statistical analyses were based on 100,000 trajectories. Throughout this work, we have used the time-dependent solubility model to simulate the developed profiles in the resist film. Solubility matrices were obtained from matrices of energy deposition by using solubility expressions of the type

$$S_{ij} = B\left(\frac{1}{M_n} + \frac{gE_{ij}}{\rho N}\right)^A$$

where A and B are empirically determined developer constants; A = 1.40, B = 1.25 x 10^3 (Å/min)

g is the radiation sensitivity, g = 0.75 per 100 eV, of a polymer (PMMA) having molecular weight M_n = 120,000. This value of M_n, the value obtained in our radiation chemical studies, is taken as representative. The value of g was also experimentally determined in our laboratory [2].

ρ is the density of the polymer, in g/cm³; ρ = 1.22.

N is Avogadro's number

E_{ij} is the two-dimensional energy matrix obtained by the Monte Carlo simulations, corrected for the beam shape (edge slope = 0.25x bandwidth)

The development program gives a variety of outputs, one of which, as we reported earlier, is a display of the contours. Here we have chosen to display the profiles at a particular time at which an opening of a specified dimension is made at the bottom of the resist film, i.e., the resist-substrate interface. In all the developed resist profiles displayed in this paper, the horizontal axis has been normalized to the beam diameter and the vertical axis represents the depth in the resist film, in micrometres. All the profiles are actual terminal outputs; no routines to fit or smooth the curves were used in processing the data. Other simulated conditions are discussed at appropriate places in the paper. The results shown here represent a few distinct cases of the parameter variations that have been examined, and are meant to be representative of the effects seen in the simulations.

Results and Discussion

1. Beamwidth Effects. In an earlier paper [7] we reported the effect of beamwidth on the final images. Here, we show the effect of beamwidth on a thick film (3-μm PMMA on silicon) at 25 kV, at a dose of 50 $\mu C/cm^2$. As was noted before, the X axis is given in normalized units of beam diameters, and the Y axis represents the depth in the polymer film, in micrometres. Comparison of Figs. 1 and 2 shows that the narrower beam suffers considerable loss of lateral resolution when it is imaging in a 3-μm layer of the resist. For both cases the simulation was stopped when the image at the substrate had reached a width equal to the beamwidth. For the narrow beam, that point is reached after 8.6 min of development, whereas for the broad beam it is reached after only 1.4 min.

Let us look at these results in another fashion. If one wants to generate two isolated beams or lines--one 0.1 μm wide, the other 1.0 μm--on a thick resist material, complete development will take about seven times as long for the narrow lines as for the broad lines. One can visualize a broad line, say 1.0 μm, as being made up of several narrow lines, say 0.1 μm each, which are added together; it is impossible to develop the narrow lines in the same length of time. This is a direct result of the intraproximity effect, discussed in greater detail in our earlier paper [7]. This effect is of great consequence for different linewidths generated on the same wafer--for example, in metallizing devices on which different metal lands are used. We will examine this effect, using another criterion, later in this paper.

2. Proximity Effects. To simulate the effect of adjacent beams on developed images of a particular line, we have taken one line as a reference line and calculated the energy deposition contours adjacent to this line. The displacement of the ancillary beams was simulated by calculating the energy depositions individually, and the effective energy densities were obtained by displacing and superposing the individual energy matrices. For accuracy, the Monte Carlo calculations were extended to about six times the beam diameter on each side. For example, the energy matrix of the primary 1-μm beam is in effect obtained at a distance of 6 μm from each side of the center. To conserve storage, beams other than the central one were usually calculated to be within this distance range. For simplicity, we assume that the central beam has an arbitrary point in the axis (X axis in all the figures) at 0.0, and the relative displacements of

the beams are given in normalized beamwidths with respect to the center. For example, the notation -2.0, 0, 2.0 describes three beams whose centers are four beam diameters apart.

Figures 3-6 show the results obtained in simulations of cases chosen as representative of the adjacency effect. Figure 3 shows how two adjacent beams affect a reference beam between them. The beams are positioned at -2.0, 0, 2.0. In the final image only the center beam is fully developed out; the others are incompletely developed. The important point is that the final image is symmetrical about the center.

In the actual situation when the beam size is 1 μm, and an imaging of three identical beams is attempted, the final images will not be seriously affected when the beams are offset by 2 μm. This would be the extreme proximity condition, in which offset effects are not prevalent. Both of the outside beams require additional development, which will result in some overdevelopment of the principal beam. This effect, which will be examined later, is not dose-dependent, as is evidenced by the data in Fig. 4, for a dose of 100 μC/cm². The most important difference between the images in Figs. 1 and 4 is that the development time is only 0.1 min for the 100-μC/cm² dose, as compared to 0.4 min for the 50-μC/cm² dose.

Figures 5 and 6 show the effect of increased adjacency, for -2.0, 0, 1.0. The differences between these cases and those discussed above are that the beam on the right is offset by 1.0 μm rather than 2.0 μm, and that the same two doses are used. As before, the development was terminated when a 1.0-μm opening occurred at the interface. The final image obtained in Fig. 5 is interesting in several respects. First, the developed images are no longer symmetrical, and the image of the central beam is widened out in the direction of the closer beam. Second, the sizes of the images at the interface are 0.65 and 1.70 μm, respectively, and only two images are obtained. Last, the distance between the two final images is about 1 μm. All these effects are due to pronounced additive energy contributions from the two beams on the right of the isolated beam on the left. In fact, the results imply that the energy deposited at the interface is spread over a 1.70-μm region, and a 1-μm image cannot be replicated at the resist-substrate interface. We believe that this is due to the significant interproximity effects. As with all previous configurations, dosage appears to have minimal effect; cf. Figs. 5 and 6. Doubling the dose to 100 μC seems to have no effect on the final profile (except perhaps in the development times), and the displacement is the same as in Fig. 4. The results of this simulation are given in Fig. 6.

We believe that the adjacency effect is due mainly to pronounced backscattering from the substrate, contributing at distances larger than the beamwidth itself. The resulting increase in exposure causes increased solubility, whose effect is seen in the final images.

This simulation can be extended to examine the effect of completely or partially overlapped exposures. Figure 7 is representative of the results of the -2.0, 0, 0.5 conditions. The development time is considerably shorter (0.19 min) than in the earlier cases (0.36 min for -2.0, 0, 2.0 [Fig. 3] and 0.31 min for -2.0, 0.1, 0 [Fig. 5]). The images also appear, as expected, to be displaced (cf. Fig. 6)in the direction of the closer beam. This condition will represent a 1-μm beam separated by 2 μm from a 1.5-μm beam. Even when the development times are increased by about 150%, the 1-μm image on the left would remain undeveloped. This condition would represent the contribution of both interproximity and intraproximity effects.

Our other simulations show similar trends. We believe that the adjacency effect is due mainly to the electrons backscattered from the substrate. Since the analytical models do not explain the backscattering effects in detail, we believe that the Monte Carlo simulation is the best method for explaining these effects. The results are very similar to those reported by Hawryluk et al. [11], who compared the experimental linewidths with calculations; also, they are substantiated by the experimental findings of Chang [9]. Chang in particular emphasized the need for exposure compensation in patterns requiring proximity, and showed excellent agreement between the experimental values and theoretical expectations. The resolution losses caused by the adjacency effects necessitate additional or compensatory exposures. Also note in Fig. 7 that the spacing between the lines is still preserved. Exposure compensation in this particular case would be about twice that needed for the broad lines (the smaller line requires a higher exposure dose). We will examine the effects of different acceleration voltages and substrates in the following sections.

3. Acceleration Voltage. Variations in the acceleration voltage affect the proximity effects discussed above. Our earlier papers [2, 5] have dealt with isolated lines in greater detail; here we shall discuss only array lines. Figures 8 and 9 show the simulations of the two conditions for a 15-kV beam. The other parameters--dose, beamwidth, etc.--are kept constant. In this section we show simulations for -2.0, 0, 1.0, and -2.0, 0, 0.5. Figure 8, for example, is a direct comparison to the data in Fig. 5. When compared in this fashion, the data show that

i) The development times for the 15-kV case, as expected, are lower. This behavior is similar to that observed for isolated lines.
ii) For 15 kV, the 1-μm image on the left, although undeveloped, suggests an increased state of development when compared to the configuration in Fig. 5. This in turn suggests an increased inter-proximity effect (cf. Figs. 5 and 8).

We believe that this effect is due to the pronounced lateral scattering of the backscattered electrons from the substrate. These electrons have lower energies than those of the 25-kV beams and suffer more lateral resolution loss as they re-enter the resist film. Figure 9 represents the data from the closer spacing and is comparable to the data in Fig. 7. All the arguments advanced for the lower acceleration voltage apply in these cases as well.

4. Substrate Effects. To compare the proximity effects on different substrates, we have chosen gold as a substrate with high atomic number ($Z = 79$) to compare with silicon ($Z = 14$). For brevity we confine our discussion to the specific cases of interest. The symmetrical case (-2.0, 0, 2.0) shows little adjacency effect; see the profiles in Fig. 10. The data should be compared with the 25-kV data for silicon (Fig. 3). At the dose used, however, 50 $\mu C/cm^2$, the overall development times drop to 0.25 min for gold and 0.3 min for silicon. The relative dimensions show that the auxiliary images are considerably smaller, 0.85 μm, than for the silicon substrate, 0.95 μm. Any additional development will obviously affect the dimensions of the principal image, causing a loss in resolution.

Figures 11 and 12 show the effects of proximity on asymmetrically spaced beams. These results should be compared to Fig. 5 (for -2.0, 0, 1.0) and Fig. 7 (for -2.0, 0, 0.5). They are identical with the results for the silicon system. The main difference is perhaps that the residual image of the resist, between the open areas, is considerably less high in Fig. 12 than in Fig. 7, showing that the gold substrate pronouncedly affects the resolution. When the 2.0-μm image is completely developed (Fig. 13), the 1.0-μm image is incompletely formed and the resolution loss is excessive. We again attribute this to backscattering from the substrate, the consequences of which have been discussed by Kyser and Murata [5] in relation to the calculation of absolute energy deposition on various substrates. A practical consequence arises in the fabrication of closely spaced lines on substrates with high or medium atomic number--for example, acoustic surface-wave devices--which require high resolution. Under these conditions, the exposure intensities of patterns have to be modified by suitable calculations of effective charge densities, such as those proposed by Chang [9]. Our conclusion is therefore that the proximity effects are greater for substrates with high atomic number and also for beams with low acceleration voltages.

5. Developer Effects. As was mentioned in the introduction, the properties of the developer are also important. Greenich [12] considered the effect of developer contrast on the time-development contours of the images in PMMA. Hatzakis [10] showed how image contrast is affected by the properties of the developer. In this section, we shall devote our attention to the simulations of proximity effects on the developer constants used in all the simulations. In the equation, for example, two developer-related constants, A and B, are used. These constants can be modified to examine the resulting variations in the final images. This equation, suggested by the work of Ueberreiter *et al.* [13], has been the basis for all the development simulations in this and earlier papers [2, 7].

So far we have dealt with results for A = 1.4. We simulated several profiles with A = 1.4 to 0.6, 1.0, 1.8, etc. The final images for the simple isolated-beam cases are given in Figs. 14-17. The images obtained with high A values show higher contrast and progressively higher development times than those with lower A values. The development times range from less than 0.6 s at A = 0.6 to]6 min at A = 1.8. By the Ueberreiter approach, the lower A values represent the solvents and the higher A values represent the nonsolvents for the polymers. These results imply that the best contrast can be obtained by using solvents having high A values, or nonsolvents as reported by Greenich [12] and experimentally supported by Hatzakis [10].

The effect of variations in B is fairly straightforward. B, given in Å/min, is related to diffusion rates. It linearly affects the development times, as the results in Fig. 18 show. The value of B here is 1.25×10^{8} as against 1.25×10^{9} used in earlier calculations. The value of A is kept at 1.8, and the results are directly comparable to those in Fig. 17.

The interrelations of developer contrast with the proximity effects are compared in Figs. 19 and 20. The results are shown for a narrow beam (0.15 µm) and for A = 0.6 and 1.8, the extremes that were examined earlier. B is kept at 1.25×10^{9}. For higher contrast (Fig. 19), the images still maintain their identity; for lower contrast (solvent developer), considerable resolution is lost, as is shown in Fig. 20. We therefore conclude that the developer contrast too must be considered in examining the effects of adjacency. From the results shown here, it appears that the preferred system for minimizing the resolution losses suffered when lines are closely spaced should be a high-contrast developer--that is, a solvent with low solubility for the polymer, or the mixtures of solvent and nonsolvent used in practice.

Summary and Conclusions

Monte Carlo simulation methods have been coupled with time-dependent solubility models to obtain resist images by simulations at various exposure conditions. The effects of different acceleration voltages, beamwidths, and substrates have been examined to estimate the effects of close spacing between lines in high-resolution electron-beam lithography. Our major conclusion is that proximity effects are dominated by electrons backscattered from the substrate. In general the effect is independent of beamwidth, but is dependent on the line spacings. The effect is aggravated by the use of a low acceleration voltage or a substrate with high atomic number. Moreover, image clarity is degraded when low-contrast solvent developers are used. For lines of dissimilar widths, both interproximity and intraproximity effects influence development times and the resolution of the images. Dose effects seem to be minimal at the dose ranges studied in this investigation.

Acknowledgments

We thank J. A. Todoroff, H. Ezrol, and R. S. Horwath for their support and encouragement in the course of this work.

References

1. R. W. Nosker, J. Appl. Phys., 40, 1872 (1969).
2. M. Hatzakis, C. H. Ting, and N. S. Viswanathan, Proc. 6th Int. Conf. on Electron and Ion Beam Sci. and Technol., Electrochemical Society, Princeton, N. J., 1974, p. 542.
3. R. J. Hawryluk, A. M. Hawryluk, and H. I. Smith, J. Appl. Phys., 45, 2551 (1974).
4. J. S. Greenich and T. van Duzer, IEEE Trans. Elec. Devices, ED-21, 286 (1974).
5. D. F. Kyser and K. Murata, Proc. 6th Int. Conf. on Electron and Ion Beam Sci. and Technol., Electrochemical Society, Princeton, N. J., 1974, p. 205.
6. R. Shimizu, T. Ikuta, T. E. Everhart, and W. J. DeVore, J. Appl. Phys., 46, 1581 (1975).
7. D. F. Kyser and N. S. Viswanathan, J. Vac. Sci. Technol., 12, 1305 (1975).
8. F. H. Dill, A. R. Neureuther, J. A. Tuttle, and E. J. Walker, IEEE Trans. Electron Devices, ED-22, 456 (1975).
9. T. H. P. Chang, J. Vac. Sci. Technol., 12, 1271 (1975).
10. M. Hatzakis, J. Vac. Sci. Technol., 12, 1276 (1975).
11. R. J. Hawryluk, H. I. Smith, A. Soares, and A. M. Hawryluk, J. Appl. Phys., 46, 2528 (1975).
12. J. S. Greenich, J. Appl. Phys., 45, 5264 (1975).
13. K. Ueberreiter, in Diffusion in Polymers, ed. by J. Crank and G. Park (London, Academic Press, 1968), Ch. 7.

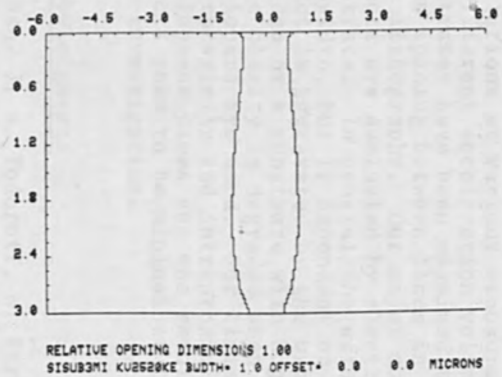


Fig. 1. 25 kV, narrow isolated beam

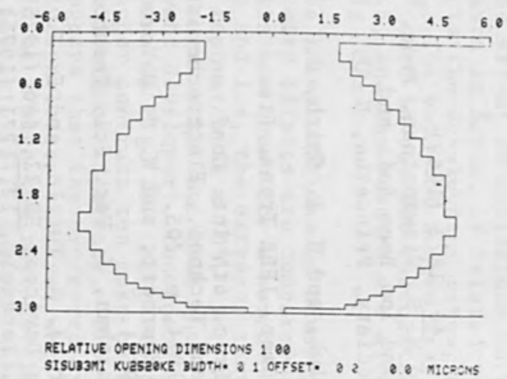


Fig. 2. 25 kV, broad isolated beam

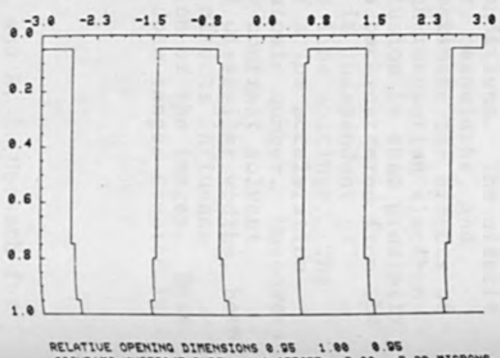


Fig. 3. 25 kV, symmetrical beams, 50 µC

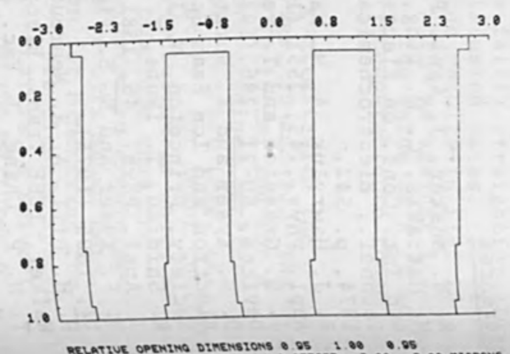


Fig. 4. 25 kV, symmetrical beams,

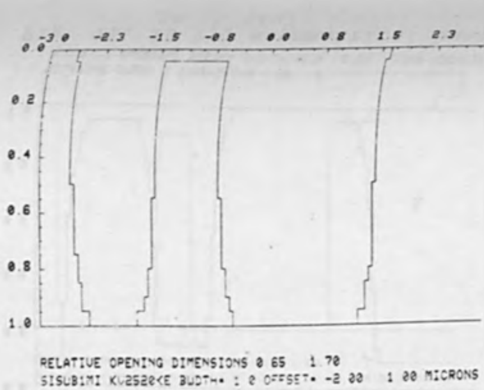


Fig. 5. 25 kV, asymmetrical beams, 50 μC

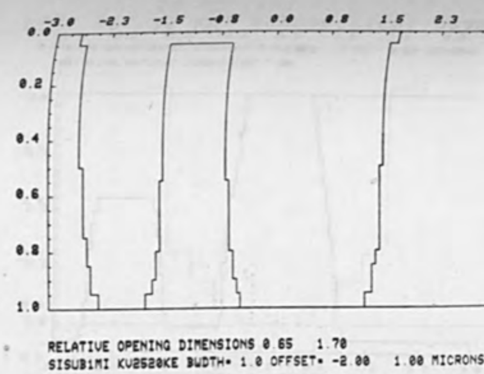


Fig. 6. 25 kV, asymmetrical beams, 100 μC

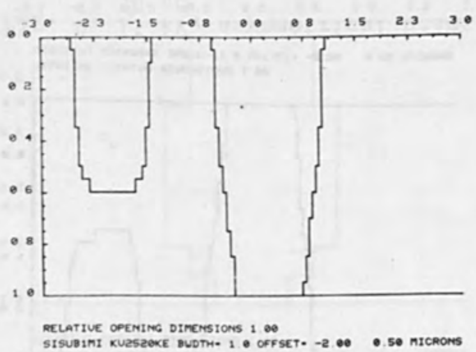


Fig. 7. 25 kV, asymmetrical beams, 50 μC

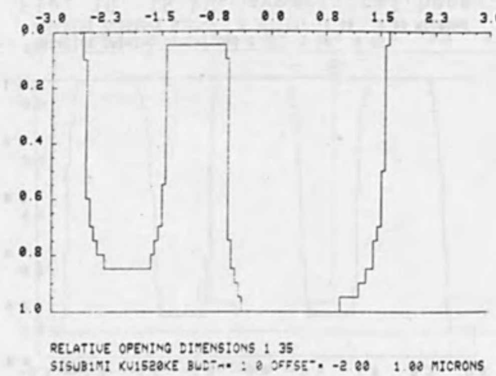


Fig. 8. 15 kV, asymmetrical beams, 50 μC

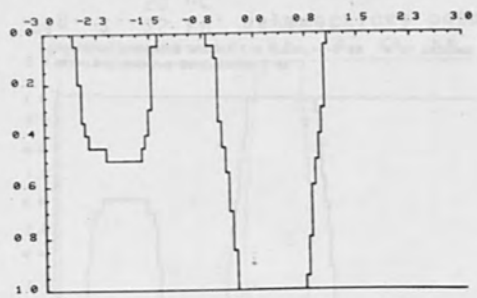


Fig. 9. 15 kV, asymmetrical beams, 50 μC

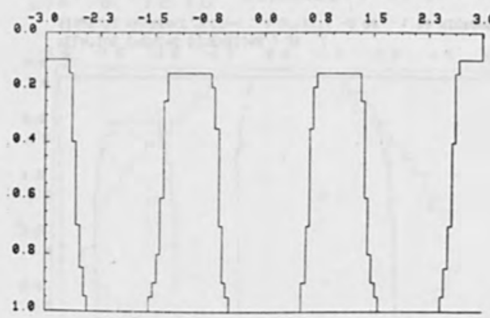


Fig. 10. 15 kV, symmetrical beams, 50 μC (Au)

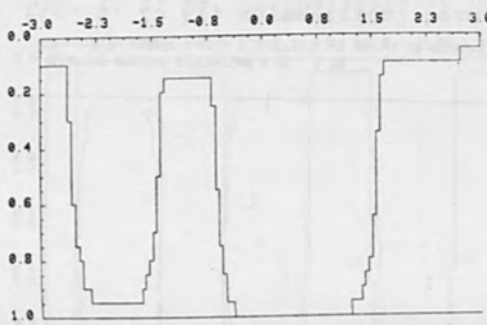


Fig. 11. 15 kV, asymmetrical beams, 50 μC (Au)

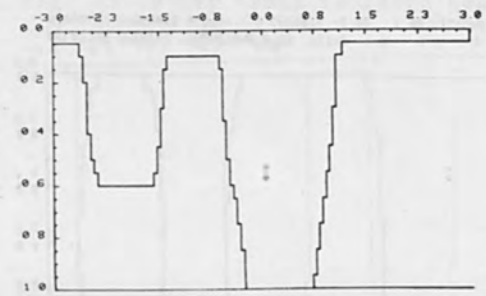


Fig. 12. 25 kV, asymmetrical beams, 50 μC (Au)

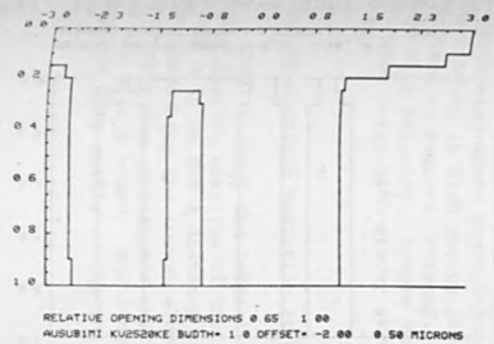


Fig. 13. 25 kV, asymmetrical beams, 50 μC (Au)

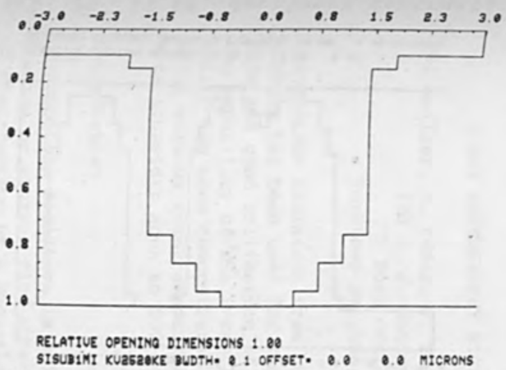


Fig. 14. 25 kV, developer contrast; A = 0.6

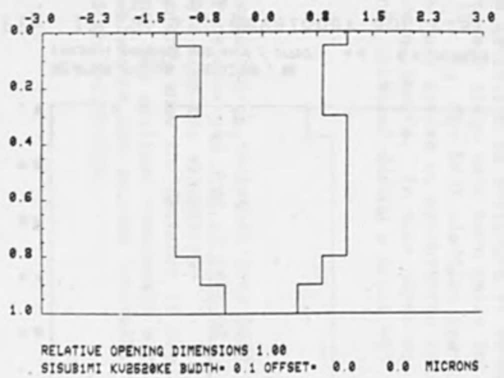


Fig. 15. 25 kV, developer contrast; A = 1.0

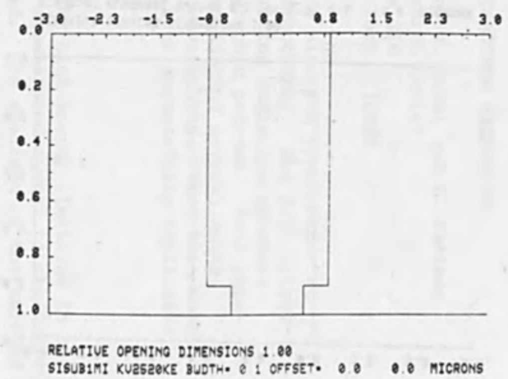


Fig. 16. 25 kV, developer contrast; A = 1.4

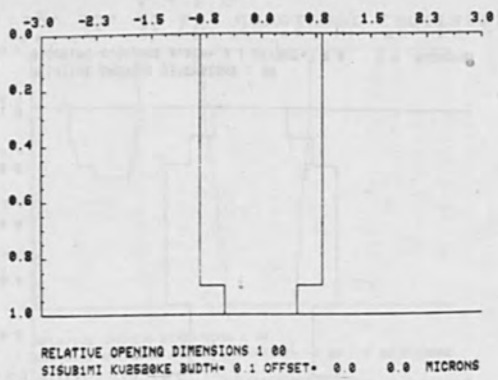


Fig. 17. 25 kV, developer contrast;
A = 1.8

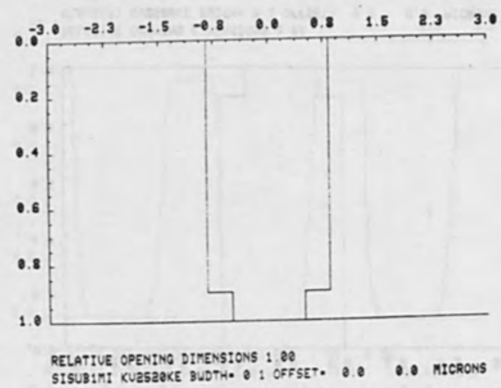


Fig. 18. 25 kV, B variations

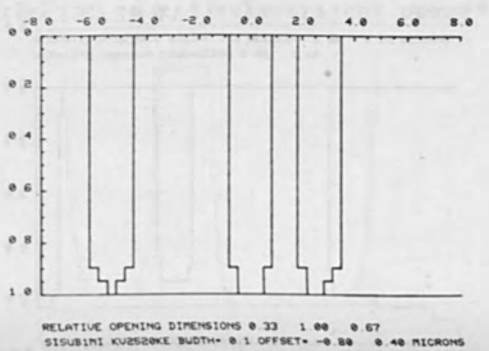


Fig. 19. 25 kV, developer contrast,
adjacent lines

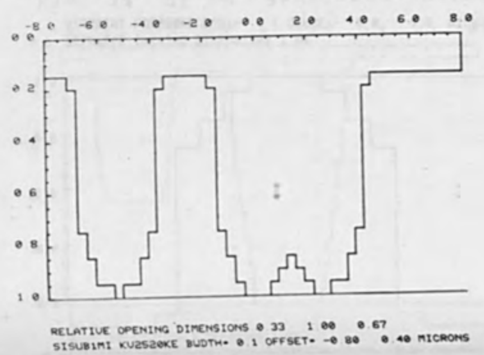


Fig. 20. 25 kV, developer contrast,
adjacent lines

X-RAY LITHOGRAPHY WITH SYNCHROTRON RADIATION

E. Spiller, R. Feder, J. Topalian, W. Gudat and D. Eastman
IBM T. J. Watson Research Center
Post Office Box 218
Yorktown Heights, New York 10598

Synchrotron radiation from the DESY electron synchrotron at Hamburg has been used for x-ray lithography. The high brightness and good collimation of sychrotron radiation enhances the capability of the x-ray lithographic process. This capability has been demonstrated by replicating several masks using various x-ray exposures and wavelengths. Very thin masks with linewidth down to 500 A have been successfully replicated.

I. Introduction

Synchrotron radiation is emitted by high energy electrons in a synchrotron or storage ring which are accelerated normal to the direction of their motion by a magnetic field. The spectrum of the emitted radiation extends from the microwave region throughout the infrared, visible, ultraviolet and into the x-ray region, decreasing in intensity below a critical x-ray wavelength λc (Å) = 186 Å/B (koe) E^2 (GeV) where E is the energy of the electrons in GeV and B is the magnetic flux in koe. This radiation which is highly collimated in the direction of the instantaneous electron motion is a uniquely bright x-ray source.

At the present time synchrotron radiation is available at several laboratories where synchrotron or storage rings have been built for experiments in high energy physics [1, 2]. The DESY electron synchrotron in Hamburg, Germany, offers easy access to synchrotron radiation with relatively simple vacuum requirements. In this paper we describe x-ray lithography experiments obtained during a visit at DESY.

II. Synchrotron Radiation at DESY

Fig. 1 shows the intensity of synchrotron radiation from DESY for different energies of the electrons averaged over an area of 20 mm x 10 mm for a distance of 37.9 m from the synchrotron [3]. The curves for the lowest electron energy show the decrease in the intensity for wavelengths shorter than the critical wavelength e.g. λc = 3 Å at E = GeV. For longer wavelengths the emitted intensity depends only weekly on the energy of the electrons.

At DESY the radiation laboratory is about 40 m away from the source and has two beam lines. One beam is a direct and the other is obtained by reflection from a gold mirror at 4° grazing incidence. The gold mirror acts as a short wavelength cutoff which eliminates the hard x-ray components ($\lambda \lesssim 10$ A) of the spectrum. Radiation safety requires that the experimental area is evacuated before the beam shutter is opened; therefore all experiments are remotely controlled. For our lithographic studies we used two wheels with eight openings to mount mask-wafer combinations, filters or detectors; each wheel being driven by a remote-controlled motor as to move samples into the beam. By using the two wheels behind each other and keeping one position in the first wheel open we are able to sequentially irradiate 15 different wafers during one pump-down cycle.

The DESY synchrotron operated at 3.5 GeV and 8 ma for most of the time during our visit. The effective exposure of the x-ray resist is determined not only by the emission from the synchrotron but also by the sensitivity of the resist, the transmission of the mask substrate and of the absorber in the mask. Fig. 2 shows the effective exposure of x-ray resist in the direct beam. Most of our masks are on 6μ thick mylar substrates; curve f shows that the effective exposure for such a substrate is centered around $\lambda = 8$ A, the wavelength most widely used for x-ray lithography with conventional sources (i.e. AlKα). The total exposure of the resist is given by the area under the curves in Fig. 2. A comparison of the areas of curves f and g shows, that resist under a 0.8μ thick film of gold still obtains a considerable exposure, mainly from the shorter wavelengths of the synchrotron radiation. A mask with a 0.8μ thick gold absorber pattern has a transmission of 0.01 for AlKα-radiation ($\lambda =$ 8.3 A), in the absorbing sections. For DESY at 3.5 GeV the effective transmission is 0.16, therefore the effective contrast of the mask is much lower than for AlKα radiation.

The contrast of all masks is drastically increased by the use of the 4° beam line, which does not give any radiation below $\lambda = 10$ A. The calculated exposure spectra in this beam line are given in Fig. 3; for the calculations the reflectivity of the 4° gold mirror has been obtained from optical constants of gold given in Ref. 4.

A comparison of the curves (a) and (b) in Fig. 2 and 3 shows that the resist is not exposed uniformly throughout its depth if exposed to the unfiltered beam from the synchrotron. There is a large contribution to the exposure from very soft x-rays, these soft x-rays, however, expose only the resist close to the top surface.

III. Experimental Results

a) Mask Replication

Mask patterns have been replicated in x-ray resist with all the mask and filter combinations given in Figures 2 and 3. Due to the good collimation of synchrotron radiation, penumbral blurring is completely eliminated and the distance between mask and wafer can be made very large. The resist patterns in Fig. 4 are obtained with a mask wafer spacing of 0.54 and 1.04 mm. These patterns demonstrate, that although weak diffraction effects are visible, usable patterns with 1μ linewidth can be obtained even for very large distances between mask and wafer.

Fig. 5 shows x-ray replications obtained in the 4° beam. The absorbing areas on the masks used are very thin and do not allow usable replications with AlKα radiation or with the direct beam synchrotron radiation. In Fig. 5a the absorber pattern in the mask consists of 0.1 μm thick gold, the smallest linewidth is 700 Å, and is replicated with vertical resist walls. In Fig. 5b, a diatom has been used as a mask and well resolved features of the order of 500 Å in width are seen in the replica.

b) X-Ray Resist

Polymethyl methacrylate is the most widely used resist for electron beam and x-ray lithography. Tests with AlKα radiation have shown that its dissolution rate increases steeply with increasing exposure. Even at the highest exposure levels no sign of any saturation in the dissolution rate versus exposure curve has been previously detected with conventional x-ray sources [5, 6]. Much higher exposure levels than with AlKα radiation can easily be obtained with synchrotron radiation and we have tested various exposures on PMMA up to 10^5 J/cm^3. At the high exposure levels of our test, the dissolution rate increases very rapidly with exposure and no developer concentration covers the entire region of exposures. We used methyl isobutyl ketone (MIBK) diluted with isopropylalcohol (IPA) as developer. Starting with low exposures and high developer concentrations we switched to a lower developer concentration as soon as our resist film (thickness about 4000 Å) developed completely upon immersion in the developer. For a developer concentration of 1:1 MIBK-IPA we find a dissolution rate of 10^4 Å/min for an exposure of 2×10^3 J/cm^3. The same dissolution rate is obtained for an exposure of 4×10^3 J/cm^3 and a developer concentration of 1:3 MIBK-IPA. We have observed the highest dissolution rate (more than 10^5 Å/min in pure IPA) for an exposure of 4×10^4 J/cm^3. The dissolution rate is decreasing at still higher exposures. We have found for an exposure of 1.5×10^5 J/cm^3 that the exposed resist is less soluable than the unexposed using concentrated MIBK as a developer. PMMA at these high exposures acts as negative resist in the same way as observed for high electron beam exposures[7].

c) Radiation damage

A mask with a mylar substrate has been exposed to the full beam of the synchrotron for a total dose equivalent to 10^4 "vertical wall" mask replications with PMMA resist. The total exposure was interrupted several times in order to replicate the mask on a silicon wafer. The standard deviation of the distance of two equivalent marks on all the silicon wafers was 0.25μ over a distance of about 1 cm and was the same as that obtained from other masks which had been exposed only to very low radiation doses. Radiation damage to the mask evident as a yellow color could, however, be detected in the mask at the end of the experiment.

d) Wafer Heating

Heating of the mask-wafer combination due to the high intensity of the synchrotron radiation and a consequent loss of resolution is a potential disadvantage of a synchrotron radiation source compared to a conventional source of much lower power. We have measured the temperature rise of a 1" diameter silicon wafer exposed to the full beam of the synchrotron (power $0.195 W/cm^2$ over 10×20 mm^2) in vacuum or in an atmosphere of 1 Torr helium. The final temperature rise for very long exposure was 2.5° and 0.2° C respectively, with a heat sink of 35g copper in contact with the back of the silicon wafer in both cases. The same heat sink was also used for all mask replications. No deleterious effect of this temperature rise on the resolution of the replica was observed in any case. The time constant of the temperature rise was longer than the required exposure time. Therefore, exposure to synchrotron radiation of higher intensity will not result in larger temperature rises during the required exposure time. In addition our temperature rises represent a worst case because we exposed the wafers to the full unfiltered beam of the high energy 3.5 GeV synchrotron, i.e. the short wavelengths which contribute only little to the exposure contributed fully to the heating. Less heating can be expected from a dedicated storage ring where the emission spectrum is tailored to the sensitivity of the resist.

Conclusion

Experiments performed at the DESY synchrotron in Hamburg have shown that synchrotron radiation is very suitable for x-ray lithography. The high intensity and good collimation of synchrotron radiation simplify the lithographic process considerably.

Penumbral blurring is completely eliminated and large distances between mask and wafer can be tolerated (as much as 1 mm for 1μ linewidth patterns). DESY is not optimized for x-ray lithography, its spectrum contains very hard radiation for which thin high resolution masks have low contrast. For this reason we produced mask replicas not only with the full beam but also modified the beam to obtain a more optimized spectrum. Even though DESY operated during our visit at a low current of 8 ma (1/10 of the maximum) and masks were 40m from the source, exposure times ranged from about 30 seconds to 3 minutes for the replication with vertical walls.

Based on our studies, we conclude that a dedicated storage ring, optimized for x-ray lithography could have a modest electron energy between 600 and 1000 MeV and still allow exposure times under one second. Such a dedicated storage ring will be considerably smaller (several m diameters) and cheaper (about \$ 2 x 10^6) than the existing high energy machines and at the same time deliver a higher flux. The brightness and collimation of the emitted radiation in the wavelength region interesting for x-ray lithography (5 Å < λ < 100 Å) would be several orders of magnitude higher than that of the best alternate source (e.g. rotating anode x-ray tube) thereby allowing shorter exposure times, larger throughput, less critical resist exposure conditions, and simplified geometrical conditions for applications requiring registration. Penumbral blurring can be completely eliminated.

References

1) M. L. Perlman, E. M. Rowe and R. E. Watson
Physics Today, July 1974, p. 30

2) C. Kunz, Phys. Bl. 32, p. 9 (1976) and p. 55 (1976)

3) E. E. Koch and C. Kunz, Synchrotronstrahlung bei DESY
DESY Report, March 1974

4) H. Y. Hagemann, W. Gudat, and C. Kunz, J. Opt. Soc. Am. 65,742 (1975)

5) E. Spiller, R. Feder, Y. Topalian, E. Castellani, L. Romankiw, and M. Heritage, Solid State Technology, April 1976, p. 62

6) H. Smith, D. L. Spears and S. E. Bernacki, J. Vac. Sci. Technol. 10,913 (1973)

7) M. Hatzakis, J. Electrochem. Soc. 116, 1033 (1969)

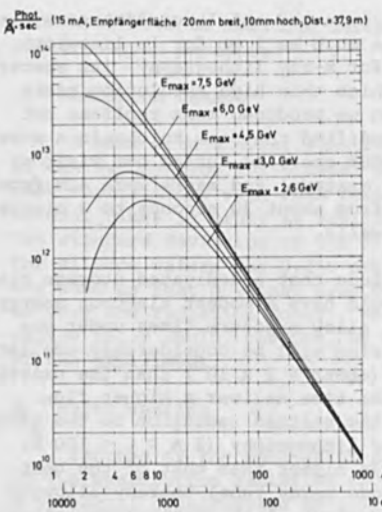


Fig 1. Intensity of synchrotron radiation of DESY (photons/A sec) averaged over an area 20 mm wide and 10 mm high at a distance of 37.9 m from the electron orbit for a beam current of 15 ma (From Ref. 3).

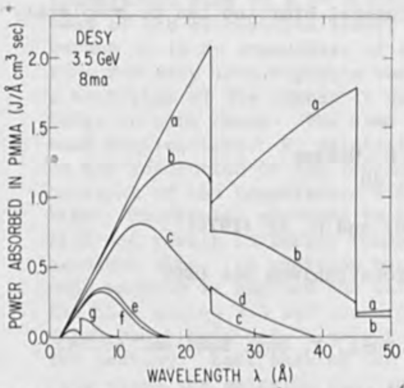


Fig. 2 Power absorbed in resist (PMMA) as a function of wavelength for synchrotron radiation from DESY (3.5 GeV, 8 ma) 40 m from electron orbit for various filters in the beam. (a) no filter; (b) 0.3 μ of PMMA; (c) 1.5 μ of Parylene N; (d) 1 μ of PMMA; (e) 12 μ of beryllium; (f) 6 μ of mylar; (g) 6 μ of mylar + 0.8 μ of gold.

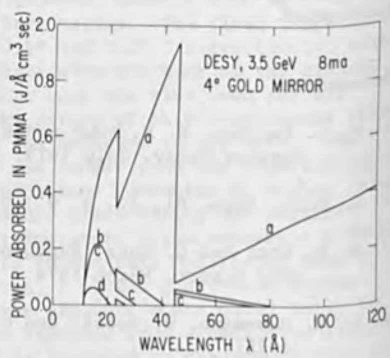


Fig. 3 Calculated power absorbed in resist (PMMA) for synchrotron radiation from DESY (3.5 GeV, 8 ma) after reflection from gold at a glancing angle of 4°. (a) no filter (b) filter of 1 μ thick PMMA, (c) 1.5 μ of parylene N (d) 2.5 μ of Mylar.

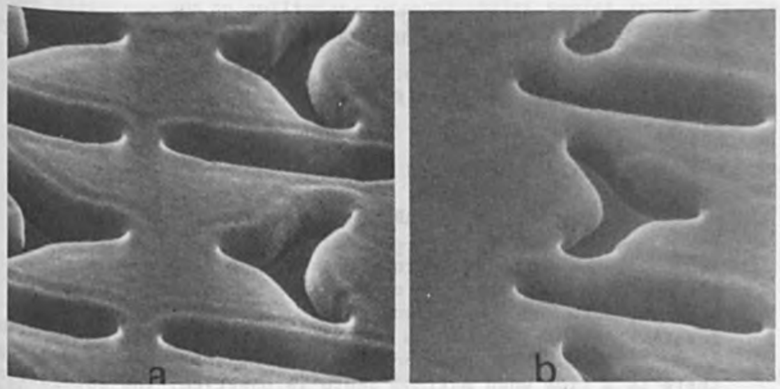


Fig. 4 Mask replication (1 μ linewidth) with large distances (a: 0.54, b: 1.04 mm) between mask and wafer. Direct beam filtered with 12 μ thick Be + 6 μ mylar. Exposure time 3 min, effective exposure 250 J/cm, resist: PMMA 2041.

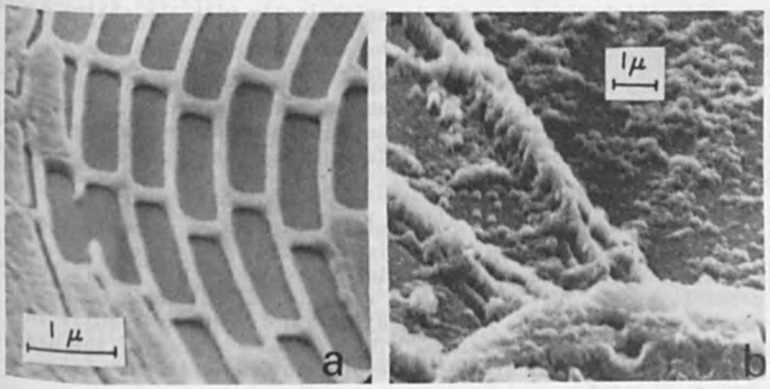


Fig. 5 Replication with soft radiation in the 4°-beam. a) Section of a Fresnel zone plate with a smallest linewidth of 700 Å. The absorber in the mask was gold of 1000 Å thickness, exposure time: 25 min, effective exposure 3000 J/cm^3, resist PMMA. b) Detail of a diatom. Exposure time 10 min, effective exposure about 1200 J/cm^3, resist PMMA.

A STUDY OF ELECTRON BEAM EXPOSURE OF POSITIVE AND NEGATIVE RESISTS

O. Ersoy
Central Institute for Industrial Research
Forskningsveien 1,
Oslo 3
Norway.

ABSTRACT

Two models based on bulk properties are developed to account for the observed behaviour of positive and negative resists. Both exposure and development are considered as a whole. In positive resists, the depth of observed profile increases with time, charge density and decreases with initial electron energy; there is a time when maximum depth is observed; the saturation of development is a strong function of the film thickness; and complete development time varies inversely with charge density. In negative resists, the insoluble thickness as a function of charge density is also a strong function of the total film thickness, and the initial electron energy. Reducing the film thickness and increasing the initial electron energy have the effect of making insolubilization more difficult. At the limit when the film thickness is very small or the initial electron energy very large, the ratio of insoluble thickness to the total film thickness increases logarithmically with the charge density.

1. INTRODUCTION

Recently the electron beam exposure of resists have been of considerable interest in order to fabricate electronic, light optical and holographic devices. This, in turn, depends on the proper combination of resist properties, electron beam parameters and development techniques.

In long-chain polymers containing thousands of chemical bonds, radiation has two effects: One is the linking together of the molecules, thereby increasing the molecular weight; this is also called cross-linking. The other is the scission of individual molecules. According to which one of these effects dominate, it results in negative and positive resists. The exposed portions of the negative resists become insoluble after development, whereas they become soluble for positive resists.

The penetration of electrons in matter has only been recently placed on a solid basis [1], [2], [3]. Earlier works produced erroneous results since they were based on the Thomson-Widdington law. Even though this law is mathematically correct, the range of electrons calculated in this way are not quite reasonable since scattering in various directions is neglected. The energy loss by the electrons due to collisions with molecular nuclei including the effect of multiple-scattering has been empirically found to be [2]

$$\frac{dE}{dz} = -\frac{Eo}{R_G}\Lambda(f) \tag{1}$$

where E_o is the initial energy of the electrons in keV, $\Lambda(f)$ is the depth-dose function given by:

$$\Lambda(f) = .74+4.7f-8.9f^2+3.5f^3$$
$$f = \frac{z}{R_G} \tag{2}$$

z is the direction of electron bombardment, and R_G is the Grun-range in microns, given by [4]

$$R_G = \frac{.046}{\delta} Eo^{1.75} \tag{3}$$

where δ is the density of the solid in g/cm^3. The depth-dose function is plotted in Figure 1.

2. POSITIVE RESISTS

The absence of image swelling and shrinking that is observed with negative resists is a definite advantage for positive resists, contributing to the excellent dimensional fidelity and sharp edge definition of the images.

Experimentally the following is observed with positive resists:

1. Thin films of resists are soluble in the developer in a saturable fashion. In other words, solubility decreases as the film gets thinner [5].
2. Solubility increases under electron exposure. The sensitivity of the resist decreases with increasing initial electron energy Eo. The saturation behaviour as a function of film thickness is also observed in this case [5], [6].

The model to be developed is based on the above observations and is represented by the model shown in figure 2. We will assume that the mth developed thickness Δ is proportional to the time t_m and the total available number of scissions. The word scission is used in the sense of certain broken bonds or rearranged molecules that make the resist more soluble in a given resist. Different numbers of scissions are effective for different solvents. Perhaps a better word is the number of available sites rather than the number of scissions. The constant of proportionality will be taken to be a function of z. We pay attention to the fact that what is already developed cannot affect what is to be developed. Thus, for the m^{th} Δ, we can write:

$$\Delta = t_m S_{(m-1)\Delta}^{dt} K(z)dz \qquad (4)$$

where K(z) can be written as

$$K(z) = K_1(z) \frac{QEo}{R_G} \wedge (f) + K_2(z) \qquad (5)$$

Q is the charge density of electrons in coulombs/cm^2, given by

$$Q = It_e/A \qquad (6)$$

I is the current of the electron microscope in amperes, t_e is the exposure time in seconds and A is the area of the beam in cm^2. The first term in equation (5) is due to electron bombardment and the second due to residual sites.

Since the total number of Δ's is assumed to be n, the total developed thickness d can be written as:

$$d = nt_n S_{(n-1)\Delta}^{dt} K(z)dz \qquad (7)$$

Since Δ's are constant we can also write:

$$\frac{tm}{tn} = \frac{S_{(n-1)\Delta}^{dt} Kdz}{S_{(m-1)\Delta}^{dt} Kdz} \qquad (8)$$

and

$$t = t_n \left[\frac{1}{S_o^{dt} Kdz} + \frac{1}{S_\Delta^{dt} Kdz} + ---- + \frac{1}{S_{(n-1)\Delta}^{dt} Kdz} \right] S_{(n-1)\Delta}^{dt} Kdz \qquad (9)$$

Using equations (8) and (9), equation (7) can be written as

$$t = \int_0^d \frac{dx}{\int_x^{d_t} K(z)dz} \tag{10}$$

This is the general expression to find the developed thickness d as a function of charge density Q, initial electron energy E_o, film thickness d_t, and development time t.

We note that equation (9) could also be written as:

$$\begin{aligned} d = t \int_d^{d_t} K dz &+ t_1 K(d_t)dz \\ &+ (t_1+t_2)K(d_t-dz)dz \\ &+ \\ &\vdots \\ &+ (t-t_n)K(d_t-(d_t-d))dz \end{aligned} \tag{11}$$

If we neglect the infinitesimal terms, we obtain:

$$d \doteq t \int_d^{d_t} K(z)dz \tag{12}$$

This expression can only be used when d is much smaller than d_t. Equation (12) can also be found from equation (10) as d approaches zero.

It is clear that equations (10) and (12) cannot be solved unless we know the functions $K_1(z)$ and $K_2(z)$. We can expect them to be nearly constant if the interactions between different sites are strong, or somewhat exponentially decreasing if the interactions are weak. Choosing K_2 to be constant gives results that match the experimental data, as shown in the next section. However, the choice of K_1 is found to be much more difficult, as discussed in the fourth section.

3. DEVELOPMENT WITHOUT ELECTRON EXPOSURE

In this case equation (10), with constant K_2, gives:

$$\frac{d}{d_t} = 1 - e^{-tK_2} \tag{13}$$

if we had used equation (12) instead we would obtain:

$$\frac{d}{d_t} \doteq \frac{K_2 t}{1+K_2 t} \tag{14}$$

For small t, we can write

$$1-e^{-tK_2} = K_2 t - \frac{(K_2 t)^2}{2} + \text{higher terms} \tag{15}$$

$$\frac{K_2 t}{1+K_2 t} = K_2 t - (K_2 t)^2 + \text{Higher terms}$$

We see that t has to be quite small for equation (14) to be reasonable.

Equation (13) was compared to Hatzakis' experimental data i polymethylmethacrylate [3]. Method of least squares and Gauss-Newton iterations were used to find the best value of K_2. Equation (15) shows that K_2 can also be estimated from the slope of the experimental curve as t goes to zero. The results are shown in figure 3.

4. DEVELOPMENT AFTER ELECTRON EXPOSURE

If K_1 is chosen to be constant, equation (10) gives:

$$\frac{d}{d_t} = \frac{e^{tc}-1}{e^{tc}+\frac{1}{1+B}} \qquad R_g >> 1.23 d_t \tag{16}$$

where

$$C = \left(1 + \frac{6.35 d_t}{R_G}\right) \frac{.74 K_1 E_o Q}{R_G} + K_2 \tag{17}$$

$$B = \frac{R_G}{3.176 d_t}\left(1 + \frac{K_2 R_G}{.74 K_1 E_o Q}\right) \tag{18}$$

According to equation (16), development will be over when $e^{tc} >> 1$. If we define complete development as:

$$\frac{d}{d_t} = 1 - \frac{1}{r} \tag{19}$$

where r is a large number, the complete development occurs when

$$t_c \doteq \frac{\log r}{C} \tag{20}$$

If K_1 is chosen to be

$$K_1(z) = K_1\ \delta(z-x) \tag{21}$$

which means assuming no interaction between sites, we get

$$\frac{d}{d_t} = \frac{2\left(\frac{.74}{d_t} + \frac{K_2R_G}{K_1QE_o}\right)}{|b - \sqrt{q}|} \quad \frac{1 - \frac{\exp(tK_1QE_o\sqrt{q})}{R_G}}{\frac{b + \sqrt{q}}{|b - \sqrt{q}|} + \frac{\exp(tK_1QE_o\sqrt{q})}{R_G}} \tag{22}$$

where

$$c = \frac{8.9}{R_G^2} \tag{23}$$

$$b = \left(\frac{4.7}{R_G} - \frac{K_2R_G}{K_1QE_o}\right) \tag{24}$$

$$a = \left(.74 + \frac{K_2R_Gd_t}{K_1E_oQ}\right) \tag{25}$$

$$q = b^2 - 4ac \tag{26}$$

In this case, complete development time occurs when

$$t_c \doteq \frac{R_G}{K_1E_oQ}\ \frac{1}{\sqrt{q}} \tag{27}$$

The behaviour of equations (16) and (22) and experimental data is shown in figure 4. It is clear that assuming K_1 constant leads to overestimating saturation phenomenon whereas the second assumption leads to the under estimation of the same phenomenon. Thus, a middle course needs to be chosen. For example, we can choose

$$K_1(Z) = K_1 \, e^{-\alpha(z-x)^2} \qquad (28)$$

However, this would mean the introduction of a third constant α. Thus, the resulting expression would be rather difficult to evaluate.

Let us go back to equation (10). When Q is sufficiently large so that $K_1 QE_o \vee / R_G \gg K_2$, we can write

$$\frac{t\,QE_o}{R_G} = \int_o^d \frac{dx}{\int_x^{dt} K_1 \, dz} \qquad (29)$$

or

$$\frac{t\,QE_o}{R_G} = f(d, d_t, R_G) \qquad (30)$$

where the function f is equal to the integral in equation (29). It decreases with R_G and d_t. However, if R_G is large, the dependence on R_G is negligible. Complete development occurs when:

$$t_c = \frac{R_G}{QE_o} \, f(d_t, d_t, R_G) \qquad (31)$$

In other words, complete development time varies inversely with charge density. Comparison with experimental data is shown in figure 5. It is this basic dependence that causes equations (20) and (22) to be insensitive to the choice of K_1. Both equations behave very much like equation (31).

When d is very small, we can use quation (12). The net observed development thickness d_{net} will be:

$$d_{net} = \frac{tQE_o}{R_G} \int_d^{dt} K_1 \, dx \qquad (32)$$

In other words, developed thickness increases less than linearly with charge density since the integral above is reduced as d increases.

Equation (10) also shows that observed profile will be negligible when

$$\frac{K_1 Q E_o}{R_G} \ll K_2 \qquad (33)$$

This should be considered as a rule of thumb since the variations of K_1 and K_2 were found to be quite different.

5. NEGATIVE RESISTS

Negative resists are more available commercially, and their properties are more extensively studied as a consequence.

The following experimental facts are observed with negative resists [7], [8], [2], [9].

1. Negative electron resists show a threshold in exposure, no insolubilized layer forming below the threshold charge density Q_s.

2. Above the threshold, an increase in the charge density Q gives an increase in the insoluble resist thickness in a saturable fashion; that is, t increases less than linearly as it approaches d_t and as Q is increased; maximum thickness is achieved at a certain charge density Q_e.

3. Increasing E_o has the effect of reducing the sensitivity of the resist; both Q_s and Q_e increase; the saturation of t with Q becomes more pronounced, and, as a result, the contrast function $\Gamma = Q_s/Q_e$ becomes less.

4. Reducing the total resist thickness d_t gives results that are somewhat similar to increasing the initial electron energy E_o.

5. Development of the exposed resist causes first swelling and then shrinking of the resist so that the maximum insoluble thickness achieved is usually less than the initial thickness of the resist.

It has been rather difficult to explain the observed properties of negative resists as given above as a function of experimental parameters on a solid theoretical basis [7], [8], [2].

For several types of the initial molecular weight distribution of the polymer substance, Flory [9] and Charlesby [10] derived theoretical formulae that give the insoluble fraction g as a function of the cross-linking index γ , defined to be the average number of cross-links per molecule, based on statistical considerations. When all molecules are initially the same size,

$$g = 1 - \exp(-\gamma g) \tag{34}$$

When the molecules follow a Poisson distribution,

$$g = 1 - 1/(1+\gamma g)^2 \tag{35}$$

However, it has been found difficult to apply equations (34) and (35) to experimental data [8].

Taking the number of cross-links to be proportional to the amount of energy lost by the electrons, Heidenreich et al [2] proposed that an infinitesimal thickness of the resist at a depth z becomes insoluble when the number of cross-links are equal to a constant:

$$\frac{QE_o}{R_G} \Lambda(z) = C \text{ (constant)} \tag{36}$$

Then we can write

$$Q_s = \frac{R_G}{E_o} \frac{C}{\Lambda(d_t)} \tag{37}$$

$$Q_e = \frac{R_G}{E_o} \frac{C}{.74} \tag{38}$$

$$\Gamma = \frac{.74}{\Lambda(d_t)} \tag{39}$$

Provided that $d_t \leq .328R_G$. When $d_t > .328R_G$ the "thick film" case is obtained meaning that a certain thickness $(d_t - .328R_G)$ becomes insoluble all at once at Q_s':

$$Q_s' = \frac{R_G}{E_o} \frac{C}{\Lambda(.328R_G)} \tag{40}$$

Equations (36), (37) and (38) do behave like the experimental facts outlined above to a certain extent, but there are 3 serious short-commings: the curve of t versus Q as obtained from equation (36) is much more steep than the experimental curve; the contrast function Γ is much larger than the experimental value; and theoretical Q_e does not change with d_t, whereas it tends to increase with decreasing d_t at low electron energies E_o. These discrepancies indicate the feasibility of a bulk effect, as was the case with positive resists.

6. THE MODEL FOR NEGATIVE RESISTS

In order to show this feasibility, we are going to make some logical assumptions. These are:

1. The incremental insoluble thickness dt at depth z is directly proportional to the total number of molecules because more molecules mean more number of available sites for cross-linking that causes insolubilization.

2. It is also directly proportional to the incremental increase in energy lost by the electrons at depth z.

3. It is inversely proportional to the average amount of energy lost by electrons per unit volume because as the number of cross-links increase, it becomes more difficult to find available sites for cross-linking that causes insolubilization, and the average number of cross-links per unit volume can be roughly taken to be proportional to the average amount of energy lost per unit volume by the electrons.

Above we were careful to differentiate between cross-links that cause insolubilisation and cross links in general because some cross-linking does not cause insolubilization [10].

Now we can write

$$dt \propto \frac{\Lambda(d_t - t)\,dQ}{\dfrac{d_t}{QS_o}\,\Lambda(z)dz}\,d_t^{\,2} \tag{41}$$

or

$$y(t) = A\log Q - B \tag{42}$$

$$y(t) = \frac{1}{d_t{}^2} \int_0^{d_t} \Lambda(z)dz \int_{d_t-t}^{d_t} \frac{dz}{\Lambda(z)} \tag{43}$$

Fitting of equations (42) and (43) to the experimental data with the help of the method of least-squares as shown in figures 6 thru 10 looks encouraging. However, figure 8 shows that as the total thickness of the resist d_t increases, some discrepancy arises, indicating the possibility of a range for molecular interactions. Figure 6 indicates that the constant A increses slowly with decreasing initial electron energy E_o. Thus we can anticipate that A is a constant in the sense of a statistical average rather than a universal one.

It is interesting to observe that in the limit $d_t \to 0$ or $E_o \to \infty$, equations (42) and 43) reduce to

$$\frac{t}{d_t} = A\log Q - B \tag{44}$$

We plotted the threshold values for KTFR as obtained from figure 6 in figure 11 and compared them with equation (37). If we accept the performance of equation (37) as reasonable, then the constant B can be approximated by

$$B \sim A\log \frac{R_G}{E_o} \frac{C}{\Lambda(d_t)} \tag{45}$$

At E_o = 5keV, the Grun range for KPR and KTFR (δ = .9g/cm^3) is .885 um so that we can expect t = .32 um to become insoluble all at once if we accept the "thick film" concept as indicated by equation (40). Figure 9 indicates this possibility for KPR, whereas figure 6 does not for KTFR. However, figures 7 and 10 show that the thickness versus charge density curves are so steep at E_o = 5keV so that any slight error in measurements near the threshold value can cause serious misjudgement.

7. CONCLUSIONS

In the course of this work, we tried to analyze the experimental behaviour of positive and negative resists as a function of film thickness, charge density and initial electron energy. We proposed that their behaviour can be explained in terms of "bulk" properties.

The basic results can be considered to be equations (10), (13), (31), (37), (42), (43), (44) and (45).

For positive resists, we can conclude that:

1. Positive resists are soluble in a developer without electron exposure, as described by equation (13). The solubility will change with the concentration of the developer.

2. Electron exposure increases the solubility in a fashion described by equation (10). Since exposed and unexposed portions develop simultaneously, there will be a time at which maximum profile depth is obtained. Further development will reduce this depth.

3. Complete development time varies inversely with charge density, as shown by equation (31).

For negative resists, we can conclude that:

1. The threshold charge density at which insoluble layer starts to form increases with initial electron energy and decreases with total film thickness as described by equation (37).

2. The relationship between the insoluble thickness and charge density is reasonably described by equation (42). The final charge density at which the total thickness becomes insoluble increases with initial electron energy and decreases with the total film thickness.

3. When the total film thickness becomes very small or the initial electron energy becomes excessively large, we can expect the ratio of the insoluble thickness to the total thickness to increase logarithmically with the charge density as described by equation (44).

As a final conclusion, we point out that:

The quantity QE_o/R_G or $Q/E_o^{.75}$ should be considered a figure of merit for resists. Sensitivity varies in a direct fashion with it. Thus, increasing Q and reducing E_o have similar effects.

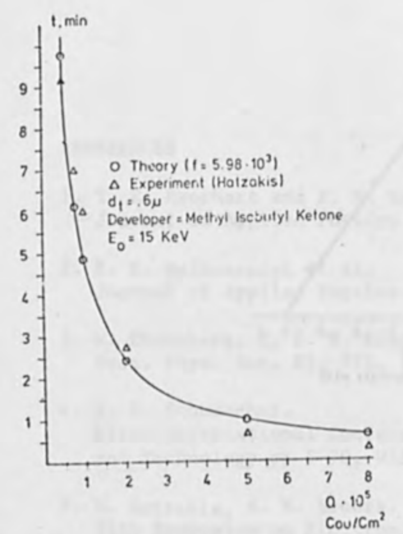


Fig. 5. Complete Development Time versus Charge Density in Polymethylmethacrylate.

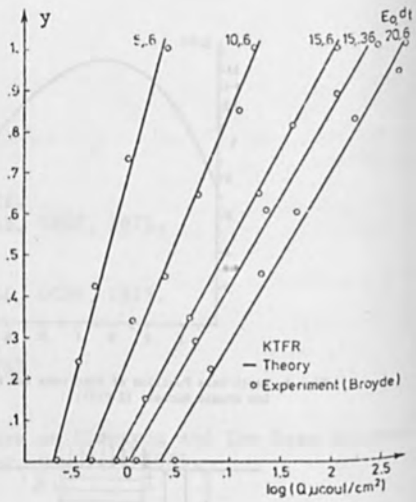


Fig. 6. Exposure of KTFR films (initial thickness .6 and .36 μm) to kilo-volt electrons of 5, 10, 15 and 20 keV.

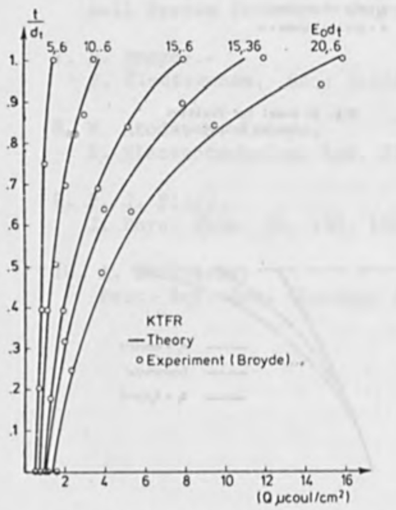


Fig. 7. Exposure of KTFR films (initial thickness .6 and .36 μm) to kilo-volt electrons of 5, 10, 15 and 20 keV.

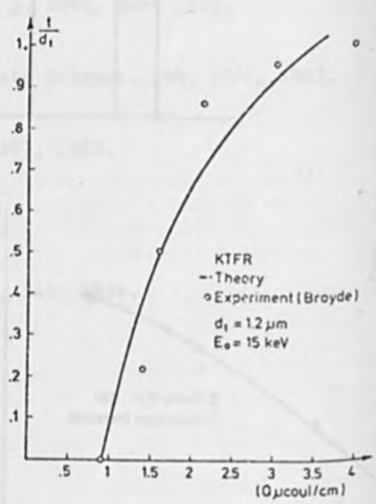


Fig. 8. Exposure of 1.2 μm thick KTFR film to 15 keV electrons.

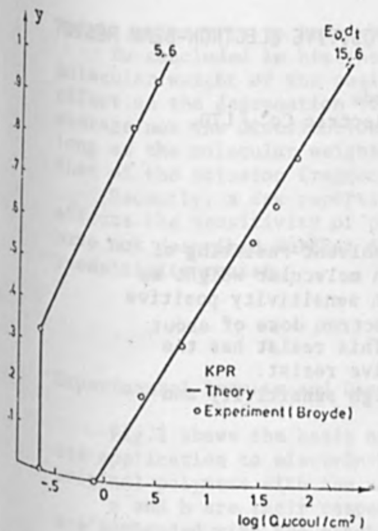


Fig. 9. Exposure of .6μm thick KPR films to 5keV and 15 keV electrons.

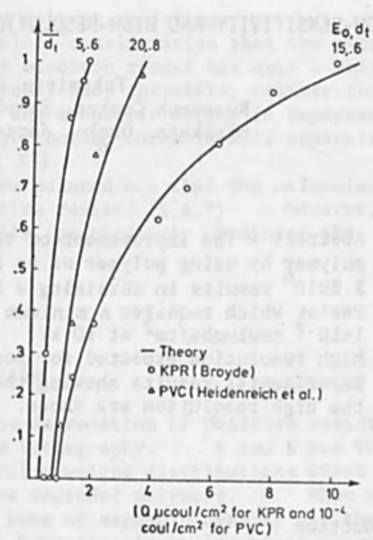


Fig.10. Exposure of .6 μm thick KPR films to 5 keV and 15 keV electrons, and 8 μm thick PVC film to 20 keV electrons.

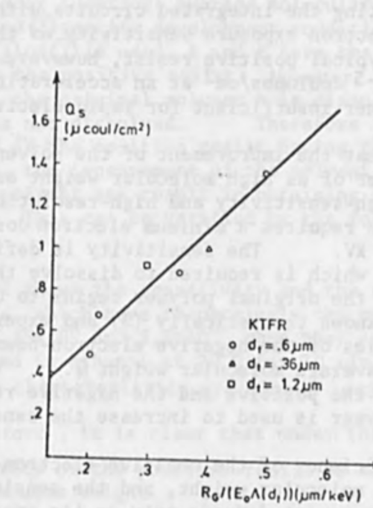


Fig.11. Threshold charge density as a function of energy lost per electron at full depth in verying thicknesses of KTFR.

HIGH-SENSITIVITY AND HIGH-RESOLUTION POSITIVE ELECTRON-BEAM RESIST

Takamitsu Kuroda
Research Center, Sanyo Electric Co., LTD.
Hirakata, Osaka, Japan

Abstract - The improvement of the solvent-resisting of polymer by using polymer of as high molecular weight as 3.8×10^6 results in obtaining a high sensitivity positive resist which requires a minimum electron dose of about 1×10^{-7} coulombs/cm^2 at 10 kV. This resist has the high resolution expected for positive resist. Experimentsl results showing the high sensitivity and the high resolution are shown.

Introduction

The positive electron-beam resists using the degrading polymers have been known to be superior to the negative electron-beam resists in resolution for fabricating the integrated circuits with submicron patterns (1,2) The electron exposure sensitivity of the polymetyl methacrylate which is a typical positive resist, however, has been reported to be about 5×10^{-5} coulombs/cm^2 at an accelerating voltage of 10 kV and still is rather insufficient for rapid electron-beam scanning (1,2,3)

This paper reports that the improvement of the solvent-resisting of polymer by using polymer of as high molecular weight as 3.8×10^6 results in obtaining a high-sensitivity and high-resolution positive electron-beam resist which requires a minimum electron dose of about 1×10^{-7} coulombs/cm^2 at 10 kV. The sensitivity is defined here as the minimum electron dose which is required to dissolve the exposed polymer completely before the original polymer begins to dissolve.

It has been already known theoretically (3) and experimentally (3,4) that the sensitivities of the negative electron-beam resists increase with increasing average molecular weight $\bar{M}$. The fundamental difference between the positive and the nagative resists when high molecular weight polymer is used to increase the sensitivity is described below.

The crosslinking efficiency of the negative electron-beam resist increases with increasing molecular weight, and the sensitivity of the negative electron-beam resist depends only on its crosslinking efficiency and is independent of the development conditions. On the contrary, the degradation efficiency of positive resist is independent of the molecular weight. The purpose of using a high molecular weight polymer is to improve the solvent-resisting.

Ku concluded in his theoretical consideration that the average molecular weight of the positive electron resist has only a very minor effect on the degradation efficiency and, therefore, neither the average nor the distribution of the molecular weight is important so long as the molecular weight distribution curve is well separated from that of the scission fragment. (3)

Recently, a few reports have pointed out that the molecular weight affects the sensitivity of positive resist. (5,6,7) However, they have not described further details, particularly obtaining the high -sensitivity resist.

Experimental Results and Discussion

Fig.1 shows the basis of the degradation of positive resist and its application to electron-beam lithography. A and B are the original polymers with the molecular weight distributions drawn below.

a and b are their respective degraded polymers. When A and B are bombarded with the electron beam of exposure density Q_1, they are degraded to a_1 and b_1, which both have the almost identical average molecular weight. When A and B are bombarded with the electron beam of exposure density Q_3, they become a_3 and b_3, which both have also the almost identical average molecular weight.

1, 2 and 3 specify development conditions. When development condition 1 or 2 is used, A and B have the same sensitivity when they are used as the positive resist. However, when development conditions 3 is used, the original polymer A is dissolved but the original polymer B is not dissolved. Therefore , the original polymer B can be used as the positive resist having the sensitivity of Q_3.

Thus , the improvement of the solvent-resisting of polymer by using high polymer may result in obtaining a high-sensitivity positive resist. This can be verified by the following experimental data.

Fig. 2 shows the sensitivity and the exposure characteristics of the polymers, which have respectively the average molecular weight $\bar{M}$ 3.5×10^5 , 1.35×10^6 and 3.8×10^6 when they are exposed to 10kV electrons and developed in ethanol at 20 C for 20 min.

These characteristics are almost identical within experimental errors.

Therefore , it is clear that under the same development conditions the sensitivity of the positive resist is almost independent of the average molecular weight.

Fig. 3 shows the effect of molecular weight on the solvent-resisting. The solvents used for determining the solvent-resisting consist of mixtures of rich solvent A and poor solvent B which were found to be suitable for developers. The mixtures are shown by the concentration of the rich solvent A in volume percentage.

The solvent-resisting was determined by measuring the dissolved thickness of the polymer film coated on a polished Si wafer. The curves show the highest solvent-resisting characteristics of the respective molecular weight polymers. The polymer of $\bar{M}$ 3.5×10^5 began to be dissolved in 5% A solvent mixture when the development was carried out at 20°C for 20 min. The polymer of $\bar{M}$ 3.8×10^6, however, was not dissolved in 30% A solvent mixture.

From these results, the higher polymer becomes undissolved in more powerful development conditions.

Fig. 4 shows the rate of dissolution of the exposed polymers and the original polymer.

The original polymer of 3.8 million molecular weight begins to be dissolved in about 22 min, while the exposed polymer of 1×10^{-7} coulombs/cm^2 begins to be dissolved in about 4 min, and the dissolution of the latter is speedier than the former. From the figure, it is clear that the exposed polymer 5000Å thick of 1×10^{-6} coulombs/cm^2 is easily developed before the original polymer begins to be dissolved.

The usable thickness is also calculated from the differece between the original polymer and the exposed polymer in dissolution characteristic.

Thus, this figure shows how the thin polymer film dissolves.

Fig.5 shows the effect of developers on the sensitivity of the positive resist. The characteristics were obtained by developing 1500±200 Å thick polymer films of $\bar{M}$ 8×10^6 at 20°C for 20 min in several kinds of developers: ethanol, solvent B, solvent mixtures containing 10%, 15%, 20%, and 25% solvent A and solvent B.

The crosslinking characteristic of KTFR (Kodak Thin Film Resist) determined under the same electron-beam exposure conditions is also shown in Fig. 5 for comparison. Although the original polymer began to be dissolved in 25% A solvent mixture, the 20% A solvent mixture dissolved completely the degraded polymer exposed with approximately 1×10^{-7} coulombs/cm^2 before the original polymer began to be dissolved. This resist is useful over the range of 10^{-4} coulombs/cm^2 to 10^{-7} coulombs/cm^2. The solvent-resisting of the polymer, however, depends on prebaking conditions, surface status of Si wafer, and other factors. Therefore, a positive resist used at the sensitivity of 1×10^{-7} coulombs/cm^2 is not yet practically stable.

However, the positive resist at the sensitivity of 1×10^{-6} coulombs /cm^2 can be used with high stability.

Fig. 6 is an optical microscopic photograph showing the high-resolution. This is a etched pattern of SiO_2 obtained by using the Sanyo Positive Resist which is exposed at 1×10^{-6} coulombs/cm^2 at 20 kV.

The left half of this photograph shows a SiO_2 etched pattern

which consists of about 0.2 μm wide lines separated by 2 μm.
The right half of this photograph shows the resist film with the developed pattern after separation from SiO_2.

Fig. 7 is also an optical microscopic photograph showing the high sensitivity. This is a developed pattern of Sanyo Positive Resist which is exposed at 1×10^{-7} coulombs/cm^2 at 20 kV.
The undeveloped space widths between the developed patterns are, respectively, 7 , 5 , 2 , 1 , and 0.5 μm.

Fig. 8 is an optical microscopic photograph of SiO_2 etched pattern obtained by using Sanyo Positive Resist. It is exposed at about 5×10^{-6} coulombs/cm^2 at 30 kV which is an optimum exposure for KTFR. The black area shows the Si surface, the gray area shows the SiO_2 surface.

Fig. 9 is an enlarged photograph of Fig. 8. The minimum line width is about 8 μm.

Fig. 10 is a photograph of the KTFR developed pattern exposed at the same exposure conditions as Fig. 9.

In the KTFR developed pattern, both spreading and narrowing are marked at the electron beam joints.
On the contrary, in the SiO_2 etched pattern obtained by using Sanyo Positive Resist, neither spreading nor narrowing is observed.

Conclusion

1 Using the high molecular weight polymer such as those of the average molecular weight about 2 to 4×10^6 is recomended for the positive resist.
2 The positive resist having a sensitivity of 1×10^{-7} coulombs/cm^2 is obtained by using the high molecular weight polymer and proper development conditions.
3 This high sensitivity positive resist has the high resolution expected for positive resist.
4 The fundamental difference between the positive and the negative resists when high molecular weight polymer is used to increase the sensitivity has been shown.

Acknowledgement

We would like to thank Dr.M.Yamano, Dr.E.Komatsu and Mr.K.Hinotani for their continuous guidance and encouragement.

We are also grateful to our colleagues Mr.H.Fukuda and Mr.A.Hakusui for their helpful assistance and discussion.

We wish to thank the Japan Electron Optics Laboratory Co. LTD for operating the Electron-beam Generator for Fig.6 and Fig.7.

References

(1) I. Haller, et al., IBM J. Research & Development, 12, 251 (1968)

(2) M. Hatzakis, J. Electrochem. Soc., 116, 1033 (1969).

(3) H. Y. Ku, et al., J. Electrochem. Soc., 116, 94 (1969).

(4) Y. Yatsui, et al., J. Electrochem. Soc., 116, 94 (1969).

(5) M. Hatzakis, et al., Sixth International Conference on Electron Ion Beam Science and Technology, 542 (1974)

(6) J. H. Rai et al., Symposium on polymer materials for microelectronic device fabrication at 170th Meeting of ACS, 252 (1975)

(7) M.J. Bowden et al., Symposium on polymer materials for microelectronic device fabrication at 170th Meeting of ACS, 258 (1975)

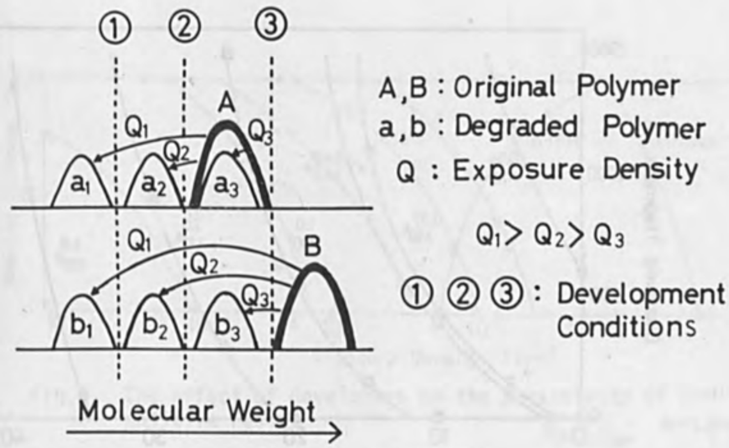


Fig.1 The basis of the degradation of positive resist and its application to electron-beam lithography

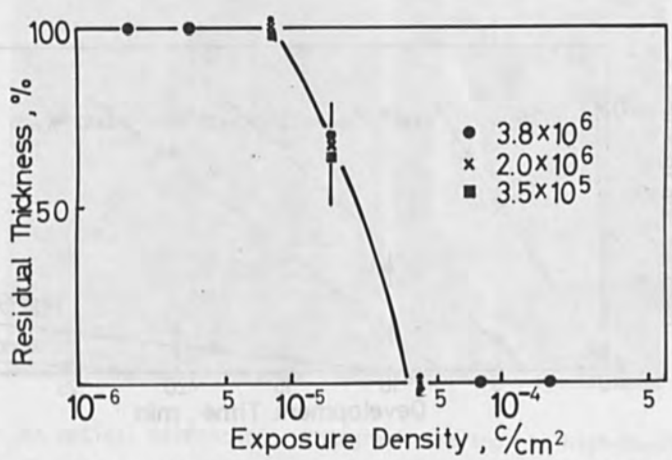


Fig.2 The sensitivity and the exposure characteristics of the polymers, which have respectively the average molecular weight $\bar{M}$ 3.5×10^5, 1.35×10^6 and 3.8×10^6 when they are exposed to 10kV electrons and developed in ethanol at 20°C for 20 min

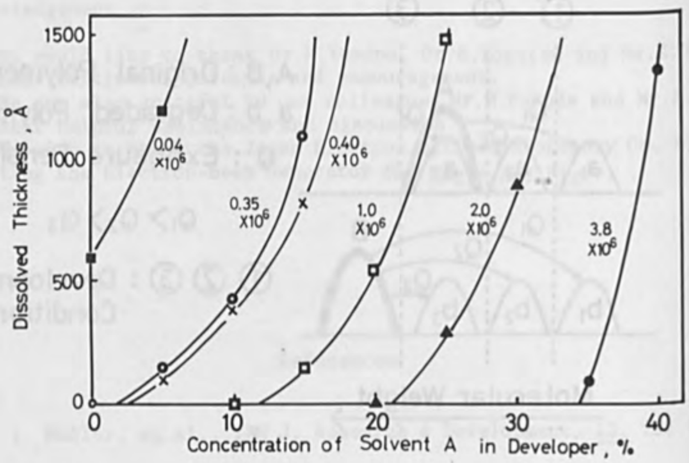


Fig.3 The effect of molecular weight on the solvent-resisting

20°C , 20 min

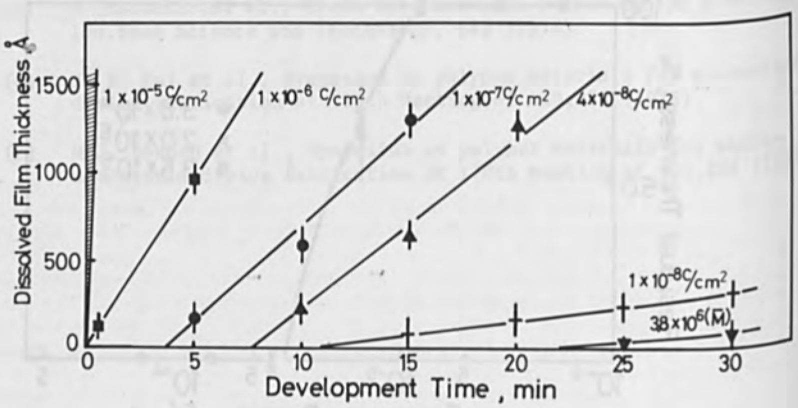


Fig.4 The rate of dissolution of the exposed polymers and the original polymer

$\bar{M}=3.8\times10^6$, 20°C , 20 % A

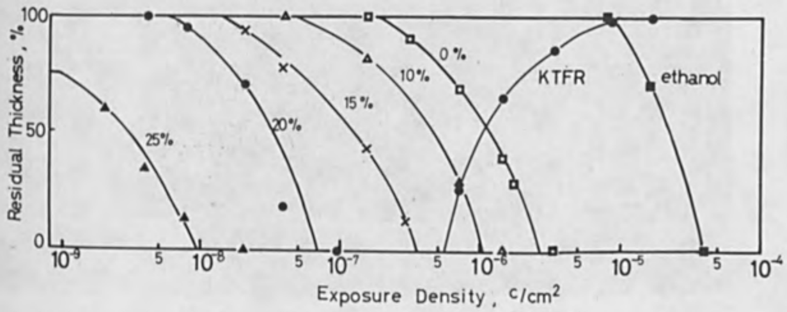


Fig.5 The effect of developers on the sensitivity of the positive resist AT 20°C FOR 20MIN $M=3.8\times10^{6}$

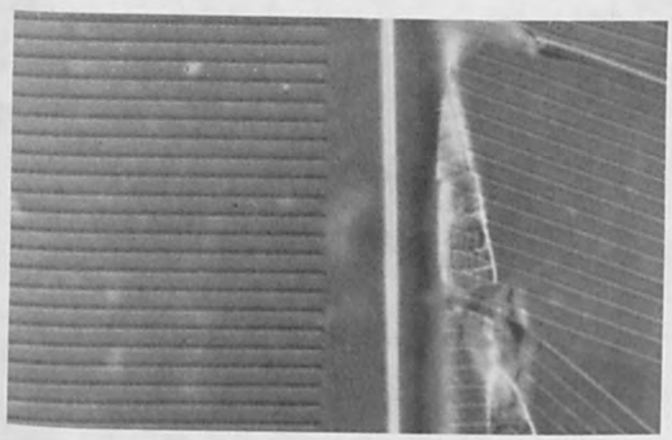


Fig.6 An optical microscopic photograph showing the high-resolution.

The left half shows a SiO_2 etched pattern which consists of about 0.2 μm wide lines separated by 2 μm. The right half shows the resist film with the developed pattern after separation from SiO_2. 1×10^{-6} C/cm² at 20kV

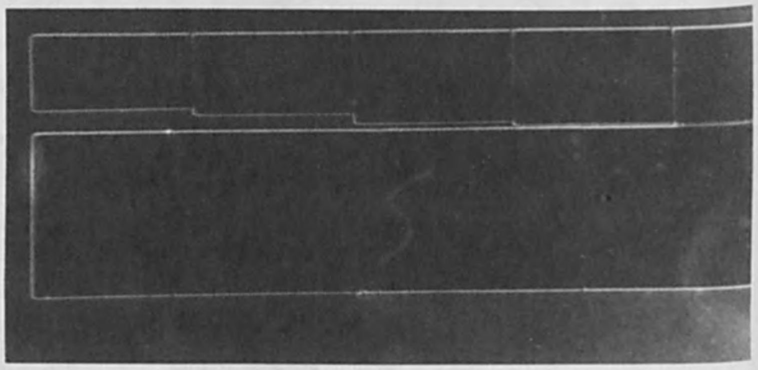


Fig.7 An optical microscopic photograph showing the high sensitivity 1×10^{-7} C/cm^2 at 20kV

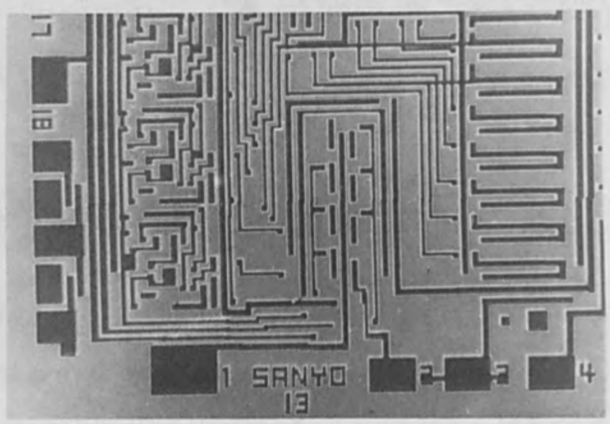


Fig.8 An optical microscopic photograph of SiO_2 etched pattern obtained by using Sanyo Positive Resist

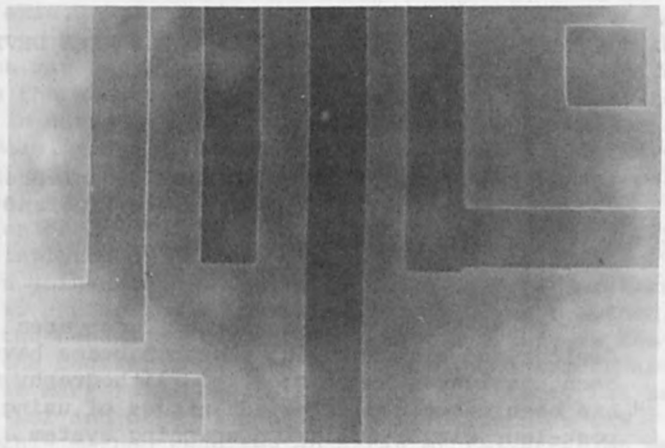


Fig.9 An enlarged photograph of Fig.8
SiO_2 etched pattern

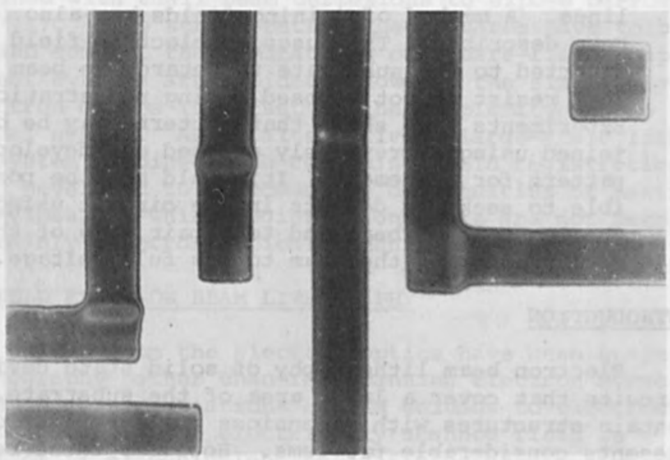


Fig.10 A photogrsph of the KTFR developed pattern exposed at the same exposure conditions

ELECTRON BEAM LITHOGRAPHY OF LARGE AREA DEVICES WITHOUT REGISTRATION MARKS

H. Ahmed and G.A.C. Jones

Engineering Department, Cambridge University,
Trumpington Street, Cambridge, England.

ABSTRACT

The problems of fabricating large area devices and circuits with electron beams have been considered. A large-field lithography system has been described. The advantages of using a post-lens single-deflexion scanning system rather than a pre-lens double-deflexion arrangement have been shown with experimental results. Critical areas of a field with an area of 1 cm x 1 cm have been covered with a test pattern which shows a resolution of 10,000 lines. A method of joining fields has also been described. This uses an electric field directed to the substrate to retard the beam so that resist is not exposed during registration. Experiments have shown that patterns may be butt joined using a previously exposed and developed pattern for alinement. It should also be possible to seek for defects in the circuit using the low voltage beam and to repair some of these after restoring the beam to its full voltage.

INTRODUCTION

Electron beam lithography of solid state devices and circuits that cover a large area of the substrate, and also contain structures with dimensions that are 1 µm or less, presents considerable problems. However, these requirements have to be met in applications such as the manufacture of SAW devices, CCDs, large memories and displays.

The basic problem may be summarised as follows. In an electron beam lithography system the diameter of the focused beam increases for large angles of deflexion from the optical axis. By restricting the deflexion angle, it is possible to expose a small area or field within which the beam size may be kept to the required specification. To increase the area the substrate is moved mechanically to centre, in turn, a number of adjacent fields on the axis of the beam. Precise alinement of the fields is achieved by locating a registration mark on the substrate and using the signal derived from this mark to add small currents to the deflexion coils. These shift the beam axis slightly to the misplaced mechanical centre. Fields of 2 mm x 2 mm have been joined with this technique to cover a large surface area. Reference 1 gives an account of this method. However the technique is wasteful since not only is the area of the registration mark lost but also a search area around the mark is exposed. A further difficulty can be that for LSI and other large area devices such as SAW devices this method requires an array of registration marks within the active device area which places a constraint upon the layout permitted to the circuit designer.

