INTERNAL VISUAL WORKMANSHIP STANDARD FOR MICROELECTRONIC DEVICES (NASA STD XX-2) AND TRAINING MANUAL

VOLUME II

TRAINING MANUAL

By

J. Gaffney, E. J. Rice, C. D. Root

September 1968

Prepared under Contract No. NAS 12-691

RAYTHEON COMPANY
SPACE AND INFORMATION SYSTEMS DIVISION
Sudbury, Massachusetts 01776

RAYTHEON

ELECTRONICS RESEARCH CENTER

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Mr. Leonard M. Pauplis
Technical Monitor
NAS 12-691
Electronics Research Center
575 Technology Square
Cambridge, Massachusetts 02139

This copy has been reproduced in black and white. The original copies are in color.

INTERNAL VISUAL WORKMANSHIP STANDARD FOR MICROELECTRONIC DEVICES (NASA STD XX-2) AND TRAINING MANUAL

VOLUME II

TRAINING MANUAL

By

J. Gaffney, E. J. Rice, C. D. Root

September 1968

Prepared under Contract No. NAS 12-691
RAYTHEON COMPANY
SPACE AND INFORMATION SYSTEMS DIVISION
Sudbury, Massachusetts 01776

ELECTRONICS RESEARCH CENTER

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

CONTENTS

TRAINING MANUAL
FOR USE WITH
INTERNAL VISUAL WORKMANSHIP STANDARD
FOR MICROELECTRONIC DEVICES
NASA STD XX-2

1. INTRODUCTION

This training manual is intended to be used as a guide to those who are responsible for implementing the Internal Visual Workmanship Standard for Microelectronic Devices. The objective in writing this manual is to provide the reader information pertaining to the physical processes that take place and the manufacturing techniques utilized in the fabrication of semiconductor integrated circuits. The manufacturing processes used in integrated circuit fabrication are, many times, relatively new, and it is essential that those given the responsibility of implementing the NASA Internal Visual Workmanship Standard for Microelectronic Devices be thoroughly familiar with modern integrated circuit fabrication methods.

It is recommended that a training program be conducted for both Government and in-plant personnel to further insure that the provisions of the NASA STD XX-2 will be complied with. A training course outline and schedule is given in appendix A.

The components NASA purchases must do the job they are designed to do upon demand and without failure. In space missions, present day systems cannot be repaired once the mission is initiated. This means the parts must perform as specified throughout the flight. Hence all visual inspection of the parts must be carried out with the utmost care, to insure that each and every component will perform satisfactorily in its intended application.

As with all introductory expositions, the material contained herein can only cover the subject in brief. For those readers who would like more information on the subject, reference material is provided in the bibliography. Some of this reference material is quite basic, and the reader is urged to make use of it wherever possible.

2. PHYSICS OF SEMICONDUCTOR TECHNOLOGY

2.1. Materials used in fabricating semiconductor devices. - In general all materials may be placed in one of three major categories - conductors, semiconductors, or insulators. These categories evolved from the consideration of the material's ability to allow an electric current to flow, which is a function of their atomic structure. A good conductor is a material that has a large number of free electrons. All metals, for example, are conductors of electricity to some extent, but some are much better conductors than others. For example, silver is a much better conductor than aluminum. An insulator is a material which has an extremely poor ability to conduct electricity. Materials falling in this class are many: glass, rubber, and certain plastics, to name but a few. While these materials do conduct electricity in extremely small amounts, for all practical purposes they are classed as insulators. Between the extremes of good conductors and good insulators are a number of materials which are neither good conductors nor good insulators. The elements germanium and silicon fall into this category and are called semiconductors. However, as will be discussed later, the resistivities of both germanium and silicon can be greatly modified by the addition of impurities. While only germanium and silicon as semiconductors are discussed here, there are many other semiconductors. However these two materials are the most frequently used in semiconductor device fabrication today.

2.2. The nature of a single crystal. - Most inorganic solids reveal what is called a crystalline structure when studied under a microscope. The atoms of a crystal are arranged in specific patterns. One of the simpler crystal patterns is that of iron. Each atom occupies a corner of an imaginary cubic structure with one additional atom located in the center of the cube. See Figure 1.

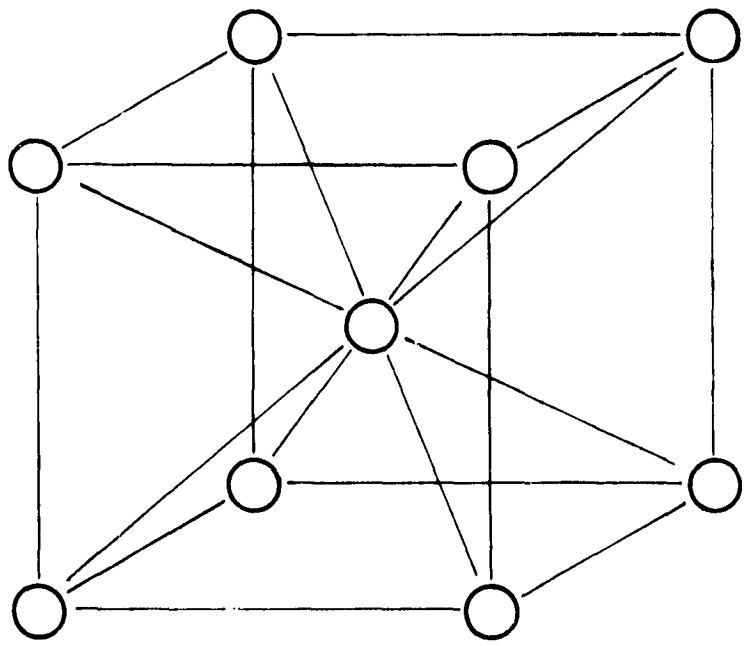


Figure 1. Atomic Crystal Lattice Structure of Iron.

If we were to stack these cubes one upon another and side by side, we would eventually build up a single crystal of iron. It is possible to grow such single crystals in the laboratory using specially designed furnaces. Most of the iron used for industrial purposes is actually polycrystalline. In polycrystalline material, crystallites are formed in a rather sporadic and haphazard fashion, but each crystallite is a single crystal. Most semiconductor devices use single crystal silicon or germanium.

Various methods are utilized in producing a single crystal of silicon or germanium, but all methods depend on a temperature gradient (i.e., temperature difference) existing across the material, and the movement of the temperature gradient through the length of the material to be crystallized. Usually

the material to be crystallized is formed into a cylindrical ingot, varying in size from 1 inch to 2 inches in diameter and upwards of 7 inches long.

2.3. Impurities in semiconductor crystals.

2.3.1. p and n type impurities. - We noted earlier that impurities (or dopants) in semiconductors modify their behavior markedly. Certain materials when added to silicon are known as n-type dopants, and silicon crystals doped with these n-type dopants are called n-type crystals. Other impurities are known as p-type dopants, and produce p-type crystals. An n-type dopant is called a donor impurity, an example of which is arsenic. A p-type dopant is known as an acceptor impurity, an example of which is boron. These nomenclatures are derived from the fact that a donor impurity has the effect of giving extra Negative charges to the crystalline material (electrons) and thus increases electron conduction. An acceptor impurity has the effect of giving extra Positive charges to the crystalline material, which increases conduction by introducing mobile, positively charged sites, called "holes".

The crystal in the pure form is said to be "intrinsic". Adding either donor or acceptor impurities to the crystal lowers the crystal's resistance to electrical current markedly. Use of this effect is made in fabricating resistors in semiconductor networks, which are discussed later in more detail.

If an n-type dopant is added to a p-type crystal, compensation of one for the other takes place, i.e., the resultant effective impurity concentration is the difference of the two concentrations. Therefore if sufficient p-type dopants are diffused into an n-type crystal it will become p-type.

2.3.2. Introducing impurities into the crystalline lattice. There are several methods of introducing impurities into the crystalline lattice to produce crystals having p and n characteristics. One method is called alloying, and was the first method utilized in the mass production of semiconductor

devices. In this method the impurity (the donor or acceptor material) is placed on a small wafer of single crystal material, and both are heated until the melting temperature of the silicon-impurity combination is reached. At this temperature alloying takes place. Some of the impurity material mixes with some of the crystalline material, the amount depending on the total volume of the impurity, and on the temperature used. On cooling the silicon, most of the impurity material is thrown out of the silicon back into the original dopant but enough remains to effectively "dope" the single crystal silicon.

A second method of introducing impurities into the crystalline lattice is by means of diffusion. If, for example, a silicon wafer is placed in a high temperature furnace whose ambient (i.e., surrounding gases) contains one of the doping elements in gaseous form (e.g., arsenic) this impurity will enter or diffuse into the crystal. The penetration depth of the diffusant into the crystal is a function of time and temperature, and can thus be controlled. The diffusant (n-type arsenic in this case) concentration decreases the deeper it penetrates into the crystal. A concentration level exists where the n-type concentration is equal to that of the original concentration (p-type in this case) of the crystal. This point separates the n material from the p material. This transition point between the n and p material is called a p-n junction. If an n-type crystal had been used we would have an area of greater n-type concentration instead of a p-n junction.

Here we see that two methods have accomplished the same result, namely, that a junction between p and n-type material has been formed in a single crystal of silicon. While there are certain advantages in the use of these and other methods, the diffusion method is preferred in the fabrication of modern semiconductor devices.

