V. PATTERN REPLICATION AND GENERATION

An integrated circuit is essentially a set of patterns of impurities, oxide layers, and conductors laid down on a semiconductor substrate. These patterns are usually placed on the circuit by pattern replication processes from masks. The masks are themselves replicas of some original pattern that was recorded as opaque and transparent areas by a pattern generation process from a description of the pattern originally conceived by a designer. Usually, pattern generation is done by a digitally controlled recording device. In this report pattern generation refers only to the process of first recording the pattern in physical form from some abstract description of it; we will not concern ourselves with how the description was obtained. Pattern replication refers to the process of copying a pattern from one physical form to another by optical, chemical, physical, electronic, or mechanical means.

PATTERN REPLICATION

In today's integrated circuit technology, very precise pattern replication processes are used. Such precision has enabled the industry to proceed to ever finer feature dimensions while simultaneously increasing the overall size of the patterns replicated. Because the component count per chip allowed by present yield limitations is much lower than the limit imposed by the available pattern replication steps, the industry has been able to produce many chips simultaneously and thus achieve very low costs. If yields permitted a component count per chip larger than the limit imposed by pattern replication steps--as may well be the case in the future--multiple pattern replication steps would be required to produce different areas of the same chip. The precision available in pattern replication provides a natural boundary for growth of component count in integrated circuits; it will be substantially more difficult, but not impossible, to produce chips more complex than is permitted by this precision.

The number of components that can be fabricated in a single circuit have a natural limit determined by replication precision. The ultimate precision available in pattern replication steps is limited by the materials used to hold the patterns. It appears that the anomalous dimensional changes in silicon subjected to the high temperatures required for integrated-circuit processing will limit precision to about 10^5 resolution elements on a side or approximately 10^8 devices per chip. Chips with larger component counts must of necessity be assembled from subunits linked together by areas permitting misregistration. The connection areas will have conductors considerably larger than those within the subunit, and the size of these conductors will limit the number of interconnections between subunits. If the precision of newer fine-line replication technologies is substantially less than that of current optical techniques, the natural subunit may have only about 10^6 devices.

Figure 1 shows the relationship between replication precision and the difficulty of replication. If the precision is adequate to cover

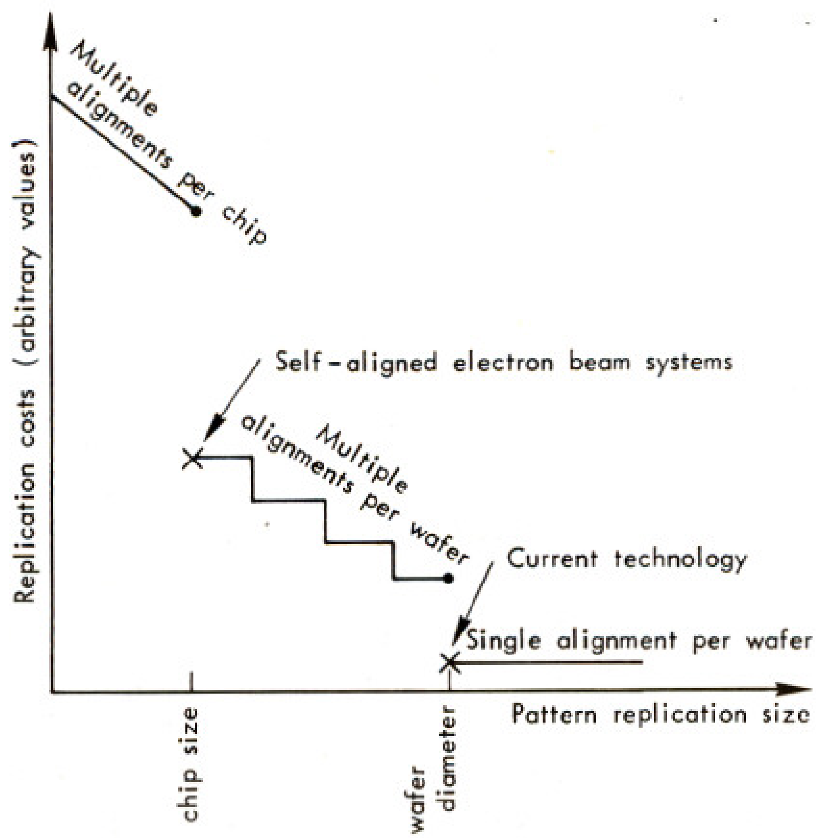

Fig. 1--Schematic representation of the effect of pattern replication precision on replication costs

