HEWLETT·PACKARD JOURNAL

Contents:

MAY 1981 Volume 32 • Number 5

In this Issue:

Shown on our cover this month is the exposure end of HP's electron beam lithography system, the subject of this issue. On the left are the column that produces the electron beam (see page 14), and below it, the X-Y stage that holds the target (see page 16). On the right is the system that transports the target objects to and from the X-Y stage (page 16).

Electron beam lithography is one of the technologies on the frontier of the business of producing integrated circuits, those tiny silicon chips that can hold enough electronic circuitry for an entire computer—HP's latest chip has 450,000 transistors on it. On page 5, Frank Ura, manager of HP's electron beam research and development team, introduces us to electron beam lithography and its advantages and disadvantages. Briefly, its main advantage is that it promises to let us put more transistors in a given area of chip and do it more accurately, so a higher percentage of good chips comes off the production line. Its disadvantage has been that it is slower than older methods, too slow for mass-producing any but the simplest chips.

The scientists at Hewlett-Packard Laboratories, HP's central research facility, thought they saw a way to make a faster electron beam system that would be useful in a production environment, and HP management supported them. That investment has paid off in the system described in this issue, which is much faster than any other system we know of. To make it work, the project team applied good, sometimes clever engineering and advanced technology, as this issue attests.

The new system will be used by HP divisions to make better HP products. There are no present plans to make it available to other companies. The electron beam team has already used the system to make some advanced devices and to develop new ways of dealing with an electron beam problem, proximity effect, or the random scattering of electrons when they hit a target. We didn't have room in this issue to talk about these topics, but we're planning articles on them for future issues.

-R. P. Dolan

A Precision High-Speed Electron Beam Lithography System

This very fast electron beam system is designed for mask making or direct writing on wafers in an integrated circuit production environment.

by John C. Eidson, Wayne C. Haase, and Ronald K. Scudder

THE INITIAL OBJECTIVE of HP's electron beam project was to build a precision high-speed direct writing electron beam lithography system that would be usable in a production environment. Later, a requirement was added that the system function as a mask maker. The system architecture and specifications were determined based on the results of experiments on field emission guns and numerous calculations and experiments on column design. The most significant of these results was that it appeared likely that, using a field emission source, a column could be built that would be capable of delivering a beam of up to 600 nanoamperes into a 0.5-micrometre spot, which could be deflected over a 5-mm-square field of view with minimal distortion.

The decision was made to design a machine with the highest possible throughput and accuracy, useful for geometry in the 0.5-to-1-μm range. The high-current beam would permit a bit-writing rate of at least 300 MHz, which was felt to be the limit that could be reached by the control circuitry. This bit rate and the desire for high accuracy and throughput led to the design of a modified raster-scan system. In this system, a mechanical X-Y stage is used to move from one 5×5-mm field of view to the next. This simplifies the fiducial mark location needed in a direct writing machine and keeps the time lost in stage motion to a reasonable fraction of the total exposure time.

The 5×5-mm field of view and 0.5-μm beam led to the adoption of an ideal pattern grid of 0.5-μm spacing or 10^8 pixels per field of view. The placement accuracy goal for individual pixels was set at ±0.1 μm, which implies a deflection accuracy of 0.1 μm in 5000 μm or roughly 16 bits. Since it was not deemed possible to design deflection circuitry that could meet both the 300-MHz bit rate and the 16-bit accuracy requirements simultaneously, the modified raster-scan scheme was adopted. The field is divided into blocks of 128×128 pixels. Each block is addressed by a high-accuracy octopole driven by circuits that can settle in about 10 μs. This is a reasonably small fraction of the time required to expose a block at 300 MHz. Within each block a raster is scanned using high-speed circuitry driving a quadrupole deflector.

Each 5×5-mm field of view is composed of 80×80 blocks of pattern data, each with 128×128 bits (pixels). During the exposure of each block this data must be delivered at the 300-MHz bit rate. To handle this data the electron beam machine has a pattern memory that can store the entire bit pattern for a 5×5-mm field of view and can deliver this data at the full data rate.

A further complication was introduced by the requirement that the machine also function as a mask maker capable of producing 10× reticle masks. This increases the number of data bits to about 10^{10}. Since the time required to load the pattern memory is long compared to the time to expose a field of view, a rather simple data compression scheme has been implemented that permits the data for the entire 10× reticle to be contained in the pattern memory.

To position each block accurately using the octopole, numerous calculations must be done to determine the voltages required by the octopole plates. The calculations must be done in real time on a block-by-block basis. Since the main system computer, an HP 1000, is far too slow for this task, a special computer called HAL was designed for this purpose.

The pattern memory, the deflection circuitry, and HAL are the electronic subsystems that are most specific to the HP electron beam system because of its field of view and data rate. The other electronic subsystems are more conventional, with stability and system compatibility being the main design constraints.

Pattern Generation and Format

Pattern information for this system is generated on one of several standard computer-aided artwork (CAA) systems. This artwork is converted by means of special routines to a standard format called electron beam external format. This conversion is accomplished by an off-line computer.

After the data is converted into the electron beam external format, it is converted by an off-line computer system to one of the two electron beam internal formats, compressed or uncompressed, and stored on disc until exposure time. For 10× masks the compressed format is always used. To expose the substrate with this data, the disc is placed on one of the system disc drives and the data is then transferred to the pattern memory. From here it is transferred to the beam blanker in synchronism with the sweep circuits.

System Architecture

Fig. 1 is a block diagram of the system. The system consists of an electron column, a target area with a cassette loading system and an X-Y stage, a computer control system, and a variety of special-purpose electronic equipment used to drive the beam and control the data.
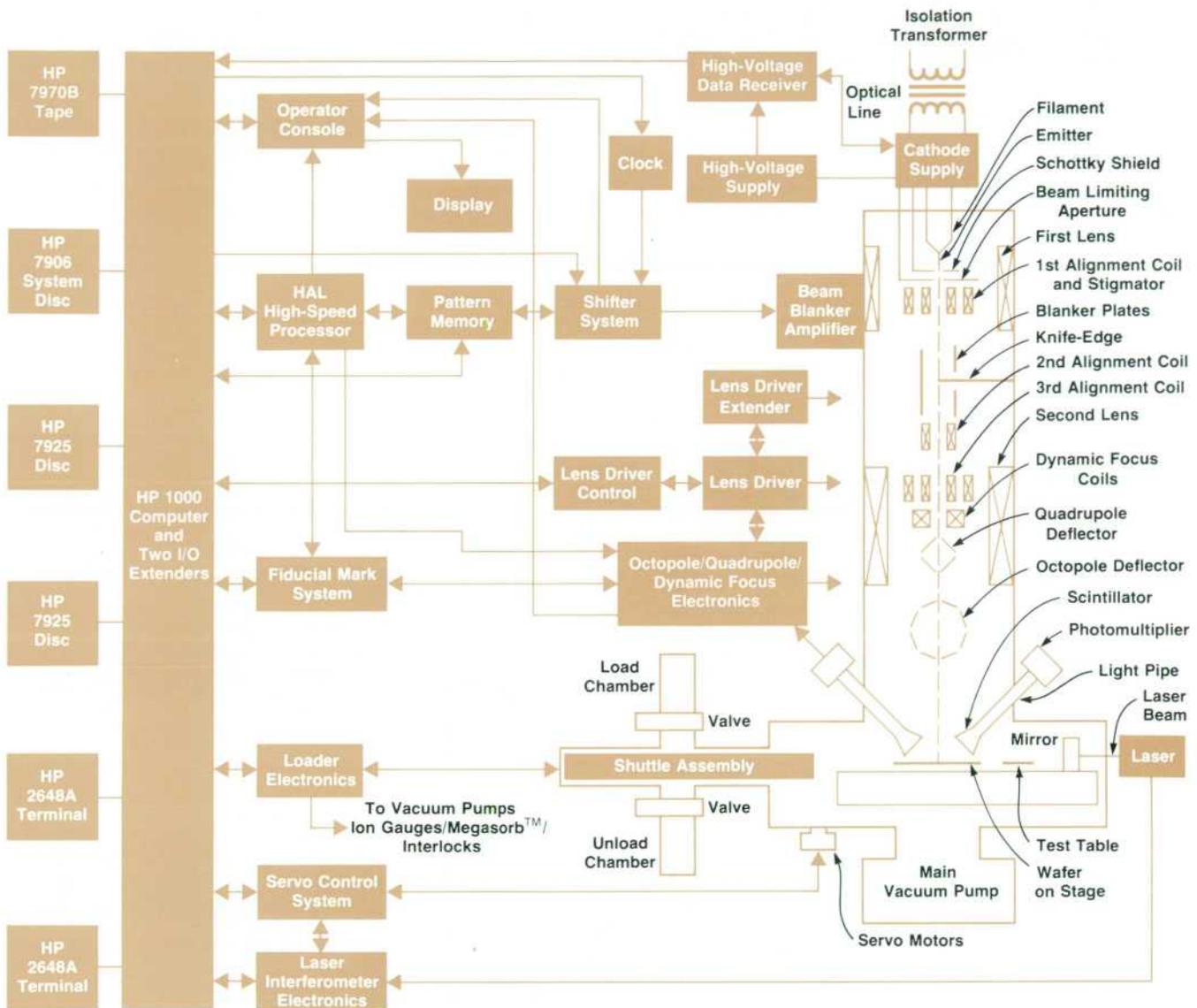
Fig. 1. *Block diagram of the HP high-speed electron beam lithography system.*

The Electron Column

At the top of the column is the electron gun, consisting of a tungsten cathode and an accelerating anode. The beam formed by the emitted electrons passes through the first lens and is focused near a knife-edge in the beam blanker region. In the region of the first lens is a stigmating system to correct beam astigmatism (i.e., to make the beam round), and an alignment coil to compensate for column misalignment. Because of machining tolerances, column alignment to the required accuracy is not possible without this electronic control.

Following the first lens is the beam blanker. Here the beam is focused just to the side of a knife-edge. When a potential is applied to the blanker, the beam is deflected a few micrometres and impinges on the metallic edge. Thus the electrons are stopped instead of continuing down the column. Because the beam focused near the knife-edge is only a few micrometres in diameter, a very small deflection is required to move the entire beam. This means that a small

voltage (about three volts) is adequate to provide the necessary blanking.

After the beam traverses the blanker, it passes through a second alignment coil and then goes through a double vacuum valve. This valve allows the column to be split apart while maintaining vacuum both on the gun and on the lower sections of the column. The ability to split the column in this fashion greatly facilitates installation of new emitters if one should fail. In this way, a standby gun can be kept ready for rapid transfer to the column, so that unnecessary down time is eliminated.

Following the valves, and immersed in the final lens, are the final stigmators and alignment coils, which focus the beam on the target. Here again is a set of alignment deflectors to aid in directing the beam to the target. Following the final lens is the deflector system. The upper deflector, a quadrupole deflector, is used to raster-scan the beam rapidly over a limited area of 128×128 pixels. The second deflector is an octopole, which has a very uniform field and

Electron Beam Lithography

The use of computers has revolutionized society. From their beginnings in the 1950s, computers began to flourish in the 1960s. In the 70s people began to understand how to use them, and in the 80s their use is beginning to explode. One of the main reasons for their widespread use is the fact that their cost for a given capability has been going down steadily. It has now dropped to the point where computers are economical for many household and business applications.

The cost of computing has dropped primarily because the complexity and capability of integrated circuit chips have increased at a constant rate for the past decade. Because of this increased capability, we see integrated circuit chips not only in computers, but also in microwave ovens, calculators, TV games, typewriters, and many other products.

A principal reason for the increase in the capability of integrated chips and the reduction in their cost has been the increased capability of the lithography machines used in their manufacture. Integrated circuits are built up on silicon wafers in a series of operations. Some of these operations involve etching the surface of the wafer to produce a particular pattern, while others require introducing certain impurities into the wafer in a particular pattern. It is lithography machines that allow us to imprint the patterns on the wafer. The wafer is first coated with a resist, a compound that is sensitive to light or to electrons. The lithography system then exposes the resist to light or electrons in the desired pattern, and the resist is developed, leaving the desired pattern as openings in the resist through which the etching or implanting can be done.

During the past 10 years, the dimensions of these patterns have been decreasing at a steady rate. This has been made possible by advanced optical aligners in which the lenses have been improved from year to year, allowing the user to imprint finer and finer geometries onto the silicon wafers. At present, the minimum geometries that are obtained in production have lines and spaces as small as two to three micrometres, and in the laboratories as small as one micrometre. The fundamental limit of these lenses is being approached; it appears to be in the area of one micrometre or slightly less. Another criterion, which is becoming more and more important as the geometry shrinks, is the accuracy with which we can place these geometries at the desired location. The best optical aligners on the market today can place geometries to an accuracy in the order of 0.3 to 0.4 micrometres. (The diameter of a human hair is approximately 100 micrometres.)

To continue increasing the capability of integrated circuit chips, designers must be able to attain finer geometries and to place them more accurately. Electron beam systems have ultimate limits in resolution and in accuracy much finer than those of optical aligners. It is for these reasons that there is great interest in electron beam lithography systems.

Another factor that is extremely important, particularly in production, is throughput of wafers per hour. This number sometimes is misleading because of the way that it is defined. Many people tend to use the number of wafers a particular system can process in an hour. More meaningful is the number of functions that can be processed per hour by a given system. The bottom line is really the cost per function of an integrated circuit chip.

Although a few companies have developed direct writing electron beam systems (systems capable of writing a pattern directly on a resist-coated wafer, using an electron beam), the only electron beam systems that are commercially available today are designed for making masks for integrated circuit production. An optical aligner must be used to project the mask image onto the wafer. This means that most of the finer resolution and better accuracy of electron beam lithography are ultimately lost because the limitations of optical lithography dominate.

Besides optical lithography, the only other technology that appears to be competitive with electron beam lithography is x-ray technology. This technology requires masks, but because the illuminating source is of much shorter wavelength than optical, the minimum geometries attainable with x-rays are much finer. X-ray lithography has been in development for a number of years now, but has not reached production stages because of two major problems, the construction of the masks and the availability of a sensitive resist. At this time, it is not obvious which technology will eventually dominate. If x-ray technology's two major problems are solved, it is likely that all three technologies will be complementary and will exist side by side. To summarize, the major advantages of electron beam lithography are greater resolution, better registration accuracies, and greater flexibility in handling the pattern data. The major disadvantage is the high cost of the instrument.

HP System

The HP electron beam lithography system described in this issue is designed for direct writing on wafers in a production environment. It is much faster than any commercially available system and is believed to be much faster than any electron beam system existing today. Its speed is due primarily to three things: the use of a field emitter electron source, high-speed electronic circuitry to process the data, and the deflection system.

As an example of the potential throughput of this system, consider HP's recently disclosed microcomputer chip that has on it approximately 450,000 transistors. Assume that an optical aligner can process about ten wafers per hour containing such chips. HP's electron beam system can potentially produce three times as many transistors on the same-size chip, but at perhaps five wafers per hour. In functions per hour, the electron beam system has a greater throughput.

The HP system has the ability to print lines with minimum dimensions in the order of 0.5 to 1 micrometre and has registration accuracies in the order of 0.1 to 0.2 micrometre. Future systems will continue to reduce these parameters.

-Frank Ura
Manager, Electron Beam
Research and Development

is used to position the beam at the center of a 128×128-pixel subfield.

For alignment purposes, it is necessary to scan various targets mounted on the stage. These targets allow the system to focus the beam, measure deflection sensitivities of the two deflectors, and get a general view of system accuracy. To detect the beam as it scans these targets, there is a system of four detectors in the target area, corresponding to the + and − X and Y axes of the system. They are made of a scintillating material that emits light when struck by backscattered electrons, a light pipe that routes the light to a convenient location, and a set of photomultiplier tubes which convert the weak light into electrical signals.

The Pallet Handling Mechanism

Substrates, (wafers, masks, or reticles) are processed semiautomatically in this system. Substrates are loaded onto a holder called a pallet, normally one substrate to a pallet, but in the case of small substrates such as 5-cm wafers, possibly more. A pallet is roughly 12 cm on a side. Pallets are loaded into a cassette. The cassette is placed in the load chamber of the pallet loading system and the chamber is pumped down to vacuum. After a sufficient level of vacuum is attained, the cassette is lowered through a valve into the central chamber, and the valve is closed, freeing the load chamber for the next cassette. Pallets are now individually loaded from the central chamber onto the X-Y stage. After exposure, the pallet is retracted and the next pallet inserted onto the stage. This process continues until all substrates have been exposed. After all of the substrates in a cassette have been exposed, the cassette is lowered through a valve into the unload chamber, where it may be removed or, as required for some electron-sensitive resists, vacuum-cured. The sequencing of valves, detection of appropriate vacuum levels, and all of the motions of the substrates are fully automatic, but the insertion of cassettes into the load chamber and their removal from the unload chamber are manual operations.

The X-Y stage is a very precise mechanism controlled by a laser interferometer and servo system. It is capable of over 13 cm of travel in both X and Y directions, and can move 5 mm and settle to less than 0.1 μm from its destination in only 250 milliseconds. In normal operation it is driven to a destination with an accuracy of 0.016 μm, or 160 angstroms! Resolution is 0.008 μm or 80 angstroms. This extreme precision and a third interferometer axis for measuring rotation make it possible to stitch together fields of view to obtain much larger patterns than the 5-mm-square single field of view.

Computer Control

Overall system coordination is accomplished via a host computer, in this case an HP 1000. This machine is a 16-bit computer with extended memory and two input/output extenders. It is used somewhat as a process controller, in that all of the beam control, deflection control, data handling, and analog processing are done with special-purpose hardware designed specifically for the system. The large number of peripheral controllers required the addition of the I/O extenders. Attached to the computer is a system disc for program storage, a magnetic tape for system backup and program transfer from other machines, several terminals used to enter operator commands, and four HP 7925 125-megabyte discs for pattern data storage. The large capacity required of the data storage discs stems from the fact that one of the two data formats used is a direct bit map of a chip. With 100,000,000 bits required for a single 5-mm field of view, disc space is used rapidly!

Noise Reduction

In a system that maintains the exceptional accuracy of this one, great care must be taken to eliminate every possible source of noise. Each electronic subsystem is powered by its own set of power supplies, and these are driven from separate isolation transformers. The main power source for the entire system is a separate winding on the transformer supplying the building housing the electron beam system. Single-point grounding is used, eliminating the possibility of ground loops. To further reduce any possibility of ground loops, all digital data and control lines are optically isolated between all electronic subsystems, including the host computer.

To reduce magnetically induced noise, transformers and other magnetically active components are not permitted near the column. Room lights are required to be incandescent bulbs rather than fluorescent fixtures, because fluorescent fixtures have magnetically active ballasts, and the tubes themselves act as unshielded current carriers. To further reduce interference from stray fields, the column is completely surrounded by two layers of mu-metal shielding spun to conform to column geometries. Critical components of the column itself are made of nonmagnetic materials.

Finally, the entire column and loading mechanism are floated on air bearings to eliminate horizontal components of floor vibration. Vertical components of vibration are

---

Proximity Effect Correction by Processing

The proximity effect is the unwanted exposure of resist caused by electrons scattering in the resist layer and from the substrate beneath the layer. This leads to pattern features different from those designed. It is a major limitation to the fabrication of high-resolution devices by means of electron beam lithography.

Three methods have been used to limit the proximity effect. The most widely used method involves preprocessing the input data and modifying the dose at exposure time. This method has been used with vector-scan electron beam systems.[1] the second method also involves preprocessing the input data; the exposure shapes are modified by the addition or subtraction of scans to achieve a uniform dose. This method has been applied to electron beam systems that are modified electron microscopes, where the smallest patterns are formed by many beam scans.

In the HP system, other methods, which do not involve preprocessing the input data, are used to limit the proximity effect. Some of these techniques involve special resist processing, hence the term: proximity effect correction by processing. These processing techniques reduce proximity effect by minimizing the effect of backscattered electrons and reducing their number. The proximity effect has been investigated by means of a computer model, and processing methods that can improve the pattern fidelity have been explored.

It has been found that significant reduction in the proximity effect can be achieved by two methods: a "small-beam" approach and multilayer resist techniques.[2] Both of these methods have been evaluated on the HP electron beam lithography system. An article detailing these studies and the results is planned for a future issue of the Hewlett-Packard Journal.

References

1. M. Parikh, "Corrections to Proximity Effects in Electron Beam Lithography," Journal of Applied Physics, Vol. 50, 1979, pp. 4371-4387.
2. P. Rissman and M. Watts, "Optimization of Beam Diameter for Proximity Effect Reduction," 16th Symposium on Electron, Ion, and Photon Beam Technology, May 1981, Dallas, Texas.

eliminated with a system of pneumatic dampers.