In an alternative method a continuously scanned table is combined with small beam deflexions to expose narrow strips (2). Here strips rather than squares have to be alined and some of the limitations of square-field step and repeat systems are avoided. However the simplicity of the electron optics is offset by the complexity of the mechanical system. As an alternative a large-field lithography system combined with retarding field registration is being developed to overcome the difficulties of making large devices, in which registration marks are not permitted, by mainly electron optical techniques.

LARGE-FIELD ELECTRON BEAM LITHOGRAPHY

In this system the electron optics have been optimised for lithography rather than for scanning electron microscopy as in most adaptations of SEM columns to electron beam lithography. The electrically scanned field is 1 cm x 1 cm i.e. equivalent to 25 fields of 4 mm^2 which would each have required separate registration marks or equivalent to a 100 µm wide strip of 100 cm length. The basic resolution requirement for the system is to produce 10,000 lines or 10^8 points in the field. A certain amount of flexibility is available in that an area of 7 mm x 7 mm may be covered with 0.7 μm lines and it is possible to slightly increase the field beyond 1 cm x 1 cm with some loss in resolution.

The system uses a post-lens single deflexion system rather than the pre-final lens arrangement of double deflexion found in SEMs. The arrangements are compared in Fig. 1. The coil design and construction minimises the deflexion abberations (3). The size of the on-axis focused beam to its 10 % amplitude points is 0.2 μm and increases to just under 1 μm at its extreme deflected positions. Dynamic correction of field curvature is relatively straightforward in such a post-lens deflexion system. We have found that self-compensation as a result of field curvature results in a beam diameter that may be kept constant to within 10 % over the whole field at an angular aperture of 1 x 10^{-3} radian, without dynamic correction.

Initial experiments showed that the beam size varied considerably over the field at different points symmetrically placed around the optic axis although the average size at diametrically opposed spots was the same. This effect depends critically on accurate centering of the coils. An alinement procedure using adjustments of the mechanical coil position and, independently, the current through any of 4 deflexion coils was adopted. The coils were run in opposition until the largest symmetrical image of a test grid was obtained on the screen corresponding to a minimum deflection on the specimen. This is a very sensitive alinement procedure and a field cancelling factor of approximately 10^4 can be obtained and hence an accurate electron optical axis established for the coils.

PERFORMANCE TESTS

Performance tests were designed to establish both the electron optical resolution of the system as well as the resolution obtainable with lines fabricated in electron resist. A grid with 5 μm bars was placed in turn at the centre of the field andthe extreme positions of the field perimeter. At each position the beam size was measured by the usual method where the signal from a scan across a bar is displayed on an oscilloscope. The distance in which the current falls from 90 % to 10 % of its final value defines the beam diameter. The results showed that 10,000 lines per field and a uniform spot over the field to within 10 % could be obtained.

The beam was then used to expose a test pattern as shown in Fig. 2(1) in a silicon substrate coated with PMMA. The positions in which the pattern was exposed are shown in Fig. 2(2). The resolved lines were measured and the results again showed that 10,000 lines/field could be obtained in the resist. Further tests have shown that for long narrow devices, typically 1 mm wide, 16,500 lines can be resolved in the resist. The whole field could not be covered because of electronic limitations and no accurate distortion measurements are available at present. However some initial measurements have shown that at the edge of a 7 mm square the maximum bowing of a full side is not more than 2 µm. For areas larger than 1 cm x 1 cm a retarding field registration method has been devised.

RETARDING FIELD REGISTRATION METHOD

The principle of the retarding field registration method is described with reference to Fig. 3. The electron beam is formed at the normal gun voltages necessary for electron beam lithography i.e. 10 - 15 kV. The beam is slowed down for registration purposes in the region between the deflexion coils and the specimen plane. The beam landing energy is determined by the difference in potential between the gun cathode and the specimen. The secondary electrons are collected by a silicon solar cell with a hole drilled through its centre and mounted on the deflexion coils facing the substrate. Since the cell is at a potential near earth while the substrate is at say, -12 kV the low energy electrons in the secondary emission are collected by the cell. The greater portion of the high energy secondary electrons in the total signal actually disappears up the hole in the coils and final lens.

Such a system has advantages over a low voltage system of conventional design. First of all the high brightness and simpler gun design of the conventional high-brightness, high voltage gun is available. The electron optical limitations of the low voltage systems are also avoided in lens design. Interference from stray magnetic fields which are very significant in low voltage electron-optical columns are not very serious except in the very last section of the column where adequate shielding can be provided. Experimentally and theoretically we have shown that the resolution of the system only varies by a factor of two, ranging from about 0.25 μm to 0.5 μm, for beam landing energies from 15 keV to 1 keV. Fig. 4(a) shows a grid (bar width = 5 μm) photographed at 15 keV landing energy and Fig. 4(b) shows the same part of the grid at a beam landing energy of 1.3 keV.

RESIST EXPOSURE AT LOW BEAM VOLTAGES

Hatzakis et al (5) have considered the effects of changing the beam voltage and charge density on PMMA resist and Heidenreich et al (6) have considered the effects similarly for a number of positive and negative resists. Their experiments were aimed at determining optimum resolution and exposure conditions. At beam voltages as low as 1-3 kV, however there is negligible exposure of resists such as PMMA for normal current densities used in our examination technique. Thus a resist coated substrate may be examined without exposure.

For registration the retarding field method may be used in any of the three arrangements shown in Fig. 5. Either conventional registration can be used with marks between 1 cm x 1 cm fields as in Fig. 5(a) or registration marks outside a 2 cm x 2 cm field and one central mark previously exposed and developed can be used as in Fig. 5(b). Finally it is also possible to butt join patterns as in Fig. 5(c). Here it is necessary to use a buffered development procedure to maintain line width. A line is first exposed with the beam energy at 15keV and a charge density less than that required for development to full line width. The developed substrate is then returned to the machine and the line is located at retarding field conditions. A second line is then exposed joining the first and the whole substrate developed to the correct amount. Results from preliminary experiments have been reported (7).

Problems are encountered with the shift in the beam position when the beam is changed from the retarding field voltage to fabricating voltage. However when the experiment is performed at two precise working voltages the shift can be accurately calibrated and compensated by deflexion current components. The accuracy with which this can be done is not yet known. A magnification change of a few per cent takes place between the two voltages. This does not seriously affect the alinement procedure. There is also a change in the astigmatism conditions and correction is needed to maintain accuracy of alinement. There is also a small change in the final lens focus which also has to be calibrated and corrected.

Further tests have been carried out on the retarding field registration method. An alinement cross of 3.5 µm width as shown in Fig. 6(a) was exposed at 15 kV in PMMA and developed in MIBK. It was then located and photographed in the fabrication system at a beam landing energy of 1.3 keV. Note that the area around the mask was not exposed during the search procedure. A pattern consisting of a set of lines was exposed with reference to this alinement mark and developed. The specimen was then returned to the system and a second pattern of lines was exposed again with reference to the alinement mark and adjacent to the first pattern. The whole slice was redeveloped. The area of the lines is shown in Fig. 6(b) which was taken with an optical microscope.

From these results it has been possible to show that even for very long exposures (several minutes) of an alinement mark the PMMA remains virtually unmarked. The misalinement of the patterns is caused by the asymmetry of the electric field in the specimen region. It is possible to correct this asymmetry either by changing the field shaping electrodes in the chamber or by adjustment currents in the scan coils. The experiments also show that it is possible to adjust the exposure and development procedure for PMMA so that a double development of the pattern does not cause a serious mismatch of the line widths in the two alined patterns. Further experiments to test whether or not this system will prove to be a satisfactory practical registration method are being carried out.

ACKNOWLEDGEMENT

The support of the Wolfson Foundations, STL (Harlow) and SRC is gratefully acknowledged.

REFERENCES

1. J.P. Beasley and D.G. Squire, "An Electron Beam Mask-maker" IEEE Trans. on Elec. Dev., Vol. ED-22, 7, 376-384, (1975).

2. D.R. Herriot, R.J. Collier, D.S. Alles, J.W. Stafford, IEEE Trans. on Elec. Dev., Vol. ED-22, 7, 385-392 (1975).

3. G. Owen and W.C. Nixon, "Aberration Correction for Increased Lines/Field in Scanning Electron Beam Technology", J. Vac. Sci. and Technol., 10, 6, 983 - 986, (1973).

4. G.A.C. Jones, H. Ahmed and W.C. Nixon, "Electron Optical System for Microfabrication with 10,000 Lines/Field" Electronics Letters, 11, 10, 214-216 (1975).

5. M. Hatzakis, C.H. Ting, N. Vishwanathan, "Fundamental Aspects of Electron Beam Exposure of Polymeric Resist Systems", Proc. 6th Int. Conf. Electron and Ion Beam Sci. and Technol. San Fransisco Ed. R. Bakish, 542-579 (1974).

6. R.D. Heidenreich, L.F. Thompson, E.D. Feit, C.M. Melliar-Smith, "Fundamental Aspects of Electron Beam Lithography 1. Depth-dose Response of Polymeric Electron Beam Resists", J. Appl. Phys. 44, 9, 4039-4051, (1973).

7. H. Ahmed and R.D. Hoare, "A New Method of Registration for Electron Beam Lithography", Electronics Letters, 12, 1, (1976).

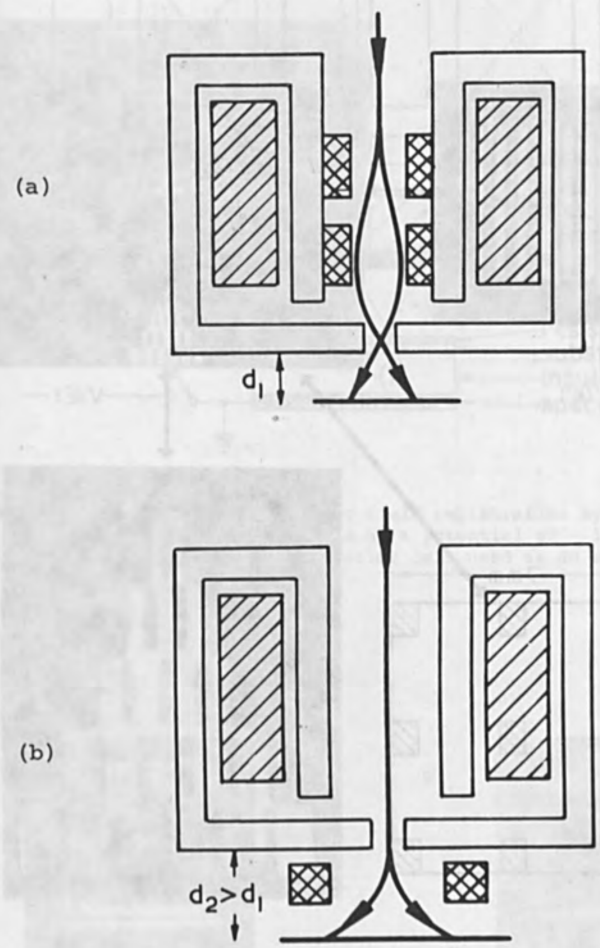


Fig. 1 (a) Double deflexion pre-lens scanning coils suitable for high resolution scanning electron microscopy.
(b) Single deflexion post-lens scanning coils suitable for electron beam lithography.

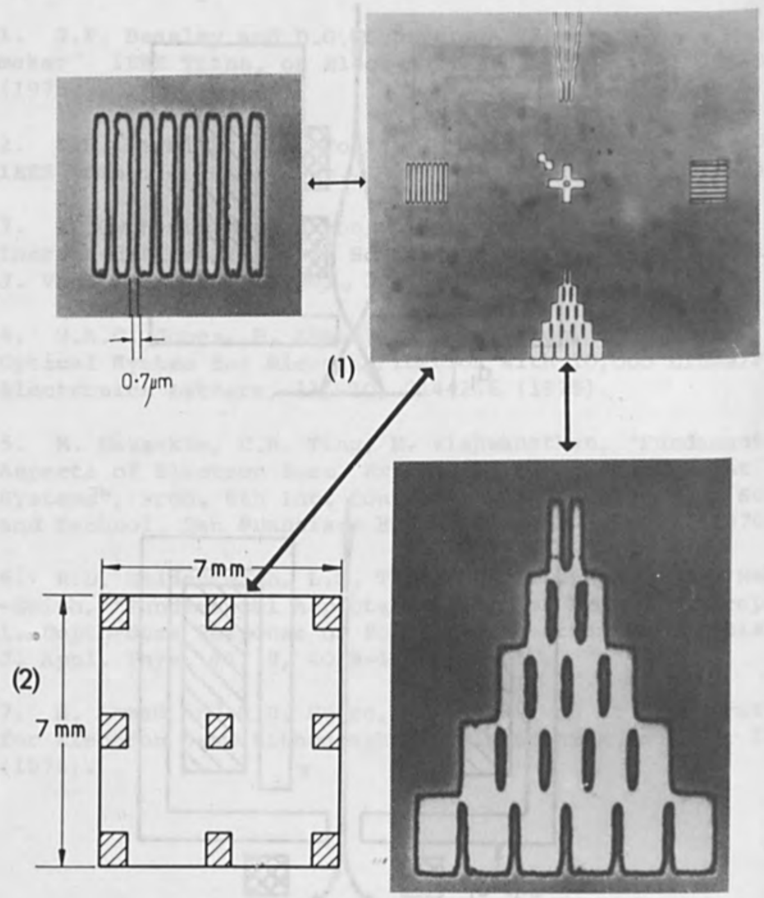


Fig.2 (1) Test pattern used to check the resolution of the large-field lithography system.
(2) The positions at which the test pattern was exposed to check the resolution capabilities within the 7 mm x 7 mm field.

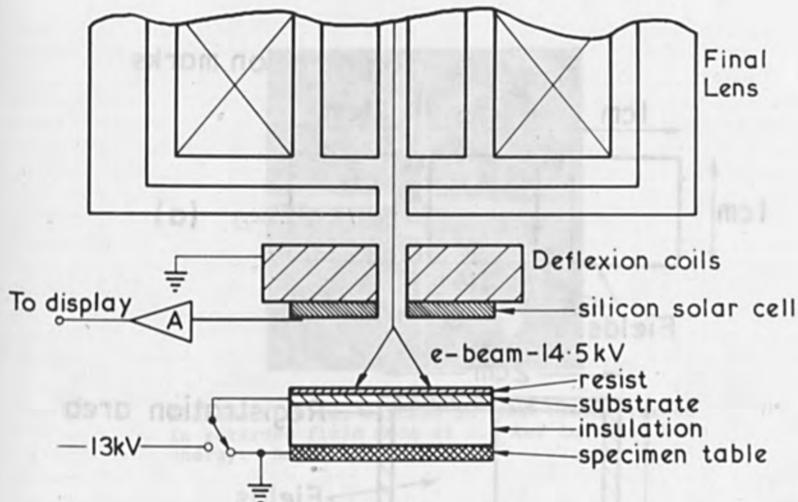


Fig. 3 The principle of the retarding field registration system for e-beam lithography. The substrate is at a potential of - 13 kV. Secondary electrons are collected by a solar cell used as an electron detector.

(a) ALIGNMENT GRID WITH 5 μm BARS PHOTOGRAPHED AT 15 keV BEAM LANDING ENERGY

(b) SAME GRID AREA AS ABOVE PHOTOGRAPHED AT 1.3 keV BEAM LANDING ENERGY

FIGURE 4

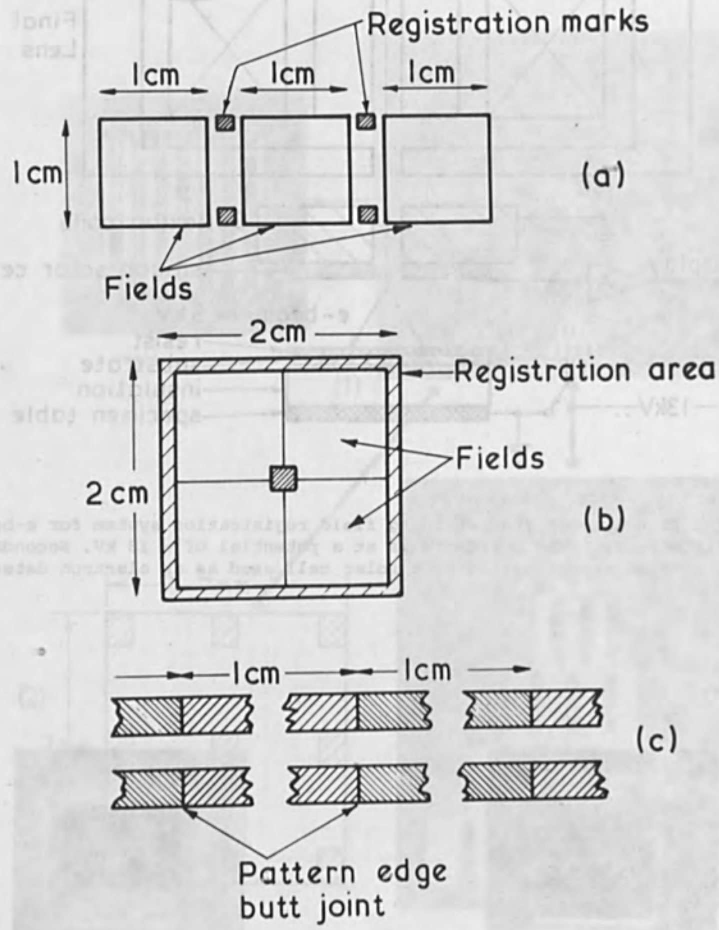
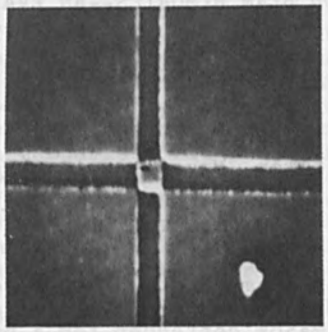
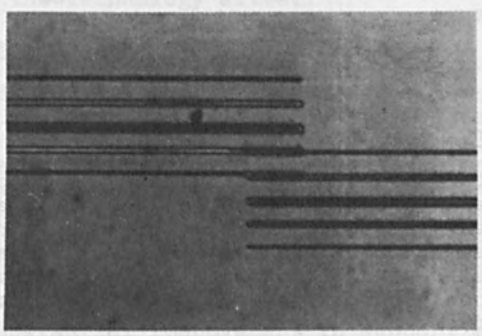


Fig. 5 Registration methods that may be used with the retarding field system.
(a) Conventional step and repeat
(b) External registration marks
(c) Butt joints on large patterns



Alinement marker exposed in PMMA resist viewed in retarded field mode at 1.3 keV incident beam energy. Bar width is 3.5 μm.



Two sets of bars a, b exposed in PMMA resist. Set a was exposed and developed first. Set b was then alined under retarding field and exposed. The whole slice was then developed again. Bar widths 0.7, 1.4, 2.1, 1.4, 0.7 μm.

FIGURE 6

ELECTRON BEAM MEMORIES--AN EVALUATION

S. P. Newberry

Micro-Bit Corporation
40 Hartwell Avenue
Lexington, Massachusetts 02173

ABSTRACT

Electron beam memories are entering the marketplace. Memory performance indicates this technology is a major contender for a position in the memory hierarchy of intermediate to large computer systems. Using high resolution electrostatic electron optics, precision electronics and semiconducting storage targets these memory systems offer megabyte capacities with costs and access times in the infamous "access gap" between expensive high speed memories and the slow but inexpensive rotating memories. We attempt to evaluate the status of this technology and project its future.

INTRODUCTION

Early History

The return of electron beam memories to computer architecture has been anticipated for some twenty years. This past year independent groups have delivered prototype electron beam memories for test and evaluation. Subject to the pitfalls of all high technology forecasts, the Micro-Bit Model 700 EBAM system, a full product version, is scheduled for delivery and test this year. It will be a plug compatible replacement for a fixed head per track disk. It will provide a much shorter latency time, i.e. delay between demand and arrivals of the first bits. A delay three orders of magnitude faster than the head per track competition.

This is a return because the first modern style computers (that is digital computers with stored programs and parallel access) employed electron beam memories.(1) The most widely used and best known was the Williams tube (2). The use by Williams and Kilburn of a slightly modified but standard cathode ray tube was doubtless a defendable choice from an economic view. It was also commendable for its ingenuity. It was, however, an unfortunate choice for the technology because unreliable performance added to the disenchantment with electrostatic storage. Even tubes specially designed for storage (but still based on the then popular surface storage of charge) did not hold up in comparison with the new ferrite core technology which displaced it. Cores had greater reliability, lower cost, faster access and nonvolatility (1). It was again unfortunate that when the cores appeared, in the early fifties, computers were primarily thought of as calculating machines. Thus the ferrite core came into its deserved position at that time as the technology for central memory (not to be successfully challenged until the advent of semiconductor memories) but also for want of understanding of requirements for noncomputational machines the electron beam store was unwisely abandoned. Some of the ongoing programs did continue an example being work based on the Barrier grid tube of Jensen et al (3) appearing as late as 1958 (4, 5) but hampered by the poor showing of their primative predecessors and not really making a sufficient break from standard cathode ray tube technology to properly exploit the power of the electron beam, they did not keep the technology alive.

Return of Electron Beam Memories

As noted above electron beam accessed memories (EBAM) have returned. The moving force for their development has been the reality of the access gap between main frame and peripheral memory (6). This "access gap" creates inordinate systems software complications and extracts a heavy toll on valuable CPU time by increasing its housekeeping chores to a sizable fraction of its activity. The dismal failure of all other technologies to bridge this gap placed the needed support behind electron beam memories. They are vastly improved over their predecessors because of substantial improvement in support technology and improved understanding of crucial design considerations required to give reliability and cost effectiveness. A graphical representation of the access gap is repeated here in Figure 1. These data are from a currently appearing paper by Speliotis et al (7). (The basic method of presenting the access gap is due to Pugh (8).) In Ref. (7) the projected performance of Micro-Bit Model 700 EBAM and projections for the next decade of both EBAM and its competition have been added. In this form of presentation the cost per bit and the access time are highlighted as the crucial parameters in the choice of a technology for bridging the access gap it being reasonably assumed that each of the technologies can be made to perform adequately in thruput rate by resorting to parallel access and can be adequate in total capacity for this application without incurring significant degradation of performance. These requirements must not be taken lightly however. As may be seen the access speed to be accommodated across the gap is four or five orders of magnitude. To cover the needs of central memory, the bridging memory must then, as a minimum, store a quantity in excess of the size of a "normal central memory access request" multiplied by this factor. It is this substantial memory size which makes the cost per bit a crucial factor and places extension of present day central memory out of consideration for bridging the gap. One should further note that if the above projections hold up, then EBAM should take over the job of movable head disk as well in the next decade thus eliminating that side of the bridge and the attendent software complications.

THE MOS MEMORY TARGET

Let us turn attention to the basic technology which has made the new generation of electron beam memories possible; that is the development of the unstructured, large area, MOS capacitor electron beam memory target (9). This is the memory medium in the only EBAM devices which have reached product stage, namely the Micro-Bit and the General Electric tubes.

The MOS memory medium is represented schematically in Figure 2. It is shown here as an n on p junction device which has become standard practice although a simpler structure of oxide on a homogenous silicon wafer also makes a good memory target. A bias control and signal detection circuit is also indicated. One should observe that the biases are switchable and are connected to the substrate permitting the metal which is exposed to the electron beam to be placed at the electron beam drift potential. Thus the beam is not exposed to changes in electric field when bias is changed.

The layer thicknesses are in the neighborhood of metal 500Au, oxide 2000Au, n layer 1 micrometer. The thickness of the silicon substrate is not critical. In general it is a high resistivity material. Considerations in the choice of thickness of the other layers are resolution loss by beam scattering, breakdown strength of the oxide, and low spreading resistance in the silicon.

Basic Storage Mechanism

The basic storage mechanism depends on the well documented observation that when ionizing radiation such as the electron beam passes through the oxide as shown in the presence of a positive bias (negative bias on semiconductor, positive on metal) a positive charge will accumulate in the oxide in the vicinity of the oxide silicon interface (10, 11). It is further known that the charge will remain there for weeks unless bombarded by the beam while a negative bias is simultaneously applied. With a substantial negative bias the charge is quickly dissipated providing for an efficient erase process. With a bias near zero volts the state of charge storage can be detected for read-out without greatly altering the "charged" and "not charged" areas providing for multiple reads (up to 100) before refresh is required. The read-out process, that is the discrimination between a charged and an uncharged area, takes advantage of the modification within the silicon of the carrier distribution and fields immediately under an oxide area containing positive charge. Under these areas, beam produced holes formed in the n region by dissipation of beam energy in the silicon are detected as current flow across the junction. By contrast holes formed in the silicon where there is no positive oxide charge are subject to recombination at the oxide-silicon interface and therefore less junction current flows. For more detailed discussion of the memory target physics the reader is referred to References (17), (9), (12), (13), (14) and (15). The purpose here is to evaluate MOS silicon operationally as a memory material.

Advantages and Disadvantages of MOS Target

Silicon is an almost ideal material for an electron beam memory. Here the term "ideal material" is used as the author defined it before this conference in 1962 (16) with additions and refinements developed through the years. An ideal electron beam storage material would have all of the following characteristics:

* Rigid and mechanically sturdy material or mounted on a mechanically sturdy support.
* Vacuum compatible with electron tube technology including high temperature bakeout.
* Present no disturbing electrostatic or magnetic fields to the beam path.

* No built-in bit location structure, i.e. either no structure or a structure much smaller than beam diameter.

* Few natural or structured defects which produce inoperative addresses.

* A low temperature coefficient of expansion so that data may be relocated despite temperature changes.

* Controllable properties which are consistent batch to batch.

* Capable of all memory processes such as recording-updating-reading etc. with submicrosecond response time and without changing vacuum environment or moving from its position.

* Not require changing beam energy or current.

* Capable of changing memory functions by a simple switching operation such as changing an electrical bias by a few volts or illuminating with radiation of different wave lengths.

* Have a built-in amplification mechanism so that the beam current need only be statistically adequate for binary (or at most ternary) discrimination and so that power for storage, erasure, and readout are derived from a power supply rather than the beam.

* Nonvolatile, with recording and erasure threshold both for beam energy and power density.

* Nondestructive readout with echo readout during recording and erasure.

* Does not fatigue or disintegrate with time or usage.

We didn't really expect to find a material with all of these characteristics but did hope to find most of them in one material. It was well known at General Electric that when the search for this material would become too discouraging Dave Skelley would amuse his fellow chemists by reciting the above list after the manner of reciting "Horatio at the Bridge."

What is not generally known is that even as late as the formation of Micro-Bit Corporation in 1969, a suitable target material had not been proven. In fact, a choice had not even been made of which material to pursue.

Measuring silicon against the above list, we find not only is it a clean vacuum compatible material with a mature fabrication technology but it also permits one to switch from store to read mode by a simple low voltage (<50 volts) bias change on the target without separate writing and reading structures. It is not even necessary to change beam energy or current. The MOS structure has further desirable characteristics. By the process of hole-electron pair production it acts as an ideal impedence match between the low current, high energy, high impedence electron beam and high current low impedence solid state circuitry. This internal transformation also is a high gain amplifier in the reading mode and to a lesser extent in the writing mode because of electron beam induced conductivity (EBIC). This is fortunate because for readout the beam current need only be large enough to statistically determine between the two memory states--the power to read the bit and partially to store it comes from the bias power supply. The use of charge storage as the basic mechanism has the advantage of orders of magnitude in sensitivity over magnetic or thermal mechanisms. Location of the charge deep within the material has the advantage over surface charge that it is not disturbed by low energy electrons or surface leakage and the beam trajectory is effectively screened from data related fields by the conducting metal film. This immunity is not generally provided in surface charge methods. Traditionally charge storage tubes have produced the charge pattern on the front surface of the target by secondary emission wherein the beam landing energy and surface potential are varied. The beam is required to be normally directed to the surface to avoid latteral sweep with voltage and it is subject to pattern dependent lateral forces from the surface charge. When recording sequentially from one side progressively toward the opposite side on surface charge devices it is not uncommon with a 10 kilovolt beam to see a shift of over 25 micrometers in beam location due to this unsymmetrical surface charge. Location of the stored charge out of view of the free electron path is a key advantage of the MOS target.

Echo readout does occur during writing of positive charge and is possible even during erasure of positive charge.

The requirement of high energy electrons to change the data releases the modern electron beam memory from the far field loss of data characteristic of the early storage tubes of the secondary emission variety; that is it does have an energy threshold. It does not, however, have a power threshold and one therefore operates the present memories in some kind of restore or periodic refresh mode because of gradual nearest neighbor disturb from beam tails and small angle scattered high energy primaries. There has as yet been no mechanism for a holding bias or holding beam suggested. One is faced then with periodic refresh, the alternate choice being loss of packing density by wider data spacing.

In the event of power failure, data is not lost provided a proper priority system for power loss and return is designed into the equipment. With no beam on, the stored charge pattern will remain for a month at least. The silicon target has a low temperature coefficient, reasonable well matched to other materials used in tube construction. It also has good thermal conductivity.

Summary of MOS Advantages

The strong points of the MOS silicon target for EBAM are:

* Is a well developed technology--well controlled and essentially defect free for EB memory use.

* Is vacuum compatible.

* Need have no built-in address structure.

* Has a low temperature coefficient.

* Capable of storage erasure and reading by simple bias change without changing beam current, beam energy, or changing vacuum environment.

* Has a built in amplification system which also acts as an impedence transformer from electron beam to solid state circuitry and provides good sensitivity.

Disadvantages of MOS

Silicon has no insurmountable problems as evidenced by the emergence of viable memory products. It does have some shortcomings and some unique characteristics which must be reckoned with in any practical memory system. The most striking characteristic is the appearance of large transients when the bias is switched to change memory functions. The easy way to avoid recording or reading under severe transient conditions is to wait a few microseconds for the transients to die down. Obviously one could not do this between bits and is therefore forced to organize data manipulations in small blocks and to employ parallel access to provide large thruput rates. One must bear in mind that similar transient delays are also inherent in settling of large voltage changes needed for deflection to random addresses. Since the two transient times are of approximately the same duration there is no immediate urge to search for material with shorter switching transients. Compared with rotating memories this delay is approximately one thousand times less, i.e. a few microseconds compared with a few milliseconds for disks and drums. Memory blocks as small as 64 bits dilute the effect of this delay to the point that it is no longer perceptible in the thruput rate and a very small buffer hides it from main memory.

Another characteristic quality, which may disappear as the technology matures, is that the electronic characteristics of silicon are a slowly varying function of integrated irradiation history. At one time it was necessary to give the material an extended radiation dosage to "activate" it to full signal level. Activation is no longer required due to intentional incorporation of interface states to support the recombination mechanism essential to read. It is likely that change of materials characteristics with usage will also disappear as the technology matures. It has been established that the silicon does not change, the problems are primarily in the interface region and the oxide. As a practical matter, despite the slowly changing characteristics of the target the useful life in a large block oriented memory is a matter of years by which time other factors may dictate replacement of a tube on a preventative maintenance basis. A problem does arise if one area of the target is used disproportionately and thus becomes incompatible with the rest of the target. This problem can be avoided by address location randomization procedures which can be built into the controller to make target usage more even. Employing such management techniques, a target life of many years is achievable.

Beyond these two considerations, MOS silicon doesn't have any unique problems or difficult operational requirements. For very large data bases where most of the store will change slowly or not at all we shall see in a latter section that silicon also is still a very attractive material but not in the MOS version.

OTHER TARGET APPROACHES

Not all current effort for electron beam memory is based on the structureless silicon target. The first proposal, made by Everhardt and MacDonald (11), to use charge storage in the oxide was a structured target. It was a well thought out concept and could be used as the basis for an attractive memory but becomes difficult or impossible as packing density is increased. At lower packing density target manufacturing cost is a more severe problem with it than the unstructured target but radiation induced changes should be less severe. Other contemporary approaches exemplified by the mu-cap approach of Stanford Research Institute (17) rely on storage of surface charge but with approaches which hopefully will avoid the severe reliability problems of the early tubes. A working demonstration model showing some promise was delivered on contract by SRI last year.

There is a large body of significant electron beam memory and electron beam recording work which cannot be covered in the scope of this paper. Unfortunately not all is in the open literature. Especially noteworthy is: the work with micro-images by Möllenstedt and his students at Tübingen; very early demonstration of electron beam recording advantages by Ardenne; ultrahigh resolution recording by Nixon and his students at Cambridge; and pioneering work by many commercial laboratories including Ampex, Bell Labs, CBS Labs, GE, IBM, Mullard, RCA, Siemans and Westinghouse. The earliest account of recording information by electron beam known to the author is work by Selenyi, starting in 1928 (18, 19).

ELECTRON BEAM ACCESSING SYSTEM

We next turn attention to what kind of electron beam system has been used with the unstructured MOS capacitor type silicon target. Electron beam systems suitable for this application are by now fairly well covered by publication. The requirements are very similar to those for microfabrication discussed in the preceding paper. An earlier paper by the preceding author is a landmark paper (20) and along with another landmark paper of the preceding generation by Mulvey (21) and the Fly's Eye work initiated by the author (22, 23) combine to give an overview of the market basket of electron optical techniques available to the memory tube designer. I would like to emphasize certain key considerations however. Of primary importance are pattern stability, spot current density, and long cathode life.

Pattern Stability

Pattern stability depends jointly on tube performance and circuit performance with the tube performance being easier to obtain since it is essentially a static requirement whereas deflection wave forms are dynamic. The primary appeal of the Fly's Eye approach is that it drastically reduces accuracy requirements on the deflection drive circuits.

To insure tube stability one is well advised to employ electrostatic optics. We have been able to demonstrate repeatedly that in a clean well shielded system the beam shape and beam position are predictable. With magnetic systems one is at the mercy of subsurface materials variation both spatially and temporially. Magnetic systems have hysteresis which will show up in the magnetic shield if not in the rest of the structure and cause pattern dependent and history dependent displacements. Since one must place the memory tube in a shielded environment, hysteresis effectively rules out magnetic deflection.

With assembly line produced electrostatic systems one may find pattern displacement occassionally. This can generally be attributed to thermal drifts or relaxation of built in stress. These are slowly varying effects and may be corrected out by occasionally applying a feedback correction to small steering plates. It is the small size and slow change in these errors which makes feedback correction practical. Real time feedback uses too much time to be permitted. To generate a feedback signal capable of directing location to a fraction of a spot diameter one must move the beam very slowly across a fidicular mark (or many times across at a faster rate). Since it is performed infrequently and requires only a few milliseconds it makes no appreciable impact on system thruput.

Spot Current Density

Spot current density is of primary concern because of need for recording and readout speed. The requirements are identical with those of pattern generation covered in the preceding paper. It is generally the deflector aberration which is limiting. Electrostatic deflectors have historically had much worse aberration than magnetic deflectors. One of the primary concerns in the memory development has therefore been the improvement of electrostatic deflectors (24). These have been built on the 8-fold deflector due to Kelly (25) and the deflectron originated by Schlessinger (26). As a reference point, the Model 700, an all-electrostatic system, has a current density of nearly 2 amps/cm^2 with a deflection of 4000 by 4000 resolved spots. This effort has required development of hitherto unavailable three-dimensional electron optical computational methods, the work of Wang, Harte, et al (27) and of Harte (28) being notable examples.

The gun brightness is equally important as the system aperture. Here again, the requirements parallel those for pattern generation. With electrostatic deflection one should note that more care must be exercised to reduce the "Boersch effect" (20) because of greater chromatic sensitivity than magnetic deflection.

Cathode Life

In opposition to current density is the requirement for long cathode life. The first-generation memory tubes make use of commercially available miniature dispenser cathodes. These commercial units can supply the cathode loading of 3 amps/cm^2 required to give the 2 Amperes/cm^2 in the spot with a heater power less than 10 watts. At this cathode heater power and loading, the cathode life should exceed four years.

Summary Electron Optical System

We may summarize the electron optical evaluation of current memories by noting there are no surprises in this part of the memory but there has been a general upgrading of the accuracy of assembly and lowering of the cost of components. Realization of the 8-fold electrostatic deflector is a direct product of this effort and availability of miniature dispenser cathodes, a key item in the success of the development, was initiated by this effort. Both current memory development teams have made improvements in design and construction of the coarse deflection structure and of the Fly's Eye Lens. General Electric has been the first to produce a sealed-off version Fly's Eye Tube while Micro-Bit, by using single-channel tubes for their first sealed-off version, thereby obtained an earlier start on plug compatible memory systems. The Micro-Bit Model 450 System was tested as a peripheral on a CDC Star Computer more than a year ago. The work of the two groups, as a total, comprise a firm demonstration of the power of the approach.

DESCRIPTION OF CURRENT APPROACHES

In this section we attempt to give a balanced, concise view of the equipments and approaches selected by the two product ready groups. For more complete description of each system the reader is referred to recent publications by the two teams (7), (15).

The heart of the system is the memory tube. Micro-Bit has chosen for its first models (Model 450 and Model 700) to provide single channel tubes i.e. one electron beam with one target in a tube. A number of these tubes are then operated in parallel to make a memory system. The advantages of this approach are:

* Easier to develop and produce--giving a head start on experience with operating systems.

* More flexibility in system size and configuration.

* Smaller systems possible.

* Yield problems less severe.

* Replacement of part of memory in event of a failure.

The alternate tube configuration available is the multi-channel approach employing a Fly's Eye Lens and a single beam directed to a required lenslet by a coarse deflection system. General Electric has chosen this approach for their first equipment and Micro-Bit has chosen it for their larger systems. The advantages of the multi-channel approach are:

* Less demanding on deflection precision.

* Capable of much larger capacity per tube.

* Less sensitive to vibration because of shorter beam sensitive region.

The Micro-Bit single channel tube is shown in Figure 3. both schematically and pictorially. The General Electric BEAMOS tube is shown pictorially in Figure 4. The Fly's Eye Lens to the left of Figure 4 is pictured separately in Figure 5 with the target removed to show the fine deflection bars. Figure 6 shows an earlier Micro-Bit Fly's Eye with the fine deflector removed to reveal the lens array. A schematic of a Fly's Eye system, furnished by Micro-Bit, is shown in Figure 7.

In Figure 8 one may see typical data outputs from the two systems at approximately 6 micrometers spacing. Both obviously have acceptable signal to noise and freedom from imprint of previous information. Multiple read is demonstrated in the lower set of traces by General Electric at 5.6 microns spacing.

Figure 9 shows the schematic diagram for the 128 million bit Model 950 tube currently under development at Micro-Bit. It is understood that similar advanced models are under consideration at General Electric.

One interesting point of divergence between the two groups is in the choice of unipotential versus accelerating lens systems. The unipotential approach maintained by Micro-Bit permits the high speed electronics to be at ground potential. The accelerating system permits lowering lens aberration by a large factor, as explained by General Electric in a subsequent paper in this Conference, but does require either the sense amplifier and fine deflection or the coarse deflection to be at a high potential with respect to ground.

Currently the published results of the two systems are comparable. Comfortable packing density is 6 micron bit spacing with approximately 6 million bits per square centimeter and bit readout at 5 million bits per second on each channel.

FUTURE EXTENSIONS OF THE ELECTRON BEAM TECHNOLOGY-- A PERSONAL PERSPECTIVE

Near Future Applications and Impact

In this section we are dealing with predictable extensions of the current technology. There are few unknowns in these forecasts. For these near-term projections no new technology or breakthroughs are required for electron beams and none is taken into account for competing technologies.

Two computer areas can expect to feel the impact of electron beams in the near future they are:

1. Main memory extensions.

2. Rotating memories.

The advantages electron beams offer for main memory extension are much reduced cost per bit and very few inter connections compared with cores or semiconductors. The penalty one must pay for these advantages is the access delay of a several microseconds described above. Fortunately this is not a serious impediment. One may use relatively small, very fast, cache memories (8) as a buffer to hide this delay most of the time since there is a high probability that the next called for information will already reside in cache. The use of a cache or other buffer also eases the problems of data rate synchronization between main memory and the electron beam memory. Models for such systems using cache concepts already existing (29) have been examined in considerable detail with favorable prognosis.

The second near-term application invades the domain of rotating memories. This has already begun with the present systems which are designed to be plug compatible with head per track disks and drums but with improved performance because of the three orders of magnitude gain in access time. A next generation of electron beam systems which are planned, and currently under development, would achieve sufficient capacity and cost improvement to compete with moving head devices but would still maintain the performance advantage in access time and the advantages of no moving parts. Replacement of moving head memories by electron beam addressed memories would completely close the access gap by essentially eliminating it. That is the gap would be reduced from four orders of magnitude to approximately one.

Long Term Developments

We shall now deal with the question of what lies beyond events predictable by direct extension of present accomplishments. That is how far can electron beam technology ultimately carry us? What are the areas which most need its unique qualities?

It appears that currently the most difficult area of systems architecture is that of proper management of very large randomly accessed data bases. It is the combination of very large size and random access which creates the problems.

Fletcher et al (30) in a review of the past 23 years of one laboratory's computer experience (Lawrence Livermore Laboratory) describe many of the problems with very large data bases (VLDB). They also note the changing emphasis toward placing larger and larger fractions of memory on line (i.e., machine access). For an appreciation of the degree of effort currently expended on VLDB management, the reader's attention is called to a session on "Data Base Management" at the 1975 National Computer Conference (31). A text book by C.D. Date (32) gives an excellent introduction to present and emerging software and operating system approaches to VLDB management.

The volume and sophistication of current work is both indicative of the extensive problems in this area and the growing competence to handle them. Hopefully we can provide parallel hardware improvements to help lessen the problems.

A good way to portray the problem of the VLDB is shown in Fig. 10 taken from the 1962 meeting (16) of this conference. Here one sees in the homely illustration of file cabinets on the desert the basic problems of very large files. Even if adequate storage capacity is available, one still has the problem of knowing where information is located and how to retrieve it quickly. The problem is further complicated if more than one customer requires the access vehicle (multiport requirement). The cost of these data bases dictate that a large number of subscribers share the cost, thus multiport, concurrent use is inevitable.

Currently, VLDB are stored on strips of tape or on microreductions such as microfiche (33). Their problems are less severe than file cabinets because they have a storage density improvement of approximately 10^4:1. They are an obvious improvement for access time over tape reels. Even so, they do require mechanical selection of a given record and transportation of that record to the reading station. Not only does this require a complex and costly mechanical device but compared with main frame access rates it is 10^6 to 10^7 times slower. This is indeed a colossal memory access gap. It is presently bridged by copying segments of the file on to disk from which a part is in turn copied into main frame memory or extended main frame. If one knows the address of required information such a hierarchial approach can be made to work with a limited number of users. However, the larger and more frequently accessed the data base, the longer and more impossible the queuing delays become. Thus, access delay becomes much worse than the few seconds to snatch a record,turning into minutes or even hours in some cases and causing the expensive CPU to spend as much as 70% of its time held up by housekeeping chores (29).

If one does not know the address of required information, then a search must be performed. Here the strip fetching technique is in serious trouble as we shall see. The most general case, and in the author's opinion, the most important case, is that in which the information is distributed through a large number of addresses whose relationship was not obvious at the time of data recording. Conceptually the most straightforward way to find such information is to scan the entire memory--a technique often employed on smaller files. An example will show that there are data bases already too large to be searched end to end for each interrogation. The Library of Congress is a well known example being of the order of 10^{15} bits. Assuming no time loss for changing tapes or film strips and one thousand channels reading in parallel at 10 megahertz rate each, an elapse time of 10 hours would be required for one reading pass. For such a large data base reading from end to end for each query or set of queries is not practical. One alternative is to try to use machine reading of memory content but let the memory content itself aid in skipping nonrelevant segments. The enormous access gap of 10^6 between strip memory and main frame effectively prevents such a mode since a reiterative loop may be reqired between current and next addresses.

What will likely be the first practical compromise is to preprocess the data base indexing it with some kind of key words or content classifiers as is currently done for smaller files. This is analogous to the card index of the ordinary library but hopefully more extensive and penetrating. What fraction of the data base should it be? Is an average of one descriptor per page a reasonable estimate? This amounts to $\sim 0.2\%$ or 2×10^{12} bits for the Library of Congress. I do not doubt that this size index can and some day will be organized for serial searching by lining up interrogations, excursion boat fashion,for 3 minute runs through the entire index memory--a very frustrating time delay for the user which does not guarentee him the required information but instead a first response in a game of 10 questions to get the answer or give up after a half an hour of frustration.

I submit that a much nicer approach would be to employ a randomly addressed index completely in main memory or extended main memory with multiport capability. What I am basically striving for is a way to get around the time limitations inherent with remote location of memory (16). I should like to employ the distributed logic techniques of Slotnick (34) for example, but in a time frame 1000 times closer to central activity than rotating or other inexpensive but remote (in time frame) memories.

Can a Fly's Eye Lens with a silicon target, managed by the software techniques reviewed by Date (32) provide such a multiport random index memory to enable the intelligent subscriber using the distributed logic techniques of Slotnick (34) to thread his way more directly to the desired information than by joining the serial search excursions? Having such a service, will he profit by it and enjoy using it? I believe he will find his infomration more readily than by serial search methods because information is not totally without organization. It is generated by human activity and tends to maintain identity with its origins so that it is highly structured in its location, being interconnected by time and association with interrelated groups of people. It is this structure which, for example, makes citation indexing appealing for scientific information retrieval. It is this same structure which makes browsing through library stacks so effective. With the machine accessed data base we can hope for machine assisted browsing through its memory without loosing the capability of searching through the entire memory when required. The popularity of the time-sharing terminal lends credence to the notion that the individual enjoys and prefers the man/machine interactive approach over batch processing, if for no other reason than it helps him refine the definition of what he truly wants to search for.

But why, you may ask, an electron beam addressed memory? Why not any randomly addressible memory? The answer is that the memory must search itself. Obviously, the human operator cannot search it directly. Not quite so obvious, and therefore a more important consideration, is that is should not depend upon a central processing unit to search it. It should be in a system with "distributed logic" (34) where logic capability is readily accessible to all parts of the memory without waiting for block transfer to an intermediate level memory. It is the possibility of a silicon target having solid state logic circuits built on the same chip as the memory and activated by electron beam semiconductor switching that gives the electron beam its inherent flexibility for this application. Arntz has further suggested ways to include beam-activated content addressible memory units directly on the wafer at a fraction of the cost of present-day CAM units. (35)

We have yet to answer whether the MOS target described earlier is well suited to this "card index" type memory application. I think not. For one thing, it requires refresh which is a severe drawback for so large a memory. For another, it does not have good file security. "What can be intentionally erased can be accidently erased also." With very large files, integrity is of paramount importance.

Finally, the one micrometer level of resolution currently predicted for it with some degree of confidence is marginal from a cost standpoint for this size memory. A resolution of 0.1 micrometer is the desired level.

Fortunately there is a branch of silicon technology which does fill this need very nicely. It is the read only version of the silicon target wherein the information is engraved into an over layer (for example, the top metal layer) on a broad area silicon detector. The roots of this approach go back to Dove (36) who proposed use of holes in a metal film as the electron beam recording media and Norton (37) who proposed that the holes be placed over a silicon detector. While none of these earlier approaches had all the parts necessary to construct an "on line" data base, all the necessary technologies are now available (for example, H. Smith's x-ray lithography (38)); no invention is needed to produce read only memory of the required size with add-on capability and with each bit of the order of 0.1 micrometers on call (without need for intermediate transfer as opposed to simply on line) in front of a channel of a Fly's Eye Lens. This kind of memory does not suffer from bias switching delays since it is in read mode all the time. Deflection electronics will have to try to minimize their settling time to be compatible.

The Final Step Required

Thus far we have considered as an intermediate step a form of VLDB management which employs an on call card index. We have shown that the electron beam addressed silicon read only target is attractive as the technological vehicle for this approach. It is, in my opinion, the only operating approach with sufficient flexibility and low cost to accomplish this goal. We now ask how one may retrieve data and relationships between data which are not represented by the index. A good example of a data base with such requirements is a national population and productivity census.

The current approach to this requirement is the classical hierarchy of memories--that is, raw data is snatched a block at a time for processing in a serial operation. For a head count type of operation this is a valid approach, except for the large overhead cost of copying data through several levels of memory. The head count process is very slow and the results are limited in scope and application. Can we handle such questions as what is the effect of a new law or what would be the effect of dropping a series of night flights? When such interrogation involving interrelations of data are required, one really needs to have the entire memory on call with sufficient distributed logic to inactivate parts of the store not concerned with the query. The problem is then reduced to a manageable sized data base approaching that of the card index or smaller. An important point is that the reduced data base should not necessarily be transfered at this level. This remaining pertinent data, although physically scattered through the file, is available either for transfer to the more sophisticated CPU or for processing "Tic Tac Toe"fashion using the address hopping skill of the electron beam and the distributed logic. One of the unique qualities of electron beam access is its flexibility in addressing, it being able to rapidly alternate between serial and parallel processing.

This capability depends upon having every part of the file on call ready for instant addressing by the beam. The message which I wish to impart is that VLD bases are of limited usefulness unless the entire file is on call. For these files, the concept of memory hierarchy totally breaks down with the very large capacities already projected. To keep cost in line as the size of the file increases, the density must also increase. To place 10^{16} bits on call with 0.1 micrometer bit spacing would require something approaching 10^6 tubes of 10^{10} bits capacity each. I believe you will agree this is an unacceptable number of tubes. At 10^{13} bits per tube, only 1000 are required, but now we are considering 100 AU type resolution. Surely at this level we must invoke imaging techniques (16) to obtain sufficient parallel input and output. The pioneering work by Möllenstedt and his students may well be the approach which makes 100 AU performance practical (39, 40). Certainly no other technology offers the promise which electron beam technology has for this accomplishment.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the many helpful discussions with my colleagúes F.O. Árntz and D.E. Speliotis. I also wish to thank C.Q. Lemmond for the pictures and discussions concerning the General Electric work. Surely the reader will appreciate that the work evaluated here represents the creative work of hundreds of people and commitment of sizable resources over many years. We should all be thankful for their successes. Finally, I wish to point out that the long-term projections are personal accessments and are not necessarily those of my colleagues.

REFERENCES

1. Lawless, W.J., "Developments in Computer Organization," Advances in Electronics and El. Phys. (Marton, ed.), X, pp 153-84 (1958).

2. Williams, F.C., and Kilburn, T.A., "A Storage System for Use With Binary Digital Computing Machines," Proc. IEEE 96, part III, pp 81-100 (1949).

3. Jensen, A.S., Smith, J.P., Mesner, M.H., and Flory, L.E., "The Barrier Grid Storage Tube and Its Operation," RCA Review 9, pp 112-23 (1948).

4. Hines, M.E., Chruney, M., and McCarthy, J.A., "Digital Memory in Barrier Grid Storage Tubes," Bell Syst. Tech. Journ. 34, pp 1241-64 (1955).

5. Greenwood, T.S., and Staehler, R.E., "A High Speed Barrier-grid Store," Bell Syst. Tech. Journ. 37, pp 1195-220 (1958).

6. Speliotis, D.E., "Bridging the Memory Access Gap," AFIPS Conf. Proc. 44, pp 501-7 (1975).

7. Speliotis, D.E., Smith, D.O., Harte, K.J., and Arntz, F.O., "Electron Beam Memories," Proc. IEEE Electro-76 Int. Conf. Boston, May 1976.

8. Pugh, E.W., "Storage Hierarchies: Gaps, Cliffs, and Trends," IEEE Trans. Magnetics, MAG-7, pp 810-14 (1971).

9. Huber, E.F., Jr., Cohen, M.S., and Smith, D.O., "Electron Beam Detection of Charge Storage in MOS Capacitors," Appl. Phys. Lett. 16, pp 147-9 (1970).

10. Snow, E.H., Grove, A.S., and Fitzgerald, D.G., "Effects of Ionizing Radiation on Oxidized Silicon Surfaces and Planar Devices," Proc. IEEE 55, pp 1168-85 (1967).

11. MacDonald, N.C., and Everhart, T.E., "An Electron Beam Activated Switch and Associated Memory," Proc. IEEE 56, pp 158-66 (1968).

12. Cohen, M.S., and Moore, J.S., "Physics of the MOS Electron-Beam Memory," J. Appl. Phys. 45, pp 535-48 (1974).

13. Ellis, G.W., Possin, G.E., and Wilson, R.H., "Diode Detection of Information Stored in Electron Beam Addressed Structure," Appl. Phys. Lett. 24, p 19 (1974).

14. Arntz, F.O., Rockstad, H.K., Smith, D.O., and Sah, C.T., "Electron Beam Accessed Data Storage in MOS Capacitors," Digest, 1976 Int'l. Solid State Conf. Philadelphia, p. 176.

15. Hughes, W.C., Lemmond, C.Q., Parks, H.G., Ellis, G.W., Possin, G.E., and Wilson, R.H., "A Semiconductor Nonvolatile Electron Beam Accessed Mass Memory," Proc. IEEE 63, pp 1230-40 (1975).

16. Newberry, S.P., "Problems of Microspace Information Storage," Proc. Fourth Symposium on Electron Beam Technology (R. Bakish, ed.), pp 81-94, Alloyd Electronics Corp., Cambridge, Mass. (1962).

17. (SRI) Feature Editor Electronics, "MOS Target in a Vacuum Tube Yields Fast, Compact Memory," Electronics 47, pp 44-5, Oct. 3 (1974).

18. Selenyi, P., "Methods, Results and Prospects of Electrostatic Recording (Electrongraphy)," Zeit. f. tech. Physik 12, pp 607-14 (1935).

19. Selenyi, P., "Zeitsch. f. Phys. 47 (1928).

20. Broers, A.N., "Electron and Ion Probes," Fifth International Conf. on Electron and Ion Beam Science and Technology (R. Bakish, ed.), pp 3-25, Electro Chem.Soc. (1972).

21. Mulvey, T., "Electron-Optical Design of an X-Ray Microanalyzer," J. Sci. Inst. 36, pp 350-5 (1959).

22. Newberry, S.P., "An Electron Optical Technique for Large-Capacity Random-Access Memories," Proc. Fall Joint Comp. Conf. AFIPS, 29, p 717, Spartan Books, Washington (1966).

23. Lemmond, C.Q., Buschmann, E.C., Klotz, T.H., and White, G.M., "Electron Fly's Eye Lens Artwork Camera," IEEE Trans. on El. Devices, Vol. ED-21, pp 598-603 (1974).

24. Wang, C.C.T., "Analysis of Electrostatic Small-Angle Deflection," IEEE Trans. on El. Devices, Vol. ED-18, pp 258-74 (1971).

25. (Kelly) Heynick, L.N. (ed.), "High Information-Density Storage Surfaces," Stanford Res. Inst., Res. & Dev. Tech. Report ECOM-01261-F, June (1970).

26. Schlessinger, K., "Internal Electrostatic Deflection Yokes," Electronics 25, p 105 (1952).

27. Wang, C.C.T., Harte, K.J., Curland, N., Likuski, R.K., Dougherty, E.C., "Computer-Aided Design and Experimental Investigation of an Electron Optical Collimating Lens," Proc. Twelfth Symposium on Electron Ion and Laser Beam Technology, pp 1110-3, J. Vac. Soc. #6 Nov/Dec (1973).

28. Harte, K.J., "Theory of Lenslet Interaction in a Fly's Eye Lens" (abstract only), Proc. Thirteenth Symposium on Electron, Ion and Photon Beam Technology, p 1160, Am. Vac. Soc. (1974), also J. Vac. Sci, Tech. 12 #6, Nov/Dec (1975).

29. Pohm, A.V., "Cost/Performance Perspectives of Paging With Electronic and Electromechanical Backing Stores," Proc. IEEE 63, pp 1123-8 (1975).

30. Fletcher, J.G., Fernbach, S., Dubois, P.J., and Boer, G.L., "Computer Storage Structure and Utilization at a Large Scientific Laboratory," Proc. IEEE 63, pp 1104-13 (1975).

31. Codd, E.F. (session chairman), "Data Base Management," AFIPS Conf. Proc. 44, pp 377-493, National Computer Conference, Anaheim, Cal. (1975).

32. Date, C.J., "An Introduction to Data Base Systems," Addison-Wesley Pub. Co., Reading, Mass. (1976).

33. Johnson, C.T., "The IBM 3850: A Mass Storage System With Disk Characteristics," Proc. IEEE 63, pp 1166-70 (1975).

34. Slotnick, D.L., "Logic Per Track Devices," Advances in Computers 10 (Franz Alt, ed.), pp 291-6, Academic Press, N.Y. (1970).

35. Arntz, F.O., private communication.

36. (Dove) Costa, J. (ed.), "New Data Recorder May Increase Memory 3 Orders of Magnitude," Electronic News, p 5, Sept. 28 (1964).

37. Norton, J.F., "Semiconductor Data Storage Apparatus With Electron Beam Readout," U.S. Pat. #3,550,094 (1970).

38. Smith, H.I., Spears, D.L., and Bernacki, S.E., "X-Ray Lithography: A Complementary Technique to Electron Beam Lithography," J. Vac. Sci. Tech. 10, pp 913-7 (1973).

39. Möllenstedt, G., Speidel, B., "Newer Developments in Microminiaturization," Third Symposium on Electron Beam, p 340, Alloyd Corp., Boston, Mass. (1961).

40. Koops, H., "Elektronenoptishe Verkleinerung durch Kombination einer Langbrennweitigen Feldlinse mit einer Starken Kurzbrennweitigen Magnetlinse," Optik 36, pp 93-110 (1972).

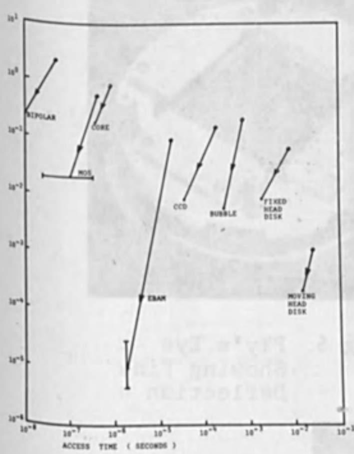


Fig 1 The "Access Gap"

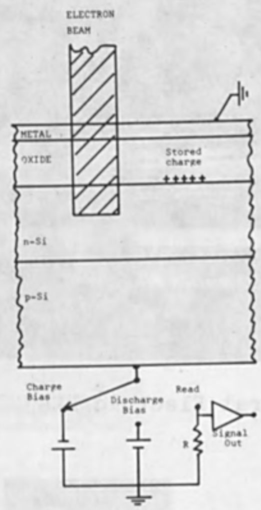


Fig 2 MCS Target

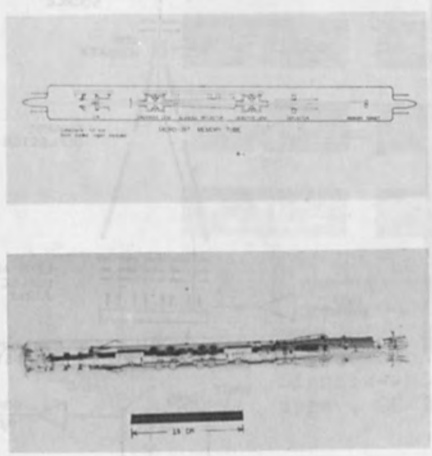


Fig 3 Micro-Bit EBAM Tube
Schematic upper, Tube lower

Fig 4 General Electric BEAMOS Tube

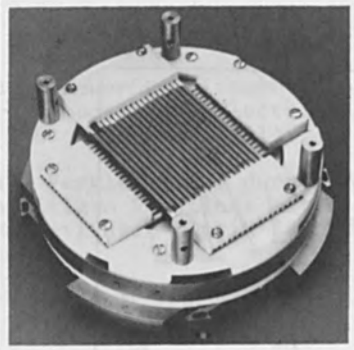


Fig 5 Fly's Eye Showing Fine Deflection

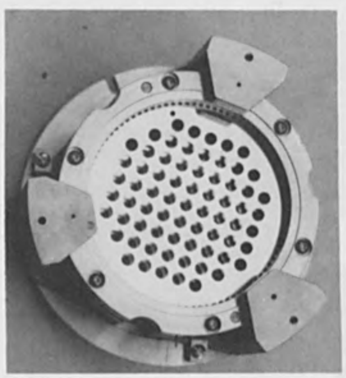


Fig 6 Fly's Eye Showing Lenslet Array

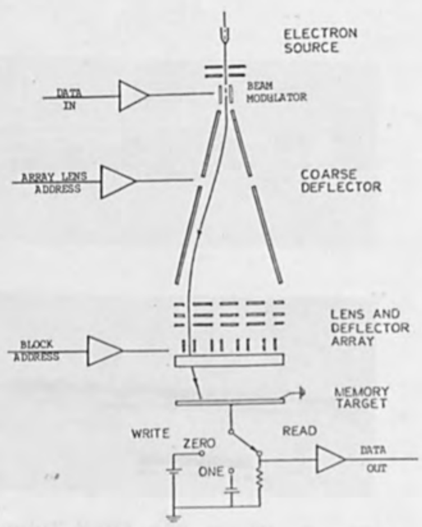


Fig 7 Fly's Tube Schematic

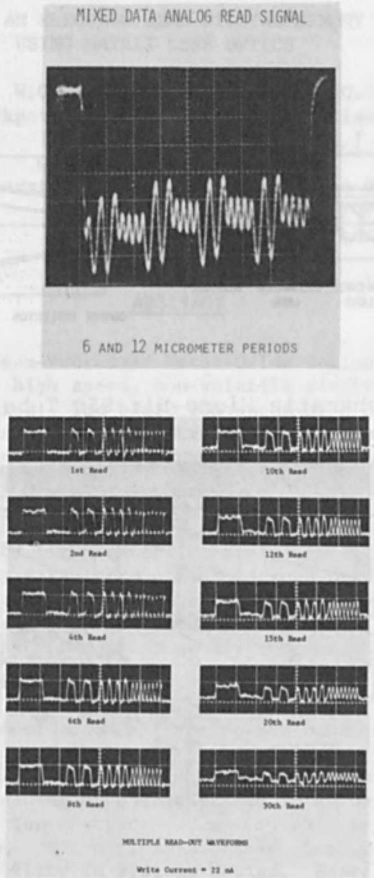


Fig 8 Typical Unprocessed Signals--Micro-Bit upper, GE lower

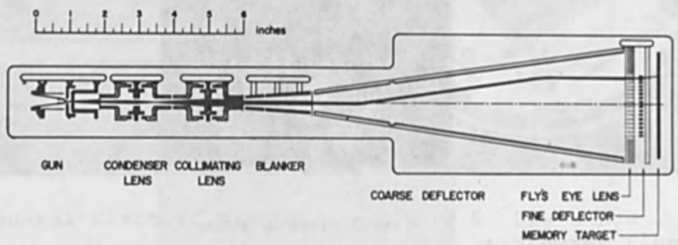


Fig 9 Schematic Micro-Bit 950 Tube

Fig 10 The Problems of a Very Large Randomly Accessed Data Base

BEAMOS -- AN ELECTRON BEAM DIGITAL MEMORY DEVICE USING MATRIX LENS OPTICS

H.G. Parks, W.C. Hughes, G.E. Possin, C.Q. Lemmond, C.G. Kirkpatrick, G.W. Ellis, R.H. Wilson

General Electric Company
Corporate Research and Development Center
One River Road
Schenectady, NY 12345

ABSTRACT

BEAMOS, Beam-Addressed Metal-Oxide Semiconductor is a large capacity high speed, non-volatile electron beam digital memory device. The basic component of a BEAMOS memory system consists of an unstructured MOS memory plane and a matrix lens electron beam accessing system, housed in a sealed electron tube, capable of $\geq 32 \times 10^6$ bits storage. This paper discusses the basic physics and operation of the memory component as a prelude to the central theme of electron optics design considerations. The influencing parameters of the various parts of the memory component and their interaction in the fundamental electron optics design is discussed in detail. The paper ends with a status report on memory component development.

Introduction

BEAMOS, Beam Addressed Metal-Oxide-Semiconductor, is an emerging memory technology for large bit capacity, rapid axis, high data rate, non-volatile storage. It is based on electron beam writing and reading of charge stored in an unstructured MOS memory plane. The unstructured memory chip and the very rapid switching time of the finely focused electron beam allow considerable flexibility in systems design. Representative performance of current BEAMOS technology is as follows:

Component capacity	32×10^6 bits
Systems capacity	32×10^6 to 600×10^6 bits or more
Access time	30 μs
Data transfer rate	10 Mbit/s per component

Previous papers[1,2] have dealt with the memory plane physics and memory system considerations. In this paper, the basic tar-- get physics and principles of operation of the matrix lens optics are reviewed. The most important beam limiting aberrations of the matrix lens are identified and general design considerations highlighting the interaction of memory plane, optics, mechanical and system requirements are discussed. Finally, a status report of the BEAMOS memory component development is presented.

Memory Component

The BEAMOS memory component has two major parts, an unstructured MOS memory plane in which data is stored as small islands of positive charge and an electron beam matrix lens accessing system for storage and retrieval. These are housed in a sealed glass electron tube approximately four inches in diameter and 17 inches long shown in figure 1.

Fig. 1 BEAMOS memory component

In this section, these two major parts of the memory component are reviewed. The basic physics and operation of the target are discussed first, leading to the requirements imposed on the optics by the memory plane. Next, the advantages and operation of the matrix lens optics are discussed. This serves as an introduction to the optical components that must be considered in the following sections.

The BEAMOS memory element[3] shown in cross section in figure 2 consists of a reverse biased np junction with an overlying MOS capacitor. The total oxide and aluminum layer thickness is thin enough to allow the addressing beam to penetrate through them and into the n-layer. The boron doped P substrate is much more lightly doped than the phosphorous doped n-layer by a factor of approximately 200.

Data is stored by biasing the aluminum layer positive with respect to the n-layer and directing the electron beam to the areas where ones are to be recorded. Under these conditions, positive charge is stored in the silicon dioxide and will remain there for months whether power is applied or not. Note that the memory plane has no fine patterning, the bit sites are defined by the writing beam itself. Erasure is carried out by applying a negative bias to the aluminum and directing the beam to the site to be erased.

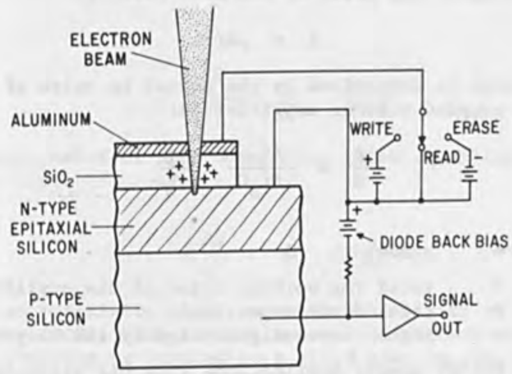


Fig. 2 Cross section of MOS memory chip

Electron beam addressing is also used for readout. The back biased diode of the memory plane acts as a current amplifier for the reading beam, providing a current gain which varies with the amount of charge stored at the addressed site. Current gain comes about because each electron penetrating into the n-layer produces a minority carrier (hole) for each 3.5 electron volts of energy. For present memory plane designs, this results in approximately 1400 holes/electron. If no charge is stored in the oxide, most of these carriers recombine with electrons at the oxide-silicon interface. About 200 escape recombination and drift to the junction, thereby producing current in the signal circuit. Under regions where positive charge is stored, the surface recombination velocity is greatly reduced and more holes are able to reach the junction. Thus, the gain varies from approximately 200 where no charge is stored, to a value approaching 1400 where there is charge. A read operation removes some of the stored charge, but many reads are possible before writing is necessary. Readout rates can be 10 megabits/sec or higher. Storage density can be 3×10^7 bits/in^2 or greater.

The spatial resolution or bit packing density of the memory plane is determined by the incident beam spot size, electron scattering in the target and lateral diffusion of the injected holes in the n-layer. Measurements of electron scattering[4] at 5 and 10 keV show that the effective increase in electron beam spot size in the target is less than 0.5 µm. Hence, electron beam spot size and n-layer diffusion are the primary considerations in determining the bit packing density.

The memory plane also interacts with system requirements to determine beam current that must be delivered to the target. The writing current is determined by required target fluence, ρ (typically 10^{-7} coul/cm^2), bit area, A, and data rate, f:

$$I = \rho A f \tag{1}$$

Reading current is determined by the signal to noise of the target and voltage coupled readout amplifier as:

$$\frac{S}{N} = \frac{\sqrt{3}}{2\pi f_b C} \quad \frac{i}{\langle E \rangle} \tag{2}$$

where:

f_b	=	bandwidth
$\langle E \rangle$	=	total rms voltage noise of the amplifier
C	=	total diode capacitance
i	=	signal current generated by the target

For present BEAMOS memory applications with bit sizes in the 2.5 to 4 μm range, S/N = 15, the write beam current required is approximately 20 times the read current.

Thus, the memory plane and system requirements specify the electron beam spot size, current, and landing energy. The limiting parameters of the matrix lens optics and design considerations including these requirements are discussed in the following sections of this paper.

The matrix lens is a unique electron optical device that achieves a much larger total field of high resolution spot diameters without the need for beam servoing. It consists of an array of electron lenses, each with its own deflection system, that are addressed by a single electron beam. In essence, a total field size equivalent to that of a single lens and deflection system multiplied by the number of lenslets can be realized.

A conventional single lens and deflection system without servo positioning can accurately locate an electron beam over a discrete number of beam diameters. The size of such a field, limited by electronic stability, can readily be determined. For this purpose, consider a square array of N x N spots of diameter and center to center distance d, addressed by single lens and deflection system. Maximum deflection over a linear edge of such an array is then approximately Nd/2 and is directly proportional to the deflection drive voltage V_D, and inversely proportional to the beam potential, V_B. Taking a tolerable position error of one-tenth of a spot diameter, and considering variations in V_D and V_B, independently leads to the voltage stability requirements:

$$\frac{\Delta V_D}{V_D} \approx \frac{1}{5N} \tag{3}$$

and

$$\frac{\Delta V_B}{V_B} \approx \frac{1}{5N-1} \tag{4}$$

For a large number of spot diameters, these equations are essentially the same, i.e.:

$$\frac{\Delta V}{V} \approx \frac{1}{5N} \tag{5}$$

Assuming an optimistic limit on voltage stability of .01%, equation 3 gives N = 2000. Thus, a single lens and deflection system without beam servoing is limited to a field size on the order of 2000 x 2000 spot diameters.

A typical matrix lens recording system is shown in figure 3.

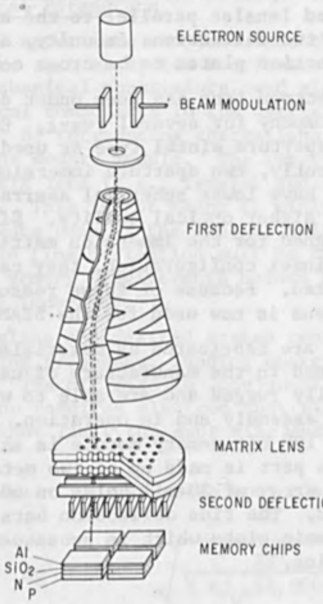


Fig. 3 Matrix lens recording system

Electrons emerging from the source are deflected by a lenslet selector (first deflection) to the desired lenslet, which focuses the beam on the recording medium. Deflection over a small subfield of the target under each lenslet is performed with the fine deflection system (second deflection). The lenslet selection is accomplished in such a way that the beam enters each lenslet orthogonally and on its optical axis. Therefore, each lenslet, which operates with a large demagnification (~ 25:1), images an on-axis object. Because of the large demagnification provided by the lenslet, beam position errors and aberrations introduced to the beam by the lenslet selector are greatly reduced at the image plane. The lenslet selector and fine deflection are therefore substantially independent of each other. Power supply and fine deflection circuit voltage stabilities of $\pm$ 0.05 percent are sufficient for positioning and holding micron size final spots to within one-tenth of a spot diameter.

Lenslet selection is accomplished with a conical double deflectron. The deflectron[5] consists of sets of interleaved sinusoidally patterned electrodes, which produce a highly uniform deflection field which results in much less scan distortion than conventional parallel plate-deflection systems. The first deflectron deflects the beam off-axis and the second unit redirects the beam parallel to the optic axis. This redirection causes the beam to enter the desired lenslet parallel to the axis. The voltage ratios between the two deflectrons is unity, allowing the upper and lower sets of deflection plates to be cross connected.

The matrix electron lens has been under development by the General Electric Company for several years. Early matrix lenslets were of the three aperture einzel type as used in an artwork camera[6]. More recently, two aperture immersion matrix lenses have been shown to have lower spherical aberration coefficients[7] and, hence, are of higher optical quality. Since only two lens plates must be aligned for the immersion matrix lens, as opposed to three for the einzel configuration, they can be more accurately and easily fabricated. Because of these reasons, the two aperture immersion matrix lens is now used for the BEAMOS memory component.

Matrix lenses are fabricated by materials and techniques similar to those used in the manufacture of ceramic metal tubes. They are mechanically rugged and are able to withstand a wide range of temperatures in assembly and in operation. The matrix lens used in the BEAMOS 32 x 10^6 bit memory module is shown in top view in figure 4. The lens part is made up of two metal plates into each of which a 18 x 18 array of 30-mil holes on 60-mil centers has been chemically machined. The fine deflection bars are formed by machining slots in a ceramic plate which is subsequently given a conductive metal coating.

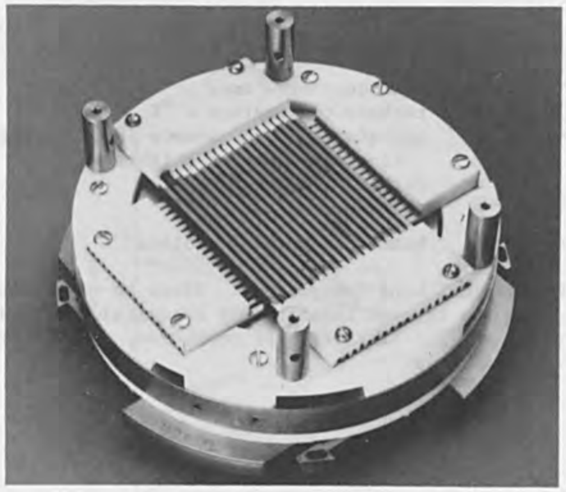


Fig. 4 Matrix lens top view

Matrix Lens Optical Parameters

Many factors enter into and influence the design of matrix lens memory optics such as system requirements, memory plane requirements, mechanical constraints, and electron optical factors. Each of the optical components in the matrix lens addressing optics has the potential of introducing aberrations that could limit the performance of the memory module. The important limiting aberrations of the matrix lens (i.e., lenslet and fine deflection system) will be assessed. This is not meant to minimize the importance of the other components such as the lenslet selector. However, the matrix lenslet and fine deflection must be compatible with, and satisfy, desired memory requirements before the effects of other components become important.

The memory plane and desired system requirements determine the bit size, operating frequency, landing energy, and beam current of BEAMOS memories. The bit size in conjunction with target scattering determines the required electron spot size. Lenslet electron optical factors that limit the spot size and maximum attainable current are:

Thermal Effects: The gaussian spot size, dg, is:

$$dg \approx \sqrt{\frac{IT}{Y\,\pi j_o\,11,600\,V_B}} \cdot \frac{1}{\varphi} \qquad (6a)$$

where:

I	=	beam current - amps
T	=	cathode temperature - ^{o}K
γ	=	efficiency of the source in producing Langmuir limiting brightness
j_o	=	cathode loading A/cm^2
V_B	=	voltage - volts
φ	=	beam half angle - radians

Geometrical Lens Aberrations: Since an on-axis object is presented to each lenslet only spherical aberration is important. The disk of least confusion, d_s, due to spherical aberration is:

$$d_s = \tfrac{1}{2} C_s \varphi^3 \qquad (6b)$$

where C_s is the spherical aberration coefficient of the lens.

Chromatic Aberration: The diameter of the disk of least confusion due to chromatic aberration, d_c, is:

$$d_c = 2C_c \frac{\Delta V_B}{V_B} \varphi \qquad (6c)$$

where:

C_c	=	chromatic aberration coefficient of lens
V_B	=	nominal beam potential
ΔV_B	=	beam potential variation

Diffraction: The minimum resolvable beam diameter, d_d, due to diffraction in angstrom units is:

$$d_d = \frac{12.3}{\sqrt{V_B}} \cdot \frac{1}{\varphi} \qquad (6d)$$

Present BEAMOS modules operate with bit sizes in the low micron range at a data transfer rate of 10^7 bits per sec. As may well be anticipated for these operating conditions, thermal effects and spherical aberration prove to be the limiting electron optical parameters for the lenslets. This is easily illustrated with a logarithmic plot of the individual limits imposed by equations 6a thru 6d versus beam half angle φ as shown in figure 5.

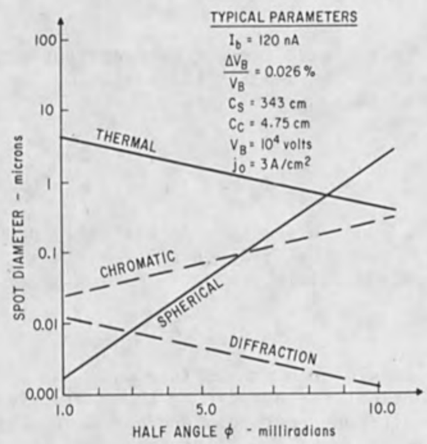


Fig. 5 Matrix lens aberrations

Typical parameters assumed in these calculations, as indicated on this figure, are based on the 32 x 10^6 bit memory module.

Since it is spherical aberration and thermal effects that limit the matrix lenslets one can define a maximum spherical aberration coefficient for given performance requirements as:

$$C_{smax} = (k_1 d)^4 \left[\frac{\beta_i}{I}\right]^{3/2} \qquad (7)$$

where:

d	=	desired spot size (cm)
β_i	=	brightness of the beam at the target (A/cm^2/str)
I	=	beam current (amperes)
k_1	=	1.3 for 90 percent beam current in spot d

Determination of C_{smax} defines an optimum half angle of convergence of the beam at the target, φ_{opt}, given by:

$$\varphi_{opt} = \left[\frac{k_2 d}{C_{smax}}\right]^{1/3} \tag{8}$$

where $k_2 = 0.9$ for a 90 percent beam current profile. The theoretical brightness of the beam at the target, β, is related to cathode parameters by the Langmuir limit[8]:

$$\beta_T = j_o \frac{11,600}{\pi T} V_L \tag{9}$$

Electron guns in practice do not fully produce this value of brightness. The actual brightness of the beam at the target, β_i, is given more accurately by:

$$\beta_i = \gamma B_T \tag{10}$$

where γ is the electron gun efficiency. The electron guns used for BEAMOS optical systems are of the Radley-Pierce construction, with barium dispenser cathodes and can have efficiencies as high as 80 percent (i.e., $\gamma = 0.8$) as previously reported by Hughes[9].

The next most important aberrations for BEAMOS matrix lens systems are those associated with the fine deflection system. A third order theory of deflection aberrations, based on a method of path integration proposed by Hutter[10], has been developed to treat fine deflection aberrations in detail. This method uses the calculus of variations and an expansion of the deflecting field to determine the deflection aberration coefficients. According to this theory, the third order deviations, $\Delta^3 x$ and $\Delta^3 y$, from an ideal point due to deflection errors produced by deflection in the y-direction, are given as:

$$\Delta^3 x = \alpha_{0010} x'_s + \alpha_{0011} x'_s y'_s + \alpha_{0110}\, y_s x'_s \tag{11a}$$

$$\Delta^3 y = \beta_{0000} + \beta_{0100} y_s + \beta_{0200} (y_s)^2 + \beta_{0001} y'_s + \beta_{0101} y_s y'_s + \beta_{0020} (x'_s)^2 + \beta_{0002} (y'_s)^2 \tag{11b}$$

where:

α's and β's = coefficients of the deflecting field and first order deflection summarized in Table 1

x'_s = slope of undeflected ray in the x-direction at the target

y_s = position of the undeflected beam in the y-direction at the target

TABLE - 1

SUMMARY OF THIRD ORDER DEFLECTION COEFFICIENTS FOR DEFLECTION IN Y-DIRECTION

$$\alpha_{0010} = \frac{YsYs}{2} + \frac{3}{4V}\int EYdz$$

$$\alpha_{0011} = 2Ys$$

$$\alpha_{0110} = -Ys'$$

$$\beta_{0000} = \frac{YsYs'}{2} + \frac{1}{V}\int EYY'dz + \frac{1}{4V^2}\int EY(z - zs)dz + \frac{1}{4V}\int EY'^2(z - zs)dz - \frac{1}{4V}\int E''Y^2(z - zs)dz$$

$$\beta_{0100} = \frac{1}{2V}\int EY'dz + \frac{1}{4V^2}\int E^2(z - zs)dz - \frac{1}{2V}\int E''Y(z - zs)dz$$

$$\beta_{0200} = -\frac{1}{4V}\int E''(z - zs)dz$$

$$\beta_{0001} = \frac{3}{2}YsYs' + \frac{5}{4V}\int EYdz$$

$$\beta_{0101} = -Ys' - \frac{1}{2V}\int E''(z - zs)^2dz$$

$$\beta_{0020} = Ys$$

$$\beta_{0002} = 3Ys - \frac{1}{4V}\int E''(z - zs)^3dz$$

Note all integrals have limits zo, zs (i.e., entrance of deflection to target).
Y,Y' = Deflected ray height and slope along the path, respectively
Ys, Y's = Deflected ray height and slope at the target, respectively
E = Deflecting field along the z-axis
V = Beam potential with respect to the cathode

These equations can be applied to each of the sequential deflection plates of the matrix lens fine deflection system. Assuming an initially point focused conical beam of half angle φ, deflected in the x- and y-directions by θ_x and θ_y, respectively, the third order aberrations for the matrix lens fine deflection system are:

$$\Delta^3 x = A_{300}\theta_x^3 + A_{120}\theta_x\theta_y^2 + A_{201}\theta_x^2\varphi + A_{111}\theta_x\theta_y\varphi$$

$$+ A_{021}\theta_y^2\varphi + A_{102}\theta_x\varphi^2 + A_{012}\theta_y\varphi^2 \qquad (12a)$$

and

$$\Delta^3 y = B_{030}\theta_y^3 + B_{210}\theta_x^2\theta_y + B_{021}\theta_y^2\varphi + B_{111}\theta_x\theta_y\varphi$$

$$+ B_{201}\theta_x^2\varphi + B_{102}\theta_x\varphi^2 + B_{012}\theta_y\varphi^2 \qquad (12b)$$

The A and B coefficients in equations 12a and 12b are evaluated from the α and β coefficients summarized in Table 1 and are functions of the deflection system parameters. Terms involving θ are distortion aberrations, i.e., the position of the beam is affected, but not the spot size. The terms involving θ and φ affect the spot size. These are astigmatism and field curvature aberrations for terms with $\theta^2\varphi$ dependence and coma aberrations for terms $\theta\varphi^2$ dependence. The A and B coefficients in equations 12a and 12b have been evaluated assuming an ideal parallel plate matrix lens fine deflection system, figure 6, for an initially point focused conical beam, figure 7. The resulting coefficients are summarized in Table 2.

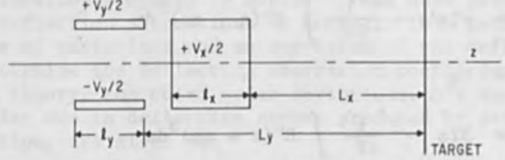


Fig. 6 Parallel plate deflection system

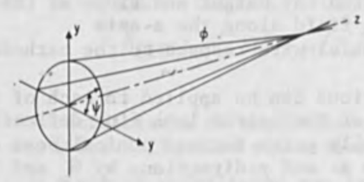


Fig. 7 Point focus concial beam

TABLE - 2

THIRD ORDER DEFLECTION COEFFICIENTS
FOR SEQUENTIAL Y - AND X-DEFLECTION SYSTEM

$\Delta^3 x$ Coefficients	$\Delta^3 y$ Coefficients
$A_{300} = \frac{Lx}{2}$	$B_{030} = L_{y/2}$
$A_{120} = 2(Lx + 1_{x/2}$	$B_{210} = L_{x/2}$
$A_{201} = \frac{L^2 x}{1_x} + 3.5L_x + 1_x \cos\Psi$	$B_{021} = \frac{Ly^2}{1_y} + 3.5\ Ly + 1Y \sin\Psi$
$A_{111} = 2(L_x + 1_{x/2})\sin\Psi$	$B_{111} = 2(L_x + 1_{x/2})\cos\Psi$
$A_{021} = \frac{L_y}{2}\cos\Psi$	$B_{201} = \frac{L_x}{2}\sin\Psi$
$A_{102} = (L_x + \frac{1_x}{2})(\sin^2\Psi + 3\cos^2\Psi$	$B_{102} = 2(L_x + 1_{x/2})\sin\Psi\cos\Psi$
$A_{012} = 2(L_y + 1_{y/2})\sin\Psi\cos\Psi$	$B_{012} = (L_y + 1_y)(\cos^2\Psi + 3\sin^2\Psi)$

Although not quantitatively exact, the aberrations represented by these equations can be used to qualitatively determine the dominant terms. Since the x plates are closest to the target, they operate with the largest deflection angle and therefore introduce the largest aberrations for deflection over a square field. For memory applications, the most concern is with the deflection aberrations that affect the spot shape, i.e., astigmatism and coma. As long as the deflection is stable and repeatable, distortion aberrations thataffect the spot position can be tolerated within reasonable limits. This is not true for artwork camera applications where straight lines and matching fields from lenslet to lenslet are required.

Considering just deflection in the x direction and using the coefficients of Table 2, the conditions for astigmatism-field curvature aberrations $\Delta^3 x_{AF}$ to be greater than coma aberrations $\Delta^3 x_C$ can be ascertained:

$$\frac{\Delta^3 x_{AF}}{\Delta^3 x_C} = \frac{k^2 + 7/2k + 1}{k + 1/2} \cdot \frac{\theta_x}{\varphi} > 1 \tag{13}$$

where $k = L_x/\ell_x$

Equation 13 can be rewritten as:

$$\frac{\theta_x}{\varphi} > \frac{k + 1/2}{k^2 + 7/2k + 1} \tag{14}$$

As defined $k \geq 0$, hence θ/φ has a maximum value of 0.5 at k = 0 and falls monotonically to 0 as $k \to \infty$ as shown in figure 8.

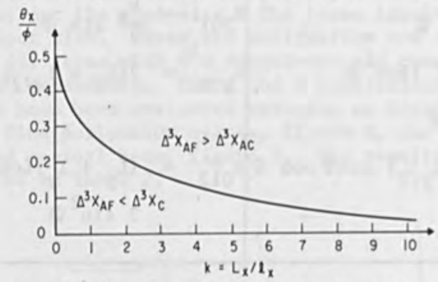


Fig. 8 Ratio of deflection angle to beam half angle versus k for equal deflection field curvature/astigmatism and coma aberrations

Therefore, astigmatism and field curvature are the dominant aberrations of the fine deflection system as long as $\theta/\varphi_{max} > 5$. This is always the case for matrix lenses as can be shown by the two hypothetical but limiting cases for θ/φ_{max}, shown in figure 9. In both of these cases, the half angle is set at a maximum by using the full lens aperture. This actually sets a larger value for φ_{max} than is possible since spherical aberration does not allow the full aperture of the lens to be used. The lenslets, as in practice, are centered with a distance of two lenslet diameters, 2D, to avoid lenslet to lenslet interaction, thus allowing a maximum deflection of $\pm$ D for each lenslet. In the first case, the deflection bars are assumed to be infinitisimally small leading to:

$$\frac{\theta}{\varphi_{max}} = \frac{D/Z}{D/2Z} = 2 \tag{15}$$

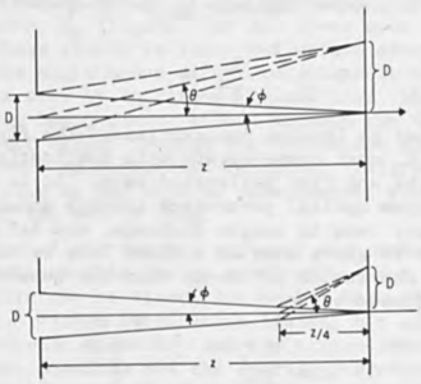


Fig. 9 Hypothetical deflection cases. Top - infinitisimal deflection bars. Bottom - halfspace deflection bars.