an entire wafer, replication is relatively easy. If the precision is inadequate for an entire wafer but adequate for an entire chip, the cost of replication will depend on how many replications are required to cover the wafer, and will thus vary in staircase fashion with changes in precision. If the precision is inadequate to cover even the size of a single chip, multiple replications per chip are possible but will be substantially more expensive. Current practice involves pattern replication over an entire wafer. Most of the self-aligned electron beam systems now under development operate with one alignment per chip.

PATTERN GENERATION

Pattern generation, unlike replication, needs to be done only to the precision required by the component count of the circuit, or part of a circuit, for which the pattern is being generated. In today's technology, patterns for a single circuit need to be generated with a precision of only about one part in 10^4. These individual patterns are then combined by a step-and-repeat process to form highly precise masks, which are accurate to about one part in 10^5 and are ultimately used to replicate the patterns onto the semiconductor. As component count increases, both the need for accuracy in pattern generation and the number of features to be recorded in a pattern will also increase.

The time that it takes a pattern generator to record a pattern depends, of course, on the number of parts that are in the pattern and the speed of the pattern generator. As will be explained below, pattern generation times increase more than linearly with decreases in feature size. Pattern replication, if it can be done at all, can be done in a time that is independent of the actual pattern replicated for a given feature size. This time is usually much less than that required for pattern generation. Moreover, equipment to generate patterns must be digitally controlled as well as precise, whereas pattern replication equipment can often be less costly analog equipment similar to a camera. To the extent that replication devices must include complex alignment procedures, however, they become more complex and costly. If the alignment procedures are very complex, the distinction between pattern generation and pattern replication steps may become diffused. Alignment procedures are an important part of the development of the new replication technologies.

The newly evolving submicron fabrication technologies are based on new mechanisms for pattern generation and replication. It is generally agreed that pattern generation is best done with computer-controlled electron beam recording, but there is essentially no agreement on the design details of such recorders. Two major categories of recording systems exist: raster scan and vector control. The raster scan systems methodically cover the area of the pattern to be generated, turning the electron beam on or off as demanded by the requirements of the pattern being recorded. The vector control systems deflect the beam to locations specified by the needs of the pattern.

Although the raster-scan technique is simple in concept, and avoids critical constraints on the linearity and hysteresis of the deflection system, it has two major disadvantages: (1) The video rate of a given pattern is very high; the rise time must be less than the pixel time. (2) The format makes it difficult to adjust exposure to compensate for local variations in pattern complexity. The vector scan, on the other hand, requires exceedingly tight control of the deflection system but preserves the locality of shape. This locality permits the beam current to be dynamically adjusted for interiors of large areas and makes adjustment for proximity effect at adjacent edges easier.

It thus appears that if the deflection problems can be solved, vector scan systems will provide about 10 or more times the throughput of raster scan systems for line dimensions of $\sim$0.3 μm or smaller.

SPEED LIMITS IN PATTERN GENERATION

Both raster-scan and vector-control pattern generators face basic limitations in operating speed. The rate at which a pattern can be generated can be limited by two factors: (1) Given that the pattern contains a certain amount of information, the bandwidth of the systems that control the electron beam place limits on how fast the pattern can be transmitted. (2) Resist sensitivity sometimes limits the writing rate of the electron beam.

Many complex phenomena are involved when electrons expose a resist by losing energy in it. The electron scatters elastically and inelastically, changing direction and losing energy. The resist changes physically and chemically, and the pattern left after development depends on all of these effects plus those introduced by the development process. However, there are certain fundamental relationships that are useful to observe.

