

Modeling of defects in integrated circuit photolithographic patterns

by C. H. Stapper

In a previous paper by the same author the foundation was laid for the theory of photolithographic defects in integrated circuits. This paper expands on the earlier one and shows how to calculate the critical areas and probability of failure for dense arrays of wiring. The results are used to determine the nature of the defect size distribution with electronic defect monitors. Several statistical techniques for doing this are described and examples are given.

1. Introduction

This paper is a sequel to an earlier one [1] on modeling of integrated circuit defect sensitivities. The earlier work described mathematical models for very small defects that cause insulator pinholes and reverse current diode leakage. It also dealt with photolithographic defects, in which case the defect size becomes of importance.

In that paper the critical area was defined as that area in which the center of a defect must fall to cause a failure or fault in the integrated circuit. This area was found to vary as a function of the diameter χ of circular defects. The magnitude of χ was used as the defect size, and we explained how long conductive lines could be used to determine the

nature of the defect size distribution. Measurements at IBM, taken over a long period of time, have suggested a $1/\chi^3$ fall-off in that distribution as the defects size χ increases.

Not covered in the previous paper were the effect of pattern "proximity" on critical area calculations, the consequences of a $1/\chi^3$ distribution, and the effect of limiting the critical area to the size of a chip or defect monitor. These three subjects are covered here, as well as the determination of the defect size distribution with defect monitors.

2. Critical area as a function of defect size

The critical area for a long conductor of width w and length L was found in [1] to be

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (1a)$$

$$A(\chi) = L(\chi - w) \quad \text{for } w \leq \chi \leq \infty, \quad (1b)$$

as a function of defect size χ . This function is depicted graphically in **Figure 1**.

For two very large conductors, separated by a narrow slit of width s and length L , the critical area was found to be

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq s, \quad (2a)$$

$$A(\chi) = L(\chi - s) \quad \text{for } s \leq \chi \leq \infty. \quad (2b)$$

This is the same function as that for open circuits except that the symbol s is used for spacing instead of w for line width. This complete duality between open and short circuit models holds even for many of the more complex circuits. Where they are the same, only the open circuit case is described in this paper.

©Copyright 1984 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

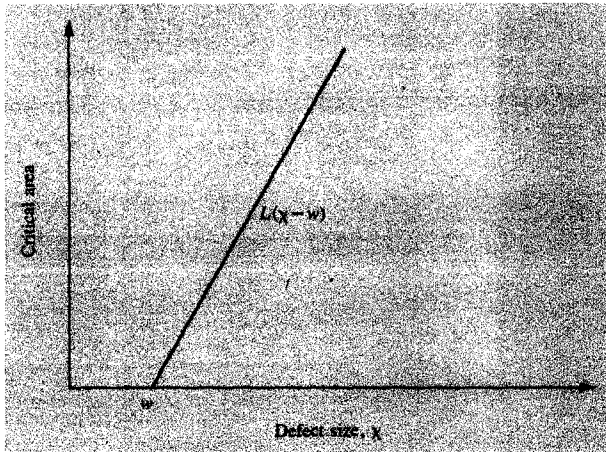


Figure 1

Critical area as a function of defect size for a very long and narrow integrated circuit conductor of length L and width w .

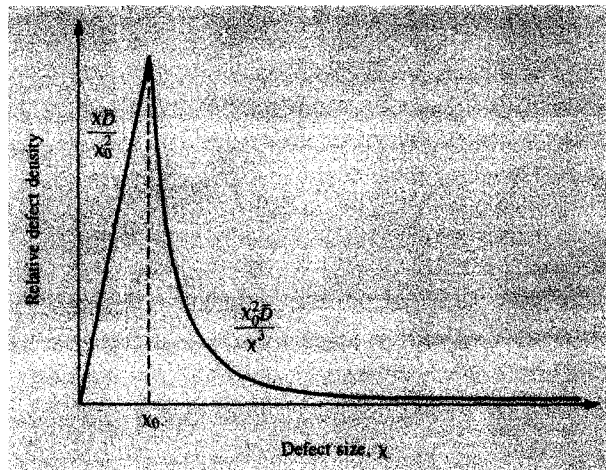


Figure 2

Hypothetical defect size distribution.

3. Defect size distribution

The critical areas as a function of defect size have to be combined with a distribution of defect densities that is also a function of defect size. Hypothetical distributions of the form

$$D(\chi) = \frac{2(n-1)\chi\bar{D}}{(n+1)\chi_0^2} \quad \text{for } 0 \leq \chi \leq \chi_0, \quad (3a)$$

$$D(\chi) = \frac{2(n-1)\chi_0^{n-1}\bar{D}}{(n+1)\chi^n} \quad \text{for } \chi_0 \leq \chi \leq \infty, \quad (3b)$$

where n is a positive number, have been used to analyze defect monitor data. This distribution has a maximum and discontinuous slope at defect size χ_0 .

A value of $n = 3$ has been found acceptable in a number of experiments at IBM, as described in [1]. This is the number that is used in this paper. It produces the distribution

$$D(\chi) = \frac{\chi\bar{D}}{\chi_0^2} \quad \text{for } 0 \leq \chi \leq \chi_0, \quad (4a)$$

$$D(\chi) = \frac{\chi_0^2\bar{D}}{\chi^3} \quad \text{for } \chi_0 \leq \chi \leq \infty, \quad (4b)$$

which has an average defect density $\bar{D}$. A plot of this function is shown in Figure 2. The graph indicates that the defect density has a maximum at χ_0 . Defects smaller than this size cannot be resolved by the optics used in the photolithographic process. The minimum dimensions of the patterns must therefore always be larger than χ_0 .

Other defect size distributions can also be used. An example of a hypothetical exponential distribution appears elsewhere in this paper.

4. Average critical area for a single long line

In [1] the failures caused by defects were called faults. The expected or average number of faults λ can be calculated by combining the critical area and the defect size distribution with the integral

$$\lambda = \int_0^\infty A(\chi)D(\chi)d\chi. \quad (5)$$

Since $\chi_0 < w$, the critical area in (1b) and the defect size distribution in (4b) are the only nonzero contributors to this integral, and therefore

$$\lambda = \int_w^\infty L(\chi - w) \frac{\chi_0^2\bar{D}}{\chi^3} d\chi. \quad (6)$$

Evaluation gives

$$\lambda = \frac{L\chi_0^2\bar{D}}{2w}. \quad (7)$$

With this result it is possible to define an average critical area

$$\bar{A}_1 = \frac{L\chi_0^2}{2w}. \quad (8)$$

This average takes the defect size distribution into account. The average number of faults in this example is now simply given by

$$\lambda = \bar{A}_1\bar{D}. \quad (9)$$

The subscript 1 has been used here to designate the critical area for a single long line. We next take a look at the critical area of two long conductive lines.

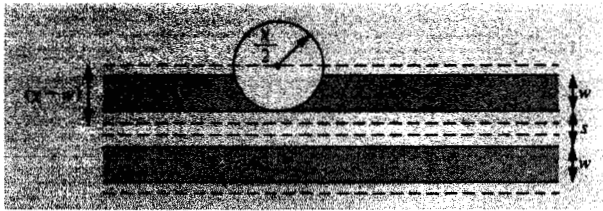


Figure 3

Two long lines. The critical areas for the lines grow independently until they meet halfway between.

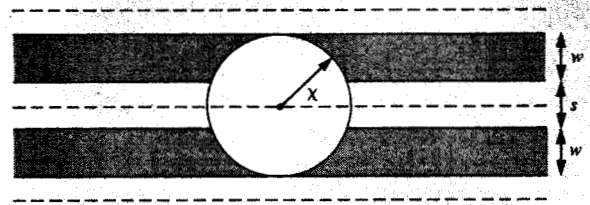


Figure 4

Defect size for which the two critical areas meet halfway between two lines is $2w + s$.