In the second case, the x and y deflection bars are assumed to each fill up half the deflection space leading to the condition:

$$\frac{\theta}{\varphi_{max}} = \frac{4D/Z}{D/2Z} = 8 \qquad (16)$$

Extensions of the simple argument can be made to the complete aberrations equations 12a and 12b, resulting in the same conclusion that astigmatism and field curvature are the dominant terms for the matrix lens fine deflection system.

BEAMOS Matrix Lens Design Considerations

The memory plane, matrix lens optics, system requirements and fabrication capabilities all interact in the final design of a BEAMOS memory component. The evolution of BEAMOS memory components has followed density increases governed mainly by target development. For a given memory plane configuration the bit size is set by the n-layer thickness. This in turn sets the electron spot size and beam landing energy through beam spreading and beam current, in conjunction with the system requirement on data rate.

Specification of the electron spot size interacts first with electronic stability considerations to determine the lenslet field size. The field size in turn specifies the lenslet center to center distance, X_L, for the matrix lens. The lenslet diameter is related to the center to center distance by the inequality:

$$D \leq \frac{X_L}{2} \tag{17}$$

to avoid lenslet to lenslet interactions. The final relationship between D and X_L must commensurate with fabrication technology for the lenslets and fine deflection bars. It is also interwoven with the electron optical parameters through spherical aberration of the lenslets, lens to target distance, and deflection aberrations. Figure 10 shows that at a fixed lens to target distance the spherical aberration is lowest when the lenslet diameter has its maximum allowed value.

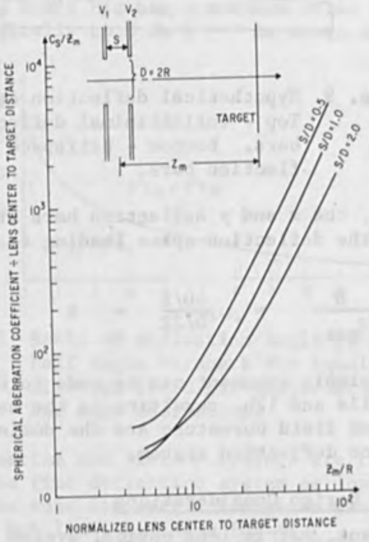


Fig. 10 Two aperture immersion lens characteristics - accelerating case $V_2 > V_1$

For given spot size, d, beam current, I_B, and landing potential, V_L, the maximum spherical aberration coefficient for the lens and optimum convergence angle are given by equations 7 and 8. The spherical aberration is related to the cathode loading and temperature by the source brightness equations 9 and 10. Higher cathode loadings produce brighter electron sources at higher cathode temperature and therefore, at lower source lifetimes. The higher brightness allows larger values for C_{smax} and therefore larger lens to memory plane distance, Z_m (figure 10) and lower beam convergence angles. Both of these result in lower deflection aberrations. The beam convergence angle has a direct relationship to deflection aberrations as seen by equations 12a and 12b. The lens to memory plane spacing has a direct effect through the deflection angle and an indirect effect through the deflection aberration coefficient dependence on deflection geometry. This dependence on geometry again ties in and must be commensurate with fabrication techniques for the matrix lens.

The interaction of the various parts of the memory component indicate several possible design trade-offs. For example, memory component life can be traded for higher data rate and/or higher bit density. Within fabrication limits, bit size may be traded for electronic stability and deflection sensitivity. Other trade-offs are possible and the overall performance requirements and application of a memory component serves to identify the most important ones. The final design then incorporates the resulting matrix lens with geometric and lenslet selector considerations in the memory component.

BEAMOS Memory Component Status

Present BEAMOS components of 32 and 100 megabits have been designed for maximum life with a data rate of 10 megabits per sec. and access time of 30 μsec. The 32×10^6 bit component, shown previously in figure 1 is being built in pilot quantities and operated in a computer controlled test system. The 10^8 bit memory component shown is quite similar in exterior appearance and dimensions to the 32×10^6 bit component. It has been successfully operated at 10 megabits per second and 2.5 μm bit spacing in limited test situations. The matrix lens for the smaller bit size operates with longer lens to target distance and deflection bars for reduced deflection aberrations. A photograph of the 10^8 bit matrix lens is shown in figure 11, and the memory component in figure 12.

Fig. 11 10^8 bit BEAMOS matrix lens

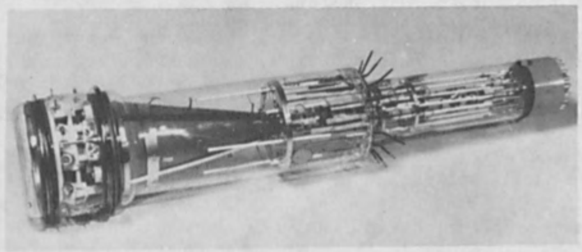


Fig. 12 10^8 Bit BEAMOS memory component

References

1. W.C. Hughes, C.Q. Lemmond, H.G. Parks, G.W. Ellis, G.E. Possin, R.H. Wilson, "BEAMOS -- An Electron Beam Addressed Semiconductor Mass Memory", IEEE Proceedings, Vol. 63, 48, pp. 1230-1240, August 1975.

2. W.C. Hughes, C.Q. Lemmond, H.G. Parks, G.W. Ellis, G.E. Possin and R.H. Wilson, "BEAMOS -- A New Electronic Digital Memory", 1975 National Computer Conference, Anaheim, California, May 19-20, 1975.

3. G.W. Ellis, G.E. Possin and R.H. Wilson, "Diode Detection of Information Stored in Electron Beam Addressed Structure", Applied Physics Letters, Vol. 24, pp. 419, 1974.

4. G.E. Possin and J.F. Norton, "Spatial Distribution of 5 and 10 Kilovolt Electron Beam Ionization in Solids", Proceedings of the Eighth Annual Scanning Electron Microscope Symposium, Illinois Institute of Technology Research Institute, Chicago, Illinois, April 1975, pp. 457-464.

5. K. Schlesinger, "Progress in the Development of Post-acceleration and Electrostatic Deflection", Proceedings IRE, Vol. 44, pp. 659-667, May 1956.

6. C.Q. Lemmond, E.C. Buschmann, T.H. Klotz, and G.M. White, "Electron Fly's Eye Lens Artwork Camera", IEEE Transactions on Electron Devices, Vol. ED-21, No. 9, September 1974, pp. 598-603.

7. H.G. Parks, W.C. Hughes, "Advances in Matrix Lens Technology", Journal Vac. Science Technology, Vol. 12, No. 6, Nov./Dec. 1975, pp. 1116-1164.

8. D.B. Langmiur, "Theoretical Limitations of Cathode Ray Tubes", Proceedings of the IRE, Vol. 25, August 1937, pp. 977-991.

9. W.C. Hughes, "Long Life, High Brightness Sources for Demountable Guns", Record of the 10th Electron, Ion, and Laser Beam Symposium, Gaithersburg, Maryland, May 1969, San Francisco Press, Inc., San Francisco, California, 1969.

10. R.G.E. Hutter, "Electron Beam Deflection, Part I: Small Angle Deflection Theory", Journal of Applied Physics, Vol. 18, August 1947, pp. 740-758.

ELECTROCOMPOSEUR, A FAST ELECTRON BEAM PATTERN GENERATOR

Jacques Trotel
THOMSON-CSF, Laboratoire Central de Recherches
B.P. 10 91401 - ORSAY (France)

When using an electron beam for microlithography one can expect two main differences when compared with visible light : an electron beam can be focused to a small diameter and it can be deflected very easily.

Thus, the specific uses of electron beams in microlithography should be foward the realization of very small geometries (taking advantage of the small beam diameter) and the fast generation of complex patterns (using high speed deflection).

In the semiconductor industry conventional mask shop equipment is well suited for the present day needs of planar technology ; for example, photo-repeaters are fast, rugged and reliable. However, a bottleneck exists with optical pattern generators, which makes it difficult to develop a new circuit in a short time. To fill this gap, we have produced an electron beam machine which is a very fast pattern generator and which can be inserted in a conventional mask shop.

The Electrocomposeur has been built to generate 10 X reticles in a production environment. Emphasis has been put on :

- Fast generation time
- High reliability
- Low defect density by using photoresist

Characteristics of the system

. Maximum plate size	6" x 6"
. Maximum exposed area	5" x 5"
. Minimum increment	2.5 µm
. Minimum linewidth	2.5 µm
. Maximum drawing time of a fully covered 3" x 3" pattern	17 minutes
. Sample (typical)	Chrome plate with AZ 1350 resist

General description

As indicated in figure 1, the pattern is described in a raster scan mode which can be compared to a television scan where line scan is done by beam deflection in the X direction and image scan is done by mechanical motion of the table in the Y direction. The movements of XY table are measured by two laser interferometers. When the beam is writing, the X motor is stopped and the Y motor drives the table at a constant speed. The Y interferometer triggers a beam scan each 2.5 µm ; the X interferometer measures the small variations due to imperfections of the movement and forces the beam to track these déviations from a perfect translation. The X interferometer is used also to stitch together the stripes which are made at each image scan. Figure 2 gives an overall view of the system.

Electron optics

The electron optical column focuses a square beam on the sample and scans this beam in only one direction. A La B6 electron gun makes it possible to write patterns on AZ 1350 resists ; a conventional tungsten filament could be used with high sensitivity resists.

Main characteristics of the electron optics are :

. Voltage : 20 kV
. Brightness of electron gun : $\approx$ 3.10^5 A $cm^{-2}Sr^{-1}$
. Beam blanking : electrostatic
. Number of lenses : 3
. Beam deflection : electrostatic : 1.25 mm amplitude
. Shaped beam : 2.5 μm x 2.5 μm

Data for description of the pattern

The pattern is defined as a set of polygons. The sides of these polygons can be parallel to the X and Y axis and to the bisectors of these axis. The pattern is divided into stripes 1.25 mm wide parallel to X axis. The angles of the polygons are then ordered according to their coordinates and are recorded on a magnetic tape. The magnetic tape is read by the machine while the pattern is being drawn. An HP computer is used mainly as a buffer memory between the magnetic tape reader and the machine : this buffer memory is necessary to match the data transfer rate of the magnetic tape to the speed of the machine.

Results

Figures 3 and 4 give some details of patterns drawn on negative resist, the scale is given by the wiggle on the edge of the lines : 2.5 μm.

From figure 4 we see that 2.5 μm lines can really be delineated and that, although the machine is designed to make 10 X reticles, in some cases 1 X masks can be directely generated at very high speeds.

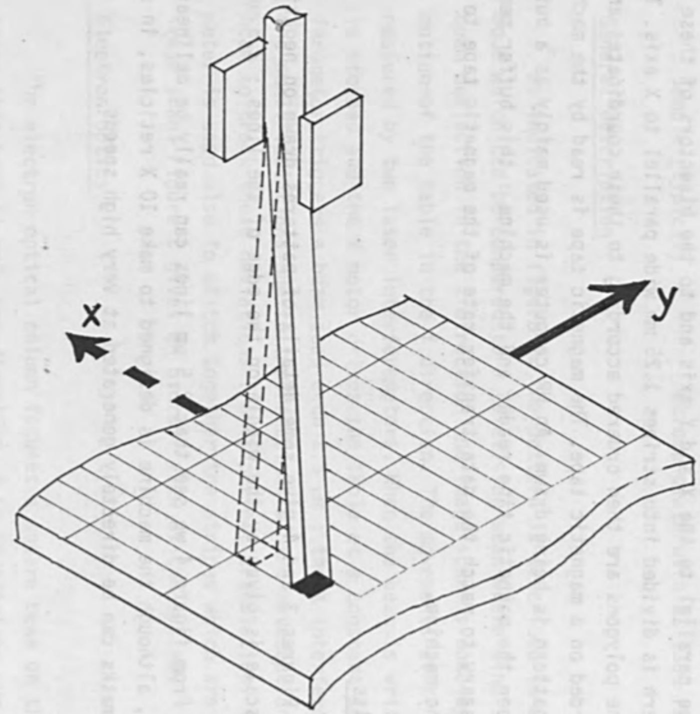


Figure 1 : Pattern is decomposed into adjacent vertical stripes 1.25 mm (9 bits) wide by 130 mm long.

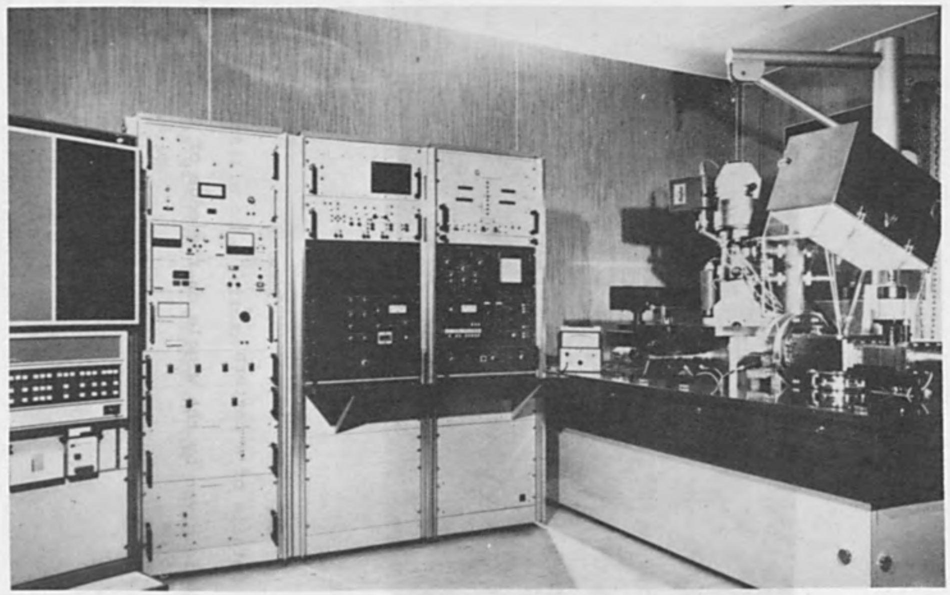


Figure 2 : overall view of the system

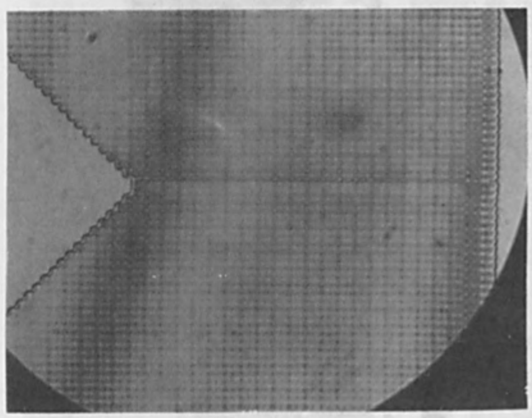


Fig. 3. Details of a pattern

The scale is given by the wiggle on the edge of the lines : 2.5 μm

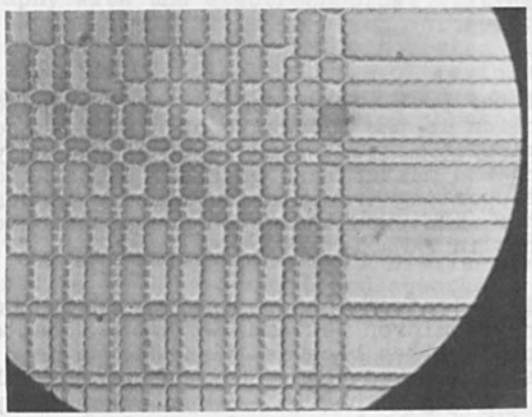


Fig. 4. Details of a pattern

The scale is given by the wiggle on the edge of the lines : 2.5 µm

TRANSMISSION MASTERS FOR DEMAGNIFYING ELECTRON PROJECTION

A. Meyer, A. Politycki, E. Fuchs

Siemens AG, Forschungslaboratorien München
Balanstraße 73, 8000 München 80, Germany

One of the problems of demagnifying electron projection is the necessity of a self-supporting transmission master. This paper shows a way of solving the master mask problem. Preparation of self-supporting transmission masters is described together with experiments to produce submicron patterns in PMMA layers on silicon.

Demagnifying electron projection (1 - 4) is a method of fabricating submicron patterns with the advantage of short exposure times because the whole pattern area is exposed using a flood beam. This contrasts with the electron beam writing method (5 - 9) in which the pattern is produced with a spot beam or shaped electron beam. One of the problems presented by demagnifying electron projection is the necessity of a self-supporting original master. Depending on the given pattern features, this transmission master must be electron opaque in some areas and electron transparent in others to enable the generation of patterns containing disconnected areas.

Our concept for solving this problem is as follows: It is well known in electron beam lithography that during exposure scattering occurs due to interaction between the electron beam, the resist and the substrate, resulting in a certain edge shift of the pattern (10 - 12). If, for example, a 10 μm wide line in a stencil is projected onto the recording plane with a demagnification of 1:10 and the edge shift due to scattering is 0.05 μm, the linewidth in the resist pattern after development will be 0.9 μm. Under these conditions a 1 μm wide line in a stencil should not appear in the resist pattern. The consequence of this applied to a self-supporting transmission master is: If the disconnected electron opaque areas are supported and locally fixed by a very fine grid, interaction during exposure causes the demagnified image of the supporting grid to be exposed by scattered electrons. After development, the demagnified image of the supporting grid does not appear in the resist pattern if the grid traverses are small enough, whereas the demagnified active pattern of the microstructure is correctly reproduced and appears with only a slight edge shift.

We have realized such self-supporting transmission masters with a fine supporting grid by combining UV photolithography with the electrochemical plating technique. Fig. 1 shows an outline of the technology: A thin film metallized glass substrate is spin coated with positive photoresist. During exposure with ultraviolet light a chromium-on-glass mask with the negative pattern of the transmission master contacts the photoresist coating of the substrate. Those parts of the substrate that are free from photoresist after development are electroplated to a thickness not exceeding that of the photoresist. After dissolving the photoresist, the electroplated transmission master is removed from the substrate by separately etching the thin film electrode. A similar method was already used in 1959 by Boersch and coworkers (13) to prepare fine nickel grids for electron microscopic purposes.

Fig. 2 shows on the left a scanning electron micrograph of the resist pattern on a metallized glass substrate. This pattern is used to prepare the fine supporting grid by plating up the spacings between the resist bumps. The picture on the right shows a SEM micrograph of the metal supporting grid. The traverses of the grid are 1 µm wide, the mesh size is 4 µm. Since the walls of the resist bumps are almost vertical the profile of the traverses appears nearly without slope. The rounded crossings of the traverses can already be seen on the chromium-on-glass mask and are reproduced during ultra violet light exposure.

As can be seen in Fig. 3, it is possible to stretch the as prepared transmission master over a brass ring and affix it. The inner diameter of the ring is 24 mm. Fig. 4 shows the transmitted light micrograph of a self-supporting master. The meander and the contact pads of a test pattern are supported by the 4 µm-mesh grid with traverses 1 µm in width. The grid thickness is approximately 1 µm.

To test our transmission masters in damagnifying electron projection we used a provisional electron projection system similar to that described by Koops (2). It

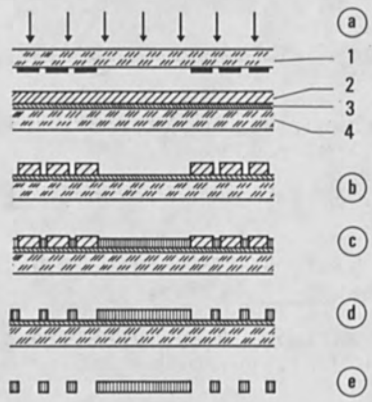
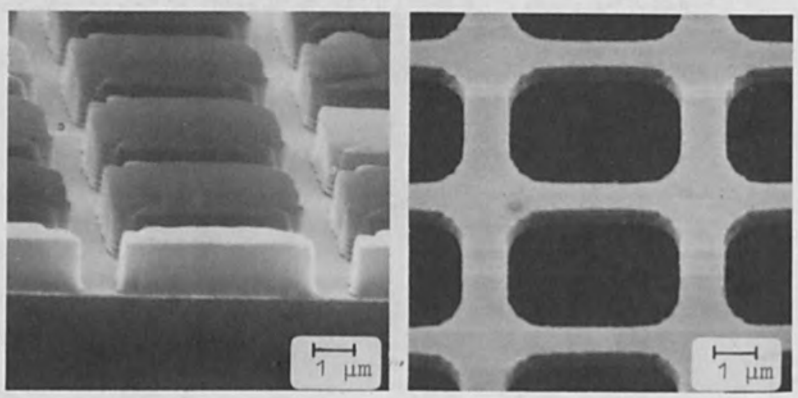


Fig. 1: Preparation of self-supporting transmission masters: (a) UV exposure, (b) resist pattern, (c) electroplating, (d) resist-removal, (e) separation. 1 mask, 2 resist, 3 thin film electrode, 4 glass substrate



Resist pattern Metal grid

Fig. 2: Supporting grid of transmission masters

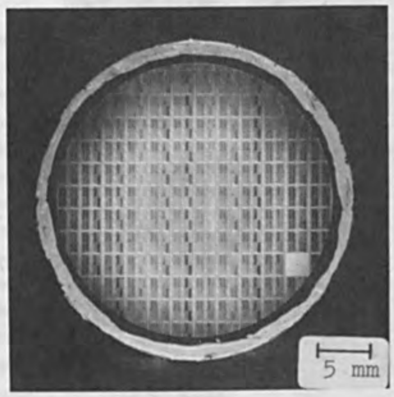


Fig.3: Self-supporting transmission master, stretched over a brass ring

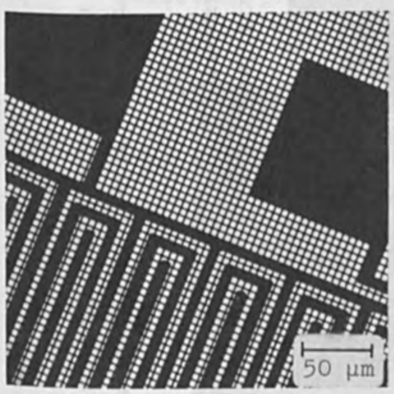


Fig. 4: Transmission master with meander test pattern

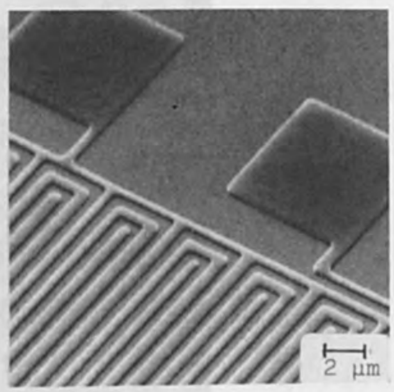
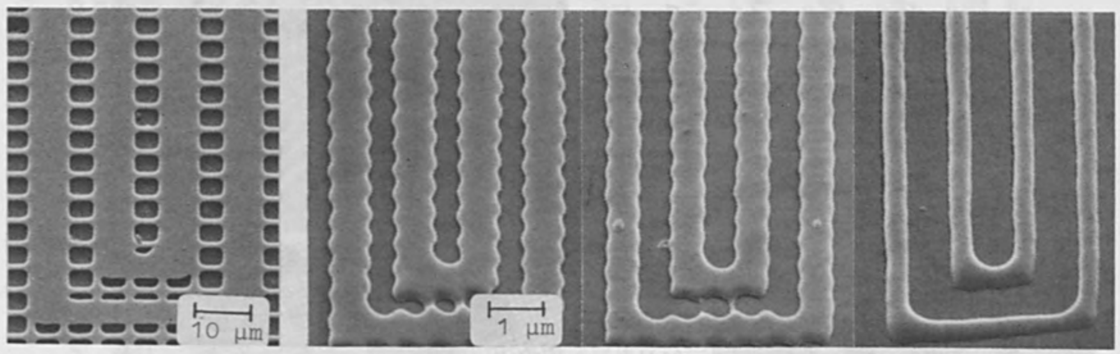


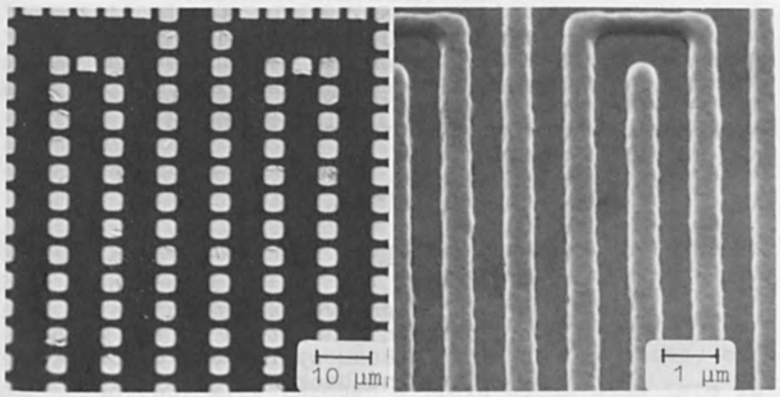
Fig. 6: PMMA resist pattern on silicon produced by demagnifying electron projection



Transmission master

Resist pattern on silicon increasing dose

Fig. 5: Demagnifying electron projection with self-supporting transmission master



Transmission master Resist pattern on silicon

Fig. 7: Submicron pattern by demagnifying electron projection

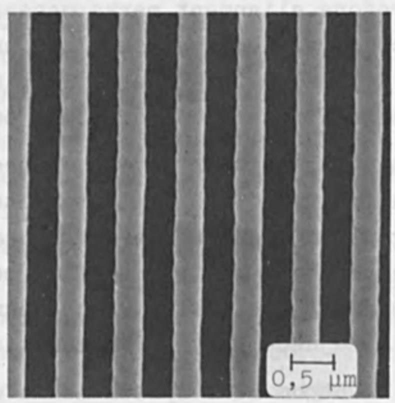


Fig. 8: Submicron pattern by demagnifying electron projection

using two different electron detectors:

- a conventional secondary electron detector especially sensitive for the electrons of low energies.
- and a silicon diode detector for backscattered electrons with higher energies.

In every case the primary energy has been 20 KeV, the slice was covered with 0.3 µm PMMA-resist and the width of the structures was 1 µm and 2 µm respectively.

Fig. 5 shows these signals, measured with a secondary electron detector for a substrate covered with 0,3 µm silicon dioxide and resist, for a substrate with 0,3 µm polysilicon and resist and for a substrate with poly-silicon, silicon dioxide and resist. Even the last one gives useful signals for detection and computing.

The symmetrical silicon-diode detector for backscattered electrons consisting of four parts has been used for the signals of the same structures in fig. 6. As one can see, these signals are better than those shown in the last figure.

Therefore one should use a backscattered electron detector for alignment mark detection and getting a processable signal which gives the additional advantage to work with a detector which is free of any potential that influences the electron beam.

The example of an n-silicon-gate process for the use of those grooved alignment marks has shown that narrow structures give very good electron signals, which are insignificantly influenced by the processing steps.

For special purpose, e.g. deminuation of the proximity effect [4], it might be necessary to decrease the primary energy. The influence of this on the detected signals is shown in fig. 7, for 10 and 20 KeV. The decrease of sensitivity for lower energies of the backscattered electron detector results in decreased back-scattered signals while the higher secondary electron efficiency leads to better secondary signals.

Fig. 8 shows a graph of the signal to noise ratio versus primary electron energy for alignment marks with and without V-grooved substrate etching. In every case the signal quality of the substrate etched structures is much better and the alternative which detector is to use for the best signal depends on the energy needed.

It has been shown in this paper that V-grooved alignment marks for a semiconductor process using electron beam technology give good and exact electron signals independent of the type of layer sandwich and which are not affected by the process but are compatible to it.

Acknowledgement

The authors express their appreciation to Mrs. Sylvia Henkel for providing the structures and helpfull assistance during the measurements.

References

[1] Dennard, R.H. et al., Design of Micron MOS-switching Devices, IEDM Washington D.C. Dec. 1972

[2] Wolff, E.D. et al. Composition and Detection of alignment marks for electron beam lithography J. Vac. Sci. Technol., 12 (1975) 6, pp 1266-1270

[3] Chang, T.H.P., An Electron Detection System for Microfabrication Registration J. Vac. Sci. Technol. 10 (1973), p 1129

[4] Zeitler, H.-U., Friedrich, H. Test Patterns and MOS-Transistors with Short Channel Lengths Fabricated by Electron Beam Lithography, Conf. on Micro Electron Beam Technology for Fabrication Recording and Dynamic Insp., Cambridge, March 1975.

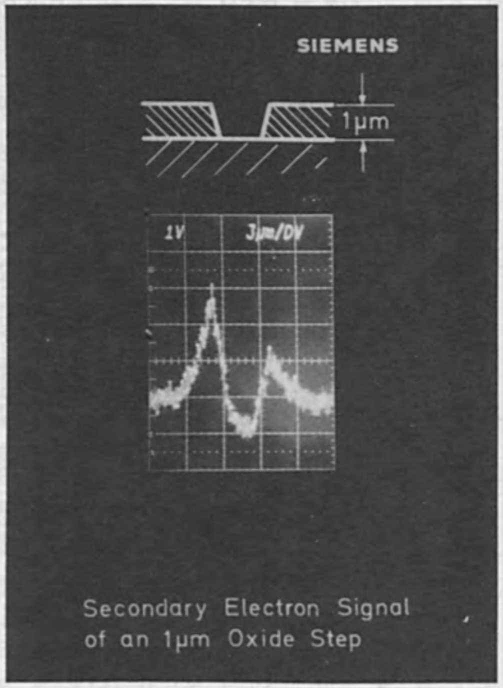

SIEMENS
1µm
1V
3µm/DV
Secondary Electron Signal
of an 1µm Oxide Step

Fig. 1

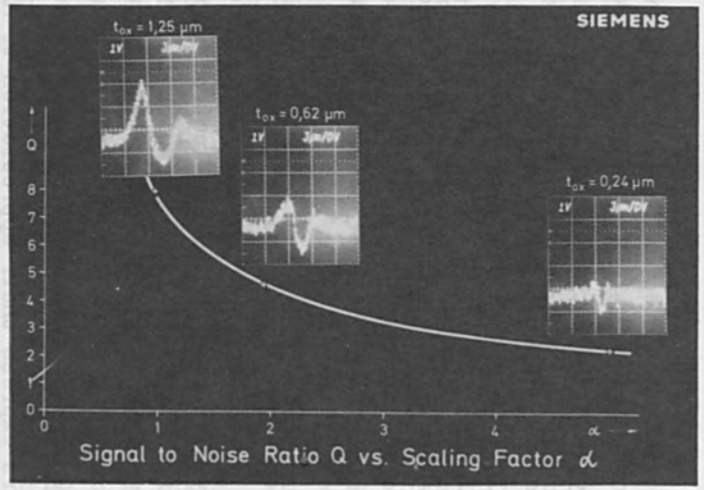
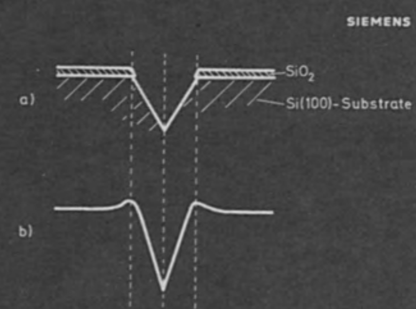
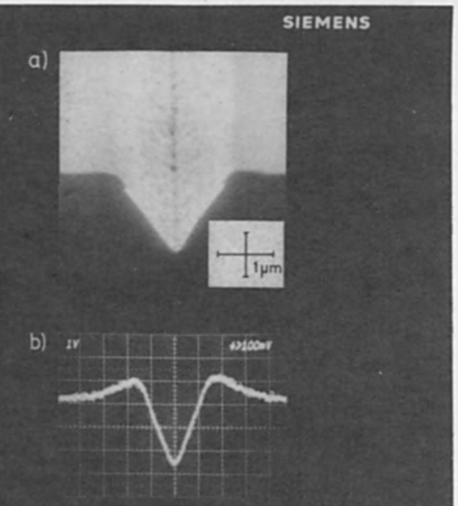

SIEMENS
t_{ox} = 1,25 µm
t_{ox} = 0,62 µm
t_{ox} = 0,24 µm
Q
α
Signal to Noise Ratio Q vs. Scaling Factor α

Fig. 2



Principal Cross Section of V-Grooved Alignment Mark (a) and the Belonging Backscattered Electron Signal (b)

Fig. 3



SEM-Micrograph of V-Grooved Alignment Mark (a) and the Belonging Backscattered Electron Signal (b)

Fig. 4

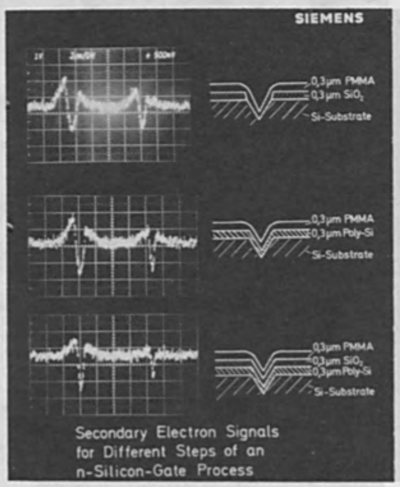

SIEMENS
0,3µm PMMA
0,3µm SiO_2
Si-Substrate
0,3µm PMMA
0,3µm Poly-Si
Si-Substrate
0,3µm PMMA
0,3µm SiO_2
0,3µm Poly-Si
Si-Substrate
Secondary Electron Signals
for Different Steps of an
n-Silicon-Gate Process

Fig. 5

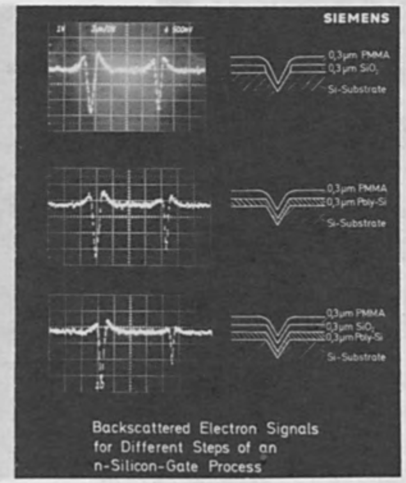

SIEMENS
0,3µm PMMA
0,3µm SiO_2
Si-Substrate
0,3µm PMMA
0,3µm Poly-Si
Si-Substrate
0,3µm PMMA
0,3µm SiO_2
0,3µm Poly-Si
Si-Substrate
Backscattered Electron Signals
for Different Steps of an
n-Silicon-Gate Process

Fig. 6

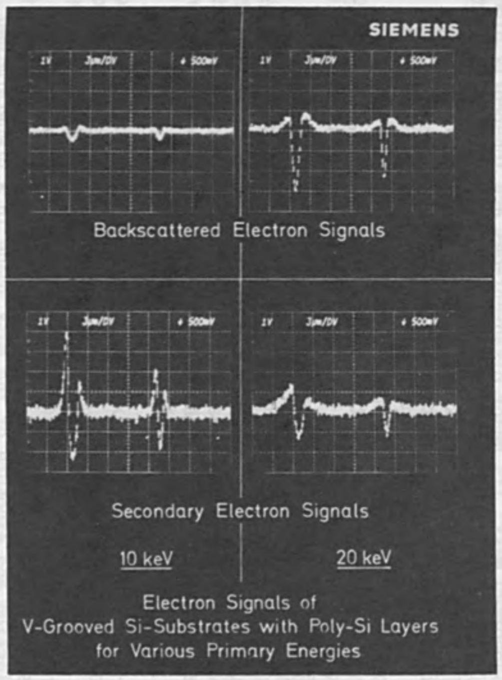

SIEMENS
1V
3µm/DIV
500mV
Backscattered Electron Signals
Secondary Electron Signals
10 keV
20 keV
Electron Signals of
V-Grooved Si-Substrates with Poly-Si Layers
for Various Primary Energies

Fig. 7

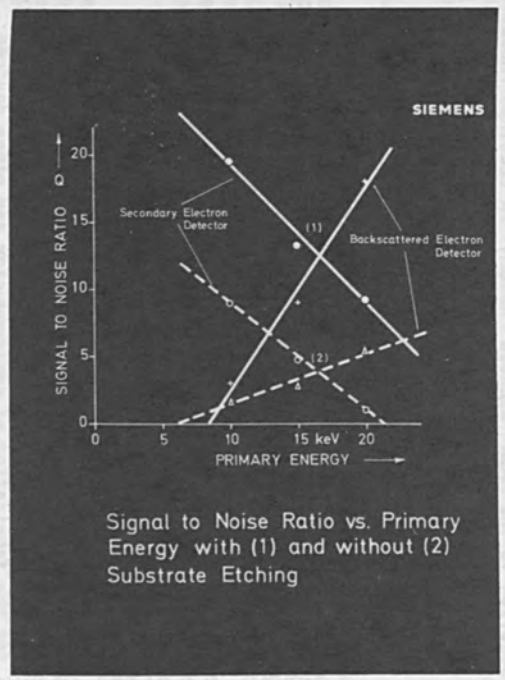

SIEMENS
SIGNAL TO NOISE RATIO Q
Secondary Electron Detector
Backscattered Electron Detector
(1)
(2)
PRIMARY ENERGY
0
5
10
15 keV
20
Signal to Noise Ratio vs. Primary
Energy with (1) and without (2)
Substrate Etching

Fig. 8

A SOLUTION TO THE MASK 'STENCIL PROBLEM' IN ELECTRON PROJECTION MICROFABRICATION

M. B. Heritage, P. E. Stuckert, and V. DiMilia
IBM Thomas J. Watson Research Center
Post Office Box 218
Yorktown Heights, New York 10598

Electron beam microfabrication systems which project a demagnified image of a self-supporting foil mask suffer from the 'stencil problem' which occurs when a desired image requires an electron-opaque area which is completely surrounded by an electron transparent area. The technique presented solves the 'stencil problem' with a mask comprised of elementary apertures. The exposure image is synthesized by multiple deflection and exposure of this mask image. Applications, examples, and limitations of this technique are discussed.

Introduction

Electron projection systems provide a high-volume low-cost exposure technique for the fabrication of semiconductor and other devices. Projection systems achieve high throughput by exposing all points in an image field simultaneously and have evolved in two distinct forms: the 1:1 cathode imaging system[1] and the demagnifying foil mask system,[2] each with its inherent advantages and disadvantages.[3,4] The major advantage of the cathode imaging system is its more readily fabricated cathode-supported mask; its major disadvantages are inherently higher image distortion and lower resolution. Demagnifying foil mask systems, conversely, have lower image distortion and higher resolution, but present formidable mask fabrication difficulties.

An obvious difficulty encountered in the fabrication of self-supporting foil masks is the 'stencil' problem, that is, the fabrication of a mask which will provide an unexposed region which is completely surrounded by an exposed region. Closely related to the 'stencil' problem is the difficulty encountered in the fabrication of a foil mask for exposing an array of long closely-spaced parallel lines, a common circuit requirement. Electrons interact so strongly with matter that the equivalent of the glass substrate of photon optics does not exist so that either type of mask, fabricated in a straightforward manner, presents mask stability problems.

The four step deflection/exposure technique which is described in the following section employs a foil mask which contains a 'transformed' image of the desired exposure image. This 'transformed' mask image is a pattern of properly located and properly dimensioned rectangular apertures from which the desired exposure image is synthesized by four-fold deflection and exposure of the mask image in a predetermined pattern. (It will become obvious that some exposure patterns, e.g., an array of parallel lines, would require only two-fold deflection and exposure. Since such special cases are subsumed by four-fold deflection and exposure, the remainder of this paper will deal with the four-fold technique only.)

The Four Step Deflection/Exposure Technique

The four step deflection/exposure technique employs the multiple exposure of elementary images which are produced by patterns of elementary mask apertures. For simplicity, the following description will be confined to rectilinear images whose dimensions are even integer multiples of the elementary aperture shown in Figure 1(a); the synthesis of rectilinear images in which one or both dimensions are not even integer multiples of this elementary aperture is described in the Appendix.

Figure 1 illustrates the four step deflection/exposure technique for one elementary mask aperture. Assuming, for descriptive purposes, a demagnification of x1, Figure 1(a) shows an elementary mask aperture with dimensions S/2 x S/2. Region 1 in Figure 1(b) is exposed in the first of four exposures. Deflection of the image by -S/2 in the Y direction and a second exposure causes region 2 to be exposed. Similar deflections and exposures at positions 3 and 4 yield the elementary exposed region of Figure 1(b) whose dimensions are S x S, as indicated.

Any desired rectilinear exposed region with dimensions NS x MS, where N and M are integers, can be synthesized by appropriate patterns of elementary (S/2 x S/2) mask apertures spaced center-to-center in X and Y on a pitch equal to S. Figure 2 shows several such patterns and their corresponding exposed regions:

- Apertures A_A are two elementary apertures which are the same as the aperture of Figure 1(a); E_A are their corresponding exposed regions.

- A_B is the aperture pattern for the exposure of E_B, two minimum width lines separated by an unexposed region of the same width. A conventional mask for this exposure would require a long cantilevered and, otherwise, unsupported tongue to produce the unexposed region T.

- A_C is the aperture pattern for the exposure of region E_C whose dimensions are 3S x 7S.

- Finally, A_D is the aperture pattern for producing an unexposed region, U, which is completely surrounded by an exposed region, E_D. Again, the mask region A_D is well-supported and an exposure which is not possible with a conventional foil mask is readily achieved.

The four step deflection/exposure technique described above permits the exposure of patterns which would be impossible or, because of mask instability or fragility, very difficult with a conventional mask. Regardless of the exposure pattern(s) required, the mask required for the four step deflection/exposure technique is relatively rugged. For the extreme case of 100% exposure of a region the mask is 75% mask material and only 25% apertures with all apertures evenly spaced in both dimensions.

The technique described above has the following limitations:

- If regions of multiple exposure cannot be tolerated, the technique can only be used for the exposure of rectilinear patterns whose dimensions are integer multiples of S, the dimensions of the elementary exposed region shown in Figure 1(b). The Appendix describes a mask design algorithm which, for rectilinear patterns of arbitrary dimensions, minimizes the extent of multiple exposure.

- For a minimum line width and a given demagnification factor, the mask aperture size and its tolerances are one-half those of a conventional mask.

- Exposure time is quadrupled. This is not a serious disadvantage because (a) exposure time in demagnifying projection systems is a performance parameter which is easily improved and (b) the times required for mechanical translation and registration frequently limit throughput.

Experimental System

The feasibility of the four step deflection/exposure technique has been demonstrated on a system which has been described previously.[3] This system provides two modes of operation, registration and exposure, and only relatively minor modifications were required to permit use of the existing alignment coils for:

- Production of the image shifts required by the four step deflection/exposure technique during exposure and

- calibration of the image shift magnitudes with the system in the registration mode of operation.

As illustrated in Figure 3, operation in the exposure mode is typical for demagnifying electron projection systems. After suitable collimation by a three lens condenser system, 20 kV electrons illuminate the entire mask. Illumination is blanked between exposure steps with the electrostatic blanking plates immediately beneath the electron gun.

Electrons which pass through mask apertures are subsequently imaged at the wafer plane with a x10 demagnification factor by the two magnetic projection lenses. A circular registration aperture and the orthogonal magnetic alignment coils are located in the plane of the electron source image, S". As was indicated above, in addition to their use in the registration process, these coils also provide the orthogonal image shifts required by the four step deflection/exposure technique. Since the image shifts required are small, drive requirements for these coils are modest and stepping may be done rapidly and accurately. The wafer is mounted on a conventional X-Y table below the second projection lens.

The mask is a square self-supporting metal foil which is 30 mm on each side and 5 μm thick. The elementary square apertures of Figure 1(a) are 5 μm on each side and are spaced on a pitch of 10 μm in patterns which will produce the desired exposure image. With the system demagnification factor of x10, the field size is 3 mm x 3 mm and the minimum line width, S in Figure 1(b), is 1 μm.

Calibration of the drives to the orthogonal image shifting coils is necessary since shift accuracies of the order of 0.05 μm are required. In the calibration mode of operation, as illustrated in Figure 4, the condenser system operation is altered so that the system operates as a large field scanning electron microscope. The demagnification of the source image at S' (Fig. 3) is increased slightly and the power of the third condenser lens is correspondingly increased to focus S' at the mask plane with a magnification of approximately unity. The probe thus formed is positioned and raster scanned across selected calibration apertures in the mask by the deflection coils which, as shown in Figure 4, are mounted within the third condenser lens. These coils are located at a position which is conjugate to the registration aperture. Operation of the projection optics is unchanged and the probe at the mask plane is imaged at the wafer plane, reduced in size by a factor of 10 and scanning 1/10 of the distance scanned on the mask. The secondary and backscattered electron detectors produce a video signal which is displayed on a CRT whose X and Y signals are the same as the signals to the deflection coils. The display is, thus, the same as that of a conventional scanning electron microscope and contains two superimposed images: the transmission image of the mask and the image of any structure on the wafer. Since the transmission image of the mask is reduced by a factor of 10, mask features and wafer features are both displayed at the same scale.

Since the superimposed mask and wafer images are at the same scale and excitation of the shift coils between the projection lenses will cause a displacement of the two images, calibration of the shift magnitude and direction referenced to the mask is possible. The preferred procedure for making this calibration is described in the following section.

Calibration Procedure

Magnetic image shift calibration is required in magnitude and direction for both the X and Y directions. Since the procedure is the same for both directions, the following description will be confined to calibration in the X direction.

For a mask with 5 μm square apertures and a demagnification factor of x10 the shift magnitudes and tolerances are:

— Magnitude: 0.5 μm ± 0.05 μm

— Direction: 0° ± 6°

These tolerances correspond to shift errors which are one-tenth of the aperture image size.

The calibration sequence is as follows:

— The wafer stage is moved so that the probe scans a featureless portion of the wafer holder which provides a good signal to the secondary electron detectors.

— The mask image is scanned and viewed as described in the previous section with electronically generated index lines superimposed, yielding the display shown in Figure 5a.

— The parallel vertical index lines are manually aligned with widely separated mask aperture images whose separation is accurately known.

— With the index lines fixed, a mask region which contains a relatively large aperture is scanned while a second site on the wafer holder, which contains a grid with good edge features, is moved into the field of view, yielding the display shown in Figure 5b.

— A signal applied to the shift coil between the two projection lenses causes the image of the grid feature to move relative to the vertical index lines. With the distance between the index lines known, the shift coil drive current is readily calibrated.

As indicated above, calibration in both X and Y is performed and the magnetic shifts are adjusted in magnitude, and by electronic rotation, in direction.

Experimental Results

An example of the use of the four step deflection/exposure technique is shown in Figures 5, 6 and 7. Figure 5a shows a mask image containing features similar to those of Figure 2. All mask apertures are nominally 5 μm square and, within each of the several structures, are spaced on a 10 μm pitch. Among others, aperture patterns for the following are included:

- an isolated elementary exposure area,
- two minimum width lines separated by an unexposed area of equal width,
- two minimum width lines separated by an unexposed area equal to one-half a minimum line width, and
- an unexposed area which is completely surrounded by an exposed area.

Figure 6 is an electron micrograph of this last area showing the apertures and the spaces in greater detail. Figure 7 shows a resist pattern after exposure by the four step deflection/exposure technique. The developed resist shows lines of 1.1 μm width. Some scalloping and a certain degree of rounding of the corners may be observed. These imperfections are relatively minor and may be attributed to the combination of mask fabrication errors and image spreading due to electrons which are backscattered from the substrate. Inspection of Fig. 6 shows the size of the mask apertures as slightly smaller than the intervening spaces. Consequently in the four step exposure the aperture images do not abut. Despite this, complete exposure is achieved, except for a small degree of edge scalloping, by electron backscattering from the substrate.

The results shown in Figure 7 demonstrate clearly that the four step deflection/exposure technique works and that it has a reasonable error tolerance. This demonstration of reasonable error tolerance tends to confirm that the exposure of areas whose dimensions are not integer multiples of the elementary exposure, as described in the Appendix is feasible.

Conclusion

A technique which overcomes the "stencil" related limitations of a self-supporting mask in a demagnifying electron projection system has been described. The technique has been verified experimentally and the calibration procedures have been outlined.

The major limitation of the technique is the requirement that exposed regions must be rectilinear. Another limitation, the severity of which is not yet resolved, is the effect of double exposures when regions whose dimensions are not integer multiples of the elementary exposure area are required. Together, the reasonable error tolerance of the technique and a procedure for minimizing the extent of double exposures may render the technique suitable for the exposure of rectilinear areas of arbitrary size.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the work of G. Wardly, L. Romankiw, E. Castellani, P. McCaffrey and A. Pfeiffer in designing and fabricating the masks used in this work. We also acknowledge the many contributions of R. Scott in building the system.

REFERENCES

1. J. P. Scott, IEEE Trans, Electron Devices, 409, 22, (1975)

2. H. Koops, J. Vac. Sci. Technol., 909, 10, (1973)

3. M. B. Heritage, J. Vac. Sci. Technol., 12, 1135, (1975)

4. G. A. Wardly, IEEE Trans, Electron Devices, 414 22, (1975)

APPENDIX

Rectilinear regions whose dimensions (L) are not integer multiples of S (see Figure 1(b)) cannot be exposed with the four step deflection/exposure technique without regions of multiple exposure. The mask design algorithm described below uses the same image deflection magnitude (S/2) and the same fourfold exposure sequence as are described in the body of this paper and yields mask aperture dimensions (A) and spacings (M). This algorithm is considered optimum because it provides a mask which yields an exposed region in which:

- all regions of quadruple exposure lie totally within the exposed region,

- the widths of all regions of double exposure along the edges of the exposed region are equal,

- the regions of double exposure along the edges of the exposed region are equally spaced along the edges, and

- the sum of the widths of the regions of double exposure is the minimum possible.

With reference to Figures A-1 through A-3, the following algorithm is applied separately to each dimension of the region to be exposed.

(a) If: $\left[\left(\frac{L}{S}\right) - \lfloor\left(\frac{L}{S}\right)\right] \leq 0.5$

Then:

$$N = \lfloor\left(\frac{L}{S}\right)$$

$$M = \frac{S}{2}$$

$$A = \frac{S}{2} + \frac{L - NS}{N}$$

$$O = A - M$$

$$O_t = N(A - M)$$

(b) If: $\left[\left(\frac{L}{S}\right) - \lceil\left(\frac{L}{S}\right)\right] > 0.5$

Then:

$$N = \lceil\left(\frac{L}{S}\right)$$

$$M = \frac{S}{2} + \frac{L - NS}{N - 1}$$

$$A = \frac{S}{2}$$

$$O = A - M$$

$$O_t = (N - 1)(A - M)$$

Where:

- S is the minimum line width as described in the body of the paper and shown in Figure 1,
- L ($\geq$ S and in units of S) is the length of a side of the rectangular region to be exposed,
- N is the number of apertures required in the direction of L,
- M is the width of the mask between adjacent apertures,
- A is the width of each aperture,
- O is the width of each of the double exposed regions along the edges,
- O_t is the sum of the widths of all the double exposed regions along the edges,
- $\lfloor$ indicates rounding down to the next lower integer, and
- $\lceil$ indicates rounding up to the next higher integer.

For magnitudes of L which are integer multiples of S this algorithm yields masks which are identical to those presented in the body of the paper.

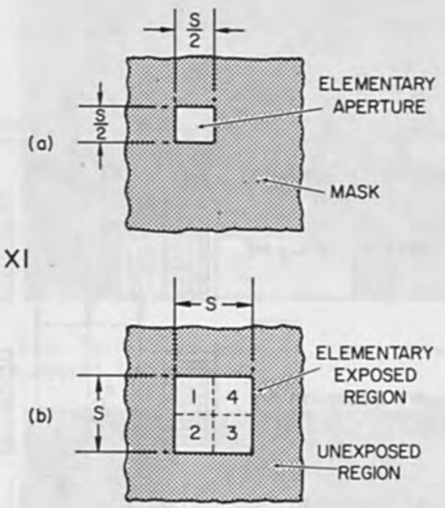
Figures A-1 through A-3, for the dimensions shown in parenthesis, also show an example of the application of the mask design algorithm. Figure A-1 shows a desired exposure region which is 3.3s x 1.8s. With reference to Figure A-2, for the X dimension and from (a), above:

$N_x = 3$
$M_x = 0.5s$
$A_x = 0.6s$
$O_x = 0.1s$
$O_t = 0.3s$

For the Y dimension and from (b), above:

$N_y = 2$
$M_y = 0.3s$
$A_y = 0.5s$
$O_y = 0.2s$
$O_t = 0.2s$

Figure A-3 shows the exposure region at the end of the four step deflection/exposure process.



XI

Fig. 1(a) Elementary mask aperture
(b) Exposure after four step deflection/exposure technique

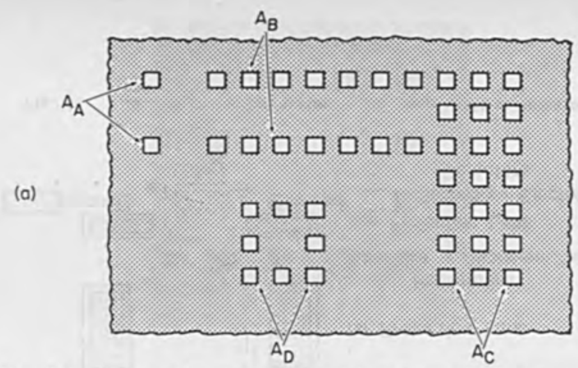
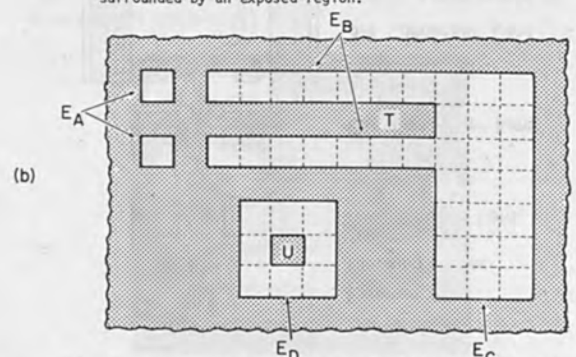


Fig. 2(a) Mask apertures to expose, a minimum element S x S, a region of 3S x 7S, two minimum width lines separated by an unexposed region of the same width, an unexposed region surrounded by an exposed region.



(b) The exposure pattern resulting from (a) after a four step deflection/exposure sequence.

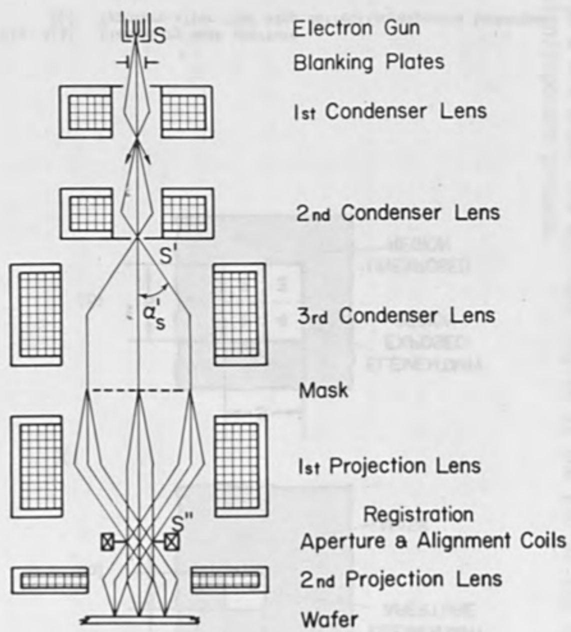


Fig. 3 The 10x electron beam projection system.

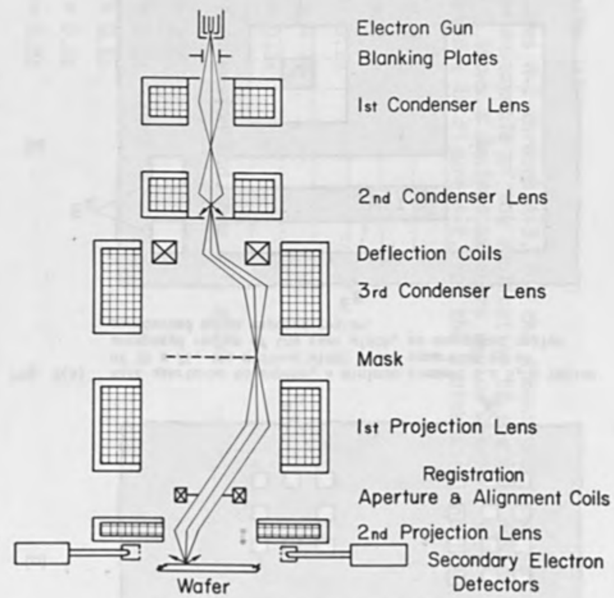


Fig. 4 10x E/B projection system in the image shift calibration mode.

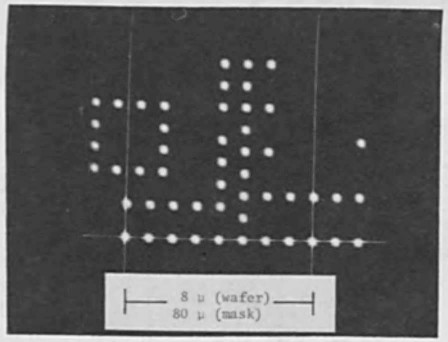


Fig. 5a. Video image of a mask area with superimposed index lines.

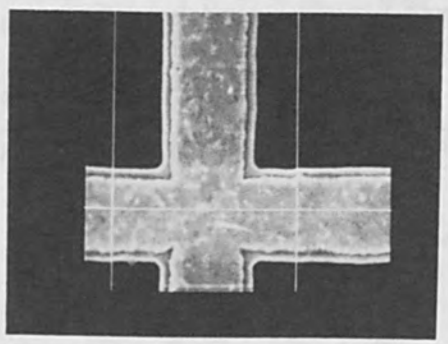


Fig. 5b. Portion of a grid in wafer plane viewed through an open area of the mask superimposed index lines.

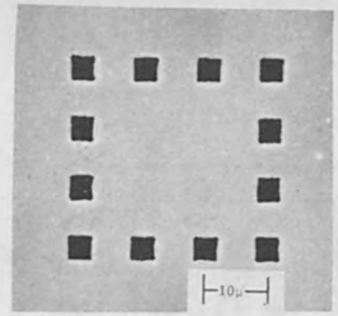


Fig. 6. Scanning electron micrograph of a mask showing an aperture array designed to produce an unexposed region totally surrounded by an exposed region.

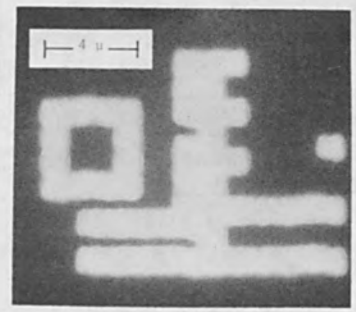


Fig. 7. A developed resist pattern exposed using the mask of Fig. 5 (a) and the four-step deflection/exposure technique.

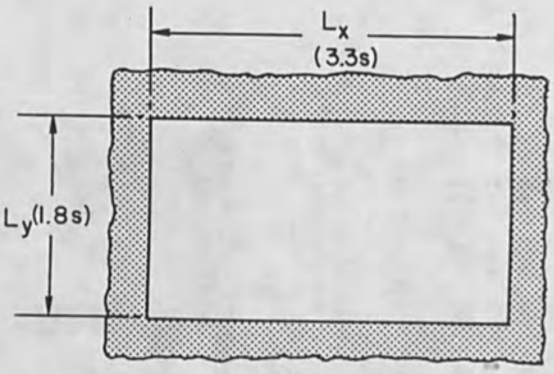


Fig. A1 Desired exposure region 3.3s x 1.8s

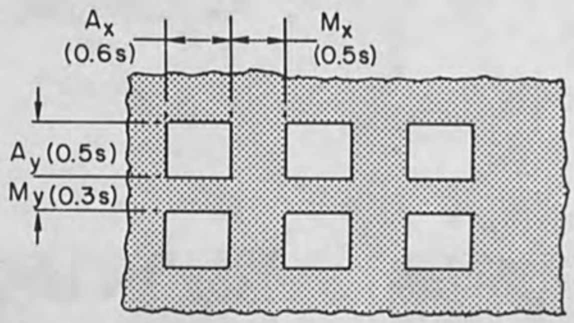


Fig. A2 Aperture array required to give exposure region 3.3s x 1.8s with minimum overlap.

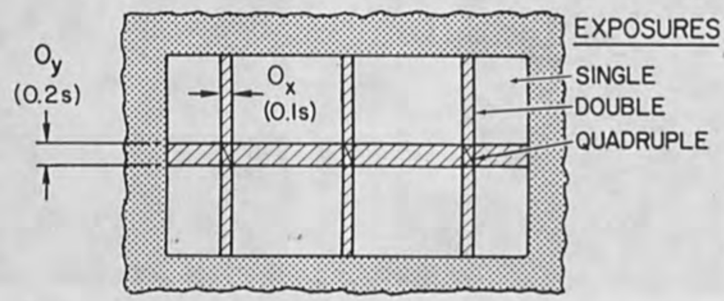


Fig. A3 Exposure region 3.3s x 1.8s after the four step deflection/exposure technique.

AUTOMATION OF VECTOR SCAN I ELECTRON BEAM LITHOGRAPHIC SYSTEM

A. D. Wilson, A. Kern, A. J. Speth, A. M. Patlach,

P. R. Jaskar, T. W. Studwell and W. L. Keller

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Vector Scan I, an automatic electron beam lithographic system, is described from an automation point of view. We divide the presentation into a number of facets: computer system, I/O channels, scanning interface unit, pattern generator, workstage controller and servo system, programmable signal enhancement registration unit, analog field correction units, operating software and console terminal. Each facet of the automation system is treated with concern for design philosophies, alternatives, performance and flexibility.

Introduction

There are several independent groups developing some form of an automatic e-beam system.[1-4] The intent of this paper is not to give a general review of the subject or to compare the many routes that have been followed in developing a precision micron and submicron lithographic system, but rather to summarize the work that we have done in the development of Vector Scan I.

In developing Vector Scan I, the authors focused on several objectives during the course of the program: 1) extension of an earlier system, 2) increased thruput. 3) modular design, 4) stand alone operation, 5) simplicity, 6) consideration of human factors, 7) maintainability, 8) ease of system debugging and 9) packaging. This paper is divided into several sections, notably: system architecture, I/O channels, workstage servo system, pattern generator, automatic registration, correction units and software system. The automated operations in Vector Scan I are pattern writing, workstage movement, registration and proximity effect compensation. However, there are a number of other necessary functions served by the control system such as operator interaction with the system via a console terminal located on the e-beam system console, management of data set files that include pattern information, wafer layout and chip parameters such as size, registration mark location, etc. Because this paper is one part of a three part presentation on Vector Scan I, considerable useful information available in the paper by Chang[5], et.al., will not be duplicated here.

System Architecture

The control system used to automate Vector Scan I is shown in

Figure 1. The functions under control are pattern-generation, workstage control, beam-current regulation, auto-registration, pattern field: offset, attenuation and rotation; operator console terminal and laser interferometer (under development). Not shown in Figure 1 are the many analog and other connections to the control units. What is shown is the digital control data flow and the system is shown as it is cabled. All of the above units have been designed and built in-house for the explicit use in Vector Scan I and thus they are tailored to do a specific task. Some use is made of commercially available subsystems such as the laser in the interferometer. The system hardware is connected to the IBM 1130 computing system via two I/O channels: OEM and SAC. The IBM 1130 system is configured with a card reader/punch, a medium speed chain printer (1403), a telecommunication interface, three 500 K word capacity magnetic memory disk drives, a system multiplexor (shown as an ellipse with an enclosed cross) and a CPU having a 2.5 usec memory cycle time with 32 K of core memory. This system has a well-supported IBM disk monitor system and has been used extensively as a process controller. It is therefore well suited to act as a controller for an experimental e-beam system. The 1130 CPU unit has an operator keyboard and a Selectric type of output printer.

Most devices which are attached to the computer, except the pattern generator, require small amounts of data infrequently and at a slow data rate. Therefore, only the pattern generator was placed on the high speed SAC channel. The OEM devices and the pattern generator have a number of operator controls that permit the operator to use the device in a manual or automatic mode of operation. Should the operator switch a device from one mode to another, this information is presented to the CPU the next time it services the device to "warn" of a possible change in the system state. All slow data transfers are channeled through the OEM scanner which will be described in a later section.

The three magnetic disks are used as follows: one for programs and "cold start" parameters and the other two for pattern data sets. Approximately 10 to 20 LSI device levels such as those of an 8 K bit FET or bubble memory chip can be stored on each disk. Data is read from disk to CPU via a cycle steal I/O operation. The computer and control system are designed to operate in a stand alone mode without assistance from a host system. Pattern and program information is, however, transmitted to the CPU via a telecommunication interface operating at 2000 baud.

I/O Channels

Two data transfer channels are used in the control system for Vector Scan I. The first is a software controlled channel mentioned above as the OEM channel. This channel has four 16-bit wide registers: data-out, data-in, control-out and status-in as well as 12 interrupts.

This channel is terminated in a scanning interface unit shown in Figure 2 where the small numbers shown by each signal line are the number of bits represented by a single line. For example, the interface is controlled by the CPU with a 4 bit wide control word indicated by the box in the uppermost left corner of Figure 2. Data is transferred to and from a device such as the attenuator unit in groups of 4 bits. Located in each device is a universal adapter board that interfaces each device to the scanning interface unit. Devices requiring service are placed into a queue that is serviced by a software system designed to be "table" driven. The status of any device can be determined at any time by software; some operations such as read and write, however, can only be done under certain well defined conditions. A total of 16 devices can be active on the scanner at any one time. Devices are connected to the channel via four daisy-chains, 4 devices long. Eight ports are currently in use. The speed of the OEM channel is dependent upon the implementation of the software to drive the channel. Because it is a very flexible part of the system and hence somewhat generalized the speed is sacrificed to retain its generalized nature. Data transfer rates correspond to about 150 usec per word. Transfer is done by interrupts.

The pattern generator is attached to the CPU via the Storage Access Channel or SAC. This channel permits direct access to the core memory by a cycle-steal operation. This is a case where if a program is being executed and a cycle steal request is made, the program would be halted until a specified number of data reads or writes from or to core memory is accomplished. Once the I/O operation has been initiated, no software control is required, or for that matter possible. The SAC transfer rate is that of the memory cycle-time or 2.5 usec. The pattern generator is buffered from the SAC by a small memory of 256 words. Initially the memory is loaded with pattern data and the pattern generator begins to write a pattern. When the memory is depleted by N words then a burst of data from the CPU main memory is transmitted to again fill the buffer. In this way 'nuisance' requests to the CPU for one or so words are prevented. Data can be fetched by the pattern generator from the cache buffer in about .4 usec per word.

Workstage Controller and Servo System

The X,Y and theta stage located in the e-beam system chamber is driven by three low-inertia DC servo motors. The stage is of a lead-screw type having a two mm pitch screw. It is a precision system having high accuracy and good repeatability. A servo amplifier is used to drive the stage motors. The control signals presented to the servo amplifier consist of operator-initiated jog and slew analog requests and digital increment requests that can be generated by either an operator or by the control computer. In addition, the stage-controller can be commanded to move the workstage to an absolute location specified in microns. A micro-processor stage-controller was developed to handle control of the three axis servo system, operator requests and main control computer requests and to compute the absolute location. The complete system is shown in Figure 3 as a block diagram. Ultrafine encoders (1000 line) on the motor shafts serve as position counting sensors as well as direction and velocity sensors. To minimize step-time and overshoot problems, the servo system is designed to drive the motors up to full speed in a fast ramp fashion and then to decelerate the motors in a controlled way as the final position is approached. The three axis position information is presented to the micro-processor via a multiplexor also shown in Figure 3. The absolute position of the stage is presented to the computer and the operator in units of microns by the micro-processor controller. The micro-processor controller also scans the operator console searching for depression of a key by the operator. Debouncing of the switch contacts is also done by the micro-processor.

In a step and repeat mode of operation, the stage system can be driven at a velocity of 1 cm/sec. At the end of the chip step, some time is required for the stage to settle down. Because the Y axis slide is carried upon the X axis slide, more settling time is required for the X motion step than for the Y motion step due in part to the extra mass carried by the X axis slide. This settling time is compensated for by an interferometer monitoring the stage absolute motion (currently being completed). From interferometer measurements, it has been determined that approximately 300 msec is required for the system to settle to a peak-peak vibration amplitude of 300 Å after a rapid X step to a new chip site. The chip step-time plus the settle-time is usually one second or less. In the open loop mode of operation, that is, without laser interferometer feedback compensation, the stage reaches an absolute location within about ± 5 microns for a travel of 10 cm. The stage repeats a location within a micron or less when approached from the same direction. Greater stage velocity can be attained if the accuracy requirements are relaxed. The tightness of the rider nut on the lead screw and the lubricant employed on the screw help determine the speed at which the stage can be moved and the accuracy attained.

Pattern Generator

Two classes of a pattern generator have been developed. One is a hard wired generator of somewhat limited flexibility but with little pattern-write overhead. The other is a flexible programmable unit which will be discussed here. The pattern generator system has a number of differing methods of filling-in a variety of predefined shapes. The generator also provides a programmable delay function to accommodate deflection system settle characteristics and provides a programmable exposure rate clock alterable on a shape-by-shape basis. In addition to the ability to vary the rate at which the beam steps, the ability to change the distance the beam stepped in one clock period is also a feature. Thus, the beam increment (step-size) can be changed. Proximity effect compensation is done with the programmable exposure clock and with the ability to selectively place additional scans along the boundary of the shape.[6]

The pattern generator is driven directly from the 1130 memory via the SAC and a cache memory. Pattern data is transferred from disk storage via the 1130 DMS to core memory. Transfers of data to the pattern generator are then started by a command from the CPU. The transfer is effected directly from core-memory on a "cycle-steal" basis, and this is done in bursts. After the pattern generator begins to write a pattern, the words are extracted from the cache sequentially. Requests for refill of the cache are automatically generated when the count of remaining words falls below 190. Replenishment can take place in the cache memory while data is being extracted. With typical pattern data, a burst mode is established, in which data transfer overhead is usually invisible to the rest of the system. Core-memory is divided into A and B sections with the disk loading section B while section A is being emptied by cache refill requests. When A is empty, a test is made to determine if B has been filled and if so, reading from B commences while A is replenished from disk storage.

From the cache memory, data is transferred to a set of registers called "word-sorter". In this section, the serial data stream is converted to a form in which data words can be accessed by the process controller in a semi-parallel fashion. The eleven 16-bit word sorter registers are: control, x position, y position, delta x, delta y, clock rate and delay, x and y increment, skew and altitude (2 words), blank and end of pattern, and a spare word. Each of the named registers has a specific destination in a normal instruction sequence and may be used repetitively without change. Thus, if a one-word transfer, for example, X position, provides adequate information for the next shape to be written, this is communicated to the process controller. Post processing of the desired pattern arranges the data set in such a fashion as to take advantage of the compaction scheme used by the pattern generator. (The post processing of the CAD pattern is done not on the 1130 computer but rather offline in a large fast processor.) The process controller is in turn informed of this transfer at which time the control memory is addressed. Resident with the control memory is the micro-code, the series of internal instructions needed to execute the particular shape requested by the control word. The control memory may have within it all of the shape micro-code or a writeable sequence that may be modified under control of the 1130 computer. Micro-code to generate a special pattern can be contained in the pattern data-stream and automatically loaded into the pattern generator. The algorithms, or micro-code, to generate the fill-in of a number of primitive shapes has been written and executed. Some of the predefined shapes and the methods of filling them in are shown in Figure 5. These shapes are rectangles filled-in by a spiral, bi-or uni-directional raster in either the X or Y direction, right triangles and parallelograms with one side parallel to the X axis or the Y axis. Fill-in of the latter two shapes is done with a raster. Because of the programmable nature of this system, the variety of possible shapes and their respective fill-in modes can be readily increased.

During the generation of a given shape, that is execution of a given algorithm, the process controller (see Figure 4) in conjunction with the control memory, places the data from the appropriate registers onto the data bus. Simultaneously, control information specified by the operation code selects a specific control function to be placed on the control bus. The period clock (exposure rate) provides the time base which controls the rate at which new position information is sent to the D/A converters. This clock is started and stopped in synchronism with beam blanking signals from the control bus, thereby providing the capability to produce a spot by spot dwell for writing a pattern. The period clock is directed to the stepper section, Figure 6, as selected by the micro-code and at the same time to a counter which has been loaded with dimensional information from the delta X and Y registers via the data bus. As the shape is executed, line-by-line, a line-length count is retained which is fed back to the process controller. When both X and Y dimensions are satisfied, a complete shape signal is generated causing new pattern data to be fetched from the cache memory; the process repeats until the entire pattern is finished.

The steppers that provide the four 15-bit wide position and increment signals are shown in Figure 6 and are built around ALU's (arithmetic logic unit). Because of space limitations, we will not be able to delve in any detail into the operation of the stepper section. The output of the stepper is a function of the A and B inputs to the stepper and this permits the programmable step increment discussed earlier in this paper. The flexibility of this counter arrangement also permits the generation of angle shapes such as parallelograms and triangles because the increment for stepping can be programmed. The combination of the programmable stepper and the micro-coded fill-in algorithms also permits the pattern generator to be used either with single channel D/A converters or with higher performance dual channel units; one D/A for position and one for high-speed increment fill-in.

The pattern generator system operates at a 10 MHz step clock rate; stepper operation at 18 MHz rate has been demonstrated. Usually two micro-code inctructions are required to execute a single line scan of fill-in. Between 0.1 and 0.4 usec is required to execute these instructions.

Automatic Registration

The system we have experimented with uses a digital signal enhancement technique to improve the video signal-to-noise ratio

(S/N) without resist overexposure.[7-9] Figure 7 illustrates the overall system in block form. Backscattered electrons are collected by a grounded pair of scintillators located directly above the wafer or other substrate. As the beam is scanned across the registration mark, a change in the backscattered signal is observed. A signal processor under control of a registration control unit (lower right box in Figure 7) collects data from the detector and stores a digitized summed result in its memory. The signal processor is shown in block form in Figure 8. The "raw" video is amplified or attenuated, as required, and then presented to a very high speed D/A converter. This is a 6-bit plus sign device. The digitized result is stored into a buffer memory of length 256. The buffer contents are multiplied by a constant scaling factor and added to the contents of a main memory in a location by location fashion. Because the memory pointer and the D/A converter are strobed in step with the scan of the registration mark, the signal grows according to the number of scans and the random noise grows according to the square root of the number of scans. Therefore, the S/N improves according to the square root of the number of scans. Once sufficiently adequate S/N is obtained, the "averaged" registration mark signal is read into the computer and the mark edges are determined by software. Each line of scan of the registration mark is conducted over a different area of the mark which averages out mark defects and prevents the resist layer covering the mark from becoming overexposed. Because the registration control unit is in close communication with the main computer, it is possible to make the system adaptive. That is, if the S/N is not adequate after a prescribed number of scans have been collected, then, without erasing the signal processor memory, additional scans can be collected and added to those already collected; another test of the S/N is then made to see if it is now adequate to determine the mark location within the specified tolerance level.

The present signal processor can collect data at a sample rate of 2 MHz and requires 100 usec per word readout time. For a chip with eight registration marks, an X and a Y mark located in each corner of the chip, the automatic registration system requires approximately 200 msec to average 64 scans of 128 steps in length and read the eight averaged results into the 1130 computer. This time can be reduced by a different organization of the signal processor. At 5 MHz and proper channel I/O organization the collect time is estimated to be about 15 msec. Computations of mark location, setting the correction units to new values and channel overheads require approximately another 600 msec. Thus the registration time per pass in on the order of 800 msec for the conditions specified. With optimized programming 100 msec will cover all overheads related to computations and correction unit I/O. Registration accuracy that has been obtained is on the order of ± 0.1 micron for a 2 mm chip. The analog/digital correction units used to obtain field overlays are discussed in the next section.

Correction Units

The field corrections necessary for registration are accomplished using analog correction devices shown in Figure 9. A linear transformation of the e-beam coordinate system is made with these units. Each unit is designed so that a minimum of ground loop problems is developed. This is accomplished by using electro-optical couplers to separate the general system digital ground from the sensitive analog deflection system ground. The attenuation of the field, or field-size selection, is determined by a series of pi networks, 18 in total, that are connected between the D/A converter output and the deflection amplifier input. These networks are connected using single-pole double-throw latching Hg wetted miniature relays. With the attenuator unit it is possible to obtain any field-size from full field down to about 3% of full field with a resolution of one part in 16383. A small overlay of any field-size is provided.

For the attenuator, and all other correction units, the operator or the computer can have full control of the unit. In either the operator or computer control case, the current status and actual device setting is always available to both the operator and the computer, as shown in Figure 10. Here the four possible control inputs to a device, e.g. the attenuator, are gated via an automatic 4-way switch. These inputs are: manual control (derived from a potentiometer--A/D converter combination that gives a good responsive analog feel to the operator), zero, center and CPU. The data flow is arranged so that the CPU data-in port sees the current valid data setting and thus the correction settings are well known at any one time to both the operator and the program. Use of this is made, for example, at initial set-up time where the operator first finds the rough correction unit settings to locate the registration marks and then signals the program to read the current settings and use these as the initial starting values in automatic registration.

Field offset is obtained using a 12-bit D/A converter isolated from the rest of the digital system and attenuated at the deflection amplifier input. The X and Y offsets are independently generated and isolated from each other electrically. Rotation and orthogonality field adjustments are accomplished using 10-bit CMOS multiplying D/A converters. A small amount of X deflection signal is added to the Y channel and a small amount of Y signal is subtracted from the X channel. The amount of signal added is determined by the MD/A converter setting. Again, the unit is isolated from the rest of the digital system and it also becomes the deflection system "mecca" ground point because this is where X and Y channels are mixed in terms of their grounds. Because only small rotation-orthogonality correction is required, the rotation-orthogonality unit output is attenuated at the deflection amplifier input by a factor of 100 as compared with the main deflection signal. This yields about a 1 degree correction effect.

Each of the field correction units is built as a separate modular unit, each attached to the OEM channel via the scanning interface unit discussed earlier. Front panel controls are available to the operator for use of the units in a totally manual mode or the units can be controlled completely by the control computer.

Software

The software for this system is outlined in Figure 11. The electron-beam system software is built around the IBM supplied 1130 Disk Monitor System (DMS). Added to this is an electron-beam loader which brings into operation programs necessary to "cold start" the electron beam system. There are three main functions handled by the software system: pattern or mask preparation, file management and system communication. We have called the data set written on a wafer or other substrate such as a mask-blank a "mask".

A file management system was developed that is easy to use, checks syntax of a mask set as it is stored, permits linking of several mask sets into one build-level and has tables of other information necessary for fabrication of an actual wafer. Most mask preparation is done offline on a company CAD system, however, test masks and other small mask sets can be directly generated and altered with the mask preparation facility.

The system communicator section is somewhat more complicated in that it has manifold tasks. One hardware item not mentioned before that is very closely connected with the software is the operator's electron-beam console terminal. Through this device the operator communicates with the rest of the system to select the kind of job he wishes to run, the changes that may have to be made during a wafer exposure, etc. It is a functional type of device in the sense that it has a 36 position keyboard, each key having a specific meaning. For example, should the operator wish to execute automatic registration at a particular location on the wafer (having not requested that function at the onset of the wafer exposure) he keys "auto-reg" and all programs, data, etc. required for automatic registration will be collected from disk storage and the function executed. If a problem should arise, an error message will be produced and printed on the 1130 console printer. The software system consists of approximately 12,000 lines of coding--that is, about 10 feet of punched cards--and has required very close development with the hardware section of the system.

The system communicator handles the e-beam console terminal requests and all communication from devices of the system such as the workstage controller, attenuator, etc. In addition, this part of the software generates all control information necessary to actually write a wafer. For example, for a registered exposure, the wafer is first positioned at a "target" location to grossly locate the registration marks. This may be an area outside the active wafer region. After locating the marks, and the operator indicating this to the program, the software moves the wafer to the first chip site and an automatic registration cycle is started. After the unit is registered, the software initiates the exposure of the requested pattern(s) on the chip site and then either pauses or goes on to the next programmed step such as move to the next chip site, etc. In other words, the software is written with a number of predefined programmed sequences that can be interrupted by the operator at any time and then a return to the established sequence can be effected.

The software system is written in Fortran and Assembly languages with about a 50/50 division. It is very modular, with most main programs being mostly "calls" to various subroutines.

References

(1) G. L. Varnell, D. F. Spicer, A. C. Rodger, "E-beam writing techniques for semiconductor device fabrication," J. Vac. Sci. Tech., 10, (1973) pp. 1048-1051.

(2) F. S. Ozdemir, W. E. Perkins, R. Yim, E. D. Wolf, "Precision electron beam microfabrication," J. Vac. Sci. Tech., 10, (1973) pp. 1008-1011.

(3) T. H. P. Chang, A. D. Wilson, A. J. Speth, A. Kern, "Electron beam lithography for complex high density devices," Proc. 6th Electron & Ion Beam Sci. & Tech. Conf., Ed. R. Bakish (Electrochemical Soc., Princeton, N. J., 1974) pp. 580-588.

(4) D. R. Herriott, R. J. Collier, D. S. Alles, J. W. Stafford, "EBES: A practical electron lithographic system," IEEE Trans. on Electron Devices, ED-22, (1975) pp. 385-392.

(5) A. J. Speth, A. D. Wilson, A. Kern, T. H. P. Chang, "Electron beam lithography using vector-scan technique", J. Vac. Sci. Tech., 12, (1976) pp. 1235-1239.

(6) T. H. P. Chang, A. D. Wilson, A. J. Speth and C. H. Ting, "Vector-Scan I: An automated electron beam system for high resolution lithography", this issue.

(7) A. D. Wilson, T. H. P. Chang and A. Kern, "Experimental scanning electron-beam automatic registration", J. Vac. Sci. Tech., 12, (1976), pp. 1240-1245.

(8) C. D. Gill and P. M. Ryan, "Electron beam registration system", U. S. patent no. 3,924,113, December 2, 1975 Assignee IBM.

(9) A. D. Wilson, T. H. P. Chang, A. J. Speth and A. Kern, "Automatic electron beam fabrication of micron size devices," Scanning Electron Microscopy/1976 (part IV), IITRI, Chicago, pp. 659-668.

Acknowledgements

The authors wish to acknowledge the contributions of W. W. Blair, T. Donohue, A. Kraft, H. Luhn, M. Parikh, C. Ting, and R. Viswanathan during the development and operation of the Vector-Scan System. The system was developed under the leadership of T. H. P. Chang. Stimulating and helpful discussions with M. Hatzakis and A. Broers have been valuable. Manuscript preparation by F. Berry is appreciated.