Let us assume that the resist requires a dose of Q coulombs/cm^2 for correct exposure. In order to be certain that a given pixel is exposed, i.e., to ensure adequate pixel signal-to-noise level, at least a minimum number of electrons, N_m, must strike and lose their energy in each pixel. This is fundamental and important. Since $Q/e \geq (N_m/\ell_p^2)$, where ℓ_p is the pixel linear dimension and e is the electron charge, Q *must increase* as ℓ_p decreases, for the probability that each pixel will be correctly exposed to remain constant. Stated another way, based on pixel signal-to-noise considerations, the minimum total number of electrons needed to expose reliably a pattern of a given complexity, i.e., with a given number of pixels, is *independent* of the size of pixels. More sensitive resists are useful for larger pixels; less sensitive resists *must* be used for smaller pixels. This argument assumes that an electron's energy is lost within a pixel, i.e., that the transverse scattering is considerably smaller than a pixel, and that the beam size is at least as small as a pixel.

Appendix C exhibits the fundamental considerations of electron beam formation and focusing that cause the time, τ, required to expose a pixel to N_m electrons to increase as the pixel linear dimension ℓ_p decreases. As shown by the left-hand curve in Fig. 2, $\tau \propto \ell_p^{-8/3}$. To correctly expose a real resist of sensitivity Q coulombs/cm^2, a fixed number of electrons per unit area must strike the resist, and the time required to expose such a resist is $\tau_R \propto \ell_p^{-2/3}$. A family of curves corresponding to such real resist exposure is also shown in Fig. 2. For a given probability that each pixel will be correctly exposed, these curves for a real resist cannot extend to the left past the limiting curve. As we proceed to the right of the limiting curve along a curve for constant sensitivity, Q, the number of electrons striking each picture