5. Critical area for two conductive lines

The diagram in **Figure 3** represents two conductive lines, each of width w , and a space s between them. The length of these lines is given as L . It is assumed that this length is much greater than the line width or the spacing so that end effects can be neglected. We determine the critical area for the condition where an open circuit in either one of the lines or both lines is considered a failure or a fault.

Let us first look at defects that are only slightly larger than the line width w . This situation is shown in Fig. 3, and the critical area associated with it is double that of a single line:

$$A(\chi) = 2L(\chi - w). \quad (10)$$

The critical areas for these two lines continue to grow this way as a function of defect size until they meet halfway between the lines, as shown in **Figure 4**. This happens for the defect size $\chi = 2w + s$.

For defects larger than $2w + s$ the critical area increases only outside the two lines. **Figure 5** shows the two dashed lines that form the boundaries of the critical areas in that case. According to the diagram the distance between these lines is $\chi + s$. The critical area is therefore given by $L(\chi + s)$.

The preceding discussion can be summarized by writing the functional dependence of the critical area as

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (11a)$$

$$A(\chi) = 2L(\chi - w) \quad \text{for } w \leq \chi \leq 2w + s, \quad (11b)$$

$$A(\chi) = L(\chi + s) \quad \text{for } 2w + s \leq \chi < \infty. \quad (11c)$$

This function is plotted in **Figure 6**.

Combining (11) with the defect size distribution of (4) gives the average number of faults

$$\lambda = \int_w^{2w+s} 2L(\chi - w) \frac{\chi_0^2 \bar{D}}{\chi^3} d\chi + \int_{2w+s}^{\infty} L(\chi + s) \frac{\chi_0^2 \bar{D}}{\chi^3} d\chi. \quad (12)$$

Evaluation of the integrals gives

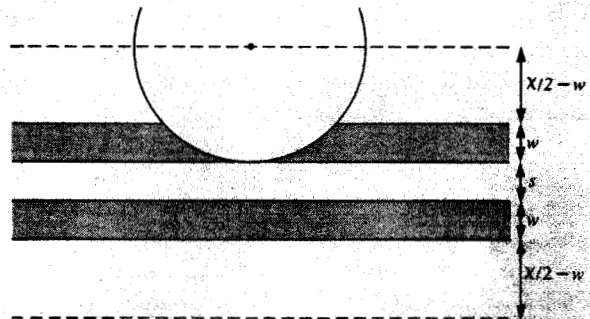


Figure 5

Critical area of two long lines for defects that are larger than $2w + s$.

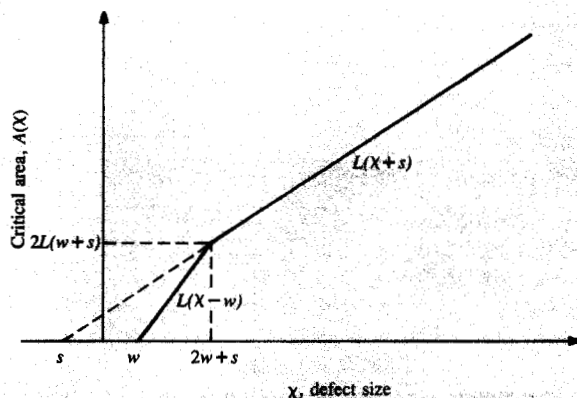


Figure 6

Critical area as a function of defect size for two long conductive lines.

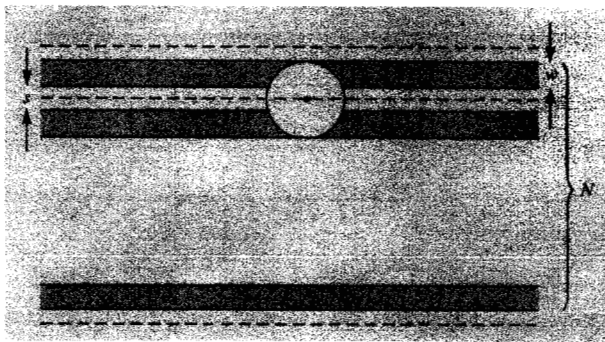


Figure 7

Array of N long conductors, each one of width w and each pair spaced a distance s apart.

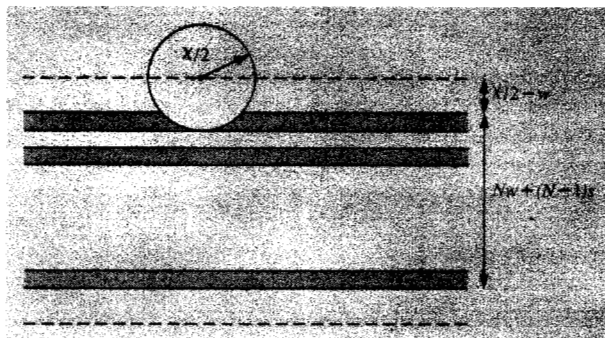


Figure 8

Diagram pertinent to the determination of the critical area of N conductors for very large defects.

$$\lambda = \frac{L(3w + 2s)\chi_0^2 \bar{D}}{2w(2w + s)} \quad (13)$$

From this result we can define

$$\bar{A}_2 = \frac{L(3w + 2s)\chi_0^2}{2w(2w + s)} \quad (14)$$

as the average critical area for two very long conductors. This is an interesting result and merits some discussion.

6. The proximity effect

Combining the formulas in (8) and (14) allows us to express the critical area for two lines in terms of the critical area of the single line. This results in

$$\bar{A}_2 = \frac{(3w + 2s)}{(2w + s)} \bar{A}_1 \quad (15)$$

When the separation s between the two conductors becomes large, the critical area of the two lines

$$\lim_{s \rightarrow \infty} \bar{A}_2 = 2\bar{A}_1 \quad (16)$$

is exactly twice the critical area of a single line. This is as expected. For most integrated circuits, however, the spacing between long conductive lines is of the same magnitude (numerical dimension) as the line width. If we therefore let $s = w$, the critical area in (15) becomes

$$\bar{A}_2|_{s=w} = \frac{5}{3} \bar{A}_1 \quad (17)$$

This shows that the critical area of two closely spaced conductors is less than the sum of the critical areas of two single conductors. We call this the "proximity effect." It is caused by the increase of critical area as a function of defect size. The critical area of each line will grow independently until these two areas meet halfway between the two lines. The merging occurs when the defect size is equal to twice the line width plus the spacing. For defects larger than this, the critical area increases as a function of defect size as it did for a single line. The effect is therefore independent of the defect size distribution. Use of the $1/\chi^3$ distribution in this example merely simplified the mathematics. Other size distributions affect only the magnitude, not the principle involved here.

The proximity effect does not imply that a manufacturer of integrated circuits must cram all the photographic patterns closer together. Although the sensitivity to defects causing open circuits may be lowered in this way, the sensitivity to defects causing short circuits is increased by the tighter spacing of the patterns. The optimal situation therefore depends delicately on the defect densities and size distribution of the defects causing open and short circuits.

7. A large number of long conductors

Most integrated circuit chips have a large number of interconnecting wires. In computer logic circuits and memory chips, such connections consist of long parallel lines. It is therefore useful to investigate the critical area of N parallel conductors, each of width w , and each pair spaced by distance s , as shown in Figure 7. All these lines are assumed to be of the same length L .

The key to critical area calculations is the failure criterion. In this case we assume that an open circuit in any of the lines constitutes a failure for the entire array. If we had assumed, for instance, that an open circuit in any pair of lines caused a failure, while open circuits in single lines are permissible, the results would come out quite differently. We investigate that situation later.

When an open circuit in any of the parallel wires results in a failure, the minimum defect size needed to cause this problem is the same as the width of the line. For defects slightly larger than this, the critical area increases linearly for each line independently. We therefore have

$$A(\chi) = NL(\chi - w). \quad (18)$$

This area grows as a function of defect size until the boundaries of the critical areas between each pair of lines meet. As was the case with two conductors, this happens halfway between the lines when the defect size $\chi = 2w + s$. The condition for this is shown in Fig. 7. Substituting this defect size in (18) shows that the critical area at this point is given by $NL(w + s)$.