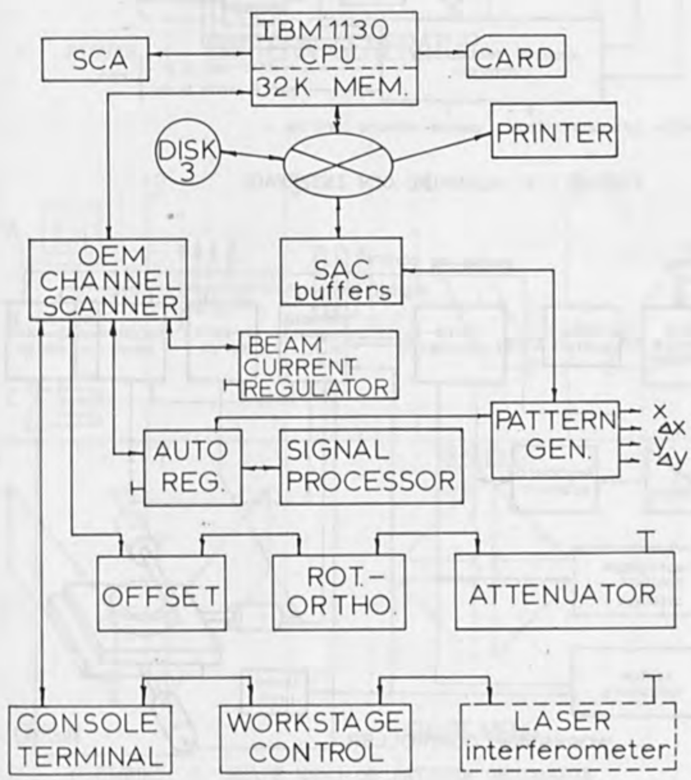


FIGURE 1 SYSYEM ARCHITECTURE

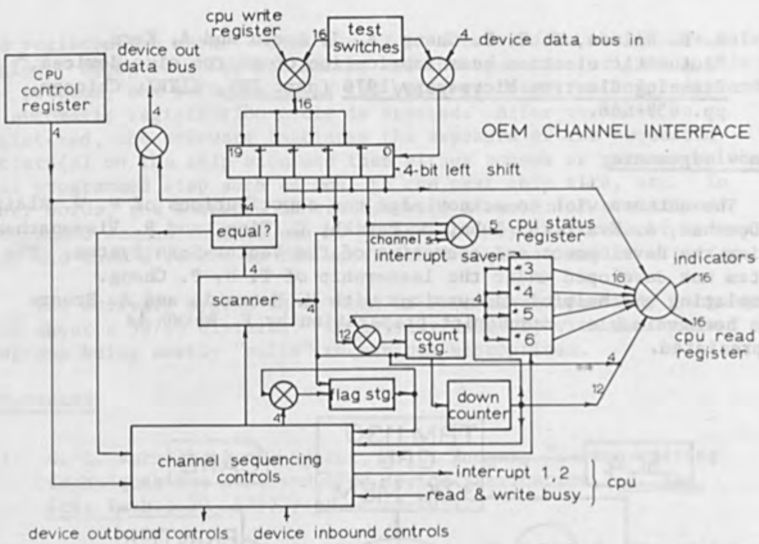


FIGURE 2 SCANNING OEM INTERFACE

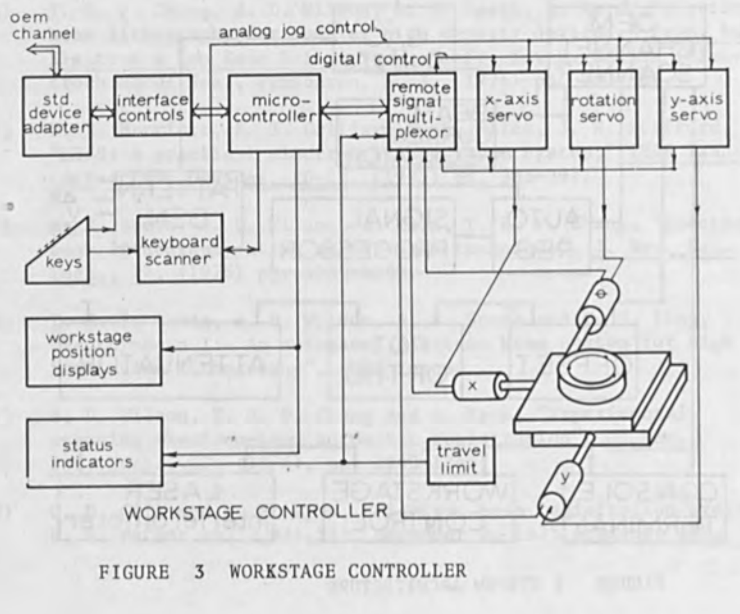


FIGURE 3 WORKSTAGE CONTROLLER

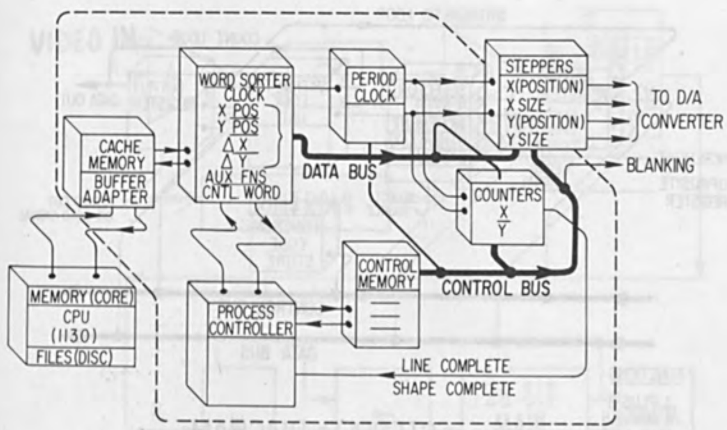


FIGURE 4 PATTERN GENERATOR

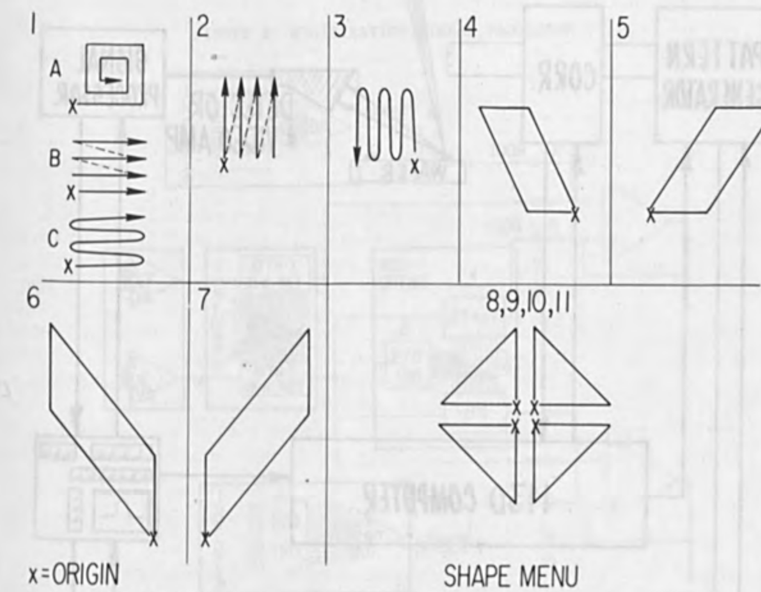


FIGURE 5 SHAPE MENU OF PATTERN GENERATOR

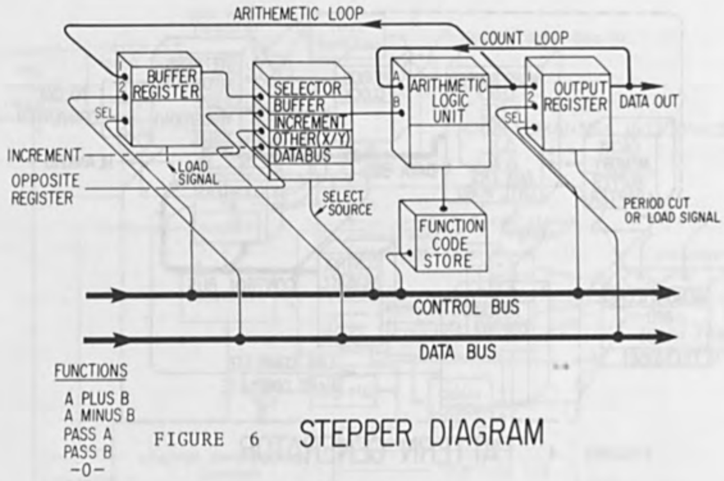


FIGURE 6 STEPPER DIAGRAM

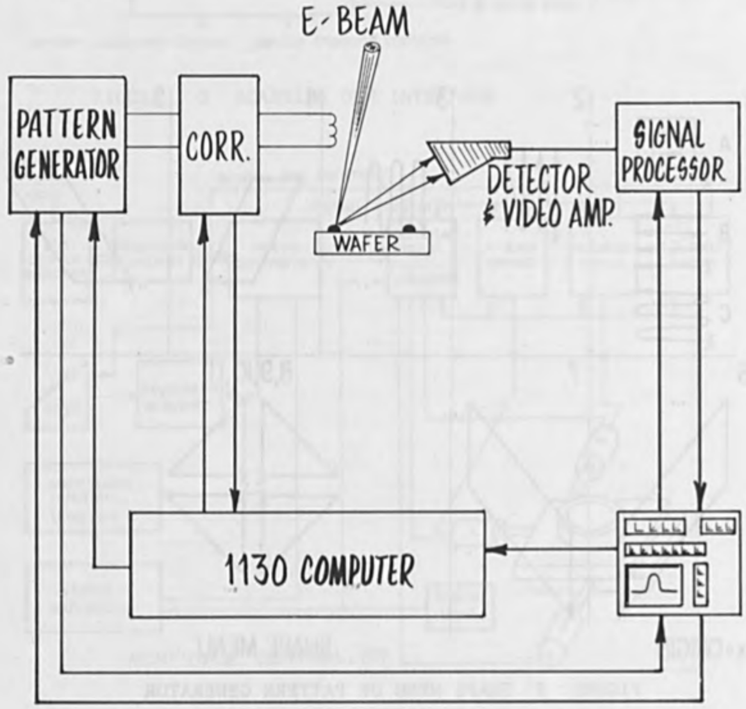


FIGURE 7 AUTOMATIC REGISTRATION ORGANIZATION

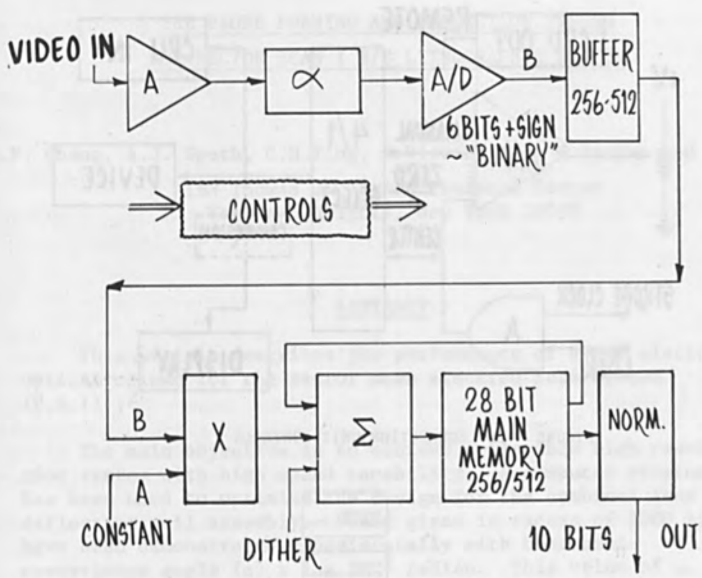


FIGURE 8 REGISTRATION SIGNAL PROCESSOR

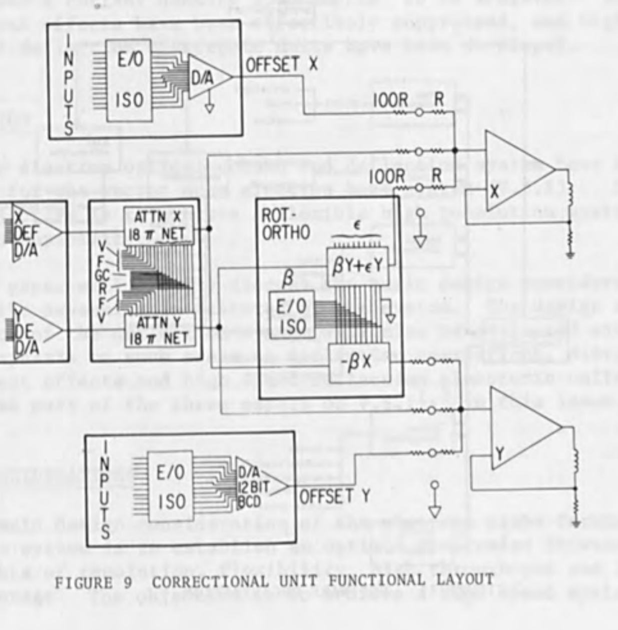


FIGURE 9 CORRECTIONAL UNIT FUNCTIONAL LAYOUT

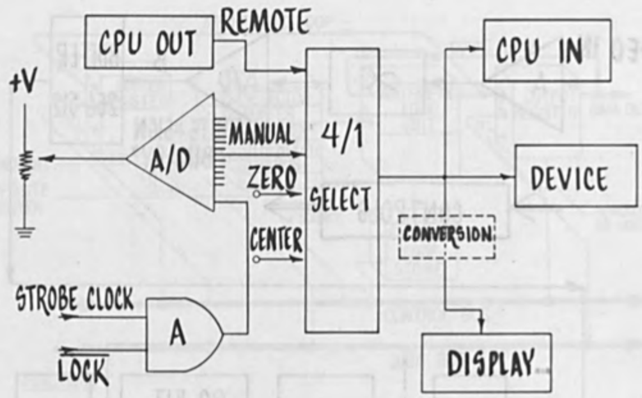


FIGURE 10 CORRECTION UNIT CONTROLS

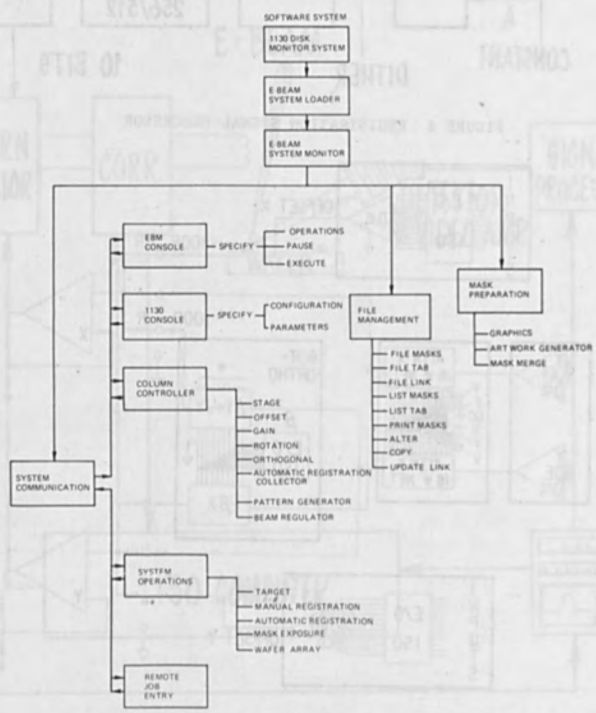


FIGURE 11 SOFTWARE ORGANIZATION

THE PROBE FORMING AND DEFLECTION SYSTEM
FOR VECTOR SCAN I E/B LITHOGRAPHIC SYSTEM

T.H.P. Chang, A.J. Speth, C.H.Ting, R.Viswanathan, M.Parikh and E. Munro

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

ABSTRACT

This article describes the performance of a new electron optical column for the vector scan electron beam system (V.S.I)[1,2]

The main objective is to achieve a flexible high resolution system with high speed capability. A computer program[3] has been used to optimize the design for the combined lens and deflection coil assembly. Field sizes in excess of 2000 lines have been demonstrated experimentally with beam semi-convergence angle (α) $\geq 1 \times 10^{-2}$ radian. This value of α, together with the use of a high brightness LaB_6 gun, have allowed a current density $\geq$300 Amp/cm^2 to be achieved. Eddy current effects have been effectively suppressed, and high speed deflection electronic units have been developed.

INTRODUCTION

A new electron optical column and deflection system have been developed for the vector scan electron beam system (V.S.I). The main design objective is to achieve a flexible high resolution system with high speed capability.

This paper will briefly discuss the basic design considerations and describe several main features of the system. The design and performance of the deflection system will also be evaluated with special emphasis on such areas as deflection aberrations, distortions, eddy current effects and high speed deflection electronic units. This paper forms part of the three papers on V.S.I[1,2] in this issue.

DESIGN CONSIDERATIONS

The main design consideration of the electron probe forming and deflection system is to establish an optimum compromise between the requirements of resolution, flexibility, high through-put and large field coverage. The objective is to achieve a high speed system that can cover a field size equal to 2000L to 4000L (L = minimum pattern linewidth) without need of dynamic focus corrections. To achieve this objective, two aspects of the system design require special attention, and they are (a) deflection aberrations and (b) eddy current effects.

(a) Deflection Aberrations

From the through-put point of view, the system should produce a high beam current for a given probe diameter. The beam current I can be expressed as:

$$I = \beta \pi \alpha^2 (\eta \frac{\pi}{4} d^2) \quad \ldots \quad \ldots \quad \ldots \quad \ldots (1)$$

where β is brightness of the electron gun
α is the semi-convergence angle of the beam
d is the diameter of the probe (neglecting lens aberrations)
η is a constant related to the gaussian distribution of the beam. If d is defined as the half-amplitude width, then $\eta = 1.44$

In the presence of lens aberrations (spherical and chromatic), and deflection aberrations, d can be replace by:

$$d^2 = d_o^2 - d_s^2 - d_c^2 \quad \text{on-axis} \quad \ldots \quad \ldots \quad \ldots (2a)$$

$$d^2 = d_o^2 - d_s^2 - d_c^2 - d_f^2 \quad \text{off-axis} \quad \ldots \quad \ldots \quad \ldots (2b)$$

where d_o is the actual beam diameter
d_s is the disc of confusion due to spherical aberration = $K_6\alpha^3$
d_c is the disc of confusion due to chromatic aberration = $K_8\frac{\Delta V}{V}\alpha$
d_f is the disc of confusion due to sum of all deflection aberrations.

A detailed analysis[3] of the deflection aberrations show the major contributions to be:

Aberration disc due to Field curvature $d_1 = K_2r^2\alpha$

Aberration disc due to Isotropic astigmatism $d_2 = K_3r^2\alpha$

Aberration disc due to Anistropic astigmatism $d_3 = K_4r^2\alpha$

Aberration disc due to Coma $d_4 = K_5r\alpha^2$

Aberration disc due to Transverse chromatic effect $d_5 = K_7r\frac{\Delta V}{V}$

where r is the off-axis distance of the deflected beam, and

$$d_f^{\,2} = d_1^{\,2} + d_2^{\,2} + d_3^{\,2} + d_4^{\,2} + d_5^{\,2} \quad \ldots \quad \ldots \quad \ldots \quad \ldots(3)$$

Taking all these aberrations into consideration:

$$I = (1.44\,\frac{\pi^2}{4}\beta)\alpha^2[d_o^{\,2} - (K_6\alpha^3)^2 - (K_8\frac{\Delta V}{V}\alpha)^2 - d_f^{\,2}] \ \ldots \quad (4)$$

For a system in which the aberrations are dominated by K_8, K_2, K_3 and K_4, (this is the case for V.S.I, as can be seen later in Table 1), eq.(4) can be simplified considerably. The optimum beam semi-convergence angle which maximizes the beam current in a probe of given diameter is then given by:

$$\alpha_{opt} = \frac{1}{\sqrt{2}}\, d_o\, (K_8\,\frac{\Delta V}{V} + \Sigma K r^2)^{-1} \quad \ldots \quad \ldots \quad \ldots \quad \ldots(5)$$

$$\text{where } \Sigma K = K_2 + K_3 + K_4 \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots(6)$$

Substituting this value of α_{opt} in eq.(4), the optimum beam current for a given probe size can be determined and this value will reflect the effects of both on and off axis aberrations.

The field size that can be achieved in a given system will also depend on the off-axis aberrations. It can be seen from eqs.(2a) and (2b) that deflection aberrations cause the beam diameter to increase with deflected distance. Normally, in order to limit the beam diameter at the fully deflected position to within a given tolerance, it is necessary to use a somewhat smaller beam at the center of the field and to choose a beam semi-convergence angle so that the growth of beam size is kept under control.

There is, however, a simple way of reducing the deflection aberrations. This involves focussing the beam at a predetermined distance off axis, so as to reduce the effect of field curvature. In this case, a relatively more uniform beam diameter can be achieved over the whole field, and the effect of K_2 is considerably reduced. At the point of focus, the contribution of K_2 is zero.

It can be seen from the above that one of the main design requirements, from both through-put and field size points of view, is to achieve a system with low aberrations, in particular low deflection aberrations. A computer program[3] has been developed for the design of the lens-deflection system and this program was used in the development of this system.

(b) Eddy Current Effects

Eddy current effects[4,5] are associated with deflection systems employing deflection coils. The magnetic deflection field induces eddy currents in adjacent metallic components and these currents in turn set up a magnetic field which opposes the deflection field. As a result, an error in the deflected position of the electron beam is introduced. Although this positional error eventually disappears as the eddy currents subside to zero, the relaxation time can, in some cases, be excessively long. It is essential to remove or reduce these eddy current effects so that the desired pattern can be written accurately at high speed.

There are, in general, two main sources of eddy currents:

* Eddy currents in thin metallic films such as the conductive film in the screening tube.

* Eddy currents in thick cylindrical shells and other solid metallic parts. (The two major areas where this situation exist are the mild steel bore of the final lens, and other metallic components adjacent to the deflection coils.)

First, consider the case of a thin conductive film in the screening tube. It has been shown[6] that when a thin film cylinder is subjected to a sinusoidal external magnetic field, B $\cos\omega t$, the resultant field inside the cylinder is given by:

$$B' = B\,2S\,(4S^2 + \omega^2\mu_o{}^2a^2)^{-1/2}\cos(\omega t - \varepsilon) \quad \ldots \quad \ldots \quad (7)$$

and

$$\varepsilon = \tan^{-1}\left(\frac{2S}{\mu_o\omega a} - \frac{\pi}{2}\right) \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (8)$$

where B = Amplitude of the applied field
B' = Resultant field inside the cylinder
S = Surface resistivity of thin film
a = Radius of the cylinder
μ_o = Permeability of free space
ω = Angular frequency of applied field

It can be seen that the effect of the eddy current is to introduce both an attenuation and a phase shift to the applied field. Both these parameters depend on the frequency of the applied field: attenuation increases with increasing frequency, while the phase shift approaches $\pi/2$ when frequency is infinite. It can be shown that an analogous RC network giving the same response, would have a time constant

$$RC = \frac{\mu_o a}{2S} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (9)$$

Thus when an external field in the form of a step signal is applied, the resultant field inside the screening tube will take the familiar transient response as given by:

$$B' = B\,(1 - \exp \frac{-t2S}{\mu_o a}) \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (10)$$

To eliminate this effect, it is necessary to find a conductive film which has a high surface resistivity and can be applied uniformly to the internal surface of the screening tube.

In the case of eddy currents in solid cylindrical shells, analytical analysis available is much less complete because of the complex nature of the problem. Jaeger[7] has studied eddy current effects induced by a step magnetic field in an infinitely long, thick, hollow conducting cylinder. He found that the rise of field inside the cylinder takes a complex form involving several Bessel functions. The rise time, which is related to resistivity, permeability and geometry of the cylinder, can be as high as several milliseconds for the case of a copper cylinder and as much as tens of milliseconds for the same cylinder in mild steel. Although Jaeger's example is not an accurate description of the case in question, it serves to demonstrate that relaxation time of eddy currents in metallic cylinders can be very lengthy. Suppression techniques are therefore required to reduce the magnitude of the leakage deflection field that couples into the metallic parts, and to reduce the relaxation time to a tolerable level.

THE ELECTRON OPTICAL COLUMN

The electron optical column consists of (1) a probe forming system to produce a focussed beam of electrons, (2) a beam blanking unit to switch the beam on and off, (3) a deflection system for the positioning of the beam, (4) a detector system to collect an electron signal for focussing and registration purposes and (5) a workstage and airlock system for the movement and handling of the workpiece. A schematic cross-sectional view of the column is shown in Fig.1.

The probe forming system is comprised of a lanthanum hexaboride gun followed by three magnetic lenses. The LaB_6 gun used is a new version of the Broers'[8] gun with improved performance in stability and warm-up time (typically less than 5 minutes). A brightness value of 1×10^6 Amp/cm^2 steradian at 25kV and a life of several hundred hours can be achieved routinely. The gun alignment system uses a mechanically movable anode for coarse adjustments and an alignment coil for fine adjustments.

Three magnetic lenses are used to form the focussed probe. The first two lenses have been designed as a pair, so as to image the source to a fixed image plane immediately below the second lens. The final lens, the objective, focusses the beam at the surface of the workpiece. By properly coupling the current controls of the two condenser lenses, the size of the probe can be varied without the need to refocus the final lens. This arrangement allows the final lens current to stay constant for all beam sizes and thus eliminates the problem of pattern field rotation associated with lens current changes.

The objective lens is designed along with the deflection coil assembly as an integral assembly. Special attention has been given to the problem of eddy currents and extensive suppression techniques have been incorporated. Results covering this area will be discussed in more detail later.

The beam blanking unit is of the electrostatic type and is normally installed in the first condenser lens, although it can also be placed in the second condenser lens as shown in Fig.1. The concept of placing the blanking unit in the plane of the condenser lens crossover, as proposed by Lin[5], has been adopted. This allows the beam to be blanked with a relatively low voltage (a few volts) and also minimizes shift in beam position during the blanking process.

The beam limiting aperture carrier has fine x and y adjustments and a selection of three apertures for experimental convenience. The apertures can be cleaned in situ by electrical heating.

The workstage is capable of 4" x 4" traverses in the x and y directions and a $\pm 5^{\circ}$ rotational adjustment. The workpiece can be introduced to the system through an airlock unit designed by G. Folchi. The pumpdown time of the airlock is typically less than 15 seconds.

The entire column rests on a pneumatic antivibration platform with a natural frequency of approximately 2 Hz.

The vacuum system consists of a diffusion pump for the chamber area and an triode-ion pump for the upper column. The vacuum level attained is typically in the 10^{-7} torr range.

THE DEFLECTION SYSTEM

This section will discuss the performance of the deflection system in which a pre-lens double deflection coil is used. Other electro-magnetic deflection structures[9,10,11] have also been used in electron beam microfabrication systems.

The deflection coils have been designed to have a very low inductance. Precision winding techniques have been developed to achieve the desired tolerances on wire position. The entire deflection coil assembly is housed inside the bore of the objective lens. A mechanical adjustment for coil rotation is provided for coarse control and an electronic rotation and orthogonality control unit provides fine adjustments. The deflection unit is designed to cover a maximum field size of 5mm x 5mm at a working distance of approximately 5 cm.

Computed Performance of Deflection Unit

The performance of the deflection unit, as computed by the computer program, is shown in Table 1. In order to achieve this performance, various optimization procedures described by Munro[3] have been applied.

TABLE I

COMPUTED PERFORMANCE OF VSI DEFLECTION SYSTEM

Components		Values of K	Aberrations @ Corner of Field Size	
			2mm	4mm
Distortion	$(K_1 r^3)$	3.76×10^{-3} cm^{-3}	0.11µm	0.85µm
Field Curvature	$(K_2 \alpha r^2)$	0.058 cm^{-2}	0.07µm	0.28µm
Isotropic Astigm	$(K_3 \alpha r^2)$	0.035 cm^{-2}	0.04	0.17
Anisotropic Astigm	$(K_4 \alpha r^2)$	0.023 cm^{-2}	0.028	0.11
Coma	$(K_5 \alpha^2 r)$	0.196 cm^{-1}	0.01	0.02
Spherical	$(K_6 \alpha^3)$	5 cm	0.01	0.01
Transverse Chrom	$(K_7 \frac{\Delta V}{V} r)$	0.07 cm^{-1}	0.01	0.02
Axial Chrom	$(K_8 \frac{\Delta V}{V})$	7 cm	0.04	0.04
Overall aberrations with no dynamic corrections			0.095µm	0.35µm
Overall aberrations with dynamic corrections			0.043µm	0.050µm

r = Radial distance from center

α = Semi-convergence angle (6×10^{-3} rad.)

$$\frac{\Delta V}{V} = 1 \times 10^{-4}$$

Table 1 also gives the computed values for deflection aberrations and distortion at the corner of two field sizes, 2mm x 2mm and 4mm x 4mm. These values were computed assuming a beam semi-convergence angle of 6×10^{-3} radian and $\Delta V/V$ of 1×10^{-4}.

Using the results of Table 1, the optimum value of α for a given beam diameter can be evaluated using eq.(5). For example, for a beam diameter of 0.25μm, which is normally used to expose 1μm pattern lines, and for a 2mm x 2mm field, the α opt values for two operating conditions have been computed.

Case 1 Beam focussed on center of field

$$\Sigma K = K_2 + K_3 + K_4 = 0.116 \text{ cm}^{-2}$$

$$\alpha_{opt} = 5.85 \times 10^{-3} \text{ radian}$$

Case 2 Beam focussed off axis to reduce the effect of K_2

$$\Sigma K = K_3 + K_4 = 0.058 \text{ cm}^{-2}$$

$$\alpha_{opt} = 9.5 \times 10^{-3} \text{ radian}$$

The computer program has proved to be a very powerful method for the design of the deflection unit. Its self optimization routines allow comparative performance for a wide range of different lens-coil designs to be evaluated readily.

Eddy Current Suppression

To suppress the effects of eddy current in metallic components adjacent to the deflection coils, the final lens has been designed to use ferrite as its polepiece material. In addition, ferrite cylinders were also used as shields over the entire bore of the lens. The ferrites have a permeability in the range of 5000 to 10,000 and resistivities $\simeq 5 \times 10^7$ ohm-cm. (approx. 6 orders higher than mild steel). The use of ferrite as a polepiece material requires the development of special machining techniques. Extreme care has to be exercised in the handling of the material to avoid accidental chipping of the precision ground edges. Satisfactory system performance has been achieved.

In addition, many other mechanical components in the vicinity of the deflection coils have been replaced by non-metallic materials. Surfaces within line of sight of the electron beam such as the inner surface of the screening tube, are coated with a special conductive film of a high surface resistivity. This coated surface is hardened by a high temperature bake and is not easily damaged by cleaning.

Deflection Electronics

A layout of the electronic components that drive the deflection coils is shown in Fig.2. A special dual channel D/A system is used to receive the data from the pattern generator. In this system, the corner address of the pattern cell, normally 14 bit resolution, is fed into a 16 bit high precision 'low speed' D/A unit, and the fill-in scan data is fed into a separate 10 bit 'high speed' D/A unit. The outputs of these two D/A units are summed and sent to a programmable resistive attenuator unit which controls the field size. The main advantage of this dual channel D/A system is that a much higher conversion rate can be attained relative to a single channel D/A system, principally due to the capability of attenuating the switching noise in the lower resolution unit. This approach places a restriction on the maximum cell size, but the inconvenience introduced is minimal.

With regard to electrical performance, the deflection driver, the deflection coils and their interconnections must be considered as a composite network. From the system point of view, two parameters are sufficient to characterize the deflection speed capability. The first relates to large signal response; the time required to accurately reposition the electron beam to within ½LSB. The second is the small signal bandwidth, which determines the lag between the output and input, and also the transient duration at the initiation of a sweep. The resonant frequency of the coil and its driver cable will, of course, set a limit to the small signal bandwidth capability.

EXPERIMENTAL RESULTS

The performance of the deflection system in terms of field size coverage has been evaluated experimentally. Figs.3a to d show enlarged views of a resolution test pattern generated by the system at the following four positions: (a) center of field, i.e. on axis, (b) at corner of a 2mm x 2mm field, i.e. 1.41mm off axis, (c) at corner of a 3mm x 3mm field, i.e. 2.12mm off axis, (d) at corner of a 4mm x 4mm field, i.e. 2.83mm off axis. The pattern consists of a series of lines and squares at different linewidths and spacings. The smallest line and gap in each case are approximately 0.5μm. These results were achieved with a beam semi-covergence angle (α) equal to 1×10^{-2} radian. The beam was focussed at the center of the field. The results show that reasonably good pattern definition for 0.5μm lines is observed up to a 3mm field, and for 1μm lines, up to 4mm field. These results agree well with the objectives for the system.

Using a beam semi-convergence angle, α, of 1×10^{-2} radian and a gun brightness, β, of 1×10^{6} Amp/cm^2 steradian, the typical beam current density (j) of the system is given by:

$$j = \beta \pi \alpha^2$$
$$\geq 300 \text{ Amp/cm}^2$$

The pattern distortion (pincushion) over a 2mm x 2mm field size has been measured by examining the results of field stitching experiments. It was found that the distortion is considerably less than the 0.1μm value predicted by the computation (Table 1), since no measureable stitching error was observed. Experiments to verify the distortion over a 4mm x 4mm field have yet to be performed.

The effects of eddy currents have also been studied experimentally. Fig.4 shows a pattern obtained using a lens-deflection assembly without eddy current suppression. The positional error introduced at the starting position of the pattern cell is as much as 5% of the jump distance (OA) and the decay time is long ($>10^{-3}$ sec.). Fig.5 shows patterns exposed after the suppression techniques have been applied. Two pattern cells are shown, each 30 beam step units wide, with jump distances corresponding to 4000 and 8000 beam steps respectively. A nominal 50μs delay has been provided for settling in each case. It can be seen that the effect of eddy currents is negligible. No measureable effect is seen in the case of 4000 step jump and an approximately 1 step error is seen in the case of 8000 step jump, which may or may not be due to residual eddy current effects. Typically, this residual error can be removed by using a slightly longer settling time than the value quoted above.

The amplifier-coil combination has an overall small signal bandwidth of 2.5 MHz. This bandwidth allows output current ramping at a rate equivalent to a 10 MHz D/A LSB update rate, with about ½LSB lag between output and input. Propagation delays in the amplifier can be accomodated by timing adjustments to the beam blanking signals.

Test exposures shown in Figs.5 & 6 in the accompanying paper[1] show two levels of an 8K FET memory chip exposed at a 10 MHz stepping rate. Good pattern integrity with no loss of definition on edges and corners is observed. Efforts are continuing to extend the speed performance of the deflection system.

SUMMARY

A high speed electron probe forming and deflection system has been developed. The performance of the system can be summarized as follows:

Column	3 lenses with LaB_6 gun
Accelerating potential	25kV
Gun brightness (β)	$\geq 1 \times 10^6$ Amp/cm^2 steradian
Beam diameter range	Variable from 0.05μm
Field distortion	2mm x 2mm field ∿0.05μm (verified) 4mm x 4mm field 0.85μm (to be verified)

Beam semi-convergence angle (α) and field size	$\geq 1 \times 10^{-2}$ radians experimentally demonstrated for 1µm lines over 4mm field ½µm lines over 3mm field with no dynamic corrections
Current Density (j)	≥ 300 Amp/cm^2
Eddy current effects	Negligible after suppression
Deflection Amplifier-coil characteristics	Large signal response: 20µs to ½LSB (~100µs at present) Small signal response: 2.5 MHz bandwidth
D/A System	Dual channel system (16 bit precision D/A + 10 bit high speed D/A) 10 MHz operation demonstrated.

ACKNOWLEDGMENT

The authors gratefully acknowledge the technical contribution of W.W. Blair, T.P. Donohue, G.A. Folchi, P.R. Jaskar, W.L. Keller, A. Kern, A. Kraft, H.E. Luhn, A.M. Patlach, T.W. Studwell and A.D. Wilson in the development of V.S.I. In particular, G.A. Folchi and W.W. Blair, in conjunction with members of the Central Scientific Services Department of IBM Thomas J. Watson Research Center, have contributed to the design and fabrication of the system's mechanical components. W.W. Blair also took part in developing the precision coil winding technique and G.A. Folchi is responsible for the development of the air lock system. Helpful discussions with A.N. Broers, M.B. Heritage and H.C. Pfeiffer are also appreciated.

REFERENCES

1. T.H.P. Chang, A.D. Wilson, A.J. Speth and C.H. Ting, "Vector-Scan I an Automated Electron Beam System for High Resolution Lithography", this issue.

2. A.D. Wilson, A. Kern, A.J. Speth, A.M. Patlach, P.R. Jaskar, T.W. Studwell & W.L. Keller, "Automation of Vector Scan I Electron Beam Lithographic System", this issue.

3. E. Munro, J. Vac. Sci. Technol. 12, 1146 (1975)

4. G.A. Wardly, J. Appl. Phys. 45, 2316 (1974)

5. L.H. Lin & H.L. Beauchamp, J. Vac. Sci. Technol. 10, 987 (Nov./Dec. 1973).

6. W.R. Smythe, "Static and Dynamic Electricity" (McGraw-Hill, New York, 1950), Chapter 11.

7. J.C. Jaeger, Phil. Mag. 29, 18, (1940).

8. A.N. Broers, J. Sci. Instrum. Ser. 2, 2, 273 (1969).

9. K. Amboss, J. Vac. Sci. Technol., 12, 1152 (1975).

10. G. Owen & W.C. Nixon, J. Vac. Sci. Technol., 10, 983 (1973).

11. H.C. Pfeiffer, J. Vac. Sci. Technol., 12, 1170 (1975).

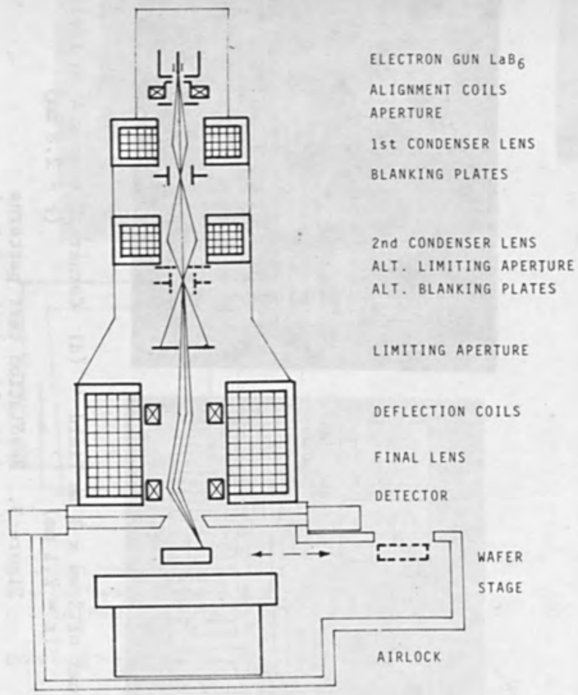


Figure 1. Schematic cross-section of the electron optical column

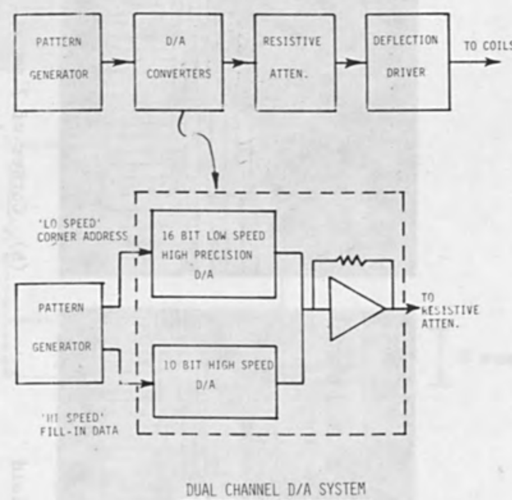
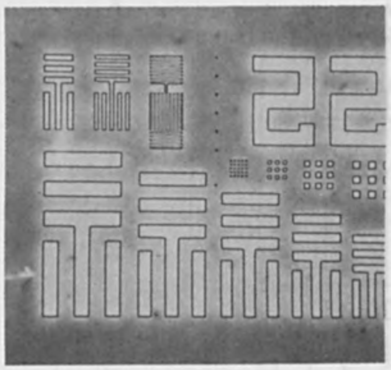
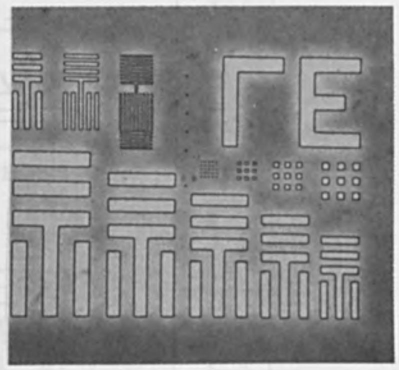


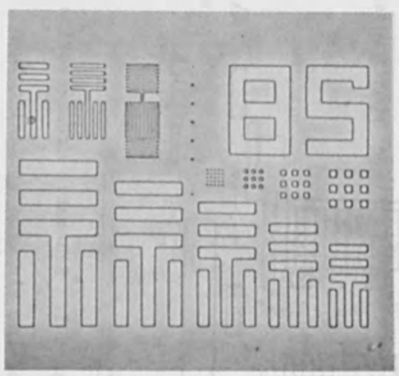
Figure 2. The deflection electronics



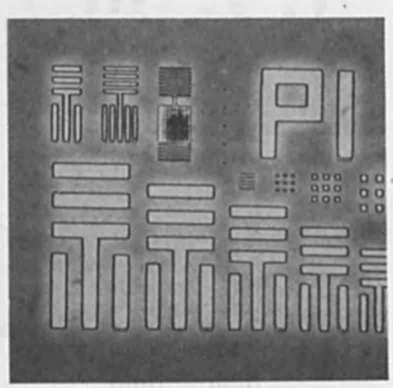
(a) Center of field



(b) Corner of 2 mm x 2 mm field
(r = 1.4 mm)



(c) Corner of 3 mm x 3 mm field
(r = 2.1 mm)



(d) Corner of 4 mm x 4 mm field
(r = 2.8 mm)

Figure 3. Resolution test patterns
(min. linewidth = 0.5 μ)

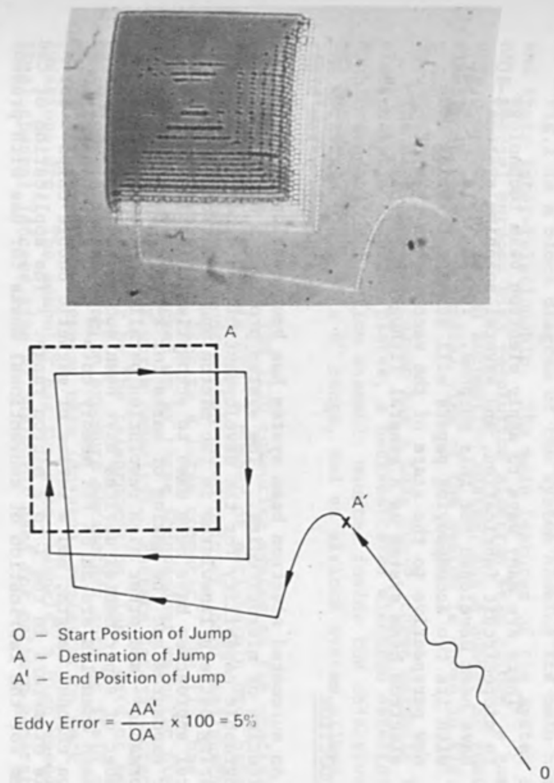


Figure 4. Pattern exposed with no suppression for eddy current effects

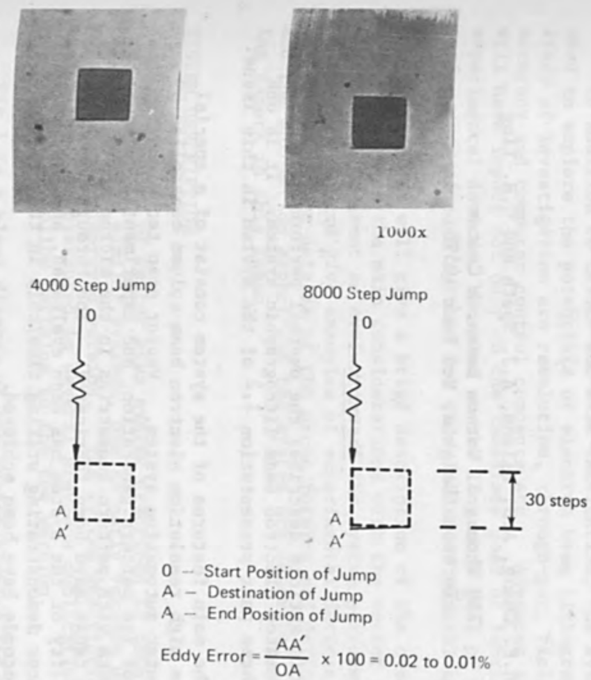


Figure 5. Patterns exposed with suppression for eddy current effects

VECTOR-SCAN I AN AUTOMATED ELECTRON BEAM SYSTEM FOR HIGH RESOLUTION LITHOGRAPHY

T.H.P. Chang, A.D. Wilson, A.J. Speth and C.H. Ting

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Abstract

This article describes the overall performance of V.S.I, an automated electron beam lithographic system. It is one of a three part presentation [1,2] of the system in this issue.

The main features of the system consist of a special purpose high resolution electron beam column controlled by a computer automation system. Vector scan technique is used for its pattern generation, and experimental devices and masks with pattern geometries in the micron and sub-micron range have been fabricated. The through-put capability of the system has been evaluated, and test exposures demonstrating writing time/chip in the order of 1 to 2 seconds have been achieved.

There are many ways to apply electron beam technology to microcircuit fabrication, and several scanning systems[3-10] have been developed for this purpose. This article, together with its two accompanying papers will attempt to bring some new perspective to the state of the vector scan type of electron beam system as a general lithographic tool.

INTRODUCTION

An automated electron beam system has been developed for the fabrication of microcircuits. The system provides a high resolution lithographic capability for the development of new experimental devices requiring pattern linewidths in the micron and submicron range. Several approaches have been used to accomplish this, including direct wafer exposures and the making of masks for x-ray[11], deep U-V conformable[12] and other high resolution parallel printing techniques. A number of experimental devices have been successfully fabricated. The main examples are an 8K FET memory chip[13] with approximately one micron channel length, and a series of magnetic bubble chips[14] for bubble diameters in the 1 to 2 micron range. The application of the system for the fabrication of conventional masks for use with present day optical printing systems has also been studied.

In addition to device and mask fabrication, the system has been used to explore the potentials of electron beam lithography. The key areas of investigation are resolution, through-put, field size, accuracy and computer control concepts etc. Advances in these areas will have impact not only on the immediate goal of fabricating experimental devices, but also on the long term goal of defining the performance of such systems for wider and more general applications.

This paper will give a brief description of the overall system and discuss some of the main considerations with the vector scan technique. It will also present an evaluation of the system performance in terms of through-put and give examples of experimental devices fabricated by the system. The details of the probe forming and deflection system[1] and the automation aspect[2] of the system are described in two accompanying papers in this issue.

SYSTEM DESCRIPTION

The results of vector scan system described in this paper are derived from two similar systems, one an early version and the other a more recent and improved version. The system description of the early system has been presented in several previous publications[15-17]. The description to follow pertains to the new version.

Fig.1 is a block diagram of the system showing its main components and their inter-relationship. The main features of the system consist of a special purpose high resolution electron beam column, a comprehensive computer controlled automation system, and a number of control electronic units. A general view of the system is shown in Fig.2.

The electron beam column has a LaB_6 gun[18], three magnetic lenses, electrostatic blanking plates, a backscattered electron detector and a double magnetic deflection assembly mounted inside the objective lens. The workstage has a 4" x 4" range, and an airlock system allows for rapid wafer changes.

The system uses an IBM 1130 computing system as the central processing unit. To this CPU, a number of peripherals have been attached and these include a high speed printer, card read/punch unit and a telecommunication link to a large computer. The 1130 has a 32K word (16 bit) memory backed up by three independent magnetic disk units (each data disk has 500K words capacity). Pattern data is entered into the computer by cards or by the teleprocessing link. Card or keyboard entry is used for setting up exposure parameters such as field size, chip array, exposure stepping rate (dosage), registration mark locations etc. The computer has two data channels, a high speed cycle-steal channel (access time $\sim$2.5μs/word) used for pattern generation and a low speed flexible software channel (access time $\sim$150μs/word) used to control the workstage, field size, registration corrections, operator console etc.

The control electronic units can be divided into two main parts; the digital control electronics and the analog electronics. The digital electronic units are all interfaced to the computer via one of the two channels mentioned earlier. They provide the logic functions required for the automatic operations of pattern generation, registration and work stage movements etc. The analog electronics are used primarily for the functioning of the electron beam column and the deflection system.

Patterns are written in the vector-scan mode using the following basic cell shapes: rectangles, parallelograms (all angles) and triangles. In addition a random shape store is provided for non-rectilinear shapes and for repetitive groups of cells. Each cell is described by up to five binary words and data preparation is attained by a proprietary CAD facility operating offline.

The pattern for each chip is defined on a grid structure containing 16,384 addressable points, i.e. 14 bits, on each exis. (The hardware is designed to handle up to 65,536 points, i.e. 16 bits per axis, for future extension). During writing, the pattern generator always positions the electron beam to the starting corner location of each pattern cell using the basic 14 bit grid structure. However, in the fill-in operation, the generator has the option of either stepping the electron beam an increment equal to the basic grid resolution or stepping an increment equal to 2 or any other integer multiple of the 14 bit basic resolution (or LSB). The larger the increment would obviously give faster writing time at the same stepping rate, but could also degrade pattern integrity if carried too far. In practice, it was found that satisfactory pattern definition can be achieved by a beam increment equal to 2 LSB and by the use of 4 beam increments to define the minimum pattern linewidth. This option is therefore normally used in V.S.I. Using these groundrules, the chip size of the pattern field is exposed using approximately 8000 beam increments per axis and as 4 beam increments are needed to define the minimum linewidth, the chip size is often referred to as a 2000 "line" field, where "line" refers to the minimum pattern linewidth. This ability to vary the beam stepping increments in a dynamic manner has several important applications. It can be used for example, to improve registration accuracy by having a finer increment for the locating of registration marks and a coarser increment for writing. It can be used to improve pattern integrity by using the finer increment for critical pattern elements and a coarser increment for less critical elements, or to define the boundary of each cell by a finer increment using a smaller beam, and a coarser increment for the interior of the cell using a larger beam size.

Data is transferred from the 1130 computer to the pattern generator via the high speed channel. A small buffer (256 words at present) is provided in the generator for high speed operation. The generator is of a programmable type with a wide range of predefined primitive cell shapes. It is also of a modular design to allow for future extension. The rate at which the individual pattern cells are exposed is controlled by a stepping rate clock. Counting circuitry in the pattern generator drives a D/A converter system. The D/A outputs, after attenuation by the field size control unit, drive the deflection current amplifier.

Automatic registration is accomplished using an enhancement technique[16] to improve the signal-to-noise ratio under full program control of the computer. A dual channel detector system[19] using grounded scintillators has been instrumented to collect the back-scattered electron signal from the registration marks. The digitized signals after enhancement are transferred to the computer and field corrections are computed and applied via the field size, offset, rotation and orthogonality correction units to the deflection system.

The work stage is controlled by a micro processor based control unit. A special servo system that uses low inertia high torque servo motors with shaft encoders is used to drive the stage lead screws in three axes, i.e. x, y and Θ. A laser interferometric measuring system is currently under development to provide a higher position precision.

The system operator communicates with the software control program of the system via a console terminal which provides an interactive control for the system operation.

In addition, the system can also be operated in the SEM mode for inspection purposes. This is particular useful for system set-up and testing. A manual registration system is also provided to give additional information in registration when required.

Briefly, the main features of the system can be summarized as the following:

Configuration:	A three-lens electron beam column controlled by an 1130 computer system.
Pattern Linewidth:	0.1μm minimum
Pattern Cell Shapes:	Rectangle Parallelogram (all angles) Triangle (all angles) Random shape store
Deflection System:	Pre-lens double magnetic deflection

Field Size:	Continuously adjustable from 0.1mm x 0.1mm to 5mm x 5mm
Chip Size:	2000 x minimum pattern linewidth (typical) e.g. 2mm x 2mm chip for 1μm lines 5mm x 5mm chip for 2½μm lines
Workpiece Size:	3½" diameter wafer 4½" mask plate
Stage Traverse:	4" x 4" in x-y directions with $\pm 5^{o}$ rotation
Automation:	A. Pattern generation B. Registration C. Stage movements D. Proximity effect corrections

VECTOR SCAN TECHNIQUE

A number of techniques[3,6,7,9] have been used to implement the exposure of pattern cells in vector addressed electron beam exposure systems. The basic principle of the technique is to dissect the pattern details within the deflection field into a composite of simple cell shapes such as rectangles and parallelograms. The electron beam exposes the pattern by serially filling in each pattern cell. This method is simple to implement from both the system and hardware point of view since algorithms for providing data of this type are normally used in existing optical lithographic systems. Also, special hardware generation of a limited number of pattern shapes relieves the source computer (or memory) of a high through-put requirement, since data access is required only once per cell. In addition, because of the repetitive nature of most processed patterns, substantial data compaction can be attained. This technique is also time efficient since the electron beam is only addressed to the pattern area requiring exposure. To demonstrate the significance of these attributes, the data capacity and pattern area coverage for an 8K FET memory chip is used here as an example. This chip has four pattern levels (A to D) consisting only of rectangular cells. Table 1 gives the number of rectangles in each pattern level, the data compaction parameters and pattern area coverage.

TABLE I

Level	No. of Rectangles	Unpacked Data*	Packed Data*	Data Compaction(%)	Pattern Area Coverage (%)
A	35,761	178,805	55,955	31	26.95
B	31,085	155,400	49,162	31.6	30.35
C	12,387	61,935	20,545	33	1.35
D	26,166	130,830	48,715	37	27.65
* (16 bit words)				Av. 33.2%	21.58%

It can be seen from this example that as a result of data compaction, an average of 1.5 words are needed to specify one rectangular cell as compared to 5 normally required. It should also be pointed out that these results are achieved with a relatively primitive sorting algorithm which by no means provides optimum performance. The average pattern area coverage is approximately 22%. This is probably a high average value as the device is of a high density memory type. A more general average value is believed to lie between 10 and 20%.

The vector scan technique also offers a high degree of flexibility. One important example is the ability to compensate for proximity effects[20] by adjusting the beam stepping rate during the exposure of each pattern cell. This compensation technique is significant for pattern geometries in the micron and submicron range.

The main difficulty in the successful implementation of the vector scan approach lies in the deflection system. Because of the vector address nature of the operation, the settling time requirement of the deflection current driver becomes very stringent. The eddy current effects associated with magnetic deflection also have to be carefully controlled and suppressed. Some very encouraging results have been achieved in V.S.I in solving this deflection problem and they will be presented in this paper and the two accompanying papers.

Techniques for generating the line-scan to fill-in a particular pattern cell vary in the combination of digital and analog techniques. Fig.3A to C show the three basic fill-in techniques most commonly used. Fig.3A uses a bi-directional raster scan. This is the simplest approach and can be obtained by using D/As only or by the use of two D/A converters per axis to establish the bounds between which a rate controllable analog ramp is generated[21]. Fig.3C is the spiral scan technique which provides a continuous scan around the boundary of the cell[7,15]. This technique has the advantage of offering better edge exposure control and can be used in conjunction with an added line at the edge to enhance the exposure as described by Speth[15]. Fig.3B is a hybrid between 3A and 3C.

Fig.3D to E show possible extension of the three basic techniques. In Fig. 3D a high frequency signal source is used to shape the beam into a small rectangle (or parallelogram) and this "shaped" beam is stepped over the cell using the bi-directional scan. The size of the "shaped" beam can be adjusted to achieve a proper fit in the boundary of the cell. The main advantage of this approach is to achieve a very high speed beam stepping rate. Fig. 3E & F both demonstrate the use of a small beam to expose the edges of the cell where the definition is critical and a larger beam to expose the interior of the cell where definition is less critical so as to achieve a higher writing speed. In practice, the advantages gained by such approaches become significant as cell size increases and it is estimated that on the average, the improvement of writing time should be a factor of 2 to 3.

In addition, there is the possibility of using different beam sizes for each cell in proportion to the cell size.

THROUGH-PUT CONSIDERATIONS

The system exposes a wafer or a mask in a step and repeat mode of operation which involves the following major steps: pattern write, stage movement and registration. The through-put of the system is therefore measured by the speed in which each of these steps are performed, i.e.

$$T = N\,(T_w + T_r + T_s)$$

where T is the exposure time per wafer
T_w is the writing time per chip
T_r is the registration time per chip
T_s is the stage move time per chip
N is the number of chips per wafer

Of special interest is the speed needed to write a given pattern T_w, and the factors covering this speed can be divided into two parts:

1. The exposure time T_e which is the actual time taken by the beam to expose the pattern elements.

2. The overhead times T_o which are the various waiting times needed in the exposure.

A more detailed discussion of these two factors will now follow.

Exposure Time/Chip

The chip size in V.S.I is defined in the following manner:

Chip size = 2000 x minimum pattern linewidth
Each min. pattern line = 4 x beam spots
Total beam spots/chip = 8000 x 8000
= 64 x 10^6 spots

The time required to write such a chip depends on the speed at which the electron beam is being moved, i.e. stepping rate, and the percentage coverage of the pattern area, i.e.

$$\text{Exposure time/chip} = \frac{(64 \times 10^6)\ K}{\text{Stepping rate}}$$

where K is the percentage pattern area coverage.

Table 2 shows the exposure time at different stepping rates for three exposure techniques, raster, vector and vector with multiple beam sizes. A pattern area coverage of 20% is assumed for the vector case.

TABLE 2 EXPOSURE TIME/CHIP (SEC.)

Stepping rate	1 MHz	10 MHz	50 MHz
Raster Scan (K=1)	64	6.4	1.3
Vector Scan (K=0.2)	12.8	1.28	0.26
Vector Scan with multiple beam size (estimated)	<5 sec.	<0.5 sec.	<0.1 sec.

It can be seen from Table 2 that a short exposure time can be attained with high beam stepping rate and it is therefore important to establish how this high stepping rate can be achieved.

The factors governing the stepping rate are manifold. First, the electron optical system has to provide a sufficiently intense beam to cause the resist coating to be adequately exposed. Second, the deflection system; D/A,deflection amplifier and coil, have to be sufficiently responsive at fast beam stepping rates,without introducing positional errors or excessive noise.

The electron optics and resist sensitivity relate to the beam stepping rate by the following equation:

$$\text{Operable stepping rate} = \frac{\text{Current density (j)}}{\text{Resist sensitivity (S)}}$$

To the first approximation, the current density can be given by:

$$j = \beta \pi \alpha^2$$

where β is the gun brightness in Amp/cm^2 steradian
α is the beam semi convergence angle in radian

This approximation assumes that the beam half-amplitude width is about equal to the stepping distance and that lens and deflection aberrations are negligible.

In V.S.I, a high brightness lanthanum hexaboride gun is used with typical brightness value (β) at 25kV of 1×10^6 Amp/cm^2 steradian. The beam semi-convergence angle (α) has been experimentally established[1] to be $\geq 1 \times 10^{-2}$ radian for a 2000 line field without the use of dynamic corrections. Using these values, the operable beam stepping rates for different resist sensitivities are as shown in Table 3.

TABLE 3

$\beta = 1 \times 10^{6}$ Amp/cm^2 steradian

$\alpha = 1 \times 10^{-2}$ radian

$j = 300$ Amp/cm^2

Resist sensitivity (C/cm^2)	1×10^{-5}	1×10^{-6}	1×10^{-7}
Stepping rate (operable)	30 MHz	300 MHz	3000 MHz

The choice of resist will depend on the application. For processes requiring metal lift-off, a positive resist has to be used. This requires a relatively high exposure dosage to give a proper resist profile and $S = 1 \times 10^{-5}$C/cm^2 is a typical workable value[22]. For subtractive etching processes, both positive and negative resists can be used and a much lower exposure dosage in the 10^{-6} to 10^{-7}C/cm^2 range, has been found acceptable[23,24]. Taking these values, one can see that the operable beam stepping rate can be at least 30 MHz and can be as high as a few thousand MHz.

The deflection system used in V.S.I is described in some detail in the accompanying paper[1] and will not be repeated here. The two main areas in this system that can impact the operable beam stepping rate are the D/A converters and the deflection driver (plus coil). Experimental results have been achieved operating the system at a stepping rate equal to 10 MHz. Fig.4 and Fig.5 show test exposures of two levels of the 8K FET memory pattern operating at this rate. In Fig.4 the smallest pattern linewidth is approximately 1μm and in Fig.5 the smallest gap between pattern line is also approximately 1μm. It can be seen that in both cases the patterns show good edge and corner definition. No loss of pattern integrity was observed when comparing these results at 10 MHz rate to similar results at a much lower stepping rate (e.g. 1 MHz). It is difficult at the present time to establish the ultimate operation limits of such system, since progress is still being made in the design of better D/As, deflection drivers and coils. It is therefore reasonable to expect that much higher stepping should be attainable in future.

Overheads

Fig.6 shows the two major overhead times associated with the vector scan technique. The first, T_{jo}, is the settling time related to jumps between pattern cells. The requirement is that the beam should settle to less than ½LSB before the start of the fill-in scan. The main factors affecting this overhead are;(1) the deflection driver settling (i.e. the large signal response characteristic of the amplifier-coil combination. This time is related to jump amplitude, i.e. larger jump distances will require longer settling times.) (2) eddy current effects caused by flux linkage of the deflection field to conducting media adjacent to the deflection coils (The settling time of this effect is also jump amplitude dependent and can be very long, typically in excess of tens of milliseconds), (3) the transfer time of pattern data from the computer memory to the pattern generator (Ideally, the data transfer for the next pattern cell should be completed during the exposure of the previous cell. To meet this requirement at extreme conditions, i.e. high exposure speed and small pattern cell size, the data transfer time should be as short as a few microseconds.)

In V.S.I the eddy current effects have been suitably suppressed[1] and can be regarded as negligible. The data transfer time can be resolved by the use of a high speed data buffer in the pattern generator. One such buffer with a relatively small capacity (256 words) has been instrumented in the system, and by increasing the size of this buffer, it is anticipated that data transfer will not constitute a significant overhead factor. The main settling time overhead is therefore contributed by the deflection driver. In V.S.I, the deflection driver-coil combination has a large signal response characteristic that should, in the limit, allow settling time to ½LSB in the 14 bit field in about 20μs. At the present time, however, a delay for a full field jump of approximately 100μs is generally used. A programmable delay time feature has been instrumented in the pattern generator and 16 delay values can be applied for the different jump distances to minimize the total settling time overhead.

The second overhead time, T_{fo}, relates to the functioning of the digital circuits in the pattern generator for the fill-in scans. For example, in the spiral scan technique, at the closing of each frame scan, some waiting time is needed for the logic circuitry to compute and implement the operation of the second inner frame and during this time the beam is blanked off. This overhead time is dependent upon the number of scan segments used and is therefore proportional to the size of the pattern cell and larger cells will require longer overheads. The actual time needed for this overhead will depend on the design of the circuit and the logic components used. In V.S.I an overhead time for each scan line segment is approximately 0.1μ sec. for pattern generator using hardwired logic. Much shorter overheads are considered possible in future developments.

The total overhead time T_o per chip can therefore be expressed as:

$$T_o = M\ (\ \overline{T}_{jo} + \overline{T}_{fo}\ )$$

where $\overline{T}_{jo}$ = Av. jump settling time/cell

$\overline{T}_{fo}$ = Av. fill-in logic overhead/cell

M = Number of pattern cells in chip

Total Writing Time/Chip

The total writing time per chip can be obtained by summing the two main components, the exposure time (T_e) and the overhead (T_o), i.e.

$$T_w = T_e + T_o$$

Practical Example

As a practical demonstration of the performance of the system in the exposure of a complex L.S.I circuit, typical values for the exposure of an 8K FET memory chip is taken here as an example.

Fig.7 shows a histogram of the distribution of jump distances and cell sizes for the A level pattern of the 8K memory chip. It can be seen that the large majority of jump distance between cells is less than 1% of the full field. However, as the present pattern exposure routine is not fully optimized for the vector scan operation, some small percentage of large jumps still exist. This has brought the average jump distance for the full chip to 2.8%. It is anticipated that improvement can be achieved by a better sorting routine. The cell size is expressed as a percentage to the maximum cell size which is 512 units wide in the system. Again the majority of the values is found to be small, <1%, and the average cell size is 0.7% of the maximum size. Fig.7 also shows the average value of jump and cell size for all four pattern levels of the 8K chip. The combined average values for the four levels are:

Average jump distance/cell = 4% full field
Average cell size = 1.5% max. cell (512 x 512)
Average no. of cells/level = 26,000
Average pattern area coverage = 22%

Using these values, and taking a conservative performance of 100μs settling time for full field jump and 10μs fill-in overhead for each cell, the total overhead can be computed. The exposure time is computed based on 10 MHz stepping rate.

Av. T_{jo}/cell @ 4% full field jump = ~4μs

Av. T_{fo}/cell @ 1.5% max. cell = ~10μs

(T_{jo} + T_{fo})/cell = 14μs

1. Av. overheads/chip = (14 x 26000)μs = 0.36 sec.

2. Av. exposure time (@ 10 MHz)/chip = 0.22 x 6.4 = 1.4 sec.

3. Writing Time/Chip = Exposure Time + Overheads/Chip
= 1.4 + 0.36
= 1.76 sec.

The result shown here is for a high density FET type memory chip and for other devices, the same computational technique has to be applied to evaluate the exposure speed. In general, one may say that the example used here is probably an aggressive representation of average L.S.I. circuits used to-day and therefore the exposure time of close to 2 sec./chip is perhaps a high value and for most other L.S.I. one expects the exposure time/chip to be somewhat less than this value.

The 10 MHz writing speed and the various overheads mentioned earlier reflect the initial test performance of the system and does not represent the ultimate capability. With the improvement in deflection electronics and with the development of new exposure techniques such as multiple beam size as discussed earlier, a significant reduction in writing time is expected.

EXPERIMENTAL RESULTS

The results of experimental devices fabricated using the system have been reported previously[15-17]. They include a number of magnetic bubble chips with different memory sizes for 2μm and 1μm bubble experiments[14]. Minimum linewidth down to 0.25μm has been demonstrated[15]. Both direct exposure and mask making for x-ray[11] and deep U.V.[12] has been performed. Fig.8 shows an example of a 16K bit bubble chip for 2μm bubble and Fig.9 shows details of the same chip demonstrating the importance of proximity effect corrections.

The fabrication of an operational 8K FET memory chip has been reported by Yu[13] et al. The channel length of the FET described in that report is approximately 1.25μm and a typical example is shown in Fig.10. More recently, an experiment has been performed to study the lithographic capability of the V.S.I system, by reducing the linear geometry of the 8K chip by another factor of 2, and the minimum pattern linewidth in this case is approximately 0.6μm. All four pattern levels have been exposed with critical registrations and all relevent processing applied to ensure good structural definition. The result is shown in Fig.11 with a fan-out contact pad pattern. The experiment demonstrated satisfactory lithographic performance for devices of this dimension (∿0.5μm) and registration accuracy of ±0.1μm is consistently achieved.

CONCLUSION

The vector scan electron beam system (V.S.I) has demonstrated the following achievements.

Firstly, the system has demonstrated a high resolution lithographic capability that can be extended into the submicron region for future L.S.I.

Secondly, the system is fully automated and can be used for direct wafer exposure and for mask making. It has an extensive pattern cell shape selection and the ability to compensate for proximity effect. All these have given the system a high degree of flexibility.

Thirdly, the system has demonstrated the capability of high through-put. Test exposures have been made with writing time/chip of 1 to 2 sec. It has been shown that significant further improvement in speed is possible by improving the deflection electronics and by implementing new exposure techniques such as multiple beam sizes technique, etc.

Fourthly, a chip size equal to 2000 x minimum pattern linewidth has been achieved without the use of dynamic corrections. Experiments have been performed to show that much larger chip size can be achieved by field stitching technique using registration marks. Similar stitching results are expected by the use of laser interferometric controlled stage currently under development.

Finally, the system is relatively simple and thus offers the potential of a low cost technology.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the technical contributions of W.W. Blair, T.P. Donohue, G.A. Folchi, P.R. Jaskar, W.L. Keller, A.Kraft, A. Kern, H.E. Luhn, E. Munro, M. Parikh, A.M. Patlach, T.W. Studwell and R. Viswanathan in the development of V.S.I. Encouragement and helpful discussions with A.N. Broers and M. Hatzakis are deeply appreciated. The FET memory work was performed in cooperation with silicon technology groups under R.H. Dennard, C.M. Osburn and H.N. Yu; and the magnetic bubble memory experiments were performed in cooperation with the magnetic group under M.H. Kryder at the IBM Thomas J. Watson Research Center.

REFERENCES

1. T.H.P. Chang, A.J. Speth, C.H. Ting, R. Viswanathan, M. Parikh and E. Munro, "The Probe Forming and Deflection System for Vector Scan I E/B Lithographic System", this issue.

2. A.D. Wilson, A. Kern, A.J. Speth, A.M. Patlach, P.R. Jaskar, T.W. Studwell and W.L. Keller, "Automation of Vector Scan I Electron Beam Lithographic System", this issue.

3. G.L. Varnell, D.F. Spicer, A.C. Rodger and R.D. Holland, Proc. 6th Internat. Conf. on Electron & Ion Beam Sci. & Technol., ed. by R. Bakish (The Electrochem. Soc. Princeton 1974), p.97.

4. A.N. Broers and M. Hatzakis, Scientific American 227, 34 (Nov.1972).

5. D.R. Herriott, R.J. Collier, D.S. Alles & J.W. Stafford, IEEE Trans. Electron Devices ED-22, 385 (1975).

6. F.S. Ozdemir, C.R. Buckey and E.D. Wolf, J. Vac. Sci. Technol. 12, 1246 (1975)

7. J.P. Beasley & D.G. Squire, IEEE Trans. Electron Devices, ED-22, 376 (1975).

8. H.C. Pfeiffer, J. Vac. Sci. Technol. 12, 1170 (1975).

9. B.P. Piwczyk and K.G. McQuhae, J. Vac. Sci. Technol. 10, 1016 (1973).

10. O. Cahen, R. Sigelle & J. Trotel, Proc. 5th Internal. Conf. on Electron & Ion Beam Sci. & Technol., ed. by R. Bakish (The Electrochem. Soc., Princeton, 1972), p.92.

11. R. Feder, E. Spiller & J. Topalian, J. Vac. Sci. Technol. 12, 1332, (1975).

12. B.J. Lin, J. Vac. Sci. Technol. 12, 1317 (1975).

13. H.N. Yu, R.H. Dennard, T.H.P. Chang, C.M. Osburn, V. Dilonards & H.E. Luhn, J. Vac. Sci. Technol., 12, 1297 (1975).

14. M.H. Kryder, K.Y. Ahn & J.V. Powers, IEEE Trans. Magnetics, MAG-11 1145 (1974).

15. A.J. Speth, A.D. Wilson, A. Kern & T.H.P. Chang, J. Vac. Sci. Technol., 12, 1235 (1975).

16. A.D. Wilson, T.H.P. Chang & A. Kern, J. Vac. Sci. Technol., 12, 1240 (1975).

17. A.D. Wilson, T.H.P. Chang, A.J. Speth & A. Kern, Scanning Electron Microscopy/1976 (Part IV) IIT Research Institute, Chicago, P.659.

18. A.N. Broers, J. Sci. Instr. Ser. 2, 2, (1969).

19. T.H.P. Chang, Proc. 8th Internal. Congress on Electron Microscopy, Canberra, Australia, 1974, Vol. 1, p.650.

20. T.H.P. Chang, J. Vac. Sci. Technol. 12, 1271 (1975).

21. D.F. Spicer, A.C. Roger & G.L. Varnell, J. Vac. Sci. Technol. 10, 1052 (1973).

22. M. Hatzakis, J. Vac. Sci. Technol. 12, 1276 (1975) and private communications.

23. L.F. Thompson, J.P. Ballantyne & E.D. Feit, J. Vac. Sci. Technol. 12, 1280 (1975).

24. M.J. Bowden, L.F. Thompson & J.P. Ballantyne, J. Vac. Sci. Technol. 12, 1294, (1975).

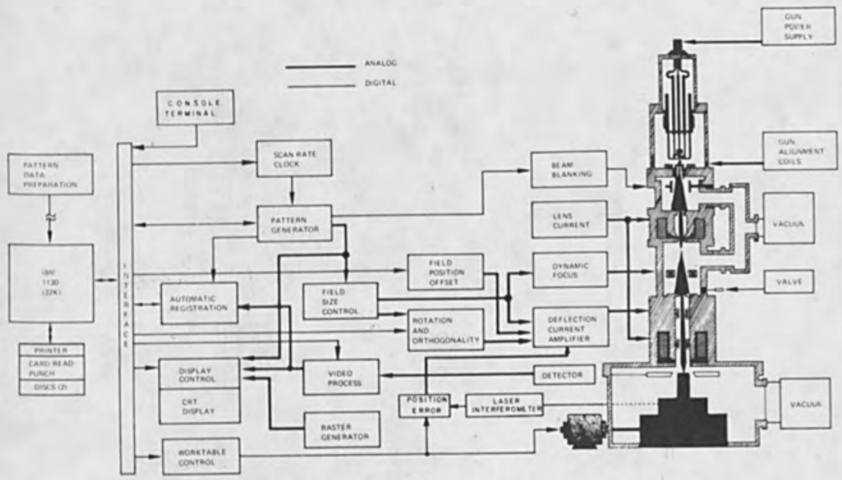


Figure 1. Block diagram of the system

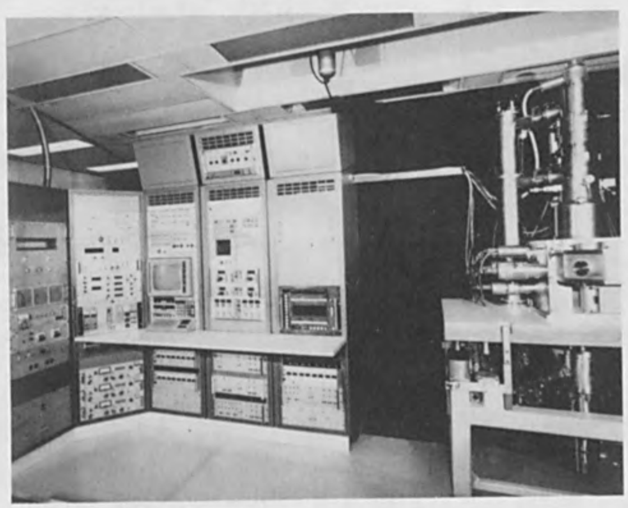
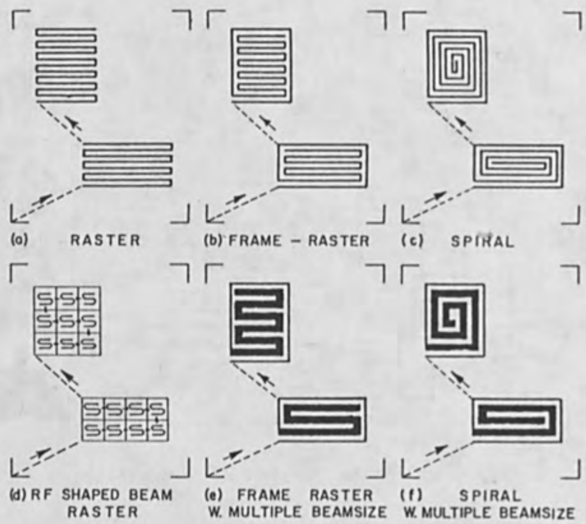


Figure 2. General view of the system



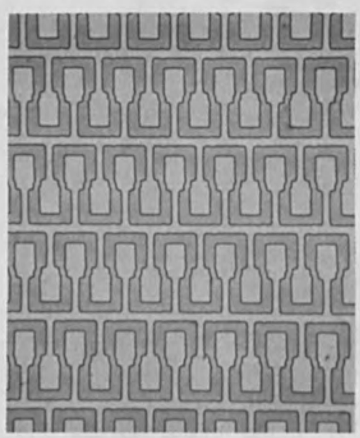


Figure 4. Level A (∿ 1 μ lines)

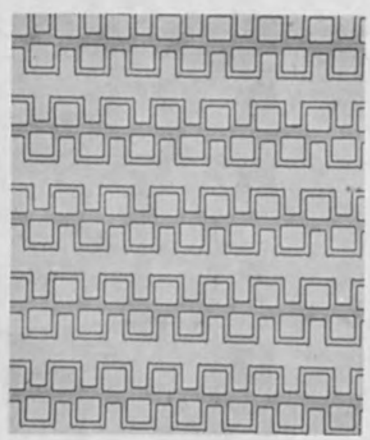


Figure 5. Level B (∿ 1 μ gaps)

10 MHz test exposures of 8K FET memory chips

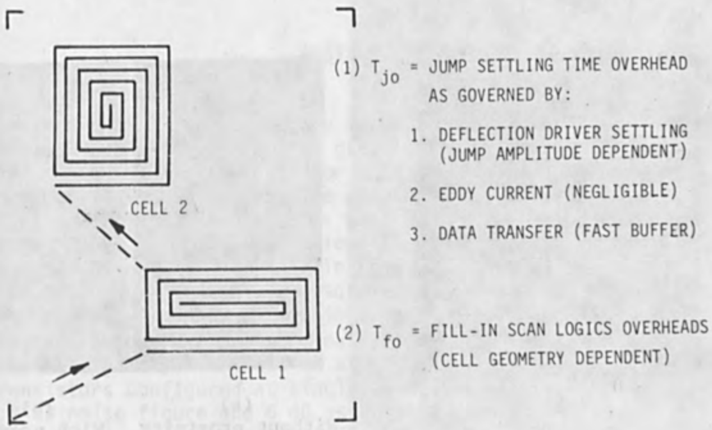


Figure 6. Overheads in vector scan technique.

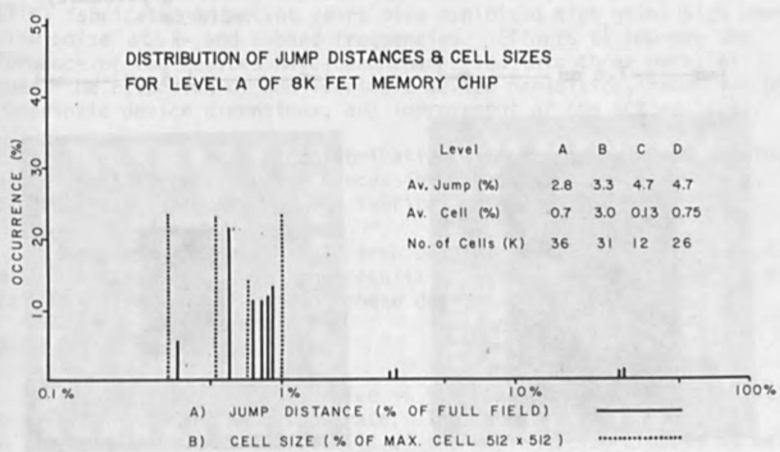


Figure 7. Histogram showing distributions of jump distances and cell sizes

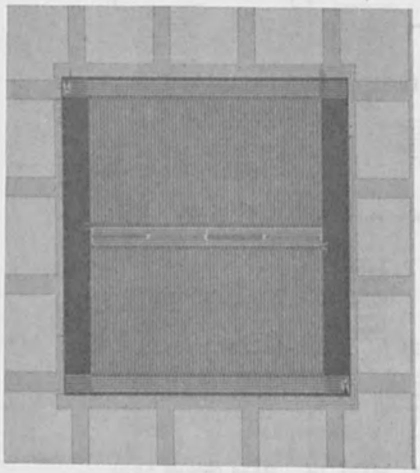
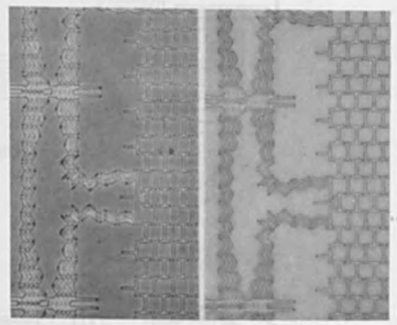


Figure 8. 16K bubble memory chip (min. linewidth = 1 μ)



(a) Without proximity corrections

(b) With proximity corrections

Figure 9. Details of bubble chip showing proximity effects

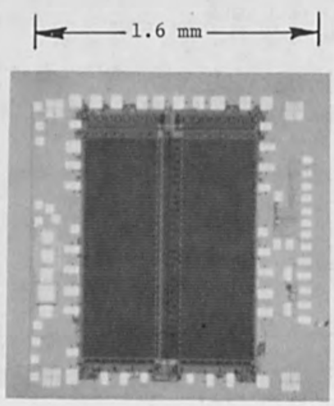


Figure 10. 8K FET memory chip (min. linewidth = 1.25 μ)

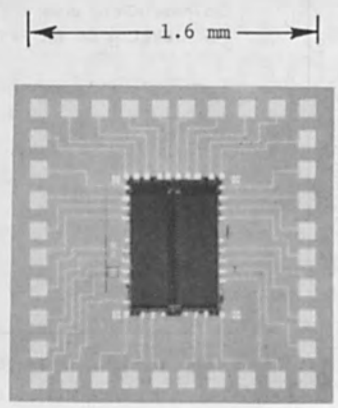


Figure 11. Half size 8K chip for lithography tests only (min. linewidth = 0.6 μ)

ELECTRON BEAM MICROFABRICATION OF GaAs SBFET's

F.S. Ozdemir, G.O.Ladd, D.D.Loper, F.W.Cleary, N.Hirsch
Hughes Research Laboratories
Malibu, California 90265

ABSTRACT

Electron beam microfabrication techniques were used to fabricate GaAs SBFET's with 0.5 μm gate lengths. The active layer was created by ion implantation into a semi-insulating GaAs substrate. The mesa structure and the final source/drain overlay metallization patterns were defined by standard photolithography. Electron beam microfabrication technqiues were used to define the source/drain and gate patterns. Low and high frequency electrical measurements have been performed. Low frequency g_m's of 80 millimhos per mm of gate width have been measured. S-parameter measurements indicate a maximum available gain of 14 to 15 dB at 10 GHz and 10.8 dB at 14 GHz. Transistors configured as single stage amplifiers have shown 5.9 dB noise figure and 6 dB associated gain at 14 GHz.

INTRODUCTION

Gallium arsenide Schottky barrier field effect transistors (GaAs SBFET's) fabricated in recent years have exhibited high gain, high power and low noise at X- and K-band frequencies. Efforts to improve the performance of the devices have been channelled into three parallel avenues: the reduction of non-intrinsic device parasitics, reduction of the intrinsic device dimensions, and improvement of the active layer.

Direct electron beam microfabrication,with which very high resolution, multi-mask semiconductor processing can and has been performed, is particularly applicable to the fabrication of GaAs SBFET's.

This paper describes the electron beam microfabrication of 0.5 μm gate length GaAs SBFET's and the results of low and high frequency electrical measurements performed on these devices.

FABRICATION

The active layer for the device is formed by sulfur ion implantation into a semi-insulating GaAs substrate. The substrate is then coated with SiO_2 and annealed. The details of the implantation process as well as some preliminary results obtained from photolithographically fabricated GaAs SBFET's have been described elsewhere.[1] After the anneal cycle the SiO_2 is removed and the carrier concentration vs depth profile is determined by capacitance vs voltage measurements made on a Schottky barrier diode. A typical profile is shown in Fig. 1.

The first lithography step is the definition of the mesa pattern by standard photolithographic techniques. The active layer outside the mesa regions is removed by wet etching. The outline of the mesa is shown in Fig. 2a. The overall dimensions of the mesa are 60 μm by 330 μm. The second lithography step is also optical and an L-shaped alignment mark pattern is defined. Lift-off Au metallization is used to fabricate this optical alignment mark which is also shown in Fig. 2a.

The third lithography step is the electron beam definition of the source/drain pattern. The wafer is coated with a positive electron resist (PMMA) and the computer-controlled electron beam is used to automatically align the electrical scan field of the beam to the mesa structure using the optically defined alignment mark. The accuracy of this alignment is not critical and an alignment error of ± 0.5 μm is quite acceptable. After the alignment process the source/drain and three additional alignment mark patterns are exposed on the substrate. Thus these new alignment marks are self-registered to the source/drain pattern. The electron beam microfabrication system used to perform these alignment and exposure steps has been described in a previous publication.[2] The details of the registration are also the subject of an earlier publication.[3] The source/drain and alignment mark patterns are shown in Fig. 2b. The channel width is 300 μm and the source/drain separation is 2.5 μm. The resist is then developed and a source/drain metallization consisting of Au/Ge/Ni is evaporated and lifted off.

The fourth lithography step is the electron beam definition of the gate pattern. The PMMA resist is again applied to the wafer and the alignment process repeated, this time using the alignment marks defined during the source/drain exposure. The alignment accuracy desired in this step is ± 0.1 μm, while errors up to ± 0.2 μm are considered acceptable. The gate pattern length is 0.5 μm and the width is 310 μm which overlaps the 300 μm channel width at both extremities of the channel. Contact to the gate structure is made at two points as shown in Fig. 2c which serves to reduce to series gate resistance. The gate metallization is Al and is defined by the same evaporation/lift-off process used for the source/drain metallization.

The fifth, and last, lithography step is the source/drain overlay metallization defined by standard photolithography in positive photoresist. Evaporation and lift-off of Au forms this overlay metallization which serves to reduce the series source/drain resistance of the device. The overlay source/drain metallization is shown in Fig. 2d.

ELECTRICAL MEASUREMENTS

I-V characteristics of the devices are evaluated at the wafer level with a curve tracer. A typical set of characteristics is shown in Fig. 3. The devices saturate at 1.3 volts and have a pinchoff voltage of

4.5 volts. The low frequency transconductance of the device is 80 millimhos per millimeter of channel width.

The wafer is then thinned to 0.008" mils, diced into 0.016" by 0.020" chips and mounted on an unsealed microstrip carrier for rf measurements. Au wire is then thermocompression bonded to the source/drain and gate contact areas. A scanning electron microscope photomicrograph of the channel area of the device is given in Fig. 4.

With the device mounted in the carrier the S parameters are measured over the range from 2 to 18 GHz using an automatic network analyzer. Noise figure has been measured by mounting the chip in a low loss microstrip amplifier circuit. An amplifier noise figure of 5.9 dB with 6 dB associated gain has been measured at 14 GHz. Network analyzer measurements indicate a maximum available gain (MAG) of 10.8 dB at the same frequency. At lower frequencies, for example, at 10 GHz, the MAG is 14 to 15 dB. These results are preliminary and a full characterization of the device will be presented at a later date.

CONCLUSIONS

The fabrication and electrical measurement results of a 0.5 μm gate GaAs SBFET have been described. The device was fabricated on an active layer created by ion implantation. Low resolution lithography steps were performed by photolithography while high resolution lithography steps utilized electron beam microfabrication.

Results obtained on a first-iteration device fabricated by these techniques have been presented in the past.[4] The work described in this paper represents preliminary results on a second-iteration device. Work is continuing to further optimize electron-beam-fabricated GaAs SBFET's to obtain lower noise figures and higher gains at frequencies well above 10 GHz.

ACKNOWLEDGEMENTS

The authors would like to thank D. M. Jamba for performing the ion implantation, R. L. Pierson for mounting and bonding the devices, P. J. Coane for Figure 5, P. V. Cooper for the noise figure measurement, and the Hughes Radar Microwave Laboratory for the S-parameter measurements.

REFERENCES

1. R. G. Hunsperger and N. Hirsch, "Ion-Implanted Microwave Field-Effect Transistors in GaAs," Solid-State Electronics, 18, 349-353 (1975).

2. F. S. Ozdemir, W. E. Perkins, R. Yim, E. D. Wolf, "Precision Electron Beam Microfabrication," J. Vac. Sci. Tech. 10, 1008-1011 (1973).

3. F. S. Ozdemir, E. D. Wolf, C. R. Buckey, "Computer-Controlled Scanning Electron Microscope System for High-Resolution Microelectronic Pattern Fabrication," IEEE Trans. Elec. Dev. ED-19, 624-628 (1972).

4. F. S. Ozdemir, G. O. Ladd, E. D. Wolf, W. E. Perkins, N. Hirsch, F. W. Cleary, "Electron Beam Fabricated 0.5 μm Gate GaAs Schottky-Barrier Field Effect Transistor," Paper No. 13.7, International Electron Devices Meeting (1974).

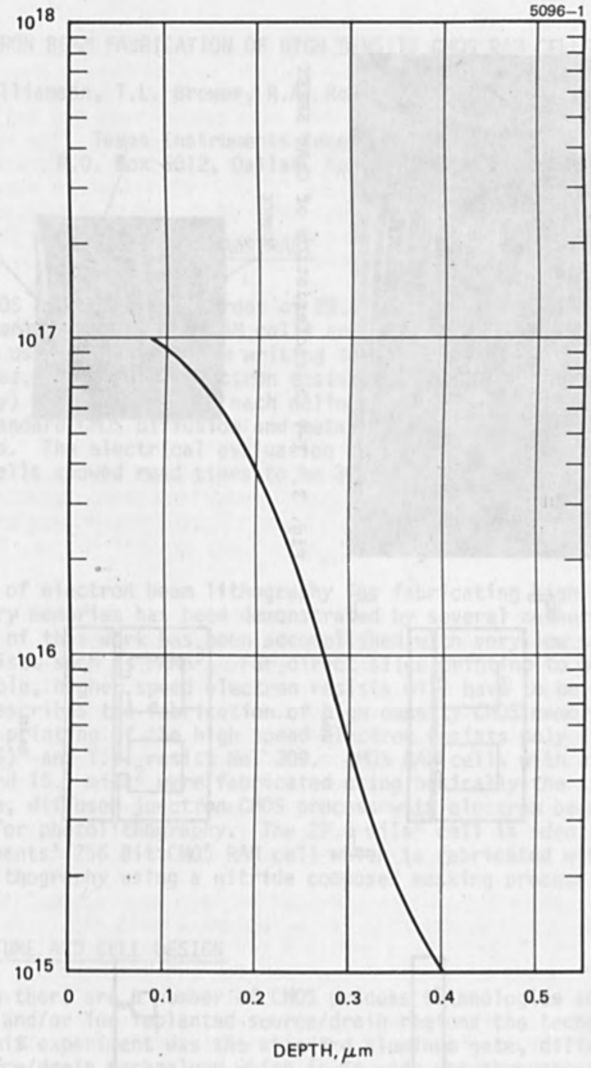


Fig. 1 Carrier Concentration Profile vs Depth

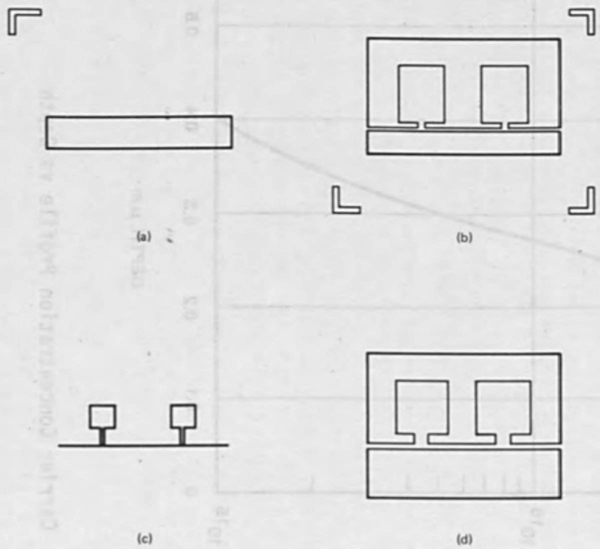


Fig. 2. Optical and Electron Beam Patterns for GaAs SBFET a) Mesa, b) Source/Drain, c) Gate, d) Overlay

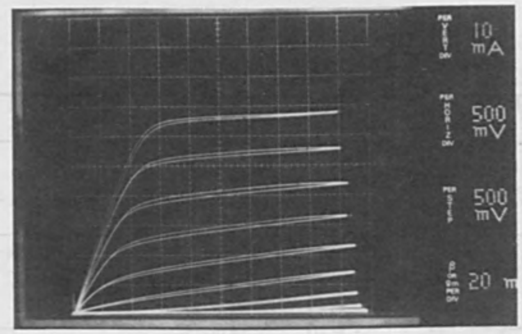


Fig. 3. I-V Characteristics of GaAs SBFET

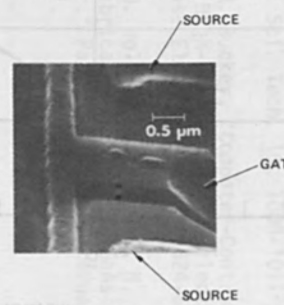


Fig. 4. Scanning Electron Microscope Photomicrograph of Device

ELECTRON BEAM FABRICATION OF HIGH DENSITY CMOS RAM CELLS*

R.A. Williamson, T.L. Brewer, R.A. Robbins, and G.L. Varnell

Texas Instruments Incorporated
P.O. Box 5012, Dallas, Texas 75222

ABSTRACT

CMOS RAM cells with areas of 29.0 $mils^2$ (identical to Texas Instruments 256 bit CMOS RAM cell) and 15.1 $mils^2$ have been fabricated using electron beam writing techniques for all lithographic processes. High speed electron resists ($\leq 3X10^{-6}$ $coul/cm^2$ sensitivity) were utilized at each delineation step. Compatibility with standard CMOS diffusion and metallization processes has been achieved. The electrical evaluation at V_{DD} = 5V of the large and small cells showed read times to be 300ns and 400ns respectively.

INTRODUCTION

The use of electron beam lithography for fabricating high density, small geometry memories has been demonstrated by several authors[1,2] but the majority of this work has been accomplished with very low sensitivity electron resists such as PMMA[3]. For direct slice printing to become economically feasible, higher speed electron resists will have to be utilized. This paper describes the fabrication of high density CMOS memory cells by direct slice printing of the high speed electron resists poly (butene-1 sulfone) (PBS)[4] and T.I. resist No. 309. CMOS RAM cells with areas of 29.0 $mils^2$ and 15.1 $mils^2$ were fabricated using basically the standard aluminum gate, diffused junction CMOS process with electron beam lithography substituted for photolithography. The 29.0 $mils^2$ cell is identical to Texas Instruments' 256 Bit CMOS RAM cell which is fabricated with conventional photolithography using a nitride composed masking process.