Pattern Memory System

The pattern memory is a very large semiconductor memory system capable of storing more than 151 million bits, which is more than enough for a direct bit map of a 10,000×10,000-pixel chip, or over three million rectangle specifications in compressed format. During exposures the memory delivers this data at a 300-MHz bit rate. This high data rate precludes the use of large disc memories for the pattern memory.

The first stage in the pattern memory is an input section which accumulates 144 bits, 16 at a time, from the host HP 1000 Computer. These 144 bits consist of two 72-bit words, each word containing 64 data bits and eight error correcting bits. After 144 bits are accumulated on the data input board, they are transferred through error correction circuitry into the memory arrays. The memory arrays store the pattern information until it is time to transfer the data to the beam at write time. This array section is composed of six modules. Each module contains eight memory boards, and each board has eight rows of 24 chips. The chips are 16K MOS RAMs. The total capacity of the memory is therefore greater than 151 megabits including the overhead for check bits. At exposure time, data is transferred again through error detection and correction circuitry from the memory modules into one of two areas. If the data is in bit map format, it is transferred directly into one of four block buffers. The block buffers hold 128×128 pixels of pattern information, which is exactly what is required to expose one 64×64-micrometre subfield or block with the raster deflection system. If the data is in compressed format it is routed through the data decompressor system and then into the block buffers after decompression.

The block buffers provide a means of synchronizing the data as it comes from the memory modules and goes to the beam blanker. The memory modules operate with a fixed clock and require refresh. Uncorrectable data errors may require several retries before data is correctly read. The beam blanker, on the other hand, may operate at blanking rates from a few kilohertz to the maximum rate of 300 MHz, so some type of intermediate storage is required to allow proper coordination.

From the block buffers the data is again transferred through error detection and correction circuitry, although at this time the decompressed data cannot be corrected because of the transformation process. The data is transmitted to the shifter system, which is a high-speed parallel-to-serial converter.

Dumping of the data from block buffers to the shifter is mediated by a control system which synchronizes the shifter, requests to dump from HAL (the high-speed processor), and the block buffers. It also can provide for block skipping or filling in the event of blank or completely filled blocks, thereby saving space in the memory modules. Line skipping may also be provided in the bit map format to save writing time. The overall control of the pattern memory system is accomplished via a programmable processor. All of the timing and control is implemented in ECL to obtain accurate timing and because the ECL logic families generate very low noise, an important consideration in this system. The pattern memory system's internal states and registers can be read, modified, or set from the host computer to facilitate troubleshooting and system debugging. Excellent support software allows very flexible troubleshooting operations.

Shifter System

The shifter system is a parallel-to-serial converter that receives 128 bits of data from the pattern memory and shifts it out at rates of up to 300 MHz to the beam blanker amplifier. Because of the extremely high transfer rates, shifter control is not implemented with combinatorial logic. Instead, the control signals are themselves generated from high-speed shift registers. Four signals are required to control the data shifter and synchronize data transfers. One signal controls the state of the main data shifter, forcing it either to load data or shift the data. The second signal a ramp reset signal, which is used to start the ramp in the raster generator or reset the ramp at the end of a scan. The third signal provides pulses that load data from the pattern memory into a temporary register, prior to the transfer of the data into the main data shifter. The fourth and last shifter generates a pulse that controls the state of all of the control
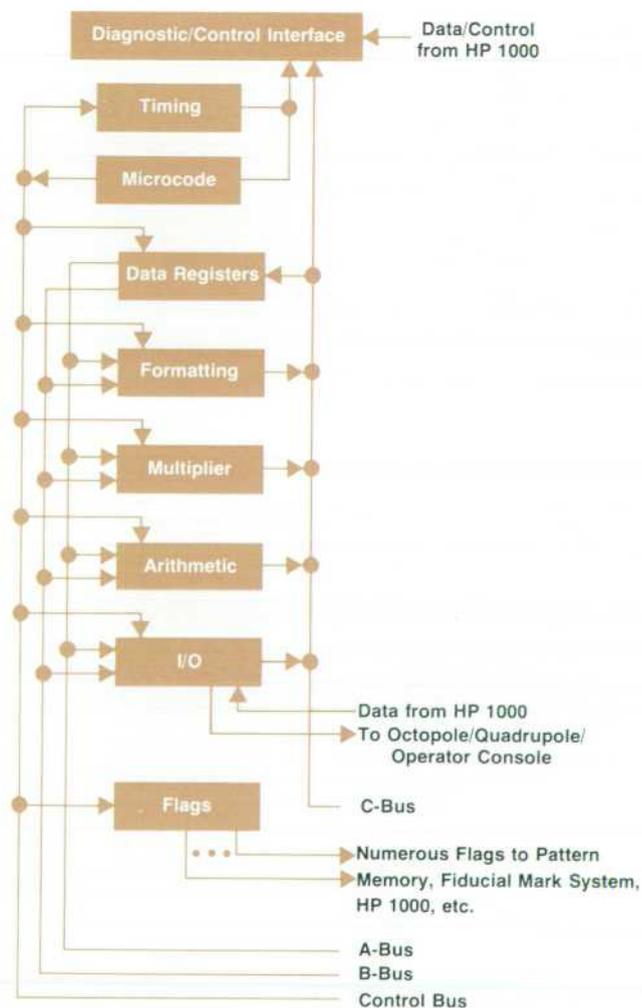

Fig. 2. HAL, the high-speed processor that calculates voltages for the octopole deflection plates, is implemented with approximately 1000 ECL circuits.

shifters, including itself. The two possible states are 1) load data from a register associated with each individual shifter, and 2) shift the data. The waveforms needed to provide the control are stored in a set of four registers, each register feeding one of the four control shifters. This data is loaded into the shifters and then shifted out. The data is then reloaded and shifted again. The process repeats indefinitely. Although this system is expensive in IC count, it provides accurate control signals at 300 MHz.

HAL System

The main HP 1000 system computer is far too slow to do the numerous calculations required to determine the voltages to be placed on the octopole plates during exposure, since these calculations must be done in real time on a block-by-block basis. These calculations are done by the HAL system, a special-purpose high-speed processor implemented with emitter-coupled logic (ECL). It is microprogrammed, and can execute a program of 512 steps. The microcode is generated using an assembler and is downloaded from the host computer. This allows flexibility in controlling the beam deflection protocol, since changes are quite easy to make. In addition to the microcode storage registers there are 128 32-bit data registers for arithmetic operations.

HAL is built around several buses, the first of which is the opcode bus (see Fig. 2). When the microcode memory places an opcode on this bus, the microcode is received at the inputs of all of the printed circuit boards that make up HAL. Each printed circuit board is designed to execute one or more instructions. When a board recognizes one of the opcodes it is capable of executing, it does so and then signals the control system that it is finished. The control circuit then fetches the next instruction. This architecture allows considerable flexibility in modifying or inserting

new instructions. To add a new instruction, one needs only to implement it on a board and insert it into the system. The control system and timing do not have to be modified, since any board can take as long as necessary to complete an instruction.

The present instruction set consists of add, subtract, multiply, a wide variety of logical operations and format conversions, and a flexible collection of special handshakes and flags. These flags are used to signal other units in the system, for example to tell pattern memory to dump a block of data or to start the fiducial mark system so that it can collect data. Some flags can be used to generate pulses of selected width, and others are used to signal HAL to start execution of special subroutines. Other flags are used to coordinate data transfers to the HP 1000 host computer.

Quadrupole Electronics

The quadrupole electronics generate raster-scan voltage waveforms for driving the quadrupole deflection plates, which electrostatically deflect the electron beam. The quadrupole electronics also provide for magnifying and rotating the raster. The deflection waveforms can be generated over a five-decade range in operating frequency.

The quadrupole electronics consist of two independent ramp and step generators. Each generator is responsible for the waveforms for one pair of quadrupole plates.

The ramp and step generator (see block diagram, Fig. 3) generates output voltages that are sums of ramps and steps. The ramp portion of the waveform is formed by an integrator composed of R1, C1, A1, and Q1. The slope of the ramp is controlled by the output voltage of a digital-to-analog converter (DAC). The ramp is reset by FET Q1. The step portion of the waveform is generated by two DACs, one to control the step size, and the second, a multiplying DAC or MDAC, to control the number of steps. The ramp and step
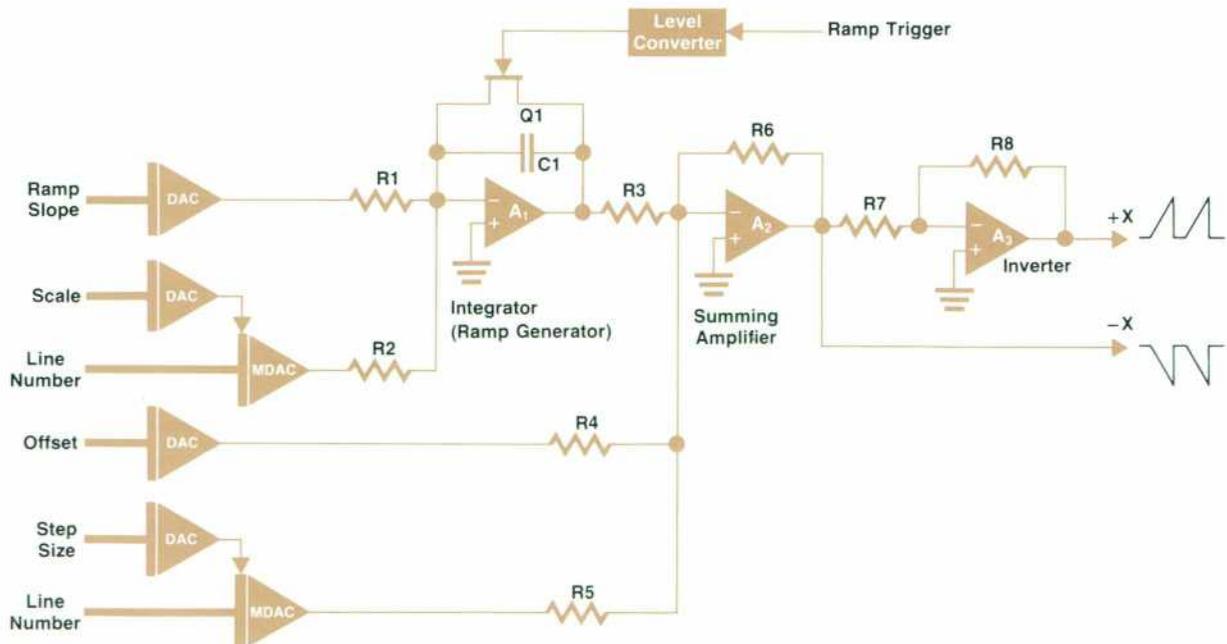

Fig. 3. *The driver for the quadrupole deflector generates raster-scan voltages to deflect the electron beam. It consists of two of these ramp and step generators.*
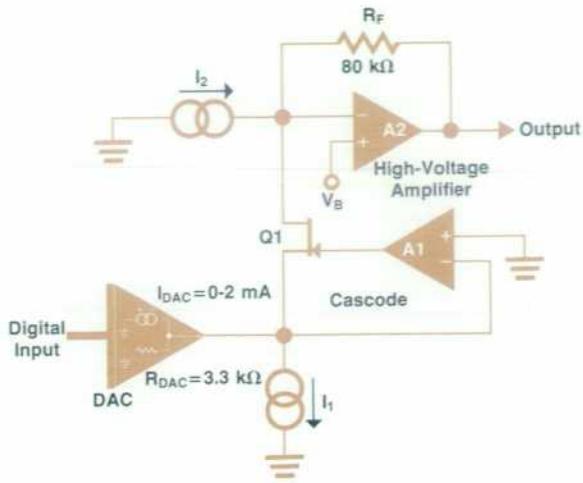
Fig. 4. Voltages for the octopole deflection plates are generated by eight of these amplifier-DAC combinations.

two decades. The outputs of the various amplifiers are then selected by multiplexers for final presentation to the deflection plates. The final waveforms generated by the quadrupole electronics are linear within better than 0.1% over the four-volt deflection range sufficient for a 64×64-μm block scan. The raster-scan signals are compatible with the 3-kHz-to-300-MHz 0.5-μm-step data rate.

Octopole Amplifier and DAC

The voltages required by the octopole deflectors are generated by eight amplifier-DAC combinations. One channel is shown in Fig. 4. To deflect the electron beam over a 5-mm$\times$5-mm field, a $\pm$80-volt capability for each deflection plate is required. Since the electron beam system's accuracy and speed are directly affected by the octopole's accuracy and speed, maximizing these parameters is essential. Accuracy is obtained by using DACs with the highest accuracy available. The best commercial units achieve 16-bit long-term stability and 17 or 18 bits short-term. The settling time for these DACs is typically 7-8 microseconds.

To convert the DAC current of 0-2 mA to the $\pm$80-volt signal required by the deflection plates, a high-voltage amplifier with an 80-kΩ feedback resistor (R_F) is used. Placed between the amplifier and the DAC is a cascode circuit formed by Q1 and A1. The purpose of this cascode is to increase the impedance seen by the high-voltage amplifier A2. If the cascode were not used, A2 would be affected by the 3.3-kΩ output impedance of the DAC (R_{DAC}). This would result in an increase in noise seen at the output by a factor of $R_F/R_{DAC} \approx 25$ and an increase in the settling time of the signal by the same factor. With the cascode, the noise gain of A2 becomes unity, reducing both the noise and the settling time. The cascode section has its own noise effect and setting time. However, the detrimental effect on the output

portions are added together by summing amplifier A2. A third input to A2 adds an offset signal, which has two main purposes: to make the center of the raster scan equal to zero on each plate, and to accommodate offset voltages in A1 and eliminate any nonlinear portion of the ramp signal. The output of A2 is used to drive one of the deflection plates and is inverted by A3 to drive the opposing plate. One extra input to the integrator is included to modify the ramp slope through R2 in case any interaction occurs between the ramp and step signals either in the electronics or in the electron optics.

To accommodate the wide range in operating frequency, three separate integrator, summing amplifier, and inverter sections are used, each optimized for operation over one or
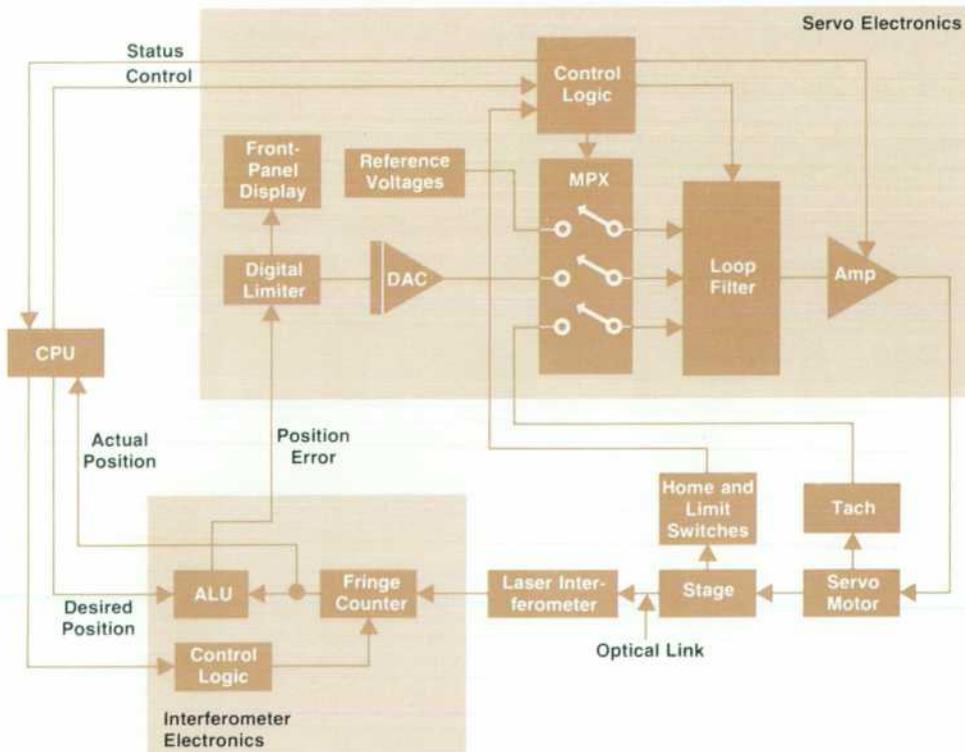

Fig. 5. Block diagram of one axis of the two-axis control system that drives the X-Y stage that carries the substrates to be exposed. The stage is positioned with an accuracy of 0.016 μm.

is considerably less than without the cascode. The overall performance (accuracy and settling times) of the cascoded amplifier circuit is determined by the DAC rather than the cascode or the high-voltage amplifier.

Blanker System

The blanker system provides the high-speed intensity modulation of the electron beam necessary to write patterns as the beam is moved across the wafer or mask in a raster-scan fashion. The blanking system is composed of three main parts: the blanker amplifier, the interconnect strip-line, and the blanker deflection plates. The blanking is accomplished by deflecting the beam into a knife-edge. The deflection center is placed at a focal point of the final lens so that, ideally, no motion occurs at the target as a result of the blanker deflection.

The blanker amplifier is a high-speed driver that converts the ECL input signals to complementary three-volt levels suitable for driving the deflection plates. The amplifier operates from dc to 300 MHz with rise times of approximately 1 ns and provides complementary outputs.

The stripline is fabricated on a Teflon™ board which is divided into two sections. A specially fabricated connector maintains the high-frequency integrity of the strip-line while providing the capability to remove the blanker amplifier without unsoldering any component. Two vacuum feedthroughs, also stripline, interconnect the Teflon™ board and the deflection plates.

Stage Control System

The motion of the electron beam system's X-Y stage is controlled by the two-axis stage control system, one axis of which is shown in Fig 5. The X-Y position of the stage is monitored by a two-axis laser interferometer. A third interferometer channel is used to monitor the rotation of the stage. The interferometer electronics, consisting of an arithmetic/logic unit (ALU), fringe counter, and control logic, processes the interferometer signals to obtain outputs of position and position error. The interferometer electronics, similar to the HP 5501A, achieves $0.008\text{-}\mu\text{m}$ resolution and is able to track stage motion up to 10 centimetres per second.

The servo electronics receives the 28-bit position error signal from the interferometer electronics and drives the motors that move the stage. The position error signal is truncated to 14 bits plus sign by a digital limiter and displayed on the front panel. The truncated error signal is converted to an analog voltage by a specially designed digital-to-analog converter (DAC) which achieves precise voltages near zero without resorting to expensive, high-accuracy DACs.

Under control of the host computer, the servo electronics operates in the following modes:
- Slew at constant velocity in positive direction
- Slew at constant velocity in negative direction
- Brake (zero motor armature voltage)
- Coast (zero motor armature current)
- Position control.

Depending on the mode, the control logic section of the servo electronics selects appropriate input and feedback signals for the loop filter. For example, in slew mode, a reference voltage of appropriate amplitude and polarity is compared to the tach voltage and the difference is integrated by the loop filter to drive the motor at constant angular velocity. Similarly, in position control mode, the DAC output and the tach signal are summed and integrated by the loop filter to drive the stage position to a zero-error condition. In position control mode the system settles to less than $0.1\mu\text{m}$ error from a 5-mm step in less than 250 ms and the final position error is less than $0.016\ \mu\text{m}$.

Lens Drivers and Control

The main function of the lens driver and control subsystem is to provide a focusing system for the electron beam. The main constraints are stability and ease of use.

The essential features of this subsystem are shown in Fig. 6. There are two parts, a control section in the main system racks near the operator, and a driver section near the column.

The control section (lens driver control) is the interface between the HP 1000, the operator, and the lens drivers. The control section permits operation in two modes: 1) under supervision of the HP 1000 and 2) as a stand-alone device. The latter mode was necessary since the lens system is also used on column and gun test stations, which are physically separate from the electron beam machine. In normal opera-
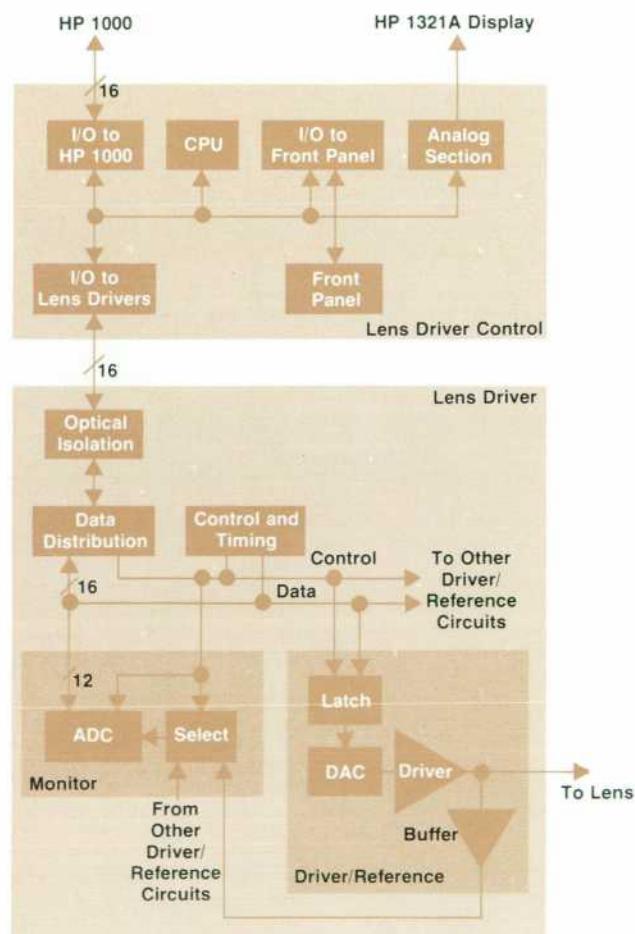

Fig. 6. *The lens driver and control subsystem provides a focusing system for the electron beam.*

tion the control section is under the supervision of the HP 1000. The HP 1000 can transmit or receive data from the lens driver control and can force the lens driver CPU to execute either resident programs or programs downloaded from the HP 1000.