What happens for defect sizes that are larger than the ones we have considered until now? In that case the critical area increases only in the region outside the parallel conductors, as shown in Figure 8. From the diagram it can be determined that the space between the top and bottom dashed line is $2(\chi/2 - w) + Nw + (N - 1)s$. It is consequently possible to write the entire critical area function as

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (19a)$$

$$A(\chi) = NL(\chi - w) \quad \text{for } w \leq \chi \leq 2w + s, \quad (19b)$$

$$A(\chi) = L\{\chi + (N - 2)w + (N - 1)s\} \quad \text{for } 2w + s \leq \chi < \infty. \quad (19c)$$

This result is depicted graphically in Figure 9.

The critical area in (19) can be combined with the size distribution in (4) to give the expected or average number of faults in these conductors. Again only part (4b) of the size distribution contributes; an evaluation of the pertinent integrals gives

$$\lambda = \frac{L\{(N + 1)w + Ns\}\chi_0^2 \bar{D}}{2w(2w + s)}, \quad (20)$$

which shows that the average critical area is

$$\bar{A}_N = \frac{L\{(N + 1)w + Ns\}\chi_0^2}{2w(2w + s)}. \quad (21)$$

This last result can also be expressed in terms of A_1 , the critical area of a single conductive line, to give

$$\bar{A}_N = \frac{(N + 1)w + Ns}{2w + s} \bar{A}_1. \quad (22)$$

When the spacing between the lines is very large, we obtain

$$\lim_{s \rightarrow \infty} \bar{A}_N = N\bar{A}_1, \quad (23)$$

which equals the total critical area of N independent single lines. This is what one would expect.

Let us next look at the condition where the space between the lines is exactly the same as the line width. In this case,

$$\bar{A}_N|_{s=w} = \frac{2N + 1}{3} \bar{A}_1, \quad (24)$$

and for a very large number of conductors we have

$$\bar{A}_N|_{N \gg 1, s=w} \approx \frac{2}{3} N\bar{A}_1. \quad (25)$$

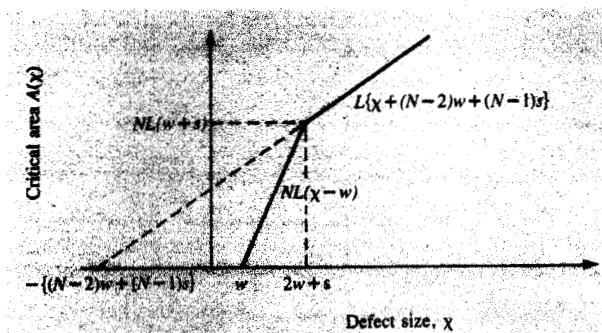


Figure 9

Critical area as a function of defect size for N parallel conductors.

This suggests that the total critical area is less than the sum of the critical areas of N conductors that are spaced far apart. This illustrates the proximity effect once more. However, it does even more than that. If we combine the results in (16) and (17), we see that the critical area for two closely spaced conductors is 5/6 times the critical area of two parallel lines that are separated by a large distance. This number is larger than the 2/3 obtained with the N parallel lines in (25). It indicates that the proximity effect becomes more pronounced when the number of conductors is increased.

8. Other critical areas associated with a large number of parallel conductors

Earlier in this paper a comment was made about the complete duality that exists between open and short circuits. This still holds for the case of N conductors. We must, however, realize that there are only $N - 1$ spaces associated with N conductors. Taking this into account and swapping the width w and spacing s gives a critical area

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq s, \quad (26a)$$

$$A(\chi) = (N - 1)L(\chi - s) \quad \text{for } s \leq \chi \leq 2s + w, \quad (26b)$$

$$A(\chi) = L\{\chi + (N - 3)s + (N - 2)w\} \quad \text{for } 2s + w \leq \chi < \infty. \quad (26c)$$

If the defect size distribution for these defects can also be described by (4), the average critical area becomes

$$\bar{A}_{s,N-1} = \frac{L\{Ns + (N - 1)w\}\chi_0^2}{2s(2s + w)} \quad (27)$$

for short circuits in N long conductors with $N - 1$ spaces.

The subscript s has been added to indicate short circuits.

Such a designation, however, is completely optional. In most practical cases an index i is used to indicate the defect type.

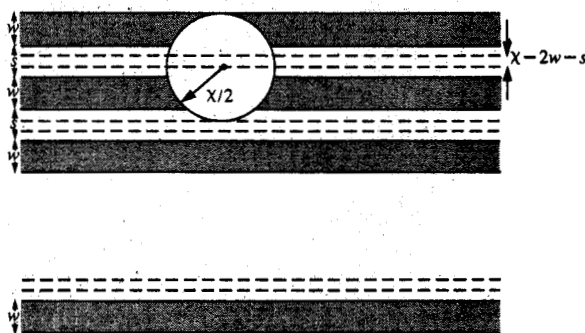


Figure 10

Critical area for open circuits grows in the space between conductors when two or more lines have to be open-circuited to cause a failure.

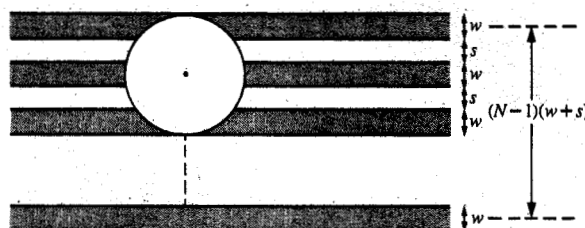


Figure 11

Defect size that will always cause two or more lines to fail when its center falls anywhere between the two dashed horizontal lines indicated at the right.

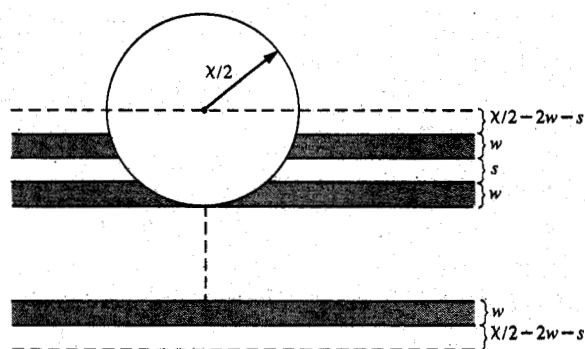


Figure 12

Upper and lower boundaries for large defects that will cause at least two conductors to be open-circuited.

This makes it possible to categorize the open and short circuit critical areas corresponding to each photolithographic process step.

Also pointed out in the previous section was the critical area dependence on failure criteria. Let us take a look at an example. Take the case where we can afford to have an open circuit in a single line, but not in two or more lines. This can happen in circuits with redundancy. Therefore, the smallest defect that causes a failure is $\chi = 2w + s$, as can be seen in Fig. 7. If this defect were placed slightly lower than shown, it would cause only one line to be open-circuited, which is permissible. Even somewhat larger defects can fall on this array of wires and cause only one line to be open-circuited.