DEVICE STRUCTURE AND CELL DESIGN

Although there are a number of CMOS process technologies such as silicon gate and/or ion implanted source/drain regions the technology chosen for this experiment was the standard aluminum gate, diffused guard-ring and source/drain technology which is in wide use throughout the semiconductor industry. A drawing of the structure achieved using this

* Portions of this work were supported by the U.S. Army Electronics Command Ft. Monmouth, N.J., Contract #DAAB07-75-C-1297.

technology is shown in Figure 1.

The basic memory cell consists of two cross-coupled inverters. The inverters are asymmetrical so that data can be read and written easily. A schematic of the memory cell is shown in Figure 2. The ratio of the widths of Q_1 and Q_2 to Q_5 and Q_6 must be such that the state of the memory cell is not changed when both the "1" and "0" lines are shorted to ground (V_{SS}). The ratio of the widths of Q_5 and Q_6 to Q_3 and Q_4 must be such that the state of the memory cell can be modified by taking either the "1" or "0" line to the supply voltage (V_{DD}).

The layout rules used in designing the two cells are shown listed in Table 1. The basic differences in the layout rules used in designing the two cells are geometry sizes and metal spacings. The diffusion spacings are the same for both cells because they have the same supply voltage requirements. Using these two sets of rules, memory cells of 29.0 mils^2 (Cell 1) and 15.1 mils^2 (Cell 2) were designed. Four cells of each size were designed in 2X2 arrays in the chip layout. All the "1", "0", column select (CS) and power supply lines were brought out to individual bonding pads. Resistor/diode input protection devices were connected to the column select lines to prevent static charge from destroying the gate oxide.

PROCESSING

The standard aluminum gate CMOS process consists of six lithographic steps: P-, P+, N+, gate, contact and leads. An additional lithographic step was added for delineating the electron beam alignment markers. The basic process used in fabricating the memory cells is shown in Table 2. The starting material was 2", (100) N-type, 3.0 - 3.5 Ω-cm silicon. All doping procedures were conventional except $P-_o$ which was implanted. Since the electron resists were thinner (~4000-6000Å) than normal photoresists, oxide thicknesses and steps were minimized to prevent any resist coverage problems. The thinner oxides did not create any field turn-on problems even at V_{DD} = 15V since guardrings were used throughout.

The lithographic steps shown in Table 3 were accomplished with two high speed electron resists: poly (butene-1 sulfone), a positive resist and T.I. resist No. 309, a negative resist. PBS and T.I. 309 both were exposed at doses of 1-3X10^{-6} coul/cm^2 for all lithographic steps. All baking in the resist processing was performed either in air or vacuum depending on the resist and process step. As shown in Table 3, all etch procedures were conventional except at N+ where plasma etching was used to insure proper oxide geometry control.

All exposures and alignments were accomplished on an electron beam writing instrument (EBMIII) developed at Texas Instruments. This instrument is similar to EBMII[5] but has several additional features designed specifically for slice printing. Fiducial markers were placed in the four corners of each bar (see Figure 3) in order to achieve precise, automated pattern registration. The alignment system on EBMIII produces level to level registration of better than ± 0.01 mils.

RESULTS

Working memory cells of 29.0 mils^2 and 15.1 mils^2 have been fabricated and photographs of each group of four cells are shown in Figures 4 and 5. The resolution of all geometries, including the 0.1 mil X 0.2 mil contacts, is excellent (see Figure 6). The most difficult processing step was lead definition where resist bridging sometimes occurred at the smallest spacing. However with proper exposure and developing conditions the 0.1 mil spacings were routinely cleared without any bridging. Figure 7 is a 10,000X SEM photograph of a test geometry used to monitor exposure and etching of the lead pattern. The metal stripes are 0.1 mil wide with 0.1 mil spacing.

A special test exerciser was fabricated to electrically evaluate the memory cells. This exerciser permits the reading and writing of information using the same input lines. During testing a 100K Ω resistor connects each "1" and "0" line to ground (V_{SS}). All testing was performed at V_{DD} = 5.0V.

Figure 8 shows an oscilliscope photograph of the "1" and "0" lines (inputs and outputs) and the column select (enable) line. Figure 9 is a read/write test timing diagram corresponding to the oscilliscope photograph. The very narrow column select write pulse is almost impossible to observe on this scale. To actually measure the read time (τ_R) and the write time (τ_W) the time scale is expanded and the relevant portion of the waveform is examined. The results for the two cells are shown in Table 4. The difference in A.C. response of the two cells was not unexpected since the W/L ratios were smaller in the 15.1 mils^2 cell. The data on Cell 1 compares favorably with data taken from the identical cell fabricated with nitride composed masking and conventional photolithography.

CV analysis showed all electron beam induced damage has all been completely removed by annealing which is in agreement with the results obtained by other[6] laboratories.

CONCLUSIONS

The feasibility of using resists at least 20 times faster than PMMA for direct electron beam slice printing of high density memory devices has been demonstrated. Because of the superior geometry resolution and level to level resolution (better than ± 0.01 mil) of the electron beam writing system at least a fifty percent cell size reduction can be achieved on CMOS memories with only the substitution of electron beam lithography for conventional photolithography. Combining electron beam lithography, with all its inherent advantages, and other advanced processing technologies such as silicon gate and ion implanted source/drains can lead to cell sizes that are substantially smaller than those achieved to date.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the significant contributions made by C.D. Winborn. Acknowledgements also go to A.R. Reinberg, J.E. Hull, J.M. Caywood, Imogene Gibson and S.R. Schenck.

REFERENCES

1. R.C. Henderson, R.F.W. Pease, A.M. Voshchevkow, P. Mallery and R.L. Wadsack, IEEE J. Solid State Circuits SC-10, 92 (1975).

2. H.N. Yu, R.H. Dennard, T.H.P. Chang, C.M. Osburn, V. Dilonardo, and H.E. Luhn, J. Vac. Sci, Technol. 12, 1297 (1975).

3. I. Haller, P. Srinivasan, and M. Hatzakis, IBM J. Res. Dev., 12, 251 (1968).

4. L.F. Thompson and M.J. Bowden, J. Electrochem. Soc., 120, 1723 (1974).

5. G.L. Varnell, D.F. Spicer, A.C. Rodger, and R.D. Holland, "High Speed Electron Beam Pattern Generator", Sixth International Conference on Electron and Ion Beam Science and Technology, San Francisco, California, May 1974.

6. K.A. Picker and L.R. Thibault, "Electron Irradiation of MIS Capacitors-Annealing and Bias-Temperature Stress Studies", Fifth International Conference on Electron and Ion Beam Science and Technology, Houston, Texas, May 1972.

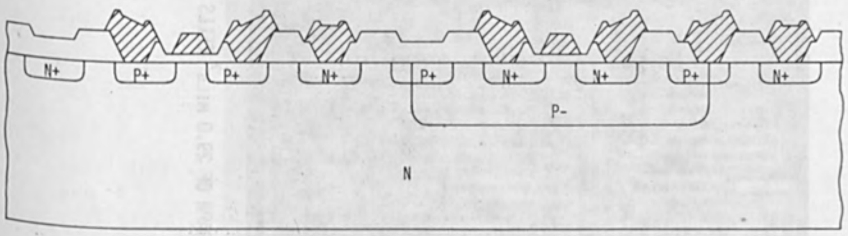


FIGURE 1: COMPLEMENTARY MOS DEVICE STRUCTURE

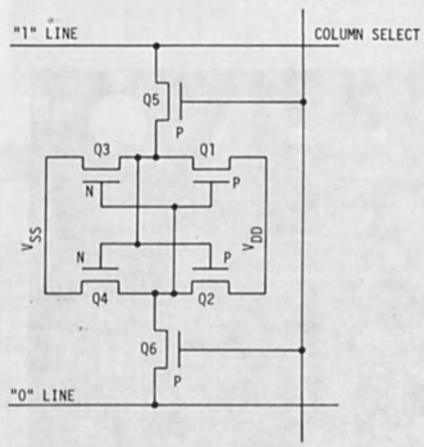


FIGURE 2: CMOS MEMORY CELL

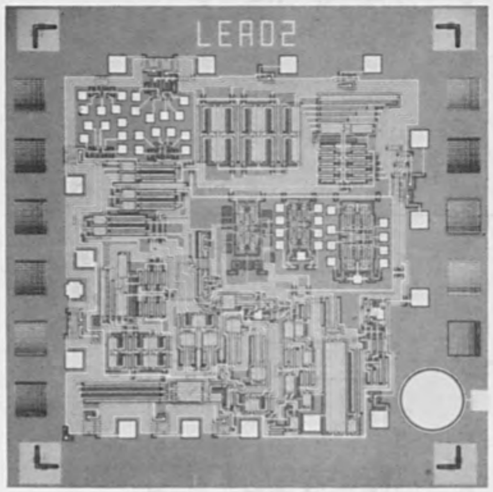


FIGURE 3: BAR PHOTOGRAPH SHOWING FIDUCIAL MARKERS

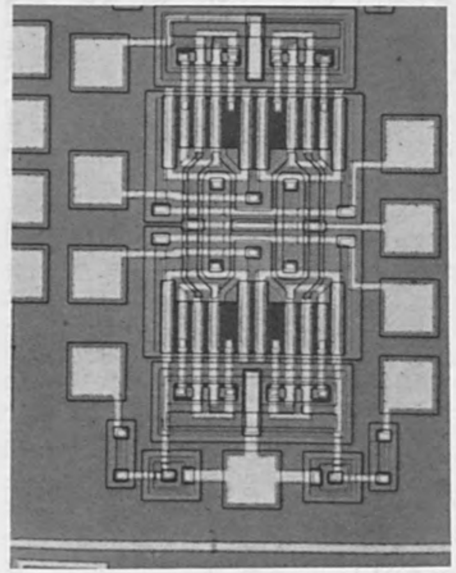


FIGURE 4: PHOTOGRAPH OF 29.0 $MILS^2$ CELLS

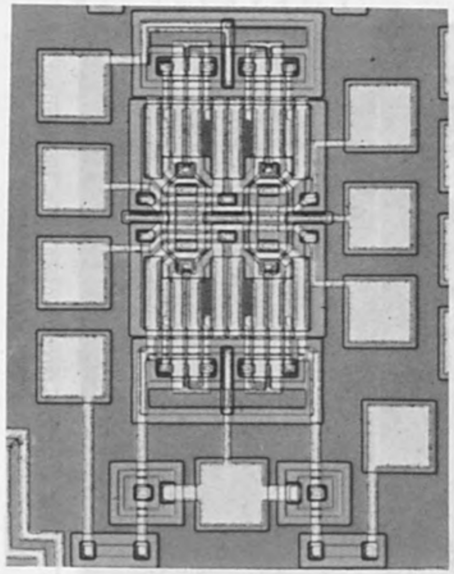


FIGURE 5: PHOTOGRAPH OF 15.1 MILS^2 CELLS

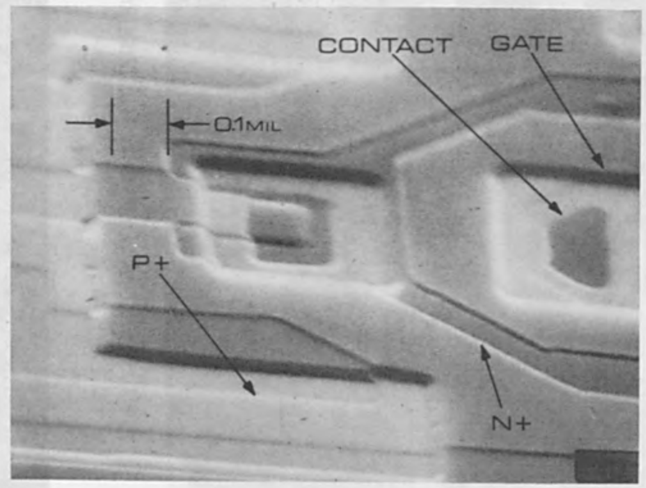


FIGURE 6: 3000X SEM OF CELL AREA AFTER CONTACT O.R.

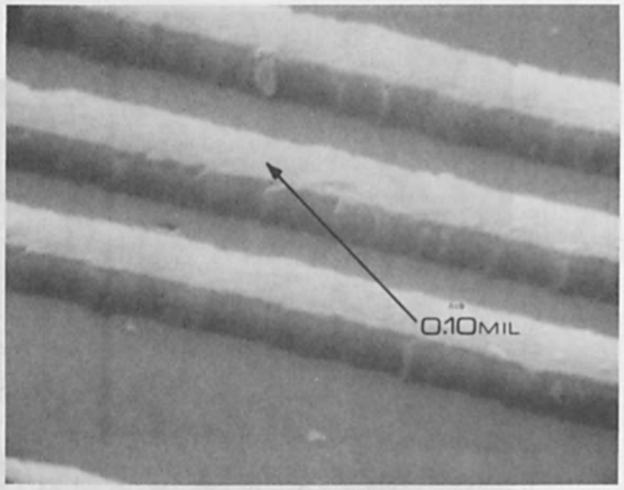


FIGURE 7: 10,000X SEM OF METAL TEST GEOMETRY

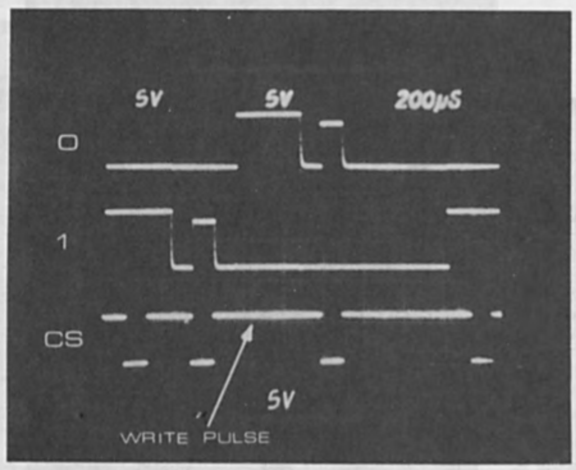


FIGURE 8: OSCILLISCOPE PHOTOGRAPH OF INPUT/OUTPUT
AND COLUMN SELECT WAVEFORMS

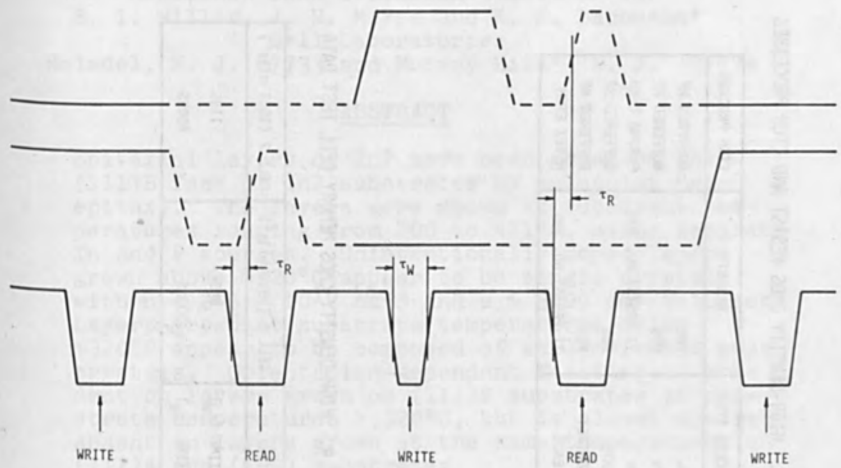


FIGURE 9: READ/WRITE TEST TIMING DIAGRAM

TABLE 1. HIGH-DENSITY CMOS RAM CELL LAYOUT RULES

ITEM	CELL 1 (29.0 mils^2)	CELL 2 (15.1 mils^2)
SOURCE-TO-DRAIN SPACING	0.3 mil	0.3 mil
N+-TO-P+ SPACING	0.3 mil	0.3 mil
METAL-TO-METAL SPACING	0.25 mil	0.1 mil
CONTACT-TO-DIFFUSION SPACING, N+	0.1 mil	0.1 mil
CONTACT-TO-DIFFUSION SPACING, P+	0.1 mil	0.05 mil
MINIMUM METAL WIDTH	0.3 mil	0.2 mil
MINIMUM O.R. WIDTH	0.2 mil	0.1 mil
MINIMUM CONTACT SIZE	0.2 mil X 0.3 mil	0.1 mil X 0.2 mil
P+, P- OVERLAP	0.1 mil	0.0 mil

TABLE 2. HIGH-DENSITY CMOS DEVICE PROCESS

STEPS

1. GROW OXIDE
2. PATTERN ALIGNMENT MARKERS
3. SILICON ETCH MARKERS
4. STRIP AND REGROW OXIDE
5. PATTERN TANK
6. IMPLANT TANK
7. DRIVE TANK
8. STRIP AND REGROW OXIDE
9. PATTERN P+
10. DEPOSIT P+
11. DRIVE P+
12. PATTERN N+
13. DEPOSIT N+
14. DRIVE N+
15. PATTERN GATE
16. GROW GATE OXIDE
17. PATTERN CONTACT
18. DEPOSIT ALUMINUM
19. PATTERN ALUMINUM
20. SINTER ALUMINUM

TABLE 3. HIGH-DENSITY CMOS RESIST AND ETCH PROCESSES

PROCESS STEP	RESIST	ETCH PROCESS
P-	PBS	BUFFERED HF
P+	PBS	BUFFERED HF
N+	TI RESIST NO. 309	PLASMA ETCH
GATE	PBS	BUFFERED HF
CONTACT	PBS	BUFFERED HF
LEADS	TI RESIST NO. 309	METEX ETCH

TABLE 4. HIGH-DENSITY CMOS MEMORY CELL TEST DATA

PARAMETER	CELL 1 (29.0 $mils^2$)	CELL 2 (15.1 $mils^2$)
WRITE TIME (τ_W)	9NS	11NS
READ TIME (τ_R)	300NS	400NS

MOLECULAR BEAM EPITAXIAL GROWTH OF InP

B. I. Miller, J. H. McFee and K. J. Bachmann*
Bell Laboratories
Holmdel, N. J. 07733 and Murray Hill*, N. J. 07974

ABSTRACT

Epitaxial layers of InP have been grown on the (111)B face of InP substrates by molecular beam epitaxy. The layers were grown at substrate temperatures ranging from 200 to 471°C, using separate In and P sources. Unintentionally doped layers grown above ∿320°C appear to be single crystal with $n = 2\text{-}5 \times 10^{16}$ cm^{-3} and $\mu \sim 3000$ cm^2/volt sec. Layers grown at substrate temperatures below ∿320°C appear to be composed of well-oriented poly-crystals. Orientation-dependent faceting is present on layers grown on (111)B substrates at substrate temperatures > 320°C, but is almost entirely absent on layers grown at the same temperatures on (111)A and (100) substrates.

I. Introduction

Molecular Beam Epitaxy (MBE) has been used successfully in the fabrication of many GaAs devices.[1] Among the many successful applications have been the GaAs/AlGaAs double heterostructure injection laser,[2] optical waveguides[3] and many microwave devices.[4,5] One unique advantage of MBE over other methods of epitaxial growth (VPE or LPE) is the ability to control the beam fluxes and thus the growth rate, layer thicknesses and doping profiles very accurately. Thus, as in the case of GaAs, it is possible to grow smooth featureless epitaxial layers as thin as 50 Å with controlled impurities.[6,7,8] Spurred on by the success of GaAs, there has been an attempt to grow other III-V compounds by MBE. Epitaxial layers grown by MBE have been reported in GaP,[9] GaAsP,[9] and InP.[10,11] There has also been some work reported on the II-IV compounds using MBE techniques.[12] Our interest is directed toward the MBE growth of InP for two reasons. First is the possible use in the CdS/InP heterojunction solar cell reported by Shay and Wagner.[13] MBE growth of InP on CdS films might be used in developing an inexpensive thin film version of this solar cell as well as in developing more complex cell geometries. Second is the potential use of InP in microwave devices such as transferred electron oscillators and amplifiers[14,15] where InP appears to yield lower noise devices than GaAs. Because the doping profile is critical in this application

MBE is especially well suited to the fabrication of these devices.

As stated earlier, some work has already been reported on MBE growth of InP. Farrow has recently reported growing InP epitaxial layers on the (100) and (111)B faces of InP by MBE using separate In and InP sources.[10,11] His unintentionally doped layers were p-type $\sim 10^{16}$-10^{17} cm^{-3}. In this work we report growth from separate In and P sources. This was done in order to obtain independent control over the In flux and the P_4/In ratio and also to take advantage of the inherently higher purities of these <u>elemental</u> sources (generally 99.9999^+ percent pure) as compared to the InP <u>compound</u> (generally $\leq$ 99.999 percent pure). We have also concentrated on the (111)B growth axis of InP due to the requirement that InP lattice match CdS for a CdS/InP solar cell. Also since many substrates such as CdS decompose at high temperatures a study of the epitaxial growth quality versus temperature was made in order to determine the lower temperature limit of good epitaxial growth.

II. <u>Experimental</u>

A. Description of MBE System

Figure 1 shows a cross section of the growth plane of our 18" UHV stainless steel bell jar. This system is similar to earlier MBE systems described by Cho[1] and others. Basically an MBE system is nothing more than a highly instrumented evaporation station and thus consists of two main parts; a growth part consisting of a substrate, substrate heater and evaporation ovens, and a monitoring section consisting of an Auger analyzer, quadrupole mass analyzer, sputter gun, reflection electron diffraction (RED), and a scattered-light monitor.

The evaporation sources consist of five small volume main ovens enclosed within a liquid nitrogen shroud and two large volume auxiliary ovens. Each main oven consists of a resistive heating element enclosing a 1 cm dia. × 4 cm long BN crucible. The auxiliary ovens are large volume ~20 cc cells having as an outlet heated snouts aimed at the substrate. The main ovens are used for the In and dopants such as Sn and Mg while the large capacity auxiliary ovens are used for P. This enables us to make about ten growth runs before having to reload either In or P.

The 3 keV electron gun shown in Fig. 1 is used as the source of electrons for RED. The resulting RED pattern is displayed on the phosphor screen. Thus we can observe the RED pattern from the substrate before a run and the RED pattern from the epitaxial layer during and after a run. The RED pattern gives information on the surface smoothness and the reconstructed two-dimensional epitaxial surface.[1,16,17,18]

The Auger spectrometer with a 3 keV sputter gun allows the surface of the substrate to be sputtered clean and monitored for any contamination prior to a run. Generally the Auger spectrometer can detect surface impurities as low as 0.1 percent.

A quadrupole mass analyzer was used to monitor the beam flux from the ovens during growth and also to measure the partial pressures of the background gases in the bell jar. The sensitivity of the quadrupole analyzer is such that at the standard epitaxial layer growth rate of .5 μ/hr. dopant fluxes of a few ppm can be detected.

In order to monitor the optical flatness of the sample surface before and during a run, a laser beam is incident on the surface. Any scattered (i.e., nonspecular) light coming from the surface is an extremely sensitive measure of the surface roughness.[19] For example, with smooth featureless epitaxial growth, the scattered light detected by a photomultiplier will generally decrease slowly during a run or remain constant. For faceted layers such as those grown on the (111)B face, the scattered light will slowly increase during a run. For polycrystalline growth the scattered light will increase rapidly at the start of growth and will continue rising during the entire run.

B. Growth Procedure

The InP epitaxial layers were grown on either n-type or semiinsulating Fe-doped LEC pulled InP single crystal substrates.[20,21] Most of the work described here was done on the (111)B face although some work was also done on the (111)A and (100) faces. Except for the (111)A face, the substrates were chemically lapped to a mirror finish in a 5 percent Br:95 percent methanol solution. Due to the unavailability of a purely chemical polish for the (111)A face of InP, a syton chemo-mechanical polish was used for this orientation. All polishing was done after the substrates were mounted on a Mo block with In solder.

After polishing, the sample is mounted in the UHV system. The system is then pumped down over night, during which time the sample is outgassed at 335°C. In the morning, prior to heating the ovens, the residual gas spectrum of the system is as shown in Fig. 2. Here the total pressure of the system is about 10^{-9} Torr consisting mostly of H_2, Ar, Ne and P. The partial pressures of H_2O, N_2 (CO) and CO_2 are 10^{-11} Torr or less.

As stated earlier, the substrates are sputtered before growth commences. A dramatic demonstration of the effectiveness of this cleaning process is given in Figs. 3 and 4. Figure 3 shows the Auger spectrum of an InP substrate before sputtering. In this case the substrate had not been subjected to the usual over night outgassing at 335°C. Besides In and P, large amounts of C and O are present on the surface together with a fair amount of Cl. Also the ratio of P/In is about 1:4, and particular attention should be paid to the shape of the P peak in Fig. 3 as compared to Fig. 4. Figure 4 is the Auger spectrum of the same substrate after sputtering. The P peak is now much larger and now P/In is about 1:1. Also all the O, C and Cl are gone.

After the surface is cleaned, the In and P fluxes are adjusted for a growth rate of ∿.5 μ/hr. The sample is rotated into growth position (see Fig. 1), the RED and scattered laser light systems adjusted. The oven shutters are then opened and growth commences. Generally the layers are allowed to grow to 2-3 μ thick so a typical run may last from 4-6 hours. After growth the samples are allowed to anneal at 330°C over night before removal from the system.

III. Results

There are two temperature regimes for the growth of InP on InP. Above growth temperatures of ∿320°C the layers are single crystal epitaxial layers exhibiting nearly bulk electrical properties. In the case of the (111)B orientation orientational dependent faceting occurs on the surface of the epitaxial layers. An example of this is shown in Fig. 5. This figure shows a Nomarski phase-contrast photograph of an epilayer grown at 410°C on a 5 mm diameter (111)B disc-shaped substrate. Chemical polishing made the substrate slightly lens-shaped (convex) so that the normal to the central region is nearly parallel to [111]B, while the peripheral regions are tilted away from [111]B by ∿2°. The morphology as revealed by the phase contrast photograph clearly reflects the three-fold nature of the [111] growth axis. The heavy faceting is strongly orientation-dependent. For example, the exact (111)B orientation at the disc center exhibits pyramidal features which are shown more clearly in Fig. 6a. However, near the rim along the three equivalent $2\bar{1}\bar{1}$ directions (Fig. 5) smooth regions appear, shown more clearly in Fig. 6b. In between these regions near the rim the morphology shows a two-fold "roof shingle" pattern as shown in detail in Fig. 6c. Sometimes a quasi-smooth area exists between the three-fold center region and the two-fold outer regions, part of this region is also shown in the upper part of Fig. 6c.

Many such lens-shaped samples were grown to make sure that these morphological features were not a experimental artifact and in each instance the resulting morphology was nearly identical to that shown in Fig. 5. As further confirmation of the orientational dependence of faceting many (100) epitaxial layers were grown simultaneously along with the (111)B layers and they exhibited shiny featureless morphologies. This type of morphology dependence on the orientation of the sample, has been seen in VPE Si by Shimbo[22] and alluded to by DeLorenzo[23] in VPE GaAs.

Evidence of the single-crystal nature of the samples grown above 320° is given in Fig. 7 for an epitaxial layer of InP grown on the (111)B face of InP at 411°C. The left hand photo shows a Scanning Electron Microscope (SEM) photograph of a typical part of the surface. At this magnification (20,000×) all faceting becomes smoothed out and the only focusing aides are dust particles on the surface. As seen, except for the dust particle and an occasional bump, the surface is uniform and grainless.

The upper right photograph in Fig. 7 shows the RED pattern taken after the epitaxial layer was grown. The streaked pattern is indicative of relatively smooth single crystal growth. The spacing of the streaks corresponds to the periodicity of a unreconstructed (111)B InP surface for an electron beam along the [110] azimuth. From the photograph no evidence of polycrystallinity or twinning can be seen. In GaAs under similar circumstances Cho sees a (111)B-$\sqrt{19}$ reconstructed GaAs surface.[16] However in his case he uses a 40 keV electron beam and is dealing with a nonfaceted surface. Our recent results on the 100 axis of InP show the same (100)-(2×8) surface reconstruction reported by Cho in GaAs.[17]

The x-ray diffraction pattern at the lower right shown in Fig. 7 was taken at glancing-incidence in the epitaxial layer (x-ray beam at ~7° from grazing incidence to maximize absorption). The x-ray was taken with a CuK_{α} line plus some white x-radiation background. If the epitaxial layers were composed of randomly oriented polycrystals, Debye-Scherrer rings should appear from the CuK_{α} line. The white x-radiation will give a Laue spot background. The lack of Debye-Scherrer rings plus a one-to-one correspondence of the Laue spots with an x-ray taken of a single crystal substrate convince us that this layer grown at 411°C is indeed a high-quality single crystal. In addition the electrical properties at 411°C for these unintentionally doped samples gives n-type $\sim 2\text{-}5 \times 10^{16}$ cm^{-3}, $\mu \sim 3000$ cm^2 V^{-1} sec^{-1}, which is comparable to bulk single crystal values.

Up to now we have been discussing typical properties of those samples grown between 470°C and 320°C. Below a growth temperature of 320°C a sudden transition occurs which we interpret to be indicative of large aligned grains possibly containing microtwins and stacking faults. Returning to Fig. 6, we see that at 290°C the surface is almost uniformly granular and appears matte to the eye. Some vestiges of the surface morphology seen at 411°C is present, for example, Fig. 6d corresponds to the same area as Fig. 6c of the 411°C sample. Also Fig. 6e is the center of the disc corresponding to Fig. 6a. At even lower growth temperatures, these small vestiges of orientation dependent morphology vanish and the growth morphology is <u>uniform</u> over the entire surface. It should be noted, however, that as the temperature is lowered the surface texture becomes smoother.

To illustrate more clearly what happens to the surface granularity as a function of growth temperature, a series of SEM photographs is shown in Fig. 8. At a growth temperature of 290°C we see the appearance of grains, as contrasted to the featureless SEM photograph in Fig. 7. At even lower growth temperatures the grains diminish in size to about 1/4 μ. Figure 9 shows the RED pattern taken at low growth temperatures. At 290°C the streak pattern still persists, but at 262°C it tends to get spotty and eventually washes out entirely at 200°C. The lack of Debye-Scherer rings, and streak spacings corresponding to the (111)B surface imply that the grains on the surface are <u>aligned</u> polycrystals. Figure 10 shows a 7° glancing incidence x-ray photograph of the sample grown at 200°C. The results are similar to the x-ray results for 411°C shown in Fig. 7, no Debye-Scherer rings are present, but a few points are broadened and a few large diffuse spots are also apparent. The absence of Debye-Scherer rings and spot patterns is an indication of alignment of the polycrystals. But the large diffuse spots in the left 1/3 of Fig. 10 are possibly indicative of defects[24] such as microtwins and stacking faults. Sloope and Tiller[25] have noted an increase in such defects as they lowered the growth temperature of evaporated Ge films.

As seen in Fig. 11, below 320°C a sharp transition also occurs in the electrical properties; the mobility decreases by an order of magnitude (from 2000 sec to 200 cm^2/volt sec) and the carrier concentration increases by nearly two orders of magnitude (from 2×10^{16} to 10^{18} cm^{-3}). The rather abrupt drop in mobility between 290°C and 262°C cannot be accounted for by decreased grain size, while the large increase in carrier concentration is certainly not due to grain size. More likely, carrier-generating defects such as stacking faults or microtwins are responsible. In fact, just the type of behavior as shown in Fig. 11 has been reported by Sloope and Tiller in evaporated Ge epitaxial films[26] and they were able to correlate the abrupt transition in the electrical properties versus growth temperature with the onset of crystalline defects. It should be noted that at temperatures below 320°, because of the high n-type carrier concentration of 10^{18} cm^{-3}, p-doping would be nearly impossible and n-doping would be possible only over a narrow range.

IV. Discussion and Conclusion

At temperatures above 320°C it is possible to grow high quality epitaxial layers of InP on InP substrates. At 410°C we have been able to grow on the (111)B, (111)A and (100) faces. On the (111)B face, pyramidal faceting of the order 3000 Å high occurs. This faceting is highly sensitive to slight misorientations away from the (111)B. It appears then that, for most applications such as microwave devices or waveguides, the (111)B orientation is far from ideal. On the contrary, the (111)A orientation appear more favorable for epitaxial growth. Shown in Fig. 12 is a Nomarski interference contrast picture of an epitaxial layer grown on a syton polished 111A substrate. Except for the pits which we feel are the result of the chemco-mechanical polish, the underlying morphology appears to be very uniform and fairly smooth. The (100) orientation which gives an essentially featureless surface and is easily polished is even more desirable. Thus the (111)A (if a good chemical etch were available) and especially the (100) would be useful orientations for epitaxial device applications.

Some preliminary work has been done on doping these samples at a growth temperature of 411°C. As of now it is possible to n-dope from 10^{16} to 10^{19} with Sn, while p-doping appears limited to above 10^{18} region. Presently we are investigating the doping properties of InP.

Acknowledgements

We would like to express our appreciation to W. Wiegmann, M. Ilegems, A. Y. Cho, A. Gossard, E. E. Chaban and J. Robinson, Jr. for help and advice in planning the MBE system. We also want to especially thank Mrs. A. A. Pritchard for technical assistance.

REFERENCES

1. A. Y. Cho and J. R. Arthur, Chapter (Molecular Beam Epitaxy) in progress in Solid State Chemistry, Vol. 10, editors G. Somerjai and J. McCaldin, Pergamon Press, New York. This review article contains 147 references.
2. A. Y. Cho and H. C. Casey, Jr., Appl. Phys. Lett. 25, 288 (1974); A. Y. Cho, R. W. Dixon, H. C. Casey, Jr., and R. L. Hartman, Appl. Phys. Lett. 28, 501 (1976).
3. A. Y. Cho and F. K. Reinhart, Appl. Phys. Lett. 21, 355 (1972); J. P. van der Ziel, R. M. Mikulyak and A. Y. Cho, Appl. Phys. Lett. 27, 71 (1975).
4. A. Y. Cho and F. K. Reinhart, J. Appl. Phys. 45, 1812
5. M. Neganuma, K. Kamimyra, K. Takahashi and Y. Sakai, Japan. J. Appl. Phys. 14, 581 (1975).
6. L. L. Chang, L. Esaki, W. E. Howard and R. Ludeke, J. Vac. Sci. Technol. 10, No. 1, pp. 11, Jan/Feb (1973).
7. A. Y. Cho, Appl. Phys. Lett. 19, 467 (1971).
8. R. Dingle, W. Wiegmann and C. H. Henry, Phys. Rev. Lett. 33, 829 (1974).
9. M. Neganuma and K. Takahashi, Phys. Stat. Sol. 31, 187 (1975).
10. R. F. C. Farrow, J. Phys. D., Appl. Phys. 8, L87 (1975).
11. R. F. C. Farrow, Extended Abstracts Electrochemical Society, Dallas, Tex., Abs. No. 214, October (1975).
12. D. L. Smith and V. Y. Pickhardt, J. Appl. Phys. 46, 2366 (1975).
13. S. Wagner, J. L. Shay, K. J. Bachmann and E. Buehler, Appl. Phys. Lett. 26, 229 (1975).
14. R. Corlett, I. Griffith and J. J. Purcell, GaAs and Related Compounds, Institute of Phys. Conf. Series No. 24, pp. 89, Chapter II, Institute of Physics, 47 Belgrave Square, London, England (1975).

15. R. C. Clarke and L. L. Taylor, J. of Cry. Growth 31, 190 (1975).
16. A. Y. Cho, J. Appl. Phys. 41, 2780 (1970).
17. A. Y. Cho, J. Appl. Phys. 42, 2074 (1972).
18. A. Y. Cho, M. B. Panish and I. Hayashi, 1970 Proc. Third Int. Symp. on GaAs and Related Compounds, Aachen, Germany, Inst. Phys. Soc. Conf., Ser. #9, pp. 18.
19. W. Wiegmann, private communication.
20. K. J. Bachmann, E. Buehler, J. L. Shay and A. R. Strnad, J. of Elect. Matls. 4, 389 (1975).
21. The preparation and properties of bulk Indium Phosphide crystals and of InP light emitting diodes operated near 1.05 μ on wavelength. K. J. Bachmann, E. Buehler, J. L. Shay and D. L. Malm, Inst. of Phys. Conf. Ser. No. 24 (1975), Chapter 3, pp. 121.
22. M. Shimbo, J. Nishizawa and T. Terasaki, J. of Cryst. Growth 23, 267 (1974).
23. J. V. DiLorenzo, J. of Crystal Growth 17, 189 (1972).
24. M. L. Hammond and D. J. Rowcliffe in Techniques of Metal Research, Vol. 2, part 2, Chapt. 18, p. 801, R. F. Bunshah, editor, John Wiley, New York, 1969.
25. B. W. Sloope and C. O. Tiller, J. Appl. Phys. 37, 887 (1966).
26. B. W. Sloope and C. O. Tiller, J. Appl. Phys. 38, 140 (1967).

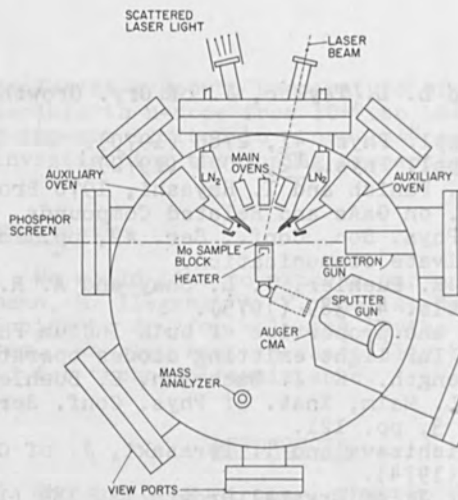


Figure 1. Horizontal cross-section of MBE vacuum system bell jar.

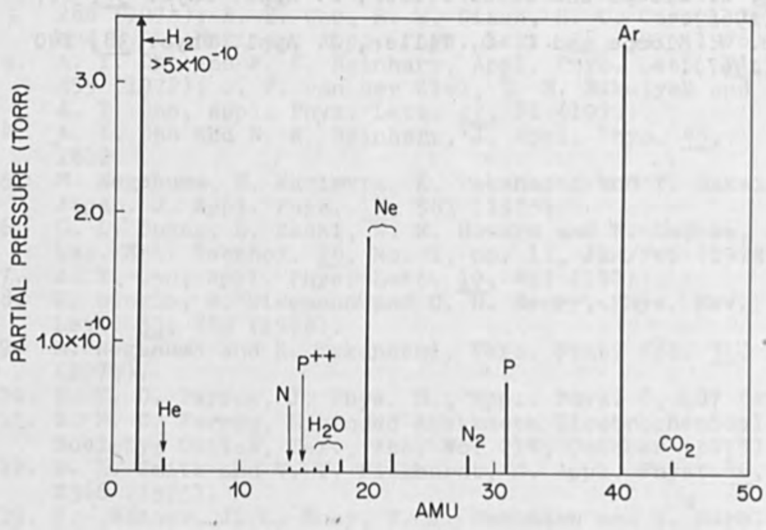


Figure 2. Mass spectrum of residual gases in bell jar following overnight bake-out but before starting layer growth. Ovens and sample are at "idling" temperatures of 440°C and 335°C, respectively. Total pressure in bell jar $\lesssim 1 \times 10^{-9}$ Torr.

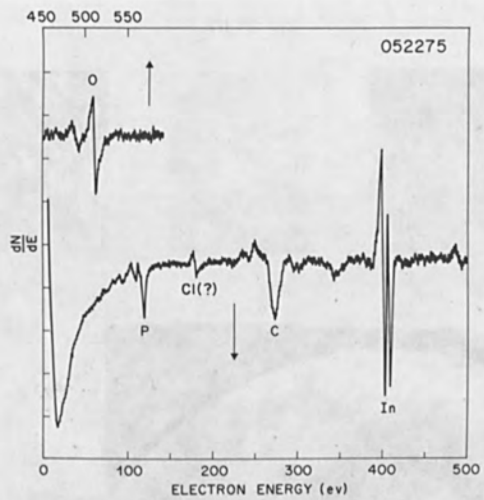


Figure 3. Auger spectrum of InP substrate surface before sputter cleaning. (Substrate has not been subjected to usual overnight outgassing at 335°C.

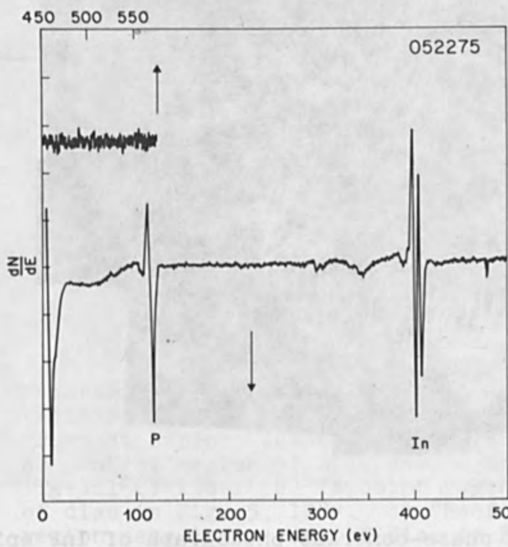


Figure 4. Auger spectrum of InP substrate after 10 minutes sputter cleaning (3 kV Ar^+, 2 $\mu A/cm^2$).

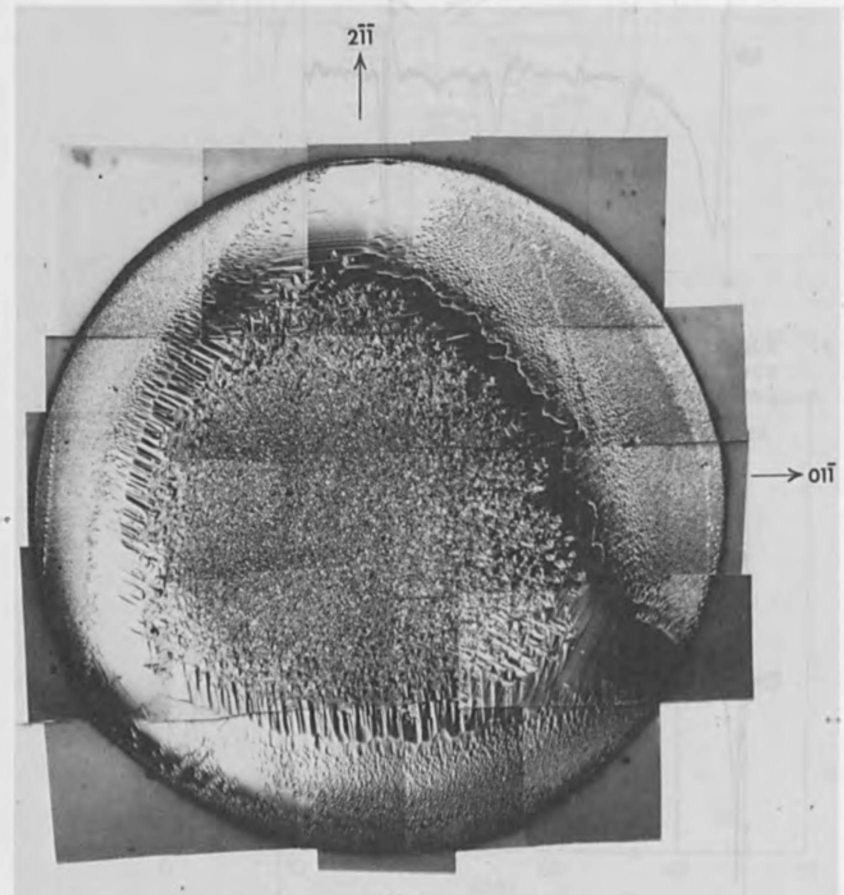


Figure 5. Nomarski phase-contrast photograph of InP epitaxial layer grown at 411°C on a (111)B disc-shaped InP substrate. Azimuthal orientation of substrate is indicated by arrows.

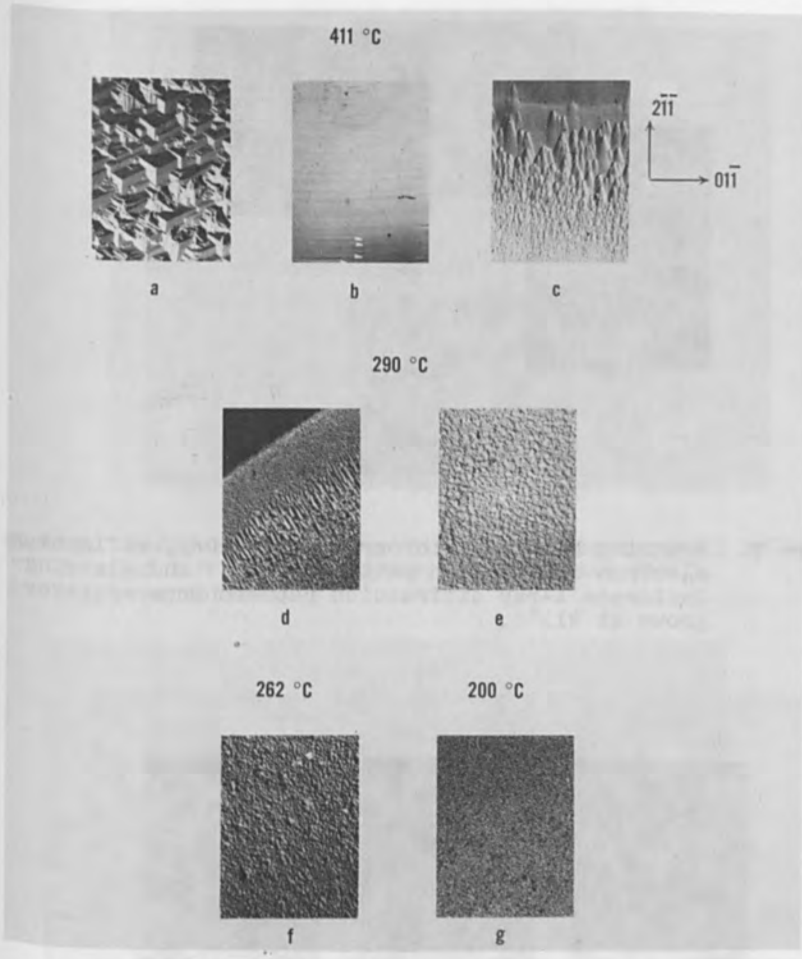


Figure 6. Nomarski phase-contrast photograph showing details of surface morphology of InP epilayers grown at various substrate temperatures, T_S. a) Central region of disc sample of Fig. 5 (T_S=411°C) 128×. b) "Smooth" region at top edge of disc in Fig. 5, 128×. c) "Roof shingle" region near lower edge of disc in Fig. 5, 200×. d) Near edge of disc (T_S=290°C), 256×. e) Central area of disc (T_S=290°C) 1000×. f) T_S = 262°C, 1000×. g) T_S = 200°C, 1000×.

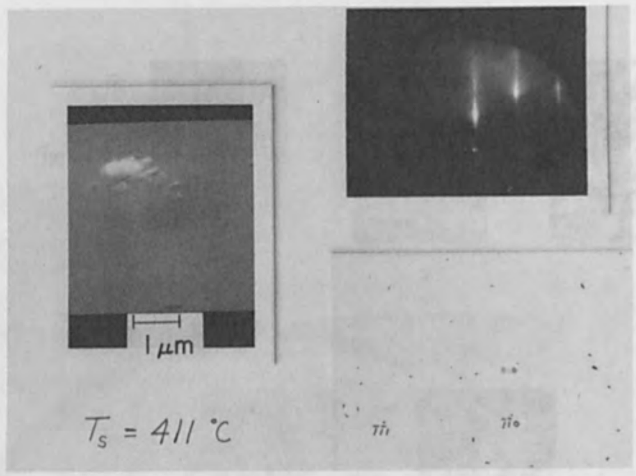


Figure 7. Scanning Electron Micrograph (20,000×), reflection electron diffraction pattern (3 keV) and glancing-incidence X-ray diffraction pattern from epilayer grown at 411°C.

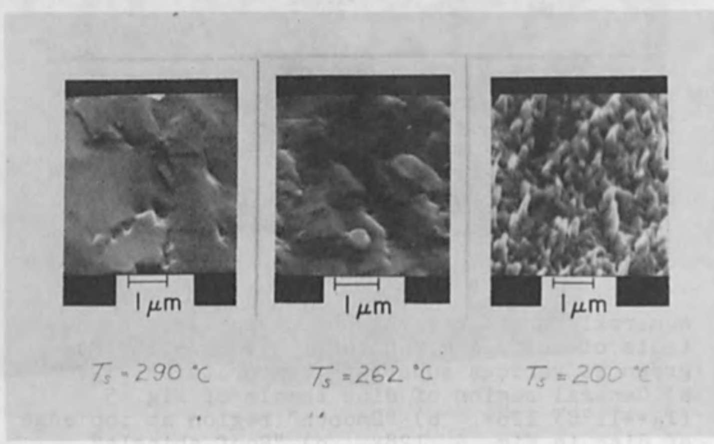


Figure 8. Scanning Electron Micrographs (~20,000×) of surfaces of InP epilayers grown at 290°C, 262°C and 200°C, respectively.

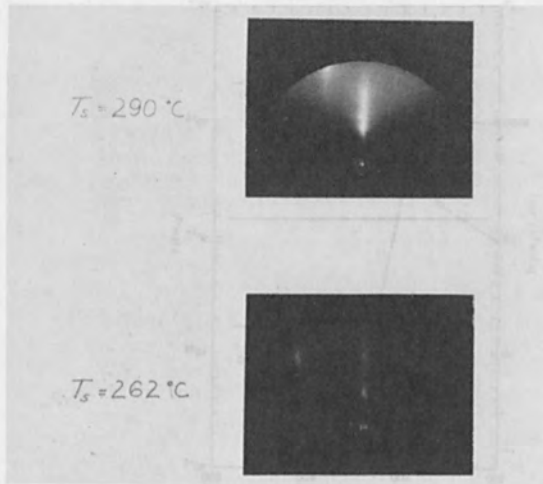


Figure 9. Reflection electron diffraction patterns (3 keV) of InP epilayers grown at 290°C and 262°C, respectively.

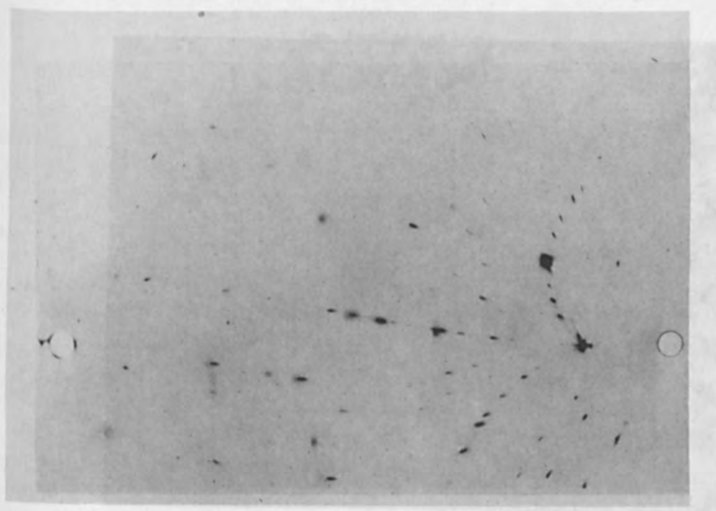


Figure 10. Glancing-incidence (∿7°) X-ray diffraction pattern from InP epilayer grown at 200°C. A cylindrical camera and Ni-filtered CuK$_\alpha$ radiation was used.

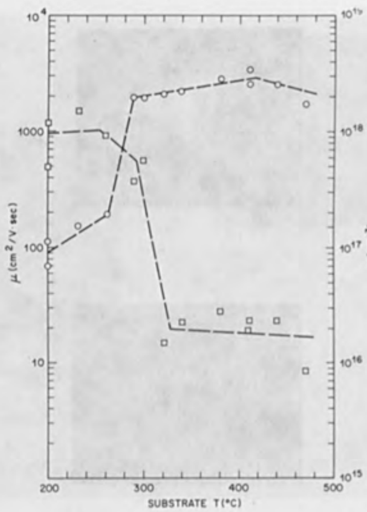


Figure 11. Mobility, μ, (circles) and carrier concentration, $n \equiv N_D - N_A$, (squares) as a function of growth temperature for unintentionally-doped InP epilayers.

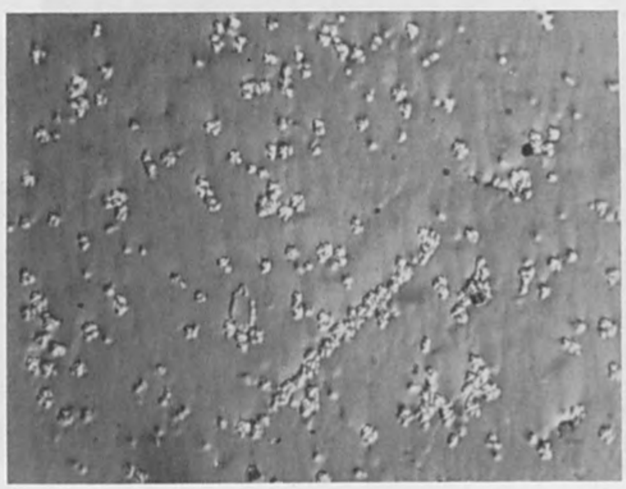


Figure 12. Nomarski phase-contrast photograph of surface of epilayer grown at 410°C on (111)A InP substrate (500×).

Section III.

Ion Implantation.

Figure 11. Mobility, μ, (circles) and carrier concentration, $n = N_D - N_A$, (squares) as a function of growth temperature for unintentionally-doped InP epilayers.

Figure 12. Nomarski phase-contrast photograph of surface of epilayer grown at 810°C on (111)A InP substrate (500×).

PR-500, A HIGH VOLTAGE, GENERAL PURPOSE ION IMPLANTATION SYSTEM

J. H. Jackson, H. M. Bird, G. J. Hofer, P. J. Mostek,
G. I. Robertson, A. F. Rodde, B. Weissman, and N. Williams
Western Electric Engineering Research Center, Princeton, N.J. 08540

ABSTRACT

The PR-500 implantation system is a general-purpose 500-kilovolt machine designed for IC device development and production. Principal system components are a high current Penning ion gauge source, a grounded 90° double focusing magnet for mass analysis, a constant gradient accelerator, and a rotating mechanical scan target chamber at high voltage. Information is transmitted between the high voltage target stage and ground by a multiplexed light beam telemetry system. Translation of the rotating target disk is controlled digitally for accurate dose control and uniformity. With the rotating scan target stage and Penning ion source, typical target currents of 400 μa ^{11}B and 600 μa ^{31}P result in throughputs of 200 two-inch wafers/hour for doses up to $2\times10^{14}/cm^2$.

1. INTRODUCTION

The PR-500 ion implantation system is a general purpose 500-kilovolt machine designed for device development and production. Like the earlier Mark I system[1], the PR-500 employs a pre-analysis configuration with source at +30kV, a sector magnet at ground for mass analysis, and a target stage at high voltage. The PR-500 system makes use of a novel rotating scan target chamber for efficient utilization of the high target currents obtainable from its Penning ion gauge source. Use of the rotating scan principle also results in a substantial reduction in individual wafer heating compared to conventional scan systems. While great flexibility in implanted doses, energies, and species makes the system a useful device development tool, its high throughput and automated control system also make it well suited for production applications.

2. SYSTEM CONFIGURATION

The overall configuration of the PR-500 system is shown in Fig.1. Ions are generated in a Penning source[2] which is held at a potential of +30kV above ground. A lens is incorporated into the source for optical matching to the rest of the system. With gaseous sources, typical target currents of 400 μa ^{11}B and 600 μa ^{31}P are achieved for accelerating voltages above 250kV. High current (1 ma) ^{75}As current operation of the source on the system is also possible. Typical source lifetime for these species is 40 hours. After extraction from the source, an einzel lens focuses the ion beam for parallel entry into the mass-analyzing magnet. A six-inch diffusion pump mounted beneath the chamber which houses the first einzel lens maintains a pressure of approximately 1×10^{-5} torr in this region.

Immediately after focusing by the first lens, a rotatable electrode machined with an eccentric aperture provides a means to reduce the relatively high ion source output currents for low dose implants. During an implant, current reaching the target is monitored and the rotating aperture is also adjusted automatically to compensate for source current drifts. In this way a near optimum target current level is ensured for the particular implant dose required. When fully closed, the variable aperture, together with a suppression electrode mounted on the exit of the first einzel lens, can be used to measure the total ion source output current.

The mass analyzer is a 90° sector magnet with a mass-energy product of 6.5 Mev-amu, allowing masses as high as 215 amu to be analyzed for a 30kV source potential. The pole tips, shaped for double-focusing, provide an approximate source size magnification of 1.5 at the mass-defining slits. For an ion of mass M, the mass dispersion of the magnet at the slits is approximately 17/M (cm/amu). The slits are made of tantalum to reduce the effects of beam sputtering and are mounted on vernier feedthroughs to allow fine adjustment of the slit opening.

After mass analysis, the analyzed beam current is measured by a movable Faraday cup which also functions as an automatic beam shutter for the target stage at the start and finish of an implant and in the event of an emergency which requires interruption of the implant. The beam is next focused by a second einzel lens and brought to a final energy of up to 500 keV by a constant gradient accelerator. Since the accelerator focal length increases with decreasing target voltage, the lens is employed to provide additional focusing at low or moderate energies. A six-inch diffusion pump mounted beneath the second lens chamber maintains a pressure of approximately 1×10^{-6} torr in this region. The accelerator consists of machined aluminum electrodes separated by glass insulating sections. Vacuum sealing at the glass-metal interface is provided entirely by o-rings and the assembly is held in compression by four spring-loaded insulating rods which run the length of the tube. The accelerator has a wide bore (20 cm.) to allow efficient pumping of the accelerator-target region by the pumping station beneath the second lens. In addition, a titanium sublimation pump attached to the target chamber provides additional pumping in the accelerator-target region.

Entering the target stage, a set of four electrically isolated electrodes intercept the periphery of the beam and provide an indication of beam focusing and alignment. The same electrodes also serve to define the beam size at the target, thereby fixing the amount of overscan required for dose uniformity. Additional electrodes provide secondary electron suppression for the target wafer disk.

3. ROTATING SCAN TARGET STAGE

Fig. 2 illustrates the principle upon which the rotating scan target chamber is based[3]. Silicon wafers are mounted in several rings on a target disk which rotates at a constant speed of approximately 850 rpm and translates across the beam at a variable speed on the order of 1 cm/sec. In order to achieve good dose uniformity, the precise speed with which the disk translates must depend upon three factors:

1. The required ion dose
2. The distance of the beam from the center of the disk
3. The instantaneous value of the incoming ion current

The disk makes one or more passes across the beam, depending on the dose to be implanted.

There are several important advantages to the use of the rotating scan principle. For a given uniformity and beam size the minimum implant time for a scan system is inversely proportional to its speed in the fast scan direction[3]. As a result, for low or medium doses, a reduction of the target current may be necessary so that the implant time equals or exceeds the minimum implant time for the desired uniformity. The rotating scan system, with its fast scan speed in the circumferential direction, permits utilization of maximum target currents at much lower doses than do conventional x-y mechanical scan systems which have fast scan speeds typically two orders of magnitude lower. In addition, with the rotating scan system individual wafer heating is, in many circumstances, substantially reduced by the distribution of beam energy over the many wafers in a single ring[4]. A marked reduction in wafer heating can also be realized when the total implant dose is accumulated over several slow scans since appreciable cooling occurs during the time each ring of wafers is out of the beam. The rotating scan mechanism is simpler mechanically than conventional x-y mechanical scans since it is much easier to generate rapid rotary motion than it is to generate rapid oscillatory x or x-y motion. When one considers that these motions take place in a vacuum without benefit of the usual lubricating agents, this advantage takes on added significance. Finally, the rotating scan system, like all mechanical scan systems, allows a shorter overall machine length than for electrical scan systems since space is not necessary for deflection plates or drift spaces. This shorter length not only conserves floor space, but also gives improved current handling capability through reduced space charge spreading.

A target disk capacity of 60 two-inch wafers arranged in three rings, high currents from the Penning source, and rapid pumpdown of the target section allows throughputs of 200 wafers/hr. for doses up to $2x10^{14}$ ions/cm^2. The disk upon which the target wafers are mounted is dished at an angle of 7° (Fig. 3) to avoid undesired axial channeling of the ions into the silicon lattice. It might be thought that the 7° angle would cause a slight variation in the orientation of the beam with respect to a wafer's cyrstallographic axes as an individual wafer traverses the beam. It can, however, be shown that as long as the disk axis is rotated in the direction parallel to the direction of the incoming ion beam, the orientation of the beam and the wafer's crystallographic axes is fixed. In addition, each wafer flat is rotated at an angle of 15° (Fig. 3) with respect to the disk radius which passes through it to prevent unwanted planar channeling[5].

A corona shield with rounded contours (Fig. 4) surrounds the high voltage target stage, eliminating sharp points and corners which are likely spots for the initiation of corona or breakdown. The corona shield is itself supported by a series of insulating legs divided by three equipotential planes. A resistive divider chain, not shown in Fig. 4, maintains an equal voltage drop across each section of the insulating legs. The field distribution of the corona shield and equipotential planes was modeled using resistive paper and the maximum electric fields were found to be 11.4 kV/cm, which represents an approximate safety factor of two for the onset of corona discharge in air. In actual operation, no problems have been encountered with operation of the target stage at potentials of up to 500kV.

Fig. 4 also illustrates the manner in which the target chamber is loaded. An entire disk of wafers is loaded onto the scan mechanism when the door is in the down positon. As the door begins to close, the target disk automatically starts to rotate and the 7° angle of the dished target disk provides a component of centrifugal force to hold the wafers securely to the disk. As a result, clamping of the wafers to the target disk is unnecessary.

Dose uniformity of the scan system was characterized by van der Pauw measurements of sheet resistivity after a suitable anneal. Single wafer 3 σ nonuniformity was 1.2% (Fig. 5) and wafer to wafer 3 σ nonuniformity within a run was 2.7% (Fig. 6). Over the past year, run-to-run 3 σ nonuniformity for 10^{14}/cm^2 boron and phosphorus implants was 6.2% (Fig. 7). If gaussian distributions are assumed, the effect of the three types nonuniformity can be combined to yield a total product 3 σ nonuniformity of 6.9%, assuring that sheet resistivities for virtually all these implants will fall within 6.9% of the mean.

4. CONTROL SYSTEM

There are two basic classes of controls in the PR-500 system (Fig. 8). The first group, accessible only to an engineer, is used to set the beam up at the beginning of the day. Controls in this category include source controls, vacuum system controls and interlocks, power supplies for the lenses and analyzing magnet, and other beam controls. After the engineer's initial setup, nontechnically skilled operators can perform implants of varying doses and energies using only the second group of controls, the implant controls. These controls are centralized on a small panel, the Operator's Panel, which allows the operator to set the dose, energy, the number of rings of wafers (one, two, or three) on the target disk to be implanted, to start the implant, and to open and close the target chamber door for disk loading and unloading.

The heart of the implant control system is a Digital Equipment Corp. PDP-16/M minicomputer. Given the dose required and the scale reading of the variable scale current integrator which measures the incoming target current, the 16/M calculates the number of scans required for the desired dose. It also translates the disk at the proper speed depending on the dose per scan required, the distance of the beam from the center of the disk, and the instantaneous value of the incoming ion current. The 16/M also signals the operator at the proper time to select the implant parameters, to start the implant, and to vent and unload the target chamber upon completion of the implant. Moreover, the 16/M ignores all control signals from the Operator's Panel unless the proper sequence is followed. For example, the high voltage cannot be set unless the target chamber is at high vacuum and the doors to the high voltage machine room are locked. Finally the 16/M monitors and displays the status of important system components (500kV power supply, dose suppression power supplies, vacuum conditions, etc.) during an implant and interrupts the implant if a malfunction is detected. Following correction of the malfunction, the implant is resumed at the precise point of interruption so that implant uniformity is not affected.

Signals from the operator's and engineer's controls are transmitted between the high voltage terminal and ground by a multiplexed optical telemetry system. There are four telemetry links, two of which are dedicated. One dedicated link transmits output charge pulses from a Brookhaven Model 1000 integrator, which measures target current in the high voltage target stage, to ground where they are counted and processed by the PDP-16/M minicomputer. Appropriate pulses from the minicomputer are then telemetered to the target stage over a second dedicated link to drive a stepping motor which translates the target disk. The remaining two telemetry channels are multiplexed to transmit a variety of digital control and status signals from the high voltage target stage to ground and vice versa.

5. CONCLUSION

The PR-500 system was designed as a general purpose 500-kilovolt machine for device development and production. High currents available from the Penning source are efficiently utilized, particularly at low or moderate doses, by its rotating scan target chamber, allowing throughputs of 200 wafers/hour for doses up to 2×10^{14} ions/cm^2. The rotating scan system also results, in many instances, in a substantial reduction in individual wafer heating by the beam energy. 3σ sheet resistivity nonuniformity across a wafer was determined to be 1.2% and wafer to wafer nonuniformity within a run was 2.7%, indicating excellent uniformity for the scan system itself. Run-to-run 3σ nonuniformity over the past year was determined to be 6.2%, resulting in a total 3σ product nonuniformity of 6.9%. While great flexibility in implanted doses, energies, and species makes the system a useful device development tool, its high throughput and automated control system permits utilization of this flexibility in a production environment.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the considerable efforts of P. D. Parry in planning and coordinating the implant uniformity characterization work.

REFERENCES

[1] A. F. Rodde, et al, Paper 222, Extended Abstracts of the Electrochemical Society Spring Meeting, May 12-17, 1974, San Francisco, Calif.
[2] J. P. Flemming, J. Vac. Sci. Technol., 12, 1369 (1975)
[3] G. I. Robertson, J. Electrochem. Soc., 122, 796 (1975)
[4] P. D. Parry, J. Vac. Sci. Technol., 13, (March-April, 1976)
[5] P. D. Parry, J. Electrochem. Soc., submitted for publication.

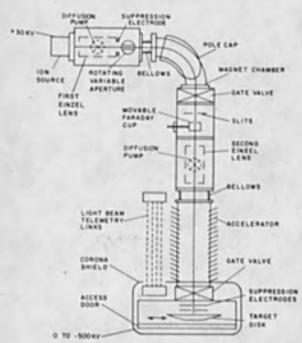


Fig. 1 PR-500 Configuration

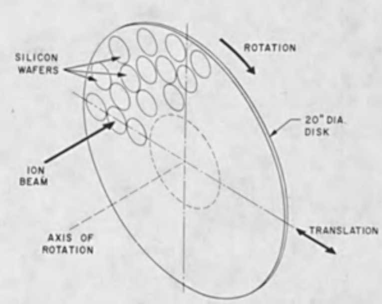


Fig. 2 Principle of rotating mechanical scan

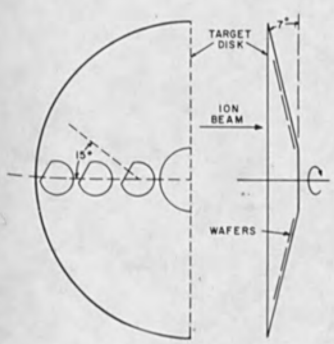


Fig. 3 Orientation of wafers on target disk

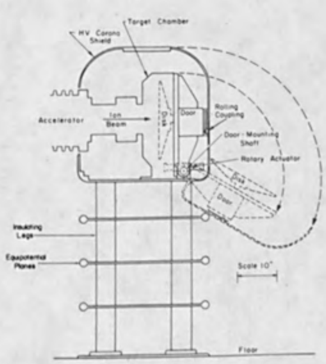


Fig. 4 High voltage target stage

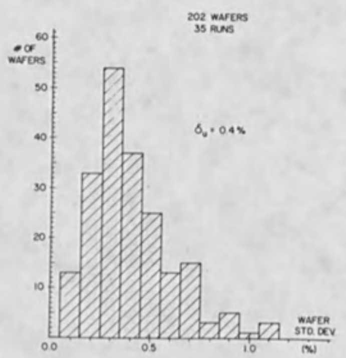


Fig. 5 Individual wafer sheet resistivity nonuniformity

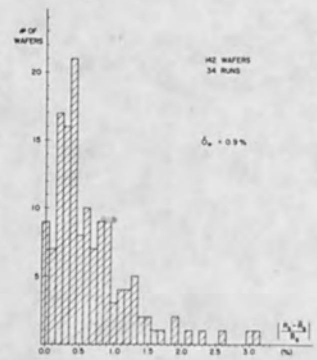


Fig. 6 Wafer-to-wafer sheet resistivity nonuniformity within an implant

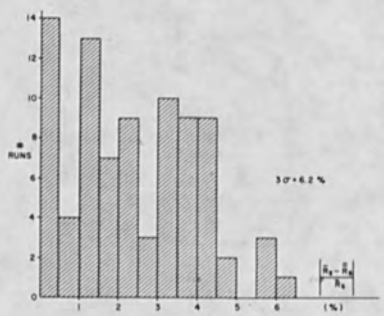


Fig. 7 Sheet resistivity nonreproducibility from implant to implant

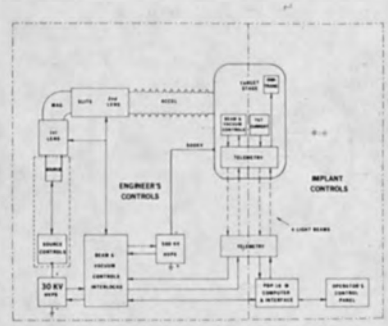


Fig. 8 Control system schematic

FACTORS WHICH DETERMINE THE EXPOSURE
TIME IN AN X-RAY LITHOGRAPHY
EXPOSURE SYSTEM

W.D. BUCKLEY
Perkin-Elmer Corporation
Norwalk, Connecticut 06856

ABSTRACT

A relation is derived for the exposure time of an X-ray lithography exposure system. It is shown that for any particular combination of vacuum window, mask membrane, photoresist and exposure atmosphere there is an optimum exposure wavelength which results in a minimum exposure time. This wavelength is calculated and it is shown that deviations from the optimum wavelength can result in a large increase in exposure time.

A typical X-ray lithography exposure system (Figure 1) consists of a characteristic X-ray source, vacuum window, exposure atmosphere, mask membrane and a substrate coated with a thin film of X-ray sensitive photoresist. The exposure time is determined by the total X-ray absorption characteristics of all the system elements, the emission efficiency of the source and by the sensitivity of the photoresist. In order to minimise the exposure time it is necessary to maximise the rate of energy absorption in the photoresist. All of the system elements between the source and the photoresist absorb X-rays which do not contribute to the imaging process. The choice of exposure wavelength is therefore a compromise between short wavelengths which decrease energy loss due to absorption in these system elements and longer wavelengths which increase absorption in the resist (Figure 2). In practical systems other parameters, such as required resolution further limit the range of exposure wavelengths available. In the following optimum exposure wavelength and exposure time will be calculated for typical systems and the factors which limit the range of available exposure wavelengths will be discussed.

Optimum Exposure Wavelength And Minimum Exposure Time

In the following discussion the X-ray source is assumed to be the characteristic emission wavelength generated by an electron beam impinging on a solid elemental target. The contribution to the exposure due to bremstrahlung radiation is ignored. In general this contribution is less than 25% of the total radiation received, and its effect is to reduce the exposure time. Hence the calculated values will err on the side of overestimation. It is important to note that short wavelength continuous radiation can affect the resolution, contrast, and background exposure and this places an upper limit on the choice of beam accelerating voltage.

It can be shown that the energy absorbed per unit area of photoresist is given by

$$\Sigma = \left(\frac{P_B}{D^2}\right)\left(\frac{\eta}{E_B}\right) T \left\{\exp[-(\mu_1 t_1+\mu_2 t_2+\mu_3 t_3)]\right\} \left\{1-\exp(-\mu_4 s)\right\}\left(\frac{hc}{\lambda}\right) \qquad (1)$$

where P_B is the beam power, E_B is the beam energy, D is the source to substrate distance, η is the electron to photon conversion efficiency and T is the exposure time. The coefficients μ_1,μ_2,μ_3,μ_4 are the linear absorption coefficients of the vacuum window, exposure atmosphere, mask membrane and photoresist respectively, and t_1,t_2,t_3 are the thicknesses of the vacuum window, exposure atmosphere and mask membrane respectively. s is the position in the photoresist relative to the irradiated surface.

In this equation the first two terms represent the conversion efficiency of electrons to photons including the D^2 term which accounts for the inverse square law reduction in photon flux. The first exponential term represents the photon losses due to absorption in the system elements and the second exponential term represents photon absorption in the photoresist. The form of this latter term arises because only a fraction of the incident photons are absorbed by the resist; the transmitted photons are ultimately absorbed in the substrate. The final term represents the photon energy.

The exposure of the photoresist is a volume effect arising from photon absorption within the film. The energy density absorbed to produce an image is the radiation dose (X) characteristic of the photoresist and defined by the relation

$$X = \frac{\partial \Sigma}{\partial s} \qquad (2)$$

Because of absorption the photon flux decreases with increasing distance into the resist from the irradiated surface. In order to produce an image it is necessary that the characteristic radiation dose is absorbed immediately adjacent to the substrate-photoresist interface where the intensity is least. The photoresist adjacent to the substrate will also be exposed by photoelectrons generated by X-ray absorption in the substrate. In this calculation this exposure is assumed to be negligable. This does not affect the generality of the discussion since significant exposure due to this mechanism has the effect of reducing the effective thickness of the resist. In most systems the phtoresist absorption is low and the optimum exposure time and wavelength are not affected significantly by this assumption.

Therefore, as a boundary condition, equation 2 must be satisfied at the photoresist-substrate interface in the exposure time τ, for a photoresist of thickness $s = t_4$. The exposure time relaion is obtained from equation 1 and equation 2 and is given by:

$$\tau = \left(\frac{X}{hc}\right)\left(\frac{D^2}{P_B}\right)\left(\frac{E_B}{\eta}\right)\left(\frac{1}{a_3\lambda^2}\right) \exp\ (\gamma\lambda^3) \tag{3}$$

In order to obtain this relation it is assumed that the absorption coefficients may be approximated by

$$\mu_n = a_n\lambda^3$$

where a_n is a constant characteristic of the absorber and whose value changes at absorption edges, λ is the exposure wavelength. The constant $\gamma = a_1t_1+a_2t_2+a_3t_3+a_4t_4$.

The value of η/E_B is a function E_B. For higher values of E_B the generation efficiency increases, however, because the beam penetrates further into the target, more of the X-rays which are generated are absorbed in the target material. An optimum beam energy exists and it has been shown (1) that for elements commonly used as X-ray lithography targets the maximum value of η/E_B is given by

$$\left(\eta/E_B\right)_{max.} = A+B\lambda^n \tag{4}$$

Where A,B, n are constants whose values are different for K-line and L-line emission. Since, in general "K" sources are more efficient than "L" sources the following discussion will be mainly confined to "K" sources.

The optimum exposure wavelength may be determined by differentiating equation 3 after substitution according to equation 4. The minimum exposure time occurs at an optimum wavelength given by

$$\lambda^3_{opt} = \frac{5}{6\gamma} \tag{5}$$

for a K-line emission source. For an L emission source the constants A,B and n have different values and the optimum wavelength is shifted slightly.

The minimum exposure time from (5) and (3) is

$$\tau_{opt} = \left(\frac{1}{hc}\right)\left(\frac{X}{a_4}\right)\left(\frac{D^2}{BP_B}\right)\left(\frac{1}{\lambda_{opt}{}^{5/2}}\right) \exp\left(\frac{5}{6}\right) \tag{6}$$

This equation embodies an important conclusion. When an optimum exposure wavelength is used the exponential term in equation 3 is reduced to a constant and the minimum exposure time increases linearly with $\lambda_{opt}^{-5/2}$. Figure 3 illustrates the variation in exposure time with optimum wavelength for a source of substrate distance of 50cm. and for a photoresist sensitivity of 10 Joules.cm.$^{-3}$. As an example a system which has an optimum exposure wavelength of 8.3A^{o}, near to the Al_k emission line, requires an exposure of 2000 Kw. seconds. This figure also illustrates the limits placed on exposure time by system constraints. X-ray absorption in the resist produces photoelectrons whose range increases with decreasing wavelength. For short wavelengths the photoelectron range in the resist limits the resolution attainable and the desired resolution sets a lower limit for the exposure wavelength. This lower limit is approximately 4A^{o} to 5A^{o} for 1μm resolution. For longer wavelengths the limit can be obtained from equation 5 from the contribution of the individual system elements to the total system absorption. A maximum wavelength can be calculated for each individual element. For example, the simplest conceivable system consists of a source and a photoresist coated substrate. For this case the optimum wavelength is determined by the photoresist and for a typical X-ray photoresist results in an optimum wavelength of 22A^{o} for a 1μm film or 27A^{o} for a 0.5μm film. This results in a minimum exposure time of approximately 170 and 120 Kw. seconds, respectively. In practice this result may be modified by the absorption edge for oxygen at 23.3A^{o} since oxygen is present in many resists. Similarly the carbon absorption edge significantly reduces photoresist absortpion and increases the optimum wavelength. However, as we shall see this can result in an increase in exposure time.

In practical systems absorption in the vacuum window, mask membrane and exposure atmosphere has a significant effect on the optimum exposure wavelength and time. As an example, Helium is often proposed as a suitable exposure atmosphere because of its low absorption, high heat dissipation and convenience. However it limits the maximum optimum exposure wavelength and hence limits the minimum exposure time. The optimum exposure wavelength for a 50cm. long column of Helium is approximately 12A^{o} and the minimum exposure time is 800 Kilowatt seconds. Other more strongly absorbing elements have a similar but larger effect.

Table I illustrates the effect of various mask membrane, vacuum window and exposure ambient combinations on the exposure time and optimum wavelength. The mask membrane considered is mylar of different thicknesses. On the long wavelength side of the silicon absorption edge, silicon and mylar have approximately the same absorption and so the results apply equally to silicon masks. A 50cm. source to substrate distance was used for calculating the exposure times. The vacuum window is a 25 mciron Be film and the photoresist is one micron thick. The effect of mask membrane absorption is evident in the first column of the table. For a system with a 6 micron mylar mask the optimum wavelelngth is 8.4A^{o}. The addition of a vacuum window to permit exposure in a helium atmosphere reduces the optimum expousre wavelength to 6.2A^{o} and increases the minimum exposure time by more than a factor of two. Unless the helium permits the same factor increase in power the exposure time will increase and the heilum atmosphere can only be justified by convenience considerations.

Exposure At Non Optimum Wavelength

Characteristic X-ray emission sources are not available for all wavelengths. This and other consideration requires a compromise wavelength to be chosen as near as possible to the optimum. The effect on the exposure time of exposing at a non opimum wavelength can be determined by the substitution in equation 6 of

$$\lambda_m = m\lambda_{opt}$$

where λ_m, λ_{opt} are the exposure wavelength and the optimum exposure wavelength respectively and m is a number. This results in

$$\frac{\tau_m}{\tau_{opt}} = m^{-5/2} \quad \exp[\tfrac{5}{6}(m^3-1)] \qquad (7)$$

This relation is illustrated in Figure 4. In general it is necessary to expose at a wavelength within 40% of the optimum to obtain an exposure time less than twice the optimum. It is important to note that fractional changes in wavelength are important. For example in a system optimised for 8A^{o} radiation, exposure at 4A^{o} more than doubles the exposure time. However for a system optimised for 4A^{o} radiation, exposure at 8A^{o} results in a factor of 60 increase in exposure time.

Absorption Edges

The effect of absorption edges can be evaluated simply by including the appropriate value for the linear absorption coefficent. In general if subscripts 1 and 2 refer to wavelengths longer and shorter than an absorption edge then

$$\frac{(\lambda_{opt})_1^3}{(\lambda_{opt})_2^3} = \frac{\gamma_2}{\gamma_1} \qquad (8)$$

and since $\gamma_2 < \gamma_1$, the effect is to increase the optimum wavelength. Similarly

$$\frac{(\tau_{opt})_1}{(\tau_{opt})_2} = \left(\frac{a_{42}}{a_{41}}\right)\left(\frac{\gamma_1}{\gamma_2}\right)^{5/6} \qquad (9)$$

where a_{41},a_{42} are the values of a_{41} in regions 1 and 2 respectively. In this case since γ contains a_4 the effect will depend on the relative contribution of the absorption edge to γ. For example, an absorption edge dominant in a system element but not in the resist will result in a decrease in exposure time. An example of this is the use of a silicon mask membrane above the silicon absorption edge. However an absorption edge dominant only in the resist will cause an increase in the exposure time determined in part by the relative contribution of the resist to the total system absorption. If the absorption edge is present in the resist and in another system element, an increase in exposure time is possible.

Approximate Optimum Wavelength And Minimum Exposure Time

In practice a source with a characteristic photon wavelength required to optimise the exposure time is rarely available. Also the optimum beam energy for maximising η/E_B may be high enough to cause serious problems due to bremstrahlung background exposure. Practical systems usually employ elemental sources at a fixed beam energy. In this case the exposure time of equation 3 may be represeneted by the relation

$$\tau = \left(\frac{1}{C}\right)\left(\frac{D^2}{P_B}\right)\left(\frac{X}{a_4}\right)\left(\frac{1}{\lambda 3}\right)[\exp\,(\gamma\lambda^3)] \qquad (10)$$

where $C = \left(\frac{\eta}{E_B}\right)\left(\frac{hc}{\lambda}\right)$ (11)

The value of C will be a constant for any source for a fixed beam energy. The value of C at 20 Kev beam energy for sources typically used for X-ray lightography was investigated by Mayden et al (2) and shown to be 60 $\pm$ 5μ watts/watt/steradian a value which is used in the following discussion.

The exposure time for $Al_K Rh_L$ and Cu_L source calculated according to equation 10 is illustrated in figure 5 for various values of the total system absorption. It is apparent that although for any source the exposure time decreases continuously with decreasing γ, the improvement below the value for which the source is optimum is limited. Also below this value a different source exists which will result in a shorter exposure time. The Rh_L source results in a larger exposure time over most of the range of γ but is preferred for high absorption systems. Calculation of λ_{opt} for the system elements described in reference 2 results in an optimum wavelength close to the Rh_L emission line which was used. The reported exposure time is in agreement with that calculated from equation 6.

The approximate exposure time relation results in an optimum wavelength given by

$$\lambda^3_{opt} = \gamma^{-1} \tag{12}$$

and hence in an optimum exposure time relation

$$\tau_{opt} = \left(\frac{1}{C}\right)\left(\frac{D^2}{P_B}\right)\left(Xt_4\right) \quad [\exp(1)] \left(\frac{\gamma}{\gamma_4}\right) \tag{13}$$

Since γ is the sum of the absorption due to the individual system elements it is possible to estimate the optimum exposure time by summing the time contribution due to the individual elements. For example, the system elements contribute to the exposure time as listed below:

Be	96	Kw. seconds. micron^{-1}. at 50cm.
Photoresist	113	Kw. seconds. micron^{-1}. at 50cm.
Mylar/Si	319	Kw. seconds. mciron^{-1}. at 50cm.
He	13	Kw. seconds. cm^{-1}. at 50cm.

This approximation is useful in assessing system tradeoffs. As an example, and exposure time budget for an X-ray lithography system is:

	Exposure Time (Kw. seconds at 50cm)
1μm photoresist	113.0
3μm mylar mask	956.0
50cm Helium	655.0
25μm Be Vacuum Window	2407.0
Total	4131.0

This estimated time is slightly higher than that calculated using the exact equation. The usefulness of the approximate equation is in indicating the most important sources of increased time. In the above example it is clear that the largest contribution to the exposure time is that due to the vacuum window. The addition of a vacuum window to permit exposure in helium has resulted in almost a factor of four increase in the exposure time. Unless this change permits a similar increase in power input the net effect will be to increase the exposure time.

Effect Of Photoresist Additives On The Exposure Time

The approximate equation 13 illustrates that the minimum exposure time is determined by the ratio of the total system absorption to the photoresist absorption. Reductions in this ratio translate linearly into reductions in exposure time. The ratio can be reduced either by decreasing absorption in the system elements or by increasing absorption in the photoresist. The latter can be achieved by adding high atomic number elements to the photoresist.

The limiting case occurs when the photoresist dominates the system absorption. In this case $\gamma = \gamma_4$ and the minimum exposure time possible is obtained. A minimum exposure time of 120 K.watt.seconds is obtained for a 1μmicron photoresist film at 50cm. source to substrate distance. In practical systems the photoresist is the least absorbing system element and $\gamma \geq 10\gamma_4$ and hence the optimum exposure time is greater than 1200 kilowatt seconds. For even the simplest, (i.e. least absorbing) system more than an order of magnitude increase in resist absorption is required to obtain $\gamma = 2\gamma_4$ and an exposure time of 240 K.watt seconds. Therefore it appears unlikely that this technique alone can result in an exposure time of less than 1000 Kilowatt. seconds at 50cm. Further reduction can only result from improvements in resist sensitivity or parallel exposures.

References

1. J.S. Greeneich, IEEE Trans on Electron Devices, ED-22,7 1975
2. D. Maydan, G.A. Coquin, J.R. Maldonado, S. Somekh, D.Y. Lou, and G.N. Taylor, IEEE Trans on Electron Devices, ED-27,7, 1975

TABLE I

MYLAR/SILICON MEMBRANE THICKNESS (μ m)	OPTIMUM WAVELENGTH (A^{o})		OPTIMUM EXPOSURE TIME (KILOWATT SECONDS)	
	MEMBRANE PHOTORESIST	MEMBRANE PHOTORESIST VACUUM WINDOW 50cm He	MEMBRANE PHOTORESIST	MEMBRANE PHOTORESIST VACUUM WINDOW 50cm He
0	21.8	-	180	-
1	14.0	6.9	560	3270
3	10.3	6.6	1200	3650
6	8.4	6.2	2000	4270
10	7.0	5.7	3150	5270

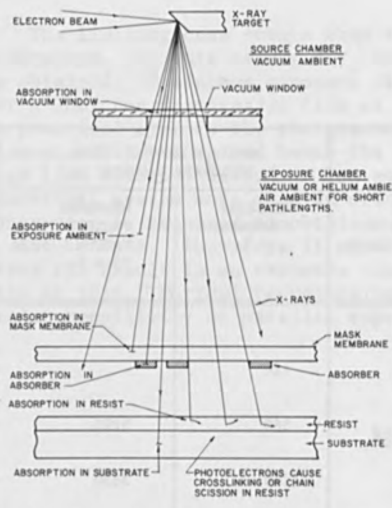


Fig. 1 - An X-Ray Lithography Exposure System

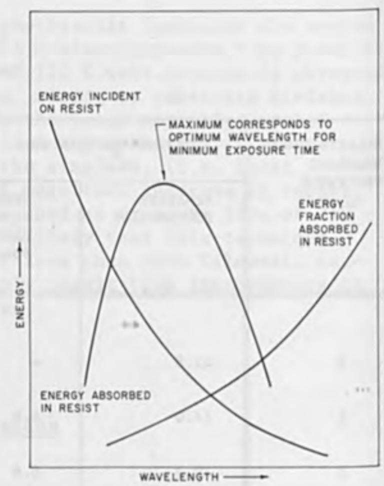


Fig. 2 - Factors Which Determine the Optimum Exposure Wavelength

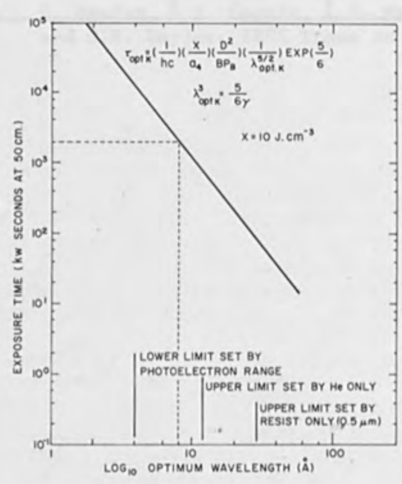


Fig. 3 - Minimum Exposure Time Vs Optimum Exposure Wavelength

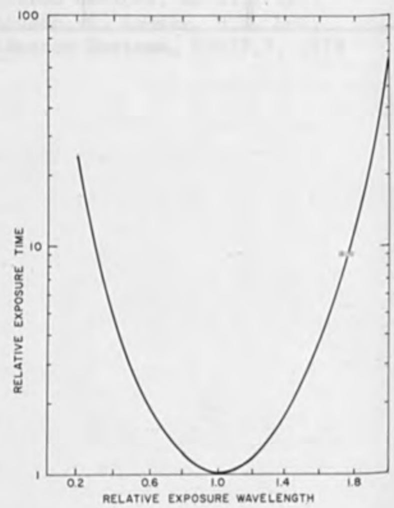


Fig. 4 - Relative Exposure Time Vs Exposure Wavelength

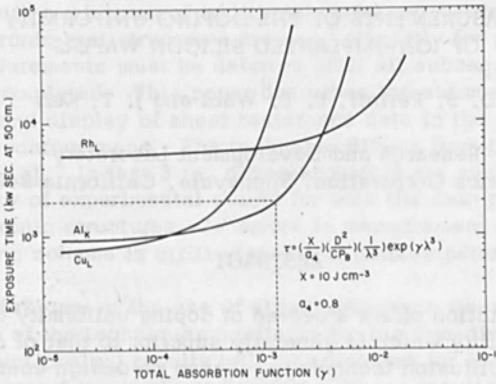


Fig. 5 - Exposure Time Vs Total Absorption of the System

MEASUREMENTS OF THE DOPING UNIFORMITY OF ION-IMPLANTED SILICON WAFERS

D. S. Perloff, F. E. Wahl and J. T. Kerr

Research and Development Laboratory
Signetics Corporation, Sunnyvale, California 94086

ABSTRACT

Ion implantation offers a degree of doping uniformity and reproducibility which is generally superior to that of conventional diffusion technology. Based on design considerations, most production-type implanters are capable of providing 2σ uniformity to better than 2% over 3 in. diameter wafers and repeatability of ±5% for the commonly used silicon dopants. This paper describes a technique which employs iso-sheet resistance topographs as a means of exhibiting the spatial distribution of impurities. Results are presented for electrostatic scanning and mechanical scanning implantation systems which demonstrate the importance of generic factors in limiting uniformity.