There are three resident programs in the 8K bytes of ROM associated with the lens driver CPU (a Z80 microprocessor). The main program allows the operator to change and monitor the settings of any of the lenses via pushbutton control from the front panel. The second mode provides an additional capability of sweeping any pair of lenses in a raster fashion and providing synchronous sweeps for driving a display scope. These two modes permit the operator to focus the column. The third mode is a test mode that allows the operator to check the analog outputs of the lens drivers.

The lens driver section has a data and control section, a monitor section, and 13 driver/reference sections. In the data and control section, the data is received from the lens driver control and is optically isolated to prevent system ground loops. The data is then distributed to the appropriate reference board on a 16-bit three-state bus. This bus is controlled by the control and timing circuit, which can direct the distribution of data or can send to the lens driver control the ADC (analog-to-digital converter) output, the contents of any of the latches, or status information from the control and timing circuit.

The ADC section is used to monitor the analog outputs of the various lens drivers and lens driver power supplies. A 12-bit ADC is used and the input is selected as directed from the lens driver control.

Each driver/reference section has a latch to hold the 16-bit data word received from the lens driver control. The output of this latch drives a 16-bit DAC which generates a reference voltage proportional to the desired lens output. The DAC
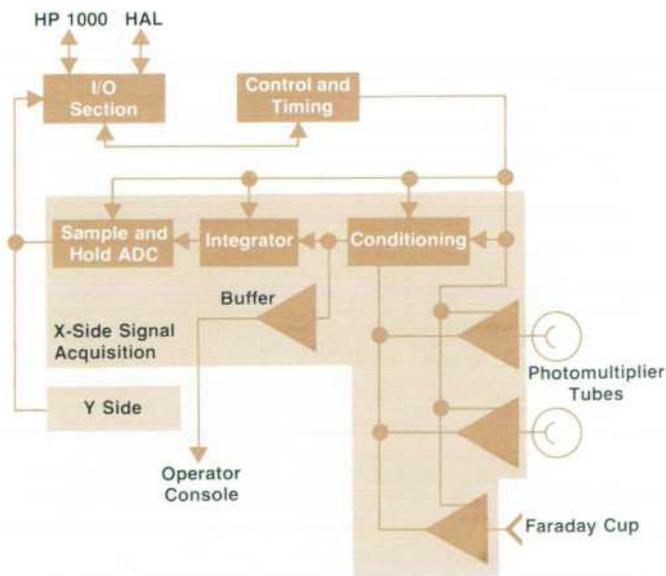
and all of the analog components were selected with stability as the overriding constraint. The drivers are current sources driven by the DAC outputs. There are several different types of drivers to match the various types of lenses. They vary from unipolar drivers with outputs of several amperes at up to 40 volts to bipolar drivers supplying only a few hundred milliamperes into a lens with only a few turns.

Loader Electronics

This subsystem is designed to control both the mechanical system that loads the substrates into the machine and the associated vacuum equipment. The main element of this subsystem is the HP 2240 Measurement and Control Processor which provides the main interface to the HP 1000. This subsystem is characterized by the large number and many types of sensors and control devices to be interfaced, which run the gamut from encoder wheels and LED sensors to step motors and valves.

Fiducial Mark Electronics

This subsystem is designed to acquire signals from fiducial marks on the substrates and transmit these signals either to the system operator console in analog form or to the HP 1000 in digital form.

A block diagram is shown in Fig. 7. The interface to the HP 1000 is provided by the I/O section. The control section provides the necessary control and timing signals for the other sections. This section is active in the nulling of the preamplifiers and the integrator under the control of the HP 1000, the selection of the signal to be acquired, and in conjunction with the control link to HAL, the acquisition of fiducial mark data.

The preamplifiers take a low-level ($<1\ \mu A$) signal either from a photomultiplier tube viewing a scintillator excited by backscattered electrons from the target or from a Faraday cup or other current-collecting electrode within the column. Each preamplifier has a 12-bit DAC that can zero the output of the preamp under control of the HP 1000.

The outputs of the preamps are sent to a conditioner section, where they are summed or differenced with provi-

Fig. 7. The fiducial mark system conditions signals from photomultiplier tubes or a Faraday cup within the column. The system computers use the conditioned signals to detect fiducial marks on the substrates so that the substrates can be precisely located.
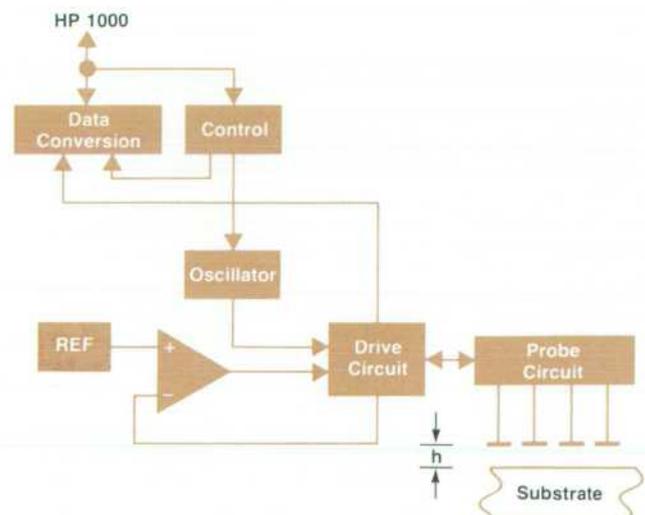

Fig. 8. The height sensor system measures the height of the substrate within a few tenths of a micrometre.

sion (via another DAC) for nulling the result. At this point a selection is made among the individual preamp outputs, the sum, the difference, or ground. The selected signal is sent to the integrator and the buffer.

The buffer converts the single-ended analog signal into a differential signal and transmits it to the operator console, where it is received with a high-input-impedance differential amplifier. This is to provide ground loop isolation. This analog signal is used by the system display in the scanning electron microscope mode.

The integrator provides the ability to average the input signals from the detectors over a period of time and adjust the amplitude to match the dynamic range of the sample-and-hold circuit and 12-bit analog-to-digital converter that follow. The integration capacitor can be selected by the HP 1000 to provide ranging over a factor of 300. A DAC is used to null the output of the integrator. The integrator time constant is set by a crystal-controlled clock section and can be varied from about 1 to 3000 μs.

Height Sensors

The height sensor system (Fig. 8) measures the height of the substrate relative to a plane determined when the machine is calibrated. The height must be known to an accuracy of a few tenths of a micrometre to meet the system specifications.

The sensor is a capacitive type. A probe consisting of a center portion and a guard ring is positioned above the substrate. These elements are excited at 1 MHz and the capacitance to the substrate is measured. The drive circuit supplies the 1-MHz signals to the probe and a dc error signal to the loop amplifier. When the loop is balanced so that the charge on the probe is constant, a dc signal proportional to the applied RF voltage and therefore the height h is sent to the data conversion section. The oscillator can be gated on or off by the control section so that during exposures there are no RF voltages on the probes to cause interference with the electron beam.

Four of these sensors are mounted around the beam axis. Thus not only height but also substrate tilt can be measured. The height sensor system is capable of measuring height variations up to 0.1 mm to an accuracy of better than 0.25 μm when properly calibrated.

Gun Supply

The gun supply powers the field emission electron gun that supplies the electron beam. The beam accelerating voltage is provided by a very well regulated, stable 20-kV supply. A small dc voltage to heat the filament and up to 4 kV for the first anode (Schottky shield) are provided by separate supplies which are floated at anode potential. Monitor circuits are provided to measure the relevant currents and voltages and transmit them to the HP 1000. Considerable effort was necessary in designing this system to reduce leakage currents and to protect the system from damage by arcs.

Operator Console

The operator console is the interface between the operator and the system when it is run as a scanning electron microscope. Inside the console are I/O, control and front-panel sections that monitor front-panel buttons and

Ronald K. Scudder

Ron Scudder has been involved with the software and electronics of HP's electron beam system since 1974, serving as designer, project manager, and now department manager. He's been with HP Laboratories since 1973 and has also helped develop an acoustically tuned optical filter. Born in Orlando, Florida, Ron attended Massachusetts Institute of Technology and received his BSEE degree in 1967. His MSEE degree was awarded by the University of California at Berkeley in 1969, and he was a PhD candidate at Berkeley from 1969 to 1973. Ron enjoys a variety of water-oriented sports, birdwatching, photography, carpentry, backpacking, reading, travel, racquetball and skiing. He's a resident of Menlo Park, California.

Wayne C. Haase

Wayne Haase is a consultant to HP's electron beam project team. A graduate of Massachusetts Institute of Technology and Stanford University, he received his BS, MS, and EE degrees in 1965, 1967, and 1969 from MIT and his PhD degree in 1978 from Stanford. All four degrees are in electrical engineering. He worked on communications satellite systems with MIT's Lincoln Laboratory from 1967 to 1970 and has been a consultant to numerous companies since then. He has co-authored six technical papers on medical electronics and electronic circuits and is a co-inventor on two pending patents dealing with aircraft instrumentation. Wayne was born in Miami, Florida. He plays several musical instruments and is a magician, tennis player, and skier. He's married, has two children, and lives in Mountain View, California.

John C. Eidson

John Eidson received his BS and MS degrees in electrical engineering from Michigan State University in 1958 and 1960, and his PhD degree from Stanford University in 1964. Now project manager for electron beam electronics, he's been with HP since 1972, contributing to the design of an optical spectrometer and the electron beam system. Before joining HP, he did research and development on mass spectrometers, x-rays, and infrared devices, and taught electronics and electromagnetic theory as a member of the Stanford faculty. He's a member of IEEE and author of eight papers. Born in Melrose Park, Illinois, he's married, has two sons, and lives in Palo Alto, California. His leisure activities include bicycling, woodworking, and youth soccer.

SAWR Device Fabrication

One of the first applications of the HP electron beam lithography system has been to produce high-resolution surface-acoustic-wave resonator devices (SAWRs).

SAWRs with 500-nm and 300-nm critical dimensions have been produced in cooperation with the Applied Physics Lab of Hewlett-Packard.[1] The photographs below show the transducer regions of two devices; they operate at 1.6 GHz and 2.6 GHz for the 500-nm and 300-nm devices, respectively.
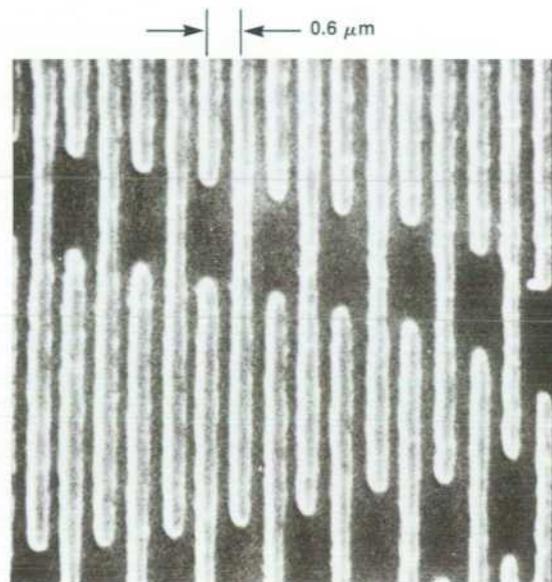
An article dealing with HP's SAWR technology is planned for a future issue of the Hewlett-Packard Journal.

Reference
1. P. Cross, P. Rissman, and W. Shreve, "Microwave SAW Resonators Fabricated with Direct-Writing, Electron-Beam Lithography," Proceedings of the 1980 Ultrasonics Symposium.
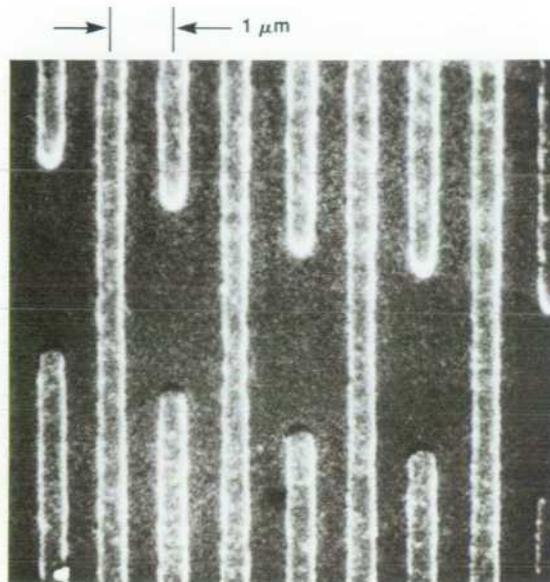
(a) 500-nm Lines/Spaces — 1 μm

(b) 300-nm Lines/Spaces — 0.6 μm

set displays.

The operator console also generates the analog signals sent to the system display. The signals are derived from signals received from other parts of the system. All incoming signals are either optically isolated or isolated by high-input-impedance buffers to break up system ground loops.

A Precision, High-Current, High-Speed Electron Beam Lithography Column

by John Kelly, Timothy R. Groves and Huei Pei Kuo

THE ELECTRON BEAM COLUMN used in the HP electron beam lithography system was developed and constructed to meet the requirements for an accurate, high-speed, direct-write production system. The requirements for this system called for a beam current of 600 nA to be focused into a 0.5-μm-diameter spot. This is sufficient current density (306A/cm^2) to expose electron-sensitive resist with a sensitivity of 1 μC/cm^2 at the system's full data rate of 300 MHz. This in itself is not difficult to achieve, but taken together with the need to deflect the beam at high speed over a substantial area, the problem becomes much tougher.

The column, illustrated in Fig. 1, is a short two-lens system which uses a field-emission cathode. This cathode is a very small, bright source of electrons, which enables the required current to be focused into a small spot and at the same time permits large deflections of the electron beam.

Both lenses are magnetic. The gun lens produces a magnified crossover, which is used for beam blanking purposes (beam switching). The second or final lens transfers the blanker crossover to the target plane with a working distance of 20.5 cm. The overall magnification of the system is 1.9×. Beam blanking is done by an electrostatic deflector.

The wide-angle, high-accuracy deflector is located just below the final lens and requires 80 volts to deflect the beam to the corner of the 5-mm-square deflector field. The high-speed deflector is positioned inside the final lens and needs only 2.5 volts to deflect the beam to the extremes of the 64-μm-square raster-scan field. Between the high-accuracy deflector and the target is the detector system for locating fiducial marks. This system detects backscattered electrons and consists of four independent scintillator, light-pipe and photomultiplier assemblies.

Alignment of the column is achieved by using three sets of magnetic deflection coils in conjunction with the movable first lens. The electron gun itself is entirely prealigned.

Dynamic corrections for astigmatism, beam focus, and position are applied to the column during operation. Astigmatism and beam position are corrected by signals applied to the wide-angle deflector. Beam focus is adjusted by means of a small dynamic focus coil buried in the middle of the final lens.
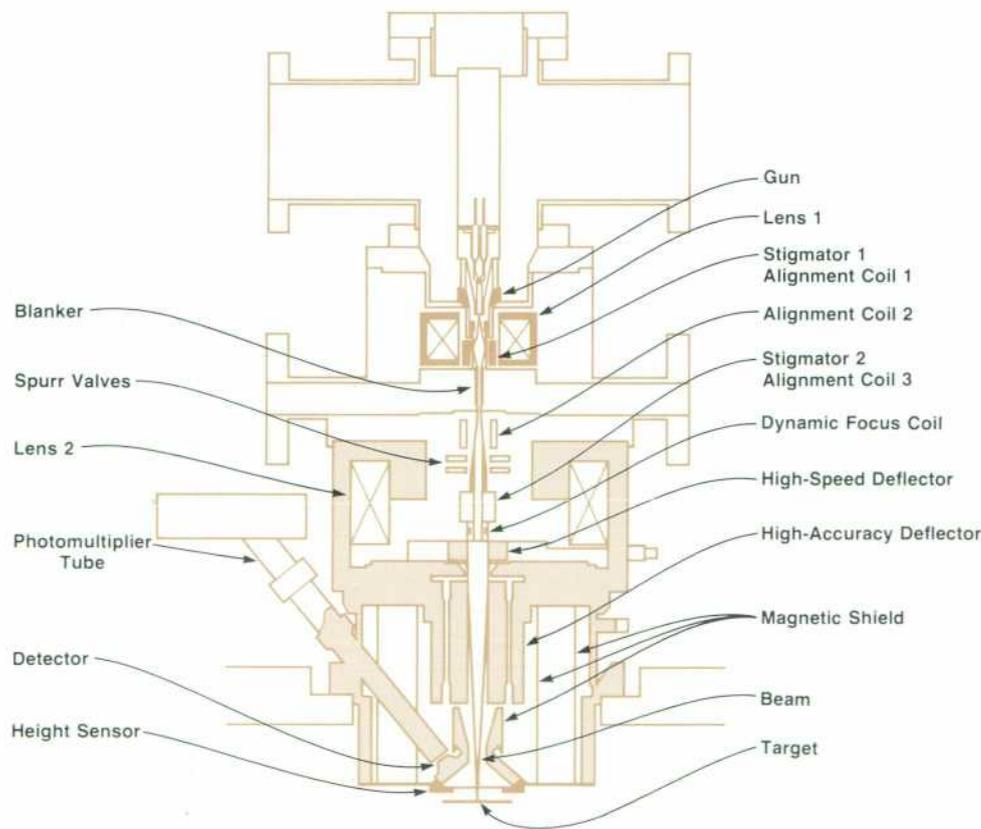

Fig. 1. *Column design for the HP electron beam lithography system uses two magnetic lenses, electrostatic beam blanking, and a field-emission cathode. The electron gun assembly can be removed from the system for servicing without exposing the lower deflection system to the atmosphere.*

Design Decisions

The cathode was a key element in the decision to make a column of this type. Assuming that a reliable field emission cathode was available, it appeared possible to make a magnifying column with a large post-lens throw (~20 cm), and obtain a fairly large, essentially aberration-free deflection field. Traditionally, field emission cathodes have had many problems that have greatly discouraged their use in practical systems. However, high brightness (>10^7A/cm²/sr) is available along with a small source size (≤35 nm).

Field emission has also been rejected traditionally as a source for all except the smallest spot sizes.[1] Apart from other practical reasons, the problem was that the specific emission of most field-emission cathodes was relatively low. Most focusing systems are severely limited by spherical aberration, necessitating the use of only a narrow cone of electrons. While this is compatible with obtaining good deflection properties, it implies a limit to the amount of current that can be accepted from the gun and cathode.

The zirconiated thermal field-emission cathode[2] now offers high specific emission (~2-4 mA/sr), good stability, low noise and long life. High specific emission is achieved because a large part of the emission is concentrated into a single forward lobe. This emitter has been extensively discussed by Swanson.[3,4]

There is no particular difficulty in obtaining an adequately small undeflected beam size. Commercial scanning electron microscopes have a spot size that is at least an order of magnitude smaller than that needed for present-day electron beam lithography. The problem is to achieve the required current and scan field simultaneously. Most, if not all, electron microscopes are limited by spherical aberration and diffraction. The current is typically limited by the cathode performance. The relative performance of various cathodes is illustrated in Table I, which gives comparative current densities into a 0.5μm diameter spot at a given beam half-angle of 2.5 mr.