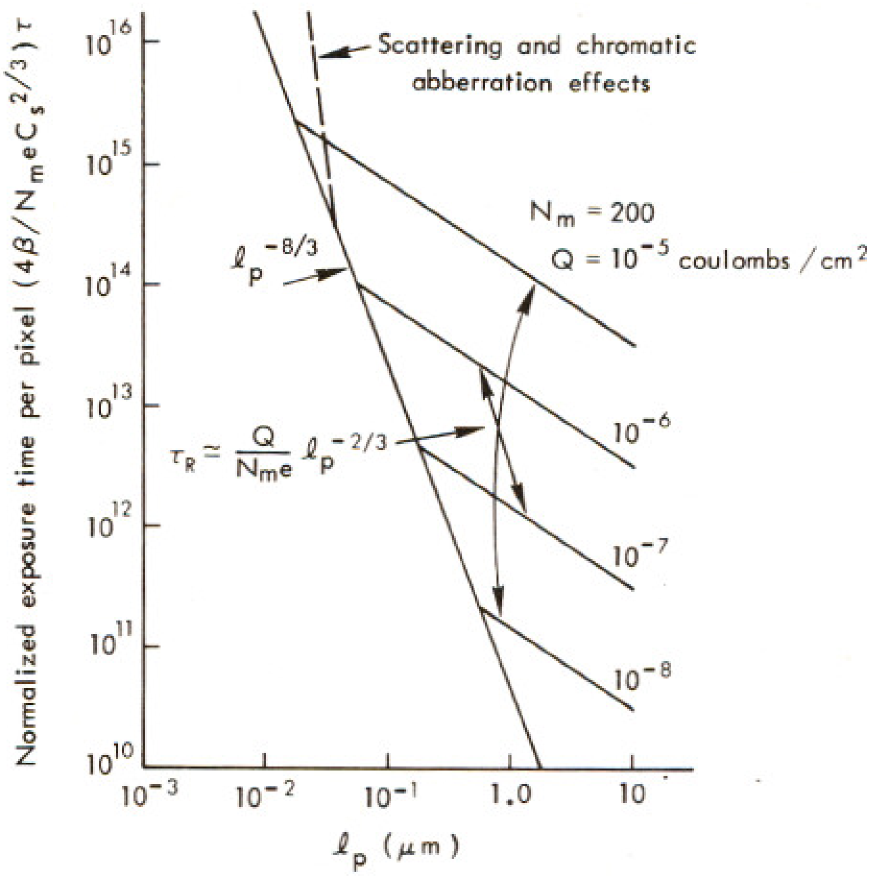

Fig. 2--Normalized exposure time/pixel vs pixel dimension

element increases, improving the pixel signal-to-noise ratio. Because the normalization factor on the ordinate of Fig. 2 includes N_m, the vertical positioning of the τ_R curves depends on the value of N_m actually chosen. For binary exposure, the probability that a pixel struck by 200 electrons is not correctly exposed is less than 10^{-12}; if struck by 100 electrons, a pixel has a probability of incorrect exposure of 3×10^{-7}, enough to cause many errors in a pattern of 10^{10} pixels. Hence, we have set $N_m = 200$ in the τ_R curves of Fig. 2.

These curves predict that for $Q = 10^{-8}$ coulombs/cm^2, pixels smaller than $\ell_p = 1.0$ μm should be possible, and for $Q = 10^{-6}$ coulombs/cm^2, pixels below $\ell_p = 0.1$ μm should be attainable, based on signal-to-noise ratio considerations alone. Resists such as polymethyl methacrylate processed for high resolution by the correct choice of developer have demonstrated line widths less than 0.1 μm. The fundamental point to emphasize is that slow resists are necessary to get higher resolution, a result familiar to all photographers. Note that if all electron energy is not dissipated within the pixel (due to lateral scattering, for example), the exposure time per pixel increases and the solid curve in Fig. 2 moves toward the dashed curve. Inclusion of quantitative information on scattering and aberrations in addition to spherical aberration will cause the actual limiting curve to move toward the right at small pixel dimensions, as shown in Fig. 2.

It is instructive to consider an example. Suppose that we want to know how long it will take to expose a chip on an integrated circuit consisting of 10^8 pixels, with a pixel dimension $\ell_p = 500$ Å. We may choose to use a tungsten hairpin cathode, a LaB_6 cathode, or a field emission cathode, as discussed by Broers (1972). Furthermore, we may choose to scan a relatively small area of the surface and move the object being exposed often (or continuously, as in EBES); or we may choose to scan a larger field and move the object less frequently. Normally, a vector scan is chosen in the latter case, and a raster scan in the former (as in EBES). If the smaller field is chosen, the electron optics can be optimized for a higher current and current density at the object being exposed.

From Broers (Fig. 1 or 10 and Fig. 11) we determine that the current in a beam of 500 Å diameter from a tungsten cathode, standard LaB_6 cathode, and a field emission gun followed by a lens is 1.3×10^{-11}, 6×10^{-10}, and 7×10^{-9} amps, respectively. For the stated resolution of 500 Å, our previous argument suggests that $Q \geq 6.4 \times 10^{-6}$ coulombs/cm^2. If, for this calculation, we do not consider the settling time of the beam and the repositioning-registration time of the stage for vector scan systems, and we assume that the rise-time of a raster scan system can be arbitrarily fast, the time that the beam must be on the chip to expose 1 pixel, 10^8 pixels, or 1 cm^2 is given in Table 3.

Many assumptions have been made in arriving at the numbers in Table 3. For example, it is assumed that the vector scan must expose every pixel, whereas in practice only 5 to 50 percent of the pixels are exposed; hence the *exposure time* for the vector scan system is very

Table 3

COMPARISON OF EXPOSURE TIMES AT 500 Å RESOLUTION FOR DIFFERENT SOURCES AND DIFFERENT ELECTRON-OPTICAL PARAMETERS

Source	Exposure Time			Conditions
	For 1 pixel $= \tau(\mu sec)$	For 10^8 pixels $= 10^8\ \tau(sec)$	For 1 cm^2 $= 4 \times 10^{10}\ \tau(hr)$	
Tungsten (hairpin)	9.7 μsec	970 sec	107 hr	*Vector Scan* C_s = 12 cm, C_c = 5 cm [Broers Fig. 11] Beam Voltage = 25 kV
LaB_6 (standard)	0.26 μsec	26 sec	2.9 hr	
Field effect (gun & lens)	0.018 μsec	1.8 sec	0.2 hr	
Tungsten (hairpin)	0.16 μsec	16 sec	1.7 hr	*Raster Scan* C_s = 1.8 cm, C_c = 1 cm, [Broers Fig. 1] Beam voltage = 25 kV
LaB_6 (standard)	0.021 μsec	2.1 sec	0.23 hr	