I. INTRODUCTION

Ion implantation was originally thought of as an efficient low temperature doping process which, because of its inherent control and spatial homogeneity, would replace diffusion technology for the fabrication of microelectronic circuits in silicon. Historically, however, implantation gained acceptance as a specialized processing tool. For example, it is used in conjunction with standard diffusion for high value resistors,[1] threshold shifting,[2] and self-aligned gate transistors.[3] In the future, as diffusion processes become inadequate for fabricating smaller device geometries, implantation is likely to assume a primary role in microcircuit fabrication. More sophisticated characterization techniques will then be required to control and monitor the implantation operation, particularly for large ($\gtrsim$3 in.) diameter wafers.

The primary method of characterizing process uniformity and repeatability involves the measurement of sheet resistance. The choice of test vehicle depends on the function which the sheet resistance data are intended to serve. For example, the four-point probe is ordinarily used for process <u>control</u>, since it enables one to characterize an impurity layer with a minimum of additional wafer processing. In contrast, microelectronic test structures are used primarily for process <u>monitoring</u>, since measurements must be deferred until all subsequent processing has been completed. This paper describes the automated collection, analysis, and display of sheet resistance data in the form of easily interpreted 1%-contour maps. The technique differs from that described by Stephen, <u>et al.</u>[4] in that 3 in. diameter wafers are evaluated using a uniform density of experimental points for both the four-point probe and microelectronic structures, 2σ errors in measurement are below 0.5%, and a coding scheme is utilized which optimizes pattern recognition.

After a review of the use of sheet resistance measurements and a discussion of the techniques available for the two-dimensional display of spatial uniformity, results will be presented for four-point probe measurements of ion implanted wafers. These results will demonstrate the value of wafer mapping as a means of displaying such generic sources of non-uniformity as geometric effects in electrostatic scanning implanters and velocity effects in mechanical scanning implanters.*

II. MEASUREMENT TECHNIQUES

1. Microelectronic Test Patterns

It is common practice to determine sheet resistance by diffusing or implanting impurities into photolithographically defined test structures such as the resistor or bridge-shaped patterns shown in Figs. 1a and 1b, respectively. These structures, however, are sensitive to dimensional variations in the photomask, as well as in the exposure and etching steps used to delineate the pattern.[5] Van der Pauw[6] has described a test vehicle (Fig. 1c) which is not sensitive to such effects. The sheet resistance for a van der Pauw pattern is

$$\rho_s = \frac{\pi}{\ln 2}\,\frac{R_A + R_C}{2}\, f\left(\frac{R_A}{R_C}\right), \tag{1}$$

where R_A and R_C are obtained using the independent current/voltage configurations shown in Table I and $f(R_A/R_C)$ satisfies the relation

*Results for conventionally diffused wafers also reveal generic sources of non-uniformity. These will be discussed elsewhere.

$$\frac{\frac{R_A}{R_C}-1}{\frac{R_A}{R_C}+1} = \frac{\cosh^{-1}\left[\frac{1}{2}\exp(\ell n2/f)\right]}{\ell n2/f} . \tag{2}$$

When a symmetric pattern is employed, $R_A \simeq R_C$ and consequently $f(R_A/R_C) \simeq 1$ in Eq. (1).

Buehler[7] has discussed the magnitude of the error introduced when using the orthogonal-type van der Pauw structure shown in Fig. 1c. Even for the limiting case in which $s \to d$, Eq. (1) is correct to better than 0.1% as long as $a/s \geq d/s$. Experience with orthogonal van der Pauw structures has shown that even minimum geometry ($d = s = 10\ \mu m$, $a/d = 3$) configurations are virtually insensitive to dimensional variations when Eq. (1) is used to determine sheet resistance.[8]

The complexity of the processing required to fabricate microelectronic test patterns frequently results in erratic readings which have an adverse effect on interpreting doping uniformity. Also, it is difficult to implement such structures for routinely monitoring doping operations, particularly when it is necessary to employ a metallization pattern for making electrical contact. For these reasons, it becomes attractive to consider the use of a vehicle which can be employed with a minimum of wafer processing, a requirement which is met by the four-point probe.

2. The Four-Point Probe

Fig. 2 shows the in-line four-point probe (FPP) described by Valdes.[9] If the average probe separation is at least twice the thickness of the doped layer, the sheet resistance ρ_s will be given by

$$\rho_s = k\frac{\Delta V}{I} , \tag{3}$$

in which k is a geometric factor, I is a current through two of the probes, and ΔV is the potential difference measured between the remaining two probes.[10] Ordinarily, one chooses the outer probes to carry current and the inner probes to measure voltage since, for an FPP having nearly equal distances of separation between the probe needles, this assignment results in the largest value of ΔV.[11]

Logan[12] has shown that a circular sample doped on both sides should behave as if its front surface were part of a continuous infinite sheet for which the correction factor is a constant value ($k=\pi/\ell n2=4.532$), independent of the location of the FPP. In contrast, a thin circular sample doped on one side will require knowledge of the orientation and position of the probe array in order to utilize Eq. (3). General expressions have been derived for the case of the FPP oriented perpendicular to the radius[12] and for the FPP oriented along the radius[13] of a thin circular sample. Perloff[14] has derived expressions for the FPP oriented arbitrarily with respect to the wafer radius, making it possible to employ automatic x-y wafer probing equipment for evaluating doping uniformity.

An alternative approach to FPP measurements involves the use of photolithographic techniques to define regions of sufficient size to accommodate the FPP array, as shown in the cross-sectional view of Fig. 3. If an automatic wafer prober is employed to position the FPP so that it occupies the same relative location in each pattern, k will be a constant over the entire wafer as long as the dimensions of the pattern are large compared to the dimensional variability discussed above. For the case of the FPP array positioned symmetrically within a rectangular test pattern, the solution for k calculated by Smits[10] may be employed.

The use of the FPP for evaluating the sheet resistance of junction-isolated implanted layers has been discussed by Stephen[15] who identified the following sources of error when making such measurements: surface effects, junction leakage, geometrical factors, magnitude of the measuring current, probe loading, and ambient temperature. In practice, by employing commercially available "precision" probe assemblies and restricting the measurement to layers greater than 2000 Å in thickness and less than 500 ohm/sq, one can achieve an absolute accuracy of ±1% and a 2σ repeatability of 0.5%. The primary source of error, under these circumstances, is associated with variation in the distances between adjacent probe needles due to mechanical tolerances of the probe assembly.[8]

III. ACQUISITION AND DISPLAY OF DATA

An automated test system is employed for the acquisition of sheet resistance data which consists of an x-y wafer prober, prober control, data terminal, recording devices, and sequencing and interface modules. Data are analyzed off-line by a minicomputer with associated program and system disk drives, tape reader, and terminal. An electrostatic printer/plotter is employed for displaying the data either as a map of per cent deviation from the wafer mean, as a two-dimensional histogram, or as an iso-sheet resistance topograph.

1. Deviation Maps

In this study, values of ρ_s are determined at 118 standard test sites located on the square grid shown in Fig. 4a. The distance between grid points is 222 mils for a 3 in. diameter wafer. The mean value and standard deviation are computed according to the expressions

$$\bar{\rho}_s = \frac{1}{118} \sum_{i=1}^{118} (\rho_s)_i \tag{4}$$

and

$$\sigma = \frac{1}{117} \sum_{i=1}^{118} [(\rho_s)_i - \bar{\rho}_s]^2 . \tag{5}$$

The relative variability is defined as

$$v^* = 2\sigma / \bar{\rho}_s . \tag{6}$$

For an ideal gaussian distribution, Eq. (6) implies that 95% of the readings fall between $(\rho_s)_{min}$ and $(\rho_s)_{max}$ where

$$(\rho_s)_{min} = \bar{\rho}_s - 2\sigma = \bar{\rho}_s (1 - v^*) \tag{7}$$

and

$$(\rho_s)_{max} = \bar{\rho}_s + 2\sigma = \bar{\rho}_s (1 + v^*) . \tag{8}$$

Fig. 4b is an example of a deviation map for the case of a 3 in. diameter phosphorus diffused p-type wafer. Values of $(\rho_s)_i$ obtained from FPP measurements (bare wafer, k = const.) were used to compute deviations from the wafer mean according to the formula

$$\Delta_i = [(\rho_s)_i / \bar{\rho}_s - 1] \times 100\% . \tag{9}$$

This method of presenting numerical data is preferable to mapping $(\rho_s)_i$ values, since portions of the wafer above $\bar{\rho}_s$ are identified by + signs and below $\bar{\rho}_s$ by - signs. Rudimentary contours defining regions of equal Δ_i may easily be sketched once the data have been analyzed and presented in this fashion.

2. Histographs

It is common practice to employ frequency plots or histograms to summarize the distribution of data over a range of values. Fig. 4c is a histograph (two-dimensional histogram) which associates each measured value, and the interpolated values discussed below, with one of seven intervals into which the range $(\rho_s)_{max} - (\rho_s)_{min}$ has been divided. These intervals, each of which contains an equal number of values, are identified by circles of various line thicknesses and the symbols –, ·, and +. Unlike the deviation map for which the +(–) sign identifies values which are above(below) the <u>mean</u>, this method of display identifies values which are above(below) the <u>median.</u>

In order to obtain the histograph shown in Fig. 4c, it was necessary to first subdivide the initial network of data points (Fig. 4a) by employing an interpolation scheme which results in a four-fold increase in the density of grid points. The distance between subgrid points is 111 mils for a 3 in. diameter wafer.

3. Iso-Sheet Resistance Topographs

In this paper, iso-sheet resistance topographs will be used to display uniformity data. Fig. 4d shows such a topograph obtained from the numerical data of Fig. 4b. The heavy contour represents the mean sheet resistance $\overline{\rho}_s$, computed using Eq. (4), while the lighter contours differ from $\overline{\rho}_s$ in increments of $0.01\overline{\rho}_s$. The symbol +(–) identifies an experimental point whose value is greater(less) than $\overline{\rho}_s$. At the lower left of each topograph is printed $\overline{\rho}_s$; at the lower right, $\pm v^* \times 100\%$, obtained from Eq. (6).

As in the case of the histographs discussed above, interpolation has been used to increase the density of grid points. Without doing so, the contours would be extremely irregular in appearance as a consequence of the use of straight line segments for their construction. In addition, a linear extrapolation scheme has been employed which allows contours to extend to the edge of the topograph. This approach considerably enhances one's ability to comprehend and interpret the patterns of uniformity which invariably exist in implantation and diffusion processes.

IV. EXPERIMENTAL RESULTS

The topographs to be presented in this section were obtained from fully annealed boron implanted wafers under the conditions described in Table II. Measurements of ρ_s were in most cases performed using a standard FPP in conjunction with a rectangular test pattern. This pattern, which is typically 110 x 220 mil^2, represents 0.34% of the area of a 3 in. diameter wafer. Similar results have been obtained with van der Pauw structures having considerably smaller active areas.[8]

1. Electrostatic Scanning Implanters

Fig. 5 depicts schematically a typical production implantation system[16] employing electrostatic deflection of the ion beam. Neutral impurities which pass through the gridded lens are not deflected by the neutral beam trap (NBT) plates. Under poor vacuum conditions, ions may be neutralized after entering the NBT plates, thereby inflicting the characteristic neutral beam pattern evident in the topograph of Fig. 6. The horizontal lobe of this pattern is associated with neutralization within the NBT plates; the vertical lobe with neutralization within the y-axis scanning plates. (In extreme cases, neutral beam doping can be as much as 50% greater than that associated with ionized impurity doping.) The topograph of Fig. 6 also exhibits a doping gradient along the x-axis which is a consequence of tilting the wafer with respect to the ion beam about the y-axis. The right side of the wafer, which is further from the scanning plates than the left side, is more lightly doped.

Fig. 7 illustrates the high degree of uniformity which can be achieved by an optimally adjusted electrostatically scanning implantation system. In this case, the same implantation unit was employed as for the wafer of Fig. 6, but was operated at a vacuum level which was sufficiently high to remove almost all evidence of a neutral beam. Since the wafer was not tilted relative to the ion beam, no gradient is present.

Three examples of the effect of wafer tilting are shown in Fig. 8 for angles of 3°, 7°, and 14°. The mean sheet resistance is observed to increase and the uniformity decrease as the angle of inclination increases. The top of each wafer also exhibits the effect of underscanning as identified by the * symbol which represents values outside of $\pm 10\sigma$ sorting limits.

The topograph of Fig. 9 is for a wafer (7° tilt) implanted under conditions of unstable beam current, achieved by detuning the ion source until the output current was caused to oscillate. No degradation of uniformity can be detected relative to the wafer of Fig. 8b, which was implanted under stable operating conditions.

The topograph of Fig. 10 was obtained immediately after venting the wafer target chamber to room conditions and then restarting the pumping cycle. Apparently only a small amount of neutralization occurred, which suggests that the vacuum level in the vicinity of the scanning plates remained relatively unaffected.

The pattern of uniformity for a severely underscanned wafer is shown in the topograph of Fig. 11. In this case, log ρ_s contours (10/decade) are employed to characterize the variation in sheet resistance. The value of ρ_s at any of the 118 test sites may be recovered by raising 10 to the specified power. For example, near the center of the wafer, log $\rho_s = 1.5$ so that $\rho_s = 10^{1.5} = 31.6$.

2. Mechanical Scanning Implanters

In recent years, interest in applications requiring implantation doses $\gtrsim 10^{15}$ cm^{-2} has led to the development of implanters which employ mechanical scanning of the wafer relative to the ion beam. Such systems permit high ion currents to be utilized by distributing the deposited energy over a number of wafers.[17] Two such implantation systems, to which the authors have had access, will be discussed in this section. Both employ mechanisms which drive a group of wafer carriers in a helical path by combining rotary and linear movement.

Track-Mounted Carriers. Fig. 12 is a topograph of a wafer (0° tilt) implanted in a system employing seven track-mounted wafer carriers.[18] The relatively uniform doping over the center of the wafer gives way to a rapid decrease in doping at the edges. This behavior is believed to arise from variations in the moment of inertia of the mechanical system which drives the wafer carriers through the ion beam. Fig. 13 illustrates the location of the seven wafer carriers when the ion beam strikes the center of wafer No. 1. At this time, the moment of inertia is a maximum since, on the average, the carriers are farthest from the center of rotation. The moment of inertia decreases abruptly as carrier No. 7 moves onto the linear portion of the track and carrier No. 2 enters the large radius portion of the track. When this occurs, the entire system is believed to speed up in order to conserve angular momentum. The leading edge of the next wafer to enter the beam, No. 7, experiences the effect of the system slowing down again as the velocity feedback mechanism begins to respond to inertial changes. The system has returned to its state of lower angular velocity by the time the ion beam strikes the center of wafer No. 7. As the wafer carriers rotate counter-clockwise in the x, z plane, the cassette body moves along the y-axis in order to distribute impurities over the entire wafer. The linearity of the features in Fig. 12 attest to the fact that the variations in velocity repeat themselves throughout the many traversals of the wafer through the beam.*

*Experiments are continuing in an effort to verify the inertial origins of this behavior.

In order to determine the sensitivity of sheet resistance measurements to small dose variations, a wafer was first implanted at 1.0×10^{15} B/cm^2 and then, after shielding the right side, implanted a second time at a dose of 1.0×10^{14} B/cm^2. The topograph of Fig. 14 shows that ρ_s increases by approximately 8%, while the dose decreases by 10%, in going from the left to right side of the wafer. This result is easily verified by counting the number of contours running normal to the wafer flat along the center of the wafer.

The topograph of Fig. 15 was obtained by implanting a stationary wafer which had previously been implanted in the conventional manner. The heavily doped rectangular area represents a distinctive imprint of the ion beam. More quantitative information about the variation of ion concentration is shown in Fig. 16 for the case of the beam focused in the plane of the wafer. From spreading resistance measurements[19] carried out at 250 μm intervals along the narrow portion of the rectangular profile, the effective width at 1% of peak doping density is 0.68 cm. These data show the usefulness of spreading resistance measurements for characterizing large local doping variations. Such an assessment would not be feasible using the techniques described in Sec. II.

Ferris Wheel-Mounted Carriers. The movement of wafer carriers in a system[20] employing a Ferris wheel mechanism[21] is shown schematically in Fig. 17. Since the carriers are maintained within the x,y plane as they traverse the ion beam, there is a velocity change which depends on the angle of the carrier with respect to r, the radius of rotation. The variation in dose obtained from the geometry of Fig. 17 is thus

$$\Phi = \frac{\Phi_0}{\sqrt{1 - \left(\frac{y}{r}\right)^2}} . \qquad (10)$$

The predicted behavior, based on Eq. (10), is shown in the topograph of Fig. 18a. A representative topograph of an actual wafer is shown in Fig. 18b. The observed 2σ uniformity of 1.00% agrees well with the predicted uniformity of 0.70%.

V. CONCLUSION

The mapping results presented in this paper demonstrate that implantation is capable of achieving extremely high spatial doping uniformity, as measured by the variation of sheet resistance over 3 in. diameter silicon wafers. Geometric effects in electrostatic scanning and velocity effects in mechanical scanning systems would appear to be the primary generic sources of variability under otherwise controlled operating conditions.

In the future, characterization techniques of the kind described in this paper may be expected to play an increasingly important role in monitoring and controlling the implantation of impurities. This will be particularly true if systems capable of handling wafers of diameter ≥ 4 in. are not designed in such a way as to minimize the generic sources of variability described above.

ACKNOWLEDGEMENT

The authors wish to express their appreciation for the technical assistance of B. Belloli, J. Conragan, K. Dettenreider, C. Gast, D. Mar, J. Reimer, C. Vorst, and J. Yeamans; the support and encouragement of Drs. J. Marley and D. Kleitman; and the cooperation of Varian Extrion, Lintott Engineering, and AERE Harwell in presenting data on mechanical scanning systems.

REFERENCES

1. J. D. MacDougall, K. E. Manchester, and P. E. Roughan, Proc. I.E.E.E. 57, 1538 (1969).
2. K. G. Aubuchon, Proc. Int. Conf. Properties and Uses of M.I.S. Structures, Grenoble, 1969.
3. J. M. Shannon, J. Stephen, and J. H. Freeman, Proc. Int. Conf. Properties and Uses of M.I.S. Structures, Grenoble, 1969.
4. J. Stephen, B. J. Smith, and G. W. Hinder, Proc. 4th Int. Conf. Ion Implantation in Semiconductors and Other Materials, Osaka,1974.
5. M. G. Buehler and J. M. David, to be published.
6. L. J. van der Pauw, Philips Res. Rpts. 13, 1 (1958).
7. M. G. Buehler, J. M. David, R. L. Mattis, W. E. Phillips, and W. R. Thurber, J. Electrochem. Soc. 122, 84c (1975).
8. D. S. Perloff and F. E. Wahl, unpublished work.
9. L. B. Valdes, Proc. Inst. Radio Engrs. 42, 420 (1954).
10. F. M. Smits, Bell Syst. Tech. J. 37, 711 (1958).
11. R. Rymaszewski, J. Scientific Instr. 2, 170 (1969).
12. L. J. Swartzendruber, Solid-State Electron. 7, 413 (1964).
13. M. A. Logan, Bell Syst. Tech. J. 40, 885 (1961).
14. D. S. Perloff, to be published.

15. G. Dearnaley, J. H. Freeman, R. S. Nelson and J. Stephen, Ion Implantation, North Holland, Amsterdam, 1973.
16. Implanter Model 200-20, Varian Extrion, Gloucester, Mass.
17. R. G. Wilson and G. R. Brewer, Ion Beams, Wiley, 1973.
18. Implanter Model 8-12, Processor Model 12, Lintott Engineering, Horsham, Sussex, England.
19. Proc. NBS Spreading Resistance Symp., NBS Spec. Publ. 400-10, 1974.
20. Implanter Model 200-1000, Varian Extrion, Gloucester, Mass.
21. A. B. Wittkower, P. H. Rose, and G. Ryding, Solid State Tech. 18 (No. 12), 41 (1975).

TABLE I. Electrode Arrangement for Measurements of the van der Pauw Configuration of Fig. 1c.

I,V Config.	(I+,I-)	(V+,V-)
A	(1,4)	(2,3)
B	(1,3)	(2,4)
C	(1,2)	(4,3)

TABLE II. Implantation Conditions for the Topographs of Figs. 6 - 18.

Boron dose: $5 \times 10^{14} - 4 \times 10^{15}$ cm^{-2}

Energy: 60 - 155 keV

Thickness of SiO_2 passivation layer: 1300 Å

Annealing: 950 C, 30 min, dry N_2

Substrate: 6 - 9 ohm-cm n-type, 111-orientation

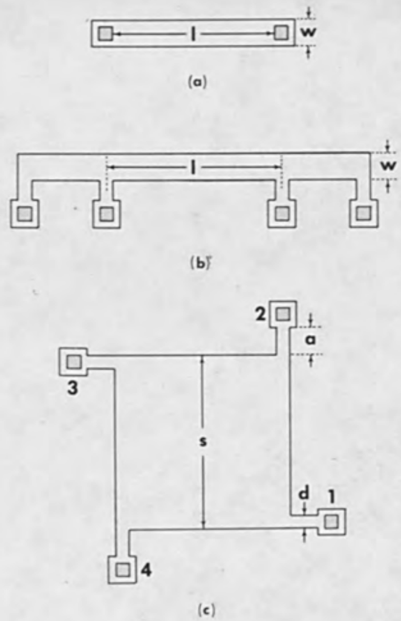


Fig. 1. Microelectronic patterns for characterizing sheet resistance. The resistor pattern (a) involves a two-probe measurement and is sensitive to variations in l and w. The bridge-shaped pattern (b) involves a four-probe measurement and is therefore able to eliminate contact resistance contributions. The van der Pauw configuration (c) also involves a four-probe measurement but is independent of the dimensions a, d, and s as long as $a/s \geq d/s$.

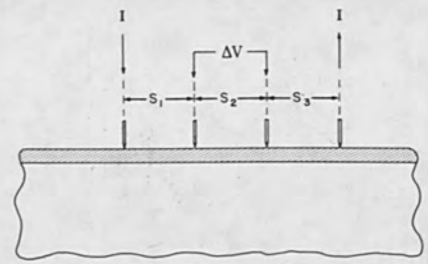


Fig. 2. Conventional current/voltage configuration for an in-line four-point probe array. The sheet resistance for a thin layer is determined from Eq. (3) when $s_1 = s_2 = s_3$.

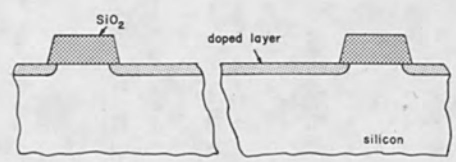


Fig. 3. Cross-sectional view of the rectangular test pattern used in conjunction with an in-line four-point probe to determine the sheet resistance of junction isolated doped layers.

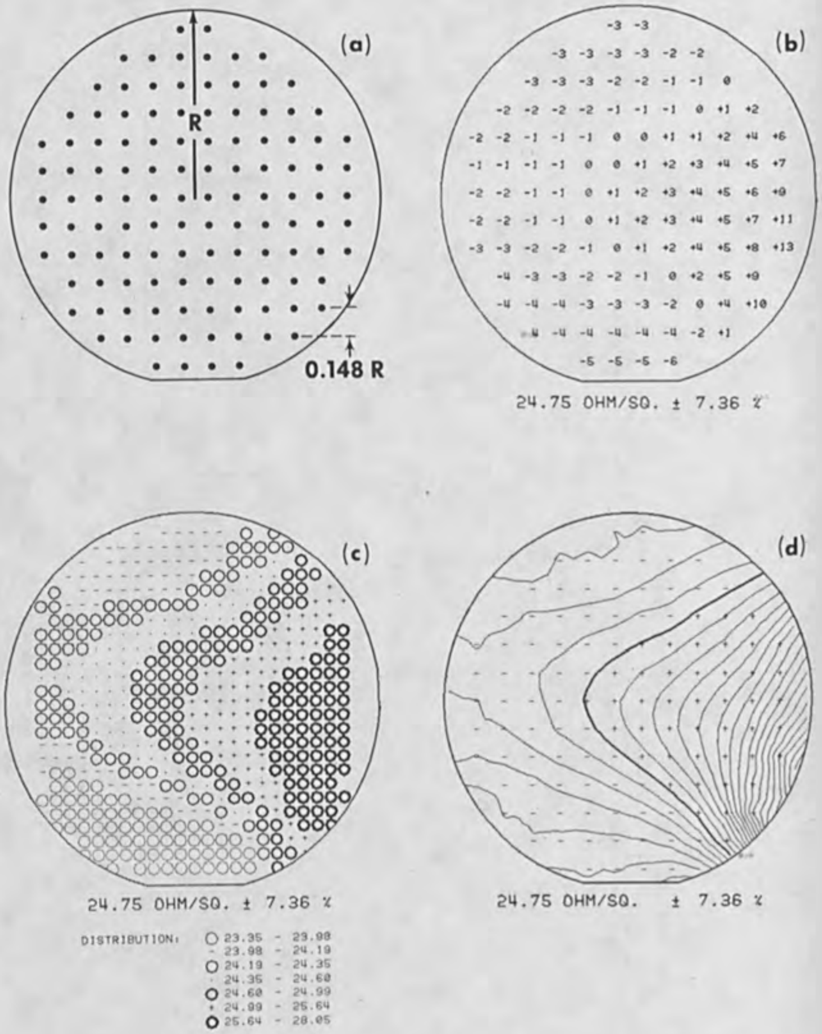


Fig. 4. Methods of displaying uniformity data: (a) standard test-site configuration (118 data points); (b) deviation map; (c) two-dimensional histogram; (d) iso-sheet resistance topograph. The data were obtained from four-point probe measurements of a phosphorus diffused p-type wafer.

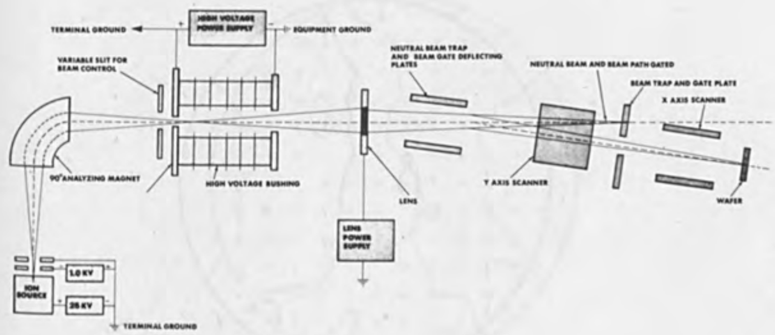


Fig. 5. Production implantation system employing electrostatic scanning of the ion beam. (Ref. 16)

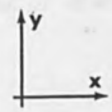
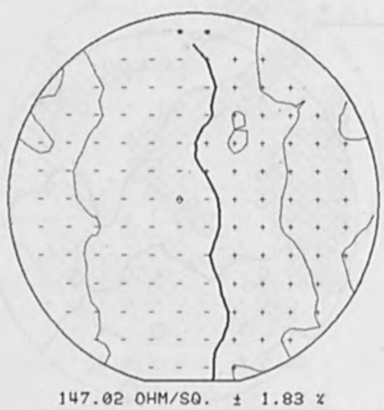


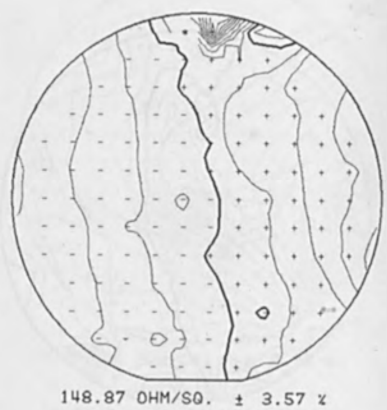
Fig. 6. Topograph exhibiting a neutral beam pattern and a gradient associated with inclining the wafer with respect to the ion beam of the electrostatic scanning implanter shown in Fig. 5.

Fig. 7. Topograph of a highly uniform wafer implanted by the same equipment as that used for the example of Fig. 6.

(a)



(b)



(c)

Fig. 8. The doping gradient obtained when wafers are tilted with respect to the ion beam of the system shown in Fig. 5: (a) 3°, (b) 7°, (c) 14°.

Fig. 9. Wafer implanted under conditions of unstable beam current.

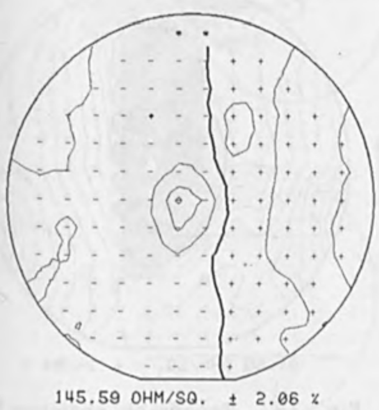


Fig. 10. Wafer implanted after venting target chamber and restarting pumping cycle.

Fig. 11. Logarithmic plot of sheet resistance for a severely underscanned wafer.

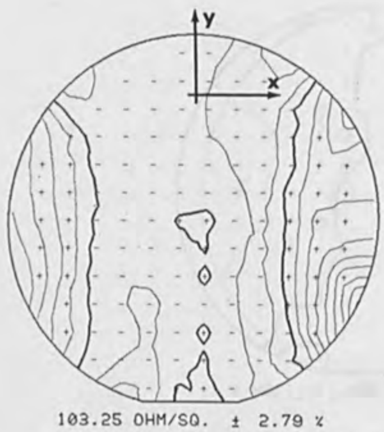


Fig. 12. Characteristic variation in doping density which arises from velocity variations in the mechanical system shown in Fig. 13.

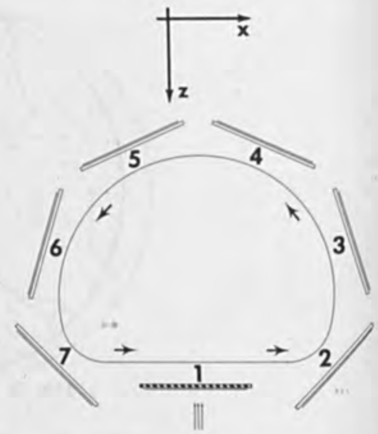


Fig. 13. Schematic representation of a target cassette used with the implanter of Ref. 18. The wafer carriers are in a configuration of maximum moment of inertia.

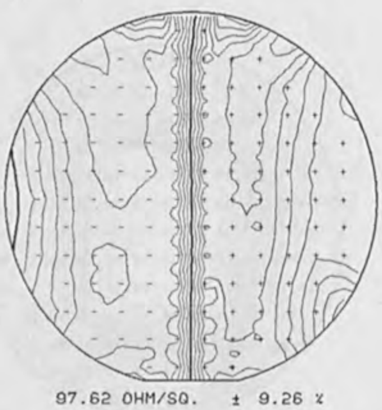


Fig. 14. Topograph showing the sensitivity to a 10% variation in boron dose. The left side of the wafer received a dose of 1.1×10^{15} B/cm^2; the right side a dose of 1.0×10^{15} B/cm^2.

Fig. 15. Topograph showing the spatial distribution of the ion beam in the implantation system of Ref. 18.

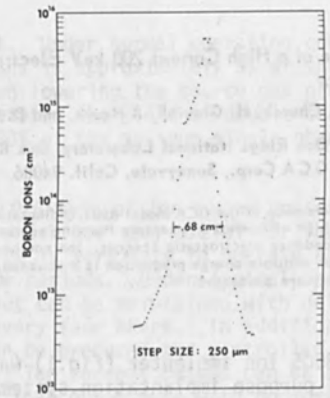


Fig. 16. Beam profile obtained from spreading resistance measurements of a boron implanted silicon wafer such as the one shown in Fig. 15.

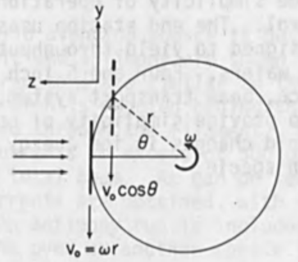


Fig. 17. The rotational mechanism employed in the implantation system of Ref. 20. The variation in linear velocity which results from maintaining the wafer normal to the ion beam causes the distinctive pattern of doping shown in Fig. 18.

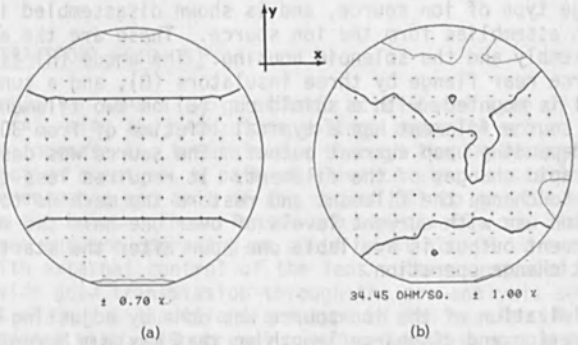


Fig. 18. (a) Simulated pattern of uniformity and (b) representative wafer implanted in the implantation system of Ref. 21.

Design and Performance of a High Current 200 keV Electrostatically Scanned Ion Implantation Accelerator

R. Allison, Jr., D. Church, H. Glavish, A Noeth, and D. Schechter*

Present address: *Oak Ridge National Laboratory, Oak Ridge, Tenn. 37830
GCA Corp., Sunnyvale, Calif. 94086

The design and performance of the GCA Model 8501 Implantation System is described. Features include a high efficiency filamentary Penning source, a variable focus analysis system, and low impedance electrostatic scanner. Ion source performance over wide current ranges and on multiple-charge production is discussed. Data on wafer heating and implant uniformity are presented.

1. INTRODUCTION

The GCA Model 8505 ion implanter (Fig.1) was designed to provide a flexible, general purpose implantation system for both production and research applications. The accelerator produces boron currents in excess of 150 uA, and phosphorus and arsenic currents in excess of 300 uA. The implanter covers an energy range of 30 to 200 keV with automatic tracking of controls when the energy is varied. Electrostatic scanning is used to provide simplicity of operations, as well as good uniformity and dose control. The end station uses batch loading and the vacuum system was designed to yield throughputs of 200 wafers per hour for 2 and 3 inch wafers. Four and 5 inch wafers can also be implanted. The ion source, beam transport system, and implanter controls were designed to provide simplicity of operation, while at the same time permit rapid changes in ion energy and current levels, as well as changes in ion specie.

2. ION SOURCE

The ion source used on the implanter is a filamentary Penning discharge type of ion source, and is shown disassembled in Fig. 2. Two main assemblies form the ion source. These are the anode-filament assembly and the solenoid housing. The anode (H) is secured to the source rear flange by three insulators (G), and a tungsten filament (F) is mounted, with a shield cup (E) on two filament posts (D). The ion source filament has a typical lifetime of from 30 to over 40 hours, depending upon current output. The source was designed to permit rapid changes of the filament. It requires less than thirty minutes to change the filament and restore the machine to normal implantation use with current levels of over one-half the maximum. Full current output is available one hour after the start of the filament change operation.

Optimization of the ion source was done by adjusting the anode configuration and discharge length so that maximum boron, phosphorus and arsenic currents were obtained. Typical spectra for these specie are presented in Fig. 3. Note the relatively high content of boron obtained using boron trifluoride, and the very low content of diatomic specie present. Under normal operating conditions the output of double charged ions is approximately 5% of the yield of the singly charged specie. Upon lowering the source gas pressure and adjusting the operating parameters, the yield of double charged current can be increased to about 20% of the maximum single charged current for each specie.

The filamentary PIG-type of ion source possesses several operating characteristics which are very useful for implantation applications. The ion source is very easy to operate and gives stable current output over long time periods. Under typical operating conditions maximum current output can be maintained with only minor source adjustments required every four hours. In addition, a very large range of current levels can be produced and controlled using only external ion source controls. The current output can be adjusted over its entire range in less than one minute with no loss of stability. When run on boron the current output covers the range of 15 nA to 150 uA, giving an effective dose rate range of 1×10^9 to 1×10^{13} ions per cm^2 per sec. Phosphorus and arsenic cover the same range with twice the maximum current output.

A solid oven ion source is being tested, and preliminary spectra for arsenic and antimony are shown in Fig. 4. The ion source discharge parameters are similar to the gas source described above. A cartridge oven is mounted behind the filament and the vaporized solid is injected through the filament. Helium is used as a carrier gas. Arsenic currents in excess of 500 uA have been achieved with 2 mA extracted total beam. As can be seen from Fig. 4, very good multiple charge currents are obtained, with very small amounts of diatomic specie. An antimony run is included, which shows the effect of running the oven on another specie (note the small amount of arsenic contamination).

3. BEAM TRANSPORT (FIG. 5)

The implanter uses a two electrode accel-decel type of extraction system (A), with the extraction voltage normally set at 28 kV. In order to increase the transmission through the analyzing magnet to the rejection aperture, a pair of electrostatic lenses (D) are used at the entrance and exit of the magnet (E) to provide vertical focusing to the beam. These lenses are essentially one dimensional Einzel lenses and are operating at about 14 kV for 28 kV extraction voltage, with external control of the lens voltage provided. These lenses provide good transmission through the pre-analysis section independent of source conditions and type of ion specie. A 65^o analyzing magnet with a 19 cm radius is used with normal entry and exit. The magnet was designed to provide adequate mass resolution while at the same time minimizing space and weight in the pre-acceleration section.

The post analysis acceleration is done using a single gap type of column (F). This section is composed of two graded ceramic columns, the second column being used to provide differential pumping of the pre-analysis section. A triplet quadrupole electrostatic lens (H) is used to focus the beam at the target position. This lens was designed to operate at 10 kV for a beam energy of 200 keV, and to produce a spot size of 1 cm at the target. At maximum current levels and low implantation energies the spot size increased due to space charge effect; however, this effect is compensated for by increasing the scan voltages to ensure good uniformity. While the implanter was designed to operate from 30 to 200 keV under production conditions, it is possible to run at lower energies and boron currents in excess of 50 uA can be transported at 10 keV.

4. SCANNING

A two dimensional electrostatic scan system is used to scan the beam over a stationary target. In addition to scanning the beam, the second set of plates is used to give a 5^{o} offset to the beam direction in order to remove neutral contaminants. With this arrangement any neutrals which do reach the target will be completely scanned and distributed uniformly across the surface of the wafer. The 5^{o} offset and scanning are done on one set of plates in order to avoid the space and neutral problems associated with using a third set of plates. However, in order to combine a D.C. offset with A.C. scanning on one set of plates, a low impedance scanner was required, especially at high beam currents. This was accomplished using the arrangement shown in Fig. 6 and provides very stable operation independent of beam current. The low impedance is achieved using two high voltage transformers which have high frequency cores. The output voltage is fed back to the driver amplifier and corrected, so as to provide a linear output voltage. The scan signal applied to the plates is linear to less than one percent and results in good uniformity across the surface of the wafer, as demonstrated by Figure 7.

Overscan is adjusted using intercepting monitor plates (Fig.5,K) and beam shuttering is done mechanically (Fig. 5,J). All monitors are biased to suppress secondary electron effects. Absolute dose accuracy is better than 5 percent and repeatability from wafer to wafer and batch to batch is better than 1 percent.

5. TARGET CHAMBER

A target chamber is shown in Fig. 8 with a loaded wafer tray (C). The tray holds twenty-five 2 or 3 inch wafers, and twelve 4 or 5 inch wafers. In addition, a test position is also included for setup and is used to check for proper focusing and overscan. After loading a tray of wafers, the operator initiates the pumpdown of the chamber and the remainder of the implant cycle is completely automatic. The chamber is first evacuated to the desired implant pressure and each wafer is then implanted to a predetermined dose; after the last wafer is completed the operator is signaled to remove the tray. The chamber can also be used for individual implants, and the machine controls were designed to permit rapid changes in dose and accelerator energy.

In order to ensure good vacuum in the implanter two 6 inch oil diffusion pumps are used. One pump is used on the ion source and acceleration region (Fig. 5, A) and the second pump is used on the target chamber and remaining beam line (Fig. 5, B). In addition, a liquid nitrogen cryogenic pump is attached directly to the target chamber. With this arrangement the chamber can be evacuated from atmospheric pressure to 4 to 8 x 10^{-6} Torr in a time of 90 to 120 sec. For most applications implantation can begin in this pressure region as the total neutral beam reaching the wafer is less than one percent, and the neutrals that do reach the wafer are distributed uniformly across the surface of the wafer due to the neutral beam trap arrangement used on the implanter.

In some applications heating of the wafer is of some concern. When operated at maximum current and energy levels, the heat input to the wafer is in excess of 0.6 watts per cm^2. Unless good thermal contact is made across the complete rear surface of the wafer, large and often erratic temperature rises can occur. In order to minimize and control the temperature of the wafer during implant a special wafer tray was developed which provides for good thermal conductivity across the back surface of each wafer. Fig. 9 shows the difference in temperature rise for the cases of good and poor thermal contact. With this wafer holder the implanter can be used at maximum current with a resulting wafer temperature during implant of less than 150°C.

6. SUMMARY

The GCA Model 8505 ion implanter is a multipurpose implantation system designed to provide reliable operation with high throughput for production applications, while at the same time provide very flexible operation over a wide dose and energy range for research use. The energy range covered is 30 to 200 keV, and lower energy operation is available with good beam transport efficiency. Dose rates from 1 x 10^9 to in excess of 2 x 10^{13} ions per cm^2 per sec. are available, and the implanter was designed to permit rapid changes in dose rates and implantation energy. The wafer processing chamber was designed for high throughput while at the same time possesses the simplicity and reliability of batch loading. Special attention has been given to the cooling of wafers during high current implants.

7. ACKNOWLEDGMENT

The authors would like to thank Dr. D. S. Perloff of Signetics Corp., Sunnyvale, California for permission to use Fig. 7.

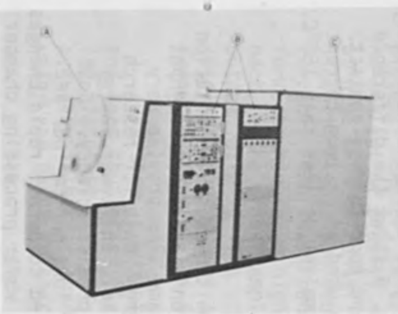


Figure 1 GCA Model 8505 Implanter

View from operator's side showing wafer chamber (A), controls (B), and high voltage enclosure (C).

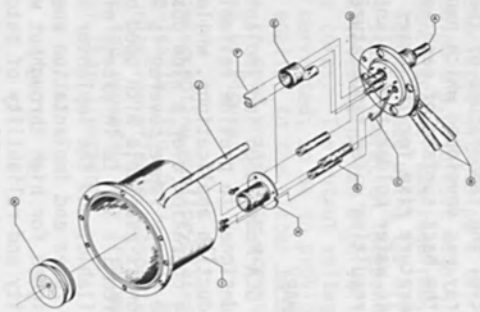


Figure 2 GCA Filamentary Penning Ion Source

Filament feed-thru (A), source cooling line (B), anode lead (C), filament posts (D), filament shield cup (E), filament (F), anode insulators (G), anode (H), solenoid housing (I), solenoid cooling (J), source aperture (K).

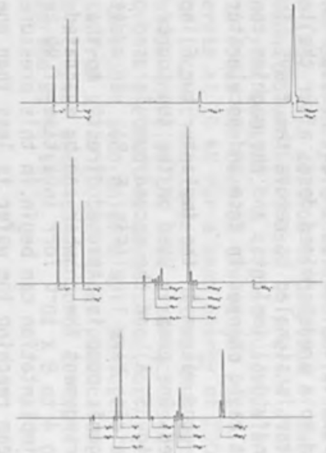


Figure 3 Ion Current Spectra-Penning Filamentary Ion Source

Top: 300 uA of $^{75}As^+$ using a mixture of 15% arsine and 85% hydrogen.

Middle: 400 uA of $^{31}P^+$ obtained using pure phosphine.

Bottom: 200 uA of $^{11}B^+$ obtained using boron trifluoride.

Spectra were taken by measuring target current vs. analyzing magnetic field. Please note spectra are not to the same scale.

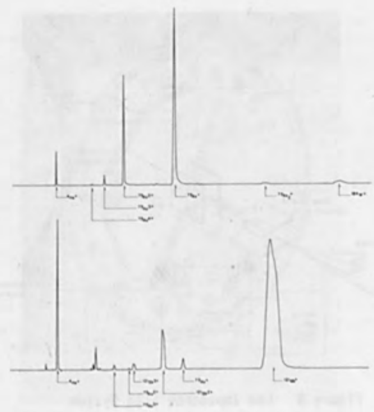


Figure 4 Ion Current Spectra - Solid Oven Ion Source

Top: Arsenic with 350 uA of ^{75}As and 215 uA of $^{75}As^{2+}$.

Bottom: Antimony with 60 uA of $^{121}Sb^{+}$.

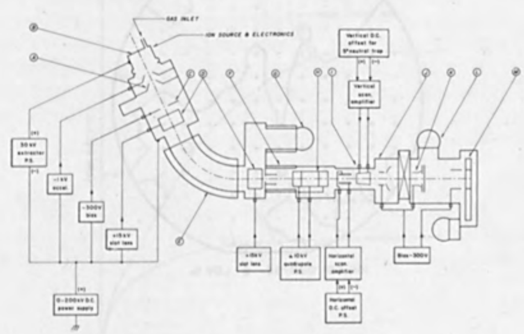


Figure 5 GCA Model 8505 Ion Implanter Beam Line Configuration

Two electrode extractor (A), ion source (B), bias plate (C), rectilinear slot lens (D), analyzing magnet (E), accelerating column (F), source diffusion pump (G), triplet quadrupole (H), deflection plates (I), shutter (J), overscan monitors (K), chamber diffusion pump (L), wafer wheel (M).

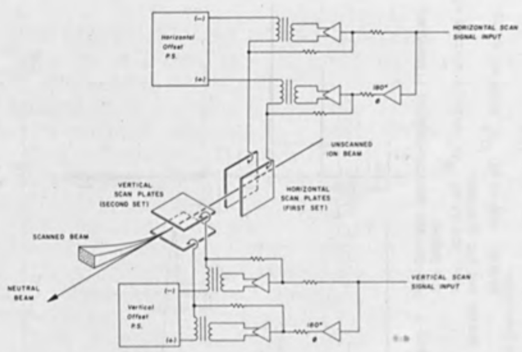


Figure 6 Low Impedance Scan System

The 5^{0} neutral beam tray is combined with vertical scan plates, and the horizontal offset is used for steering correction.

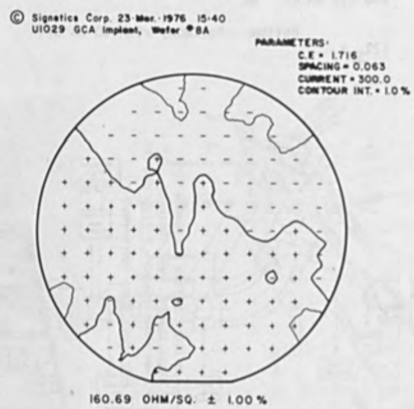


Figure 7 Uniformity profile, Boron Implant at 150 keV. Total dose 5 x 10^{14} ions per sq. cm.,dose rate 8 x 10^{12} ions per sq. cm per sec.

The heavy contours indicate the mean sheet resistance and the lighter contours differ from the mean in incrementsof 1%. The symbols +, - signify the location of experimental points with resistance values greater or less than the mean value. Twice the relative standard deviation was $\pm$ 1%.

Figure 8 GCA Model 8505 Ion Implanter 3 inch Wafer Chamber

Chamber door (A), wafer in spring loaded holder (B), carousel wheel (C), operator implant control (D).

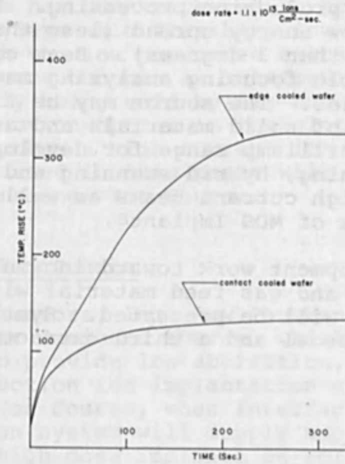


Figure 9 Wafer Heating Curves - Arsenic Implants at 110 uA scanned current and 170 keV.

These results show the difference in temperature rise when good and poor thermal contact is made to the wafer.

LOW PERVEANCE ION SOURCE BRIDGES LOW AND HIGH INTENSITIES IN ION IMPLANTATION

F. D. Orr
D. Mayhall

Accelerators, Inc.
P. O. Box 3293
Austin, Texas 78764

The Low Perveance Ion Source developed by Accelerators, Inc. offers the Semiconductor Industry the advantage of processing medium to high intensity implants on a system which will also implant 200 to 300 wafers an hour at MOS doses.

Stable source beam currents can be varied over three orders of magnitude by variation of a single source parameter. This source uses a new computer designed Low Perveance extraction optics which is completely new to the Ion Implantation Industry. Test data and calculations will be shown which define the versatility of this system. Scanned currents from 1 microamp to 400 microamps allow for a variety of production processing. Beam characteristics feature low energy spread (less than 10eV) and low divergence (less than 3 degrees). Beam control optics consist of a double focusing analyzing magnet and two triplet quadrupoles. The source may be fitted with an oven for feeding of solid materials and analyzed beam currents in the milliamp range for development purposes. The batch processing, hybrid scanning end station is most applicable for high current beams as well as high volume batch processings of MOS Implants.

Results of development work toward increased currents using both solid and gas feed material with the Low Perveance source will be presented. System improvements including Accel-Decel and a third extraction element will be discussed.

Introduction

The need for higher beam currents in ion implantation systems has been an ever increasing one since the technology was first used by device designers. In the search for higher current sources, Ion Implant system designers have usually been attracted by the ion sources used in isotope separators in the nuclear field. These sources are generally arc type sources or sources which dissipate extremely large amounts of power. The chief disadvantage (besides power dissipation) of these sources is plasma instability which causes large fluctuations in beam current both in the short term and the long term. Both of these draw-backs have complicated the design of higher current implanters therefore making them very costly.

In this paper we will discuss the evolution of another ion source which is now capable of being used in high current implanters. This source does not have the disadvantages of an arc source while exhibiting all of the advantages of a Penning or Oscillating Electron ion source.

The evolution of this source is as follows: 1) The original source has been marketed by Accelerators, Inc. (AI) for five years as a flexible, low current ion source; 2) Fitted with the low perveance extraction optics the low current source becomes an intermediate current ion source presently being sold by AI; and 3) Using the low perveance extraction optics in the slot configuration, this source becomes a high current ion source. The source initially discussed in this paper will be the intermediate current Low Perveance Ion Source.

Low Perveance Ion Source

The Accelerators, Inc. Low Perveance Ion Source is a special computer designed source, which was developed specifically to provide low aberration, high current ion beams for production ion implantation applications. The Low Perveance Ion Source, when interfaced with a suitable ion implantation system will supply ample beam currents for medium to high dose implants as well as being capable of low dose, high throughput applications. When operated in the Decel-Accel mode, it will provide full current over the complete voltage range of the implantation system (25 to 200 keV). For BF_3 feed gas, this current is 200 µA of analyzed B_{11}^+ ions on target. For PF_5 feed gas, this current is 300 µA of analyzed P_{31}^+ ions on target. For AsF_3 (or AsF_5) feed gas, this current is

400 μA of analyzed As_{75}^{+} ions on target. At the exit of the source, the beam has a very low energy spread of 10 eV or less and a low half beam divergence of 3^{o} or less. The source may be operated with solid as well as gaseous feed materials. Large quantities of solid feed material can be placed within the oven for sustained production implants.

The AI Model HC200 Ion Implanter, in which the Low Perveance Ion Source with 20 to 80 kV extraction optics is used, basically consists of a 120 keV accelerator tube, a double-focusing beam switching-analyzing magnet, two electrostatic triplet quadrupole lenses, beam defining slits, an electrostatic beam scanner, an electrostatic beam deflector, and a wafer processing end station.

The Model Low Perveance Ion Source is a hot filament, electron bombardment, Penning ion source fitted with a computer designed, low perveance, low aberration ion extraction system. A drawing of the source is shown in Fig. 1. The plasma generating portion is an improved version of the low current hot filament source marketed by Accelerators, Inc. for five years. The difference between the two sources is associated primarily with the extraction system. This is an important point to consider due to the large amount of existing data available on the ionization characteristics of this source with respect to numerous feed materials. The filament (A) is a helical filament of 0.010 inch diameter tungsten wire. The filament is surrounded by the cylindrical stainless steel anode (B). The front source cathode is formed by the back side of the source screen electrode (C). The ion source components are mounted on the source base plate (G), which carries electrical feedthroughs for the filament current (L) and for the anode voltage (M). The source base plate and the components on it form the ion source insert which is slipped into the stainless steel cylindrical source housing (N) and vacuum sealed by the O-ring (O). This feature allows for the filament to be removed easily and rapidly for convenient replacement. The source base plate may be equipped with two more electrical feed-throughs for the optional large solid feed material oven. A solenoid assembly (F) slips over the outside of the source housing, which is inside the stainless steel re-entrant tube (P). The solenoid furnishes the axial magnetic field for the source.

The ion source extraction optics consist of the screen electrode (C) and the extraction electrode (E). A positive extraction potential of 20 to 80 kV relative to the extraction electrode is placed on the screen electrode, which is at the high voltage terminal ground potential. These electrodes are insulated from each other by the cylindrical ceramic insulator (H). Aluminum potential grading rings (J) are mounted on the source which provide voltage grading across the ceramic insulator. The voltage divider resistors (K) divide the potential drop from the housing plate to the main mounting plate evenly between the potential grading rings.

The electrode (D) is an intermediate extraction electrode, which is used to alter the characteristics of the extracted ion beam. The intermediate electrode is normally operated at a potential of 20% or less of the extraction potential.

Ion Creation Process

The feed gas from the gas feed tube or the vaporized solid from the solid charge holder enters the source interior through the aperture in the source base plate. Current is passed through the tungsten filament, which causes it to emit electrons thermionically. The anode is biased at a positive voltage of 0-200 V with respect to the cathode. The thermionic electrons from the filament are accelerated toward the positive anode by the radial electric field and collide with the gas molecules. These collisions cause dissociation and ionization of the various dissociation products. The resulting ions and electrons constitute a plasma that is contained within the anode tube. The axial magnetic field from the solenoid retards the movement of the thermionic electrons toward the anode by significantly increasing their path length. This increased path length allows the thermionic electrons to have more collisions with the gas molecules and hence produce more ions before they reach the anode. The axial magnetic field also contains the ion-electron plasma within the center of the source. The screen electrode is made of a magnetic stainless steel so that the magnetic field is concentrated at the extraction aperture in the screen electrode. This concentrates the plasma ion density at the aperture so that a maximum number of ions are made available for extraction from the source plasma. Electrons which have axial velocities and attempt to leave the source discharge region by moving toward the back of the source cathode or the screen electrode, are reflected back into the plasma discharge region by the cathode's ground potential. Due to the magnetic field produced by the solenoid and this reflection process, electrons with predominantly axial velocities move back and forth axially or oscillate axially. This is called the reflex or Penning effect. It increases the number of collisions that the electrons have with the gas molecules before they are ultimately lost. The number of ions in the source is dependent upon the input gas flow rate, the filament current, the anode voltage, and the axial magnetic field strength. The shape and position of the plasma meniscus at the screen electrode aperture is influenced by the discharge parameters, as well as the value of the extraction potential. The axial magnet field strength is the parameter which is used to vary the current extracted from the source. The extracted current may be varied over three orders of magnitude while maintaining stable beam current. This allows for implants over an extremely wide dose range to be made with a minimum of set up time between implants.

Design of the Low Perveance Extraction System

Extraction systems for electron and ion sources are frequently characterized by a parameter called the perveance $\underline{P}$. The perveance is the proportionality constant between the total space charge limited current transmitted between the screen and extraction electrodes and the extraction voltage to the three halves power. For two parallel plane electrodes a distance $\underline{d}$ apart with a circular screen electrode aperture of radius $\underline{r}$, the extracted current $\underline{I}$ is given by

$$I = \frac{1.71 \times 10^{-7}}{M^{\frac{1}{2}}} \left(\frac{r}{d}\right)^2 V^{3/2} \quad \text{amp-volt}^{3/2}, \quad (1)$$

where $\underline{M}$ is the extracted particle mass number and $\underline{V}$ is the extraction voltage. The perveance $\underline{P}$ is then given by

$$P = \frac{1.71 \times 10^{-7}}{M^{\frac{1}{2}}} \left(\frac{r}{d}\right)^2 \quad \text{amp-volt}^{-3/2}. \quad (2)$$

If $\underline{M}$ is taken as the electron mass number $\frac{1}{1840}$, the perveance is called the electron perveance P_{e^-} and is given by

$$P_{e^-} = 7.34 \times 10^{-6} \left(\frac{r}{d}\right)^2 \quad \text{amp-volt}^{-3/2}. \quad (3)$$

The electron perveance of an extraction system depends, to the first order, only on the ratio of the screen aperture radius $\underline{r}$ to the electrode separation $\underline{d}$. The electron perveance may be reduced by decreasing the ratio $\frac{r}{d}$, which may be done by decreasing the aperture radius $\underline{r}$ or by increasing the electrode separation $\underline{d}$. High perveance extraction systems may be considered to be those with $P_e - > 10^{-7}$ amp-volt$^{-3/2}$, which corresponds to $\frac{r}{d} > 0.369$.

Space charge computer analysis of high perveance systems, in which $2\underline{r}$ is approximately equal to $\underline{d}$ and the maximum extraction voltage is consequently low (due to voltage break down characteristics), has shown that they are characterized by a large number of trajectory cross-overs or non-laminar flow as shown in Fig. 2. Such non-laminar flow results in a strongly diverging beam a short distance downstream from the extractor electrode. Such a beam is highly unsuitable for an ion implanter, which requires transmission distances of the order of 10 to 20 feet. A great deal of current will be lost in transmission or the difficult and expensive problem of beam refocusing must be solved which adds to the system complexity and cost.

Fig. 2 shows, that for the high perveance case, the equal-potential lines between the screen and extraction electrode are very much distorted in the region of the extraction apertures. This is due to the large ratio of $\frac{r}{d} \approx 0.50$. It is easily seen that with the electric field lines between the electrodes of this system ($\frac{r}{d} \approx 0.50$) will be much more perturbated than in a system with a lower $\frac{r}{d}$ ratio. In the high perveance case, the plasma meniscus is very much concaved. It is well known that the extracted ion leaves the meniscus of a plasma perpendicular to the plasma surface, causing a very non-laminar flow of the beam as it is extracted. In the low perveance case where $\frac{r}{d}$ is low (approximately .1), the discontinuity of the electric field generating plates has a much smaller effect on the field lines distortion. This results in a more planar plasma meniscus and a well formed beam due to better laminar flow of the ions of the beam.

Space charge analysis of low perveance, high voltage ion extraction systems ($P_e - = 1\text{-}4 \times 10^{-8}$ amp-volt$^{-3/2}$), in which $\frac{r}{d} \approx 0.125$, has shown that these systems produce highly laminar, narrow angle ion beams. Such high quality beams may be easily transported over the distances encountered in ion implanters.

The Low Perveance Ion Source was designed by a space charge compensating computer program in which an optimum initial plasma sheath shape and the electrode boundaries were assumed as input data. Ion trajectories, current density profile, equipotentials, and the perveance = I $V^{-3/2}$ were then calculated in a self-consistent manner assuming space charge limited flow at the emitting surface. A planar emitting surface was assumed. The screen and extraction electrodes were specially shaped to create a potential distribution which closely approaches along the beam edge the theoretical Langmuir potential for a spherical diode. The calculated current density variation across the emitting surface was less than 5%. The resulting beam is quite laminar and has a divergence half angle of 3°.

The Low Perveance Ion Source is designed to theoretically produce a total beam current of 4.0 mA at 100 kV extraction voltage. In actual practice, this extracted current is reached at approximately 80 keV.

Ion Source Current Characteristics

The measured current characteristics of the Low Perveance Ion Source are shown in Fig. 3. The total beam current extracted from the source is plotted versus the extraction voltage. The feed gas was BF_3. The source is seen to furnish 4.0 mA total beam current at 85 kV. The other data points fall very close to the predicted design perveance line indicated on the figure.

The measured current characteristics of the Low Perveance Ion Source, when operated on an AI Model HC200MP Ion Implanter with an X-Y scanner using BF_3 feed gas, is shown in Fig. 4. For these measurements, no high voltage was applied to the ion implanter accelerator tube. Three currents are plotted versus the source extraction voltage. The top curve in the figure is the total beam current before mass analysis. At 80 kV extraction voltage, 3.25mA of ion current was obtained. The middle curve is the analyzed, unscanned B_{11}^+ current at the wafer processing end station. At 80 kV extraction voltage, 200 µA of B_{11}^+ was obtained. The bottom curve is the X-Y scanned B_{11}^+ current at the end station.

Figure 5 shows the measured current characteristics of the Low Perveance Ion Source on the same ion implanter using AsF_3 gas feed. No high voltage was applied to the implanter accelerator tube. The top curve is the analyzed, unscanned As_{75}^+ current at the wafer processing end station. 450 µA of As_{75}^+ was obtained at 90 kV extraction voltage. The middle curve is the total beam current at the magnet chamber zero degree port without mass analysis.

At 90 kV extraction voltage, 4 mA of total beam current was obtained. The bottom curve is the analyzed, X-Y scanned As_{75}^+ current at the implanter end station.

Low Energy Beam Currents in Decel-Accel Mode

Several electrical arrangements of the accelerator tube electrodes have been investigated in order to extract high current, low energy ion beams from the Low Perveance Ion Source. They essentially consist of extracting the ion beam from the source at 50 to 80 kV of extraction voltage, then decelerated through an equal voltage with part of the accelerator tube followed by acceleration to the desired beam energy with the remaining portion of the accelerator tube. The beam is thus extracted from the source at its rated current for a specific extraction voltage then decelerated to very low energy, then accelerated to the desired energy with little loss of its original extracted current. The results for the beam energy range from 25-40 keV, with BF_3 gas feed, are shown in Fig. 6 as curve B. The total beam current measured at the magnet chamber zero degree port is plotted versus the ion beam energy. The accelerator tube arrangement is shown by the small schematic diagram labelled "Arrangement B".

The results for the beam energy range from 10-20 keV are shown in Fig. 6 as curve A. The accelerator tube arrangement for these results is shown by the schematic diagram labelled "Arrangement A". The lower curve on Fig. 6 is the predicted total beam current for the ion source without the decel-accel arrangement, that is, using only the source extraction gap. It is thus seen that the Decel-Accel arrangements greatly increase the beam current at the lower beam energies.

It is important to note that the process of switching from standard configuration to the Decel-Accel mode requires fifteen to twenty minutes time including tune up time. This change, however, allows total output currents to be used over the entire energy range of the system.

Ion Beam Characteristics for the Intermediate Extraction Electrode

The intermediate electrode assembly (D) is shown in Fig. 1. The intermediate electrode is shaped to fit an equipotential surface external to the beam edge at a given axial distance from the assumed position of the plasma emission surface. The intermediate electrode may be operated so as to have two effects on the extracted ion beam.

1) If the ion source is operated on the current-voltage curve given by the perveance relation for the source, which is similar to Eq. 1, the intermediate electrode has an additional focusing effect on the beam. This focusing of the beam essentially reduces the beam divergence half angle to less than one degree. This focusing effect is shown by Fig. 7, which is with BF_3 and AsF_3 gas feeds. The experimentally measured beam divergence half angle is plotted versus the intermediate electrode voltage with the extraction voltage as a constant parameter. The nominal beam divergence half angle for operation of the ion source without the intermediate electrode is the horizontal line at 3°. For each value of extraction voltage, the beam divergence half angle decreases with decreasing intermediate electrode voltage. For those curves with a sufficient number of data points, a minimum value of beam divergence half angle is obtained and then the beam divergence increases with decreasing intermediate electrode voltage.

2) For a given extraction voltage, as the intermediate electrode voltage increases, the total beam current increases. The total beam current is shown by the scale at the top of Fig. 7. Fig. 7 shows that the intermediate electrode decouples the magnitude of the beam current from the extraction voltage-beam current relationship given by the perveance equation. For a given extraction voltage, an extended range of beam currents are now available, instead of just a narrow range about the beam current value predicted by the perveance relation. Of course, it must be noted that, for a given extraction voltage, the beam divergence varies with the beam current. Within limits, the use of the intermediate electrode thus allows the beam current to be varied for a fixed extraction voltage or the extraction voltage to be varied for a fixed beam current.

Here again the effect of this element is to flatten the current versus voltage characteristic of the implantation system. While this focusing element does not quite achieve the effect that the Decel-Accel mode does, it is somewhat more easily operated.

High Current Slot Extraction Sources

The need for B_{11}^{+}, P_{31}^{+}, and As_{75}^{+} beam currents of one to three milliamp for production batch processing of high dose implants is a certainty. In order to use gas feed materials, a total beam current of 16 to 20 mA must be extracted from the ion source. The existing ion sources capable of this level of extraction are high power dissipation isotope separator sources. The high power dissipation (5 to 15 kVA) very much complicates the design for very high current implantation systems.

The initial calculation on the plasma capability of the Low Perveance Ion Source shows that with little modification, it will support a rich enough plasma to extract a 20 mA beam. This is accomplished with less than 1500 watts being dissipated by the source. With all other parameters being equal, the extracted beam current is proportional to the area of the plasma meniscus formed at the exit aperture of the screen electrode. The optimum method of enlarging this area is to use slot apertures rather than larger diameter apertures. There are two reasons for this: 1) The obvious control of space charge in a ribbon shaped beam is much easier than in a cylindrical beam; 2) The slot apertures represent a much lower vacuum conductance path than does a larger diameter hole of the same area. This provides better vacuum isolation of the plasma generating portion of the source from the rest of the implant system.

The first test were performed on a standard Low Perveance Ion Source with modified screen and extraction electrodes. This source shown in Fig. 8 is an axial extraction source because the beam is extracted along the axis of the magnetic field. The area of the slots were four times that of the standard ion source and the aspect ratio was approximately five. In all other respects the geometrics of the screen and extraction electrodes were identical to the standard system.

The tests on this source showed that it performed identical to the standard source and the extraction current was approximately three times that of the standard source for a given extraction voltage. In measuring the current distribution along the long axis of the ribbon beam, it was found that the distribution was somewhat gaussian. This was believed to be caused by the distribution of the plasma within the source itself because of its spherical shape.

Because of the above assumptions, calculations, and tests, another version of the source was built. In this source the plasma generation area has been modified to present a more uniform plasma density to the slot aperture. This source is presently under test at AI and is capable of extracted beam currents of 20 mA. Although the tests have not been completed, it is felt that this source, with less than 1500 watts of power dissipation, will be the next truly new source for high current Ion Implantation systems.

Conclusions

Accelerators, Inc. has followed a logical step by step approach to developing a source to supply the needs of today's and tomorrow's ion implanters. The original hot filament source fitted with high perveance extraction optics was capable of less than one milliampere of extracted current. The Low Perveance Ion Source with computer designed extraction optics increased the extracted current to four and one half milliamperes. Now, this same basic source with slot, low perveance extraction optics and modified plasma generator is capable of supplying currents of approximately twenty milliamperes. This has been accomplished with only a factor of three increase in the power dissipation of the original source. The real advantage of this source is not only its low power dissipation, and small size, but the fact that all of the existing ionization data gathered for over five years is applicable to this new source.

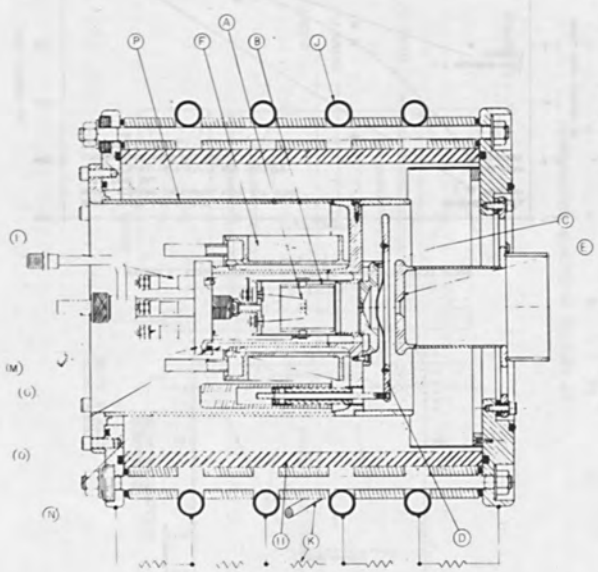


FIGURE I LOW PERVEANCE ION SOURCE ASSEMBLY

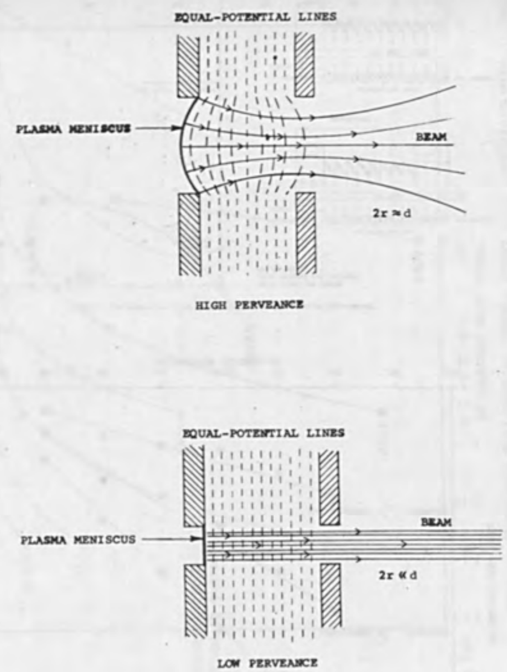


FIGURE 2 HIGH AND LOW PERVEANCE SYSTEMS

FIGURE 3 CURRENT-VOLTAGE CHARACTERISTICS OF MODEL HC200 ION SOURCE WITH BF_3 GAS FEED

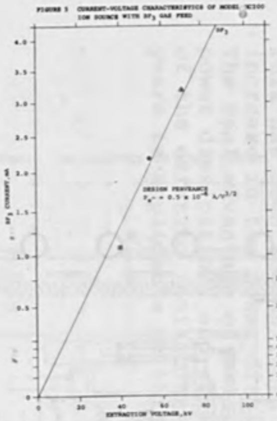


FIGURE 4 CURRENT CHARACTERISTICS OF PRODUCTION MODEL HC200 ION SOURCE WITH BF_3 GAS FEED

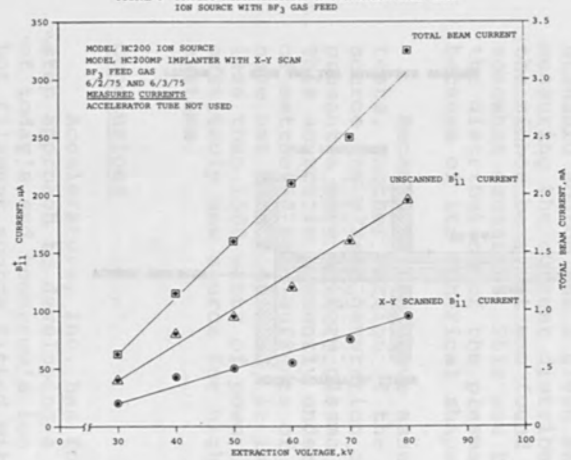


FIGURE 5 CURRENT CHARACTERISTICS OF PRODUCTION MODEL HC200 ION SOURCE WITH AsF_3 GAS FEED

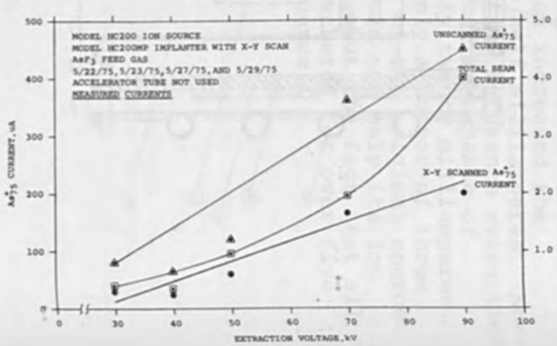


FIGURE 6 LOW ENERGY TOTAL BEAM CURRENT FOR DECEL-ACCEL EXTRACTION ARRANGEMENTS

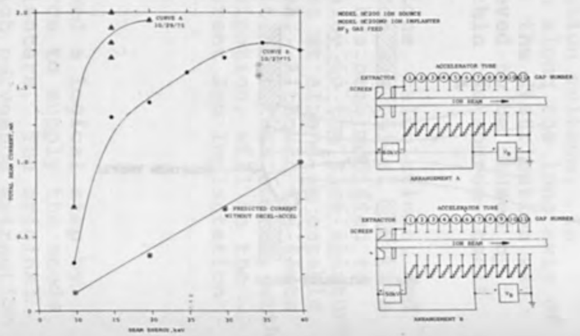
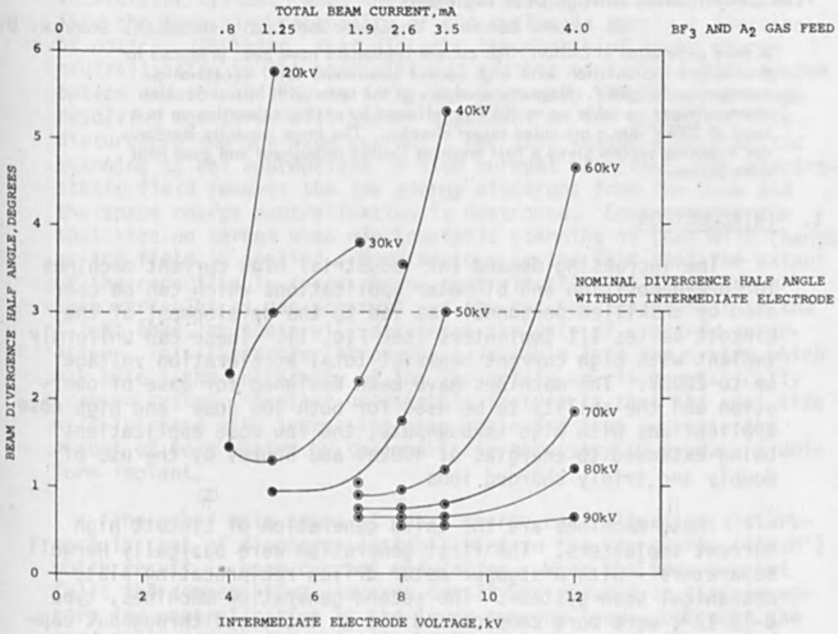


FIGURE 7 BEAM DIVERGENCE HALF ANGLE VERSUS INTERMEDIATE ELECTRODE VOLTAGE



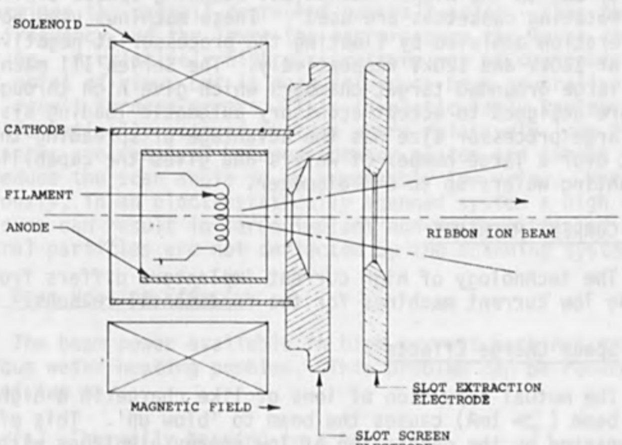


FIGURE 8 SLOT EXTRACTION SOURCE

The Lintott Series III High Dose Implanters

D. Aitken, Lintott Engineering Ltd., Horsham, Sussex, U.K.

A third generation of Lintott High current implanters have been produced for the uniform implantation with high current beams and total acceleration voltages up to 200kV. Magnetic analysis of the beam with initial acceleration voltages variable up to 40kV is followed by post-acceleration up to a total of 200kV into a grounded target chamber. The large capacity mechanical scanning system gives a fast machine limited throughput and good heat distribution.

1. INTRODUCTION

The increasing demand for industrial high current machines for pre-deposition and bi-polar applications which can be operated by unskilled personnel has led to the development of the Lintott Series III Implanters (see Fig. 1). These can uniformly implant with high current beams at total acceleration voltages up to 200kV. The machines have been designed for ease of operation and the ability to be used for both low dose and high dose applications with high throughputs, the low dose applications being extended to energies of 400keV and 600keV by the use of doubly and triply charged ions.

These machines are the third generation of Lintott high current implanters. The first generation were basically Harwell separators [1] with a stepper motor driven reciprocating plate mechanical scan system[2]. The second generation machines, type 8-12-12[3], were more compact and had an increased throughput capability due to the higher mechanical scanning speeds available when rotating cassettes are used[4]. These machines used post-acceleration achieved by floating the processor at negative voltages of 160kV and 120kV respectively. The Series III machines have large grounded target chambers which give high throughputs and are designed to accept accessory automatic loading systems. The large processor size has the advantage of spreading the heat input over a large number of wafers and gives the capability of implanting wafers up to 6½" diameter.

2. HIGH CURRENT MACHINES

The technology of high current implanters differs from that of the low current machines for two fundamental reasons.

i) Space Charge Effects

The mutual repulsion of ions of like charge in a high current beam ($>$ 1mA) causes the beam to 'blow up'. This effect is compensated by the generation of low energy electrons within the beam due to collisions with neutral gas atoms or due to collisions with apertures and other solid components in the accelerator system. When the beam is fully space charge neutralised the focussing capability of the system is purely a function of optics. Anything that disturbs the degree of space charge neutralisation in the beam will add another variable to the system optics and the ability to focus the beam onto targets or through resolving apertures will become dependent on the extent of this disturbance and the beam current. For this reason electrostatic scanning is not appropriate in high current machines. The electrostatic field removes the low energy electrons from the beam and the space charge neutralisation is destroyed. Consequently the spot size on target when electrostatic scanning is used will change as the field is applied. More serious is the fact that the extent of the reduction in space charge neutralisation will depend on such variables as gas pressure and the gas species present in the flight tube, as these will determine the rate of electron generation. This therefore introduces a variable into the system which may effect the uniformity of the implant in an electrostatically scanned system. Another undesirable feature is that the spot size will increase with increasing scan angle and thus increase the amount of beam wasted by the over-scan necessary to produce a uniform implant.

The other main cause of space charge neutralisation disturbance is that of discharge instabilities in the ion source ('hash') which result in beam current modulation. When the beam current falls the compensating electron density is reduced to that necessary for neutralisation at the lower current level, but when the current increases the rate at which electrons can be generated determines the overall degree of neutralisation. Thus the higher the frequency and the lower the gas pressure the worse the effect will be. A relatively high gas pressure may be desirable from this point of view, but it does introduce beam neutralisation which will result in dose errors. This is particularly serious in electrostatically scanned systems where relatively long flight distances are required between scanner plates and target in order to reduce the scan angle to an acceptably low value. More seriously, in an electrostatically scanned system a high system pressure can result in large implant non-uniformities as the neutral particles are not deflected by the scanning system.

ii) Beam Heating Effects

The beam power available in high current machines creates a serious wafer heating problem. This problem can be reduced in one of two ways:-

(a) Thermally bonding to a heat sink

Unfortunately, no suitable heat transfer material appears to be available that is acceptable to the semiconductor manufacturing process, and there are likely to be problems associated with the application and removal of the bonding material together with the possibility of contamination of the silicon if this material is not completely removed. Even if a suitable material were available there are likely to be doping rate effects if the full beam current is scanned across a single wafer.

(b) Batch implantation

This is the technique that has generally been adopted in high current machines. The heat load is spread over a number of wafers thus reducing the average rate of heat dissipation on any individual wafer. As one must assume that the thermal contact between wafers and the plates on which they are mounted is poor, in the absence of any suitable bonding material, then radiation is the only heat loss mechanism that can be reliably assumed to exist. Therefore the importance of having a low rate of heat input combined with the need for high throughputs and therefore the use of high current beams all point to the ideal requirement of very large batch sizes. A compromise thus has to be reached between this requirement and the need for a reasonably steady flow throughput which is more convenient in semiconductor production.

A detailed analysis of the ion beam heating situation will not be given here as it has been considered in some detail elsewhere[5], but later in this paper the results will be applied to typical implant situations in the Lintott Series III Implanter.

3. MECHANICAL SCANNING TECHNIQUES

Having concluded that for both ion optical and heat dissipation reasons that mechanical scanning is the most appropriate scanning system for high current machines, it is now relevant to consider the types of mechanical scanning systems available.

The first generation of Lintott machines based on the Harwell separator had a reciprocating plate scanning system[2] which severely limited the throughput available due to the limitation to the speed with which the fast scan could be executed both for mechanical and wafer stessing reasons. In consequence any high throughput system is likely to employ rotation rather than reciprocation for the fast scan component of the scan raster. Four such systems are shown in Fig. 2. The first is a simple circular carousel rotating about a central axis. This system has the advantage of simplicity but suffers from two deviations from the ideal scan required. The angle of the wafer to the beam changes as the wafer passes through the beam and there is also a cosine error introduced in the scan velocity across the beam. This error is

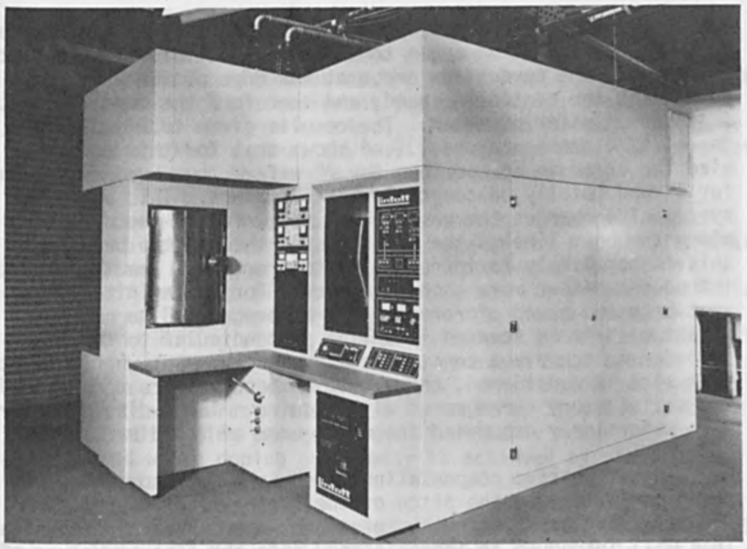
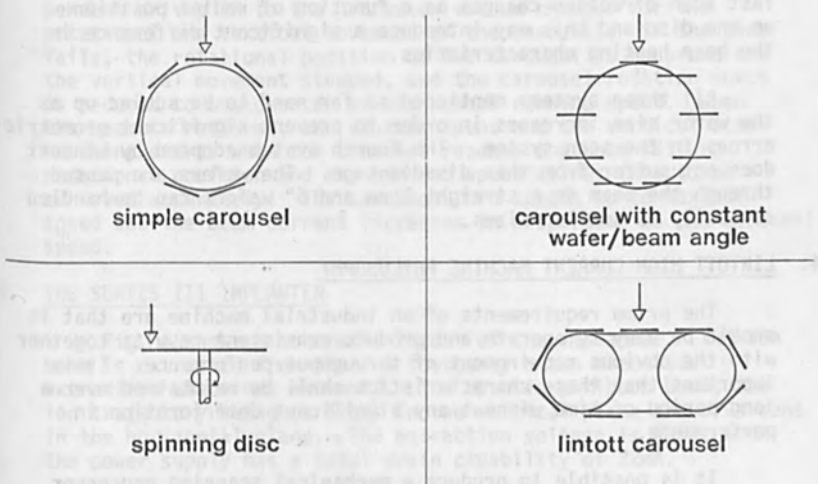


FIG. I



mechanical scanning systems

Fig. 2

more evident in the second system (Ferris wheel) where the wafers are kept at a constant angle to the beam, and this is illustrated in Fig. 3. The percentage error at the edge of the wafer compared with the centre in speed, and therefore the dose, is a strong function of wafer diameter. The results given in Fig. 3 are for a 24" diameter carousel, and shows that for this carousel size the error is not serious for 2" wafers, but becomes serious for 4" and totally unacceptable for 6" wafers. This scanning system also changes the position of the wafer in the direction along the beam line as the wafer passes through the beam, but this is not likely to introduce serious errors. The third system introduces rather more complex errors. For an infinitely small spot beam the speed of rotation can be programmed to change as the rotation shaft is scanned in a line perpendicular to the beam in order to achieve a constant scan velocity. But when a finite beam size is considered, the situation becomes more complex and the satisfactory arrangement will require a large disc diameter with wafers only occupying the outer edge of the disc.

An alternative compensation which is easier to achieve in practice is to vary the pitch of the raster, i.e. to programme the slow scan speed with the rotational speed remaining constant[6]. This will introduce an inconsistency into the fine structure of the uniformity pattern which would probably not be detectable in many uniformity measuring techniques. More serious is the fact that the rate of traversal of the beam across the wafer in the fast scan direction changes as a function of radial position on the disc and this may introduce a significant difference in the beam heating characteristics.

All three systems mentioned so far need to be scaled up as the wafer size increases in order to prevent significant geometric errors in the scan system. The fourth system adopted by Lintott does not suffer from this disadvantage. The wafers are passed through the beam in a straight line and 6" wafers can be handled as easily as smaller sizes.

4. LINTOTT HIGH CURRENT MACHINE PHILOSOPHY

The prime requirements of an industrial machine are that it should be easy to operate and produce consistent results together with the obvious requirement of throughput performance. It is important that these characteristics shall be maintained over a long period of time without any significant deterioration in performance.

It is possible to produce a mechanical scanning processor that runs extremely smoothly, but it is better if this is achieved with a control system that does not assume that it has been achieved. In this case under normal running conditions the processor carousel rotates at the nominal maximum speed of 80 r.p.m. The actual rotation speed achieved is measured and the processor control unit measures the ratio beam current to carousel rotation speed. The requirement for good doping uniformity is that this ratio shall be constant. Therefore the beam current is also controlled so that any deviations from the ideal mechanical characteristics of the processor can be compensated.

The control of beam current has another important aspect relevant to doping consistency. Because of the possibility that the resultant doping effect in the silicon wafer is not only a function of dose but also the dose rate and its related variable temperature, it is therefore important for the maintenance of a consistent doping effect to always implant under identical conditions. This inevitably means a constant beam current and therefore a constant implantation time. This rules out systems where the beam current is allowed to drift as dictated by the various conditioning processes that inevitably occur in ion sources and where doping uniformity is achieved by changing the scanning rate.

Although a well designed implanter should not give unexpected beam failures in the middle of an implant, this should not be assumed to be the case. Therefore the system must be able to continue after a beam failure with the minimum discontinuity possible to the scan raster. This is achieved by having independent drives in to the processor vacuum chamber for the rotational and vertical movement of the carousel. If the beam fails, the rotational position of the carousel is recorded and the vertical movement stopped, and the carousel rotation speed is reduced to one quarter of the normal running speed. When the beam is restored, the shutter opens and the vertical scan movement starts when the carousel reaches the beam failure rotational position and the reduced speed allows this to be achieved accurately. The carousel then speeds up to maximum speed and the beam current increases in proportion to the carousel speed.