Since the field emission source is capable of producing such high current densities, one finds that the design considerations are no longer the traditional ones. Early experience with columns of this type showed a steady increase in spot size with current. Space-charge effects in high-density
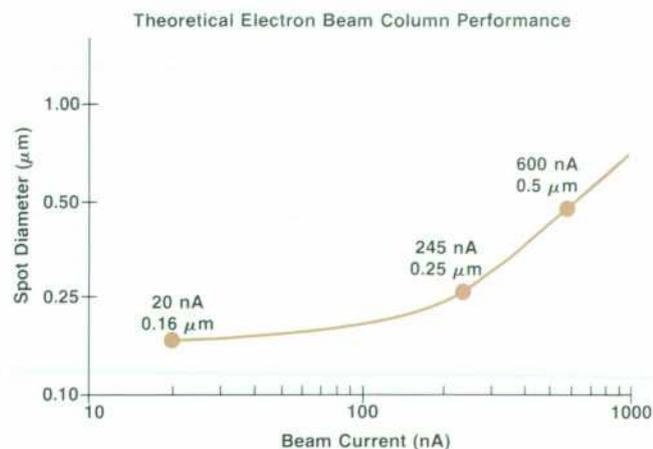
Table I

Cathode Type	Brightness (A/cm²/sr)	Beam current (nA) into 0.5 μm at 2.5 mr	Current Density (A/cm²)
Thermionic (oxide)	10^4	0.30	0.2
Short-life tungsten	10^5	3.9	2.0
LaB₆	10^6	39	20
Zirconiated tungsten TF field emitters	1.5×10^7	580	300

beams are well known, but simple calculations using the traditional theories show very small effects. Nor did the so-called Boersch effect[5] appear to offer a satisfactory explanation of the experimental data, since the measured energy spread in the beam was always less than 1 eV. It was recognized that the motion of the electrons through the column is a stochastic process and was amenable to simulation on a computer.[6]

The resulting Monte Carlo calculations gave spot sizes that were found to be many times larger than that predicted by the traditional space-charge-cloud theory. From these calculations it is clear that each electron acts on its neighbors and causes small angular deflections, which build up along the length of the column. This form of beam spreading or broadening occurs when the electrons are still sufficiently separated in space so that there is really no space-charge cloud at all. The electrons in our column at a beam current of 600 nA are in fact quite well separated along the beam axis; 25 μm is a typical spacing.

This beam-broadening theory has been well tested experimentally, and now forms the basis for the column design. In addition to the broadening, the usual beam aberrations and effects must be considered. This makes four major contributions to the undeflected spot size:

- Beam broadening
- Magnification of the source
- Spherical aberration
- Chromatic aberration.

The list does not include diffraction, which is unimportant for the relatively large spot size needed in electron beam lithography.

Analysis of the spatial broadening data shows that it decreases as the beam half-angle is increased. This is opposite to spherical aberration which increases as the third power of the beam half-angle. Since these effects work in opposite directions a compromise is required. The spatial broadening is also proportional to the length of the column and the amount of beam current. Thus, we have the ground rules defined. We need to find the shortest column and the largest half-angle beam that are consistent with optimization of the spherical aberration and the required deflector performance. In addition to this we must provide an intermediate crossover at which blanking can be done.

A theoretical performance curve for the column is shown in Fig. 2. It illustrates very clearly how the spot size grows with beam current above 50 nA. Individual contributions to the on-axis spot size at 2 mr beam half-angle are as follows:

Theoretical Electron Beam Column Performance

Fig. 2. *Theoretical plot of electron beam diameter versus current for the column used in the HP system.*

A Precision X-Y Stage and Substrate Handling System for Electron Beam Lithography

by Earl E. Lindberg and Charles L. Merja

Crucial to the performance of an electron beam lithography system is the accuracy with which the substrate to be exposed can be positioned relative to the area scanned by the beam. High accuracy requires an X-Y stage that can be precisely located and a pallet that permits mounting the substrate in a stable and repeatable fashion.

The stage performs the step-and-repeat function of the electron beam system by positioning the mask or wafer for each field of exposure of the column. The stage positioning accuracy has stringent requirements dictated by the small-geometry capability of the system.

The pallet is a frame that holds the substrate. It provides a means of locating the substrate on the stage, protecting the substrate during handling, and making electrical contact from the writing surface to the stage.

Mechanical Configuration

In HP's electron beam lithography system, the key mechanical components are mounted on a thick (50 mm) horizontal steel plate as shown in Fig. 1. The electron beam column is placed on the top face of the plate and the stage and associated interferometers are attached to the bottom face. This plate is also the top wall of the main vacuum chamber.

The stage is mounted to the bottom of the plate by means of angled beams. The beams are formed and machined at each end to allow for misalignment and to avoid stressing the base of the stage. The lower roller slide on the stage provides the Y motion and is driven by a nut supported on a shaft by eight rollers, four on each end of the nut (Fig. 2). The rollers are placed at a small angle with respect to the smooth shaft. The angle is chosen so that the rollers progress along the shaft in a helical path and move the nut 5 mm for each revolution of the shaft. The rollers are forced against the shaft by an elastic arrangement, equally loading all rollers with sufficient force so that the coefficient of friction is large enough for a stage acceleration of 1 g. This nut arrangement is compatible with the vacuum environment and the requirement for a backlash-free drive for the closed-loop servo system.

The X motion is provided by the upper set of roller slides. To eliminate carrying the drive motor on the lower slide as is often done on conventional X-Y stages, the X-drive nut and lead shaft are mounted in the base on a separate small slide and are connected by an elastic arrangement to another small drive slide mounted on the top plate of the stage. The latter small slide accommodates the Y motion of the stage and permits coupling the X motion of the stage to the X-motion drive assembly.

The servo motors that drive the X and Y motions of the stage are mounted outside of the vacuum chamber, and the shaft motion is transferred through the chamber walls via ferromagnetic-fluid shaft seals.

Three plane-mirror interferometers modified for high-vacuum service are used for measuring the stage location: one for Y, one for X and one to measure small amounts of yaw. The X and Y interferometers are mounted on the thick steel plate as close as possible to the axis of the column. The X and Y interferometers are also mounted so that measurements are made perpendicular to the plane mirrors and aligned with the column axis to reduce Abbe
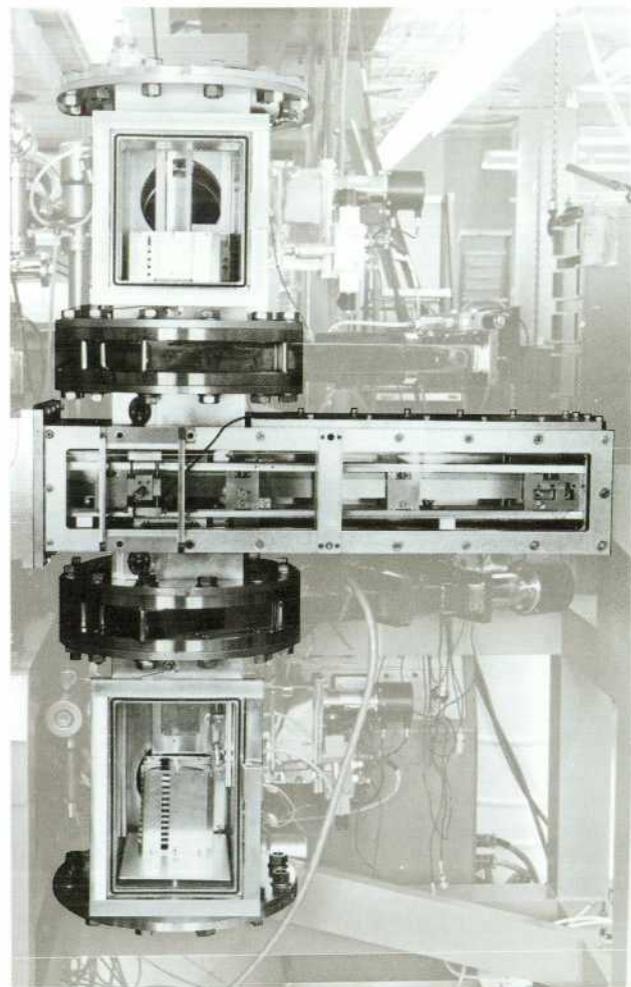
Fig. 1. *Cassette handling system for the HP electron beam lithography system. Cassettes holding pallets with substrates to be exposed are loaded into the vacuum lock chamber at upper left. The shuttle mechanism across the center of the photograph loads each pallet in turn onto the X-Y stage located to the left of the picture. Pallets with exposed substrates are then collected in another cassette located in the lower left vacuum chamber. This chamber can hold two cassettes to allow adequate time for vacuum curing of the exposed resist.*

Fig. 2. *This lead screw design, used to move the X-Y stage, works well in a vacuum environment and is easy to repair.*

error offset.* Abbe error offset causes a significant error if the alignment is not correct to within one millimetre.

Home Plate and Substrate Pallet

The accuracy of the electron beam system is influenced by the stability of the substrate position with respect to the stage. All substrates are mounted in the pallets with respect to their top surface. Square glass substrates (photomasks) are positioned by three pads and wafers are contacted by the pallet about their entire circumference. Much thought and care was taken to assure that the X and Y motions are orthogonal and coupled closely to the pallet assembly.

The X and Y interferometer mirrors are flat within an eighth of a wavelength and are mounted orthogonally within a small fraction of an arc-second to a titanium home plate. The quality of the X-Y location depends upon the flatness and orthogonality of the mirror

*Abbe error occurs when a linear measurement is performed at a distance from the dimension being measured.

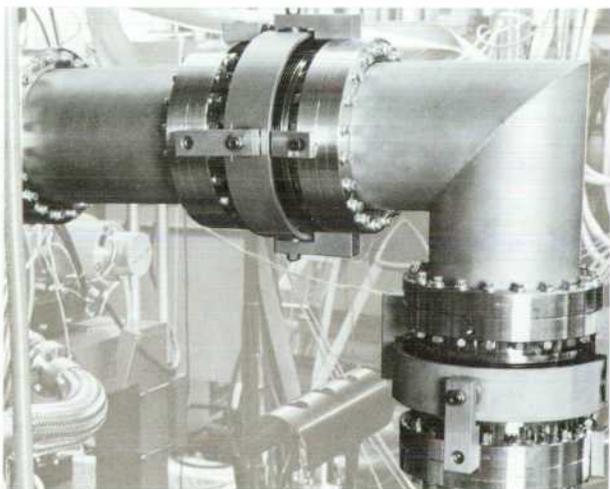**Fig. 3.** *Several of these assemblies are used in the vacuum line from the main cryogenic vacuum pump and the electron beam system to isolate the vibration caused by the pump. The U-joint design protects the bellows inside from distortion and collapse caused by atmospheric pressure and mechanical misalignment.*

assembly. This plate is secured to the X-Y stage by flexible mounts that hold the home plate slightly above the stage. This provides sufficient clearance for the mechanism that loads the pallets on the home plate.

The location of the pallet is determined by five points on the home plate: three gravity-loaded horizontal pads and two pins that engage a vee and a flat on the pallet edge. The pallet is held against these pins by a spherical contactor, whose force is applied via a toggle mechanism.

Other items located on the home plate are targets which permit calibration of the beam-deflection system in the column and a Faraday cup for beam current measurement. Height pads are also a part of the calibration system. These permit checking the operation of the height sensors, which are four capacitive transducers mounted in a square pattern on 25-mm centers attached to the lower part of the column. The height sensors determine the height and attitude of the mask or wafer so the deflection magnitude and focus of the beam can be adjusted. This allows positioning of the beam pattern within $\pm0.07\mu m$ without having to mechanically set the distance between the mask or wafer to the column within an impossible-to-attain tolerance.

Loader System

The loader system delivers pallets to the X-Y stage for exposure, using cassettes that can hold up to ten pallets each. A maximum of four cassettes can be in the loader system at any given time. The system consists of three chambers, each containing a separate mechanism.

The shuttle chamber is directly connected to the main chamber and is maintained at approximately the same vacuum. Cassettes are passed into and out of the shuttle chamber via two 25.4-cm gate valves which are used as air locks to the load and unload chambers. These two chambers have a separate, high-speed vacuum pumping system. Cassettes travel up or down in the chambers on V-shaped ways and rollers.

Cassette position is monitored by infrared light-emitting diodes and detectors. Once a pallet has been initially positioned it is delivered to the home plate on the X-Y stage by the jaw-and-tongue assembly, which operates like a slide projector mechanism. The jaw and tongue have V-shaped cuts that match with chamfers on the pallet. This allows the pallet to be lifted slightly as it is grasped by the mechanism, assuring that the pallet doesn't slide on the cassette lands or the locating pads on the home plate. The jaw-and-tongue assembly rides on ball bearings to minimize friction. The assembly has an integral linkage mechanism which actuates the clamping mechanism on the stage.

Each mechanism in the loader system is powered by an external step motor, and the rotary motion is transferred through the chamber walls via ferromagnetic-fluid shaft seals.

Vibration Isolation

The concept of fastening the critical mechanical components to a thick steel plate eliminates, to a great extent, the need of placing the electron beam lithography system in a seismically quiet environment. Vertical vibration is adequately reduced by placing simple air-filled rubber vibration isolators under each of the legs. Horizontal vibration isolation, which is more critical for a system of this type, is provided by externally pressurized air bearings placed between the floor and the vertical isolators. A cryogenic vacuum pump that is part of the system is isolated from the main system by mounting the pump to a massive iron plate in direct contact with the floor. The vacuum connection is mechanically isolated by using one horizontal and one vertical section of pipe, each section having flexible bellows at each end. A roller-bearing universal joint built around each bellows (Fig. 3) prevents its collapse when atmospheric load is applied upon pumpdown. The

resulting assembly is extremely compliant in all degrees of freedom and effectively eliminates a potentially serious vibration problem.

Acknowledgments

The authors would like to acknowledge the work of Walt Neumeyer, Bob Toda, Jack Foster and Joe Franklin who contributed much of the original loader system design. Rick Huff did the stage design and Bob Hirsch designed the cryogenic vacuum pump vibration isolation system. Edith Pope and Graham Siddall contributed to the calibration plate design and fabrication. Thanks also to those who worked so hard to fabricate, assemble and debug the entire assembly, including Ken Winkleblack, Ron Lackey, Eric Johnson, Emmett Gerrity, Glenn Weberg, and the model shop under the supervision of Len McQuaid. And finally to Rich Hoogner for the industrial engineering.

Earl E. Lindberg

Earl Lindberg attended the University of Nebraska, earning the BSEE degree in 1949. He came to HP in 1971 after serving as a senior engineer at other companies. His responsibilities at HP have included work on surveying instruments, precision mechanical engineering, and, most recently, mechanical designs for the electron beam lithography system. Earl has presented two papers on roundness measurements and is co-author of two papers on surface finish specifications. He is a member of the American Society for Metals and is chairman of the American Society of Mechanical Engineers Standardization Committee B89.3.4 on axes of rotation. Earl served in the U.S. Navy Submarine Service from 1942 to 1945. He is a native of Maxwell, Nebraska and now lives in Sunnyvale, California. He is married, has two sons and a daughter, and is interested in farming and personal transportation and communication.

Charles L. Merja

Chuck Merja came to HP in 1979 and designed the automatic loading system and did the initial production engineering on the column for the electron beam lithography system. He received the BSME degree from Stanford University in 1978 and is a member of the American Society for Metals and the American Vacuum Society. A native of Great Falls, Montana, Chuck and his wife recently moved to Sun River, Montana where they purchased a farm. Chuck enjoys skiing, flying, photography, sports, and tinkering. He intends to work on personal projects and do some consulting during the farm's off season.

Total chromatic aberration:	0.091 μm
Total spherical aberration:	0.084 μm
Geometrical spot size:	0.09 μm
Low-current spot size:	0.16 μm
Broadening at 600 nA:	0.48 μm
Spot size at 600 nA:	0.51 μm

Cathode, Field Emission Gun and First Lens

The complete gun structure is shown in Fig. 3. The cathode is operated at −20 kV with respect to the anode, providing a field of approximately 2×10^7V/cm at the tip of the cathode. The emission is controlled by an electrode placed around and just above the cathode tip. The voltage applied to this electrode causes a corresponding change in the field at the tip, and hence controls the emission current. A bias of 0 to −2 kV is sufficient to vary the current over three orders of magnitude. Since this control electrode also serves to reduce thermionic and Schottky emission, it is usually called the Schottky shield.

The gun must be a simple, stable, prealigned structure for controlling and accelerating the emission. It is desirable to get the beam to the blanker crossover without contributing significantly to the spot-size growth other than by geometrical magnification of the source size. To do this the electrode structure has been designed to reduce spherical aberration. A major contributor in this regard is often the presence of an aperture in the high-field region. In this gun the anode aperture is set in a well, almost entirely out of the accelerating field. This aperture is also used to define the beam half-angle and to prevent unusable beam current from providing unwanted beam-broadening effects. The total emission current is typically 300 μA and is reduced to the 600 nA required at the anode aperture.

The beam emerging from the gun is focused by a magnetic lens with large bore and gap, and forms a crossover 6.5 cm from the emitter with a magnification of about 1.5×. The gun is immersed in the lens field to achieve some reduction in the spherical and chromatic aberrations.[7] The lens can also be moved along the optical axis to change the magnification. Horizontal motion and tilt of the lens permit alignment of the gun and lens.

Final Lens and Dynamic Focus Coil

The final lens forms an image of the blanker crossover in the target plane with a magnification of approximately 1.8×. This lens must have sufficiently low spherical aberration and be reasonably easy to make. Low spherical aberration and long working distances do not normally go together. Given fixed conjugate points and magnification, the best lens is a large one. Clearly there are limits on how size and convenience are related. In this column a large-diameter (10 cm) upper pole piece is combined with a smaller-diameter pole piece which limits the field entering the deflector.

The dynamic focus coil is a separate, small, 20-turn winding centered in the upper part of the lens. This coil poses some interesting design problems. It must be precise so that it does not contribute unwanted aberrations that are uncorrectable. Most of all, the coil, in conjunction with the driving amplifier, must have a fast settling time. Dynamic focus-coil adjustments are made in the interval between high-speed block scans. The high-accuracy deflection circuits take approximately 10 μs to settle at a new block location and the dynamic focus must also settle within this
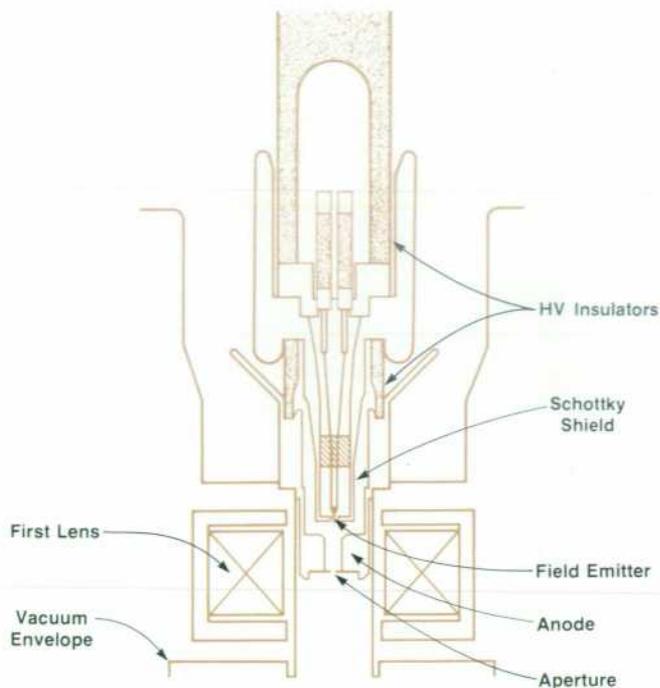
Fig. 3. *Cross-section of the electron gun and upper lens assemblies.*

HV Insulators

Schottky Shield

First Lens

Vacuum Envelope

Field Emitter

Anode

Aperture

time. The major settling problems are eddy currents in the surrounding conducting material and, since the coil is outside the vacuum, the time required for the field to penetrate through the vacuum chamber wall. Careful design has resulted in a worst-case settling time for the magnetic field of 7 μs.

Deflection System

The wide-angle precision deflector is an octopole, or eight-fold deflector (Fig. 4). It is situated in the field-free region between the lens and the target and is a means of obtaining a precisely controlled field distribution.[8] Used in its normal mode, the octopole produces a field that is very close to being uniform. The deflector has low aberrations and can be tailored to minimize one or more defects. In this design the major advantages are that it is a reasonably sensitive deflector which has low aberrations, provides a common center of deflection for both X and Y directions, and can be made precisely. Present deflectors are made with an accuracy of better than 0.01 mm throughout. The deflector is 10.16 cm long and has a 15.87-mm internal diameter. All critical surfaces are coated with gold to eliminate charging effects.

The octopole deflector can scan a full 5-mm-square field with only small dynamic corrections. Typical uncorrected distortions are close to 2 μm in a well-made deflector. Astigmatism causes the beam diameter to grow to 1.5 μm until the necessary corrections are applied.

The high-speed deflector must scan the beam over relatively small distances (~64 μm), and hence requires an accuracy of 0.1% to achieve 0.1-μm tolerances. This deflection acts like a pertubation on the main deflection field. The deflector is a simple quadrupole with the four plates arranged in a cylindrical configuration. It is located in the

center of the magnetic focusing field and is rotated to compensate for lens rotation effects. Capacitance between the plates is of key importance, and is normally less than 10 pF.