The critical area in which the center of these defects must fall will again increase linearly with defect size. This critical area, however, increases between adjacent conductors, as indicated by the dashed lines in Figure 10. From geometrical considerations it is possible to deduce that the distance between these dashed lines is $\chi - 2w - s$. As we saw before, there are $(N - 1)$ of these spaces between the conductors. The critical area therefore becomes $L(N - 1)(\chi - 2w - s)$. This critical area will grow until it corresponds to a defect size that will always open up two lines when its center falls within the array of wires. The smallest defect for which this condition occurs is shown in Figure 11. It has a diameter $\chi = 3w + 2s$. For defects larger than this, the critical area can grow only outside the wires. This is illustrated in Figure 12. In this case the boundary of the critical areas is formed by the locus of centers of defects that just open-circuit the top two or bottom two conductors. If the center of the defect shown had fallen slightly higher, it would have open-circuited only one line, which is allowed according to the failure criterion. By adding all the vertical dimensions in Fig. 12 it can be determined that the distance between the top and bottom dashed line is $\chi + (N - 4)w + (N - 3)s$. With the preceding results the critical area can be expressed as

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq 2w + s, \quad (28a)$$

$$A(\chi) = L(N - 1)(\chi - 2w - s) \quad \text{for } 2w + s \leq \chi \leq 3w + 2s, \quad (28b)$$

$$A(\chi) = L\{\chi + (N - 4)w + (N - 3)s\} \quad \text{for } 3w + 2s \leq \chi \leq \infty. \quad (28c)$$

When this function is combined with the defect size distribution in (4), we obtain

$$\bar{A}_{N2} = \frac{L\{(N - 1)w + Ns\} \chi_0^2}{2(2w + s)(3w + 2s)} \quad (29)$$

for the average critical area. This result can be compared to the average critical area for N lines in (21) to give

$$\bar{A}_{N2} \Big|_{N \gg 1} \approx \frac{w}{3w + 2s} \bar{A}_N, \quad (30)$$

when N is much larger than 1.

The critical area in (29) is therefore smaller than the critical area of N long lines in (21). It demonstrates that critical areas differ if different failure definitions are associated with the same photolithographic patterns. Often changes in test conditions result in radical changes of the critical areas. The failure criteria must therefore be clearly known before the critical areas can be determined for any integrated circuit pattern.

9. A single conductor on a chip

The critical areas that we have considered until now all become very large for large defects. Experience has furthermore taught us that the likelihood of such large defects occurring is very small, as exemplified by the $1/\chi^3$ distribution. In practice it has therefore proven useful to limit the maximum critical area to the chip area. A couple of examples are treated in this section.

Consider a chip of length L and height H with a single conductor, as shown in Figure 13. This conductor is again a line of width w and length L . It is centered on the chip so that its center line is exactly $H/2$ from both the top and bottom sides of the chip. Any open circuit in the pattern is considered a failure. The critical area is restricted to those defects whose centers fall inside the chip. This restriction has no effect on defects that are slightly larger than the line width w . For them the critical area grows as a function of defect size just as it did for the unbounded line in (1b). However, when the boundaries of the critical areas coincide with the edges of the chip, the growth stops. At this point the critical area is equal to the chip area. The defect size that corresponds to this area can be determined from the dimensions in Fig. 13. With its center at the top edge of the chip, such a defect must be large enough to cause an open circuit in the line. This happens when the radius is equal to $H/2 + w/2$. The defect size $\chi = H + w$, therefore, belongs to a defect which always causes an open circuit in the conductor, no matter where its center falls on the chip. Defects larger than this size will have the same property.

The critical area for the conductive line on a chip can consequently be written mathematically as

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (31a)$$

$$A(\chi) = L(\chi - w) \quad \text{for } w \leq \chi \leq H + w, \quad (31b)$$

$$A(\chi) = LH \quad \text{for } H + w \leq \chi < \infty. \quad (31c)$$

A plot of this function is shown in Figure 14.

The result in (31) can be combined by integration with the defect size distribution to give an average critical area

$$\bar{A}_{1c} = \frac{LH\chi_0^2}{2w(H + w)}. \quad (32)$$

The superscript 1 again indicates that this is the critical area for a single conductor, while the subscript c is used to indicate that the critical area is restricted to the chip size.

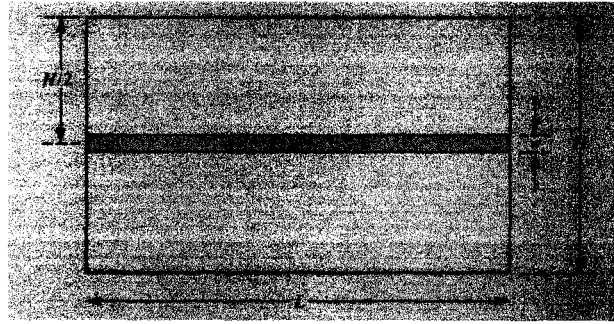


Figure 13

Chip with a single conductor. Critical areas are limited to the chip area.

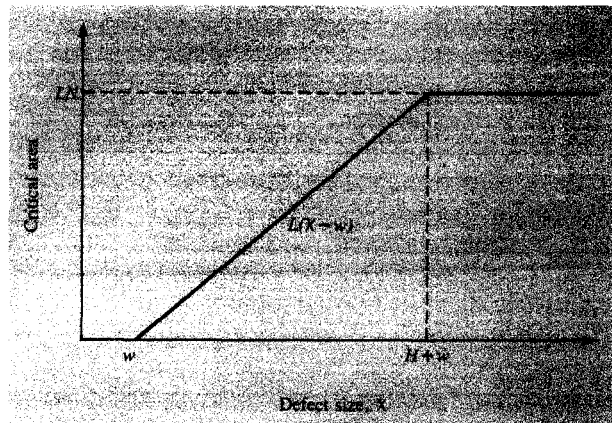


Figure 14

Critical area of a single conductor on a chip.

Expressing this result in terms of $\bar{A}_1$, the unrestricted or unbounded critical area of a single line of (8) gives

$$\bar{A}_{1c} = \frac{H}{H + w} \bar{A}_1. \quad (33)$$

Typical chip dimensions are $H = 5$ mm and line widths w are usually about $2 \mu\text{m}$. As a result $\bar{A}_{1c} \approx 0.9996 \bar{A}_1$. This shows that for all practical purposes the difference between $\bar{A}_{1c}$ and $\bar{A}_1$ is indeed negligible.

In the preceding example the conductor was centered halfway on the chip. We next look at the situation where such a conductor is offset from center, as shown in Figure 15. For small defects the critical area again grows as a function of defect size, as it did for the single line in (1b). It does so until the boundary of this area reaches the nearest

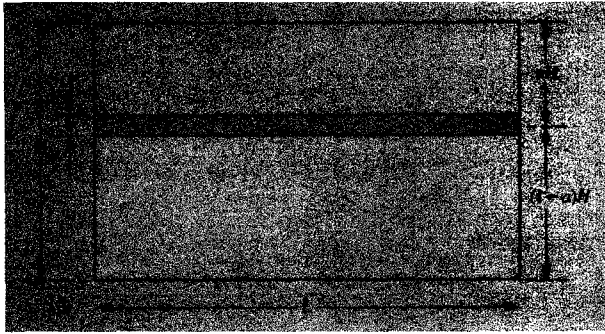


Figure 15

Chip with an off-center conductor.

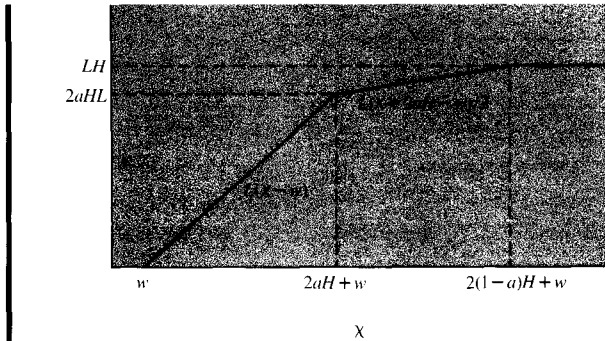


Figure 16

Critical area of an off-center conductor.

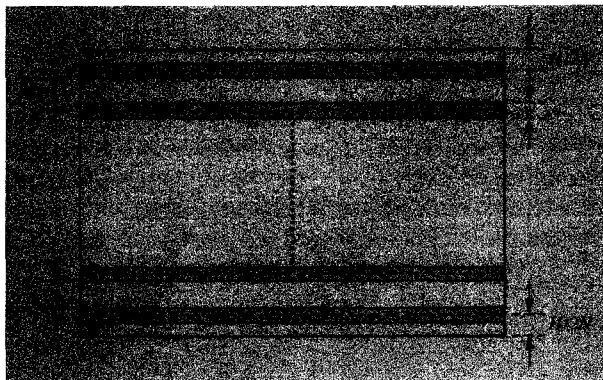


Figure 17

Chip with a large number of conductive lines.

edge of the chip. It can be deduced from Fig. 15 that this happens for a defect of radius $aH + w/2$. For defects larger than this the critical area grows only in one direction. This growth continues until the boundary of the critical area coincides with the bottom edge of the chip in Fig. 15. The defect radius corresponding to this condition is $(1 - a)H + w/2$.