5. THE SERIES III IMPLANTER

Fig. 4 shows a schematic layout of the implanter. The ion beam is extracted from the ion source by a two electrode extraction electrode assembly. The beam is a ribbon beam, the ion trajectory being parallel in the vertical plane and divergent in the horizontal plane. The extraction voltage is 40kV, and the power supply has a total drain capability of 32mA.

The ion source is similar to that used in the Harwell separator [7] but extra screening has been added to protect the filament insulators from metallisation.

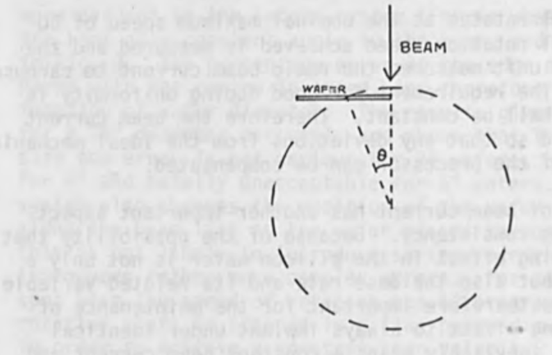
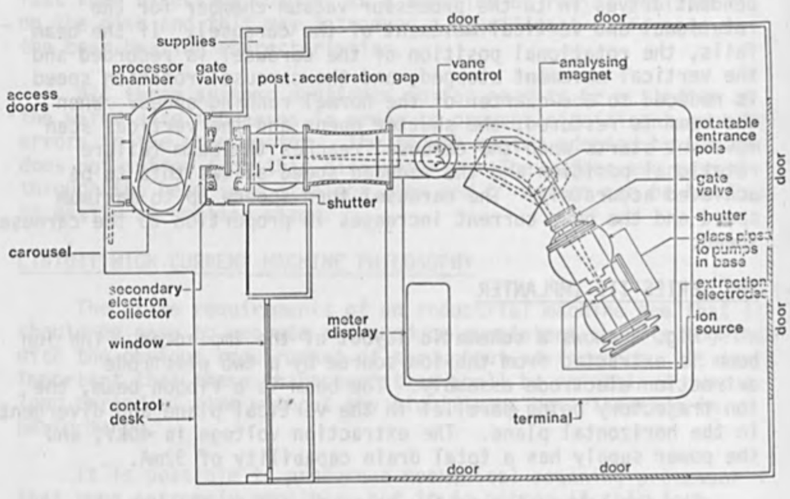


Fig. 3 FERRIS WHEEL ERRORS

For 24" diameter carousel :-

Wafer diameter (ins)	Θ	% error
2	4.78	0.35
3	7.18	0.78
4	9.59	1.40
6	14.48	3.18



schematic layout of implanter

Fig. 4

The arc chamber is a graphite block and there are three feed tubes into the block. One carries the gas feed which is usually argon and/or boron trichloride and the other two are furnace supplies for the evaporation of solids. These normally contain red phosphorus and arsenic.

The divergent beam from the ion source enters the analysing magnet, which has a rotatable entrance pole. This enables the beam to be focussed into the resolving slit without the use of any electrostatic focussing lenses. The magnet is fed from a 50V, 100A supply and is capable of bending masses up to 130 at the full extraction voltage of 40kV and larger masses at decreased extraction voltages.

Immediately after the beam has passed through the analysis flight tube it then comes to the vane unit which controls the beam current reaching the processor. This position was chosen for the vane unit for several reasons. By placing it after the analysing magnet it only sees the analysed beam, which for example in the case of boron is less than 10% of the total beam leaving the ion source. This reduces the erosion of the graphite vanes to insignificant levels. This position in the flight path is also advantageous because the beam width is large thus reducing the current density falling on the vanes and giving fine control of the transmitted current. Another possible position for the vane unit is after the post-acceleration gap, but this would have involved controlling the beam at a point where the energy and current density were high giving higher erosion rates at a point sufficiently close to the wafers that there would be the danger of carbon contamination on the wafers. For this reason the machine is designed so that the beam does not strike any surface between the resolving slit and the carousel in order to prevent sputtered contamination reaching the wafers.

A mechanical control system has been adopted primarily because it does not interfere with the optics of the system and is entirely predictable in its behaviour. Electrostatic or magnetic techniques for reducing the fraction of the beam transmitted through the resolving slit either involve defocussing which effects the resolution and will not have such predictable characteristics or modulation of beam direction prior to resolution which inevitably gives an A.C. component to the beam reaching the processor and this could cause doping non-uniformities if the carousel rotation time were an integral number of modulation cycles. The other possibility is direct ion source control, but again there are optical problems. For example it is often possible to increase the current through a high current machine by decreasing the arc current in the ion source. The reduction in an arc current may reduce the source output but it can also change the ion optics situation due to changes in the plasma meniscus geometry.

For the reasons outlined above a mechanical beam control system has been adopted. The vane unit is driven by a high speed stepper motor and can change the current by approximately 0.1% per step, the step time being one millisecond. The vane unit together with the analysing magnet and the ion source are in the accelerator terminal (see Figs. 5 and 6) which can float at up to 160kV to give the post-acceleration of the beam. Communication between the terminal and the control system at ground potential is achieved by the use of fibre optics.

The post-acceleration is achieved across a single gap as with previous Lintott machines. The novel feature in the Series III machines is that the acceleration gap is at the centre of a stainless steel chamber and this enables the X-ray emission associated with the post-acceleration up to 160kV to be shielded close to the origin of the X-rays. This arrangement also allows the vacuum pumps to be placed close to the post-acceleration gap. Immediately after the post-acceleration gap is a magnetically supressed shutter for beam current measurement prior to entry into the processor chamber.

The processor consists of a large stainless steel chamber with two large access doors. The access door close to the control panel (see Fig. 7) is for manual loading of wafer plates into the processor. The other access door can also be used for manual loading but is primarily available for an automatic loading accessory. The wafers are mounted on nine plates each of which can take 12 2" wafers, 6 3" wafers, 3 4" wafers or 1 6" wafer. Thus the maximum load of wafers is 108, 54, 27 or 9 respectively.

The large implantation area of this system means that even in the extreme case of heat loss by radiation only [5] and a beam of 4mA at 200keV then the maximum temperature rise for an infinite dose is still less than 200°C. In reality there is some thermal contact between the wafers and the plates and this together with the large thermal capacity of the carrier plates will limit the temperature to a substantially lower level.

The Series III machine vacuum system consists of four stages of differential pumping with the option of a five stage system for special applications. Each pump system consists of an Edwards DS100 diffusion pump and a liquid nitrogen cryopump/trap. The function of the diffusion pumps is primarily to maintain a low air partial pressure in the system and the cryopumps are the primary pumps of the system when used for semiconductor applications. All the source feeds used for the common dopant elements are cryopumpable at liquid nitrogen temperatures (e.g. BCl_3 has a partial pressure of $\simeq 10^{-10}$ torr at - 196°C) and consequently

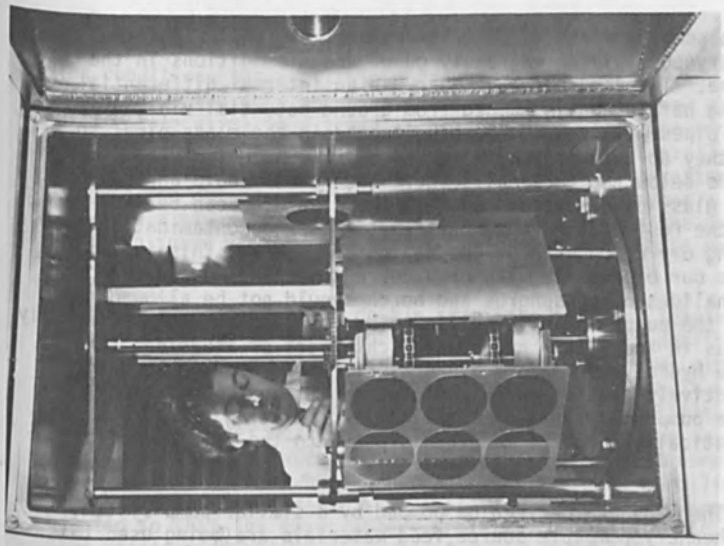


FIG.7

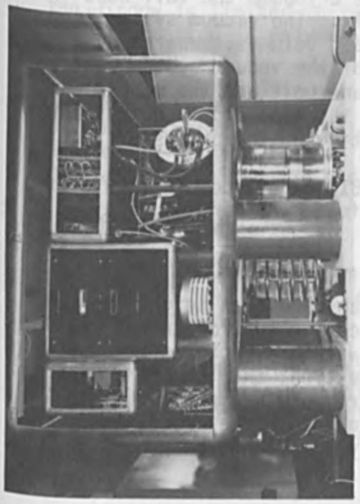


FIG.5

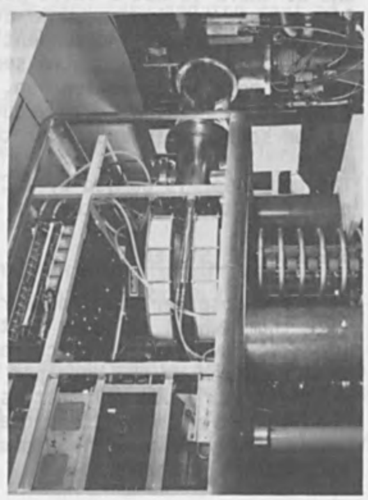


FIG.6

a high pumping speed of 6000 litres/sec. is available from each cryopump giving extremely good vacuum conditions in the machine. The ion source chamber has an internal differential pumping baffle and is pumped from ground potential via two 12" glass pipes with the cryopumps situated in the glass pipes so that they do not contribute to the overall height of the pump systems below the terminal. There are valves above and below these glass pipe sections so that the cryopumps can be isolated from the rest of the vacuum system to allow decontamination by blowing dry nitrogen over the heated cryopumps. This is important as our experience with high current machines has shown us that halides of phosphorus and boron should not be allowed to reach the pumps in any significant quantity. This is particularly serious in the rotary pumps where chlorides or flourides tend to become hydrolysed to form hydrochloric and hydrofluoric acids respectively causing corrosion of the pump and degradation of the pump oil. This decontamination operation is carried out automatically when the machine is put into the 'standby off' mode.

The vane chamber can be pumped by a similar pump system, but when cryopumpable source feed materials are being used this is totally unnecessary. The extra stage of differential pumping may be useful for machines used for research purposes, but for semiconductor production applications it is totally unnecessary.

The post-acceleration chamber has a similar pump system (without the 12" glass pipe as this chamber is at ground potential) and as there is not a valve above the cryopump the full 6000 litres/sec. pumping speed is available. The vacuum system is divided into three independently pumped volumes, separated by gate valves, and the centre section of the vacuum system containing the post-acceleration gap, the vane unit and the magnet flight tube remains under vacuum all the time and is not disturbed by ion source maintenance or the venting of the processor.

The processor is pumped by a CVC - PB1000 pump combination giving pump down to 5 x 10^{-6} torr in approximately 5 minutes.

6. THE CONTROL SYSTEM

The dose per scan, number of scans and scan height are set with thumb-wheel switches on the processor control unit. Provided the available beam current is within a factor of two of that required by the dose per scan setting, then with the machine in the 'Auto' mode all that is required of the operator is to load the wafers into the processor and close the processor door. The machine will automatically pump down the processor and implant the wafers.

The control logic is shown in Fig. 8. Under normal operating conditions the available current will be greater than that required by the dose per scan setting and the processor carousel will rotate at maximum speed. The vane unit will then reduce this current to the required level. If the current available is not sufficient then the vane control will open to attempt to produce the demand current. When the vanes open beyond a certain aperture a signal is sent back to the processor and the carousel rotation speed is reduced so that the demand current is reduced to meet that available.

A scan consists of a double scan, the carousel leaving the low position in the processor and returning to this position. At the end of the first half of the scan, the carousel direction is reversed and the vertical drive disengaged for a half revolution. The second part of the scan raster therefore fills in between the raster pattern for the first half scan due to the 180^{o} phase change. When the carousel starts the next complete scan there is 0.2 rev phase change. This produces a rolling raster effect for multiple scan implants. Carousel reversal is adopted so that the scan rasters are always parallel. If the same rotation direction were used for both half scans then a 'Moiré' fringe effect would limit the resultant uniformity.

The control face of the machine (Fig. 9) consists of a tinted window for observation of terminal parameters, an oscilloscope for beam focussing, vacuum controls, implant parameter controls and terminal controls. The information visible through the window is limited to the minimum necessary and the rest of the information is only displayed when the appropriate controls are operated.

The terminal controls include three automatically selected ion masses normally boron, phosphorus and arsenic, but any species up to mass 130 can be selected. The total acceleration voltage is displayed on a digital meter.

7. PERFORMANCE

Table 1 gives the beam current specification of the machine and Table 2 the throughput capability as a function of dose and scanning conditions.

The uniformity and beam heating measurements are in progress and the results will be available in the near future. The limited uniformity information available at present suggests that the provisional specification of 1.0% can be significantly improved upon.

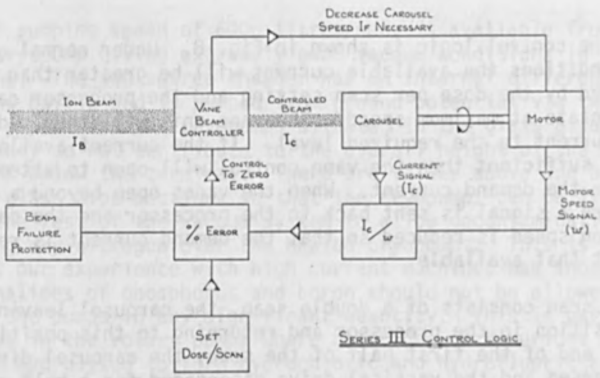


Fig. 8

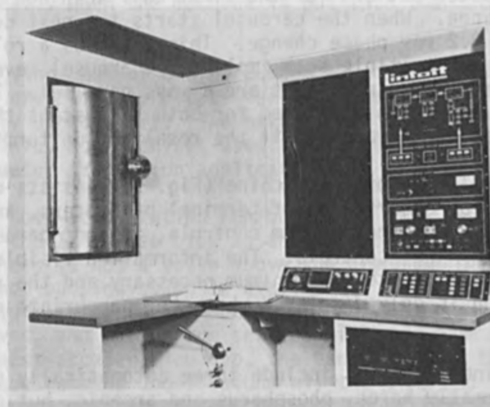


FIG.9

TABLE I — SERIES III IMPLANTER PERFORMANCE

ENERGY (keV)	ION	FEED	CURRENT	PARTICLE CURRENT	MAX. DOSE FOR MAX. THROUGHPUT ($ions/cm^2$)
200	B^+	BCl_3	1mA	1mA	4×10^{14}
"	P^+	red P	4mA	4mA	1.6×10^{15}
"	As^+	As	4mA	4mA	1.6×10^{15}
400	B^{++}	BCl_3	6μA	3μA	1.2×10^{12}
"	P^{++}	red P	400μA	200μA	8×10^{13}
"	As^{++}	As	400μA	200μA	8×10^{13}
600	P^{+++}	red P	4.5μA	1.5μA	6×10^{11}
"	As^{+++}	As	4.5μA	1.5μA	6×10^{11}

ACKNOWLEDGEMENTS

The author would like to acknowledge the contribution to the design of this machine from E. Bamsey, P. Lush, J. Glover and K. Knox the draughtsmen involved with the project, and C. Wright, the electronic engineer responsible for the electronic design.

REFERENCES

1. J. H. Freeman, Proc. Int. Mass Spec. Conf. (Kyoto, Japan)
   also A.E.R.E. Report R6254
2. J. H. Freeman, L. R. Caldecourt, K.C.W. Done and R. J. Francis A.E.R.E. Report R6496
3. Lintott Specification Sheets 1038B and 1248A.
4. Lintott Patent No. 1389294.
5. G. Dearnaley, J.H. Freeman, R.S. Nelson and J. Stephen. Ion Implantation p. 423. Defects in Crystalline Solids Series, North Holland, 1973.
6. G.I. Robertson, J. Electrochem. Soc., Solid State Science and Technology, June 1975, p. 796.
7. J. H. Freeman, A.E.R.E. Report R6138.

TABLE 2 WAFER THROUGHPUT FOR SERIES III MACHINE

Throughput in 3" wafers/hr (for 2" wafers multiply by 2) (for 4" wafers divide by 2)

DOSE ions/cm^2	BEAM CURRENT mA	SINGLE SCAN	DOUBLE SCAN	MULTIPLE SCANS
$10^{10} - 4 \times 10^{14}$	1.0	200	150	
$4 \times 10^{14} - 8 \times 10^{14}$	"		150	
2×10^{15}	"			85
10^{16}	"			22
$10^{10} - 1.6 \times 10^{15}$	4.0	200	150	
$1.6 \times 10^{15} - 3.2 \times 10^{15}$	"		150	
10^{16}	"			72
10^{17}	"			9.3

VARIAN/EXTRION HIGH DOSE IMPLANTERS

G. Ryding, A.B. Wittkower, P.H. Rose

Varian/Extrion Division
Blackburn Industrial Park, Gloucester, Mass. 01930

ABSTRACT

Two new implanters are described. For high doses using high wattage beams, milliampere currents are used in conjunction with a fully mechanical end station. Wafer heating is minimized by distributing the high beam power over a large effective area corresponding to 26 three or four inch diameter wafers per batch. At intermediate doses, it is advantageous for reasons of increased throughput to use up to 400μA of scanned beam, implanting one wafer at a time.

1. INTRODUCTION

This paper describes the philosophy behind the design and construction of two recent Varian/Extrion high dose ion implantation systems. One of these systems (Model 200-20AF) was designed for use at intermediate power levels; this machine implants each wafer independently and is fed from standard commercial linear wafer carriers or directly from a conveyor belt. The other system (Model 200-1000) is a batch loaded machine for use at intermediate and high power levels where wafer heating may become a problem. Both systems will accommodate wafers up to 4 inches in diameter.

The issue of beam heating was of prime importance in the decision to build two essentially different machines. A description of the beam heating effects, and the design of these two machines follows.

2. BEAM HEATING EFFECTS

The energy deposited in a wafer during implantation causes the wafer temperature to rise to a level which depends on the thermal capacity of the wafer, the doping rate and the rate of heat loss. Calculations, made by assuming that the heat loss is by radiation only, help in understanding the reasons behind machine design concepts. (For a more detailed description of this process see Ryding[1].) Clearly these calculations represent the upper limit of temperature that a wafer can achieve, and in practice, wafer temperatures will always be somewhat lower. Some recent experiments by Chen[2] have shown that when slices were held against a massive metal substrate by a number of clips around the circumference, the measured temperatures were a factor of two lower than calculated. However, earlier results by Freeman[3,4] indicate temperatures measured to be closely similar to those calculated. Clearly it is difficult to control heat loss by conduction from a wafer during implantation because of the fragile nature of the wafer and the difficulty in obtaining a good thermal contact between surfaces in vacuum.

We have therefore designed the Varian/Extrion High Dose Implanters so as to satisfy the worst case conditions when heat loss is by radiation only. In this case the maximum temperature rise possible can be calculated and Figs. 1 (a,b,c) illustrate the rise in temperature of a 3 inch silicon slice, assumed to be 0.4mm thick, emissivity E = 0.5, as function of the doping level for ion beams of different energies between 20 and 200 KeV. At a fixed current density, the implant time is directly proportional to the implanted dose and this time is also illustrated.

When photoresist is used as a mask material, the maximum temperature should be limited to 140°C; for implantations using oxide or metal as a mask, higher temperatures may be tolerated. For those implant conditions where temperature is not a limitation, it is advantageous in terms of throughput to implant individual wafers using X-Y electrostatic beam scanning because it is a continuous process which lends itself well to automation, and the scan frequencies used (often in the range of 50-1000 Hz) enable uniform implants to be obtained using implant times as low as 3-5 seconds. The Varian/Extrion Model 200-20AF described in Sec.4 has been built to exploit this high throughput machine up to the limitation of wafer heating.

TABLE I. Current density as a function of beam current for electrostatically scanned and mechanically scanned machines.

Beam Current μA	X-Y Electrostatic Scan on a single stationary 3" Wafers $\mu A/cm^2$	X-Y Mechanical Scan over a batch of 26 3" Wafers $\mu A/cm^2$
75	1.25	.025
200	3.33	.067
500	8.33	.167
1000	16.7	.333
2000	33.3	.667

Table I (columns 1 and 2) lists the beam currents which are equivalent to those beam densities used to develop Figs.1. Using Table I in conjunction with Figs. 1, implant conditions which cause the wafer temperature to rise to maximum temperature limits can be selected. If an implant requires beam currents or energies above these maxima, electrostatic X-Y scanning of a single wafer at a time is probably unsuitable and the beam power should be distributed over a number of wafers simultaneously. In this way, high system throughput is maintained by simultaneously implanting many wafers and so reducing the effective beam power per square centimeter to acceptable levels. (See Table I, column 3).

We have built a machine of this type, Model 200-1000, described in Sec. 3, which uses a fully mechanical X and Y motion of a batch of 25 three or four inch wafers, while maintaining a constant implant angle over the entire wafer surface. The throughput of this mechanical motion system is less than that for electrostatic scanning because the requirements of high uniformity using scan speeds in the order of 100 cm/sec^{-1} requires a minimum implant time per batch of about 10 minutes (25 seconds/wafer) in contrast to the inherently faster scan speeds of approximately 3.10^3 cm/sec^{-1} provided by a typical beam scan system.

We have rejected the concept of hybrid scanning where one scan is electrostatic and the other is mechanical because this system appears to have disadvantages of both fully electrostatic and fully mechanical systems without sufficiently compensating benefits (see Table listing advantages and disadvantages of various scanning systems in Ref.1).

3. THE MODEL 200-1000 IMPLANTER (Photograph 1)

A general view of this machine is shown in Photograph 1 and the performance specifications are listed in Table II.

TABLE II. Basic Specifications of Model 200-1000 Ion Implanter

Energy Range:	5-200 KeV			
Beam Current on Wafer:	$^{11}B^+$	1mA	$^{31}P^+$	2mA
	$^{75}As^+$	2mA	$^{121}Sb^+$	0.75mA
Wafer Size:	up to 4 inch			
Implant Angle:	Adjustable 0-10° with manual wafer flat orientation.			
Machine Throughput:	3 inch wafers/hour			

Doping Level ions cm^{-2}	Without Lock 1mA	Without Lock 2mA	With Lock 1mA	With Lock 2mA
10^{12} - 1.5×10^{15}	53	53	100	100
2×10^{15}	50	50	80	100
5×10^{15}	28	43	37	67
1×10^{16}	17	28	20	37

Dose Uniformity and Reproducibility:	95.5% of readings lie within $\pm 2\sigma$ where $\sigma = 0.75\%$, across a wafer and from wafer to wafer.

The description which follows is most easily understood by reference to the schematic of the system shown in Fig.2 which illustrates the following basic components:

a) Ion Source System (40 KeV)
b) Analyzing Magnet
c) High Voltage Terminal
d) Post-analysis Acceleration (40-200 KeV)
e) Production Wafer Handling System
f) Control System

a) Ion Source

Ions are generated in a cylindrical graphite chamber by an arc discharge (typically 100V and 2A) which is sustained between the anode chamber wall and an axial hot tungsten cathode. Ionization efficiency of the electrons in the discharge is enhanced by the application of a weak (~100 G) axial magnetic field. Beam stability is maximized by closed-loop regulation of the arc current using feed-back control of the cathode heater current. Source gas is fed from a manifold containing three independent gas feed systems. Each system contains a gas cylinder, regulator and a combination leak/shut-off valve. This valve gives independent control of 'leak' and 'shut-off' functions and enables 'pre-set' optimum gas flow to be achieved with minimum set-up time as well as providing "safety interlocked" operation of the gas feed system.

In addition to the gas feed system, a vaporizer module operating at controlled temperatures up to 800°C can be used for low vapor pressure materials in conjunction with a suitable carrier gas such as nitrogen or argon if required.

Particular attention has been given to minimizing down time when source maintenance or change is required. In this regard, the thermal capacity and the source volume was minimized and the source region was isolated from the rest of the system by an integrated gate valve in order to facilitate rapid vacuum cycling (10 minutes).

Ions are extracted from a 40mm by 1.5mm slit in the graphite chamber using a +40kV, 20mA regulated power supply. A conventional two electrode system; with the extraction electrode biased at -2kV to minimize space charge expansion of the beam and reduce 'back' electrons, is used to produce a beam which diverges in the horizontal plane with an angle of $\pm 1.5°$ (vertical source slit) and is essentially parallel in the vertical plane.

Fine steering of the beam is accomplished by lateral and rotational movement of the electrode assembly and in addition, the ground electrode is water cooled and contains a baffle system to trap condensible vapors emitted by the source, thereby minimizing contamination of the vacuum system.

b) Analyzing Magnet

A 90° double focusing magnet is used both for mass analysis and imaging of the ion beam extracted from the source. The entrance and exit magnet shim angles were computer designed to minimize aberrations and to focus the beam in both the vertical and horizontal (dispersive) planes. In this way, beam transmission through the magnet and to the target is close to 100%. Using an asymmetric configuration with a shorter object distance than image distance it has been possible to transform the beam into a symetrically converging cone with an essentially circular image diameter of approximately 6mm. The magnet has a radius of curvature of 33cm, a mass range of up to 130 at 40 KeV, and a resolution of 200 (based on full width at half height). It is constructed with the magnetic field return yoke on the inside to facilitate easy removal of the magnet vacuum chamber and the stainless steel vacuum chamber is provided with 6mm thick aluminum liners to prolong clean operation of the analyzer region.

The amount of beam current allowed to reach the target is controlled by a water cooled variable graphite aperture which is located just behind the water cooled resolution aperture of the analyzer system. The variable aperture is controlled by a slow servo-loop from the measured beam signal and provides the operator with an optional method of beam control in addition to the conventional technique of ion source parameter adjustment.

c) High Voltage Terminal

The machine features a'pre-analysis' configuration in which the beam is accelerated to the final on-target energy after analysis of the beam. In order to benefit from the many advantages of a target at ground potential, the entire ion source, analyzer, and associated vacuum systems are self-contained in a compact high voltage terminal (1.6 m x 1.6 m x 1.6 m) which is elevated to high potential (160 kV) for post-acceleration of the beam. Isolation power to the terminal is supplied by a 15 KVA motor generator set through an insulating shaft. Cooling to the terminal is supplied by a closed loop Freon or de-ionized water system.

d) Post-Analysis Acceleration

The acceleration tube is a three electrode system with a resistive divider network to give equal acceleration across the two gaps (80 kV per gap maximum). It has been designed to minimize space charge expansion of the beam. A demountable construction has been used and a re-entrant configuration permits the strategic location of lead 'X'-ray shielding outside the vacuum system, but close to the acceleration region. The acceleration voltage is provided by a highly regulated air-insulated 160 kV, 4mA R.F. power supply.

Recent experiments in which the machine was operated in a deceleration mode for low energy implants, indicated that useful beam currents of 1 milliampere of B^+, P^+ and As^+ were available on target at energies down to 5 KeV.

e) Production Wafer Handling System

The Model 200-1000 incorporates a mechanical scan system in which wafers are scanned repetitively in two directions through the stationary beam. As described earlier, the choice of a mechanical scanning technique enables 26 three inch wafers to be processed simultaneously through the beam so that the effective beam power on each wafer is reduced to an acceptable value. Alternatively, if a smaller number of wafers need to be implanted, a half batch of up to 13 wafers may be processed in roughly half the time.

The basic scan mechanism operates on a parallelogram linkage concept as shown schematically in Fig. 4 and Photograph 2. The axis of rotation is horizontal and the resulting wafer movement can be compared to that of a pleasure ground 'Ferris' wheel except that the angle of the wafer to the beam is held precisely constant and can be readily adjusted as shown.

Wafers are mounted in the mechanism on thirteen aluminum wafer holders, each holding two wafers. After the wheel is loaded it is given a programmed acceleration to a constant velocity of 30 r.p.m. using a precision d.c. motor/tachometer servo system. This rotary motion is transmitted into the vacuum system through a differentially pumped double shaft seal.

The horizontal motion is provided by a precision stepping motor and transmitted into the vacuum system through a welded stainless steel bellows assembly.

The beam current is measured on target using a conventional 'deep Faraday' system with electrostatic electron suppression. A water cooled aluminum ground mask defines the maximum possible beam size on target and ensures correct overscan of the wafers. Normally this mask intercepts less than 10% of the total beam. The mechanism itself with an appropriate water cooled shield is electrically isolated and forms the back of the Faraday system so that the total on target beam is monitored throughout the implant. A water cooled aluminum beam gate is located just in front of the mechanism and can be used for initial beam set-up and interruption of the implant.

The scan pattern generated by this mechanism is controlled by a microprocessor unit (see Section f). The required doping level is selected on thumb wheel switches and the horizontal scan velocity is calculated so that the wafers receive the precise doping level after the wheel has made an exact number of traversals (usually 10). Since the beam current may vary slightly during the implant, the transverse scan velocity is automatically controlled in constant proportion to the on-target beam current in order to compensate for any fluctuations in beam intensity.

The actual uniformity achieved is a complex function of the scan pattern and the beam intensity profile. Any variation in beam intensity profile during implantation will cause systematic errors in the uniformity, but these effects can be minimized by allowing the wafers to make several "coarse" scans through the beam, rather than a single "fine" scan. In this system, a high uniformity is ensured by control of the minimum implant time based on both calculations and actual measurements. Some typical results are displayed in Fig.5 which shows the uniformity profile measured after a 150 KeV 4 X 10^{15} ions/cm^2 boron implant on the Varian/Extrion High Current Implanter. The heavy contours represent the mean sheet resistance ρ_s while the lighter contours differ from ρ_s in increments of 1%. The symbols + (-) signify the location of experimental points whose resistance values are greater (less) than the mean value. 118 data points were taken. Below the wafer map the mean sheet resistance is shown on the left and plus (and minus) twice the relative standard deviation is shown on the right.

Throughout the implant, the microprocessor monitors and ensures correct operation and sequencing. If at any time a fault occurs, the system is placed in a 'hold' condition. At this time the beam gate in front of the wafers is closed and the horizontal scan motion is interrupted. The fast rotary motion is however, maintained so that a 'hot spot' is avoided when the implant procedure is continued.

During the load and unload procedure, the rotary drive system is used in a position servo mode in conjunction with a shaft encoder to index each wafer holder to a predetermined load/unload position. A fully automated system which will load and unload wafers through a vacuum lock is now under construction and will be completed later this year.

f) Control System

The control system of the 200-1000 Implanter contains two novel features. The first one, already referred to, is the use of a solid state programmable microprocessor (Intel 8080 C.P.U. Chip) as the center of the doping level control system. This system performs doping level calculations, directs all end station sequencing and scan motions, and provides for sub-system interlocks. It is believed to be the first application of this rapidly growing technology to a production ion implanter.

The second feature is the use of a fiber-optic data communication system between the high voltage terminal and the control console at ground potential. All commands to the terminal power supplies and controls are transmitted using binary coded signals and all terminal readout signals are transmitted to the operator using a pulse width modulation technique. In this way, insulating control rods and windows have been eliminated and a building block has been established for the fully automated computer control of an ion implanter.

4. THE VARIAN/EXTRION MODEL 200-20AF ION IMPLANTER (Photograph 3)

This machine employs electrostatic scanning in both X and Y directions to achieve dose uniformity on a single wafer. The basic specifications of this machine are listed in Table III, and a beam line layout is shown in Fig.6.

TABLE III. Basic Specifications of Model 200-20AF Ion Implanter

Energy Range:	35 - 200 KeV (decreasable to 10 KeV at reduced current levels)					
Beam Current:	Scanned on wafer 35-100 KeV					
	$^{11}B^+$	200µA	250µA	$^{31}P^+$	250µA	400µA
	$^{75}As^+$	250µA	400µA	$^{121}Sb^+$	100µA	150µA
Wafer Size:	up to 4 inch					
Implant Angle:	Adjustable from 0 - 10° Maximum angular scan angle< $\pm$ 2.8° (Automatic wafer flat orientation optional)					
Maximum Doping Rate at 400µA Scanned Current:	4 inch 2.5×10^{13} ions/cm^2/sec. 3 inch 4.0×10^{13} ions/cm^2/sec.					
Maximum Throughput for 5 Second Implants:	250 wafers/hr, 3" or 4"					
Dose Uniformity and Reproducibility:	95.5% of readings die within $\pm$ 2σ, where σ = 0.75%, across a wafer and from wafer to wafer.					

The description which follows can most easily be understood by reference to this figure. The basic components of the system are:

a) Ion Source System
b) Extraction System (25 KeV)
c) Analyzing Magnet

d) Post-Analysis Acceleration (25-200 KeV)
e) Quadrupole Triplet Lens
f) X-Y Electrostatic Scanning with Neutral Beam Trap
g) Wayflow Production Wafer Handling System

In many respects the terminal of this machine (components a) - c) and the post-analysis acceleration section (d) are conceptually similar to the comparable components of the Model 200-1000 previously described. However, in order to restrict the beam currents to suitable values, the ion source has been scaled down in size (the exit slit has been reduced to 10mm x 1.5mm) and because of this, downstream components are also smaller. The 90^{o} analyzing magnet pole piece gap has been reduced to 3.8cm and the radius of curvature chosen to be 25cm with fields capable of analyzing beams of ion mass up to 75 amu (increasable to 140 amu) at the selected extraction energy of 25 KeV. Typical spectra are shown in Figs. 7a and 7b. The post acceleration of the beam, from 25 to 200 KeV, is achieved by using a highly stabilized air insulated 175 kV 2mA R.F. power supply. The beam scan and target chamber of this machine is totally different from that of the Model 200-1000 and will now be described in detail.

e) Quadrupole Triplet Lens

As shown in Fig. 6, the beam leaving the acceleration tube is divergent; a quadrupole triplet lens is used to refocus the ion beam at the wafer. The voltages applied to the lens element automatically track with the beam energy. The beam size at the wafer is a complex function of source emittance, optical aberrations, beam current, ionic mass, and energy. Typically at 25 KeV, the beam diameter on target is about 2cm, whereas at 200 KeV, it is about 0.5cm.

f) X-Y Electrostatic Scanning with Neutral Beam Trap

The beam scanning system provides the capability to cover a relatively large area with small diameter intense ion beams. The beam scanner system consists of two pairs of orthogonally mounted electrostatic deflector plates driven by triangular voltage waveforms.

The beam is scanned in a random manner so that path retracing is minimized and a uniform dose can be obtained over the desired area. The high frequency scan is fixed at 1000 Hz and the low frequency is adjustable from 10 Hz to 100 Hz.

Both plates are balanced + and - in voltage. The non-linearity of the waveform on each plate is less than 1% over the useful portion of the signal. The non-linear turn-around portion, which is less than 5% of the total amplitude, is not used to scan the beam across the target; a knife-edge mask, correctly sized to compensate for the scanned beam divergence angle is mounted in front of the target. The scanner plates are contoured to minimize aberrations.

The neutral particles formed in the accelerator tube and in the beam line beyond the analyzing magnet must be removed from the ion beam reaching the target. This is achieved by deflecting the ion beam 7° from its previous axis by superimposing a dc voltage to the horizontal scanner plates. The neutral beam, unaffected by the electrostatic deflecting field, remains on its original path and is trapped on a collector plate located just beyond the scanner. The advantages of this scan system have been discussed previously[5]. The dc voltage applied to the horizontal deflection plates, together with the triangular waveform, is programmed to track with the beam energy and can be used over the entire rated energy range of the implanter with no separate adjustment required.

The doping level is determined by a precision current integrator which controls a beam gate. Fast gating action (less than 1 msec.) is accomplished by superimposing a dc voltage to the vertical scan plates which deflects the beam away from the target region.

g) Wayflow Production Wafer Handling System

Vacuum lock systems have been in use on our machines for nearly four years and on over 150 systems. However, in these earlier systems it is necessary to manually load or dump transfer from the standard linear wafer carriers generally used in the industry into a machine compatible wafer carrier. Increased automation makes it essential that ion implantation systems operate directly from conventional linear carriers (operator loaded) or from conveyor systems directly, and the newly designed Wayflow End Station has been designed to meet this requirement. The operation can best be understood by reference to Fig. 8 which shows the form of the end station when fed from a conveyor. When fed from a carrier, feed block (A) is replaced by a standard wafer carrier, Photograph 4.

One implanted wafer at a time enters the entrance lock (B) with the upper seal retracted. The upper seal is closed and the vacuum lock, which has a volume only slightly greater than the wafer volume, is evacuated in 5 seconds. After evacuation, the lower seal plate is retracted and the wafer slides onto the wafer holder (C) which is in position X. The wafer holder rotates to position Y which may be set externally at any angle between 0° and 10° to avoid channeling. In position Y, the wafer on its holder closes off the main Faraday cage (not shown) and is implanted. While each wafer is being implanted, the entrance vacuum lock receives the next wafer and is pumped to a rough vacuum so that machine dead time between implants is kept to a minimum, typically 3 or 4 seconds. After implantation, the wafer holder moves to position Z; there the implanted wafer slides into the exit vacuum lock (D). After venting to atmospheric pressure, the wafer moves onto the conveyor belt (E) or into a standard commercial wafer carrier.

The end station has a sophisticated optical wafer sensing system which positively identifies the transit of a wafer through each section of the end station, thus eliminating sequencing errors. By using non-vertical movements, wafers move gently from position to position, greatly reducing the chance of edge damage. Standard, well protected "O" rings are used in order to maintain the greatest cleanliness for long term seal reliability. The end station is extremely simple and reliable. All operating mechanisms are external to the vacuum system and the end station can be set up and maintained with ease.

ACKNOWLEDGMENT

The wafer map shown in Fig. 5 was provided by D.S. Perloff and F.E. Wahl of Signetics Corp., Sunnyvale, California. The authors gratefully acknowledge their permission to use this information.

REFERENCES

1. Ryding, G., Proc. International Microelectronics/Semi-Conductor Conference, Anaheim, 1974, P.107.

2. Chen C.S. Determination of Wafer Temperature Rise During Ion Implantation, Burroughs Corp., San Diego, Calif., Private communication.

3. Freeman, J.H., Caldecourt, L.R., Done, K.C.W. and Francis, R.J., Proceedings of European Conference on Ion Implantation, 1, Reading, 1970.

4. Freeman, J.H., Chivers, D.J., Gard, G.A., Hinder, G.W., Smith, B.J. and Stephen, J., Proc. of the 4th International Conference on Ion Implantation, Osaka, 1974, 555.

5. Ryding, G., Wittkower, A.B., IEEE Trans. on Manufacturing Technology, MFT-4, No.1, 1975, 21.

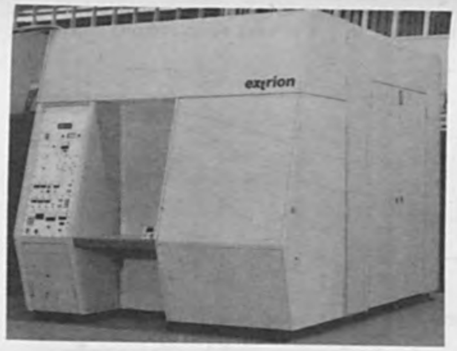


Photo 1. Model 200-1000 Ion Implanter

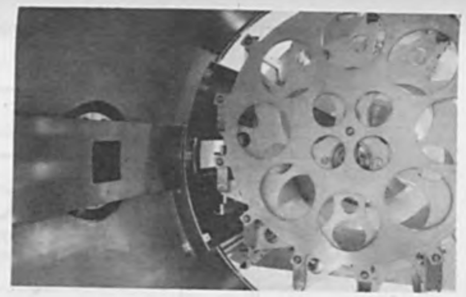


Photo 2. Photograph of the Model 200-1000 End Station showing the parallelogram "Ferris Wheel" mechanism exposed.

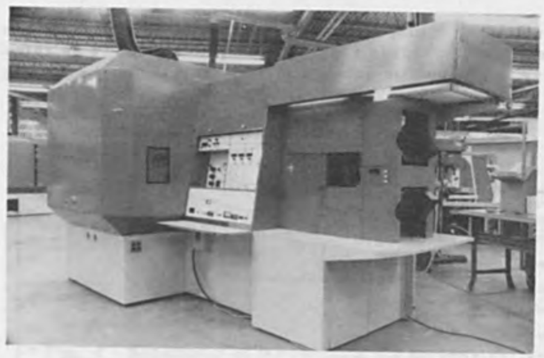


Photo 3. Model 200-20AF Ion Implanter with Wayflow End Station

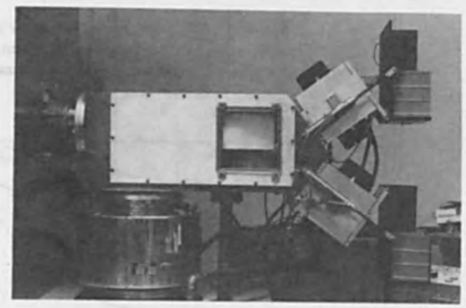


Photo 4. Photograph of Wayflow End Station (exposed) showing one wafer carrier holding unimplanted wafers mounted above entrance lock and wafer carrier in receiving position.

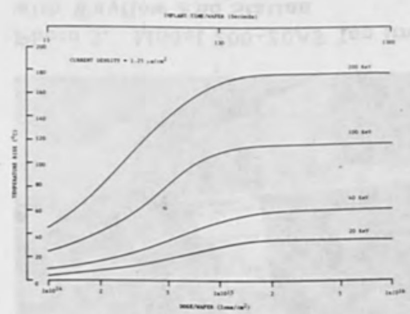


Fig. 1a - Temperature rise of a 3 inch silicon wafer as a function of dose at various beam energies, at a current density of 1. 25μa/cm^2.

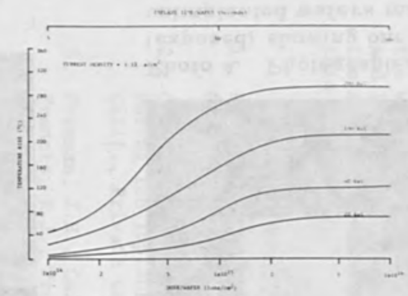


Fig. 1b - Temperature rise of a 3 inch silicon wafer as a function of dose at various beam energies, at a current density of 3. 33μa/cm^2.

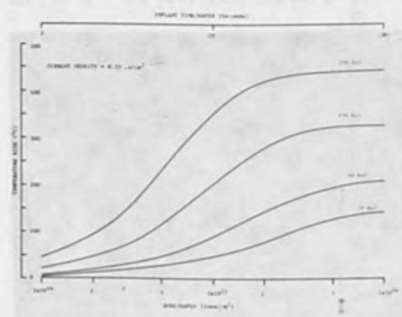


Fig. 1c - Temperature rise of a 3 inch silicon wafer as a function of dose at various beam energies, at a current density of 8. 33μa/cm^2.

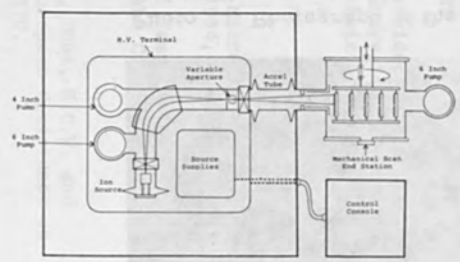


Fig. 2 - Model 200-1000 Beam Line Layout

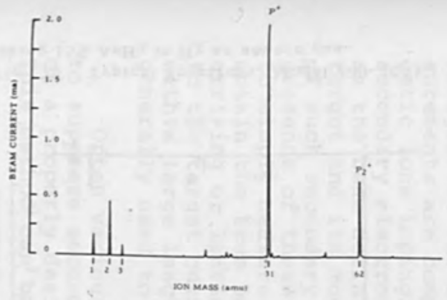


Fig. 3a - Typical Spectrum, Model 200-1000 using 15% PH_3 in H_2 as source gas.

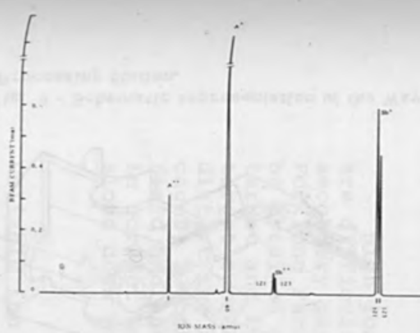


Fig. 3b - Typical Spectrum, Model 200-1000 using solid antimony with argon as carrier gas.

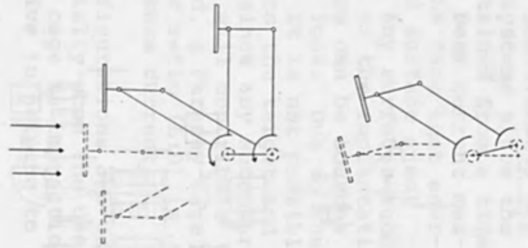


Fig. 4 - Schematic of the rotor in the mechanically scanned target chamber showing the linkage which holds the wafer angle constant.

Fig. 5 - Uniformity profile measured after a 150 KeV 4 x 10^{15} ions/cm^2 boron implant in the Model 200-1000 Implanter.

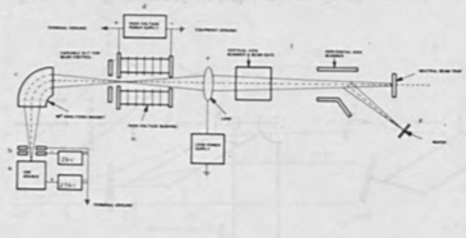


Fig. 6 - Model 200-20AF Beam Line Layout

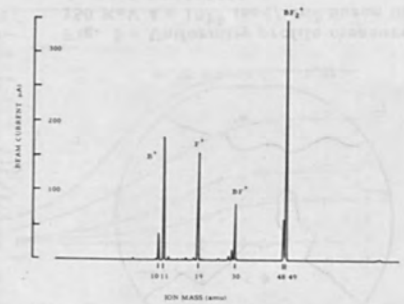


Fig. 7a - Typical Spectrum, Model 200-20AF using BF_3 as source gas.

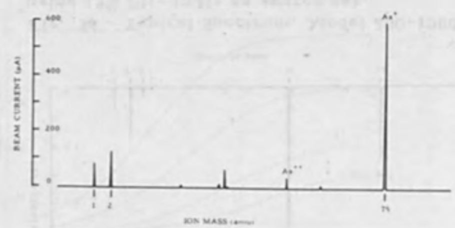


Fig. 7b - Typical Spectrum, Model 200-20AF using 15% AsH_3 in H_2 as source gas.

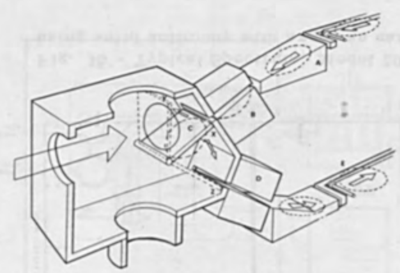


Fig. 8 - Schematic representation of the Wayflow Processing Station.

HIGH CURRENT DENSITY ION BEAM MEASUREMENT TECHNIQUES

Wen C. Ko and Erich Sawatzky

IBM General Products Division, San Jose, Ca 95193

ABSTRACT

High ion beam current measurements are difficult due to the presence of the secondary particles and beam neutralization. For long Faraday cages, true current can be obtained only by negative bias on the target and by summing the cage wall and target currents, otherwise, it will greatly distort the beam. For short Faraday cages, a combination of small magnetic field and the negative target bias results in correct beam current. Either component alone does not give true current.

INTRODUCTION

Reliable ion beam current measurements are necessary in ion implantation systems since the implanted dose is generally obtained from a time integral of the beam current. Beam current measurements are complicated by the fact that energetic ions impinging on a solid surface eject secondary electrons and ions. Any surfaces such as the beam defining apertures or the implantation target and its mounting hardware can be sources of such secondary electrons or ions. Due to the presence of these secondaries, it is not possible to simply measure the current to the target and obtain the true beam current, since any secondaries arriving or leaving the target will contribute to the target current. Instead, a Faraday cage with a large length to diameter ratio (L:D) is generally used to measure the beam current.

Often various biasing configurations are used to suppress secondaries, especially when the use of a properly designed Faraday cage is impractical. This method can be very effective in helping to obtain the true beam current, especially at low current densities. At high beam current densities biasing must be applied very judiciously as it can significantly disturb the beam profile.

The paper describes beam current measurements in a high current (>1mA) test stand, with the objective of accurately measuring beam currents in a high current density beam with a non-ideal Faraday cage arrangement.

EXPERIMENT AND DISCUSSIONS

The measurements were performed in a Faraday cage configuration where the wall of the cage consisted of two sections so that a "good" Faraday cage (9:1) could be converted to a "poor" Faraday cage (1.5:1) without breaking vacuum or shutting the beam off. The currents to the target and the Faraday cage wall were measured separately. Both parts could be biased with respect to ground. Two Langmuir probes 3.5" apart were mounted in a single vacuum feedthrough in such a way that they could be moved through the beam simultaneously, one probe being about 1/2" in front of the target while the other is 4" upstream. These probes were used to measure the beam profile under various biasing conditions.

For example, a 1mA, 50 keV argon beam has a FWHM of 0.4" at the target without any biasing. When the wall is biased either negatively or positively by more than 15 volts, the beam expands to approximately 2". A similar blow-up is observed when the target is biased positively. However, when the target is biased negatively, the beam size does not increase. Similar effects were observed with a 1mA, 50 keV arsenic beam. Again, only negative bias on the target does not cause beam blow-up. The beam expansion due to biasing depends on target material because the secondary electron yield is different for different targets. The effect of biasing on beam size is approximately the same when the full length of the wall is biased or either section alone is biased.

In a "good" Faraday cage (L:D = 9:1) almost all secondary electrons and ions are contained within the cage and biasing has a negligible effect on net ion current measurements. Only the beam profile is affected by biasing, except for a negative bias on the target.

When the Faraday cage is shortened to L:D = 1.5:1, secondaries created inside and outside such a structure cannot be contained in their respective spaces. With this configuration, a negative bias on the target alone does not give the true beam current. However, a small magnetic field near the target in addition to the negative bias results in measuring the true beam current with the short tube. A magnetic field alone does not give the true current, the magnetic field has no effect on beam size. It appears that the true current in a "poor" Faraday cage configuration in high current density beam system can only be obtained with combination of negative target bias and magnetic field. The magnetic field used here was essentially in the target plane and varied from $\sim$5 to $\sim$10 Oersted across the 2" target. A variation in magnitude by a factor of $\sim$3 and in direction by $\pm$30 degrees still gave the true beam current.

Using this Faraday cage, the secondary electron yields were measured on various ion and target combinations. The secondary electron yields initially decrease with time and hence with dose and then approach a steady state value. This is believed to be related to the fact that the target surfaces are initially saturated with absorbed gases. As the dose increases, the surface contaminants are removed due to sputtering and a rise in target temperature and the yields of clean surfaces are approached. In the steady state regime, the secondary electron yields vary linearly with incident ion energy below about 50 keV. At higher energies, the yields increase less rapidly with ion energy.

PRECISION MASK ALIGNMENT FOR X-RAY LITHOGRAPHY

John H. McCoy and Paul A. Sullivan
Hughes Research Laboratories
Malibu, California 90265

ABSTRACT

The application of x-ray lithography to the fabrication of high density integrated circuits and microwave transistors requires precision mask alignment. This is an updated report on a laboratory system which detects misalignment using x rays and provides correct alignment to within 0.1 μm. The results of an analysis of the system using either transmitted x rays or fluorescent x rays is given. A comparison of performance shows three viable alignment detection schemes. An important conclusion is that hard (1 to 2 Å) x rays should be used for transmission detection. In addition, we give results of tests of the system using optical simulation of alignment detection.

INTRODUCTION

X-ray lithography is a technique for replicating submicron linewidth patterns over large areas. It has been applied to the fabrication of bipolar[1] and MOS transistors,[2] surface acoustic wave devices,[3,4] and magnetic bubble devices.[5,6] The fabrication of very high density integrated circuits and microwave transistors[7] with submicron linewidths is a goal of electron beam and x-ray lithography which is being explored in several laboratories. Achieving this goal will require a capability for precise alignment of an x-ray mask to existing patterns on a semiconductor wafer, as well as improved processing techniques for large area, defect-free masks with good dimensional control. Both of these requirements become increasingly difficult as the linewidth is reduced. In addition, x-ray lithography becomes economically attractive for this application only if resist materials are available which have very high sensitivity and good dimensional stability for submicron linewidth patterns. This paper considers only one of these requirements, i.e., a precision mask alignment system for x-ray lithography. For the purposes of this discussion we assume that alignment accuracies are not limited by distortions or dimensional stability of the mask or wafer.

Commercially available mask aligners claim positioning accuracies of 0.25 μm; one such system uses stepping motors with 0.125 μm increments. The alignment accuracy of these machines is limited by their ability to hold the set position as the wafer is clamped in preparation for exposure. The accuracy of alignment is also limited by the operator's ability to distinguish when a feature on the mask is located exactly in the center of a feature on the wafer. In proximity or contact mask alignment systems there is a depth-of-field limitation on high power microscope objectives. This limitation may be overcome by the use of dual focus lenses. The low contrast of some wafer patterns is also a severe limitation when examining small features. Some mask aligners offer automatic alignment capability[8] which has reduced the burden on the operator. Although these machines do not resolve the image of the alignment mark directly, they still suffer the limitations of visible light optics, i.e., low resolution and contrast due to reflection, diffraction and scattering of light.

X-ray lithography is suitable for replicating submicron linewidth patterns from an off-contact mask because reflection, scattering, and diffraction effects are negligible. A resolution of 0.1 μm has been demonstrated,[6] although practical semiconductor devices may be limited to linewidths of 0.4 μm.[9] Depending upon device design and runout tolerances of the mask and wafer we believe that 0.1 μm alignment accuracies may be required in future systems. This is beyond the resolving capability of an optical system; consequently x-ray detection techniques are considered in this paper.

The alignment of an x-ray mask to existing features on a wafer requires both the detection and physical correction of the relative position of the mask and wafer. Our approach has been to detect special alignment marks on both the mask and wafer with x rays in order to achieve what will ultimately be better alignment than is possible by optical means. The following sections will describe the limitations on x-ray alignment which do exist and the electromechanical system we have developed to accomplish the alignment.

ALIGNMENT SYSTEM

An alignment system for laboratory use has been built for investigating the concepts of positioning with piezoelectric transducers and detecting alignment marks with x rays. The alignment system is shown in a simplified cross section in Fig. 1. A modified microscope stage is used to provide coarse alignment in the traditional x, y, and θ axes. In addition, it provides tilt and height adjustment so that the top surface of the mask can be made parallel with the top plate. The microstage consists of a lightweight vacuum chuck suspended on four piezoelectric ceramic cylindrical transducers. The entire microstage rides atop the mechanical stage. Properly matrixed drives to the transducers[10] provide orthogonal x and y movement of ± 10 μm and rotation of ± 0.2 mR. The wafer carrier also employs vacuum hold-down, and the differential screws are preadjusted to provide mask-wafer separation. One type of alignment detection is illustrated in Fig. 1. In this transmission detection scheme the left and right x-ray detectors are masked to intercept only x rays transmitted up through the mechanical stage, the mask, and the wafer in the area of each alignment mark set.

The control of the alignment system is provided by an electronic feedback circuit essentially the same as previously published.[10] Oscillation or dither of the mask allows phase sensitive detection of the x-ray signals for the derivation of x, y, and θ correction signals. Experiments have shown that dither up to 170 Hz is possible before mechanical resonances are encountered, and large amplitude and phase deviations occur. Therefore, we have analyzed the feedback system without taking into account the dynamics of the microstage. The loop diagram for one dimension is shown in Fig. 2. The summing junction for the complementary system results in an error signal x from the misalignment of the mask position x_m and the wafer position x_w. These are the average or undithered positions. It can be shown that the phase sensitive detector output y is linear with x and does not depend upon the dither amplitude, but the expression does not hold when the dither amplitude is less than the error x. The detector gain A_o is derived in the following discussion on alignment detection and the noise limited performance.

Alignment Detection — Two general types of alignment detection schemes using transmitted x rays or fluorescent x rays have been analyzed and both appear feasible. In the transmission system x rays are transmitted through complementary absorbing marks on the mask and wafer as shown in Fig. 1. A straightforward analysis of the alignment mark, x-ray detector, and synchronous demodulation shows that the gain A_o is equal to $2\pi JL\,(1 - \epsilon_A)$ where J is the effective photon flux as passed through the wafer and the transparent part of the mask, L is the effective length of the mark set (in one dimension), and ϵ_A is the fraction of the x rays transmitted through the absorber.

The noise-limited performance of the system is determined by the shot or quantum noise of the detected signal. This noise can be indicated by the added noise n at the second summing junction. The remaining forward loop consists of an integrator, the electronic amplifiers, and the microstage transfer function, all of which are lumped together in A_I/s where the Laplace operator s indicates the integrator. With no noise, such a servo system has no steady-state error, and the step input time response constant is $1/A_oA_I$.[11] The noise introduced as n does, however, cause an rms steady-state error which can be compared with x by the division $n_r = n/A_o$. The chief noise source is the shot noise in the background x rays transmitted through the mask and wafer and through the aligned absorber marks. This background is limited by the aperture area A of the detector. The noise limited performance for the transmission system is then

$$n_r = \left(\frac{\Delta f}{J_o}\right)^{1/2} \frac{1}{\pi 2^{1/2}} \frac{\epsilon_A}{(\epsilon_M \epsilon_W)^{1/2}\,(1 - \epsilon_A)} \frac{A^{1/2}}{L} \qquad (1)$$

where Δf is the bandwidth; J_0 is the x-ray flux incident upon the mask-wafer set; and ϵ_M and ϵ_W are the transmission fractions of the mask substrate and the wafer, respectively.

The selection of a proper x-ray wavelength for detecting alignment can be made from Fig. 3 which shows the transmission factor in Eq. 1 (the function of ϵ_A, ϵ_W, and ϵ_M) plotted as a function of wavelength. This x-ray transmission factor was calculated using the wavelength dependence of the mass absorption coefficients from Heinrich as updated by the authors.[12] The curves shown are dominated by the values of ϵ_A and ϵ_W. The ϵ_M factors for both silicon masks or Mylar masks give the same curves. This transmission factor must be minimized, and only by thinning the silicon wafer to 100 μm can one achieve a factor under 200 at the Al K wavelength. However, normal thickness 200 μm wafers can be used in the 1 to 2 Å region. The $(1 - \epsilon_A)$ factor describes the absorber contrast and causes the rise below 1 Å. Since special thinning of the wafer is undesirable, the use of a short wavelength radiation such as Cu K (1.5 Å) is warranted. Such x rays would cause little exposure because of poor absorption in the resist.

A possible source of short wavelength x rays is the continuum from a soft x-ray source.[12] Thus alignment could be provided during exposure to correct for drift and vibration. Alternatively, two low power cources of Cu K x rays could be used to provide collimated x rays only over each mark. The use of a special source other than the exposure source requires that the wafer and mask be clamped together after alignment. Then, however, several clamped assemblies could be exposed simultaneously by one high power soft x-ray source. Alternatively, a copper source could provide L x rays (13.3 Å) for exposure and K x rays (1.5 Å) for alignment.

A second alignment scheme uses fluorescent x rays which are detected back through the thin mask. One of several possible systems is shown in Fig. 4, which shows only the mask-wafer portion of Fig. 1, with the two parts interchanged. To eliminate unwanted background, pure secondary fluorescent P K x rays are excited by Cu K x rays which are produced by an electron beam. The P K x rays then excite Si K x rays when the complementary absorbing marks are misaligned. A nonfluorescing thick aperture mask such as Ta is required to limit the area illuminated by the P x rays to the active marks. The two phosphorus targets provide high efficiency point sources of x rays just over each mark set. The particular scheme requires a nonsilicon mask (Mylar or Kapton) and a silicon or SiO_2 coated wafer.

The noise performance of the fluorescent scheme can be analyzed in a way similar to the transmission case but with added factors,

$$n_r = \left(\frac{\Delta f}{J_o}\right)^{1/2} \left(\frac{4\pi}{D_F F_F}\right)^{1/2} \frac{\epsilon_A{}^{1/2}}{1 - \epsilon_A} \frac{A^{1/2}}{L} . \tag{2}$$

The solid angle of the detector is D_F and the fluorescence factor F_F accounts for the quantum efficiency for the fluorescent system.

Again the chief noise is contributed by the background of incident P K x rays and emergent Si K x rays which are transmitted through the gold and tantalum absorber marks. The fluorescent system must be chosen carefully. A system at shorter wavelengths may be used, such as the detection of tantalum L fluorescence from a tantalum mark on the wafer which is identical in shape to the gold mark on the mask; but unwanted fluorescence from gold would require a wavelength or energy selective x-ray detector. Such detectors will not work at the high counting rates (above 10^6 cps) required for a reasonable servo bandwidth Δf. The advantage of the system shown in Fig. 4 is that the P K x rays and the Si K x rays are below the threshold for exciting Au M x rays.

The theoretical performance of these alignment schemes is summarized in Table 1. For all three, the noise limited alignment parameter n_r is chosen to be 0.1 μm; the mask absorber is 0.5 μm gold, and the wafer absorber 0.5 μm tantalum. In each case a tradeoff must be made between the servo bandwidth and the acquisition range. This latter factor determines the maximum initial misalignment or the maximum misalignment caused by a disturbance from which the system can recover. For the mark shown in Fig. 5, the range is the dimension of a small square R. The $A^{1/2}/L$ factor must be minimized by proper design of the alignment mark while maintaining a reasonable value of R. The area A is actually the area of the absolute absorber aperture. Thus, the mark should be larger than the aperture area so that aperture prealignment is not critical.

Different tradeoffs are obviously possible other than the ones shown in Table 1. The 50 Hz bandwidth is desirable for obtaining continuous alignment during exposure, but the longer alignment times associated with a low bandwidth of 2 Hz would be acceptable for align and clamp operations. The table shows that transmission alignment of GaAs wafers is also possible with the necessary bandwidth reduction. The back fluorescent system works equally well with all types and thicknesses of wafers.

EXPERIMENTAL RESULTS

The system described in Fig. 1 has been tested using light field-dark field emulsion plates as mask and wafer, visible light instead of the x rays, and photomultiplier detectors. A simple mark consisting of one 5 μm by 60 μm line in each direction was used ($A^{1/2}/L$ = 0.28). A noise limited performance of 11 Å was calculated for this system because a high photon flux was used. The electronic system achieved 3 axis lock-in alignment, but the system was susceptible to large disturbances which caused occasional loss of alignment.

A new set of alignment marks similar to the ones shown in Fig. 5 has been installed in order to provide a large acquisition range of 25 μm and the entire system has been mounted on an air piston vibration isolation table. The results are very encouraging and steps are being taken to implement the transmission x-ray detection system.

CONCLUSIONS

The results of a thorough analysis of x-ray mask alignment using x-ray detection of alignment has been presented. It must be concluded that such x-ray detection will provide an ultimate accuracy of alignment better than possible by optical means and which will be compatible with the high resolution capability of x-ray lithography for making electron devices. The detection of x rays passing through complementary alignment marks and through a normal thickness wafer is practical if hard x-rays less than 2 Å are used. Moreover, these x rays can be the continuum from a high brightness Al K x-ray source. Alternately, two low-power Cu K x-ray sources can provide the transmission alignment source efficiently. A second concept in x-ray alignment detection uses fluorescent x rays which are detected back through the thin mask. A special fluorescent x-ray source is required as is a nonsilicon mask, but all limitations of wafer material and thickness are removed. The comparison of performance shows that with proper tradeoffs of servo bandwidth and acquisition range, three specific schemes of alignment are viable. The results of tests of the electromechanical laboratory alignment system using optical alignment detection indicate the practicality of implementing the x-ray mask alignment system complete with x-ray alignment detection.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the special efforts of Eugene Stevens of this laboratory for the mechanical design and fabrication of the x-ray mask alignment system.

REFERENCES

1. S.E. Bernacki and H.I. Smith, "X-Ray Lithography Applied to Silicon Device Fabrication," Proc. of 6th International Conf. on Electron and Ion Beam Science and Technology, (Electrochemical Society, Princeton, 1974), p. 34.

2. S.E. Bernacki and H.I. Smith, "Fabrication of Silicon MOS Devices Using X-Ray Lithography," IEEE Trans Electron Devices, ED 22, 421 (1975).

3. H.I. Smith, D.L. Spears, and S.E. Bernacki, "X-Ray Lithography: A Complementary Technique to Electron Beam Lithography," J. Vac. Sci. Technology 10, 913 (1973).

4. P.A. Sullivan and J.H. McCoy, "Optimized Source for X-Ray Lithography of Small Area Devices," J. Vac. Sci. Technology 12, 1325 (1975).

5. D. Maydan, G.A. Coquin, J.R. Maldonado, S. Somekh, D.Y. Lou, and G.N. Taylor, "High Speed Replication of Submicron Features on Large Areas by X-Ray Lithography," IEEE Trans. Electron Devices, ED 22, 429 (1975).

6. R. Feder, E. Spiller, and J. Topalian, "Replication of 0.1 μm Geometries with X-Ray Lithography," J. Vac. Sci. Technology, 12, 1332 (1975).

7. F.S. Ozdemir, G.O. Ladd, D.D. Loper, F.W. Cleary, and N. Hirsch, "Electron Beam Microfabrication of GaAs SBFET's," This conference, paper 198.

8. K.G. Clark, "Automatic Mask Alignment in MSI/LSI Processing," Solid State Technology, p. 48, Feb. 1971.

9. B. Hoeneisen and C.A. Mead, "Fundamental Limitations in Microelectronics -I. MOS Technology," Solid-State Electronics 15, 819 (1972).

10. J.H. McCoy, and P.A. Sullivan, "Progress in X-Ray Lithography," Proc. of 6th International Conf. on Electron and Ion Beam Sci. and Technology, (Electrochemical Society, Princeton, 1974) p. 3.

11. C.J. Savant, Jr., Control System Design 2nd Ed., (McGraw-Hill, New York, 1964) p. 84.

12. P.A. Sullivan and J.H. McCoy, "Determination of Wavelength and Excitation Voltage for X-Ray Lithography," IEEE Trans. Electron Devices, April, 1976.

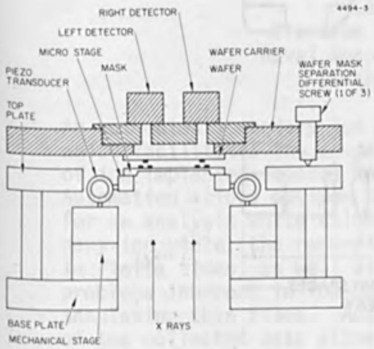


Fig. 1. Outline drawing of the electromechanical system for x-ray mask alignment showing transmission x-ray alignment detection.

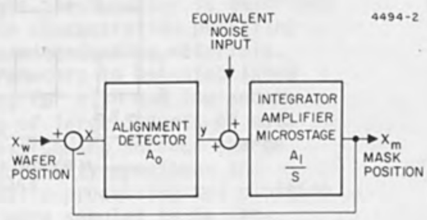


Fig. 2. Servo loop diagram of one axis control system for x-ray mask alignment.

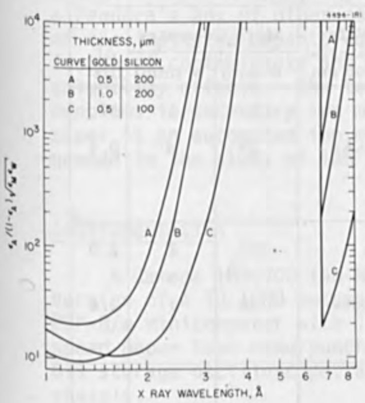


Fig. 3. Wavelength dependence of x-ray transmission factor for detection of complementary x-ray alignment marks.

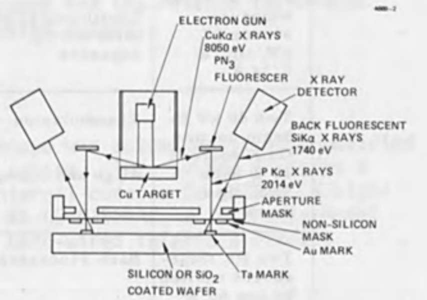


Fig. 4. Back fluorescent detection of alignment using P K x rays to excite Si K fluorescence through complementary absorbing marks.

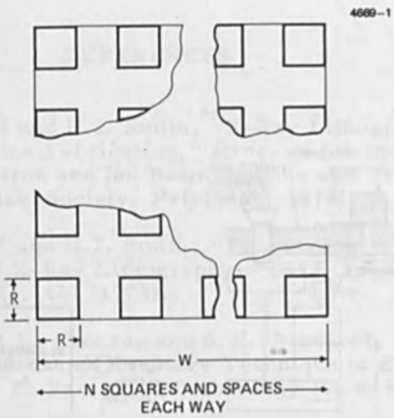


Fig. 5. Alignment mark geometry for which $A^{1/2}/L$ and acquisition ranges R are calculated in Table 1.

Table 1. Alignment Mark Requirements and Resulting Servo Bandwidths for 0.1 μm Accuracy of Alignment for Three X-Ray Detection Schemes.

X-Ray Source	Detection System	Wafer Material 200 μm	Mark Geometry R (μm)	Mark Geometry W (mm)	Δf Hz
Continuum from 20 kV Al exposure source 50 $\mu W/cm^2$ at 8.34 Å	Transmission	Si	25	4	50
	Continuous alignment during exposure	GaAs	25	4	0.4
Two 40 kV E-beam excited Cu K sources 10 W each 3 cm from mask	Transmission	Si	200	2	2.0
	Align and clamp	GaAs	40	4	2.0
Two PK fluorescers excited by one 65 W E-beam excited Cu K source at 40 kV, 3 cm	Back Fluorescent				
	Align and clamp Nonsilicon mask	Si or SiO_2 Coated	200	2	2.0

AUTOMATED SIMS CONCENTRATION PROFILING OF ION IMPLANTED DOPANTS IN SEMICONDUCTORS

Larry E. Plew
Electron Optics Laboratory
Naval Weapons Support Center
Crane, IN 47522

An automated Secondary Ion Mass Spectrometer is described which facilitates the in-depth concentration profiling of ion implanted dopants in semiconducting materials. Automation allows optimum parameters to be established for an analysis while allowing for high and low secondary ion yield, the recording of large amounts of data in finite times, as well as overcoming surface charge problems inherent in low-conductivity specimens and insulating thin films. Automatic processing and plotting of the collected data allows more samples to be profiled at less cost with a faster turn-around time for finalized reports.

Introduction

Accurate control of doping level, uniformity and reproducibility has led to the popularity of ion implantation. Implantation induced damage requires annealing at elevated temperatures. This opens a Pandora's Box of other associated problems such as: redistribution of the implanted dopant; outdiffusion of substrate elements; indiffusion of contaminants and encapsulating material elements; and solubility effects. The technique most suited to the study of these problems is secondary ion mass spectrometery. Reported in this paper is an automated ion microscope and the related techniques needed in the study of ion implanted dopants.

Instrumentation

A Cameca IMS-300 Ion Microscope was automated using a modified version of a TD-1000 automation system. This system includes a PDP-8/e minicomputer with 16K internal core, a DECWRITER, a high-speed paper tape read/punch, an RK-05 DISCPAK, a Tektronix Model 611 storage oscilloscope, and a hard-wired interface circuit chassis.

The Cameca IMS-300 Ion Microscope is a secondary ion mass spectrometer employing direct imaging mass spectrometer optics. Profiling consists of the continual removal of surface material by sputtering while performing a mass analysis of the sputtered charged particles, referred to as secondary ions. A primary ion beam of O^+_2 is normally used to enhance the secondary ion yield in comparison to the noble gasses. Concentration profiles as a function of depth are obtained by monitoring the secondary ion current versus sputtering time. A more detailed review of the instrument and the SIMS technique can be found in Ref. 1.

The minicomputer is interfaced to the SIMS through the hardwired interface circuit chassis which senses the magnetic field with a Hall cell. This hardwired interface circuit chassis acts as an intermediary between the computer, which can communicate only by digital signals, and the ion analyzer, which primarily uses analog signals. A number of elementary operations are then carried out through the utilization of programmable electronic circuits. The interface is connected to: (1) The computer by way of one input/output line carrying 98 digital signals for various codings, controls, numeric acquisitions, etc.; one line of 8 analog measurement channels going to the computer A/D converter and one line programming the X and Y deflection signal as well as the Z intensity signal for display control. (2) The analyzer by way of one magnetic field control and coupling line, one digital and analog ion current measurement line, and one mass number display line. (3) The oscilloscope by way of one memory display control line. The chassis' control board includes four sets of potentiometers to allow the pre-setting of up to four masses of interest, a digital switch to offset the voltage on samples retaining a charge, a dynamic display keyboard to manipulate the display on the oscilloscope and a function control keyboard to store, normalize and print data, stop the run and calibrate the ion analyzer, determine the process to be used or repeat a series of mass scans.

Data obtained during the various processes of analysis is displayed in the form of graphs or alphanumeric characters on a Tektronix 611 oscilloscope capable of working in either live or memory mode. At the start of and throughout PROCESSES 1 and 3 all data channels are displayed. The vertical scale is such that the maximum data point is always in the upper half of the Y-axis. The acquired cycle number and the scale of the y-axis are also displayed. At the end of the analysis, the x- and y-scales can be expanded or contracted, the curves can be translated, the display switched to memory mode, and in the case of PROCESS 3, other accumulation cycles can be added by actuating the REPT key.

Software

Software has been developed for data acquisition and display processes. These include PROCESS 0, which allows manual control with direct display of the mass number of the element being analyzed, PROCESSES 1 and 2 which allow the in-depth analysis of up to four elements by automatic peak switching with and without scanning through the peaks, and PROCESS 3, which allows programmed scanning across a mass spectrum with provision for accumulation.

A memory display program allows a more complete display of the results. For PROCESSES 1 and 2 the curve or curves determined by the operator on the DECWRITER appear as continuous traces on graphs with the amplitudes corresponding to each element analyzed for the first, middle and last channels displayed. For PROCESS 3 the spectrum is displayed as a continuous trace in front of which appear the values of the whole mass number and the amplitude. The results displayed in the three processes described above can be printed out on a DECWRITER and punched on paper tape in ASC II Code by actuating the PRINT key. These results can be standardized by actuating the NORM key and typing the mass in relation to which the standardization is computed.

Process 0 displays the mass number on L.E.D. elements. The data processing system receives values of the ion acceleration high-voltage and of the magnetic field of analysis. The computer calculates the mass by the relationship:

$$M/e = H^2/K^2V + \Delta M/e.$$

The term $\Delta M/e$ must be adjusted throughout the mass range due to possible deviations of the Hall cell from linearity. This adjustment is performed by linear interpolation between the calibration peaks (a maximum number of 32). A specimen is selected to provide a convenient spectrum spread of mass peaks in both positive and negative secondary ions. The calculation assumes that the mean energy of the ions corresponds to the applied acceleration voltage. In the case of insulating specimens, a correction is introduced manually with digital switches located on the interface circuit chassis. The calibration is obtained by setting the instrument successively to the various peaks provided in the calibration table, pressing the CALIB key, and then typing the mass value. In addition, the computer controls one of the galvanometers of the oscillographic recorder to serve as a mass marker during total manual mass spectra scans.

PROCESS 1 allows from one to four masses to be analyzed by selecting four magnetic field values with the four sets of potentiometers located on the front panel of the interface circuit chassis. There is no scanning about the mass value, so the process is suitable for examining conductive specimens with moderate mass resolution. During the analysis the dwell count time on each peak is selected as a multiple of 16 msec. The product of the number of cycles by the number of analyzed elements (≤ 4) is limited to a maximum of 512.

PROCESS 2 is preferred for materials having a high concentration of analyzed elements or having low conductivity and therefore apt to reach a surface potential differing by a few volts from the applied potential. Again from one to four whole masses can be selected by the operator. After the manual potential (ΔV) correction is set by the operator, each peak is scanned by means of a gradual variation of the reference potential of the magnetic field. The scan from one-half of a mass unit below through the mass peak to one-half of a mass unit above allows the total ion current detected to be integrated.

Another feature which has been incorporated into PROCESS 2 is the ability to compensate for a changing sample charge. This is done by establishing a reference peak, monitoring the shift of this peak due to the change in the magnitude of the surface charge, and then correcting the peaks to be analyzed by a computed correction. This is the only technique to properly monitor masses while sputtering through insulating films.

PROCESS 3 collects and displays the mass spectrum of a sample. The operator, by a dialog on the DECWRITER, can select a number of peaks or zones of the spectrum to be examined successively with a speed and the extent of scanning around each designed for the particular case. The number of isolated peaks or zones that can be programmed is such that three times the number of zones plus the total number of masses or peaks within these zones must not exceed 64. The masses chosen can be scanned once or repeatedly with the amplitudes of the various peaks accumulated channel by channel. The number of scanning cycles, mass resolution, masses involved as well as a scanning speed reduction factor of 1, 10 or 100 can be selected by the operator responding to questions appearing on the DECWRITER.

Once data from one of the previously mentioned processes has been obtained and punched on paper tape, it is necessary to present the data in more convenient forms such as graphs. A software system called IPLOT has been developed to allow the automatic reduction and plotting of data off-line on a separate system. IPLOT is stored in compiled form and can be executed via the Fortran Run Time System on the PDP-8/e minicomputer and interrogates the experimenter for additional information about the sample, operating parameters of the instrument and the types of graphs or plots desired. All graphs of the data are then plotted via a MOHAWK OS/2000 plotting program on a COMPLOT plotter, using data supplied from the paper tapes produced by the previously described processes, via a Honeywell 2070 Ion Analyzer Plot Program Version II TK1P09.

Discussion and Results

As in any technique there is a trade-off in the mass and in-depth resolution versus sensitivity. Thus in PROCESS 1 where the instrument is switched from peak to peak, maximum sensitivity is encountered. Figures 1(a) and 1(b) display the concentration profiles for two samples of gallium arsenide implanted with Be^+ at 40keV and a fluence of $1 \times 10^{19}/cm^2$. These samples are different only in the fact that Run No. 444 was annealed at 800C for 90 minutes while Run No. 446 was unannealed. In comparison Figures 2(a) and 2(b) display the concentration profile using PROCESS 2 which scans through the peak. These samples are silicon implanted with 200keV Al^+ and annealed at 700C for 30 minutes. The difference between these two samples is that Run No. 122 was implanted at a fluence of $2 \times 10^{15}/cm^2$ while Run No. 123 was implanted at a fluence of $1 \times 10^{16}/cm^2$. The curves in Figure 2 are smoother than in Figure 1 as faster sputtering rates were used and the peaks were scanned instead of being set for the counting interval. Thus integration under a scanned peak gives smoother statistics than for those which accumulate the count with a setting at the maximum peak intensity.

PROCESS 1 is desired over PROCESS 2 for elements implanted close to the surface, displaying low secondary ion yields or implanted at low fluences. PROCESS 2 is desired for elements implanted deeper into a substrate, displaying a high secondary ion yield, implanted at high fluences or whenever better counting statistics are desired. PROCESS 2 must be used whenever profiling through insulating films, interfaces or layers displaying varying conductivities. This optional technique of PROCESS 2 of monitoring the shift of a reference peak to compensate the scans of masses of interest has been developed but will be reported in a separate paper at a later time.

PROCESS 3 allows the masses to be scanned and displayed as a mass spectrum on the oscilloscope. Figure 3(a) and (b) show the mass spectrum of a gallium arsenide substrate. In Figure 3(a) the isotopic ratios for ^{69}Ga and ^{71}Ga are easily seen. Using the hardwired interface circuit chassis keyboard previously described allows the mass spectrum to be expanded as in Figure 3(b).

The real power of the technique is shown in Figure 4 which is a final plot from data processed through the IPLOT system into a graph ready for submission into a final report.

Conclusion

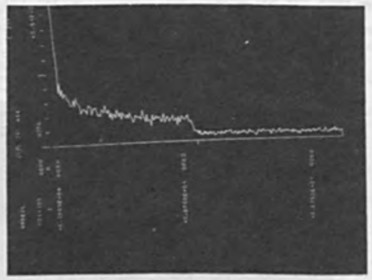
Results have been presented showing the power and flexibility in data collecting, processing and presentation using a SIMS. Due to the large capital expenditure for SIMS instrumentation it is absolutely necessary to utilize it to its maximum capability. As a result of high data output it is imperative that it be handled automatically unless 5 or 6 scientists are available for data interpretation and graphics. Even though automation may add from \$75K to \$100K to an initial investment of \$250K, it must be brought out that the savings in sample thoroughput and data analysis will pay for the cost of automation within 18 months.

Acknowledgements

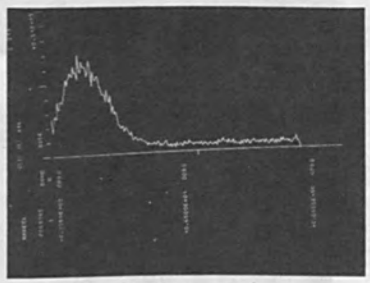
The author wishes to thank Dr. Jim Comas and Dr. Harry Dietrich of the Naval Research Lab at Washington, D. C. for supplying the ion implanted samples.

References

1. J. M. Morabito and R. K. Lewis, Anal. Chem. 45, 869 (1973)

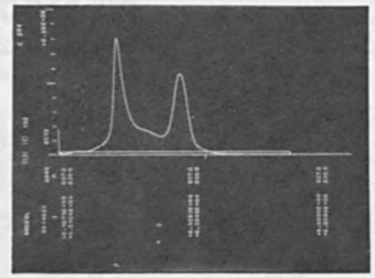


(a) Run No. 444

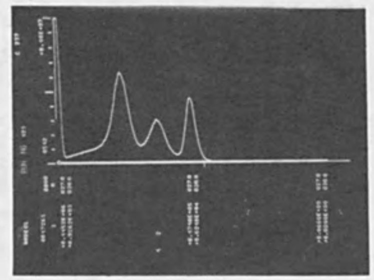


(b) Run No. 446

Figures 1(a) and (b). SIMS profiles of Be^+ implanted into gallium arsenide at 40keV and a fluence of $1 \times 10^{14}/cm^2$. Run No. 444 was annealed at 800°C for 90 minutes. Run No. 446 was unannealed. PROCESS 1 was used to obtain the profiles by setting on the mass 9 peak of Be for intervals of 4 seconds.

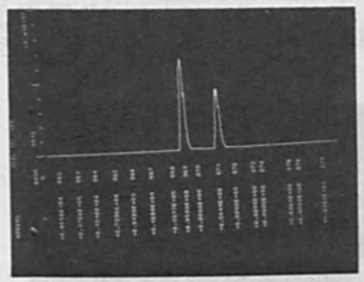


(a) Run No. 122

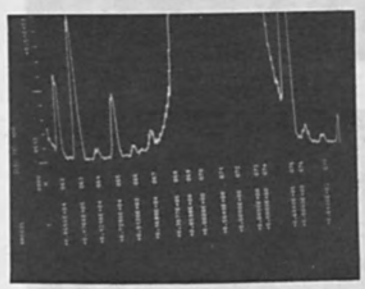


(b) Run No. 123

Figures 2(a) and (b). SIMS profiles of Al^+ implanted into silicon at 200keV and fluences $2 \times 10^{15}/cm^2$ and $1 \times 10^{16}/cm^2$, respectively. Both samples were annealed at 700°C for 30 minutes. PROCESS 2 of scanning through the mass 27 peak of Al and mass 30 peak of silicon was used.



(a)



(b)

Figures 3(a) and (b). Mass spectrum of a gallium arsenide sample obtained by PROCESS 3. The isotopic masses of ^{69}Ga and ^{71}Ga are seen in (a) while less intense masses are observed in (b) after expanding the scale with the hard-wired interface circuit chassis keyboard.

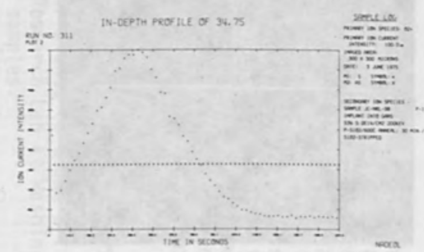


Figure 4. Final graph from the IPLOT program processes off-line and plotted with a COMPLOT PLOTTER. Run number 311 is an in-depth profile of ^{34}S implanted into gallium arsenide at 200keV and a fluence of 5.0E 14/cm^2.

EFFECTS OF HIGH DOSE NITROGEN IMPLANTATION IN SILICON

R.J. Dexter*

Texas Instruments Incorporated
Dallas, Texas 75222

ABSTRACT

Ion implantation and low temperature infrared resonance absorption has been used to establish the ionization energy of the nitrogen donor in silicon at 0.046 eV below the conduction band. In addition very high dose nitrogen implantations have been used to study the formation of buried layers of Si_xN_y compound structures. These nitride layers are shown to exhibit breakdown voltages of $7X10^5$ V/cm and a refractive index of 2.08 at 6328Å. Ellipsometry measurements gave a nitrite layer thickness neary twice that found by optical and electron microscopy and proton channeling measurements, indicating that a large fraction of the layer is made up of isolated clusters of nitride.

INTRODUCTION

Group V elements normally enter the silicon crystal lattice substitutionally as donor impurities. The electrical and optical properties of all these elements except nitrogen have been extensively studied and are well understood. In the past few years, several workers have used ion implantation to study the properties of nitrogen in silicon. Early work by Kleinfelder[1] suggested that the ionization energy of nitrogen was 0.142 eV below the conduction band but in later work Zorin, et. al.[2] and Roughan, et. al.[3] obtained values of 0.033 eV to 0.045 eV. Very recent work by Mitchell et. al.[4] gives 0.017 eV as the ionization energy. A common result in all of these papers with the exception of the work of Roughan et. al.[3] is that less than 5% of the total implanted nitrogen is on substitutional sites after annealing with the remainder occupying non-substitutional sites even for doses as low as $1X10^{14}$ N/cm^2.

Mitchell et. al.[5] has shown by a combination of backscattering and use of the $^{15}N(P, \alpha)^{12}C$ reaction that for doses of $1X10^{14}$ N/cm^2 essentiality complete annealing occurs at 700^oC except for a small surface disorder. However, for doses above $5X10^{15}$ N/cm^2 complete annealing does not occur even at 900^{o}C. These measurements suggest that nonsubstitional nitrogen is trapped as nitride precipitates or on damage clusters.

Work by Schwuttke et. al.[6], Borders et. al.[7], Dexter et. al.[8] and others have shown that for much higher dose implantations (>5X10^{16} N/cm^2) chemical complexing does take place. Dexter et. al.[8] reported that after high dose implantation and prolonged high temperature anneal a buried silicon nitride layer is formed below the surface and that the surface layer is sufficiently damage free to allow subsequent epitaxial silicon growth.

This paper will report an attempt at establishing the ionization energy of nitrogen by observation of electronic transitions of the nitrogen donor at low temperature by infrared resonance absorption and further detailed characterization of the buried silicon nitride layers.

THEORETICAL MODEL

In the slowing down of energetic nitrogen ions incident on a silicon crystal, the initial energy loss will be primarily due to electronic excitation process[9]. At greater depths, after losing sufficient energy, the energy loss in atomic collision processes increases while the electronic component decreases. Since radiation damage in silicon results from the atomic collisions, the defect density profile is peaked near the ion profile at the end of the ion path. If the nitrogen concentration in the peak region is high, a thin region will exist in which, after annealing, a stoichiometric change will take place resulting in isolated clusters of silicon nitride. Higher doses will result in higher cluster density and eventually a continum of silicon nitride. A small fraction of the nitrogen will find its way into substitutional sites and result in donor type electrical activity.

EXPERIMENTAL PROCEDURE

Infrared transmission measurements can be used to gain insight into a number of chemical and physical properties. Infrared radiation can both excite fundamental vibrational modes and produce electronic transitions between the discreet hydrogen-like energy levels in the forbidden band of silicon. At sufficiently low temperatures, donor electrons are found in significant numbers only in the ground state of the system from which transitions to excited states may be induced by infrared radiation. The resonances observed in the absorption spectra may then be used to determine the excited state energy levels relative to the ground state, and their transition probabilities. From this one can infer the position of the ground state or ionization energy relative to the conduction band. Figure 1 shows the energy level structures observed by Bichard and Giles[10] compared with those predicted theoretically by Kleiner[11].

Samples for the infrared absorption measurements were prepaired by cutting one to two millimeter thick silicon wafers from a 500 Ω cm P-type silicon crystal and mechanically polishing both surfaces. These wafers were then implanted on both sides by using the 3.5 MeV Van de Graaff at the Virginia Polytechnic Institute and were vacuum annealed at 1200°C for 0.5 to 6 hours and mounted in a liquid helium cryostat. The temperature was maintained at 10°K during the low temperature infrared measurements. A identical unimplanted wafer was used as a reference so the resulting spectra appear as relative to this wafer.