Still another (and most recent) method of p-n junction fabrication is by means of epitaxial growth. In this method a

layer of single crystal silicon (doped either n or p) is grown directly on a silicon crystal substrate (which can also be doped either n or p). The silicon epitaxial material is initially in a gaseous phase; the parent silicon wafer and the epitaxial material are made to react in an oven at about 1200°C. The free silicon atoms existing in the gaseous phase deposit themselves on the solid silicon substrate, finding correct positions in the silicon lattice structure so as to maintain the same crystalline orientation as the silicon substrate. To form a p-n junction a p type silicon gas (silicon and boron) is deposited on an n-type silicon substrate. As simple as this process appears to be, it has only been successfully used since 1960.

2.3.3. The non-linear electrical characteristics of a p-n junction. - A p-n junction, similar to that which has been described above, has very non-linear voltage current characteristic curves. An example of such a curve is given in Figure 2.

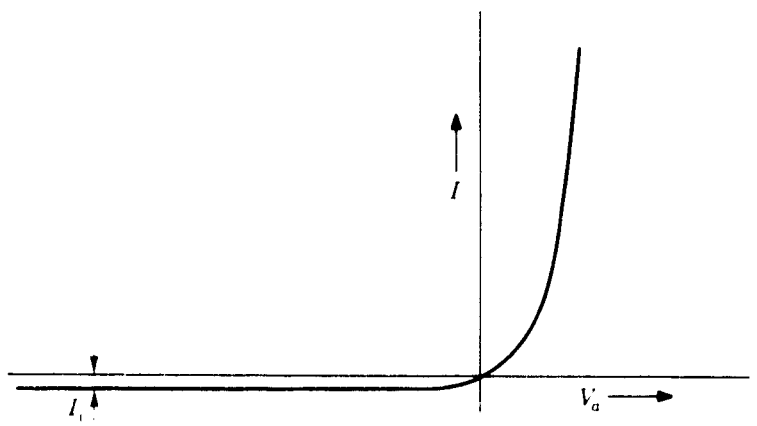


Figure 2.
The current-voltage characteristics of an ideal junction.

Note that in one direction, commonly called the forward direction, the junction resistance is very low, characterized by high currents being available at very low voltages. The reverse characteristic is obtained when the voltage across the p-n junction is reversed. The reverse behavior is characterized by a very low current (in the order of microamps) being obtained at very high voltages. Thus the reverse direction of the p-n junction is characterized by a high resistance. Such a p-n junction has been called a diode and provides rectification when alternating currents are placed across the diode.

2.3.4. Transistor fabrication. - A transistor consists of two p-n junctions back to back, formed within a semiconductor (see Figure 3). The transistor consists of an n-type crystal having a p-type and an n-type region diffused into the original n-type crystal. Referring again to figure 3, it is seen that three regions are defined: the emitter region, the base region and the collector region. The emitter junction is on the top and the collector junction is on the bottom. In operation, excess carriers are injected into the base region through the emitter to base junction and are collected at the collector junction. In normal operation the emitter to base junction has voltage applied in the direction of easy current flow (forward bias) and the collector to base junction has voltage applied in the opposite sense (reverse bias). Small variations in base current produce relatively large variations in collector current. Thus the base current may serve as a modulator for the emitter to collector current.

3. MANUFACTURING TECHNIQUES

3.1. The diffused planar transistor. - In order to fabricate a diffused transistor, it is necessary to restrict the areas of gaseous diffusion into the crystal. Thus regions where impurities can diffuse into the crystal must be restricted by covering most of the surface with a layer of material which is impervious to the diffusant. Silicon dioxide (SiO_2) is a

convenient layer which can act as a diffusion mask, and has the added advantage that it can be grown directly on the silicon surface by oxidation of the silicon. This oxide is sufficiently impervious to impurities so that diffusion will proceed only where one cuts windows through the oxide. Fortunately the cutting out of the windows can be accomplished by using photographic emulsions and etching techniques, so that extremely small areas can be opened in the oxide. The primary size limitation is that of the resolution of the lenses used in the photolithographic process. By successive oxidation, window cutting, diffusion, oxidation, window cutting and diffusion one obtains the transistor as shown in cross-section in Figure 3.

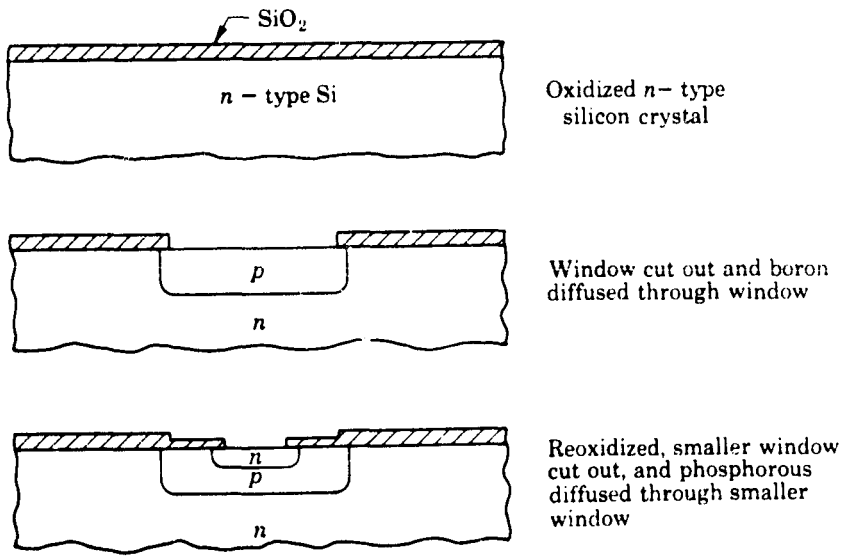
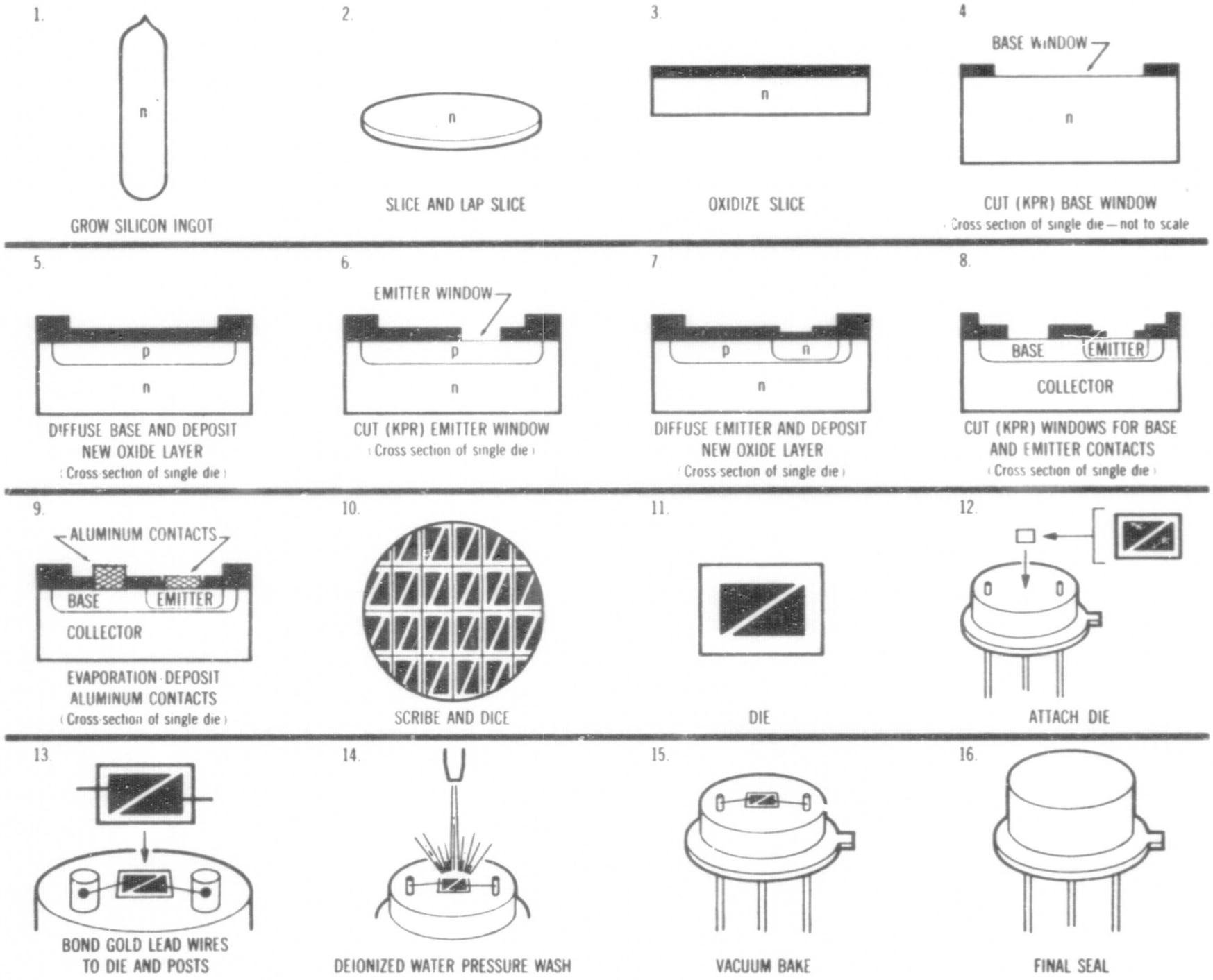


Figure 3.
Formation of a planar transistor.