The above approach can be used to determine that the critical area in this case becomes

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (34a)$$

$$A(\chi) = L(\chi - w) \quad \text{for } w \leq \chi \leq 2aH + w, \quad (34b)$$

$$A(\chi) = L(\chi + 2aH - w)/2 \quad \text{for } 2aH + w \leq \chi \leq 2(1 - a)H + w, \quad (34c)$$

$$A(\chi) = LH \quad \text{for } 2(1 - a)H + w \leq \chi < \infty. \quad (34d)$$

The graph of this function is plotted in Figure 16.

Combining the areas in (34) with the defect size distribution gives an average critical area

$$\bar{A}'_{ic} = \frac{LH\{w + 4a(1 - a)H\} \chi_0^2}{2w(w + 2aH)\{w + 2(1 - a)H\}}. \quad (35)$$

The prime has been used to indicate that this result differs from (32), even though for $a = 1/2$ the two results are the same. It can furthermore be seen that

$$\lim_{H \rightarrow \infty} \bar{A}'_{ic} = \frac{L\chi_0^2}{2w}, \quad (36)$$

which indicates that this critical area also approaches the critical area of the unbounded single conductive line (8).

10. A large number of conductive lines on a chip

Let us next investigate the critical area for a large number of parallel conductive lines on a chip. With equal spacing between N lines the problem can be simplified by assuming that the center of the top and bottom lines is spaced a distance $H/2N$ from the top and bottom edges of the chip. This is shown in Figure 17. The pitch, or center-to-center spacing of the lines, in this case is H/N .

As in Section 7, we assume that an open circuit in any of the conductors will cause the chip to fail. For defects slightly larger than w the critical area varies linearly as it does in (18) and (19b). This linear growth in area as a function of defect size continues until the critical areas between the lines merge, or until the critical areas associated with the top and bottom lines reach the top or bottom edge of the chip. With the dimensions chosen in this example all these conditions occur for the same defect of size $H/N + w$.

The critical area therefore is simply

$$A(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (37a)$$

$$A(\chi) = NL(\chi - w) \quad \text{for } w \leq \chi \leq H/N + w, \quad (37b)$$

$$A(\chi) = HL \quad \text{for } H/N + w \leq \chi < \infty. \quad (37c)$$

This function looks the same as that of the single line in Fig. 14 except that the slope of the ramp is N times steeper. When the critical area in (37) is averaged over the defect size distribution, the result becomes

$$\bar{A}_{Nc} = \frac{LH\chi_0^2}{2w(H/N + w)} \quad (38)$$

for the average critical area.

If we take the space between the lines in Fig. 17 to be s , then the pitch of the pattern is

$$H/N = w + s. \quad (39)$$

As a result we can write (38) as

$$\bar{A}_{Nc} = \frac{LH\chi_0^2}{2w(2w + s)}, \quad (40)$$

which in terms of the average critical area of N unbounded lines (21) becomes

$$\bar{A}_{Nc} = \frac{H}{(N + 1)w + Ns} \bar{A}_N. \quad (41)$$

Once more using (39) allows us to write this result as

$$\bar{A}_{Nc} = \frac{H}{H + w} \bar{A}_N. \quad (42)$$

This is the same relationship that existed between the bounded and unbounded critical area of a single conductor in (33). It therefore demonstrates once more that we introduce a negligible error when the critical area calculation is restricted to the chip's surface. This appears to be valid as long as the dimensions of the photolithographic patterns are small compared to the chip dimensions, even when the chip is crammed full with those patterns.

The preceding results are due to the $1/\chi^3$ defect size distribution. They would also hold for a $1/\chi^4$ distribution, but not for a $1/\chi^2$ distribution. For the latter case the unbounded critical areas become infinitely large and the mathematics becomes somewhat more difficult. Nevertheless, a $1/\chi^2$ size distribution would have great practical value and it is unfortunate that it has not yet shown up in the real world. If it did exist, it would become possible to increase yield by shrinking the dimensions of the integrated circuit photolithographic patterns. We investigate the effect of pattern shrinking on critical areas next.

11. Consequences of the $1/\chi^3$ distribution

Competition continually forces manufacturers of integrated circuit components to find ways to lower cost and to increase productivity. One way to achieve this is by decreasing the dimensions of the photolithographic patterns and increasing the number of circuits per wafer. Let us do this to the pattern in Fig. 17. If we shrink all the dimensions by a fraction " a ," we have the following relationships:

$$w' = aw, \quad (43a)$$

$$s' = as, \quad (43b)$$

$$L' = aL, \quad (43c)$$

$$H' = aH. \quad (43d)$$

The prime indicates the shrunken dimensions. The critical area for the smaller chip becomes

$$\bar{A}'_{Nc} = \frac{L'H'\chi_0^2}{2w'(2w' + s')}. \quad (44)$$

When the fractions in (43) are substituted in (44), the result gives

$$\bar{A}'_{Nc} = \bar{A}_{Nc}. \quad (45)$$

This is an amazing result. Equal critical areas imply the same average number of failures $\lambda' = \lambda$ since $\bar{A}'_{Nc}\bar{D} = \bar{A}_{Nc}\bar{D}$. The resulting yields of the large and small chips are therefore equal.

The above phenomenon can readily be explained. The smaller chip has fewer defects on it because of the reduced chip size. These smaller patterns, however, are sensitive to smaller defects. With a $1/\chi^3$ defect size distribution, the increase in small defects exactly equals the decrease in chip area. For a $1/\chi^2$ distribution the average number of faults would go down and the yield would increase when we shrank the photolithographic patterns. For a $1/\chi^4$ distribution the opposite would happen; the yield would be lowered when we decreased the dimensions.

12. Probability of failure

The probability of failure has been discussed in [1]. It relates to the critical area by

$$A_c = \theta A, \quad (46)$$

where A_c is the critical area, θ the probability of failure, and A the chip area. The probability of failure is therefore obtained by dividing the critical area by the chip area:

$$\theta = A_c/A. \quad (47)$$

If, for example, we divide the critical area of the chip with N conductors by the chip area $A = LH$, the probability of failure becomes

$$\theta(\chi) = 0 \quad \text{for } 0 \leq \chi \leq w, \quad (48a)$$

$$\theta(\chi) = N(\chi - w)/H \quad \text{for } w \leq \chi \leq H/N + w, \quad (48b)$$

$$\theta(\chi) = 1 \quad \text{for } H/N + w \leq \chi \leq \infty. \quad (48c)$$

This result states that the probability of failure is zero for defects of size $\chi < w$, the probability of failure is one for defects of size $\chi > H/N + w$, and the probability of failure varies linearly with defect size for defects of all other sizes, viz., $w < \chi < H/N + w$.

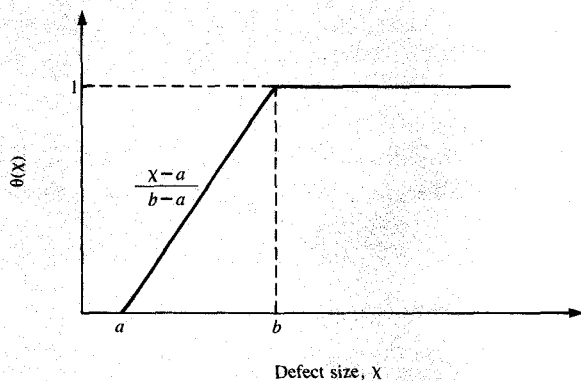


Figure 18

A general probability of failure curve.