Source	Exposure Time	Conditions
Conventional X-ray	1000 sec	100 mA, 10 kV 30 cm working dist
Synchrotron radiation	1 sec	10 mA 500 MeV beam

pessimistic (i.e., too long). The exposure times can be shortened by increasing the beam brightness. For a given resolution in the resist, the resist sensitivity has a maximum value $Q_m = N_m e/\ell_p^2$; using resists more sensitive than this will decrease resolution or the confidence in exposure, and hence the yield. Thus a more sensitive resist than $Q = 6.4 \times 10^{-6}$ coulombs/cm^2 in the above example is not useful, and will not usefully decrease exposure times.

ALIGNMENT LIMITATIONS FOR PATTERN REPLICATION

Pattern replication for submicron dimensions, unlike pattern generation, does not suffer any important fundamental speed limitations, but it does suffer various kinds of resolution and precision limitations. We have become accustomed to pattern replication processes with accuracies on the order of one part in 10^5, a level difficult to maintain with some of the new methods. Noncontact pattern replications are highly desirable, because placing a mask in contact with the integrated circuit damages the mask and forces one to use multiple generations of masks, which interposes stages of replication otherwise unnecessary. While conformable masks alleviate the problem to some extent, they do not eliminate it. Noncontact printing requires highly collimated sources of radiation and sophisticated alignment schemes. So far, the most ideal radiation source seems to be ultraviolet synchrotron radiation. However, both ultraviolet and X-rays require masks that are much thinner than the glass plates currently in use, and the dimensional stability of such masks is not likely to be any better. Decreased accuracy in replication means that smaller areas will be exposed at a single replication step. If wafer size is maintained, this implies multiple exposures of each wafer.

Alignment for the replication steps is now done manually by a light microscope. For alignments down to 0.1 micron or even finer, it still appears possible to use light as an alignment mechanism. Many systems use scanning electron beams in various configurations for automatic alignment. In addition to an evacuated chamber, they require fairly complicated sensing and actuating systems and are thus quite expensive.

Some of the electron-beam pattern-generation systems now being used for making masks can also be used to expose wafers directly. Point-by-point serial electron-beam writing equipment references the beam to registration markers on the wafer so that the orientation and size of the pattern can be held within acceptable tolerances during exposure. By using a laser interferometer to measure beam position with respect to a given origin, it is possible to expose an area of $\gtrsim 10^5$ pixels $\times \gtrsim 10^5$ pixels ($\gtrsim 10^{10}$ pixel^2). Without a laser interferometer, areas of $\sim 10^6$ pixel^2 to 10^8 pixel^2 can be exposed after each registration, depending on the stability of the electrical signals used to accelerate and deflect the beam, the electron optical corrections, etc.

The time required for each registration determines how many registrations are economically feasible within the processing time devoted to a wafer. As feature sizes continue to decrease, exposure times in electron-beam pattern-generation equipment of necessity increase, so that automatic alignment becomes a small fraction of total exposure time. Just as a constant number of electrons per pixel are required to provide an adequate signal-to-noise ratio for the self-alignment process, so are a certain number of electrons required to reduce dimensional signal-to-noise ratios to a desired level. Thus, for patterns of more than a certain number of pixels, self-alignment should be acceptable for each pattern independent of feature size.

If the fractional substrate distortion,

$$\left| \frac{r_m - r_m'}{r_m} \right| ,$$

where r_m is the measured distance between two features and the prime indicates a measurement after processing, exceeds the ratio of feature size to pattern size, misregistration is bound to occur. To avoid this, local registration is essential for smaller circuit dimensions. For patterns exceeding 10^3 to 10^4 pixels on a side, laser interferometry servo control will be required in addition to electron-beam-deflection control. For patterns simpler than 10^3 or 10^4 pixels on a side, open-loop electronic control following alignment will suffice and is substantially faster and less expensive than servo control.