The step by step process for fabrication of a silicon planar transistor is shown in Figure 4. It is instructive to go through the sequence of steps involved in the fabrication of a silicon planar transistor in detail, since this element is basic to the integrated circuit.

Referring again to Figure 4, in step 1 an ingot of single crystal silicon is grown and subsequently sawed into wafers as shown in step 2. After mechanical and chemical polishing, silicon dioxide is grown thermally on the surface of the crystal, as shown in step 3. (To obtain the SiO_2 cover, silicon is made to react with oxygen at a temperature of 1,000-1,200^{o}C.) The process of cutting holes in the desired parts of the oxide layer is accomplished by utilizing a layer of resist such as Kodak photoresist (KPR) or metal etch resist. Both of these are utilized in device fabrication. The photoresist is spread on the surface of the slice either by spraying or by applying the photoresist in small droplets. The slice is then spun at a very high velocity in order to leave a uniform thin layer. This thin layer is hardened by heating. A mask with a prescribed photographic pattern is placed in contact with the slice and exposed to ultraviolet light. The photoresist (which is actually a photographic emulsion) is then developed. In the developing process that material exposed to light undergoes a physical change. Regions which were <u>not</u> exposed to the light are removed; those regions which are exposed to light remain. (In some other photoresists, the resist that <u>is</u> exposed to the light is removed in the development process.)

The silicon slice is then subjected to a fluoride etch which attacks the silicon dioxide but not the photoresist. In this manner it is possible to cut holes in the oxide corresponding to the photographic pattern which had been previously selected in the mask. After this is accomplished all the remaining photoresist is removed from the slice leaving a window in the silicon dioxide covering as shown in step 4.



Sequence of steps involved in the fabrication of a silicon planar transistor.

Figure 4.

Usually the oxide layer has a thickness in the order of 10,000Å (angstroms) or 1 micron. However for some integrated circuits it could be considerably thicker than this, possibly as high as 19,000 Å.

Into the n-type crystal is then diffused p-type material. This is shown in step 5. Note that the region of intersection of the p-n junction and the surface of the die is covered by silicon dioxide. In other words in the diffusion process diffusion took place underneath the base window perimeter, as well as into the crystal itself. This fact is very important, for it results in the p-n junction being covered by silicon dioxide on the surface of the die, with the beneficial result that reverse currents for such a p-n junction are extremely low, and in general transistors fabricated by this process are exceptionally stable.

A second smaller window is cut in through the silicon dioxide by methods described previously. This is shown in step 6, and will serve as a window for the emitter diffusion. Note that since it is required to regrow the silicon dioxide after each step, the initial uncut areas also increase in thickness, giving a stepped appearance pattern to the silicon dioxide covering. N-type material is now diffused into the crystal via the emitter window thus forming an NPN transistor structure. Windows are then cut into the base and emitter regions for the purpose of making electrical contacts to these areas. The metal used for this purpose is usually aluminum, although other metals such as molybdenum are also in use. The entire die is covered with metal by evaporation, and then again utilizing photoresist methods, unwanted portions of the aluminum are etched away. Finally a good contact of the metal to the silicon is produced by heating the die in order for metal-silicon alloying to take place.

Although we have been discussing a single transistor, it should be noted that many transistors are made simultaneously

on one slice. Step 10 shows an enlarged area of such a slice. Each transistor is separated from the wafer by a process called scribing and dicing. Essentially the process is very much like cutting a pane of glass: a scribe line is made along a parallel to the crystal lattice and the die is broken on the scribe line. The result is an individual transistor as shown in step 11. Note that in the planar process the two metallized areas are contact areas to the base and emitter regions. The bottom of the die serves as the collector. Since such a die is extremely small (on the order of .040" x .040") it must be attached to a larger surface area (called a header) if it is to be utilized as an electronic component.

In step 12, the die is attached to the header by a soldering process in which the header, previously gold plated, is heated to an elevated temperature. When this temperature is reached, the gold alloys with the silicon and when the header is cooled, the die becomes firmly attached to the gold plated header. (Several processes are used to attach the die to the package bed, of which the gold-silicon soldering process described is but one method. In some cases a glass frit is used as the attaching agent). Wires are next bonded from the metallized areas to insulated posts, which pass through a glass-to-metal seal on the header. The external collector lead is butt welded to the package and thus makes electrical contact to the collector region of the die. Subsequent washing and baking procedures prepare the transistor for the final seal. In step 16, the transistor package is hermetically sealed by welding a can on the header.

3.2. <u>The physical appearance of a planar transistor.</u> - While electrically the planar transistor is characterized by extremely low leakage currents and overall electrical stability, there are also important visual differences between the planar and non-planar transistor when the transistor is viewed with a microscope having vertical illumination. A color photograph of

a planar transistor is shown in Figure 5. In particular various colors following a well defined pattern are noticeable. These colors, and the areas they define, give us a clue as to how the transistor was fabricated. Figure 6 shows a color schematic of a cross-section of a planar transistor similar to that depicted in Figure 5.

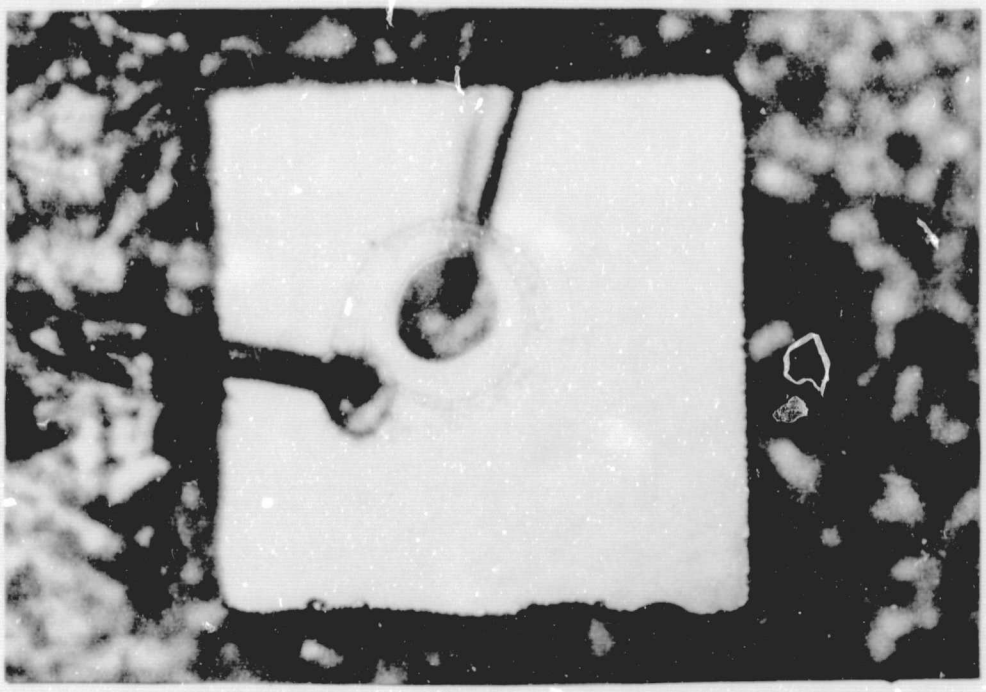


Figure 5 - Planar transistor.

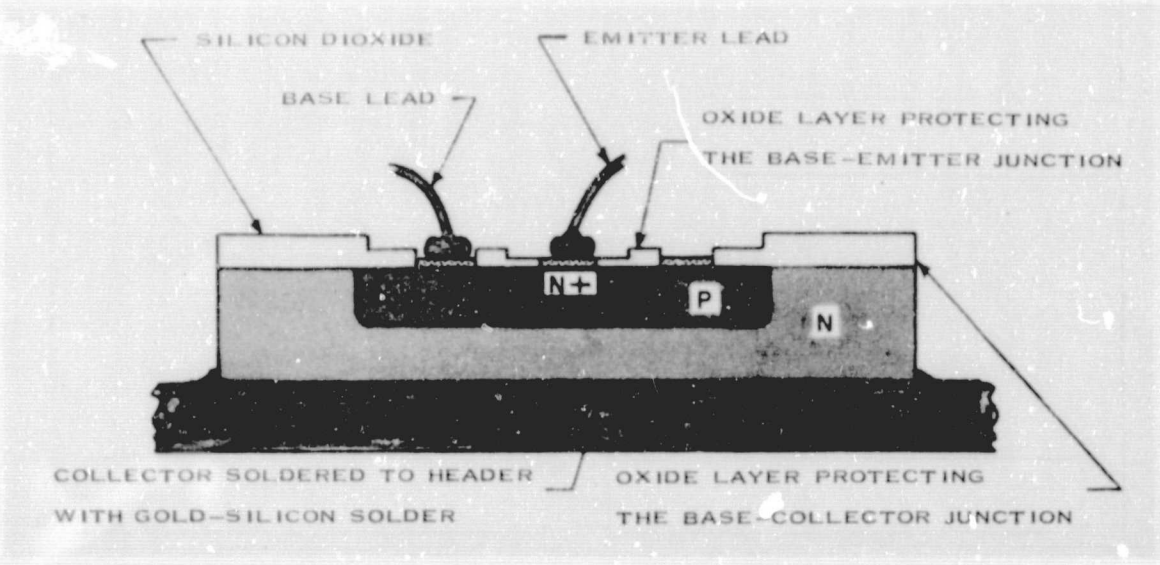


Figure 6 - Illustration of cross section of Planar transistor.