The average probability of failure can also be obtained with (47). Thus, dividing the average critical area in (40) by the chip area LH gives the result

$$\bar{\theta} = \frac{\chi_0^2}{2w(2w + s)}. \quad (49)$$

It should be noted here that for the shrunken chip defined in (43), (44), and (45) of the previous section the average probability of failure becomes

$$\bar{\theta}' = \bar{\theta}/a^2. \quad (50)$$

Since a is smaller than one, this result implies that $\bar{\theta}'$ is larger than $\bar{\theta}$. The average probability of failure therefore increases when we decrease the size of the photolithographic patterns. This stands to reason. The smaller patterns are sensitive to small defects and there are more of these. Since the defect size distribution is incorporated in (49) and (50), this effect is properly accounted for.

The concept of probability of failure as a function of defect size has advantages over the direct determination of critical areas. In computer logic and memory chips identical circuits are used a large number of times. The probability of failure as a function of defect size, as well as the average probability of failure, for one circuit is exactly the same as that for ten thousand circuits or more. It is therefore possible to determine the probability of failure for a single circuit or a small number of circuits and use this result to determine the critical area for a large number of circuits.

A number of computer programs have been developed at IBM to determine the probability of failure curves for various circuits. Analytical programs that calculate the increase of probability of failure as a function of defect size were developed by R. W. Bartoldus and N. F. Brickman at the IBM Laboratory in East Fishkill, New York, and by G.

Guhman at IBM Burlington, Vermont. Also in Burlington, W. N. Kuschel and D. H. Withers developed a simulation program for determining the probability of failure. In their approach, circular defects of different sizes are superimposed in random locations on the actual design patterns. The computer program then checks for open or short circuits in the patterns. The fraction of defects, of a given size, that cause a failure equals the probability of failure for that defect size. This approach has been used in subsequent programs by J. Carter of IBM Burlington and K. Barkley of IBM East Fishkill. The latter's program is now being used to minimize the defect sensitivities of new integrated circuit chip designs.

The probability of failure as a function of defect size is also useful in another way. The probability of failure curve shown in Figure 18 is general. The average probability of failure obtained from this, and the defect size distribution in (4), is

$$\bar{\theta} = \frac{\chi_0^2}{2ab} \quad (51)$$

if $a > \chi_0$. We can therefore simplify the critical area calculation for a large number of cases that have the same probability of failure curves as in Fig. 18.

It is also possible to combine the probability of failure curve shown in Fig. 18 with other size distributions. For the $1/\chi^n$ defect size distribution in (3b) the average probability of failure is

$$\bar{\theta} = \frac{2(b^{n-2} - a^{n-2})\chi_0^{n-1}}{(n+1)(n-2)(b-a)a^{n-2}b^{n-2}}, \quad (52)$$

provided $a > \chi_0$. Another defect size distribution formerly investigated by this author is the exponential distribution

$$D(\chi) = \frac{\bar{D}e^{-\chi/l}}{l}, \quad (53)$$

where l is a parameter. When this is combined with the probability of failure curve of Fig. 18, the average probability of failure becomes

$$\bar{\theta} = \frac{l(e^{-a/l} - e^{-b/l})}{b-a}. \quad (54)$$

Other results are possible with different hypothetical defect size distributions. The question, of course, is which of these distributions fits the actual data best. How this is determined is the topic of the next section.

13. Verification of the defect size distribution

The nature of the defect distribution can be investigated with defect monitors. Such monitors have to be designed to detect open or short circuits; they are known as open circuit detectors and short circuit detectors. A defect monitor is essentially an array of a large number of parallel conductors like the ones described in Sections 7 and 10. To make an open circuit detector, the ends of the parallel wires are

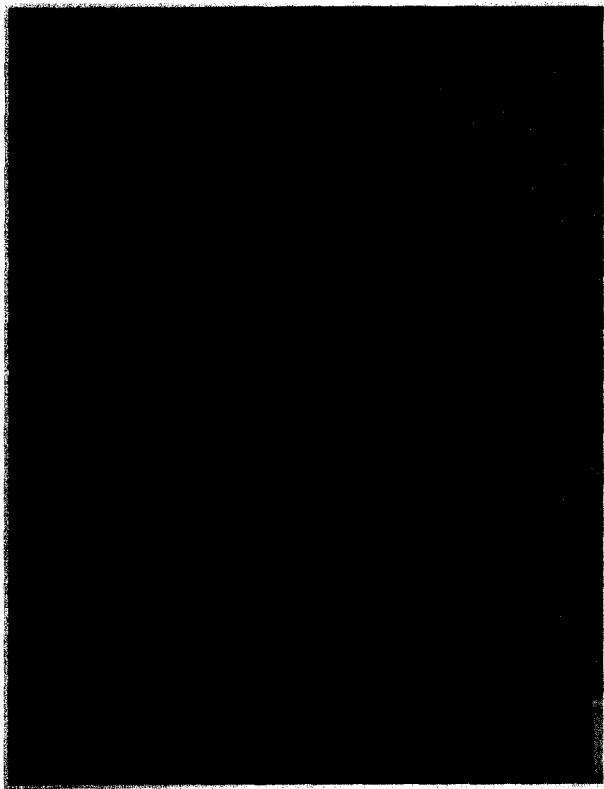


Figure 19

Open circuit defect monitor.

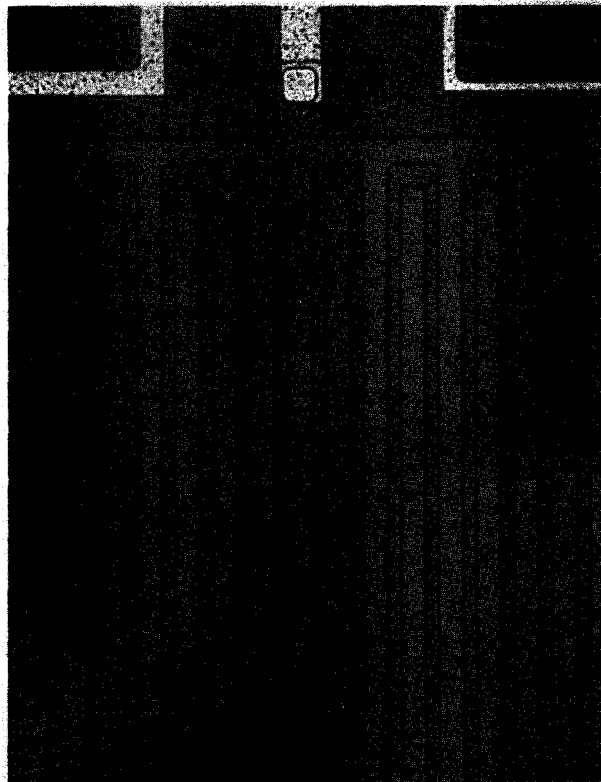


Figure 20

Short circuit monitor.

connected alternately to form a serpentine line like the one shown in **Figure 19**. The monitor in that figure was designed by D. Thomas of IBM Burlington and was part of an experimental chip named "YATS" or "yield analysis test site." The yield theory developed by this author and described in [2, 3] was based on data obtained with the defect monitors on that test site. Similar open circuit detectors have been described by A. C. Ipri and I. C. Sarace of RCA [4].

The YATS chip also contained short circuit monitors. In this case the ends of the parallel conductors are connected in such a way that two interspersed comblike structures are formed. These structures have also been referred to as interdigitated fingers. A YATS short circuit detector of this type is shown in **Figure 20**. Ipri and Sarace had similar short circuit monitors [4].

The probabilities of failure and the critical areas for these defect monitors are the same as those for the array of parallel conductors on a chip. These were described in the previous sections. Therefore, for an open circuit monitor of length L and height H , the critical area is identical to the one

given in Eq. (38). This, of course, presumes that the $1/\chi^3$ size distribution is valid, as it appeared to be in the case of the YATS data. The critical areas for different size distributions are discussed later in this section.