Blanker

The beam blanker was probably one of the hardest components of the system to design, and it is certainly one of the most important. To achieve patterning of the wafer or mask, the beam must be turned on and off at a rate much faster than the data rate of 300 MHz. With a pixel exposure time of 3.3 ns and a need to define the exposure level within 2%, the problem is a difficult one. Actually the toughest problem is to guarantee that the exposure falls in the right place during the transitional periods when the beam is switched. In addition, a beam-current attenuation factor of at least 10^5 is required to guarantee that unwanted exposure does not occur on the mask or wafer during settling periods.

A proprietary blanker design has been developed which meets the required performance specifications. The design uses an electrostatic deflection system to deflect the beam to a knife-edge. To facilitate stationary blanking the blanker plates are positioned symmetrically about a crossover. Viewed toward the blanker from the target, the crossover appears to be stationary when the beam is deflected toward the knife-edge. When the final lens forms an image of the crossover on the target, the image is also stationary.

Fig. 4. *Octopole assembly. This unit deflects the beam with high accuracy and is used to direct the beam to each of the 64-μm-square raster-scanned blocks in the system's 5-mm-square field-of-view.*

Mechanical Design

The mechanics of the column and its magnetic shielding are fairly complex. Because of the short length of the column, all of the parts interact. For example, the lens fields extend through the whole of the column, except for the octopole region. The structure does not permit magnetic shielding to be positioned inside these components where it would interfere with the lens fields. In addition, magnifying columns are usually difficult to shield and this one is no exception. A shielding factor of 50 dB is required to guarantee stability in an ambient ac field level of 2×10^{-7} Wb/m^2. This is achieved by using a double-layer magnetic shield, one layer of which passes through the vacuum chamber wall to shield the lower column.

In the center of the column below the blanker there are two valves. These provide a 3.18-mm clearance hole for the beam to pass through when open, but, when closed, permit the column to be split in two for servicing. The normal mode for electron-gun servicing is to replace the upper column. The spare gun and its cathode are activated and adjusted on a separate test station.

The upper column contains no O-rings and the vacuum wall is constructed of welded and brazed stainless steel. The electron gun is evacuated by two 60-l/s vacuum pumps and the blanker section by two 30-l/s pumps. The entire lower column is evacuated by the 400-l/s main-chamber ion pump.

Acknowledgments

The authors would like to acknowledge the continued support and active interest of Barney Oliver, Don Hammond, Jim Boyden, and Frank Ura. Also we would like to acknowledge the early work carried out by Patrick R. Thornton, especially with regard to the design of the first-generation system.

Thanks go to Earl Lindberg, Jack Foster, Per Gloerson, Hal Hiner, Bob Hirsch, Chuck Merja, John Spurr, Jerry Rector, Joe Franklin, Rick Huff, and Joe Kral for their help in the design and its implementation; also Geraint Owen, Steve Palermo, and Anatoly Rabinovich for their help with the magnetic shielding, the electron gun, and the field emitter, respectively; and to Emmett Gerrity, Ron Lackey and Glenn Weberg for assembling and ironing out many difficulties.

References

1. M. Drechsler, V. Cosslett, and W. Nixon, Proceedings of the Fourth International Conference on Electron Microscopy, 1958, p. 13.
2. United States Patent 3374386, March 19, 1968, Charbonnier, et al.
3. L. Swanson, "Comparative Study of Zirconiated and Built Up W Thermal-Field Cathode," Journal of Vacuum Science and Technology, Vol. 12, No. 6, November/December 1975.
4. D. Tuggle, L. Swanson, and J. Orloff, "Application of a Thermal Field Emission Source for High Current E-Beam Microprobes, Journal of Vacuum Science and Technology, Vol. 16, No. 6, 1979, p. 1699.
5. H. Boersch, Z. Phys. 39, 115, (1954).
6. T. Groves, D. Hammond, and H. Kuo, "Electron-Beam Broadening Effects Caused by Discreteness of Space Charge," Journal of Vacuum Science and Technology, Vol. 16, No. 6, November/December 1980.
7. H. Kuo and B. Siegel, "A High Brightness High Current Field Emission Electron Probe," Proceedings of the Eighth International Conference on Electron and Ion Beam Science and Technology, Electrochemical Society, 1978, p 3.
8. J. Kelly, "Recent Advances in Electron Beam Addressed Memories," Advances in Electronics and Electron Physics, Vol. 43, 1977, p 43.

Huei P. Kuo

Huei Kuo was born in Pingtung, Taiwan and attended National Taiwan University, earning a BS degree in physics in 1968. He then attended Cornell University and received the PhD degree in applied physics in 1975. Huei joined HP in 1977 and works on electron beam column designs. He designed the gun, blanker, and lenses for the lithography system. Huei has authored several papers on column design and is a co-inventor for several related patent disclosures. He is a member of the IEEE and the American Vacuum Society. Huei is married, has two sons, and lives in San Jose, California.

Timothy R. Groves

Tim Groves is a native of Chicago, Illinois and attended the University of Chicago where he earned the MS and PhD degrees in physics in 1970 and 1975, respectively. Earlier Tim received the BS degree in physics from Stanford University in 1968. He came to HP in 1978 and has worked on the design and testing of the electron beam column. His previous work experience involved electron microscopy, CRT manufacturing, and teaching physics at the U.S. Naval Academy Prep School. Tim has written eleven papers on electron optics and microscopy and is a co-inventor for three patent disclosures on electron beam devices. He is a member of the IEEE, he's married, and he lives in Palo Alto, California.

John Kelly

John Kelly is a native of Bristol, England. He attended Trinity College at Cambridge, England, receiving the BA and MA degrees in natural sciences in 1959 and 1966, respectively. He also earned the MS degree in physics in 1964 from London University. John came to HP in 1975 after considerable experience in electron beam system design elsewhere. He has been involved with several projects related to electron beam systems and is now department manager for the electron optics and mechanical design group. John has written several papers, is a co-inventor for three patents, and is a member of the IEEE. He is married, has a son and a daughter, and lives in Palo Alto, California. His interests include coaching and playing soccer, Scottish dancing, hiking, and camping. John also has designed and nearly finished an unusual home in the Sierra Nevada mountains of California.

Software Control for the HP Electron Beam Lithography System

by Bruce Hamilton

THE MANY ELECTRONIC and mechanical pieces of HP's new electron beam system are unified and offered to the operator by a large and complex software package. This package gives the operator several modes of system operation and performs event sequencing and system calibration.

Operator Interface

The operation of the electron beam system is almost completely interactive. At the top level, the user is presented with a prompt, and can request any of the operating modes (discussed below) or can type HELP. In the latter case the user is given a list of legal commands. The operator need type only enough of any command to distinguish it from the others.

The following is an example of operator-system interaction. The portion of the left column printed in black is typed by the computer. The operator's responses are printed in color. The right column tells what is happening at each step.

:RU,SKYHAWK	SKYHAWK is the entrance to the system.
Hello—hope you are having a good day.	
SKYHAWK> CONFIGURE	At this level the operator is prompted with SKYHAWK>. The operator chooses CONFIGURE mode (which manages pattern files).
CONF> HELP	Configure mode prompts with CONF>. The operator asks for a list of commands.

The user works in this style down through several levels, until there are few choices left, but the system still needs more information. At that point the system takes the initiative and prompts the operator for information. For example, here is what happens when the operator asks for an exposure:

# SUBSTRATES TO BE PROCESSED:	The number of masks or wafers of a given type.
RESIST SENSITIVITY IN MICROCOULOMBS/SQ CM:	The sensitivity of the electron resist on the substrates.
MASK SET NAME? (UP TO 6 CHARS)	The name of a file describing the layout of chips on the substrates.

Hardware Interface

The control software is built on HP's RTE (Real-Time Executive) operating system for HP 1000 Computers. That system provides convenient and controlled access to all standard peripheral devices, and we use it for that purpose. For custom-built devices it is more convenient to go "behind the operating system's back" and make our own access

directly. RTE makes this fairly easy.

Each piece of hardware has a software driver which conceals any complexity in the way the device communicates and offers a friendly protocol to the system software. For example, to move the X-Y stage to a given position requires these primitive operations:

- Tell the interferometer we are going to load an X-position.
- Load the X-position.
- Tell the interferometer we are going to load a Y-position.
- Load the Y-position.
- Tell the stage servo to track the interferometer position in both the X and Y directions.

The stage/interferometer driver performs all these actions in response to the command, Move To (X,Y). In the same operation it also returns the new stage position and status and the yaw of the stage at the new position.

These drivers constitute a layer of the system software that serves two isolation purposes. It conceals the complexity of the hardware, and it isolates the portions of the electron beam system that operate outside of the RTE operating system. Everywhere else the hardware appears to have a simple and logical instruction set, and the operating system's rules for I/O are obeyed. By staying within the operating system we have enjoyed relatively troublefree compatibility through three generations of RTE (II, III and IV), with attendant improvement in development and debugging tools.

Operating Modes

There are several modes of operating the electron beam system, each oriented toward an operator's task, and each entered by typing its name to the top level of the system software.

Expose Mode

The first mode, expose mode, is the mode in which a set of wafers or masks is written. The operator is asked for the sensitivity of the resist and the name of a file describing the layout of chips, and then the system follows these steps:

1. If necessary, the system allows the operator to load a cassette of substrates into the load airlock.
2. A substrate is carried from the cassette to the stage.
3. If this is the exposure of the second or higher level on a wafer, a locator pattern is found on the wafer, showing the wafer's position and orientation.
4. A survey is made of distortions in the substrate surface.
5. Correction factors are computed to rectify the distortions.
6. The electronics are configured for the proper exposure dosage.
7. If necessary, the pattern to be written is read from disc

into pattern memory.

8. The X-Y stage is moved to center the first chip's location under the beam axis.

9. If this is the exposure of the second or higher level on a wafer, fiducial marks are found around this chip, giving its exact position and orientation.

10. The exposure takes place as follows:
    a. The location of the first 64-μm block is computed, taking distortions into account.
    b. That address is sent to the octopole plates.
    c. The pattern is made ready at the output of the pattern buffer.
    d. The quadrupole sweep is started.
    e. While the sweep is underway the location of the next block is computed.
    f. The process from (b) is repeated until the entire octopole field of view is exposed.

11. The process from (6) is repeated until the entire substrate is exposed.

12. The substrate is unloaded from the stage and returned to the cassette.

13. The process from (2) is repeated until the entire cassette of substrates has been exposed.

14. The operator is allowed to remove the cassette from the unload airlock.

Pattern Format Conversion Mode

The circuits that the electron beam system produces are designed on other computer systems and stored on disc files in standard formats. In pattern format conversion mode the operator asks for a circuit pattern file to be converted from Electromask format* into electron beam internal format. That process is discussed in the article on page 24.

In this mode the operator can also move pattern files from medium to medium and machine to machine, and can purge files.

* Developed by Electromask Corp.

Substrate Layout Mode

This mode allows the operator to specify before an exposure which patterns are to go on a substrate, and where. The substrate can be covered with a single pattern repeated in a rectangular or circular array. In addition, the operator may ask that test patterns be substituted into the array.

The array may be shifted in any direction, as may any elements in it. The array, or any elements in it, may be expanded or shrunk by a small amount, and the exposure levels of the elements can be varied. One option provides a progression of exposure levels across the array. This is useful for finding the optimum exposure for a particular resist or process.

This mode is the only one that is not interactive. The operator prepares a file in an English-like language, describing the layout of the substrate and gives it to the system. The machine produces a picture of the substrate for the operator and a file for the machine to read at exposure time.

Scanning Electron Microscope Mode

In this mode the beam is scanned over its field of view and the intensity of backscattered electrons is displayed on a CRT. The scan is performed by the high-speed processor (HAL), and the processing of the signal intensity is done by the fiducial electronics. The system software sends parameters to HAL and the fiducial electronics in response to the setting of the operator panel.

Fig. 1 shows the operator panel. The system continuously polls the panel, and when any settings change, sends new parameters to HAL. The scan then changes on the next cycle.

The operator also has control of the stage in this mode, and can send it to an address entered by the operator or to a prestored address, or some distance from the current address. The current position can be displayed or stored for future use.

This mode is implemented as three concurrent processes. One reads the terminal, one polls the operator panel and controls the scan parameters, and one drives the stage. The

Fig. 1. *The operator panel of the HP electron beam lithography system.*
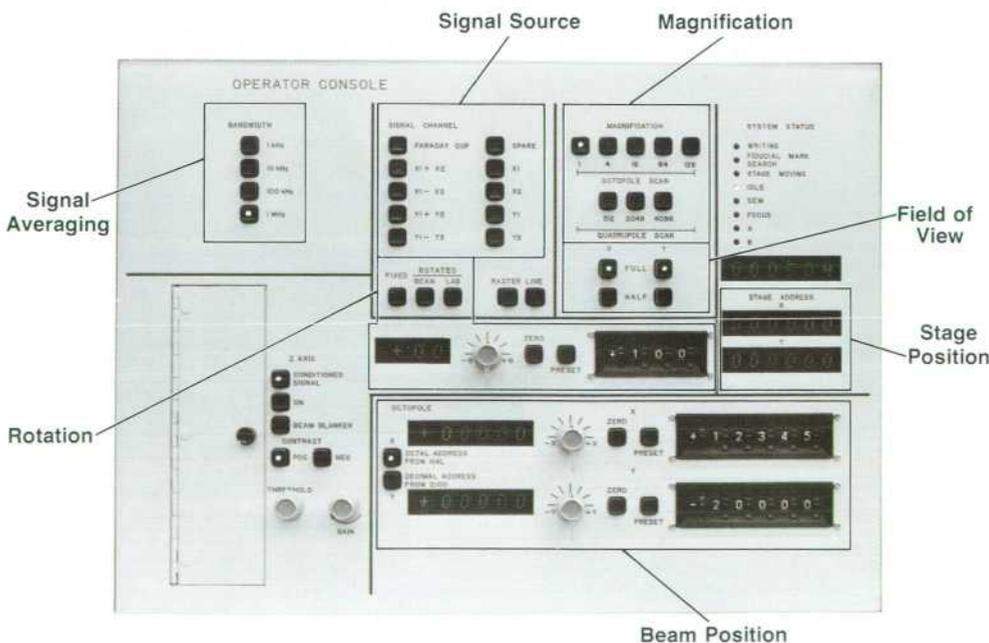
process that polls the panel runs continuously. The process that reads the terminal is always present, but is usually waiting on the terminal. The process that drives the stage is spawned only when the operator requests a stage move, and it dies when the move is complete.

Stage moves are complicated by the fact that we want the scan to stop while the stage is moving. Not only do we not wish to see a blurred picture, but sometimes we want to look on either side of a region we do not want to sweep the beam across. When the operator wants to move the stage, this is the sequence of events:

1. The operator requests a stage move at the terminal.
2. The terminal process interprets the command and spawns the stage process, passing it the desired address.
3. The stage process comes to life and sends a message to the scan process, requesting permission to move the stage.
4. The scan process, on receiving the message, waits for the current scan to complete, then stops scanning and sends a message to the stage process, granting permission to move the stage. It keeps polling the panel.
5. The stage process moves the stage and displays the new stage address at the operator's terminal.
6. The stage process sends a message to the scan process saying that the stage has finished moving.
7. The scan process resumes the scan.
8. The stage process dies.
9. The terminal process issues a new prompt and waits for the operator to type something new.

When the operator types END, the terminal process sends a message to the scan process. That process waits for the current scan to complete and then dies. The terminal process then surrenders control to the top-level program.

This mode provides both a highly linear scanning electron microscope and a valuable instrument for testing the machine itself.

Calibrate Mode

In this mode the operator can ask the system to run any of the calibration routines. At the end of a calibration the results are displayed, and if the operator so chooses, the system parameters are updated with the new results.

This mode serves not only to collect the calibration operations into one place for operator convenience, but also to protect the calibration data. This data can be changed only through overt action by the operator, and only in a mode intended explicitly for updating those parameters—this mode and system parameters mode, discussed below.

System Parameters Mode

Global information about the system is maintained by a package called the system information facility. This information includes such things as the correction function for the octopole, the stage address of the substrate center, and the beam voltage. It is kept in a central location so that when it changes only a single copy needs to be updated.

The parameters can be read and written by system programs, and also by the operator. The operator's access is useful for looking at the state of the system and for testing.

Unlike the usual forms of shared data, the system information facility provides substantial control over the information it manages. Access to a data item is made by an explicit call to an access manager giving the name of the data item and the name of the requesting program (or the operator's initials). This makes it unlikely that the program will refer to a global item by mistake, or to the wrong item, and it shows us who last changed each item. The information is kept in a push-down stack, giving us a short history and letting us undo changes if necessary. A data item can be locked; when any unauthorized program asks for access to it, that program is frozen and the operator is notified. Finally, the facility can be asked to keep a log of all requests, together with the name of the requestor, the value of the data item, and the date and time. The operator can enter comments into this log. The log and the ability to lock an item are powerful debugging tools.

Start/End Mode

Upon system startup or termination a user-definable collection of programs is run. This lets the user update records, open or close files, etc. It allows us to start and finish with the hardware in a known state.

Acknowledgments

The software team has put a tremendous amount of thought and work into this project, not only in the package described here, but also in previous versions and in support of hardware testing. Credit must go first to Sam Gebala, Nancy Kendzierski and Jim Stinger, who conceived and built the first version of the system. Its structure and much of its code stand today. The present event sequencing and data paths in the expose mode are the work of Faith Bugely, Daffydd Jones, and Sherry Ramsey. The pattern data path was built by Mike Cannon, Lee Casuto, Bob Lewis, and Randy Shingai. Test plot programs were written by Tim Crew and Terry Ligocki. System calibration is the work of Jim Hsu, Ian Osborne, and Bob Schudy, with considerable help from John Eidson and Geraint Owen. The hardware interface owes much to Walt Garms, Bob Lewis, and Rich Yamamoto. Doug Earl designed the system information facility, and John Belleville designed the mask layout package.

Bruce Hamilton

Bruce Hamilton is software project manager for the HP electron beam project. He received his BS degree in electrical engineering from Swarthmore College in 1970 and his MS in computer science from the University of California at San Diego in 1973. He joined the HP electron beam project in 1978 as a software engineer with several years' experience in software development for satellites and control systems. He's a member of ACM and IEEE. Bruce was born in San Diego and now lives in Menlo Park, California. His major interest, folk dancing, consumes about five nights of each week, and he has taught English and Scottish dancing at workshops in several California cities. He's also interested in folk and classical music and high-fidelity equipment.

Pattern Data Flow in the HP Electron Beam System

by Michael J. Cannon, Howard F. Lee, and Robert B. Lewis

I N HP'S ELECTRON BEAM LITHOGRAPHY system, the task of the format conversion path is to transform a pattern that is specified in some external form into a stream of bits that will control the electron beam exposure at each point, or pixel, on the target mask or wafer. The number of bits required for a 5-mm×5-mm square (one pass, or octopole field of view) is about 10^8 (12.5 megabytes) and is about 4×10^{10} bits (5.5 gigabytes) for a 10× mask the size of a full 13-cm wafer. The sheer bulk of the pattern data creates problems with its storage and transfer, and the speed of the transformation is also critical.

Data storage and transformation from the external representation is done by a computer system that includes an HP 1000 F-Series Computer with HP 7906 and HP 7925 Disc Drives, an HP 7970 Tape Drive and about a million words of main memory. The conversion software runs under HP's RTE-IVB operating system and is written mostly in C and ALGOL.

Fig. 1 shows the major steps in the process. The first step involves copying the pattern data from a nine-track magnetic tape to files on the 7925 Disc Drives. This is done under control of an access method software package that maintains general housekeeping information about the patterns. The remainder of the path is designed to accommodate two basic factors: the sequence of the raster-scan exposure process and the pattern data compression format.

The hierarchy of pattern exposure is shown in Fig. 2. A pixel is the smallest addressable area that can be exposed. A block is a quadrupole field of view and is 128 pixels square. A stripe is a column of up to 78 blocks, and a pass (octopole field of view) is a row of up to 78 stripes. The order of exposure is pixels within a block, blocks within a stripe, stripes within a pass, and passes within a chip.

The input format for the initial implementation is the Electromask optical pattern generator format.* Although a number of external formats are used in the industry to describe patterns, this format is widely used within HP. The Electromask magnetic tape input format specifies a pattern as a series of rectangles, or flashes, of varying locations, widths, lengths and rotations.