The various colors seen in Figure 5 are produced by a light interference phenomenon as a result of the different thicknesses of oxide layers. In general the thinner the oxide layers the brighter the colors. Within any one thickness range the color and corresponding oxide thickness is given as follows:*

COLOR	THICKNESS (IN MICRONS)			
Gray	0.01			
Tan	0.03			
Brown	0.05			
Blue	0.08			
Violet	0.10	0.28	0.46	0.65
Blue	0.15	0.30	0.49	0.68
Green	0.18	0.33	0.52	0.72
Yellow	0.21	0.37	0.56	0.75
Orange	0.22	0.40	0.60	
Red	0.25	0.44	0.62	

The different color areas, then, define regions where the oxide was etched out to allow for subsequent diffusions. Note however that since the diffusion extends under the window perimeter the distinct color lines do not define the actual junction locations, but slightly overlap these areas. The overlap is usually very small, and in all practical considerations, they can be said to outline the junction areas. (However to compensate for difficulty in precisely locating a junction, the accompanying specification has defined the junction as the area between the oxide step and a distance 1/4 mil. away from the last diffusion.) For a correct interpretation of the microelectronic visual standard it is very important that the reader thoroughly understands the relationship between oxide thicknesses and junction areas.

*More complete color relationships can be found in ASD-TDR-63-316, Volume XII.

3.3. The monolithic integrated circuit.

3.3.1. Resistor/transistor isolation in a monolithic integrated circuit. - A discussion of the construction of an integrated circuit follows directly from the above description of the planar process. A monolithic integrated circuit consists of various transistors that have been diffused into a silicon die, and interconnected by a metallic overlay. We have seen in previous sections that the introduction of impurities into a pure silicon wafer markedly effects its resistivity. Utilizing this fact, it is also possible to introduce resistors into the transistor network, by means of the diffusion process.

Compared to a discrete device, the only new feature of an integrated circuit is in the fact that isolation of the various discrete devices, whether they be resistors or transistors, must be in some way effected. Examination of the diode characteristics shown in Figure 2 shows how this can be accomplished. Note that as long as the diode is biased in the reverse direction the effective resistance of the diode is very high. Thus if a diffused element, such as a resistor, is surrounded by silicon material of opposite doping type, and voltages are so arranged that the diode formed during the diffusion process is back biased, virtually all the current will flow in the resistor, (i.e., in the region isolated by the p-n junction). This is illustrated schematically in Figure 7. In the fabrication process a window is cut in the oxide which accurately defines the required length and width of the resistor to be diffused into the silicon. After the diffusion has taken place silicon dioxide is regrown over the surface of the resistor, and two windows are subsequently cut through to allow the making of an ohmic contact to either end of the resistor. The metallization is usually aluminum, and is deposited by vacuum deposition. Note that the metallization is insulated from the silicon surface by the silicon dioxide everywhere except where the contact windows were cut through.

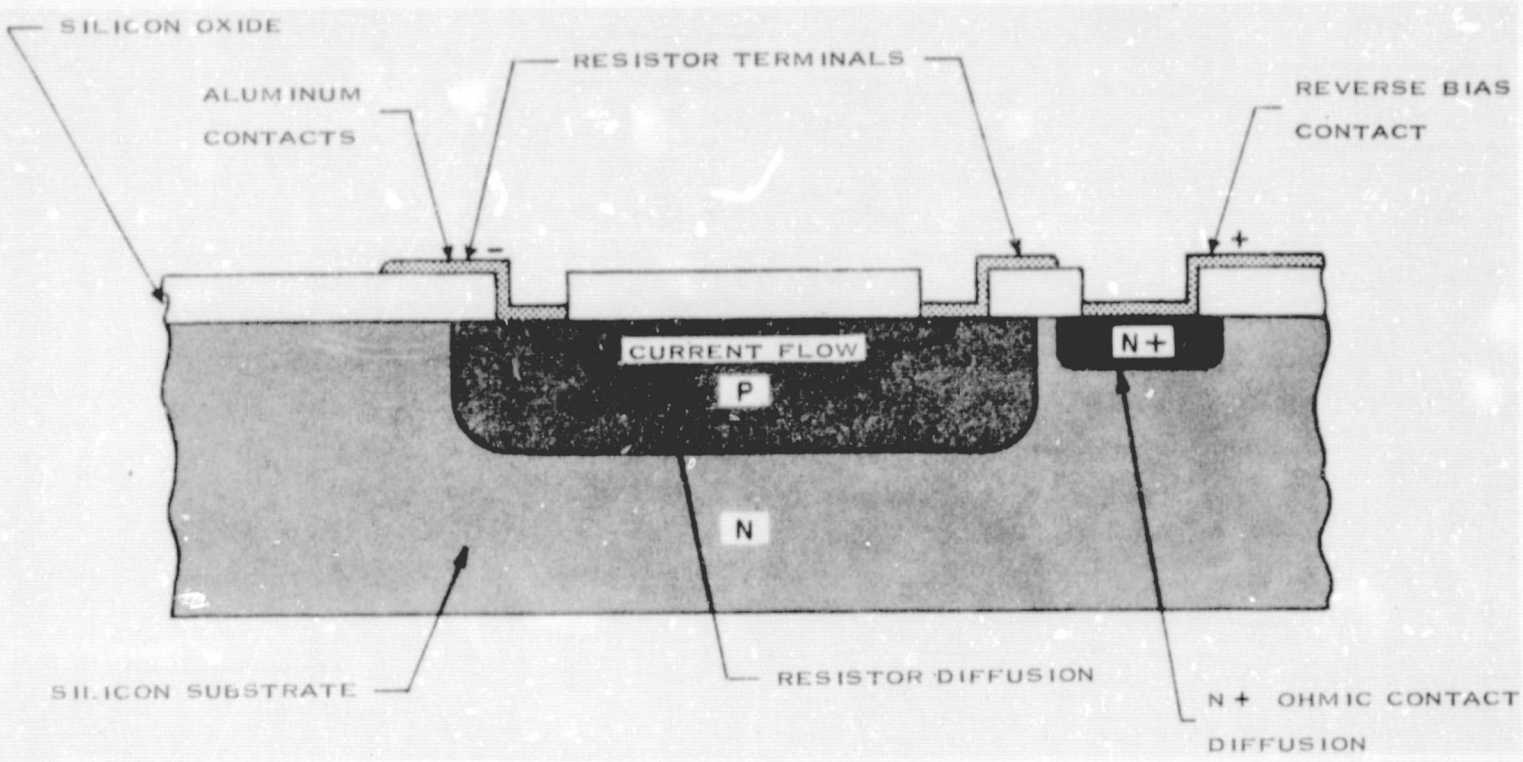


Figure 7 - Illustration of cross section of a diffused resistor.

For the resistor in question (which is a p-type resistor diffused into an n-type silicon) if the surrounding n-type silicon is made to be a positive potential with respect to the p-type silicon resistor, the resistor periphery (a p-n junction) is reversed biased. Hence practically all the current flowing into one contact of the resistor will flow out the other and thus the resistor is isolated from the rest of the surrounding silicon. Usually windows are cut through to the n-type material and metallization deposited so that electrical contact can be made to this area. Thus the proper voltage polarity can be applied to achieve the resistor isolation just described. Exactly the same scheme is used to isolate a transistor. Isolation may also be achieved by dielectric isolation methods which use dielectric materials such as glass as insulation, but to date this process is used extensively by only a few manufacturers.

3.3.2. Fabrication of a monolithic integrated circuit. - Figure 8 is a cross-sectional view of an integrated circuit which consists of one transistor and one resistor.