The critical area of a short circuit monitor is the same as that of the open circuit monitor except that the symbol s for spacing is swapped with w for line width to give a critical area

$$\bar{A}_s = \frac{LH\chi_0^2}{2s(2s + w)}. \quad (55)$$

In this case it is presumed that the defect size distribution again peaks at defect size χ_0 and falls off as $1/\chi^3$ for larger defect sizes.

To determine the defect size distribution of a photolithographic process requires the use of defect monitors with different line widths and line spacings. We now look at a hypothetical example that shows how this is done in the case of open circuits. It is assumed that the minimum line width is a and that we have monitors with line widths a , $1.5a$, $2a$, $2.5a$, $3a$, and $6a$. The physical area of these

Table 1 Physical area, line width, and relative critical area of six open circuit monitors.

i = Monitor number	A_i = Monitor area	W_i = Line width	Relative critical area
1	A	a	1
2	A	$1.5a$	0.4444
3	A	$2a$	0.25
4	A	$2.5a$	0.16
5	$1.5A$	$3a$	0.1667
6	$2A$	$6a$	0.0556

Table 2 Data from monitors with different line widths.

i = Monitor number	N_i = Sample size	u_i = Number of failing monitors	U_i/N_i	Relative average
1	1721	61	0.0354	1
2	1732	27	0.0161	0.4398
3	1730	13	0.0091	0.2120
4	1725	11	0.0061	0.1799
5	1741	11	0.0060	0.1783
6	1741	3	0.0020	0.0486

monitors is shown in **Table 1** and varies between A and $2A$. According to (40) the critical area is given by

$$\bar{A}_{mi} = \frac{A_i x_0^2}{2w_i(2w_i + s_i)}, \quad (56)$$

where the subscript i indicates the monitor number, A_i the physical monitor area, w_i the line width, and s_i the spacing between lines. If the line spacing is designed to be equal to the line width, the critical area is given by

$$A_{mi} = \frac{A_i x_0^2}{6w_i^2}. \quad (57)$$

Multiplying this quantity by $6a/A$ produces the relative critical area given in Column 4 in **Table 1**.

The average or expected number of failing monitors $\bar{\lambda}_i$ is given by

$$\bar{\lambda}_i = \bar{A}_{mi} \bar{D}, \quad (58)$$

where $\bar{D}$ is the average defect density associated with the photolithographic process that is being monitored with these defect detectors. Since this defect density is the same for all monitors, it follows that the relative average number of faults per monitor $\bar{\lambda}_i$ should be the same as the ratio of the critical areas. Let us see how this works out.

We assume that there are 109 of each one of the monitors 1–6 on a wafer. A batch (or lot) of 16 of these wafers therefore gives us 1744 of each of these monitors. When these are tested, not all the data are valid for analysis.

Stapper [3] and O. Paz and T. R. Lawson, Jr. [5] described methods for eliminating gross yield failures from such monitor data. These are the gross failures that affect entire areas of wafers. Furthermore, in actual data there is the possibility of invalid information caused by tested errors, probe damage, and misprobing. All such faulty information must be removed from the data before it is analyzed. This often requires multiple testing and visual inspections of the wafers. As a result the valid monitor sample size is usually less than the total number of monitors that are available. An example of this can be seen in the typical monitor data tabulated in **Table 2**. Column 3, with the number of failing monitors U_i , contains the valid failure data only. The monitors with invalid data have been excluded from the sample to give the same size N_i shown in Column 2.

The monitor yield for the data is given by

$$Y_{mi} = \frac{N_i - U_i}{N_i} \quad (59a)$$

$$= 1 - U_i/N_i. \quad (59b)$$

If we use Poisson statistics, this can be set equal to

$$Y_{mi} = e^{-\lambda_i}. \quad (60)$$

If $\lambda_i < 0.05$, we can furthermore make the approximation

$$Y_{mi} \approx 1 - \lambda_i. \quad (61)$$

Combining (61) and (59b) gives

$$\lambda_i \approx U_i/N_i, \quad (62)$$

which is tabulated in Column 4 of **Table 2**. If we multiply these results by N_i/U_i , they will be normalized in the same way as the critical areas in Column 4 of **Table 1**. And indeed the numbers look similar.

How good is the relationship between the relative critical areas in Column 4 of **Table 1** and the relative average number of failures per monitor of Column 5 in **Table 2**? A number of statistical techniques are available for quantifying this. The simplest method is use of the correlation coefficient. Programs for calculating this quantity are part of every statistical library for use on digital computers, and even hand calculators are available with the capability to determine it. The closer the correlation coefficient is to a value of one, the better the agreement is between the two sets of data for which it is calculated. In the example given here it comes out to be 0.998. This implies excellent agreement between the two sets of numbers. As a result, the $1/\chi^3$ defect size distribution is an excellent model for these data.

There is a danger in using the correlation technique to verify statistical models. When we obtain a good result we do not know how this compares with the results of other models. Even though in the preceding discussions we have obtained excellent agreement between data and models,

there may be even better models that give even higher correlation coefficients when data and model results are correlated. Exploring this possibility in more detail is the subject of the next section.

14. Optimization of defect size distribution

To optimize the defect size distribution requires a more flexible critical area calculation than the one obtained with the $1/\chi^3$ distribution. Both the $1/\chi^n$ distribution in (3) and the $e^{-\chi/l}$ distribution in (53) are possible candidates. Other distributions may be needed if the data require it. This, however, is beyond the scope of this paper.

For the $1/\chi^3$ distribution the probability of failure in (52) can be multiplied by the monitor area A_i to give the critical area. For an open circuit monitor the quantity a in (52) must be replaced by the line width w_i . Similarly, b in (52) must be replaced by $2w_i + s_i$ so that the average critical area is given by

$$\bar{A}_{mi} = \frac{2(2w_i + s_i)^{n-2} - w_i^{n-2} \chi_0^{n-1} A_i}{(n+1)(N-2)(w_i + s_i)w_i^{n-2}(2w_i + s_i)^{n-2}}. \quad (63)$$

In the case where $w_i = s_i$ the average critical area is given by

$$\bar{A}_{mi} = \frac{(3^{n-2} - 1)\chi_0^{n-1} A_i}{(n+1)(n-2)3^{n-2}w_i^{n-1}}. \quad (64)$$

With this we can calculate the relative critical areas for the monitors in Table 1 for different values of n . When we determine the correlation coefficient for each value of n , it results in the curve shown in Figure 21.

The result shows that the optimum correlation coefficient of 0.9983 occurs for $n = 3.07$. The $1/\chi^3$ model is therefore indeed an excellent model for the data. However, even for $n = 2.5$ and $n = 3.7$ the correlation coefficients exceed 0.995 and therefore indicate that size distributions of $1/\chi^{2.5}$ and $1/\chi^{3.7}$ are also acceptable models for these data.

The preceding results may appear disturbing. If we have such small distinctions in the correlation coefficients, how can we ever establish the nature of the defect size distribution with any degree of accuracy? This can only be done with a more sensitive technique. A straightforward method that can be used for this is a nonlinear least square fitting technique. In this method the difference between the calculated and measured values for each observation is squared. The object then is to find the right combination of parameters in the model to minimize the sum of these squares. The value n in (64) and the defect density D in $\lambda = A_m(n)D$ are the two parameters used in the minimization process. This result gives $n = 3.02$ for the minimum. How this minimum is approached as a function of n is shown in Figure 22. The vertical scale is logarithmic so that in this case the minimum is sharp and well defined.

We now take a look at what happens with the exponential defect size distribution $e^{-\chi/l}$ of (53). Using the probability of

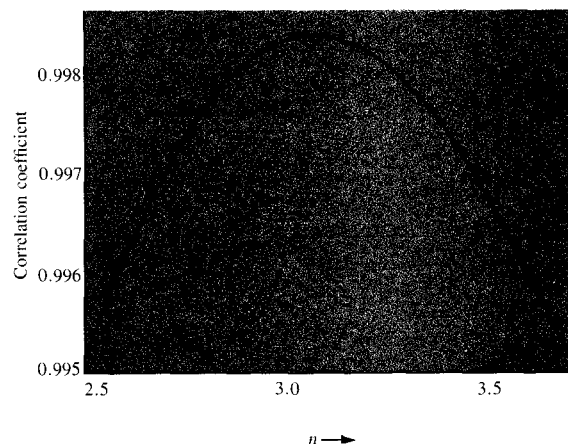


Figure 21

Correlation coefficient as a function of n in the $1/\chi^n$ defect size distribution.