Clipping

In the second step of Fig. 1, Electromask flashes are first broken up (clipped) into objects that are contained within stripes. Details of the Electromask format are isolated upstream of this point; this will make accommodation of alternate input formats a relatively simple task. When the Electromask flashes are square with the X-Y axes the intermediate objects are also rectangles whose sides are square with the axes. Otherwise, the clipped objects are a restricted form of trapezoid that have left and right edges parallel to

*Developed by Electromask Corp.

stripe boundaries. Strictly speaking, this special form would not be necessary since the trapezoids could be replaced by a set of rectangles of unit pixel width, but the trapezoids provide a more compact representation for the intermediate form.

As each object—rectangle or trapezoid—is produced by the clipping routine it is placed in a bin according to its pass and stripe. Each bin has an internal memory buffer, and when the buffer for a stripe fills up, it is written out to a file on one of the 7925 Disc Drives. Bin buffers need to be large enough for at least a disc sector (128 words), and for efficiency we would like them to be several times larger than that. RTE-IV gives user programs access to memory outside their normal 32K-word address space by providing functions to remap parts of the virtual address space to an extended memory area (EMA) of the physical memory. The bin buffers are kept in an EMA partition, so the number and sizes of the buffers are limited not by the 16-bit virtual address space but by the size of the memory partition the conversion program runs in.

Bin buffers for each stripe of a full pass require at least 10K words (78×128 words). For a pattern that spans only a few passes this is easy to accommodate, but in the case of a 10× mask that may span all 441 passes we would need more

EBX = Electron Beam External Format
EBI = Electron Beam Internal Format
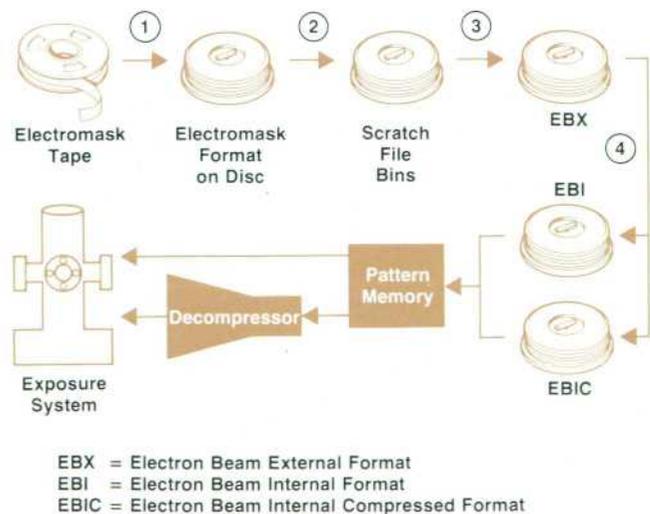EBIC = Electron Beam Internal Compressed Format

Fig. 1. *The format conversion path in HP's electron beam lithography system converts an integrated circuit pattern data file into a bit stream that turns the electron beam on and off to write the pattern on a resist-coated substrate.*

than four million words of buffer space. In this case we do the clipping by pass and stripe as before but use coarser bins (say, by pass only) as we write the objects to disc. An additional step is used to make the finer bin separations as a preliminary to the next step of the conversion. In this next step, the rectangles and trapezoids for each pass and stripe are put in a format called electron beam external (EBX), which is the input to the next phase of the process, CRUNCH.

Electron Beam Internal Formats

The average feature size for most patterns is large compared to the pixel size, especially for 10× masks, so it is more efficient to store patterns in a compressed form. In this form, called electron beam internal compressed (EBIC), patterns are represented as a set of rectangles, which are converted in real time to a stream of bits that controls the exposure. The decompressor, the hardware that does this conversion, is described later in this article.

EBIC rectangles are always described in terms of integer numbers of pixels. The compression scheme works within stripes and the restrictions on EBIC rectangles are that the sides be parallel to the X and Y axes, that each rectangle be contained within a single stripe, and that within stripes rectangles be ordered by their leading edge (lowest Y-value). Rectangles may not overlap.

As pattern geometries become finer, compressed format begins to lose its advantage over a pure bit map representation. The crossover point comes at about 2.4 million compressed format rectangles in a 9984×9984-pixel octopole field, which corresponds to a checkerboard with features 4.6 pixels on a side. Although a checkerboard is not a typical situation, patterns with angled figures do tend to generate a lot of tall, thin compressed-format rectangles. Beam and pixel sizes and pattern geometries can be expected to shrink, and it remains to be seen whether it will be advantageous to use an uncompressed format for some patterns.

Conversion to Compressed Format

In the part of the data conversion known as CRUNCH (step 4 in Fig. 1), three things are done:
1. Trapezoids are translated into a series of rectangles.
2. Overlaps are eliminated.
3. Rectangles are converted to EBIC format.
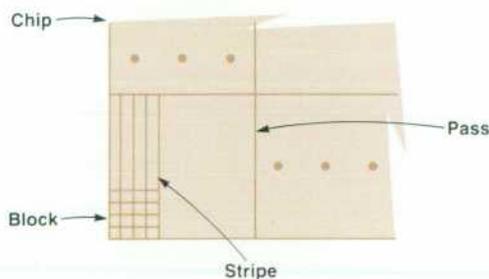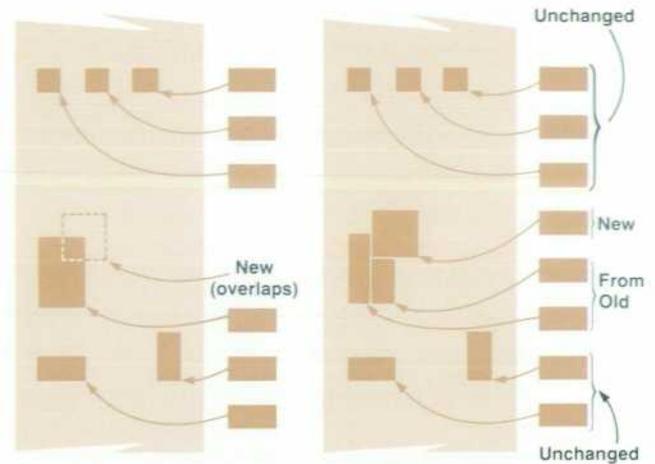A rectangle is described by an X-start, an X-end, a Y-start

Fig. 3. *An example showing how a new rectangle that overlaps rectangles already stored causes the splitting of the new and/or old rectangles to eliminate the overlap.*

and a Y-height. Y-start and Y-height are fourteen-bit quantities, while the X dimensions require seven bits each for a total of 42 bits per rectangle.

During conversion from EBX to EBIC all figures within each stripe are arranged in a list in ascending order by leading edge. The list of rectangles is maintained in EMA. As each new figure is inserted into the list, it is checked for possible overlap with figures already in the list. If an overlap exists, then either the figure in the list or the new figure is split up to eliminate the overlap. Cases of overlap are categorized into one of 16 types and are handled accordingly. One example of an overlap is shown in Fig. 3a. The new rectangle being added is shown in dashed lines. In Fig. 3b the overlap is resolved by splitting the existing figure into two and the new figure is inserted in the list.

Because the fragments of the split rectangles are parts of a rectangle already in the list there is no overlap with other rectangles in the list so they can be put in the list without further checks. The new figure must be checked against the rest of the list for overlap.

EBIC rectangles are written to a pattern disc file and from there into the 128-bit-wide electron beam pattern memory in sets of three. Each rectangle is represented by 42 bits, so three rectangles fit into 128 bits or eight 16-bit computer words.
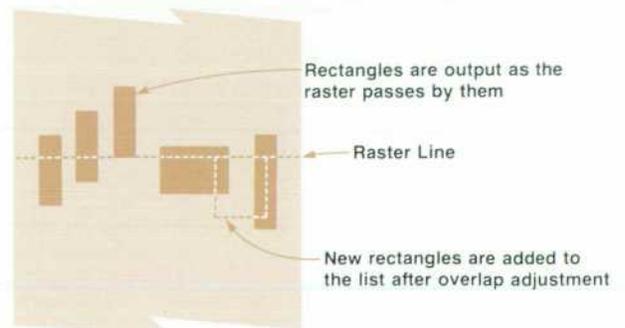

Fig. 2. *The order of pattern exposure is pixels within a block (128×128 pixels), blocks within a stripe (a column of up to 78 blocks), stripes within a pass (a row of up to 78 stripes), and passes within a chip. Each block of 128×128 pixels is 64 μm square.*

Fig. 4. *Another algorithm for overlap removal uses a raster scan method.*

As we begin to deal with larger numbers of compressed format rectangles per pass the present algorithm for overlap removal may prove to be too slow. The number of comparisons between rectangles is about $n^2/4$ where n is the number of rectangles in a stripe. A new algorithm has been developed and will be implemented when needed. This algorithm requires only about $n^{3/2}/78$ comparisons. This procedure requires that the rectangles first be sorted by Y-start, but the time needed for this is only a linear function of the number of rectangles. A coarse radix sort is followed by a finer distribution count sort.[1] The rectangles for an entire stripe can be kept in EMA buffers so no extra I/O is used for the sort. To check for overlaps a raster is passed from the top of the stripe to the bottom, and as this is done a list is maintained of the rectangles crossed by the raster (see Fig. 4). As the raster is moved down, new rectangles touched by the raster are checked for overlap with figures already in the list and are added to the list. Rectangles in the list that are no longer crossed by the raster are removed from the list and output. This gives a stream of nonoverlapping rectangles ordered by Y-start locations. The order is descending, rather than ascending as eventually presented to the decompressor. The list of current rectangles can have no more than 128 elements, and can be kept within the virtual address space of the conversion program so that checks within the list are quicker than if the list were kept in EMA.

The Decompressor

The decompressor takes the 42-bit rectangle declarations of EBIC and generates the corresponding 128-bit wide stripes. Since the decompression has to be done in real time, a parallel approach is used to generate the required bit pattern. A bank of 128 14-bit presettable up/down counters is used to construct the bit pattern. The counters are connected as presettable down counters. The carry-out signal for each counter remains high if the content of the counter is nonzero.

A block diagram of the decompressor is shown in Fig. 5. The 128 data counters and the Y-address counter are initialized to zero at the beginning of each stripe. Each input 42-bit rectangle declaration is separated into its four com-
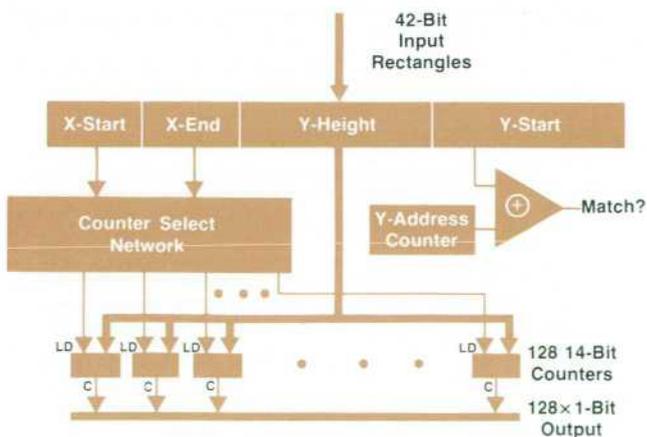

Fig. 5. *The decompressor takes rectangles in electron beam internal compressed format and generates 128-bit-wide stripes.*

Robert B. Lewis
Bob Lewis came to HP in 1973 with 14 years' experience in manufacturing systems software development. He's done software development for computer-aided design and common accounting systems, and for the pallet loader and pattern data conversion systems of the HP electron beam system. A graduate of Stanford University, he received his BSEE degree in 1957 and his MBA degree in 1959. Bob is a native of Santa Barbara, California. He's married, has seven children, lives in Los Altos Hills, California, and enjoys skiing and running.

Howard F. Lee
Howard Lee helped design the high-speed pattern data generator for the HP electron beam system. With HP since 1975, he's served as a development engineer and project leader for disc controller design and has been involved in the development of new data storage technology. He graduated from Cornell University in 1973 with a BSEE degree, received his Master of Engineering degree in electrical engineering and computer sciences from the University of California at Berkeley in 1975, and in 1980 received his MBA degree from the University of Santa Clara. Howard was born in Philadelphia and now lives in Cupertino, California. He's married, has a son, and enjoys working on audio equipment, bicycling with his family, cooking spicy foods, and working on Porsches.

Michael J. Cannon
Mike Cannon graduated from the University of Oregon in 1973 with a BS degree in computer science, and received his MS and PhD degrees, also in computer science, from the University of Illinois in 1977 and 1979. With HP since 1979, he's developed software for an OROM writer and for electron beam pattern data conversion, and served as a member of the HP electron beam task force for data conversion. He's co-author of a paper on multisensor navigation. Born in Moscow, Idaho, Mike is a cook, a runner, a cyclist, and a skier, and lives in Menlo Park, California.

ponent fields. The Y-start field of the current rectangle is compared to the value of the Y-address counter, which points to the current row—the next 128-bit data pattern to be output. If there is no match a 128-bit data pattern word from the counter carry outputs is sent to the shifter/buffer, the Y-address counter is incremented, and all the data counters with nonzero contents are decremented.

This continues until a match exists between the Y-start field of the next rectangle and the Y-address counter. Then the data counters from the X-start to the X-end are loaded with the Y-height. The other data counters' contents are unchanged. A new rectangle declaration is then fetched from pattern memory and is checked for a match against the Y-address counter. This sequence continues until the Y-address counter reaches the end of the stripe.

Reference

1. D.E. Knuth, "The Art of Computer Programming, Vol. 3: Sorting and Searching," Addison-Wesley, 1973.

Calibration of the HP Electron Beam Lithography System

by Faith L. Bugely, Ian F. Osborne, Geraint Owen, and Robert B. Schudy

THE HP ELECTRON BEAM LITHOGRAPHY SYSTEM has a specified writing accuracy over a 125-mm-square substrate of ± 0.07 μm. Calibration of the X-Y stage and the electron beam deflectors is necessary to meet the positional accuracy requirements. Calibration of the deflectors is also necessary to maintain the beam spot diameter of 0.5 μm over a 5-mm-square field. There are five major factors that make such calibration necessary.

- Nonorthogonality and curvature of the X-Y stage interferometer mirrors
- Errors in interferometer beam alignment
- Vertical motion of the stage as it moves horizontally over its ways
- Electron optical aberrations of the octopole deflector
- Gain and offset variations, which change as functions of writing speed, in the high-speed quadrupole deflection circuits

The primary objective of the calibrations is to map the various fields of view onto the mask or wafer in a way that minimizes butting errors between the fields. Other objectives are repeatability of the written patterns in spite of mechanical and electrical variations, and absolute accuracy of the electron beam position relative to the mask or wafer.

The calibration procedures must be fast to reduce downtime. This rules out any manual scheme, hence the one used is completely automatic. Besides speed, automatic operation ensures that the calibration process is repeatable and operator-independent.

One of the fundamental calibration problems is that no two-dimensional measurement standard of sufficient accuracy is available. This problem has been solved by producing such standards at Hewlett-Packard using the electron beam system. The standards consist of 125-mm-square plates on which a matrix of marks has been written, together with an accurate description of the positions of the marks. The isotropic scale of this description has been made consistent with the one-dimensional length standard of the United States National Bureau of Standards.

Calibration of the electron beam system is a sequential process: the stage is calibrated relative to the calibration plate, the octopole deflector is calibrated relative to the stage, and then the quadrupole is calibrated relative to the octopole.

During the exposure of a substrate, the calibration functions described here are applied. In addition, it is necessary to measure and correct for stage height, tilt, and yaw fluctuations, together with substrate distortion if the second or higher level of a semiconductor wafer is being written directly. These additional corrections are not described in this article.

Calibrating the Stage

If one views the stage as a rigid body, its position in space has six degrees of freedom. These degrees of freedom are represented in terms of X, Y, and height positions, and pitch, roll, and yaw angles. All are sensed, and the X and Y positions are servo controlled.

The stage calibrations determine the calibration functions which map the sensed values for X, Y, and yaw into the true X, Y, and yaw values. The stage calibrations also characterize the pitch, roll, and height variations in the stage as it moves in X and Y.

The six functions necessary to calibrate and describe the stage, one for each degree of freedom, are described as polynomials in the X and Y coordinates of the stage. The stage is calibrated in X and Y by fitting polynomials (P_x and P_y) to the difference between the true stage coordinates (x_t, y_t) and their apparent values (x_a, y_a). Thus $x_t = x_a + P_x (x_a, y_a)$ and $y_t = y_a + P_y (x_a, y_a)$. Because the calibration functions vary slowly, and the calibration is small compared to the stage dimensions, $P_x (x_a, y_a) \approx P_x (x_t, y_t)$. That is, the difference between the apparent and calibrated stage coor-
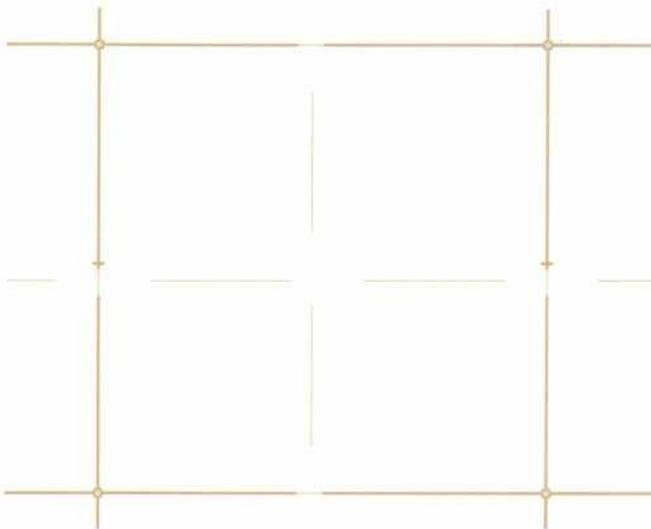
Fig. 1. *Magnified portion of the calibration plate. This array of fiducial marks is used to map the area covered by the X-Y stage for calibration of stage position and yaw.*

dinates does not matter with regard to the calibration functions. Then the calibration functions can be inverted so that $x_a = x_t - P_x(x_t, y_t)$ and $y_a = y_t - P_y(x_t, y_t)$. These expressions are used by the stage calibration software to obtain corrected values from desired true values.

The X-Y stage calibration takes place in three phases, a bootstrap calibration, a scale correction, and a routine calibration. The bootstrap calibration simultaneously calibrates the stage and determines the locations of fiducial marks on a stage calibration plate. A second calibration determines the average scale error in the stage. The results of the bootstrap (described later in this article) and scale calibrations are combined to produce a correct description of the location of the fiducial marks on the calibration plate. This plate and its description are used as the two-dimensional metrology standard in a simpler procedure used for routine calibration of the stage.

Calibration Plate

The main design requirement for the calibration plates is geometric stability. Geometric stability during the bootstrap calibration procedure is the basis of that procedure. The scale calibration depends on the stability of the plate, because the plate is used to transfer the NBS measurements to the electron beam system. After the plate is calibrated, it must also exhibit long-term geometric stability to preserve the accuracy obtained with these two procedures.

The calibration plate, which is supported at three points to minimize mechanical deformation, has a long thermal time-constant when it is under vacuum in the system. A thermal expansion as small as one part in ten million will degrade the calibrations. In the absence of a means to control or monitor the calibration plate temperature, selection of a substrate material with a very low thermal expansion coefficient is the only way to maintain plate dimensions. The substrate material for the X-Y stage calibration plate is a ceramic glass that exhibits a thermal expansion coefficient within ±0.05 ppm/°C near room temperature.

The stage calibration plate is also used to determine the pitch, roll, and height variations in the stage, so it must be very flat. Its working surface is polished to the quality of a good optical flat. Maintaining this flatness in spite of gravitational sag and the stresses introduced by the mechanism that holds the plate to the stage requires that the plate be as thick (typically 6.35 mm) as the machine will accommodate. In some circumstances special holders are used to allow sufficient thickness.

A pattern of gold fiducial marks on a chromium background is fabricated on the flat surface of the calibration plate using standard thin-film technology methods. A portion of a calibration plate is shown in Fig. 1. The stage calibration uses three sets of features in this pattern. The coarsest lines in the pattern form the coarse-scan fiducial marks. These marks are easily found by the fiducial-mark finding routines (see box on page 34), even though large uncertainties exist in the position of the marks. The arms of the coarse patterns are interrupted at their intersection by an octagonal beam-shape-measurement structure (Fig. 2), which is not used in the stage calibration. Within this octagon lie fiducial marks with 5-μm-wide arms. The X-Y stage calibration depends upon these marks.

Midway between each of the large fiducial arms are lines broken where they would cross other perpendicular lines. These interrupted crossings are optical fiducial marks. The two segments of the interrupted line serve as alignment marks pointing to a segment of the unbroken line between them. Some of the stage calibration plates are sent to the NBS, where the distances between these marks are determined by semiautomatic one-dimensional measuring machines of known scale. These same distances are measured again by the electron beam system to determine the isotropic scale error.