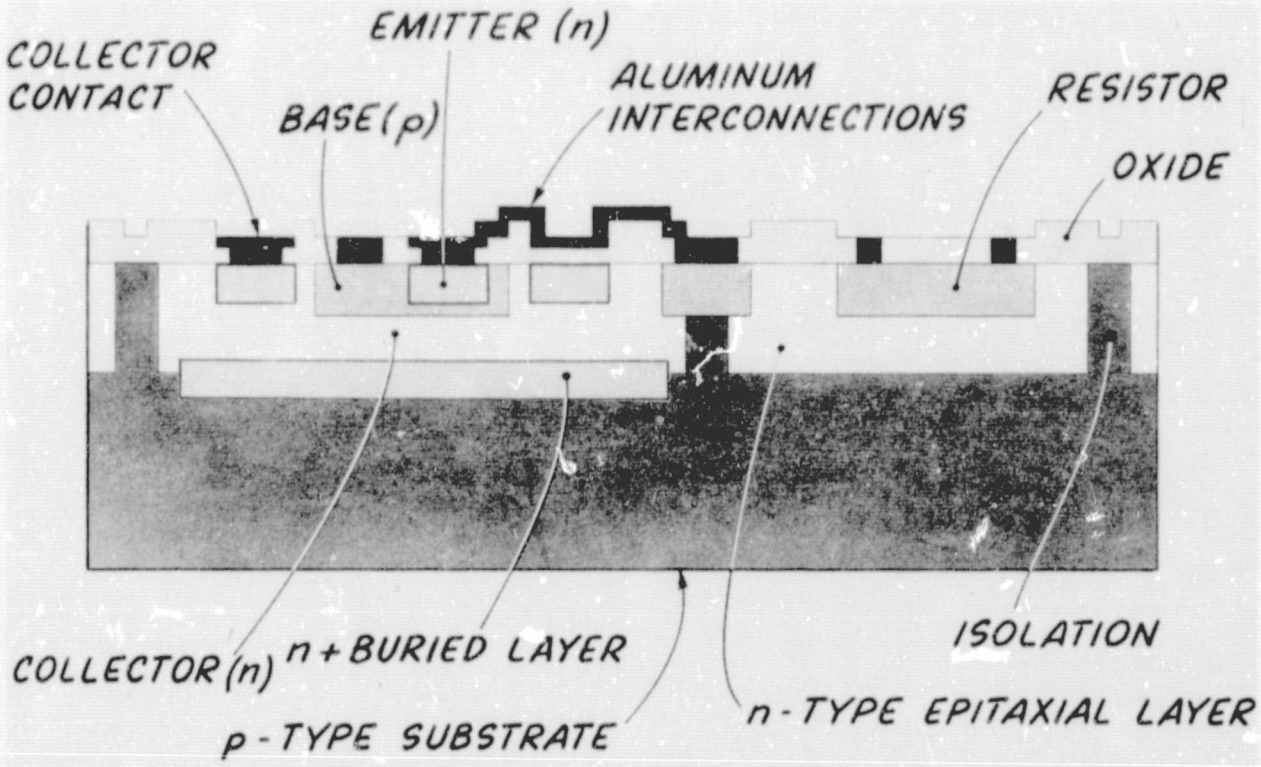


Figure 8 - Illustration of cross section of an integrated circuit.

Note that to fabricate such a structure use is made of the epitaxial deposition process described earlier. In addition, at least three separate diffusions are required. This basic integrated circuit is fabricated in the following manner. (Remember that while only <u>one</u> integrated circuit fabrication is discussed, many circuits are processed simultaneously on one silicon wafer.)

Step 1. A strong n-type layer (called n+) is positioned in such a way as to be under the collector region and is called the n+ "buried" layer. This layer may be diffused into the p-type substrate before growing the n-type epitaxial layer. Another method selectively grows the n+ regions using masked epitaxial techniques, and then continues on by growing the n-type collectors epitaxially. The outline of this "buried" layer is difficult to see on the surface of the completed device because there is no variation in oxide thickness and color. However a faint shadow may be observable at the outline. The buried layer is used to reduce the collector series resistance.

Step 2. An n-type epitaxial layer of 10-20 microns thick is deposited on a p-type silicon substrate. (The vertical gray areas labeled "isolation" are added later.)

Step 3. After suitable window cutting operations (i.e., photoresist application, exposing and developing the photoresist and etching windows through the silicon dioxide) the base and resistor elements are diffused in. These are p-type elements, hence require a p-type diffusion.

Step 4. Silicon dioxide is regrown again to cover the base and resistor windows, and a new window is opened for the emitter. Also at this time an n+ diffusion is made into some point of the collector. This is to assure that a non-rectifying ohmic contact results when the metallized contact to this region is attached.

Step 5. In order to create n-type islands in a p-type base (and thus isolate one element from another) a p+ diffusion is made around each element of the integrated circuit. These are the vertical rectangular gray areas mentioned in step 2 above. Usually the most negative part of the circuit is connected to the substrate, and in this circuit the emitter is actually connected to the substrate through the center p+ diffusion.

Step 6. Using techniques described earlier, the metal interconnecting pattern is laid down, completing the fabrication process. Die-to-header bonding and lead attachment follows, which, after hermetic sealing (capping), completes the package.

3.3.3. Physical appearance of a monolithic integrated circuit.- To clarify more thoroughly just what has been presented it would be instructive to examine a color photograph of an actual integrated circuit and to point out the various elements. Figure 9 is a color photograph of a simple dual nor gate. This circuit has been chosen because of its simplicity, yet it adequately demonstrates all of the characteristics contained in the previous discussion. The electrical schematic of this integrated circuit is shown in Figure 10. It is seen that this integrated circuit consists of six transistors paralleled in two sets of three transistors in each set. Each set of transistors has a common collector and a common emitter. Each transistor, however, has a separate base input and a series base resistor. Two collector resistors, each resistor serving three of the transistors, completes the circuit. Every element of this transistor-resistor network can be seen visually. Referring again to Figure 9 we see that the active circuit area is outlined by a violet line (1) approximately .0005" wide. This is the main isolation diffusion. Other isolation diffusions surround each set of three transistors (2) . The collector-resistor has a common point at the B+ connection (3) and proceeds to the right and the left around the rectangular area to the collector metallization. Each separate base resistor can be seen as other violet lines located between the outer metallization area and the metallization area directly connecting the base area (4) . The metal interconnects to the two common collectors (5) and the common emitter (6) can be clearly seen. Note that the metallization lies on top of the silicon dioxide area except where contact cuts to the silicon surface are made, and the silicon dioxide covering directly

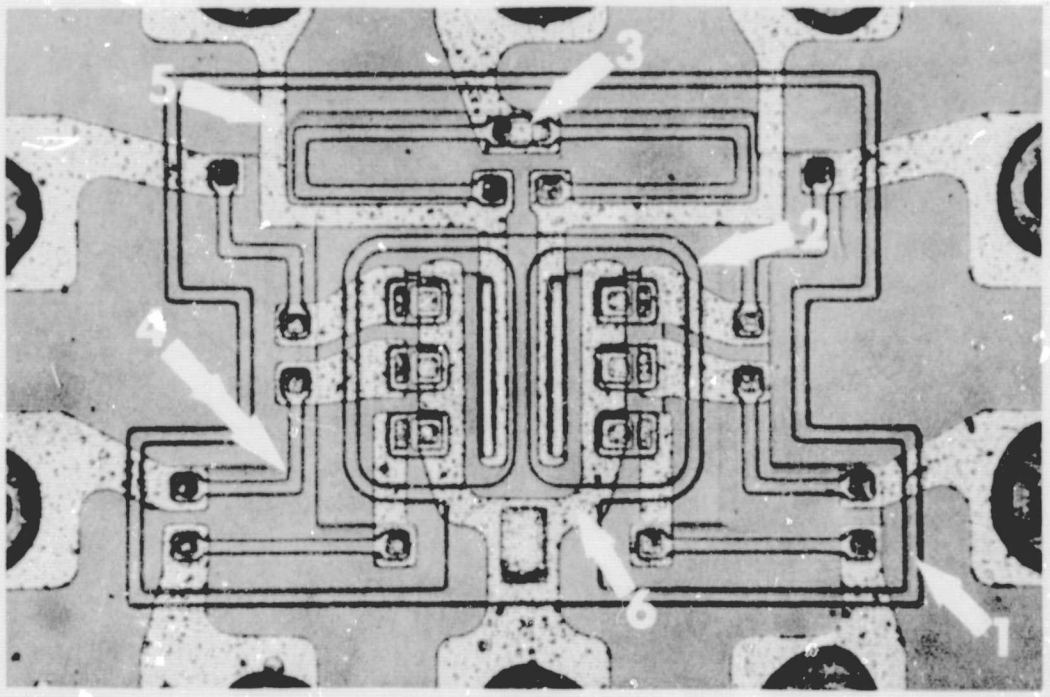


Figure 9 - A microphotograph of a dual nor gate integrated circuit, bright field illumination.

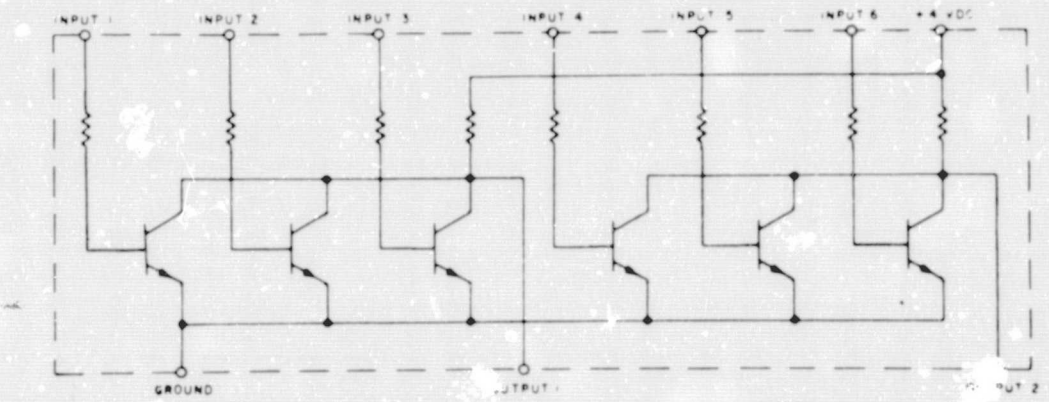


Figure 10 - Schematic diagram of the above dual nor gate integrated circuit.

under the metallization provides the necessary electrical insulation. For lead attachment purposes the metallization is widened out at the die edge, permitting the internal leads to be attached. (These areas are commonly referred to as bonding pads.) Thus the metallization in this case is a link between the heavier lead-in wire and the active circuit area.

3.3.4. Die attach and lead bonding. - The integrated circuit die, not too much larger than the transistor die, is attached to the package by methods discussed previously. After this has been accomplished it is necessary to make electrical connections between the bonding pads of the circuit and the package leads. Various bonding methods are in use today, but the most common types are thermocompression, ultrasonic and resistance weld.