Samples for the silicon nitride studies were 0.020 in thick, n or p-type, 1 to 10Ω-cm silicon wafers and were polished on both surfaces. The implantations were made at 150 keV in a conventional ion implantation machine and annealed at 1200°C in dry N_2. Total implanted doses were in the range $1X10^{15}$ to $1X10^{18}$ N/cm^2.

RESULTS

IONIZATION ENERGY

The samples were typed by a hot probe and the sheet resistance measured after each implant and again after each anneal. The samples remained p-type immediately after implantation and the sheet resistivity was found to be extremely high. Conversion to n-type occured after annealing, however, the resistivity remained so high that measurements were not reliable.

Figure 2 shows the low temperature infrared transimssion spectra of a sample implanted with 3.5 MeV nitrogen with a dose of $1X10^{16}$ N/cm^2 on each surface and annealed at 1000°C for 30 minutes.

The $1S \rightarrow 2P^0$ transition is the lowest energy transition and further assignments are made on the basis of decreasing transition probabilities (ie. $1S \rightarrow 2P\pm$ etc.). These assignments are labeled $2P^0$, $2P\pm$, $3P^0$, $3P\pm$ in Fig. 2. Adding the energy relative to the conduction band given in Fig. 1 for each state observed to the observed transition energy to that state, the ionization or ground state energy of the nitrogen donor in silicon is calculated to be 0.046 eV below the conduction band.

SILICON NITRIDE FORMATION

Heavier doses at lower energies were used to stimulate the formation of silicon nitride layers. Fig. 3 shows a room temperature infrared spectra of a wafer implanted with $1X10^{17}$ N/cm^2 at 150 keV (A) before annealing, and for anneal times of (B) 1, and (C) 3 and 6 h. Before annealing a broad absorption band is observed between 700 and 900 cm^{-1}. This band shifts to higher wave numbers and sharpens into a complex spectra with increased annealing consistant with previous studies[7]. The strongest absorption occurs in a band at 845 cm^{-1} and has general features similar to those observed for silicon nitride layers formed by low temperature rf plasma techniques and are assumed to be associated with Si-N bond stretching frequencies. Since there is no change in the spectra after 6 and 6 h anneal, it is assumed that compound formation is complete after 3 h. The smaller band at 485 cm^{-1} is attributed to single-phonon absorption[12] in regions of damaged/strained silicon adjacent to and in the nitride layer.

A planar silicon etch which does not attack silicon nitride was used to remove part of the thin silicon surface layer and expose the buried nitride layer. Tallystep measurements then showed the nitride layer to be 0.2μm from the surface. Ellipsometry measurements using a laser source at 6328Å gave an average nitride layer thickness of 0.41 μm with an average refractive index of 2.05. These same measurements carried out on slices implanted with lower doses showed a corresponding thinner compound layer and a refractive index closer to that of silicon.

The isolation characteristics of these compound layers were studied by etching through the layers in a nitride etch (Bell II) to form 100 mil diameter mesas. Metal contacts were formed to the mesa and to the back of the wafer and the IV characteristics were measured. A maximum voltage of 30V could be applied before breakdown through the layer occurred indicating a breakdown field strength of $7X10^5$ V/cm compared to 10^6 V/cm for thermally deposited silicon nitride layers. Smaller doses showed very high leakage occurring at lower voltages indicating poor layer integraty probably due to the layer being silicon rich or the existance of "pinholes". For higher doses, an interesting phenomena occurred as illustrated by the photograph in Fig. 4. Small cratered regions about 10μm in diameter appear within which a small "explosion" seems to have taken place. Note the small skin of material that has lifted from the center of the crater. The assumption is that excess nitrogen diffuses to small voids or faults the silicon and these pockets burst during the anneal. For these doses, the I-V characteristics again indicated high leakage currents presumably due to the presence of the craters.

Fig. 5 is an SEM photograph of a structure formed by implanting through a SiO_2 layer which had been patterned and etched so that the patterned edges were sloped at about 45°. This pattern caused the implanted nitride layer to slope toward the $Si-SO_2$ interface at the same angle. After the implant, the SiO_2 was removed and the slice annealed. The wafer was then placed in a nitride etch for several minutes and cleaved through one of the implanted regions. The height of the void left by etching of the nitride was 0.22μm. Comparing this to the nitride layer thickness measured by ellipsometry indicates that the transition region from the nitride layer to the surrounding silicon is made up of small clusters of silicon nitride and nitrogen doped silicon.

In a previous publication[8] this author and co-workers used wafers containing these implanted nitride layers to grow 2-6μm thick epitaxial silicon layers. This work demonstrated that epitaxial layer quality was significantly better than that resulting from heteroepitaxy of Si on sapphire or spinal. The proton channeling and backscattering spectra in Fig. 6 shows the defect density in the epi-layer to be essentially indistinguishable from bulk silicon down to the depth of the buried nitride layer, increasing sharply through the implanted layer and returning again to that of bulk Si below the layer. This measurement also gave the implanted layer thickness to be ∿0.2μm. The dip in the non-channeled spectra at the location of the nitride layer is due to the fact that the average density of the nitride layer is somewhat less than that of bulk Si.

SUMMARY

This work shows the ionization energy of the nitrogen donor in Si to be 0.046 eV compared to values of 0.017 to 0.142 eV reported by other authors. This wide variation in results only indicates that the electrical behavior of this ion is not well understood. It is indeed clouded by the readiness of the nitrogen to form compount structures with the Si. Silicon nitride layers approximately 0.2μm wide were formed below the surface of the Si wafer with dielectric characteristics approaching that of thermally deposited Si_3N_4.

REFERENCES

* Part of this work was done at the Virginia Polytechnic Institute for the author's Ph.D. dissertation.

1. W.J. Kleinfelder,"Properties of Ion-Implanted Boron, Nitrogen and Phosphorus in Single-Crystal Silicon", Stanford Electron. Lab. Tech. Report #K701-1, Stanford, California (March, 1966).

2. E.I. Zorin, P.V. Pavlov, and D.I. Tetel'baum, Sov. Phys. Semicond. 2, (1968).

3. P.E. Rougham, A.H. Clark, J.D. Macdongall, K.E. Manchester, and F.W. Anderson, 133 rd Meeting Electrochem. Soc. Abstracts, Boston, p. 74 (1968).

4. J.B. Mitchell, J. Schewchun, D.A. Thompson, and J.A. Davies, J. Appl. Phys., 46, 335 (Jan. 1975).

5. J.B. Mitchell, J. Schewchun, D.A. Thompson, and J.A. Davies, J. Appl. Phys., 46, 332 (Jan. 1975).

6. G.H. Schwuttke, K. Brock, E.E. Gardner, and H.M. DeAngelis, in Radiation Effects in Semiconductors, edited by F.F. Vook, Plemium Pub. Corp. (New York, 1967) pp 407-417.

7. J.A. Borders and W. Beezhold, in Proceedings of the Second International Conference on Ion Implantation in Semiconductor, edited by I. Ruge and J. Groul (Springer Verlag, New York, 1971) p. 241.

8. R.J. Dexter, S.B. Watelski, and S.T. Picraux, App. Phys. Let. 23 455 (1973).

9. J. Linhard, M. Scharff, and H.E. Schiott, K. Dan. Videnski Selsk. Mat. - Fys. Medd. 33, No. 14 (1963).

10. J.W. Bichard and J.C. Giles, Cand. J. Phys. 40 , 1481 (1962).

11. W.H. Kleiner, Phys. Rev. 97, 1722 (1955).

12. M. Balkanski and W. Nazarewicz, J. Phys. Soc. Jap. Suppl. 18, 37 (1963).

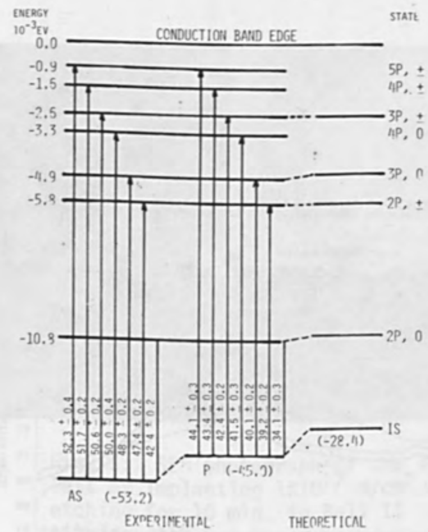


Fig. 1: Energy level diagram for arsenic and phosphorus in silicon. Experimental levels due to Bichard and Giles (10). Theoretical predictions due to Kleiner (11).

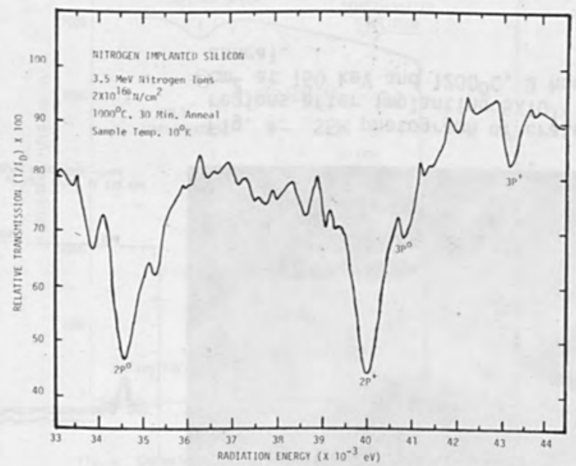


Fig. 2: Infrared resonance absorption spectra of electronic transitions from the ground state to excited states of the nitrogen donor in silicon.

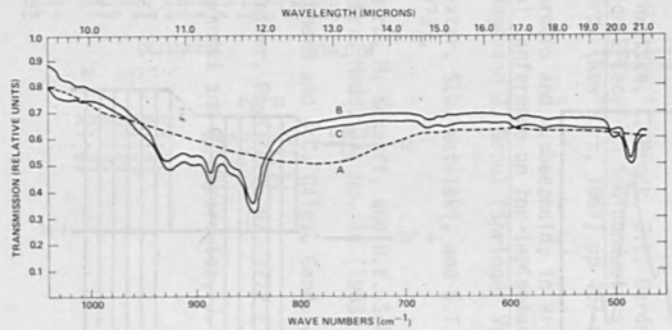


Fig. 3: Relative I-R transmission spectra of a silicon sample implanted to a fluence of 10^{17} N/cm^2 at 150 keV (A) before anneal, (B) after 1 h at 1200^{o}C in dry N_2, and (C) after 3 and 6 h at 1200^{o}C in dry N_2. From Dexter et al (8).

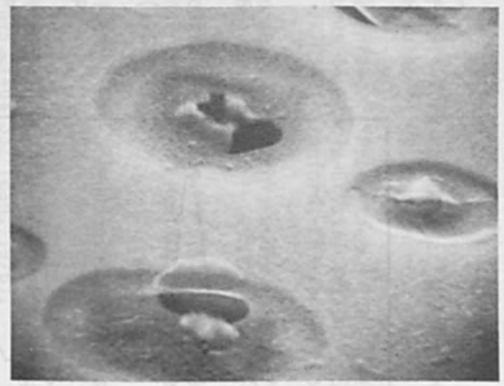


Fig. 4: SEM photograph of crater regions after implanting 5X10^{17} N/cm^2 at 150 keV and 1200^{o}C, 3 hour anneal.

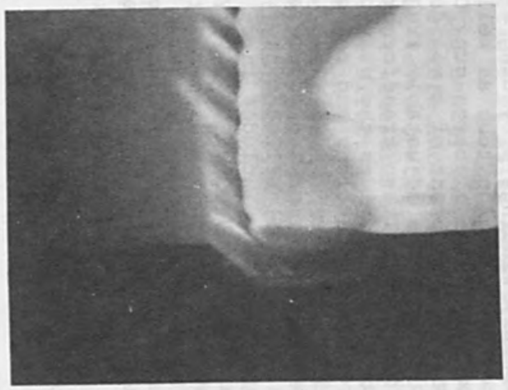


Fig. 5: SEM photograph of the void left by implanting 1X10^{17} N/cm^2 and etching for 10 min. in Bell II nitride etch.

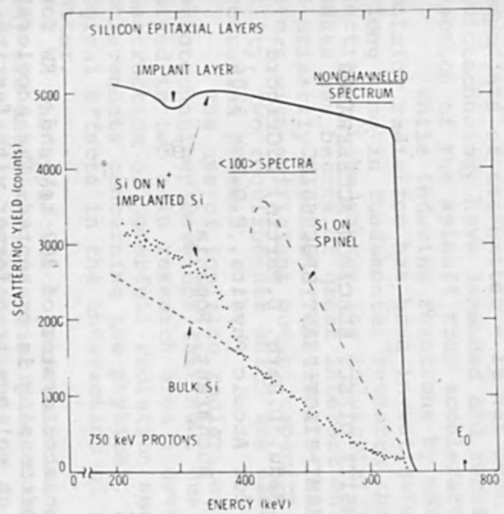


Fig. 6: Channeled <100> and nonchanneled backscattering spectra for 720-keV incident protons on epitaxial silicon with buried nitride layers are shown by dots and solid line, respectively. Portions of a <100> spectra for heteroepitaxial silicon layer on a spinel substrate (dot-dashed line) and for bulk silicon (dashed line) are also shown. All spectra are taken for 6-μC proton fluence. From Dexter et. al. (8).

HIGH INTENSITY DIRECT ELECTRON ACCELERATOR
FOR SEMI-PILOT-PLANT PROCESSING

G.Baciu, A.Radu, D.Barb, D.Martin, D.Grigore
Institute for Atomic Physics, P.O.Box 5206
Bucharest, Romania

Abstract

An electron accelerator of 300 keV and 3 KW for semi-pilot-plant processing is presented. The accelerator is powered by a high voltage transformer with ferrite core and multiple secondary windings providing a pulsed voltage to an evacuated acceleration tube as well as the power handling and control of the electron gun. Of particular importance is the development of a solide - state line - type modulator to generate adjustable frequency and voltage to power the high voltage transformer. One button digital control system provides the beam parameters preselection, protective functions and reliable operations.

The high voltage transformer and acceleration tube are contained in a metal oil filled tank.

The beam transport system includes three short magnetic lenses and a scanner subjected to a varying magnetic field causing the beam to sweep through an angle up to 60^{o}.

The conveyor system with a continuously variable velocity is correlated to the integral preselected dose.

Introduction

Although the stimulus for invention of accelerators have always, in the past, arisen from the man's curiosity about nature - about its fundamental laws, the origin of its forces and the properties of the basic components of the matter - it has been axiomatic that radiotechnology have demanded and received a substantial fraction of the spinoff from accelerator technology.

While inducing reactions by means of energy-rich ionizing radiation has been investigated intensively for a long time in fundamental research the exploitation of electron - beam energy for industrial irradiation processes has become of great interest only recently. This is primarily because of the lack of suitable beam generators with sufficient output capacity and high reliability, two requirements which are essential for the production process.

The use of ionizing radiations in the field of radiotechnology develops fast now, due to the progress gained in two main research area : development and construction of powerfull radiation sources and the achievements concerning the physical, chemical and biological effects in the interaction of radiation with matter.

The application possib ilities of radiotechnology are unlimited, in principle, although their practical use in an industrial scale is so far not spectacular because of still high price of any irradiation process.

The energy range of different radiation sources used in radiotechnology is comprised between some hundereds of keV and some MeV. The sources are either isotopic or accelerated particle beams.

The direct type electron accelerators could be acomplished in different technical solutions as for in stance electrostatic (Van de Graaff), cascade (Cockroft-Walton and Dynamitron), isolated core transformer, resonant transformer, etc.

It is to be underlined that the Van de Graaff accelerators, contrary to the others, are of high internal impedance which determines a strong correlation between the beam power and energy ; the need of bigger power imposes higher energy. In the other direct type electron accelerators the power is increased on the basis of higher intensity of the accelerated beam ; this is an important advantage allowing to produce higher power beams at lower energies and

Parameters and major components of the ADECA

The ADECA is a direct type electron accelerator delivering a pulsed electron beam with the following main parameters :

Energy	3oo keV
Intensity in the pulse	1 A
Length of the pulse	5o μs
Repetition frequency	2oo Hz
Average intensity of the beam	1o mA
Average power	3 kW
Radiation dose at 1o cm from the window .	1oo Mrad/sec
Maximum irradiated surface	5ox1oo cm^2

- The high voltage system.

There were developed two different pulser circuits

a) a solid state line-type modulator using DC link ;

b) a solid state frequency-changer system using a 8ooo Hz a single phase motor-generator set.

These circuits were used in order to establish the effects of the various circuit parameters on the shape of the output pulses of the high voltage transformer with high transforming ratio. Much attention has paid to the principal requirements for such systems including capability of delivering loads to any power factor and power conversion efficiency.

The prototype pulse modulator uses a SKT 45/16 thyristors string, operating as a A class inverter. The most critical characteristic is blocking voltage and dynamically, high values of di/dt. In order to minimize jitter, and the delay and rise time, the gate of SCR's are driven as hard as the ratings permit . Protection against forward break over and excessive reverse power is obtained by careful selection of the "snubber" circuitry.

The pulse forming circuit consisting from two network of equal impedance of 60 ohm connected in series operates at 50 μs.

The pulse forming network is nonresonant charged to maximum 7 KV from an controlled DC power supply.

Triggering of the discharge SCR's discharges the line with maximum 7 KV and 14o A into the high voltage transformer with ferrite core (the volum of the core is $V = 43.68o$ cm^3 and transformer set up ratio $n = 5o$).

Excellent high voltage pulse specifications have been obtained with the generating system using a single phase 8ooo Hz, 5 UKVA motor-generator set. In the frequency changer system, a single phase rectifier circuit is equipped with two sets of SCR's rings which would give the same output polarities in the high voltage transformer primary. Thus, by alternately triggering SCR rings at a frequency lower than the supply frequency a series of unidirectional pulses will flow in the high voltage transformer primary. The high voltage pulses across a resistive matched load represents damped oscillations. For this area of applications a rectangular high voltage pulse waveform is not necessary because it is possible to control the beam energy spread by controlling electron injector grid.

Better utilization of the unique properties of the thyristor switch can be gained by taking advantage of its phase control capability during the high voltage starting interval. The phase control capability of the SCR's rings provides a ready means for continuously controlling the high voltage and power. Used as protective devices, SCR's rings can clear faults in less one cycle.

In both cases, the power is switched "on" and "off" by the series thyristors so that special attention has been given to the method of gating [1,2] the generation of spurious frequencies [3] and to the important determining load and circuit power factors. Thus the solid state frequency changer adoptation involves the power factor-correction capacitors utilization.

- The control system.

The accelerator control system using solid state logic modules provides the necessary sequencing and interlock functions to give completely automatic fail-safe operation. Operator controls consist of an on-off key switch and re-set buttons only. The Figure 2 shows a block diagram of control system and basic accelerator circuits.

As shown in Fig.2, the high voltage transformer secondary windings are used to supply and control the electron injector equipment. A special secondary windings pair is provided for problem-solving in the fault control of the injector equipment parameters.

The thyristor devices, of modular construction, are used to obtain the control of the filament grid and anod injector supplies as well as the grid injector pulses amplitude.

Typical rectifying control system are provided to supply focus coil and scanner waveform generator. In each case the thyristor device is supplied single phase 50 Hz, 22o V. The modular concept offers distinct advantages. This allows the design optimization of the entire power conditioning system and ease maintenance. Each module contains not only the SCR's but also all of the comutation transient suppression circuitry required to ensure reliable SCR operation. An important consideration is the ciency of irradiation technology we have elaborated a work-program of research and new developments in the application of ionizing radiation to waste treatment including:

a) Research and development work for increasing the energy and power of the ADECA to a level adequate for desinfection of liquid waste.

b) Studies on physical, chemical, and biological effects of ionizing radiation related to liquid waste treatment.

c) Studies on secondary effects of radiation and other treatments, e.g. heart, ozone, chlorination, mixing the liquid manure during irradiation, etc.

d) Development of dosimetry methods for large volumes of solution irradiated by high density accelerator beams.

References

1. Dickey, D.L., Jones, M.C., Solid-state switching for Aircraft Electric Systems, Power Semiconductor Applications 2, IEEE Press, New York (1972) 58.
2. Dewey, G.C., Ellert, J.F., Lee, N.T., Titus, M.C., Development of experimental 2o-kV, 36 MW Solid-State Convertes for NVDC Systems, Power Semiconductor Applications 2, IEEE Press, New York (1972) 76.
3. Cauper, A.H.,Jr., Harnden, D.J., McQuarrie, M.A., Power Supply Aspects of Semiconductor Equipment, Power Semiconductor Applications 1, IEEE Press, New York (1972) 9.

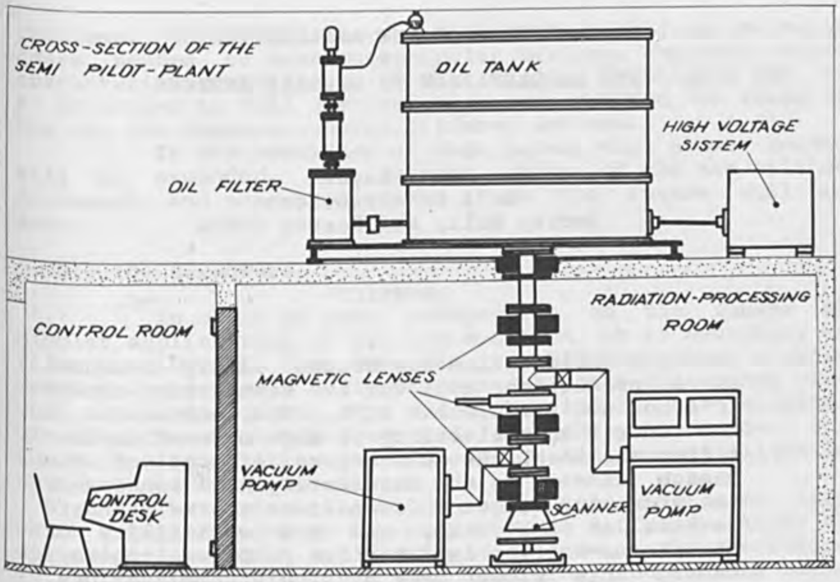


FIG. 1. CROSS-SECTION OF THE SEMI-PILOT-PLANT

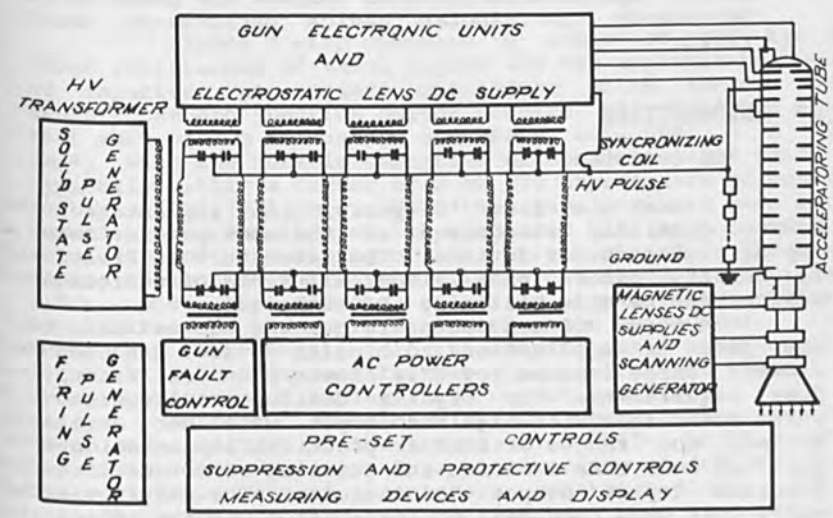


FIG. 2 ACCELERATOR UNITS AND CONTROLS

A REVIEW OF THE APPLICATION
OF
ION IMPLANTATION TO BIPOLAR DEVICES

R. S. Payne
Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

Ion implantation has become widely accepted as a production tool for the manufacture of MOS IC's but not for bipolar IC's. The combination of the general unavailability of high current implantation machines, the deleterious effects of the damage caused by the necessarily high ion doses, the short ion ranges in comparison to the standard devices of the time, and the variability of phosphorus-emitter implantation profiles together caused this slower pace of bipolar applications. Nevertheless, the problems have been overcome and from state-of-the-art devices to simple predepositions, implantation has been shown to make improvements in bipolar device performance and yield.

I. INTRODUCTION

A. An Overview

Over the last 10 years, ion implantation has become generally recognized as the most precise means of doping semiconductor devices. The literature is filled with feasibility proofs that almost all types of semiconductor devices can be made utilizing the technique.

While MOS manufacturers quickly adopted this tool, the pace of applications to bipolar devices has been much slower. Three reasons for this slower pace are first, the dose requirements for bipolar devices are large and high current machines with high throughput were not available. Second, the ranges of ions at practical implantation energies are less than 1 micron and existing common devices had junction depths deeper than 1 micron. The result was that early work was focussed on sub-micron depths where the variability of the phosphorus profile caused control problems. Third, the heavy damage, resulting from the heavy doses needed to fabricate bipolar devices, degraded device characteristics to the point that special techniques had to be developed to make devices which even matched the state of the art for chemical-source diffused devices.

In the remainder of this paper, the above points will be expanded, the present state of the art will be discussed, and some projections into the future will be made.

B. The Structure

In order to gain perspective on the course of bipolar applications of ion implantation, it is necessary to first consider the structure itself. Figure 1 shows a cross sectional view of a fully processed bipolar integrated circuit transistor of the type one might find in the millions of TTL IC's which have been sold over the past decade. For the sake of simplicity the dielectrics and the metallization have been left off the figure.

The device is built on a p type substrate into which a buried layer or sub-collector is diffused. After an epitaxial layer of 5 to 15 microns is grown, the structure is isolated with a heavy p type diffusion, and in some cases a collector contact diffusion is then done to lower the collector series resistance. Then, the base region is formed and finally the emitter is diffused to complete the structure.

Figure 1 also contains a table of the typical sheet resistances of these layers and the approximate number of impurites per square centimeter which these sheet resistances and junction depths require. It should be noted that most of the diffusions are heavy, i.e. over 1×10^{16} cm^{-2}, with the notable exception of the base layer which is typically within a factor of 2 of 200 ohms/square or equivalently 5×10^{14} cm^{-2}. Note that the intrinsic base layer formed by compensation of the extrinsic base by the emitter is 1/100 of the doping of the extrinsic base . The heavy doping of most of the layers will be discussed further (Sec. I.D.) when the effects of early implantation machine capabilities on the early research are considered. Also note that the layers are all deeper than 1 micron. This factor will be discussed further (Sec. I.C.) vis-a-vis the sub-micron range of implantation dopant profiles.

Further discussion will be concentrated on the base and emitter of the transistor. While bipolar IC's, especially linears, contain a wide variety of devices such as resistors, lateral transistors, reference diodes, JFETs, and others to which ion implantation has been successfully applied, this author has chosen to cover the active part of the vertical bipolar structure in a thorough way rather than do a less thorough job on the whole spectrum of applications.

C. Implantation Range Considerations

When one implants energetic ions into silicon the result is, to first order, a gaussian distribution of ions with a mean range Rp and a standard deviation delta Rp. The lighter the ion, the deeper the range will be. Figure 2 shows the range as a function of energy for B^{11} into silicon as presented by Seidel [7]. Note that for energies below 500 keV, which is a practical upper limit for ion implantation systems, the range is less than 1 micron. All other dopant species are heavier and have proportionately shorter ranges. This fact had a very important effect on the early work aimed at applying ion implantation to bipolars. Most of the early work focussed on microwave transistors where the structures are very shallow, and where it was believed that ion implantation could enable an advancement in the state of the art. The standard IC transistor which was shown earlier was largely ingored since it was much deeper than the practical range of the implanted ions and the promised advantages of improved parameter control, were tighter than required for high yield on existing circuit designs. Furthermore, the efforts were largely funded by government contracts with state of the art devices (in terms of high frequency performance) as the primary goals.

D. Hardware Considerations

Another factor which determined the directions of the first research was the beam current available on the machines of the day and its effect on the time neccessary to perform an implantation. Table I shows the implantation time as a function of the dose and the beam current. Remembering that most of the diffusions which made up the bipolar IC structure are in the 1×10^{16} cm^{-2} range, and noting that beam currents in the range of about 10 microamps were typical for early machines, it would require many hours or even days to dope a single wafer were all of the doping done by implantation. This machine limitation caused many to shy away from looking into bipolar applications and caused those who did undertake developments to look at the base of the transistor first. Similarly, for all implanted structures, the focus was on structures which could be achieved uniquely through the use of implantation, since hardware problems loomed ominously burdensome to any high volume applications.

II. The Early Work

The submicron range of ions at reasonable energies, the extended machine times required to fabricate bipolar devices, and the economies of contractural research all motivated the pioneering workers to look first at making microwave transistors. The early workers in the field [1-5,7,10,11,13-16] soon found that the problems associated with making bipolar devices were a good deal more complex than those associated with the MOS applications. The heavy implantations and the resultant damage associated with making bipolar transistors caused most to observe, as did Gibbons [3]: "The best results to date were with sample 3 where a small signal beta of 5 to 8 was obtained."

Assemat [11] reported: "These first transistors have relatively large leakage currents (~200ua) and widespread gains.....hfe: 2 to 8." Other investigators reported similar results and some investigators, myself included, reproduced those poor devices and did not report their efforts. The standard double diffused transistor was difficult to match by simply replacing the chemically diffused dopants with ion implanted dopants and annealing the structures to high enough temperatures to electrically activate those ions.

Importantly, the early workers found that if they implanted only the base or in some cases only the emitter, they could make devices which were as good as the double diffused devices to which they were compared. This suggested that the problem involved an interaction of the damage in the structure with the device performance.

It is interesting to note that all of the early investigators with the exception of Fujinuma, sought to fabricate npn devices with phosphorus emitters and boron bases since that combination was the industry standard for the conventionally fabricated devices. The results of Fujinuma were excellent but he used a chemical-source diffused-arsenic emitter.

The work of Tokuyama et al [15] represents the most comprehensive exploration of the phosphorus-boron combination. They looked at all combinations of implanted and diffused structures. Importantly, they realized that the damage caused by the emitter implant was a primary source of the poor results and studied, with some limited success, the utilization of high temperature implants, since it was known that a good deal of self annealing of the implantation damage occurs when the substrate is implanted while at an elevated temperature.

Table II contains a summary of their results. Their best devices were made with implanted emitters, formed after the base . It is important to notice that the hot substrate implanted emitters were much better than the cold subtrate implanted emitters. Their encouraging results for the diffused base-hot implanted emitter transistors and implanted base-hot implanted emitter transistors suggest very clearly that limiting the damaging effects of the heavy emitter implantation would be the key to successful all-implanted devices.

Figure 3 shows the results which they published for the dependance of the current gain on the annealing time after the emitter implant for the all implanted transistor. The data show clearly that annealing times beyond what is neccessary to merely activate the dopant ions were required to produce better devices. They noted that the time/temperature necessary to make good devices was sufficient to cause a non-negligible diffusion of the phosphorus emitter implant. In short, when the emitter was diffused out of its original implanted zone, better and better devices resulted. It would seem that if they had pursued the study beyond shallow devices, i.e. diffused both the base and the emitter more deeply after the implantations, they would have produced high quality devices. But, as with most others, they sought high frequency devices, not just all-implanted structures.

III. Successes

A. High Frequency Devices

Figure 4 shows a 100 keV phosphorus profile which has been redrawn based on the work of Moline and Reutlinger [9]. The details of the profile suggest why those who tried to use phosphorus implanted emitters met with generally poor results. Comparing the experimental profile (solid curve) with the first order gaussian prediction (dashed curve) one can see that the profile deviates markedly from the gaussian shape. Much of the dopant penetrates more deeply. Furthermore, Dearnley and coworkers [12], Moline [9] and for that matter all who have studied phosphorus implantations have found that the long tail is material dependant in both its slope and its magnitude. In short, the tail wags all over the place. This makes it very difficult to control structures which depend intimately on the implanted profile as does any shallow microwave transistor.

Re-examining Figure 1, note that typically 99% of the base dopant is compensated by the phosphorus emitter. The long variable tail of phosphorus implantations clearly make it an unfavorable choice for the emitter dopant in shallow, ion implanted bipolars.

Arsenic on the other hand has certain very desirable attributes. Because of its heavy mass, it renders the crystal amorphous at low doses (about 1 X 10^{14} cm^{-2}) and the tailing phenomenon saturates once the amorphous state is reached [6]. Because of its concentration dependant diffusion coefficient, it diffuses into an abrupt profile as seen in Figure 5. At Bell Labs [17,18,31] this highly favorable emitter profile, diffused out of the zone into which it was originally implanted, was combined with a base made up of two implantations, and the results were high quality, high performance devices.

It should be noted that for this structure the deeper of the base implantations, which determines the doping in the intrinsic base region is compensated less than 50% by the emitter in sharp contrast to the 99% compensation of the standard double diffused transistor. This scheme combines the elements of profile control plus dopant control to achieve highly reproducable results.

During the early 70's several groups [17,18,20,28,29,31,39] reported successful applications of ion implantation to bipolars of the high frequency type. The work at Bell Labs [17,18,31], Fairchild [22], HP [28] and Siemens [39] all utilized implanted arsenic for the emitter. The work at Bell Labs and Fairchild emphasized the controllability of this all implanted process, while HP and Siemens focussed on performance.

An example of the control available is contained in Figure 6 [17] where a percentile plot of the d.c. current gain for two different wafers illustrates a previously unheard of uniformity in gain over the comparable double diffused device. Note that the standard deviation of the current gain on these wafers was close to 1% as was the difference between their median values.

The results of Sigmon [28] represented similarly impressive characteristics for state-of-the-art devices. He reported tight gain control (±15%) and excellent yield (up to 85%) for devices as fast as any made conventionally (Ft up to 8 GHz and Fmax up to 25 GHz).

The promise of ion implantation in the field of bipolars became fulfilled. As long as the emitter is driven out of the damaged region, excellent devices can, and have been made. The abrupt arsenic emitter, when coupled with the double peaked base profile [17] allowed a very versatile transistor to be realized, where the transistor properties can be controlled with the knobs of an implantation machine. When combined with a single, shallow base implantation [28] state-of-the-art devices were realized. Also during this time period of 1971-1975 the successful inclusion of ion implanted transistors into integrated circuits were reported [19,25-27,37].

B. Novel Structures

In general the most successful way to apply new processing technology is in completely new approaches, i.e. to get in the door first. Two such new structural or new circuit approaches to IC's where ion implantation has been recently applied are in the rejuvinated triple diffusion process, a non-epitaxial approach to bipolar IC's recently discussed by both Bell Labs [32] and TRW [35,36], and in the development of high performance Integrated Injection Logic as discussed in papers from Bell Labs [38] and Fairchild [41].

The triple diffused structure is shown in cross section in Figure 7. The expensive epitaxial layer and the deep, space consuming isolation diffusion are replaced by a lightly doped n region. This region was difficult to make with chemical source diffusion since it required good dopant control at low (on the order of 1×10^{16} cm^{-3}) surface concentrations. For implantation, the control exists at all concentrations. The low surface concentration translated into a low dose requirement which made it an excellent candidate for implanting. The base region, similarly, is not overly burdensome on machine time and its utilization for circuit resistors makes it a good candidate for being implanted as well. Both Bell Labs [32] and TRW [35,36] used an implanted predeposition of phosphorus and a subsequent high-temperature drive-in to form the collector region. The base has been implanted by both as well with TRW using the same predeposition drive-in technique and Bell using the in-situ, double peaked base implantation similar to the one shown in Figure 4.

C. Predepositions

The base and emitter of the TRW transistor looks very much like the base and emitter of the standard epitaxial transistor. They have applied ion predepositions combined with thermal drive-in diffusion to achieve added control for deep layers. This kind of application is of the sort that is foreseen becoming more and more common, where an ion predeposited layer replaces a chemically predeposited layer. As reliable high current machines become the rule, more and more people will appreciate the ease of maintaining processes where the dopant control is maintained by ion implantation.

The results in this area have not been as dramatic as those for the all-implanted microwave transistors nor were they as straightforward as one might have imagined. Several authors [30,33-35,40] noticed poor results upon straight replacement of a chemical predeposition with an ion predeposition.

It is easy to appreciate why the replacement was not straightforward by looking at Figure 8 which shows the boundary between an implanted and a non-implanted region after a steam oxidation [35]. The wafer has been etched to depict the large number of stacking faults which result from the immediate reoxidation of the implanted silicon. It is well known that such large numbers of stacking faults can have a degrading effect on junction quality. However, when high-temperature neutral-ambient annealing is done prior to reoxidation, the defects do not form and high quality junctions can be made. The improvements that can be realized on one parameter, transistor gain, are shown in Figure 9.

As can be seen from these data, control of the sheet resistance of the base predeposition in deeply driven structures does indeed translate into improved control of current gain.

IV. Conclusion

Summing up the recent work, we've seen that order of magnitude improvements have been realized in the control of device parameters for high frequency discrete devices, that state of the art devices have been achieved, that problems associated with implanted predepositions have been understood and predepositions utilized in both novel and in old fashioned device structures with positive results. Ion implantation machines have come a long way in terms of reliability, beam current capability and throughput. That leaves the key question: What will the future bring?

Three areas look very promising for continued applications of ion implantation to bipolar devices.

In the first place, state of the art microwave transistors should rely heavily on implantation.

In the second place, as high current high throughput machines become commonplace, more and more chemical predepositions will be replaced by ion implanted predepositions because of the way it translates into lower costs.

Finally, as the second and later generations of devices are developed specifically for the new circuit technique of injection logic, it appears highly likely that implantation will be utilized to improve the performance and manufacturability of those devices.

REFERENCES

1. W. J. King, J. T. Burrill, S. Harrison, F. Martin, C. Kellett, IEEE IEDM, Washington, Talk 9.1 (Oct 65).

2. C. M. Kellett, W. J. King, F. A. Leith et al, AD635267 AFCRL-66-367 (May 66).

3. J. F. Gibbons, "International Conference on the a Application of Ion Beams to Semiconductor Technology", P. Glotin ed., (OPHRYS), (1967).

4. J. A. Kerr and L. N. Large, Ibid, also see SERL Tech. Jour., 7, 12.1 (1967).

5. W. J. Kleindfelder, AD665965 SU-SEL-67-015 (Mar 67).

6. J. M. Fairfield and B. L. Crowder, Trans. Met. Soc. AIME, 245, p. 469 (1969).

7. T. Ikeda, T. Tokuyama, and T. Tsuchimoto, ECS Fall Meeting 1969, Abstract 170 (1969).

8. T. E. Seidel, "II. International Conference on Ion Implantation in Semiconductors", I. Ruge and J. Graul eds., (Springer-Verlag) p. 351 (1970).

9. R. A. Moline and G. W. Reutlinger, Ibid, p. 58 (1970).

10. A. Morizot, A. Dubee, and A. Cornette, Ibid, p. 345 (1970).

11. J. L. Assemat, Ibid, p. 351 (1970).

12. G. Dearnaley, M. A. Wilkins and P. D. Goode, Ibid, p. 439 (1970).

13. K. Fujinuma, T. Sakamoto, T. Abe, K. Sato and Y. Ohmura, Proc. 1st Conf. on Solid State Devices, Tokoyo, 1969- Suppl. to Jour. of Japan. Soc, of Appl. Phys., 39, (1970).

14. J.A. Kerr, AD 882524, (Oct 70).

15. T. Tokuyama, T. Ikeda, T. Tsuchimoto, 4th Int'l. Cong. on Microelectronics, (Munich, Germany), (Nov 70).

16. M. Barnoski and D. Loper, AD 891790 AFAL-TR-366, (Nov 71).

17. R. S. Payne and R. J. Scavuzzo, IEEE IEDM, Washinton, Late News Talk 17.6 (Oct 71).

18. R. J. Scavuzzo, R. S. Payne, K. H. Olson, J. M. Nacci and R. A. Moline, IEEE IEDM, Washington, Talk 8.3 (Oct 72).

19. J. Agraz-Guerena and R. S. Payne, IEEE IEDM, Washington, Talk 8.6 (Oct 72).

20. M. Ono, S. Kibuchi and T. Tsuchimoto, IEEE IEDM, Talk 8.5 (Oct 72).

21. J. Archer, Elec. Let., 8 (Oct 72).

22. V. G. K. Reddi and A. Y. C. Yu, Solid State Tech., p.35 (Oct 72).

23. M. K. Barnoski and D. D. Loper, S. S. Elec., 16, p. 433 (1973).

24. M. K. Barnoski and D. D. Loper, S. S. Elec. 16, p. 441 (1973).

25. R. Edwards, R. S. Payne, C. F. Gibbon, M. L. Olmstead, R. A. Moline, D. L. Kushler, G. R. Weber and R. T. Kraetsch, "Semiconductor Silicon 1973", (1973).

26. S. Ueno, M. Nakatsuka, S. Saudo and T. Irie, NEC R and D, No. 30 (Jul 73).

27. W. J. Evans, A. R. Tretola, R. S. Payne, M. L. Olmstead and D. V. Speeney, IEEE JSSC, SC-8, No. 5 (Oct 73).

28. T. W. Sigmon, IEEE IEM, Washington, Talk 19.4 (Dec 73).

29. S. Ludvik and P. Froess, Elec. Let., 10, No. 4 (Feb 74).

30. S. Prussin, J. A. P., 45, No. 4 (Apr 74).

31. R. S. Payne, R. J. Scavuzzo, K. H. Olson, J. M. Nacci and R. A. Moline, IEEE Trans. E. D., ED-21, p. 273 (Apr 74).

32. P. T. Panousis and R. C. Prichett, IEEE IEDM, Washington, Talk 22.4 (Dec 74).

33. T. Koji, NEC R and D, No. 36, p. 37 (Jan 75).

34. K. H. Nicholas, R. A. Ford, P. J. Daniel, C. W. Sullivan, P. Sant, C. Bull and G. R. Booker, A. P. L., 26, No. 6 (Mar 75).

35. S. Prussin and A. M. Fern, J. Electrochem. Soc.: S. S. Sci. and Tech. (Jun 75).

36. J. Buie, Electronics (Aug 75).

37. J. Graul, H. Kaiser, W. J. Wilhelm and H. Ryssel, IEEE JSSC, SC-10, No. 4 (Aug 75).

38. F. W. Hewlett, IEEE JSSC, SC-10, No. 5 (Oct 75).

39. J. Graul, A. Glasl and H. Murmann, IEEE IEDM, Washington, Talk 20.2 (Dec 75).

40. T. E. Seidel, R. S. Payne, R. A. Moline, W. R. Costello, J. C. C. Tsai and K. R. Gardner, IEEE IEDM, Washington, Talk 25.5 (Dec 75).

41. R. E. Crippen, D. Obrien, K. Rallapalli and P. W. J. Verhofstadt, IEEE ISSC, Talk 9.4 (Feb 76).

IMPLANTATION TIME (3" WAFER)

	Beam Current			
Dose	1 mA	100 µA	10 µA	1 µA
10^{16} cm^{-2}	1.6 mins.	16 mins.	2.7 hrs.	1.1 days
10^{15} cm^{-2}	10 secs.	1.6 mins.	16 mins.	2.7 hrs.
10^{14} cm^{-2}	1 sec.	10 secs.	1.6 mins.	16 mins.
10^{13} cm^{-2}	-	1 sec.	10 secs.	1.6 mins
10^{12} cm^{-2}	-	-	1 sec.	10 secs.

TABLE I

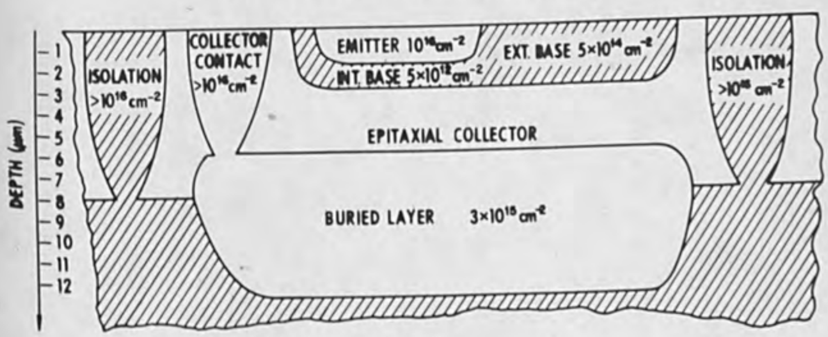
IMPLANTED/DIFFUSED TRANSISTOR PROPERTIES

Process	Implantation Temperature	h_{FE}	I_{CEO}	E-B Junction	Structure
Diff.B. + Diff.E.	-	100 - 200	Small	Good	Push Out
Diff.B. + Impl.E.	r.t. - 350°C 400 - 600°C	- 100 - 200	Very Large Small	Good Good	- Push Out
Impl.B. + Diff.E.	r.t. - 600°C	10	Large	No Good	Nearly Flat
Impl.B. + Impl.E.	r.t. - 200°C 400-600°C (E)	- 60 - 80	Very Large Small	No Good Fair	- Push Out
Diff.E. + Impl.B.	r.t. - 200°C 600°C	50 - 60 5	Slightly Large Large	Fair No Good	Flat Suck In
Impl.E. + Impl.B.	r.t. - 600°C	-	Very Large	No Good	Suck In

(REFERENCE 15)

TABLE II

STANDARD NPN IC TRANSISTOR



	RS	ϕ
ISOLATION	<10 Ω/□	$>10^{16}$ cm^{-2}
COLLECTOR CONTACT	<10 Ω/□	$>10^{16}$ cm^{-2}
BURIED LAYER	20Ω/□	3×10^{15} cm^{-2}
EXTRINSIC BASE	200 Ω/□	5×10^{14} cm^{-2}
INTRINSIC BASE	5k Ω/□	5×10^{12} cm^{-2}
EMITTER	10 Ω/□	10^{16} cm^{-2}

FIGURE 1

EXPERIMENTAL RANGE STATISTICS OF B IN Si
(REFERENCE 8)

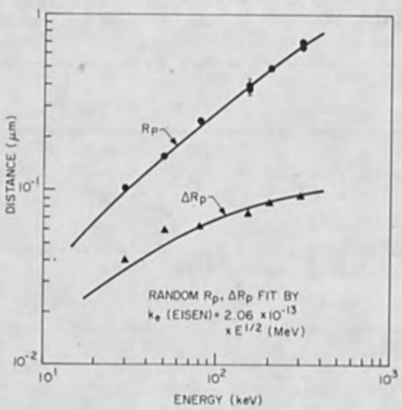


FIGURE 2

VARIATION OF h_{FE} WITH ANNEALING DURATION
IN DOUBLE IMPLANTED TRANSISTOR
(REFERENCE 15)

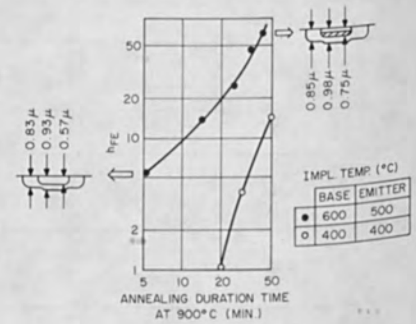


FIGURE 3

PHOSPHORUS PROFILE IN SILICON
(REFERENCE 9)

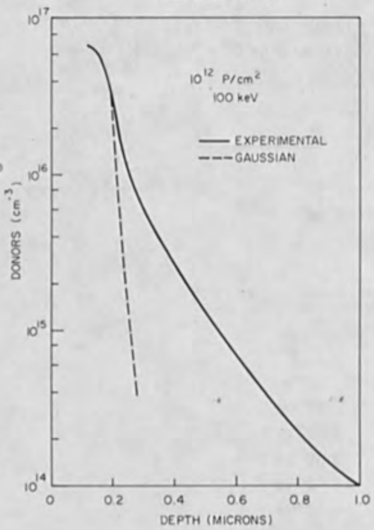


FIGURE 4

A VERSATILE TRANSISTOR
STRUCTURE
(REFERENCE 18,31)

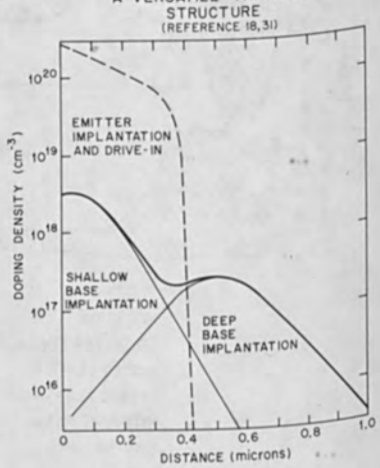


FIGURE 5

DISTRIBUTION OF h_{FE}
(REFERENCE 17, 18, 31)

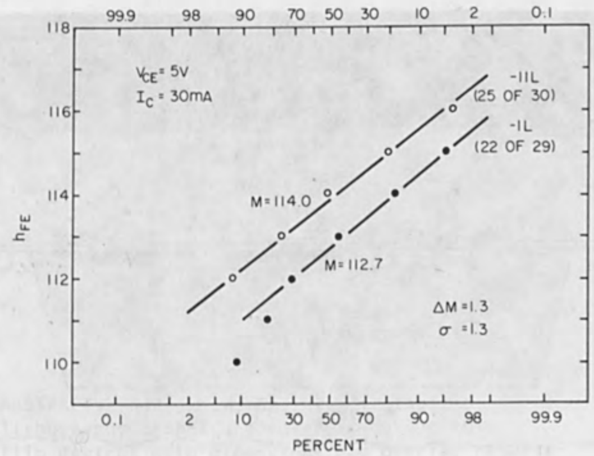
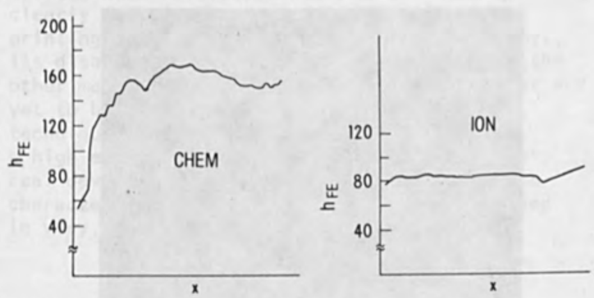


FIGURE 6

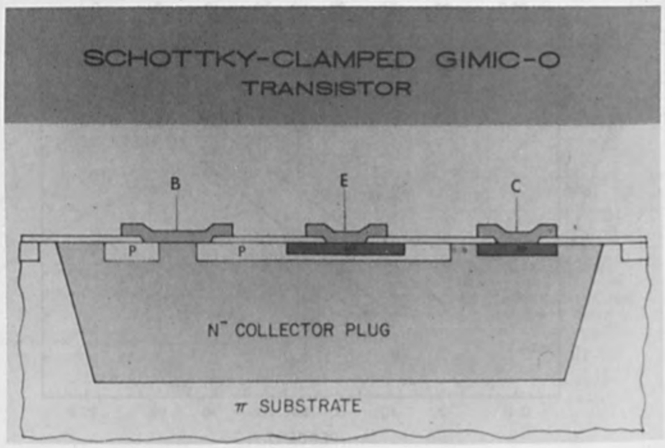
ION PREDEPOSITED BASES
h_{FE} UNIFORMITY



KOJI 1975 (33)

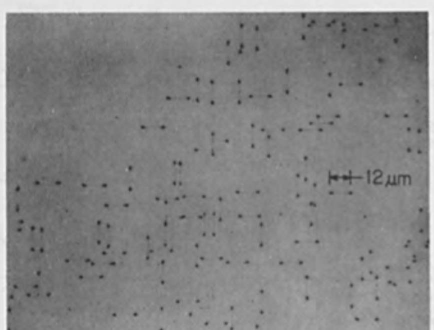
	CHEM h_{FE}	CHEM $\sigma(h_{FE})$	ION h_{FE}	ION $\sigma(h_{FE})$
PRUSSIN AND FERN 1975 (35)	27	± 26%	34	± 15%
SEIDEL et. al. 1975 (40)		± 19%		± 15%

FIGURE 9



(REFERENCE 32)

FIGURE 7



SPECIMEN A1, 10^{14} B11 cm^{-2} 30 Kev, (100) WET OXIDIZED AT 1100°c for 1h, SIRTL ETCHED 30 secs, MAGNIFICATION 405.5x

(REFERENCE 35)

FIGURE 8

RECENT LITHOGRAPHY TRENDS
A Critical Review of High Resolution Lithography Techniques

Alec N. Broers

IBM Thomas J. Watson Research Center, Yorktown Heights, NY

ABSTRACT: Advantages and disadvantages of lithography techniques for fabricating thin film devices with dimensions and overlay requirements beyond those of conventional UV contact/proximity printing are discussed. Only a brief description is given for each of the methods and the reader is referred to a comprehensive list of references for background information. The only method which has been used to fabricate devices with resolution and overlay which clearly exceeds those of conventional contact printing is scanning electron beam lithography. Its disadvantage at present is cost. Many of the other methods have shown higher resolution but are yet to be combined with adequate alignment techniques. Others have not been brought to a high enough state of development to fabricate real devices. X-ray lithography has the unique characteristic that image fidelity is preserved in very thick resist layers.

Miniaturization of electronic devices has proceeded rapidly because it offers advantages in speed and cost. Up until recently, these devices have been fabricated predominantly with ultraviolet ($\lambda \sim 3500Å$) contact printing. This technique is well established simple and low cost, but has the following limitations which restrict further progress.

1. The smallest structures that can be made are limited to lateral dimensions of about 2μ by diffraction effects which occur between mask and wafer. Diffraction also limits the accuracy with which one pattern can be aligned to another.

2. The percentage of good devices made is limited by damage which results from bringing the mask and wafer into close proximity.

Several methods are being explored to overcome these problems. They employ deep UV ($\lambda \sim 2000$Å), x-ray ($\lambda \sim 4$-45Å) and electron beams in addition to conventional UV radiation. This paper briefly describes these new approaches and discusses some of their advantages and disadvantages. Scanning electron beam lithography is treated in more detail than the other methods. Resists and processes are not discussed.

Deep UV Conformal Printing

The simplest methods for improving the resolution of the standard contact printing process are to reduce the wavelength of the exposing radiation and/or to reduce the spacing between mask and wafer. These approaches can be used independently to advantage,[1,2] but highest resolution is obtained when they are used together. 0.5μ linewidth structures have been made with the method shown in fig 1.[3,4] Deep UV radiation ($\lambda \sim 2000$Å) exposes the resist and the sample holder ensures intimate contact between mask and wafer. Resists originally developed for electron beam exposure are used instead of photoresists. The mask is aluminum or chromium on quartz and is made with a scanning electron beam system. The substrate can be thin in instances where the sample is rigid and it is necessary to conform the mask to the wafer, or thick when it is possible to conform the sample to the mask. A mercury arc is used as the illumination source, and exposure times of a few minutes for a 3" wafer can be obtained with suitable illumination optics.[4] The wavelength range used for exposure is determined at the upper end by the electron resist which is insensitive above about 2600Å, and at the lower end (2000Å) by lack of transmission through the quartz optical components and the mask substrate.

It will be difficult to apply conformal printing to semiconductor production because mask wafer damage remains a problem. There may be significant advantages to be gained, however, by using deep UV radiation and electron resist without the need to bring mask and wafer into contact. For example, it should be possible to obtain

1μ linewidths with a mask to wafer spacing of 8μ.[4] For acoustic wave devices, integrated optics, and some single layer magnetic bubble devices where the benefits of simplicity, high resolution and favorable resist profiles are important, it may be advantageous to tolerate a moderate yield loss and use both aspects of the process.

Alignment of one level to another has not been demonstrated for micron and submicron devices. To do this it will probably be necessary to follow new basic approaches because it will be difficult to reach the required accuracy with conventional alignment methods. This problem and the fact that silicon wafers distort both laterally and vertically when they undergo hot processes (such as recessed oxide growth) are additional difficulties preventing the immediate application of this technique to integrated circuit production. They also apply to x-ray proximity printing.

X-ray Proximity Printing

In x-ray proximity printing[5-12] (figure 2) the wave length of the exposing radiation is reduced to 3Å-45Å and diffraction effects become negligible, at least down to linewidths of 0.25μ. The mask can no longer be made on a quartz plate because such a plate would absorb the soft x-rays. Instead it is formed on a thin (2-10μ) membrane such as silicon or mylar. The pattern is defined in a layer of gold or similar high atomic weight material which heavily absorbs the soft x-rays. Although basic feasibility of the method for unaligned exposures has been established much work remains to be done to find the best combination of substrate material, substrate thickness, absorber material, absorber material thickness, and the wavelength of the x-ray radiation. Defect levels remain relatively high in many instances because the processes used to fabricate masks are still relatively new, but it should be no more difficult to overcome this problem than it has been to reduce defect levels in semiconductor processes.

The most remarkable feature of x-ray lithography is that the pattern definition is preserved deep into the resist. This is clearly demonstrated by the 1μ linewidths obtained in a 15μ thick resist layer, reported in reference 11. The resist is believed to be exposed (polymerized or depolymerized) by photoelectrons produced by the x-ray photons and not by the x-ray photons themselves. X-ray photons are not scattered in the way electrons are scattered, and the photoelectrons have little range in the resist because their maximum energy is only equal to the incoming photon energy (280-1500eV). The exposed pattern, therefore, does not become blurred by scattering as is the case with electron exposure, or by diffraction as in UV exposure, even for very thick resist layers.

In the limit the range of the photoelectrons determines the edge sharpness. This is about 600Å for rhodium radiation (λ =4.6Å), 400Å when aluminum characteristic x-rays are used (λ=8.3Å) and 50Å for carbon (λ =44.8Å). For high resolution it is obviously better to use softer radiation, but this requires a very thin mask substrate which is difficult to handle and may not be dimensionally stable. Shorter wavelengths and thicker substrates will probably be chosen for semiconductor applications because the minimum linewidths under consideration are still relatively large (1μ). Another advantage with x-ray exposure is that dust particles on the mask do not create as serious a problem as they do with UV masks because most particles are transparent to the x-rays.

Exposure speed has been a problem with x-ray lithography, however, recent improvements in source output and resist sensitivity promise that it should be possible to make it economically competitive with conventional UV techniques.

Methods for aligning mask and wafer are being investigated in several laboratories.[7,13,14] They utilize UV illumination, transmitted x-rays, characteristic x-rays, or photoelectrons to detect sample position. All are feasible from a theoretical point of view given adequate time. To date, sufficient accuracy for micron devices has not been demonstrated, nor has it been clearly shown that alignment can be achieved without a significant increase in the cost of the process. Overlay errors arising from lateral distortion of the sample or from lack of sample flatness also have to be investiaged before a clear committment to use x-ray lithography for semiconductor devices can be made. To overcome these errors it would be necessary to use a step and repeat approach which would seriously increase the cost of the process. As mentioned above for deep UV printing the decision is easier for single layer processes.

The ultimate source of x-rays from the point of view of throughput is the electron storage ring.[15,16] The radiation is so intense that it has been estimated that the lithography requirements for the largest predictable industry requirement could be met by a fraction of the output from a single optimized storage ring. Another advantage of synchrotron radiation is that it is highly collimated (divergence 10^{-3} radian). This means that alignment errors which arise when the substrate is not flat are no longer important as they are for a conventional point source placed close to the sample. The major problems with the storage ring are high cost and 'overcapacity' which will make it difficult to use efficiently.

Optical Projection

Figures 3 and 4 show two forms of optical projection camera that have been used successfully to fabricate semiconductor devices. Their primary advantage over UV proximity printers is that the mask and wafer are completely out of contact and the potential for damage is eliminated.

In the case of reduction projection (fig. 3) the image is considerably smaller than the sample amd exposure has to be made in a step and repeat manner. A lens that could project, for example, a 3" diameter image with adequate resolution for 1 micron linewidths cannot be built because it would be too large and the required tolerances would be too difficult to meet. For 1 micron dimensions the maximum field diameter that has been achieved with adequate contrast (MTF=60%) is 1.4cm. The depth of field is very small ($\sim 1\mu$), however, which will make refocusing for every chip essential. A pattern to wafer alignment is generally required for each exposure although in the future laser interferometry may be used to keep track of sample position after an initial single alignment. The mask has larger dimensions than the final pattern and is therefore relatively easy to make and to keep clean. The resolution can be better than for UV contact printing and it should be possible to fabricate 1μ linewidth devices if the resist process is kept under adequate control.[18]

With the scanning projection system shown in fig. 4 the complete sample is exposed after a single alignment by simultaneously scanning the mask and sample past the optical system. Only a crescent of the mask is illuminated at one instant and this portion of the pattern is transferred to the sample with a parity that allows mask and wafer to be scanned in the same direction and to be mounted on a single holder. A complete description of this system is given in reference 19. The magnification of the mirror optical system is 1X so the mask has the same dimensions as the pattern. The resolution at this time is adequate to reproduce 2μ linewidths, or slightly better than conventional UV proximity printing.

Although the scanning projection system has poorer resolution than reduction projection systems, its cost to throughput ratio is generally better because only one alignment is required for each sample and the time taken to achieve adequate alignment accuracy is a predominant cost factor. Such tradeoffs can only be made with validity, however, for a given type of device with a given set of dimensions and tolerances.

Electron Beam Projection

Electron beam projection systems have been developed to take advantage of both the higher resolution of electron optics, and for repetitive memory devices to avoid the high cost per exposure of today's scanning electron beam systems. Cost is potentially lower than scanning systems for the following reasons:

A. Pattern information is stored in a mask and not in the form of digital data which must be fed serially to a scanning electron beam system for each exposure. The cost of the electronic equipment necessary to supply these data is thus avoided. The cost of the mask is negligible provided a sufficient number of identical devices is required.

B. The field that can be covered in a single exposure is larger than for scanning systems and the exposure time per unit area is considerably shorter.

The two forms of E/B projection systems that have been developed specifically for semiconductor device fabrication are discussed here.[20-25] They are shown in figs. 5 and 6. Other forms of projection systems have also been developed, for example, for the fabrication of arrays of field emitters, or x-ray deffraction gratings.[26-32]

Figure 5 shows the reduction type of projection system. This system which is the electron optical analogy of reduction optical projection cameras is described in reference 20. It is particularly unique as a precision electron optical system because the beam diameter is very large (approximately 3cm at the mask), which requires unusually large electron lenses, and because the electrons pass at large angles to the system axis. Because of the latter its design necessitated the derivation of aberrations previously unavailable because they were not needed to theoretically predict the performance of electron microscopes. The mask is a freely suspended metal foil and its fabrication for general pattern shapes is the major problem to be overcome before this type of system can be successfully applied to device production. In particular the difficulty of supporting isolated opaque areas is of concern and two papers in this proceedings address this topic.[35,36]

The system has a scanning mode of operation which is used for alignment. In this mode the illuminating beam is focused on to the mask rather than flooding it as in the case of image projection. The focused beam is scanned across the mask and because the operating condition of the projection lenses remains unchanged, an image of this focused beam scans across the sample. Scattered electrons are collected from the sample to detect sample position and correction is made by shifting the projected image with deflection coils placed between the two projection lenses. The accuracy of alignment can in principle be the same as for scanning systems, however, the field to be covered is much larger and therefore to obtain the same resolution the beam angle must be smaller, thus reducing beam current. S/N of the deflection signals also has to be better because of the larger field. Despite these restrictions it should be possible to obtain an accuracy of 0.1μ.

Field size is larger for reduction projection systems than for scanning systems mainly because the beam half angle can be much smaller in the projection case where millions of points are exposed simultaneously and the current density in the image does not have to be as high for a given exposure rate. The system described in reference 20 is electron optically capable of projecting a 3mm diameter image with a theoretical total aberration disk of 700Å and a current density of 10^{-4} A/cm^2. The optimum beam half angle is approximately 10^{-4} radian. Patterns with 0.5μ linewidth have been successfully exposed over a 3mm field with an exposure of 0.1 sec.

1:1 electron projection systems have been under development for many years.[21-25] They employ a photocathode which is masked with a thin metal pattern. A uniform magnetic field focuses photoelectrons from the cathode onto the sample with unity magnification. Sample and wafer can, in principle, be as large as desired and 5cm diameter samples have been successfully exposed. In the system shown in fig. 6 image position is detected by collecting characteristic x-rays from marks on the sample.[25] The photocathode is masked during this process so that only alignment marks are illuminated. Magnetic deflection is then used to position the pattern with an accuracy of 0.1μ. Image current density is about 10^{-5} A/cm^2 (1 second exposure for 10^{-5} C/cm^2 resist sensitivity) for cesium iodide photocathodes which have the best overall lifetime and poisoning characteristics. The dominant aberration limiting resolution is chromatic aberration and theoretical estimates of minimum linewidth vary from 0.5μ to 1μ. So far, however, the technique has only been developed to the extent that operating semiconductor devices have been fabricated with linewidths of 2μ.

Because the sample forms part of the imaging system its flatness is very important in determining pattern distortion and for silicon wafers an electrostatic chuck may be essential for satisfactory performance. Another difficulty arises because scattered electrons from the sample are accelerated back onto the sample into unwanted locations giving rise to a background exposure which reduces the effective contrast of the image.[33,34]

Proximity effects which arise because electrons are scattered in the resist layer and back scattered from the sample are difficult to compensate with any full pattern electron beam projection systems. Correction has to be made by preadjusting the shapes in the mask because it is not possible to adjust the electron charge per unit area from point to point in the image as can be done with scanning E/B systems.

Scanning Electron Beam Systems

The most direct method for producing submicron structures is scanning electron beam lithography. In this instance the pattern is written with a small electron beam which is generally controlled (deflected and turned on and off) by a computer. It is the only method used so far to successfully make silicon devices with a capability that exceeds conventional contact printing in terms of linewidth and overlay tolerance. Masks made by scanning electron beam are also needed if high resolution is to be obtained from deep UV conformal printing, x-ray lithography and 1:1 electron beam projection.

Conventional masks can be made with scanning electron beam but the true potential for scanning electron beam cannot be realized in this way because the overall resolution and tolerance becomes limited by the replication process. Another key advantage with scanning E/B, elimination of the mask, is also lost. Scanning E/B allows direct software control of the pattern and therefore provides the designer with the ability to make changes without the need to fabricate a new mask set.

The major problem with scanning E/B for large scale integrated circuit production is cost and it is only recently that it has been economic even to fabricate masks and specialized devices with this method. However, many new methods are available for reducing system cost and increasing throughput, and it is possible with some optimism to predict that the cost will eventually approach that of conventional UV methods.

Another significant problem which affects electron beam methods is proximity effect.[37] As already mentioned this is a blurring of the image which occurs because electrons are scattered in the resist layer and also back from the underlying sample. The scattered electrons partially expose the resist up to several microns from the point of impact of the beam. In practice corrections for proximity effect which take into account the local variations in pattern density have to be made when dimensions reach one micron, particularly when it is necessary to maintain a steep profile in the developed resist pattern. For dimensions of 0.25μ, proximity effects become so predominant that dense patterns can only be reproduced in very thin resist layers and steep resist profiles are only possible for isolated structures. The only way to produce dense structures with dimensions of 0.1μ or below is to use thin substrates. In this case devices requiring bulk substrates (most silicon devices) cannot be fabricated directly. Instead a mask will have to be made with electron beam on a thin substrate,[39,40] and the mask then reproduced with x-ray lithography. Alignment will be a severe problem in this case. Very careful control of resist development can be a help in compensating for proximity effects.[38]

There are two basic methods in which the beam writes the pattern in scanning electron beam systems; raster and vector. Combinations of the two have also been used and in one instance mechanical scanning is also employed. With raster scanning the beam repetitively scans the entire pattern area and the beam is turned on where required. In vector pattern generation the beam is directed only to points where exposure is required. Vector scanning is more efficient because no time is wasted scanning areas not requiring exposure, however, provided adequate beam current is available raster scanning places less stringent requirements upon the deflection system in terms of dynamic accuracy and in particular eddy current errors. Raster scanning also makes it easier to apply corrections for deflection aberrations and pattern distortion. From the point of view of data volume the vector scan mode is generally much more efficient because the pattern can be optimally divided into shapes and each shape can be described by a few data words. For the raster mode data have to be provided each time the beam is turned on or off and although line to line correlation can be used it is usually not possible to achieve the degree of compaction obtained with vector scanning.

Computer aided design of lens deflection - yoke combinations has recently produced great advances in the performance of the electron optical columns used for scanning electron beam fabrication systems.[42,45] Two particular configurations, double deflection before the lens, and deflection in the center of the lens, have yielded the highest performance. An example of the performance that can be obtained with double deflection before the lens is given in this proceedings.[46] A 0.25μ diameter beam is deflected over a 4mm x 4mm field with a growth of 0.3μ without dynamic focus or astigmatism correction and 0.1μ with dynamic corrections. The beam half angle is $\leq 10^{-2}$ radian which for example, allows a beam current of 1.5×10^{-7} ampere for an electron gun brightness of 10^{6} A/cm^{2}-steradian (20KV). When operated in a vector scan system this beam could write a typical 1μ minimum linewidth integrated circuit pattern [field size=2mm or 2000 x the minimum linewidth, 20% of the area requiring exposure, 10^{-5} C/cm^{2} resist sensitivity] in 0.5 seconds provided the computer control system could direct the beam with sufficient speed and accuracy. This is not the case at present because of other factors in the computer control system. An example of deflection in the lens is given in the discussion of shaped beam systems which follows.

In cases where the current available from a round beam electron optical column is not adequate, large improvement can be gained by using a shaped beam.[44,45,47,48] Shaped beam columns use a physical aperture as the source rather than the electron gun cross-over as is used in most round beam systems. The lenses which produce the demagnified image of the source are similar in both cases and the resolution and field size are determined by the characteristics of the final lens and deflection yoke. In the case of shaped beam the aberrations give rise to a finite edge slope to the spot. With the round beam the aberration disk is combined with the Gaussian image of the source to give the overall beam diameter. The aperture, which can have any shape which corresponds to a basic repeat element in the pattern, is illuminated by a conventional electron gun generally using Köhler illumination. The total current drawn from the cathode in most electron guns is much greater than the final beam current and it is possible to evenly illuminate relatively large beam systems without sacrificing brightness (see footnote). This means that to a first order approximation the current density in the shaped beam is the same as for a round beam system using the same electron gun, final lens and deflection yoke, and the current gain is therefore directly proportional to the beam area. A square beam column with the deflection yoke inside the final lens has been described which operates in this manner and which focuses $3\mu A$ into a $2.5\mu \times 2.5\mu$ spot. Edge slope (10% to 90%) on the spot is 0.5 and the beam can be deflected over a field 5 mm x 5 mm.

A problem with shaped beam systems is that it is not possible to generate pattern shapes such as angle lines and circles without double exposing parts of the pattern. For silicon devices such shapes are not generally essential, however, so it should be possible to take advantage of the gain in electron current to increase throughput. An additional advantage with shaped beams is that the beam address increment is greater than that of a round spot system, and where throughput is limited by this aspect of data transmission, a gain in throughput can be realized.

The method used for alignment is perhaps the most significant factor which determines the overall configuration of a scanning electron beam system. Two basic approaches have been taken. In the first

Footnote: This is not the case when the beam current becomes a significant fraction of the total current drawn from the cathode, as it does with field emission cathodes, or where the emission pattern from the cathode is uneven as it is with LaB_6 and field emission cathodes. The current available without loss of brightness and consequently uniform illumination (for Köhler illumination - critical illumination alleviates the problem to a limited extent) is very approximately as follows: tungsten filament gun $20\mu A$, LaB_6 cathode gun $5\mu A$ (increases to $20\mu A$ for single crystal perfectly polished cathodes) and 10^{-7} ampere for cold field emission cathodes.

method a direct beam to sample reference measurement is made before each complete section (chip) of the pattern is exposed, and exposure is made with a single two dimensional scan. The reference measurement checks for position, rotation, orthogonality, and calibrates scan amplitude. For relatively high throughput systems, where writing time is a few seconds per chip, this measurement is made every few seconds. The system shown in fig. 7 operates in this manner as do several other systems which have been developed using the same basic principle.[49-54] In the second method laser interferometry is relied upon to maintain positional reference between the sample and the beam after an initial, or occasional, direct beam to sample reference measurement. The reference measurement again checks for errors in position, rotation, scan amplitude, and orthogonality. The systems shown in figs 8 and 9 both rely on laser interferometry. Combinations of the two methods are obviously possible and for mask making in particular laser interferometry may be utilized to advantages in the first type of system.

In the EBES system shown in fig 8 the beam is only scanned in one direction and continuous mechanical movement is used for the other direction.[55-57] Chips are written strip by strip, the same strip on every chip being written before proceeding to the next strip. Errors in sample position are detected using laser interferometry and corrections made by deflecting the electron beam.

In the system shown in fig 9 the electron beam is scanned in 2 dimensions over a small area.[58,59] The small areas are then stitched together to build up chips. An initial reference measurement is made for each chip and then the composite pattern is stitched together using laser interferometry as the reference.

A common advantage of the systems shown in figs 8 and 9 is that the performance of the overall electron beam deflection system is not as critical as it is for the vectorscan type of system where a relatively large area has to be scanned in two dimensions. Both electron optical and noise-bandwidth criteria are important in this respect. The disadvantage with the present EBES concept is that errors arising because of process induced distortions of the sample, or due to beam drift either because of electronic instabilities or due to electrostatic charging fields along the beam path are not compensated for from point to point on the sample. To compensate for such errors it would be necessary to detect alignment marks during the continuous mechanical movement writing mode. The marks would be distributed over the sample surface as they are for chip scanning systems. Errors due to lack of wafer flatness are not important with the EBES system because the deflection system is telecentric.[57]

Although there does not seem to be any fundamental reasons why the two basic alignment methods should not produce comparable overlay accuracies, provided the target is dimensionally stable, at present the highest accuracies have been obtained with the systems in which a reference measurement is made between the beam and the sample before each chip pattern is written.

Throughput for scanning electron beam systems is determined by electron optical performance, resist sensitivity, noise-bandwidth characteristics of the deflection system, digital to analogue conversion rates, data transfer rates, switching speeds for digital pattern generation hardware, alignment time, mechanical movement of the sample, and time taken to load samples into vacuum. The relative importance of these different factors depends on the overall system configuration. Assessment of system capital and operating cost is similarly complex. It is therefore not possible from a cost-performance point of view to compare individual features, for example raster and vector, round beam and shaped beam, laser reference and direct reference, in an absolute manner. Instead it is necessary to consider the overall system. An approximate assessment for today's systems can be made by considering data recently presented for two complete systems. The EBES system produces approximately 2 4"x4" masks in one hour with minimum linewidths of 2 $1/2\mu$ and an overlay accuracy of $\pm 0.5\mu$. A resist sensitivity of 10^{-6} C/cm^2 is required. The Vectorscan system described in this proceeding has demonstrated the ability to expose a chip 2000 times the minimum linewidth in size in less than 2 seconds. This corresponds to about 5 minutes per wafer for a 3" diameter wafer with 2 $1/2\mu$ linewidth, $\pm 0.5\mu$ overlay accuracy and 150 chips per wafer. A resist sensitivity of about 2×10^{-5} C/cm^2 is required. For smaller linewidth the overlay accuracy of the Vectorscan system is $\pm 0.1\mu$ and for large complexity chips patterns can be stitched together.[60]

The performance of both types of system is already attractive for many applications and in both cases there are several areas in which it is possible to predict considerable improvement.

REFERENCES

1. H. I. Smith, N. Efremow, and P.L. Kelly, J. Electrochem. Soc., 121, (1503 (1974)

2. H. I. Smith, Rev. Sci. Instr. 40, 729 (1969)

3. B.J. Lin, J. Vac. Sci. & Technol. 12, 1317 (1975)

4. B.J. Lin, IBM J. Res. & Devel. 20, 221 (1976)

5. D.L. Spears and H.I. Smith, Electron. Lett., 8, 102 (1972)

6. D.L. Spears and H.I. Smith, Solid State Technol., 15, 21, (1972)

7. H.I. Smith, D.L. Spears, and S.E. Bernacki, J. Vac. Sci. Technol. 10, 913 (1973)

8. J. H. McCoy and P.A. Sullivan, Proc. 6th Internat. Conf. on Electron and Ion Beam Sci. & Technol., R. Bakish, Ed. (Electrochemical Soc., Princeton, N.J.) p. 3 (1974)

9. R.Feder, E.Spiller, and J. Topalian, J.Vac. Sci. Technol. 12, 1332 (1975)

10. D.Maydan, G.A.Coquin, J.R. Maldonado, S.Somekh, D.Y. Lou, and G.N. Taylor, IEEE Trans.Electron Devices ED-22, 429 (1975)

11. E. Spiller, R. Feder, J.Topalian, E.Castellani, L.Romankiw, M.Heritage, Solid State Tech. 19, 62 (1976)

12. E.Spiller, R.Feder, J.Topalian, W.Gudat and D.Eastman, This Proceedings

13. D.Maydan, G.A.Coquin, J.R.Maldonado, S.Somekh and G.N. Taylor, Device Res.Conf., Salt Lake City (1976)

14. J.H. McCoy and P.A. Sullivan, This Proceedings

15. E.Spiller, D.E Eastman, R.Feder, W.D.Grobman, W.Gudat, and J. Topalian, This Proceedings

16. S.Doniach, I.Lindau, W.E.Spicer and H.Winick, J. Vac.Sci. Technol. 12, 1123 (1975)

17. J.Wilczynski and R.Tibbets, IBM Thomas J.Watson Research Center, Yorktown Heights, N.Y. Unpublished. See also J.Wilczynski and R.E.Tibbets, IBM J.Res.Develop 13, 192 (1969)

18. F.H.Dill, A.R.Neureuther, J.A.Tuttle and E.J.Walker, IEEE Trans. on Electron Devices ED-22, 456 (1975)

19. D.A.Markle, Solid State Technol., p.50 (June 1974)

20. M.B.Heritage, J.Vac.Sci.Technol. 12, 1135 (1975)

21. T.W.O'Keefe, J.Vine, and R.M.Handy, Solid State Electron, 12, 841 (1969)

22. P.R.Malmberg, T.W.O.'Keefe, and M.M.Sopira J.Vac.Sci. Technol. 10, 1025 (1973)

23. J.P.Scott, Proc. 6th Int.Conf. on Electron and Ion Beam Sci. and Technol., R. Bakish Ed. (Electrochemical Soc., Princeton, N.J.) p. 123 (1974)

24. B.Fay, Ref. 23, p. 527

25. J.P.Scott, IEEE Trans. on Electron Devices ED-22, p. 409 (1975)

26. H.Koops, G.Mollenstedt and R.Speidel, Optik 28, 518 (1968)

27. H.Koops, Optik 29, 119 (1969)

28. H. Koops, Optik 36, 93 (1972)

29. H. Koops, J. Vac.Sci.Technol 10, 913 (1973)

30. L.N.Heynick, E.R.Westerberg, C.C.Harteliun, and R.E.Lee, IEEE Trans. on Electron Devices ED-22 p. 399 (1975)

31. S.P.Newberry, T.H.Klotz, Jr., and E.C.Buschmann, Proc. Nat.Electronics Conf. 23 (1967)

32. H.G.Parks and W.C.Hughes, J.Vac.Sci.Technol. 12, p.1161 (1975)

33. G.A.Wardly J.Vac.Sci.Technol. 12, 1313 (1975) and J.P.Scott same reference p. 1309

34. G.A.Wardly IEEE Trans. on Electron Devices ED-22, p.414 (1975)

35. M.B.Heritage, P.E.Stuckert, and V.DiMilia, This Proceedings

36. A.Meyer, A.Politycki, and E.Fuchs, This Proceedings

37. T.H.P.Chang, J.Vac.Sc . & Technol., 12, 1271 (1975)

38. M. Hatzakis, J.Vac.Sci. & Technol., 12, 1276 (1975)

39. T.O.Sedwick, A.N.Broers, B.J.Agule, J. of Electrochem.Soc. 119, 1769 (1972)

40. A.N.Broers, W.Molzen, J.Cuomo, and N.Wittels, Appl.Phys. Lett. To be published

41. A.D.Wilson, A.Kern, A.J.Speth, A.M.Patlach, P.R.Jaskar T.W.Studwell, and W.L.Keller, "Automation of Vector-scan I Electron Beam Lithographic System". This proceedings.

42. E.Munro, J.Vac. Sci. & Technol. 12, 1146 (1975)

43. E.Munro, Optik 39, 450 (1974)

44. H.C.Pfeiffer, 8th Internat. Congress on Electron Microscopy, Canberra, Australia, p. 56 (1974)

45. H.C.Pfeiffer, J.Vac.Sci.Tech. 12, 1170 (1975)

46. T.P.Chang, A.J.Speth, C.H.Ting, R. Viswanathan, M. Parikh, and E.Munro, "The Probe Forming and Deflection System for Vector-scan I E/B Fabrication System". This proceedings

47. H.C.Pfeiffer and K.H.Loeffler, 7th Internat.Cong. on Electron Microscopy. Grenoble p. 63 (1970)

48. J. Trotel, Oral presentation. This conference.

49. T.H.P.Chang, A.D.Wilson, A.J.Speth and C.H.Ting, "Vector-scan I An Automated Electron Beam System For High Resolution Lithography". This proceedings.

50. G.L.Varnell, D.F.Spicer, A.C.Rodger, and R.D.Holland, Proc. 6th Internat.Electron & Ion Beam Sci. & Technol., R.Bakish, Ed. (Electrochem.Soc., Princeton, N.J.) p. 97 (1974)

51. N.C.Yew, Ref. 50, p. 111 (1974)

52. J.P.Beasley and D.G.Squire, IEEE Trans. on Electron Devices, ED-22, p. 377 (1975)

53. J.Pasiecznik, J.Frey, J.Vac.Sci.Technol. 10, 1012 (1973)

54. E.S.Ozdemir, W.E.Perkins, R.Yim and E.D.Wolf, Vac. Sci. Technol. 10, 1008 (1973)

55. D.R.Herriott, R.J.Collier, D.S.Alles and J.W.Stafford, IEEE Trans. on Electron Devices ED-22, p. 385 (1975)

56. R.F.W.Pease, J.P.Ballantyre, R.C.Henderson, A.M.Voshchenkov, and L.D.Yan, Ref. 55, p. 393 (1975)

57. D.S.Alles, F.R.Ashley, R.J.Collier, E.A.Geve, D.R.Herriott, A.M.Johnson and M.G.R.Thomson. Late News Paper. Internat. Electron Devices Meeting, Washington D.C. (1975)

58. O.Cahen and J.Trotel, Proc. 4th Int. Conf. on Electron & Ion Beam Sci. & Technol., R.Bakish Ed. (Electrochemical Soc. Princeton, N.J.) p.581 (1970)

59. O.Cahen, R.Sigille, J.Trotel, Proc. 5th Int.Conf. on Electron & Ion Beam Sci. & Technol., R.Bakish Ed. (Electrochemical Soc., Princeton N.J.) p. 92 (1972).

60. A.J.Speth, A.D.Wilson, A.Kern and T.H.P.Chang, J.Vac. Sci.Technol., 12, 1235 (1975)

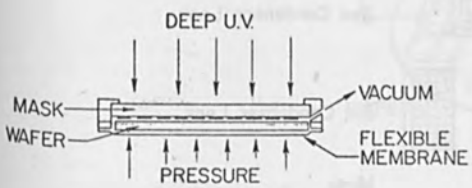


Fig. 1. Deep UV conformal lithography. References 1-4

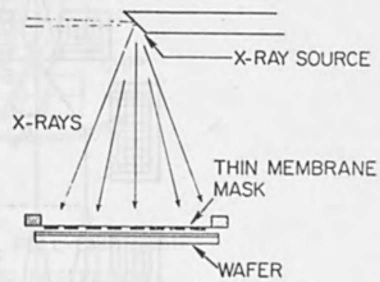


Fig. 2. X-ray lithography. Refs. 5-14. High output sources have rotating anodes.

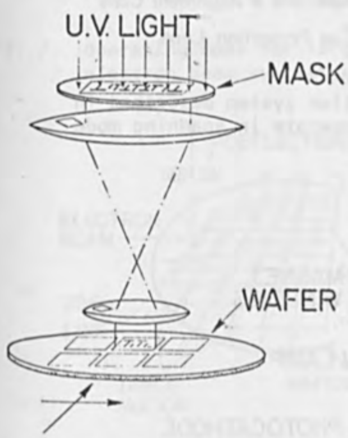


Fig. 3. Step and repeat optical projection concept. Multiple exposures are needed to cover a full wafer.

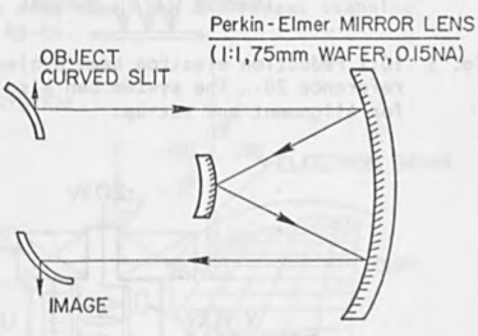


Fig. 4. 1X scanning optical projection camera. Actual camera has a roof prism before the sample so that mask and sample can be scanned in the same direction (ref. 19).

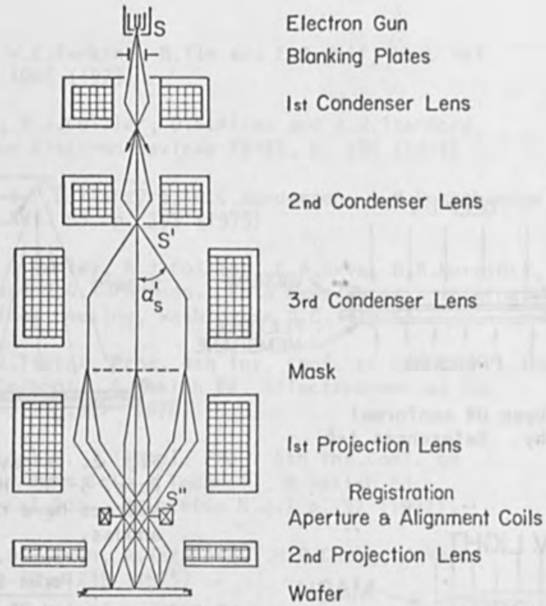


Fig. 5 10:1 reduction electron beam projection system described in reference 20. The system can also operate in scanning mode for alignment and set-up.

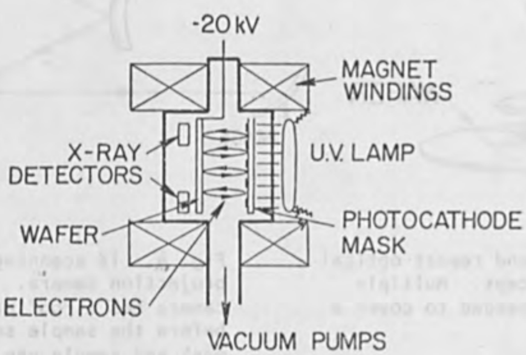


Fig. 6. 1:1 photocathode electron beam projection systems of type described in references 21-25. In the system shown alignment marks are detected by collecting characteristic x-rays from the sample.[23]

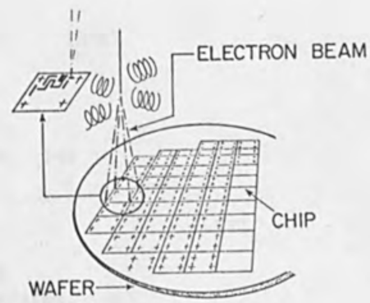


Fig. 7. Operating mode for large area scan step and repeat scanning electron beam systems. 49-54.

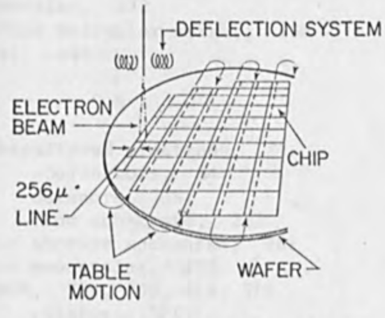


Fig. 8. Operating mode for EBES electron beam system 55-57.

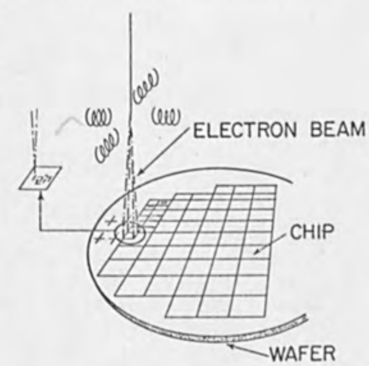


Fig. 9. Operating mode similar to that used in system described in references 58 and 59.

INDEX

A

B

C

D

E

J

K

L

M

S

T

U

V

W