Bootstrap Calibration Procedure

The essence of the bootstrap calibration procedure is to use the same measuring machine (X-Y stage, interferometers and electron beam) to measure the same object in different orientations. When the object is moved with respect to the stage, apparent differences in measurements are reflections of the inconsistencies in the measuring machine because the object, the stage calibration plate, is the same.

If the substrate could be placed on the stage in arbitrary positions and orientations, this method would make it possible to calibrate the stage in every respect except scale, which must, of course, be derived from a standard. The stage will accommodate maximum-size square substrates in four positions that differ by 90° rotations. As a result the bootstrap calibration procedure cannot detect, and hence cannot correct, fourfold rotationally symmetric errors in the stage. Increasing the number of different orientations of the calibration target on the stage increases the order of symmetry of undetected errors. Applications requiring calibration of less than the full square area of the stage can use the same bootstrap software system with translations and/or more rotations.

The calibration target bears a square matrix of 25 rows of 25 replications each of the fiducial mark pattern. The 625 patterns are measured automatically using the beam and the fiducial mark system. Using the stage, each mark is moved beneath the beam axis and the octopole is used to

locate the mark relative to its expected position under the axis. The stage position and beam deflection at which each mark is found are recorded. The calibration plate is removed from the machine, rotated, reinserted, and the apparent positions of the fiducial marks recorded. This procedure is repeated for all four 90° rotations.

Computer programs are used to accept the data from these views and determine not only the positions of all of the marks on the calibration plate, but also calibration functions for the stage. In the example above, this involves the determination of 1250 fiducial mark coordinates and 50 stage calibration function parameters. The calibration can be viewed as a minimization of the apparent calibration error as a function of these variables. The quantity minimized is the sum of the squares of the distances between the apparent and model-predicted locations of the fiducial marks, in all of the orientations.

The details of the algorithm are beyond the scope of this article, but the results of the calculation are straightforward and the important point is that, while the results of the calculation are identical to those acquired by the minimization of squared-distance error, the calculation is direct rather than iterative, and can analyze the bootstrap stage calibration measurements in less time than it takes to collect the data.

Routine Stage Calibrations

In contrast to the bootstrap procedure, which requires rotating the calibration plate on the stage, the routine calibration is readily integrated into the processing stream by substituting a previously calibrated calibration plate for a substrate to be written. The routine calibration can restore more effectively the metrology of the stage than the bootstrap procedure can determine it absolutely, because the routine calibration, which is based on the calibration plate as a standard, has no undetected error components such as the fourfold rotationally symmetric error component of the bootstrap procedure.

The stage rotates slightly as it moves along the ways, and this stage yaw is a nonrepeatable function of the stage position. The yaw angle must be determined each time the stage is moved so that the octopole and quadrupole deflection can be rotated to match the stage. This yaw angle, the angle between the octopole coordinate system and the stage coordinate system, is also measured and corrected in the octopole calibration. The stage calibration includes components of the general system calibration that are dependent on stage coordinates. If the stage-dependent component of the yaw angle is represented as a polynomial in stage coordinates, then the constant term belongs in the octopole calibration because it does not depend on stage coordinates. In other words, the average yaw angle is part of the octopole calibration, and the stage-dependent yaw variations are part of the stage calibration. Suppressing the constant term in the stage yaw calibration has the desirable consequence of making the yaw angle reported by the stage driver independent of the particular octopole calibration at the time the stage is calibrated.

When the two translational degrees of freedom of the stage are calibrated relative to the beam axis, the angle of the stage relative to the octopole can be measured directly by moving the stage to a number of positions in the octopole

---

Software for Octopole Calibration

The calibration of the octopole (including deflection, astigmatism correction, and focus correction via the dynamic focus coil), is performed automatically by a program that is parameter-driven. This allows flexibility in specifying to the program such data as target dimensions, target construction, initial target location and various search and measurement criteria (scan lengths, etc.). It is assumed that approximations for the major terms (A_{10}, A_{01}, B_{10}, B_{01}) of the deflection function exist. A primitive version of the program computes these.

The program reads from mass storage the standard run parameters and allows the operator to make any changes dynamically. Using the stage, the program then moves a calibration target to the beam axis. The fiducial scan program is invoked to locate the target, and after calculation of the actual offset from the beam axis, the target is relocated by a stage move.

The program then repeats the following sequence of operations for each coordinate point within the field of view.
1. Using the stage, move the target to the new coordinate point.
2. Using the fiducial scan program, locate the target at that point. The fiducial scan program calculates the octopole voltages for the search pattern and applies any known corrections for astigmatism and focus errors.
3. Measure the beam spot diameters at the coordinate point by scanning the beam over the target. If the major axis of the ellipse is greater than a given tolerance, a through-focal series is performed iteratively to correct the beam spot size. The through-focal series is a sequence of beam spot measurements taken at various settings of the dynamic focus coil distributed about the presumed focal point of the beam. The results of the series are used to compute the corrections to be applied to the stigmator (V_a, V_b) and to the dynamic focus coil (V_f). The fiducial scan program is used to locate the target with the corrected beam.
4. The values used for deflection (V_x, V_y), astigmatism correction (V_a, V_b), and focus (V_f) are recorded in mass storage with the point coordinates.

Upon completion of the sample measurements, the file containing the stored values is read by the program and functions are fitted to the deflection, astigmatism, and focus data values. The new function coefficients are then used to update the system data. A summary of the run including input parameters and results is produced.

---

field of view and computing the average angle between their apparent coordinates in the octopole field and their known stage coordinates. This angle measurement is repeated with each of the fiducial marks in the field, and a smooth angle function fitted to the stage coordinate data. This function ($P_a(x, y)$) is a polynomial similar to $P_x(x, y)$ and $P_y(x, y)$.

This simple procedure is complicated by the fact that the stage yaws during the calibration. The solution is to rotate the calibration data to a standard apparent yaw angle, and then, when the yaw angle is reported, to add to the smooth angle function a correction A_a derived from the deviation of the X' laser from its expected value. The two X interferometers normally differ from each other by $<1 \mu m$, so a satisfactory approximation is that the X and X' lasers

are identical. This yaw reference condition will not be free from yaw, but the yaw calibration will correct for any yaw present in this standard approximation. In a manner similar to that for x_t and y_t above, the apparent yaw angle $a_a = a_t - P_a(x_t, y_t)$.

The angle calibration data is collected using the same calibration plate in only one orientation. For each fiducial mark on the target, the stage is moved to a number of different positions in the periphery of the octopole field of view. Let $\mathbf{d_{mp}}$ be the deflection vector for the mth fiducial mark in the pth position. Let $\mathbf{D_p}$ be the mean vector over fiducial marks of vectors $\mathbf{d_{mp}}$. An estimate A_m of the yaw angle of the stage at mark m relative to the mean yaw angle over the marks is given by

$$A_m = \frac{\sum_p \left[(\mathbf{d_{mp}} - \mathbf{D_p}) \times (\mathbf{d_{mp}}) \right]}{\sum_p \left[\mathbf{d_{mp}} \bullet \mathbf{d_{mp}} \right]} \qquad (1)$$

If A_{ma} is the apparent yaw angle derived from the X and X′ interferometers at mark m, let $\Delta a_m = a_m - A_{ma}$. Then, fitting a polynomial $P_a(x, y)$ to the Δa_m as a function of the stage coordinates x and y of point m, one obtains $a_t = a_a + P_a(x, y)$.

In summary, when the stage reports its position, it has an X value that is the sum of the value from the X interferometer and the X-calibration function. The stage reports a Y value equal to the sum of the Y interferometer value and the Y-calibration function evaluated at its current position. The angle is computed as the sum of the stage yaw calibration function and the current apparent stage yaw as computed from the difference between the X and X′ laser interferometers.

Octopole Calibration

The octopole deflector deflects the electron beam within a field that may be as large as 5 mm square. Within this area, the placement of the beam must be within 0.07 μm of its required position, and its diameter must be maintained at 0.5 μm or less. Calibration of the octopole is necessary to fulfill these requirements.

To a close approximation, octopole deflection is a linear function of applied voltage, and the maximum error due to nonlinearities is about 2 μm over a 5-mm-square field of view. However, this level of accuracy is not sufficient to meet the electron beam system specifications, and therefore the nonlinearities must be characterized so that they can be compensated during exposure. This characterization is done during the octopole calibration procedure.

Without octopole calibration, not only would beam misplacement occur, but also the beam diameter would vary over the field of view. Typically it can increase from 0.5 μm at the center of a 5-mm-square field to 1.5 μm at the edges. The other purpose of calibration is to characterize the electron optical aberrations that cause this increase in diameter so that they too can be compensated during exposure.

The octopole has four entry ports. Two of these, V_x and V_y, are deflection voltages; their major effect is to deflect the electron beam to a position (x,y) in the writing plane. However, another significant effect of V_x and V_y is the introduction of aberrations, thereby enlarging the electron beam diameter.

The remaining two octopole entry ports are the stigma-

Table I
Octopole Plate Voltage Components

Plate	Total Applied Voltage
1	$V_x + \alpha V_y + V_a$
2	$\alpha V_x + V_y + V_b$
3	$-\alpha V_x + V_y - V_a$
4	$-V_x + \alpha V_y - V_b$
5	$-V_x - \alpha V_y + V_a$
6	$-\alpha V_x - V_y + V_b$
7	$\alpha V_x - V_y - V_a$
8	$V_x - \alpha V_y - V_b$

tion voltages V_a and V_b. Their major effect is to alter the beam shape, thus providing a means of compensating for the aberrations introduced by V_x and V_y. The side effect caused by V_a and V_b is a small change in beam position. Table I shows how the voltages V_x, V_y, V_a and V_b are distributed among the octopole plates. The deflection voltages V_x and V_y are applied antisymmetrically, producing a highly uniform deflection field when $\alpha = \sqrt{2} - 1$. The stigmation voltages V_a and V_b, on the other hand, are applied to the plates symmetrically, thus producing a highly nonuniform field that can correct for astigmatism. For a 5-mm-square deflection field, the maximum required deflection voltage is about 80V and maximum required stigmation voltage is about 1V.

The dynamic focus coil is controlled by the voltage V_f. Its major effect is to alter the beam shape, and its side effect is to cause small changes in the beam position.

The octopole calibration derives a transfer function for the octopole. Then, given the required deflection coordinates (x,y) of the electron beam, and the constraint that the beam diameter maintain its undeflected value at any point within the deflection field, it is possible to compute the required values of V_x, V_y, V_a, V_b and V_f. There are good physical reasons for assuming the following forms for the calibration functions:
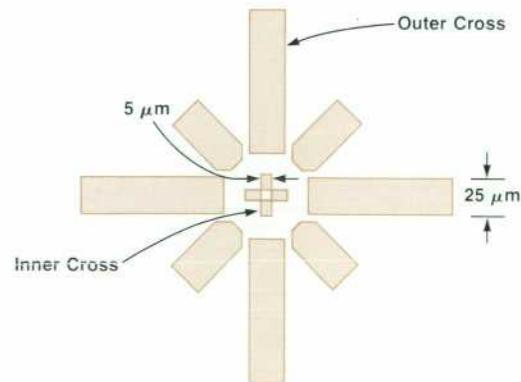

Fig. 2. *Individual pattern on calibration plate. The outer cross is used for coarse beam position measurements and the small inner cross is used for X-Y stage calibration and for fine beam position measurements. The octagonal arrangement is used for measuring the beam shape.*

$$V_a = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \quad (2a)$$

$$V_b = b_{00} + b_{10}x + b_{01}y + b_{20}x^2 + b_{11}xy + b_{02}y^2 \quad (2b)$$

$$V_f = f_{00} + f_{10}x + f_{01}y + f_{20}x^2 + f_{11}xy + f_{02}y^2 \quad (2c)$$

$$V_x = A_{00}x + A_{01}y + A_{20}x^2 + A_{11}xy + A_{02}y^2 + A_{30}x^3 + A_{21}x^2y + A_{12}xy^2 + A_{03}y^3 \quad (2d)$$

$$V_y = B_{10}x + B_{01}y + B_{20}x^2 + B_{11}xy + B_{02}y^2 + B_{30}x^3 + B_{21}x^2y + B_{12}xy + B_{03}y^3 \quad (2e)$$

Thus the output of the calibration procedure consists of values for the a_{jk}, b_{jk}, f_{jk}, A_{jk} and B_{jk} coefficients that allow the electron beam to be deflected accurately and without enlargement within the 5-mm-square deflection field.

In the calibration procedure, a test target on the calibration plate is moved to a series of known locations in the writing plane, using the X-Y stage. At each location, the corresponding values of V_x, V_y, V_a, V_b and V_f are measured and stored. The calibration coefficients are then evaluated by forming a least-square-error fit to the empirical data.

Fig. 2 is a drawing of the test target. The target material is silicon, onto which a 100-nm-thick layer of gold has been evaporated and patterned to form two concentric crosses. The resulting video signal when the target is scanned by the electron beam is largely due to the difference between the backscattered electron coefficients of silicon and gold.

The target may be used to measure the values of V_x and V_y
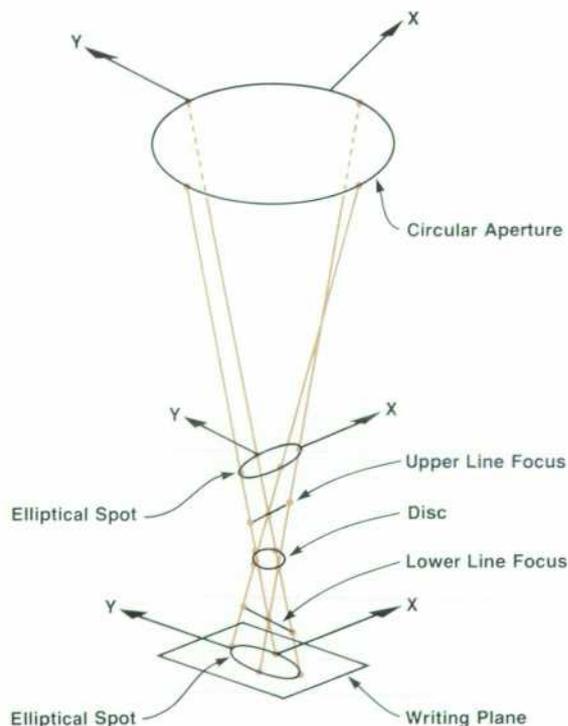

Fig. 3. The initially circular electron beam is distorted to an elliptical cross-section by astigmatism and defocusing. The magnitude of these effects is measured by adjusting the dynamic focus coil voltage, which shifts the appropriate portion of the beam envelope up or down the beam axis to the measurement target plane.

corresponding to a given X, Y position by scanning the electron beam around the perimeter of a square that intersects the four arms of either the outer cross (for a coarse measurement of beam position) or the inner cross (for a fine measurement). In general, the test target is found by a two-stage process consisting of a coarse search followed by a fine search. The search routine makes extensive use of fiducial search programs developed by Hsu (see page 34 of this issue).

The test target is also used to measure beam shapes. To a close approximation, the beam cross-section is elliptical and may be quantified by its major diameter a, minor diameter b and orientation angle φ between the major diameter and the x-axis of the stage. To measure its shape, the beam is scanned radially over the eight silicon-gold edges at the interior of the large cross. The beam diameter in each orientation is deduced from the maximum gradient of the video signal as the beam traverses the silicon-gold boundary. By symmetry, it is to be expected that the diameters measured on opposite edges are equal. The eight measurements may therefore be reduced to four by averaging opposite pairs: these four correspond to 0°, 45°, 90° and 135° scans and are denoted by d_1, d_2, d_3 and d_4. It may be shown that a, b and φ are then given by:

$$\left.\begin{matrix} a \\ b \end{matrix}\right\} = \left[\frac{A/2 \pm \sqrt{B^2+C^2}}{2} \right]^{1/2} \quad (3a)$$

$$\varphi = \tfrac{1}{2}\tan^{-1}(C/B) \quad (3b)$$

where $A = d_1^2 + d_2^2 + d_3^2 + d_4^2$

$B = d_1^2 - d_3^2$

$C = d_2^2 - d_4^2$

Fig. 3 shows the envelope of an electron beam under the influence of astigmatism and defocusing. The electrons are focused to two vertically separated, perpendicular line foci, and the cross section of the beam is, in general, elliptical (the exceptions to the elliptical shape are at the two line foci and at the circular disc which lies halfway between them). The magnitude of the defocus aberration is quantified by the height of the circular disc above the writing plane. The astigmatism is quantified by a magnitude and an angle: the magnitude is the separation between the upper and lower line foci, and the angle is the orientation angle of the upper line focus.

These quantities are measured by building up a three-dimensional picture of the beam envelope from a series of cross-sections parallel to the writing plane, and at different heights above and below it. Each cross-section is measured by energizing the dynamic focus coil appropriately (this shifts the beam envelope along its axis) and simultaneously scanning over the target to make beam diameter measurements. Typically, the through-focal series consists of 11 such measurements which may be processed to determine the magnitudes of the defocus and astigmatism aberrations, and hence the required values of V_a, V_b and V_f.

To calibrate the octopole deflector, a number of points (typically 25) distributed within the octopole field of view are sampled. The test target is moved to each point in turn.
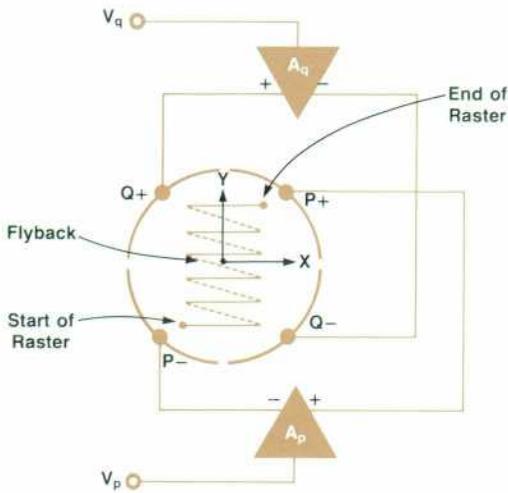
Fig. 4. *Simplified schematic of the four quadrupole deflection plates and their drive circuitry.*

Fig. 5. *If the quadrupole is calibrated, four points, one in each corner of the raster scan, will define a square centered about the optic axis as shown in (a). However, if the quadrupole is uncalibrated, the same four points will define a parallelogram whose center is offset from the optic axis as shown in (b).*

At each point, a through-focal series of beam diameter measurements is made, and the results are used to derive the required values of V_a, V_b and V_f. These defocus and astigmatism correction voltages are then applied, and the values of V_x and V_y needed to center the beam over the test target are measured.

The result of the sampling procedure is a set of 25 values of x and y, and 25 corresponding values of V_a, V_b, V_f, V_x and V_y. The data is fitted to functions of the form of equations (2), and optimum values of the calibration coefficients are derived. The software used to perform this calibration procedure is described in the box on page 29.

Quadrupole Calibration

The quadrupole deflects the electron beam in a raster over a 64-μm square field. Because the quadrupole field of view is small, the corresponding electron optical aberrations are negligible, so three simplifying assumptions may be made:
- The behavior of the quadrupole is independent of the octopole deflection
- Quadrupole deflection causes no growth in beam diameter, and
- The relationship between applied voltage and quadrupole deflection is linear.

Fig. 4 is a simplified view of the quadrupole and its drive circuitry. The quadrupole consists of four plates, labeled $P+$, $P-$, $Q+$ and $Q-$. Plates $P+$ and $P-$ are connected to the symmetrical outputs of one driver amplifier A_p and the Q plates are driven in a similar way by the other driver amplifier A_q. The driver amplifiers have input voltages V_p and V_q.

During exposure, the electron beam is scanned over a 64-μm-square area following the raster pattern indicated in Fig. 4. V_p and V_q are generated digitally in such a way that each raster line consists of 128 points. The 128-line raster therefore consists of $(128)^2 = 16384$ separate points. By using the beam blanker, the electron beam may be switched on or off at any of these points.

The frame-scan and line-scan times of the quadrupole raster depend on the clock frequency in use during a given exposure. This, in turn depends on the beam current, beam
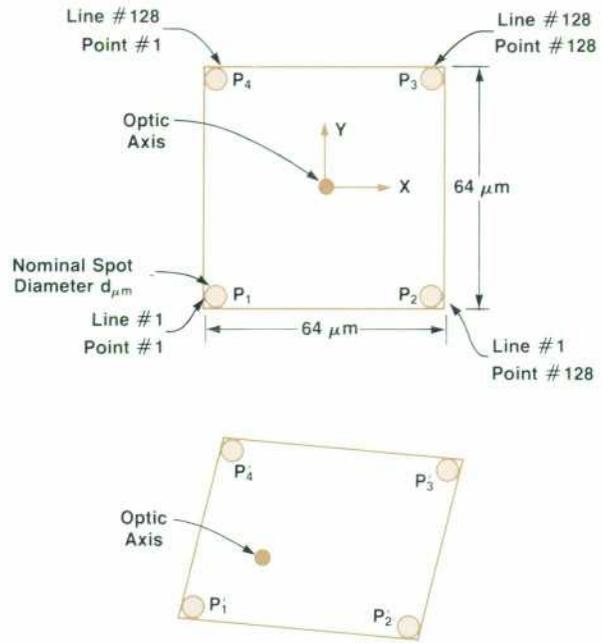
diameter, and resist sensitivity, and in general varies from exposure to exposure.