3.3.4.1. Thermocompression bonding. - Thermocompression bonding (also called ball bonding, stitch bonding and wedge bonding) is effected by the use of both force and temperature through a bonding tool that is forced against the wire which is to be bonded to the circuit. Although both ball and wedge bonds are thermocompression type bonds, they utilize different tools and will be discussed separately.

3.3.4.1.1. Wedge bonding. - The wedge bond is illustrated in Figure 11. The attachment of the wire to the metallized area is accomplished by the use of a wedged tool and employment of both force and temperature. Initially both the header and the mounted integrated circuit die are heated to an elevated temperature. Next the wire to be bonded is positioned over the desired contact area. The bonding wedge is then brought down upon the wire with a controlled amount of force, which brings the wire in intimate contact with the metallized area. Since the header is kept at an elevated temperature, an intimate union of the two metal surfaces takes place, with the end result that the wire is permanently bonded to the metallization. A similar bond is then made to the lead-in post, which

completes the bonding operation. Both aluminum and gold wires are used.

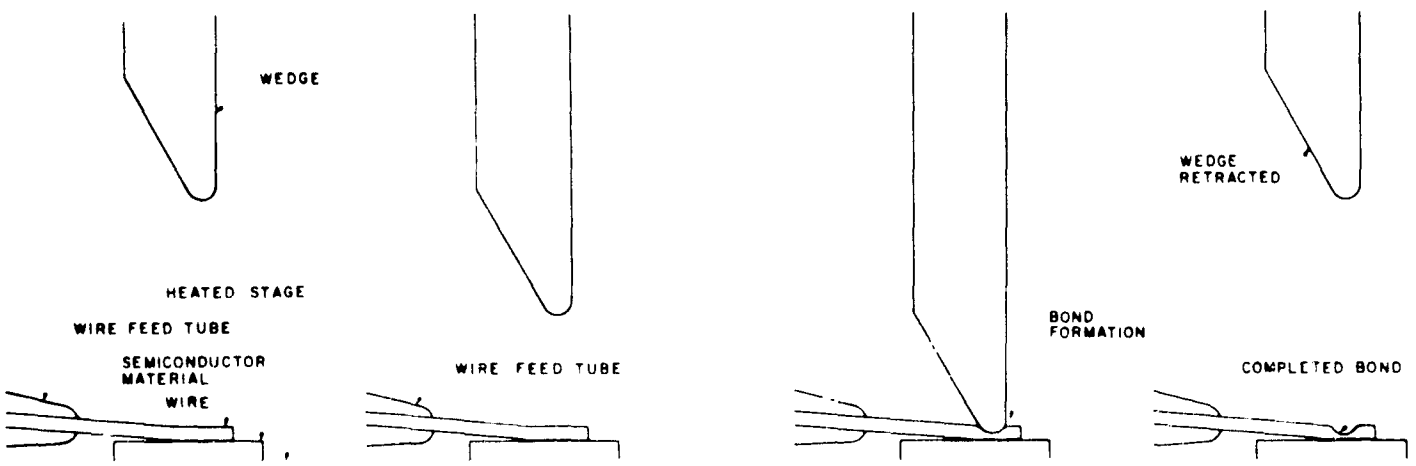


Figure 11.
Wedge bonding technique.

3.3.4.1.2. Ball bonding. - Ball bonding is also used for bonding small gold wires to thin film metal contacts. Bonding of this type is effected by threading the wire to be bonded through a quartz or tungsten carbide thick walled capillary. The capillary has a tapered end of only a few mils diameter on the outside. This capillary is mounted in a mechanical fixture so that it can be moved vertically. Initially a small amount of wire is forced through the end of the capillary. The wire is then cut by a hydrogen flame which results in the wire having a ball formed on one end. The wire, fed from a roll, is free to move in the capillary. Bonding is effected by lowering the capillary annulus onto the ball and applying pressure, while at the same time applying heat. Thus since temperature and pressure have been applied during the bonding operation, the process has been given the name thermocompression. The process is illustrated in Figure 12. The bond to the post or package land is not a thermocompression ball bond, but is made

with the same tool, the end of which is used as a wedge bonder.

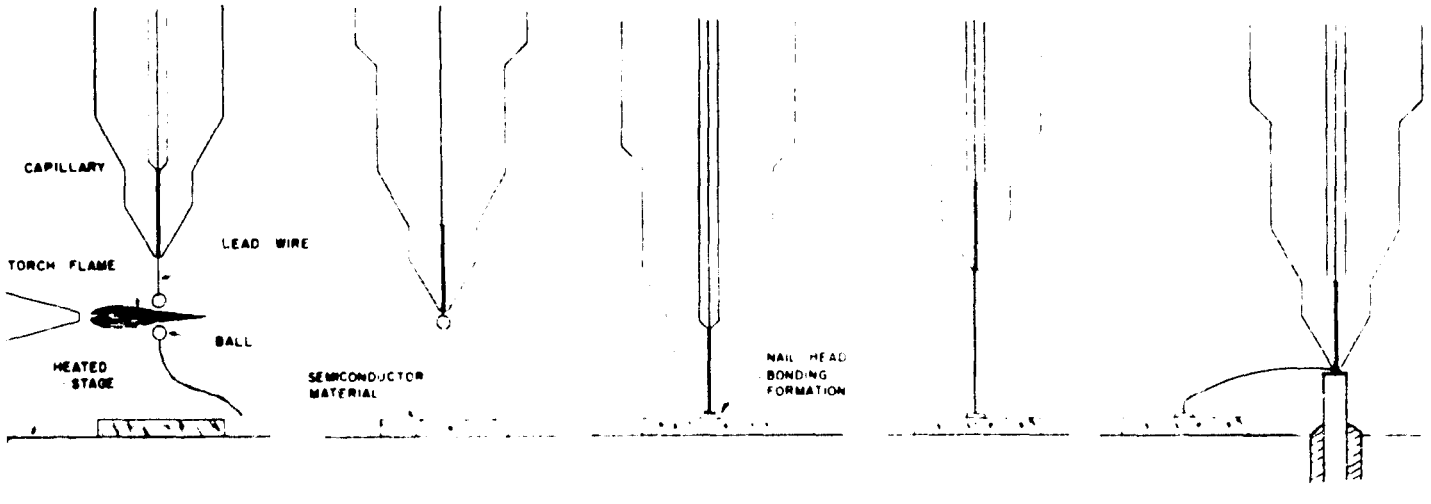


Figure 12.

Ball bonding technique.

This method of bonding has been very popular in the past, and probably will continue to be popular. However usually the ball diameter is from 3 to 5 mils, and thus the technique is difficult to apply when used in conjunction with very small bonding pads. Also intermetallic compounds of gold and aluminum can form, which could lead to the bond becoming brittle and fragile.

3.3.4.2. Ultrasonic bonding. - Ultrasonic bonding has received much interest currently, because it affords the possibility of bonding aluminum wires directly to aluminum metallization areas, thus eliminating any intermetallic compounds which may form when dissimilar metals are bonded together. The technique is roughly similar to the wedge bond discussed earlier except that the bonding tool is made to vibrate at an ultrasonic frequency (usually around 20 KC) and affords a scrubbing action which effectively breaks through the aluminum oxide (which surrounds the wire and covers the metallization pad) thus bringing about

a clean metal-to-metal interface. The ultrasonic bond has the advantage that it can be used to bond fine aluminum wires to very small metallized areas. Although the bond strength of an ultrasonic bond is typically 3 to 5 times weaker than a gold thermocompression bond (ultrasonic aluminum bond strengths are nominally about 2.5 grams) the weight of aluminum is only 1/7 that of gold. Thus it will take 7 times the acceleration to break an aluminum bond over the gold bond of equal strength. Ultrasonic bonds are relatively new to the industry, as of this writing, and you will note that in the Internal Visual Workmanship Standard much attention has been given to the physical appearance of the bond itself. Fortunately most of the bonding machines are almost entirely automatic and contribute much to the elimination of many of the variables associated with the bonding process.

3.3.4.3. Resistance welding. - Resistance welding is employed when the wire and the element to be bonded are accessible to welding tweezers. It is usually used when a wire is to be bonded to a lead-in post. The bonding wire is squeezed against the post by the tweezer tips, so that a large current can be passed through both the wire and the post simultaneously. The large current melts both the wire and a section of the post material, resulting in a welded structure.

This type of welding is less frequently used in the fabrication of integrated circuits.

4. INTERPRETING THE INTERNAL VISUAL WORKMANSHIP STANDARD FOR MICROELECTRONIC DEVICES - NASA STD XX-2

4.1. Introduction. - The preceding sections outlined the basic physical principles and the modern manufacturing techniques used in the fabrication of the silicon monolithic integrated circuit. In this section a study of the NASA STD XX-2 will be made, and much of the material previously presented will be utilized.

4.2. Instruments used in the assessment of workmanship quality.

4.2.1. The microscope. - Microscopes are used to enlarge the object viewed, so that clearer details in regard to the object's appearance can be observed. The optical principles of all light microscopes are the same, and though relatively simple, will not be discussed here. The important differences are in the amount of magnification obtained. The practical upper limit to microscopic magnification is in the region of 1500X. Macroscopic viewing usually is the term applied to magnifications of 10X to 80X. Microscopic viewing usually makes use of optical systems giving magnifications of 100X or more. The microscope may be monocular (eyepiece for viewing) or binocular (two eyepieces). Binocular microscopes provide separate light paths from the viewed object to each eye, and thus a stereoscopic (3rd dimensional) effect can result. The viewed object is lighted either by a simple side lighting arrangement or by light normal (perpendicular) to the specimen. Sometimes a microscope may combine both lighting arrangements. Macro type microscopes almost always employ side lighting.