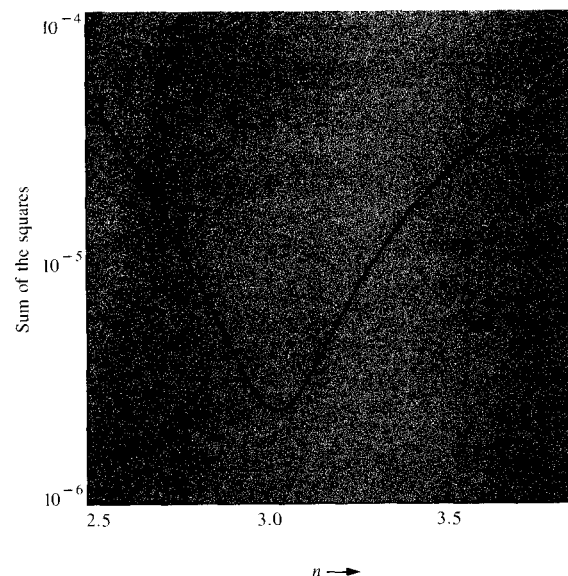


Figure 22

Sum of the squares as a function of the power n in the defect size distribution.

failure in (54), the monitor area A_i , and the relationships $a = w_i$ and $b = 2w_i + s_i$ results in the critical area

$$\bar{A}_{mi} = \frac{LA_i \{e^{-w_i/l} - e^{-(2w_i+s_i)/l}\}}{w_i + s_i}. \quad (65)$$

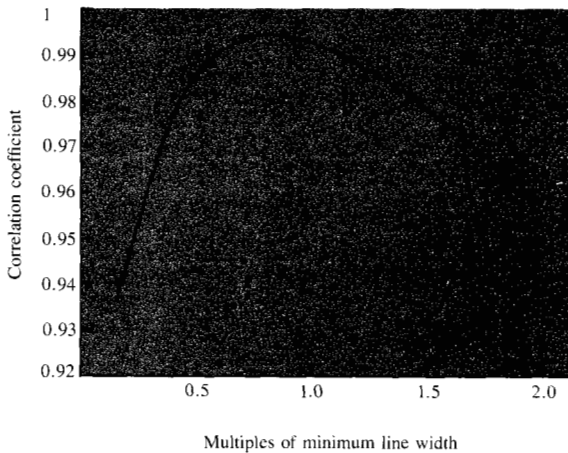


Figure 23

Correlation coefficient as a function of the length parameter in an exponential defect size distribution.

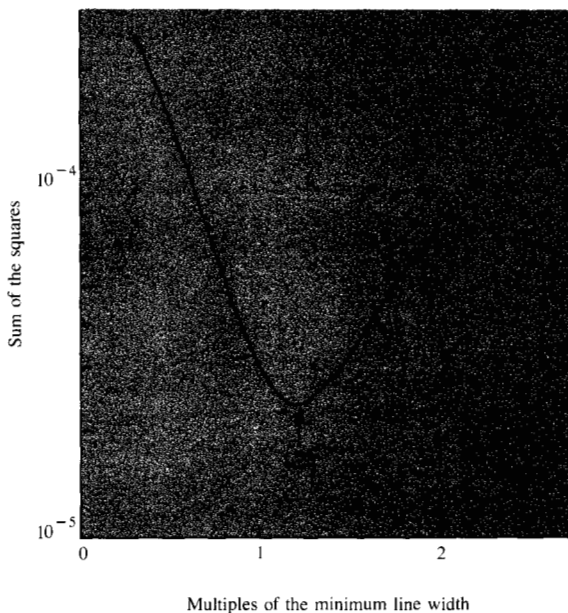


Figure 24

Sum of the squares determination of ℓ for an exponential defect size distribution.

Under the condition where $w_i = s_i$, this result simplifies to

$$\bar{A}_{mi} = \frac{LA_i e^{-w_i/l} (1 - e^{-2w_i/l})}{2w_i} \quad (66)$$

To use this expression in a correlation coefficient comparison required that we express l in multiples of the

minimum line width a in the form $l = ma$. It is now possible to optimize the correlation coefficient as a function of the parameter m . The result of such an exercise is shown in **Figure 23**, where a maximum correlation coefficient of 0.9946 is reached for $l = 0.83a$. This maximum, however, is lower than any of the correlation coefficients shown in **Fig. 21**. This therefore leads to the conclusion that the $1/\chi^3$ defect size distribution is a better model for these data than an exponential distribution.

The above results can also be checked with the sum of the squares. The minimum sum of the squares as a function of l in terms of multiples of line width a is shown in **Figure 24**. The minimum occurs for $l = 1.23a$ and the sum of the squares is 2.36×10^{-5} . This is an order of magnitude higher than the minimum of 2.5×10^{-6} in **Fig. 22** obtained for a $1/\chi^3$ defect size distribution. This allows us once more to conclude that the $1/\chi^3$ distribution is a better model for these data than the exponential distribution.

The reader has probably noted that the two techniques used here for determining the parameter l gave results that differed by approximately 40 percent. This is due to the exponential nature of the distribution and the fact that no statistical technique is perfect. However, we can conclude that for these data a value of $l \approx a$ is near optimum.

15. Conclusions

In this paper the critical areas for large arrays of wiring and defect monitors have been derived. In doing so the "proximity" effect was described and its effect on critical area calculation was evaluated. It was also shown that if we shrink the patterns of long parallel wires in width, spacing, and length, the $1/\chi^3$ defect size distribution results in the same random defect yield for the large and shrunk patterns. Determination of the nature of the defect size distribution is therefore crucial if we want to decrease the cost of integrated circuit manufacture by shrinking the size of the integrated circuit patterns.

Experimental techniques for evaluating the defect size distribution have been described and an example with its results has been discussed.

References

1. C. H. Stapper, "Modeling of Integrated Circuit Defect Sensitivities," *IBM J. Res. Develop.* **27**, 549-557 (November 1983).
2. C. H. Stapper, "Defect Density Distribution for LSI Yield Calculations," *IEEE Trans. Electron Devices* **ED-20**, 655-657 (July 1973).
3. C. H. Stapper, "LSI Yield Modeling and Process Monitoring," *IBM J. Res. Develop.* **20**, 228-234 (May 1976).
4. A. C. Ipri and J. C. Sarace, "Integrated Circuit Process and Design Rule Evaluation Techniques," *RCA Rev.* **38**, 323-350 (September 1977).
5. O. Paz and T. R. Lawson, Jr., "Modification of Poisson Statistics: Modeling Defects Induced by Diffusion," *IEEE J. Solid-State Circuits* **SC-12**, 540-546 (October 1977).

Received November 14, 1983

Charles H. Stapper *IBM General Technology Division, Burlington facility, Essex Junction, Vermont 05452.* Dr. Stapper received his B.S. and M.S. in electrical engineering from the Massachusetts Institute of Technology in 1959 and 1960. After completion of these studies, he joined IBM at the Poughkeepsie, New York, development laboratory, where he worked on magnetic recording and the application of tunnel diodes, magnetic thin films, electron beams, and lasers for digital memories. From 1965 to 1967, he studied at the University of Minnesota on an IBM fellowship. Upon receiving his Ph.D. in 1967, he joined the development laboratory in Essex Junction. His work there included magnetic thin-film array development, magnetic bubble testing and device theory, and bipolar and field effect transistor device theory. He is now a senior engineer in the development laboratory, where he works on mathematical models for yield and reliability management. Dr. Stapper is a member of the Institute of Electrical and Electronics Engineers and Sigma Xi.