Because of the simplifying assumptions, the transfer function connecting the input voltages V_p and V_q and the resulting beam position (x,y) is linear:

$$V_p = P_{10}x + P_{01}y + \Delta V_p \tag{4a}$$
$$V_q = Q_{10}x + Q_{01}y + \Delta V_q \tag{4b}$$

The quadrupole calibration procedure determines the parameters P_{10}, P_{01}, Q_{10}, Q_{01}, ΔV_p and ΔV_q. Because of the high-speed circuitry involved, these parameters are functions of clock frequency, a fact that has two consequences. The first is that the calibration may have to be repeated as often as once per substrate, since different substrates will in general require different dosages. The second consequence is that the calibration must be performed while the quadrupole is deflecting the beam in a raster at the same frequency as that used for substrate exposure.

It has already been mentioned that the electron beam may be switched on or off by the beam blanker at any of the 16,384 points within a 64-μm-square block. To calibrate the quadrupole, all but four of the points within a block are switched off. Fig. 5a shows the four points that are on, shaded and labeled as P_1, P_2, P_3 and P_4. If there is no octopole deflection, these four points will be arranged around the optic axis. If the quadrupole is exactly calibrated, they form a square with 64-μm sides centered about the optic axis. If, however, the quadrupole is uncalibrated so that the values of P_{10}, P_{01}, Q_{10}, Q_{01}, ΔV_p and ΔV_q are not known accurately, the four points form a displaced parallelogram whose center is offset from the optic axis as illustrated in exaggerated form in Fig. 5b.
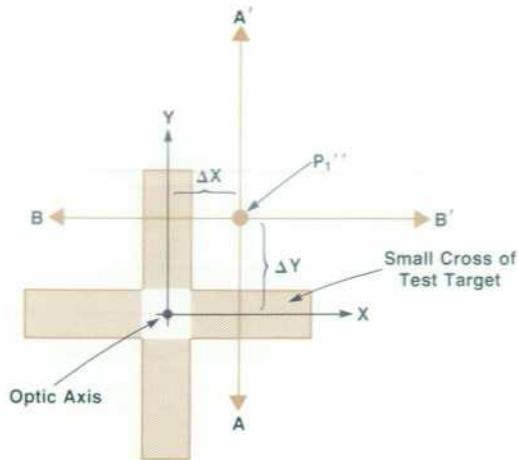
Fig. 6. *When the beam is deflected from P_1 in Fig. 5 to a position near the optic axis (center of the fiducial mark), the error $(\Delta x, \Delta y)$ between its actual position P_1'' and the optic axis can be measured by scanning the beam across the mark along A-A' and B-B'.*

To calibrate the quadrupole, it is run with the test target centered about the optic axis and at the required clock frequency in the way described above. If the quadrupole is correctly calibrated, point P_1 (Fig. 5a) can be brought to the optic axis by shifting the scanned raster a distance $(32-\frac{1}{2}d)$ μm in the X direction and $(32-\frac{1}{2}d)$ μm in the Y direction, using the octopole deflector (d is the nominal spot diameter in μm). With an uncalibrated quadrupole, however, shifting the scan in this way would bring point P_1' (Fig. 5b), not to the optic axis, but to a point P_1'' as shown in Fig. 6. The coordinates of P_1'' may be measured by scanning the raster in the directions A-A' and B-B' with the octopole, and using the fiducial-mark locating routines to evaluate Δx and Δy. In this way the unshifted coordinates of P_1' may be measured.

In a similar manner, the coordinates of P_2', P_3' and P_4' may be measured. The measured coordinates of these points, together with the corresponding values of V_p and V_q may be fitted to equations 4a and 4b to give the values of the parameters P_{10}, P_{01}, Q_{10}, Q_{01}, ΔV_p and ΔV_p.

Robert B. Schudy

Bob Schudy has been a member of the technical staff at HP Laboratories since 1979. He developed mathematical models for the electron beam system and calibration software based on those models. Bob received his Bachelor's degree in physics from the University of California in 1975. He then earned the MS degree in computer science at the University of Rochester, New York in 1977. He worked on algorithms for image processing prior to joining HP and is the author of four papers on modeling heart motion from cardiac ultrasound data. He is a member of the Association for Computing Machinery and the IEEE. A native of the state of Washington, Bob now lives in Mountain View, California with his wife and young son. His outside interests currently center on his family and finishing his PhD thesis.

Ian F. Osborne

Ian Osborne came to HP in 1979 and was responsible for implementing the octopole calibration programs for the electron beam lithography system. Before joining HP, he worked on software system support and development in England. Ian is a member of the British Computer Society and the Association for Computing Machinery. He is a native of London, England, married, and makes his home in Bracknell, Berkshire, England. Ian has been working in HP's laboratories in Palo Alto, California and will return to England this summer. His outside interests include the theatre, ballet, listening to music, photography, and various sports (squash, cricket, soccer, tennis, skiing, and backpacking).

Faith L. Bugely

Faith Bugely came to HP in 1979 and is an engineer and programmer for the electron beam lithography system. She previously worked as a systems analyst and studied mathematics at Rutgers and Columbia Universities from 1960 to 1965. A native of Plattsburgh, New York, Faith now lives in San Jose, California and enjoys woodworking, backpacking, and mountain climbing.

Geraint Owen

Geraint Owen earned the BA and PhD degrees in electrical engineering from Cambridge University, England in 1971 and 1975, respectively. After doing postdoctoral research in electron optics, he joined HP in 1979 as a member of the electron optics group working on the electron beam lithography system. Geraint has authored twelve papers and is a member of the IEEE.

Digital Adaptive Matched Filter
for Fiducial Mark Registration

by Tsen-gong Jim Hsu

Fast, accurate detection of fiducial marks on substrates is necessary to attain high levels of throughput, pattern registration, and equipment calibration in an electron beam lithography system. This is generally done by processing an electronic waveform $y(x)$ derived from the detection of the electrons backscattered by the fiducial mark and its substrate (Fig. 1) as the electron beam is scanned across the region where the fiducial mark is estimated to be. By correlating the occurrence of the signal caused by the mark with the deflection voltages required to position the electron beam at the location where the signal is detected, the pattern coordinates can be established accurately.

Discerning the portion of the waveform caused by the fiducial mark is often difficult because considerable levels of noise are also present, the mark may be covered by a layer of resist, and the contrast between the mark and its surroundings may be low. This difficulty is overcome by recognizing that the desired signal should be an odd or even function with a center of symmetry corresponding to the location of each arm of the fiducial mark. Also, because the substrate is mechanically registered on the system's X-Y stage, the approximate location of the fiducial mark is already known. This reduces the search area and hence the time required for detection.
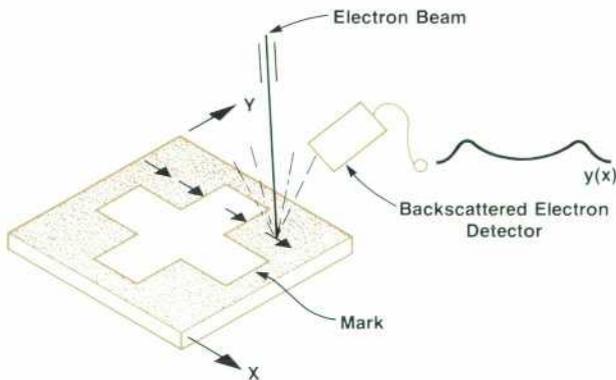

Fig. 1. *The signal for detecting the fiducial mark is generated by scanning the electron beam through the vicinity of the mark and detecting the number of electrons backscattered from the mark and its substrate.*

Signal Detection

Thus, the problem is to detect a partially known signal in the presence of additive noise. The signal $f(x)$ is considered to be an even or odd function with a center of symmetry located at $x = x_c$ and to have a value of zero for values of x outside the range from $x_c-w/2$ to $x_c+w/2$. Here x_c is the centerline coordinate of the arm of the fiducial mark and w is the width of the search area. The noise $n(x)$ is estimated to be stationary white noise. Therefore, the waveform $y(x) = f(x)+n(x)$.

Adaptive matched filters can be used to detect the even or odd signals embedded in noise. If $y(x)$ is applied to a space-variant linear system whose block diagram is shown in Fig. 2, it can be shown that the output of the system is

$$z(x) = \int_{-w/2}^{w/2} y(x-x')y(x+x') \, dx' \qquad (1)$$

when the system has an impulse response of $y(x+x')$ at x for $-w/2 \leq x' \leq w/2$.

By multiplying the signal and noise components in equation (1), the average output of the detector can be obtained as the sum of the averages of four product terms. Because of the nature of $n(x)$, it can be shown that the average of the noise product term is zero. Also, the averages of the signal-noise product terms will cancel each other, leaving only the average of the signal product: $f(x+x')f(x-x')$.

To analyze the nature of the average of the signal product term, define a function $g(x';x) = f(x+x')$ within the interval w centered at x and equal to zero elsewhere. Separate g $(x';x)$ into its even and odd parts. The even component $g_e(x';x)$ has mirror symmetry about x and the odd component $g_o(x';x)$ has rotational symmetry about x. Fig. 3 shows examples of $g(x';x)$ separated into their respective even and odd parts for various values of x. Note that the even and odd parts are orthogonal over the interval from $x-w/2$ to $x+w/2$ and that

$$f(x+x') = g(x';x) = g_e(x';x)+g_o(x';x) \qquad (2)$$
$$f(x-x') = g(-x';x) = g_e(x';x)-g_o(x';x) \qquad (3)$$

for $-w/2 \leq x' \leq w/2$.

Thus, by multiplying equations (2) and (3) and remembering that the average value of the product of two orthogonal functions is zero, the output of the system is equal to the difference between the energies of the even and odd components. That is

$$Z(x) = E_e(x) - E_o(x) \qquad (4)$$

Since the average even and odd component energies are both positive for all values of x, it follows from equation (4) that the expected value for the output at the point of symmetry, $x = x_c$ is equal to $E_e(x_c)$ if the signal is an even function, and is equal to $-E_o(x_c)$ if the signal is an odd function.

Also, the energy detected for $x = x_c$ is greater than the energy detected for any other value of x. Therefore, it follows that, when $x = x_c$, the output of the detector is a maximum if the signal is an even function or a minimum if it is an odd function.

It also can be shown that the signal-to-noise ratio at the location with maximum $|z(x)|$ is $\frac{1}{2}\sqrt{E_s/N_0}$ where E_s and N_0 are the total energy of the signal $f(x)$ and power spectral density of the stationary white noise $n(x)$, respectively.

One advantage of the adaptive matched filter is that having complete information about the signal is not necessary to detect it. That is, one doesn't need to know the exact waveform of the signal, just its approximate length and that it is an even or odd

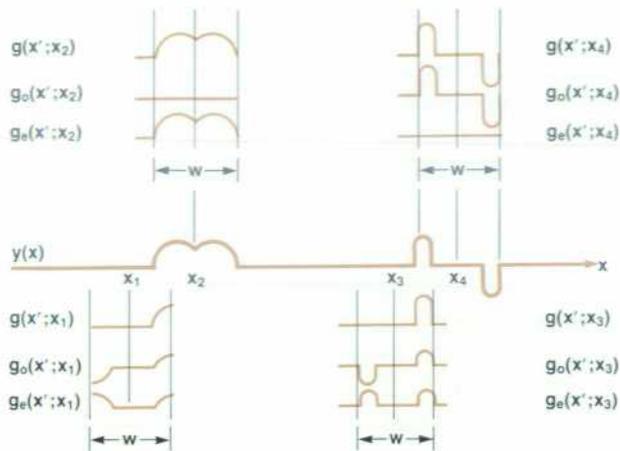**Fig. 2.** *Matched filter block diagram for detecting signals that are even or odd functions.*

Fig. 3. *Examples of a function detected over an interval w, centered at a value x_n and separated into its even and odd components.*

function is sufficient. This makes an adaptive matched filter very attractive for finding fiducial marks in an electron beam lithography system. As the electron beam scans across a line or an arm, the waveforms of the detected backscattered electron signals always vary because of varying resist thickness and contrast between the marks and the substrate. However, the signals remain either even or odd functions, and since only the centers of the fiducial mark arms need to be determined, an adaptive matched filter is obviously a good technique to detect them.

Another favorable property of the adaptive matched filter is its processing speed. Since the integrand $f(x-x')\,f(x+x')$ is also an even function, one can take advantage of this by integrating over one half the interval and doubling the result. This further increases the processing speed.

Digital Approach

Because digital systems are more easily implemented and controlled, a discrete version of the above analysis was done. Briefly, equation (1) can be realized by applying discrete input samples $y[i]$ to a digital filter whose impulse response at i is $y[i+j]$ for $-K/2 \leq j \leq K/2$ where K is the window of the filter (digital equivalent of w). The responses are then summed and averaged and the resulting output is

$$z[i] = 2\sum_{j=0}^{K/2} y[i-j]\,y[i+j] \qquad (5)$$

To have a normalized measurement of the output, the following modified detector was introduced where $E[i]$ is the energy and $A(y[i])$ is the mean value of $y[j]$ for $i-K/2 \leq j \leq i+K/2$, respectively.

$$r[i] = \frac{z[i]/K - A(y[i])^2}{E[i]/K - A(y[i])^2} \qquad (6)$$

where $E[i] = \sum_{j=-K/2}^{K/2} y[i+j]^2$ and $A(y[i]) = 1/K \sum_{j=-K/2}^{K/2} y[i+j]$

Therefore, when $i = i_c$, $r[i]$ is maximum for an even function signal and minimum for an odd function signal. Also, because the absolute value of $r[i]$ should be less than or equal to one for all i, it can be used as a good measure of output reliability.

This approach has considerable flexibility because various parameters, such as K, can be varied to adapt to different situations. For example, if the initial estimate of the fiducial mark location is inaccurate, the search can be easily broadened by increasing K. Then, when the mark is found, a more accurate estimation of its location can be made and the value of K can be decreased.

Results

Two sets of real data are shown in Fig. 4. There is an even function signal and an odd function signal embedded in the waveforms of Fig. 4a and Fig. 4b, respectively. The corresponding outputs of the adaptive digital filter system are shown below each input waveform. It is clear that the maximum output of Fig. 4a and the minimum output of Fig. 4b indicate the centers of the respective even and odd signals.

Fig. 5 shows three sets of real data containing odd-function signals as a function of increasing resist thickness over the fiducial mark. Despite the increased thickness, the minimum of the output still indicates the center of the mark.
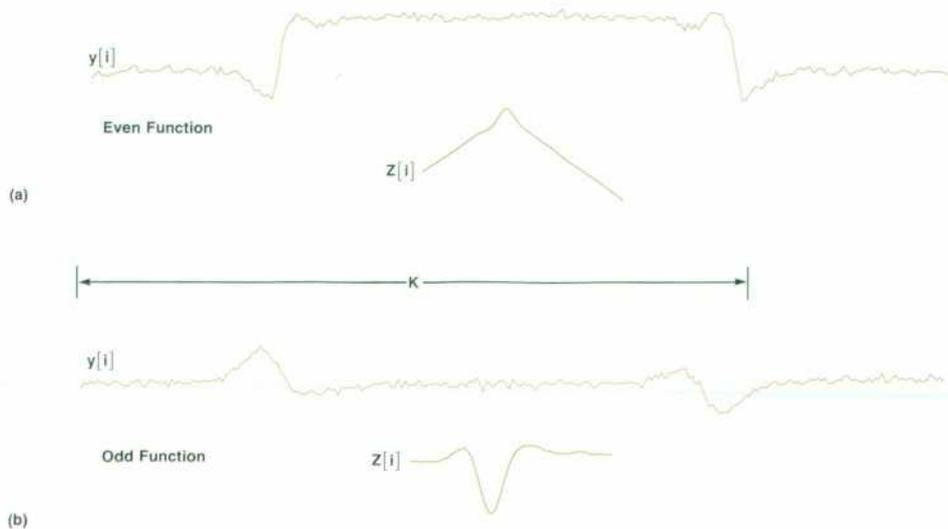

Fig. 4. *Detection of actual even (a) and odd (b) function signals embedded in random noise. The corresponding detector output $z[i]$ in (a) is maximum at the center of the function and in (b) is at a minimum.*
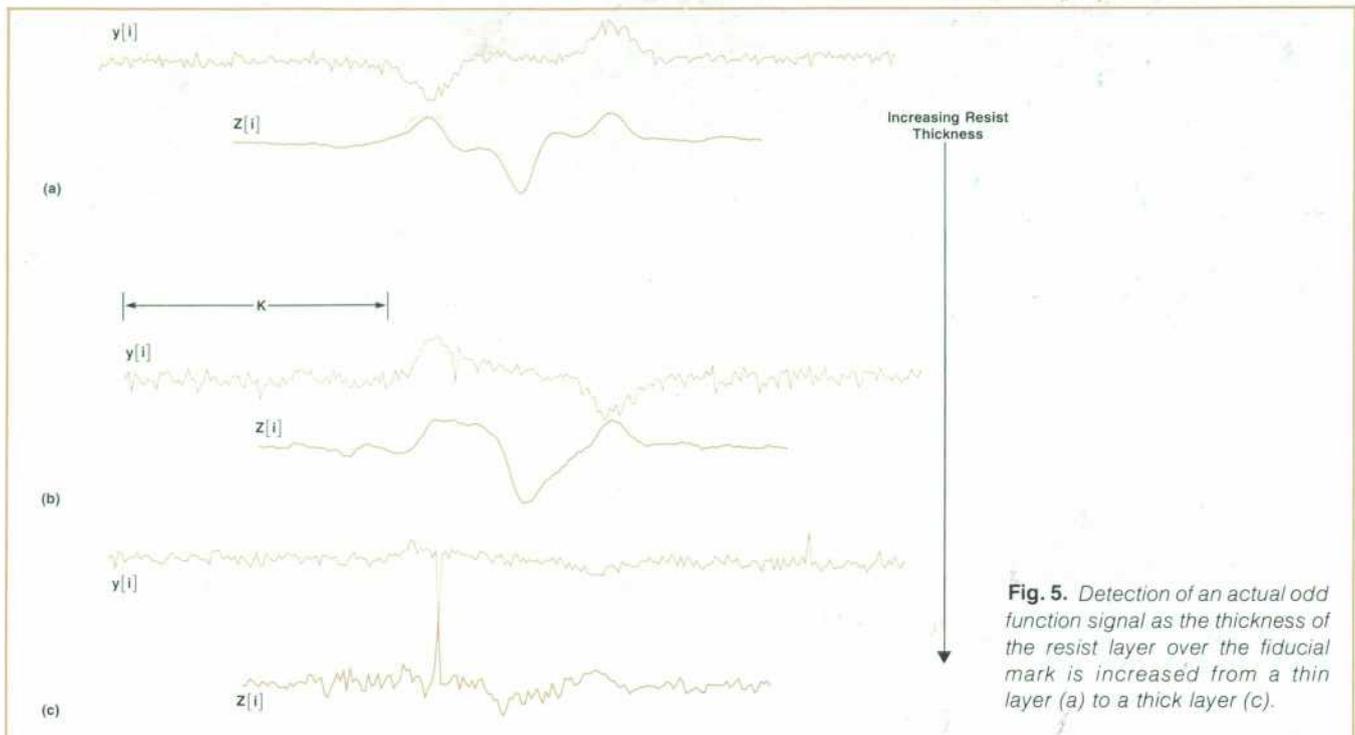
y[i]

z[i]

Increasing Resist
Thickness

(a)

|←———— K ————→|

y[i]

z[i]

(b)

y[i]

z[i]

(c)

Fig. 5. *Detection of an actual odd function signal as the thickness of the resist layer over the fiducial mark is increased from a thin layer (a) to a thick layer (c).*

Tsen-gong Jim Hsu

Jim Hsu attended Chung Kung University in Taiwan where he earned the BSEE degree in 1966. He did graduate work at the University of California at Davis, receiving the MSEE degree in 1970. He also has done most of the work for his PhD degree there. Jim joined HP in 1979 and is a member of the software group working on the electron beam system. He is a member of the IEEE and enjoys hiking and swimming. He is married, has two sons, and lives in Cupertino, California.

CHANGE OF ADDRESS: To change your address or delete your name from our mailing list please send us your old address label. Send changes to Hewlett-Packard Journal, 1501 Page Mill Road, Palo Alto, California 94304 U.S.A. Allow 60 days.