The term "Metallurgical Microscope" is reserved for an instrument that has a high magnification capability, and that employs normal bright field illumination. It is used extensively by metallurgists to study metal specimens. Due to its direct applicability to semiconductor viewing problems, it is now utilized in many production stations in the semiconductor manufacturing line.

Figure 13 is a diagram showing the elements of a Stereoscopic Microscope. The object to be viewed is illuminated from the side by a lamp (which usually has light focusing capabilities). The light passes through two separate and optically identical paths to each eye. Both the objective lenses and the eyepiece lenses are usually capable of being changed. The total magnification is the product of the individual lens magnifications. Thus an 8X objective and a 10X eyepiece will result in the object being magnified 80 times.

Focusing is accomplished by raising or lowering the whole lens system until the object appears in sharp detail. This type of microscope is most useful in viewing gross imperfections. NASA STD XX-2 specifies that such a microscope and lighting arrangement be utilized in such viewing tasks as "Die Mounting and Orientation" (section 6.1) "Wire Bonding" (section 6.7) "Wire Routing" (section 6.8) "Contamination" (section 6.9) and "Package Condition" (section 6.10).

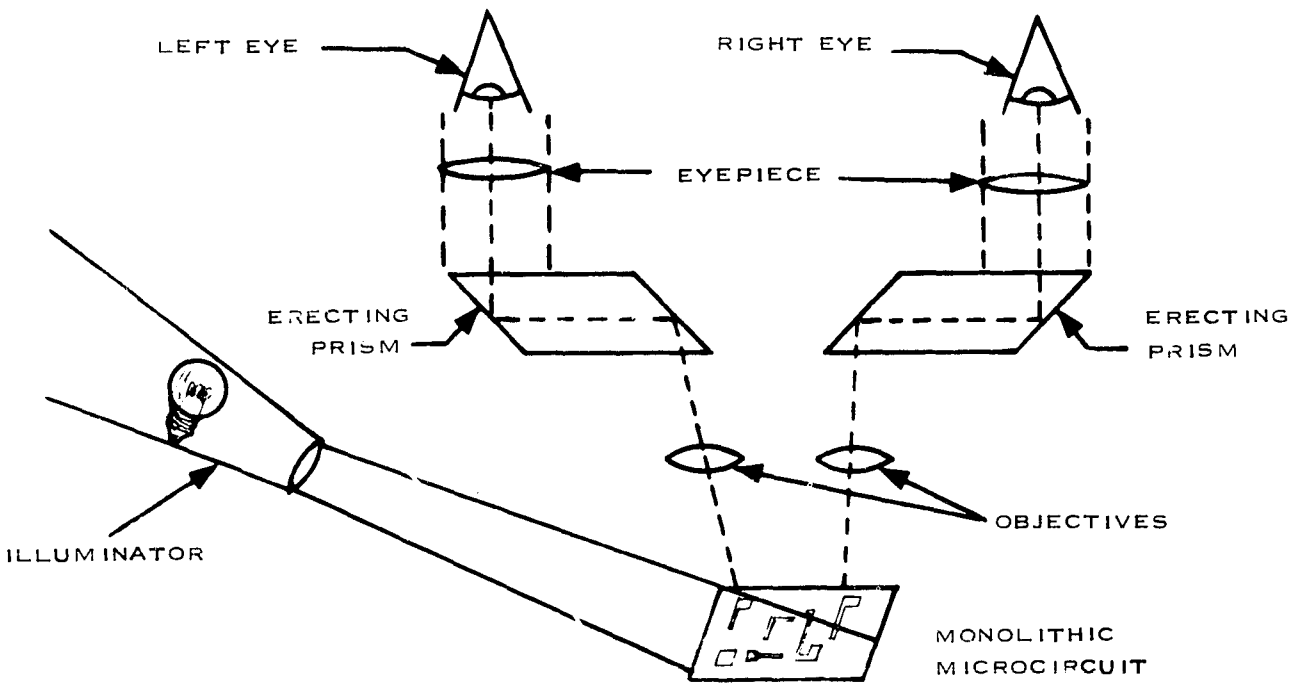


Figure 13. Elements of a stereoscopic microscope.

Figure 14 is a diagram showing the elements of a Metallurgical Microscope. This microscope differs from the low power microscope previously described mainly in three areas. First, the lighting arrangement, by means of a half silvered mirror, illuminates the object to be viewed normally, yet allows enough light through to the eyepiece from the object so that the object may be seen. Secondly, the lens quality and magnification capability is far superior to that of the lower power microscope. Due to the normal lighting arrangement, it is this type of microscope that must be used to obtain the sharply defined colors seen on an integrated circuit, and hence

is used for viewing anomalies that exist as a result of oxide thickness variations, or stains and residues left on the chip surface. Third, the metallurgical microscope has a far smaller depth of field than the lower power microscope. This term is used to denote the range of in-focus objects at any one setting of the microscope. Thus in viewing an integrated circuit with the low power microscope, both the die surface and the die mounting surface on the package may be in focus. Thus the depth of field of the lens system is approximately .010". The depth of field of the metallurgical microscope may be only a fraction of this. Depth of field is inversely proportional to the magnification, i.e., the higher the magnification the smaller the depth of field.

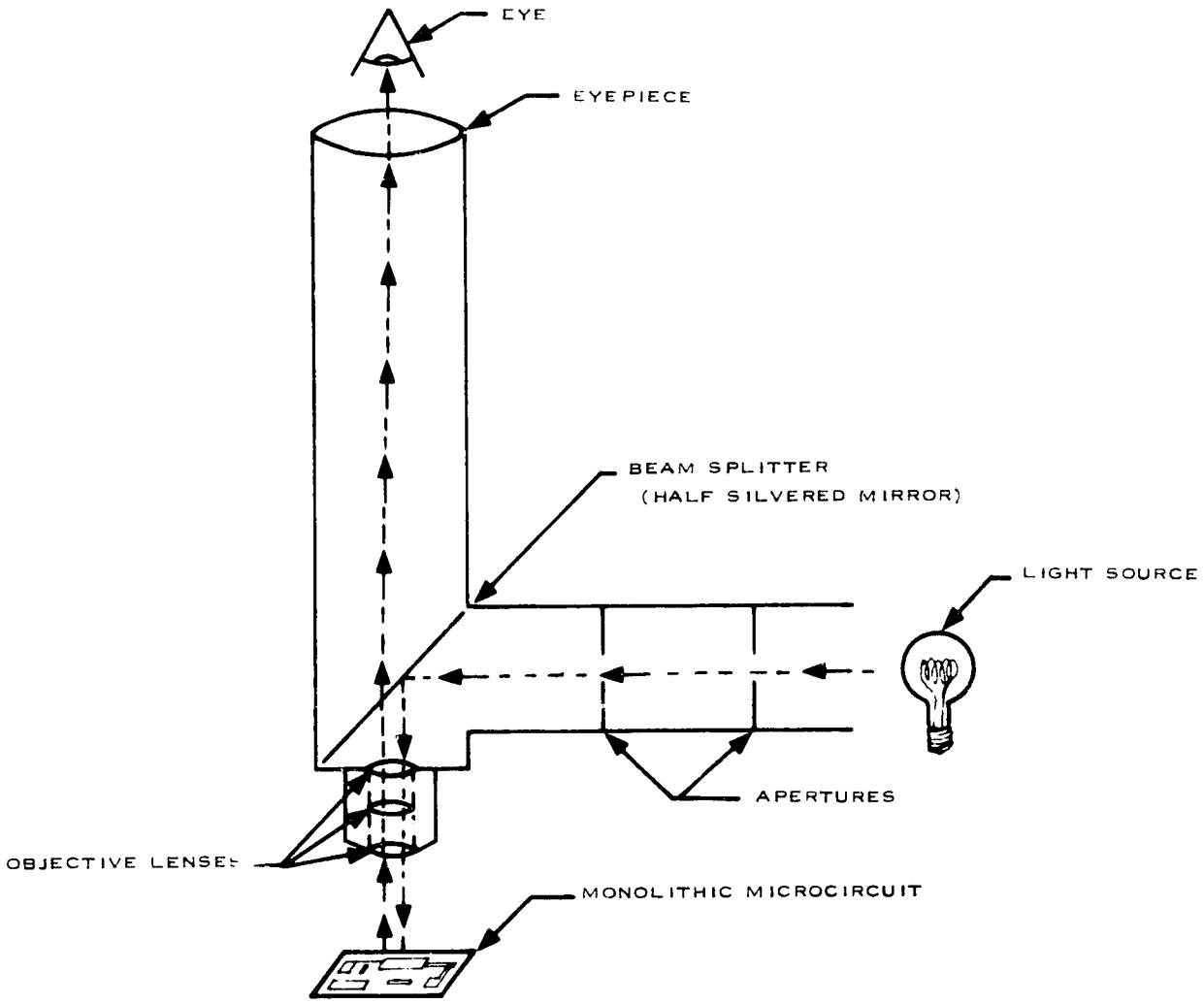


Figure 14. Elements of a metallurgical microscope.

The metallurgical microscope also has a much smaller lens to object distance, thus making it difficult and at times impossible to focus on an integrated circuit die located in a package. The walls of the package surrounding the die may be too high, or the lead-in posts may be too long to allow the lens to be brought into focus on the die surface, if a high magnification is required. In this case the only alternative is to use a lower magnification. NASA STD XX-2 specifies that this type of lighting and degree of magnification be used in "Scribing" (section 6.2) "Alignment" (section 6.3) "Cracks" (section 6.4) and in other sections. Figures 15 and 16 are additional views of a low power and metallurgical microscope respectively.

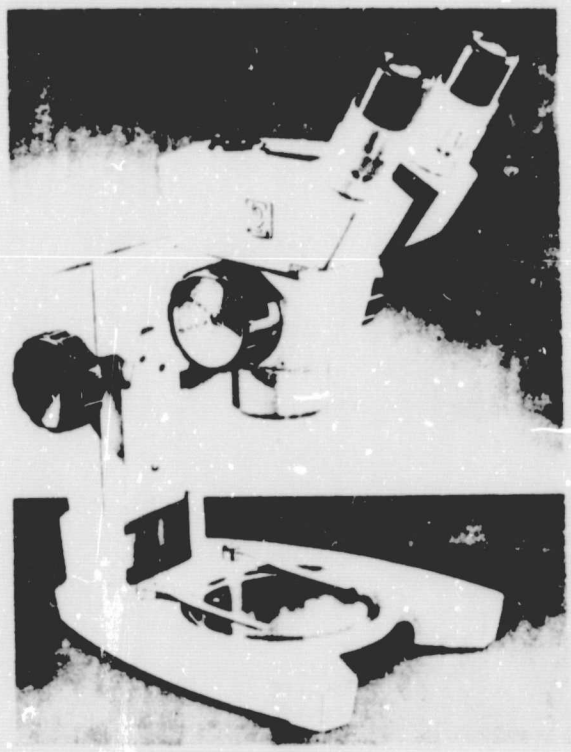


Figure 15.
Low power microscope.

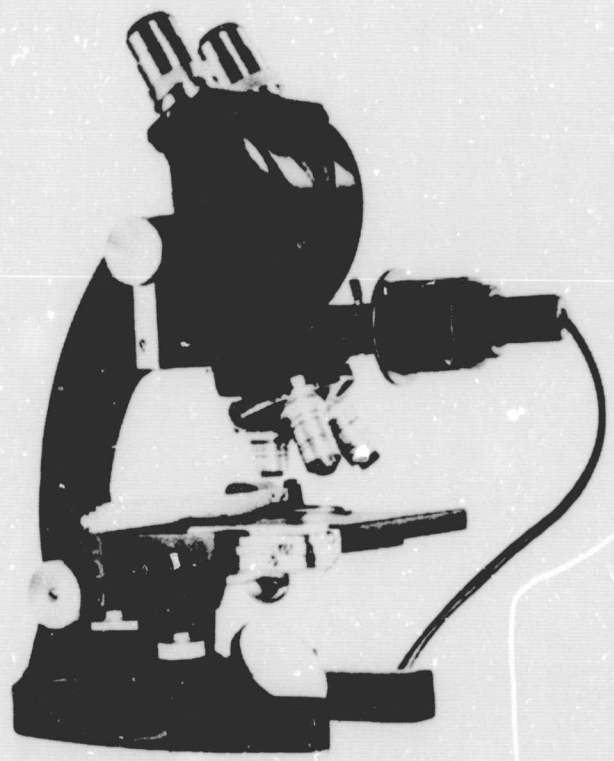


Figure 16.
Metallurgical microscope.

With practice, viewing objects through a microscope becomes a routine procedure for the inspector. Important factors which affect the viewing capability are lighting and the position of the specimen on the microscope stage. Even though the NASA STD XX-2 specifies the minimum magnification to be used in each case, many times additional information can be gained by changing to another magnification or even to another microscope. Thus in section 6.10 "Package Condition", magnification of 30X is specified as a minimum. However judging whether or not a package contains a possible leak through a pinhole may require a magnification of 150X. Each case is unique, and requires separate consideration.

It should be noted that in order to see in full brilliance the colors given off by an integrated circuit under a microscope the surface of the die must be perpendicular to the light beam. Normally, the surface of the die is parallel to the bottom surface of the package, requiring no leveling.

Due to the normal lighting arrangements found on most metallurgical microscopes, sometimes true color is not observed on raised objects such as intermetallic growths. In these cases employment of side lighting might be helpful. Some metallurgical microscopes have an attachment that affords this capability. Otherwise a focused lamp light source may be used. The lamp should be brought as close to the object to be viewed as possible. However due to the short lens-to-object distance at high magnifications, such a technique may be impossible to employ.

Before using any binocular microscope the operator should adjust the eyepieces for interpupillary distance and for focus. In the former case the adjustment is easily accomplished by moving the eyepiece body tube laterally, either in or out, until the operator sees one image. In the latter case, one of the eyepieces is always individually adjustable. Thus proper focus of the fixed eyepiece is first obtained, with the other

eyepiece covered, by adjustment of the microscope focusing knob. Next the other eyepiece is focused by turning the adjusting ring located around the eyepiece.

4.2.2. The filar eyepiece. - The capability of providing an absolute measurement of closer than .00025" is often required. For instance, section 6.5.6.4 of the Workmanship Standard reads "cause two diffused areas to be within 1/4 mil (.00025 inch) of shorting together." The filar eyepiece, also called a filar micrometer, is used in making such lateral measurements (Figure 17).

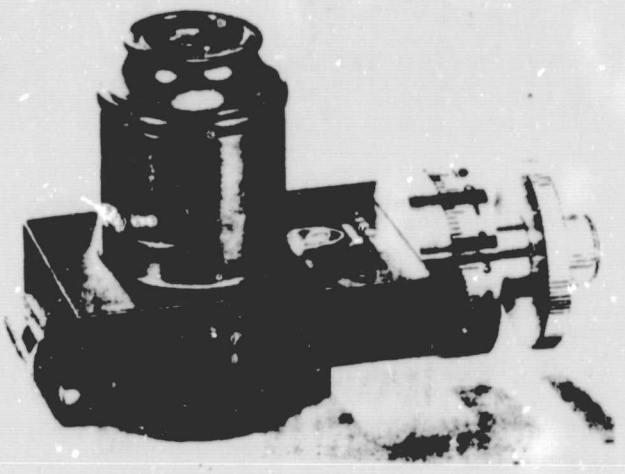


Figure 17. Filar Eyepiece.

This microscope accessory is used in place of one of the eyepieces of the microscope. (Viewing is thus monocular when this eyepiece is used.) Looking through the eyepiece, a series of fine equally spaced lines is observed, in addition to a movable cross hair. When the microscope is properly focused on the object to be viewed the equally spaced lines and the movable cross hair of the eyepiece will be superimposed on the image. The cross hair is moved by rotating the wheel seen on the right side of the eyepiece. The eyepiece must first be calibrated,

since the calibration is dependent on the objective lens used. Calibration is accomplished by first focusing the microscope on a scale of known dimensions (known as a "stage micrometer") and moving the cross hair from one line on the scale to another. The known interval is then compared with "before and after" readings on the filar eyepiece wheel, which has evenly spaced and numbered markings. Once calibrated, the eyepiece can be used for direct measurement.

4.2.3. Depth measurements. - Most metallurgical microscopes have a depth measuring capability built into the instrument. Usually the fine vertical height adjustment knob serves this purpose. This knob has a calibrated scale engraved on the perimeter which permits quantitative measurements to be made. Usually the limit of resolution of such indicators is about 0.4 mils. Oxide thicknesses are in the neighborhood of 10,000Å (.04 mils) or less, and are outside the measurement capabilities of the vertical height adjustment scale. Interferometric methods must be used to measure thicknesses of this magnitude, and are not usually adaptable to routine depth measurements.

4.2.4. Tweezers, probes, and other aids. - Handling a packaged integrated circuit poses no special problems for the inspector. Most of the inspections specified in NASA STD XX-2 can be carried out with the device to be inspected lying flat in a suitable container. When it is necessary to handle the device, the device should be handled using the exterior package walls as hold points, rather than the package leads. Most package leads are in the order of .004" thick and are easily damaged. Tweezers are usually of the all-metal type, but bone tweezers are sometimes used to reduce the possibility of scratching the device.

4.3. Defect interpretation.

4.3.1. Defect occurrence. - A table showing the incidence of the cosmetic type defects noted in the NASA STD XX-2 is given below. This table was derived using actual case histories from failure

analysis data cards.

The parts listed in this table all passed several electrical screening and burn-in tests before being assembled into equipment. However after assembly these units failed due to defects which could have been eliminated by visual inspection. This points out the importance of visual inspection to eliminate potential failures not detectable by x-ray, centrifuge, electrical and other type tests.

FAILURES IN PRODUCTION AND FIELD, TRANSISTORS AND IC'S

	DEFECT CLASS	NO.	%
6.7	Wire bonding	316	48
6.9	Contamination	72	11
6.4	Cracks	55	8
6.6	Metallization	54	8
6.1	Die mounting	52	8
6.8	Wire routing	38	6
6.5	Oxide defects	35	6
6.10	Package defects	26	4
6.3	Alignment	5	1
6.2	Scribing	0	0
		653	100

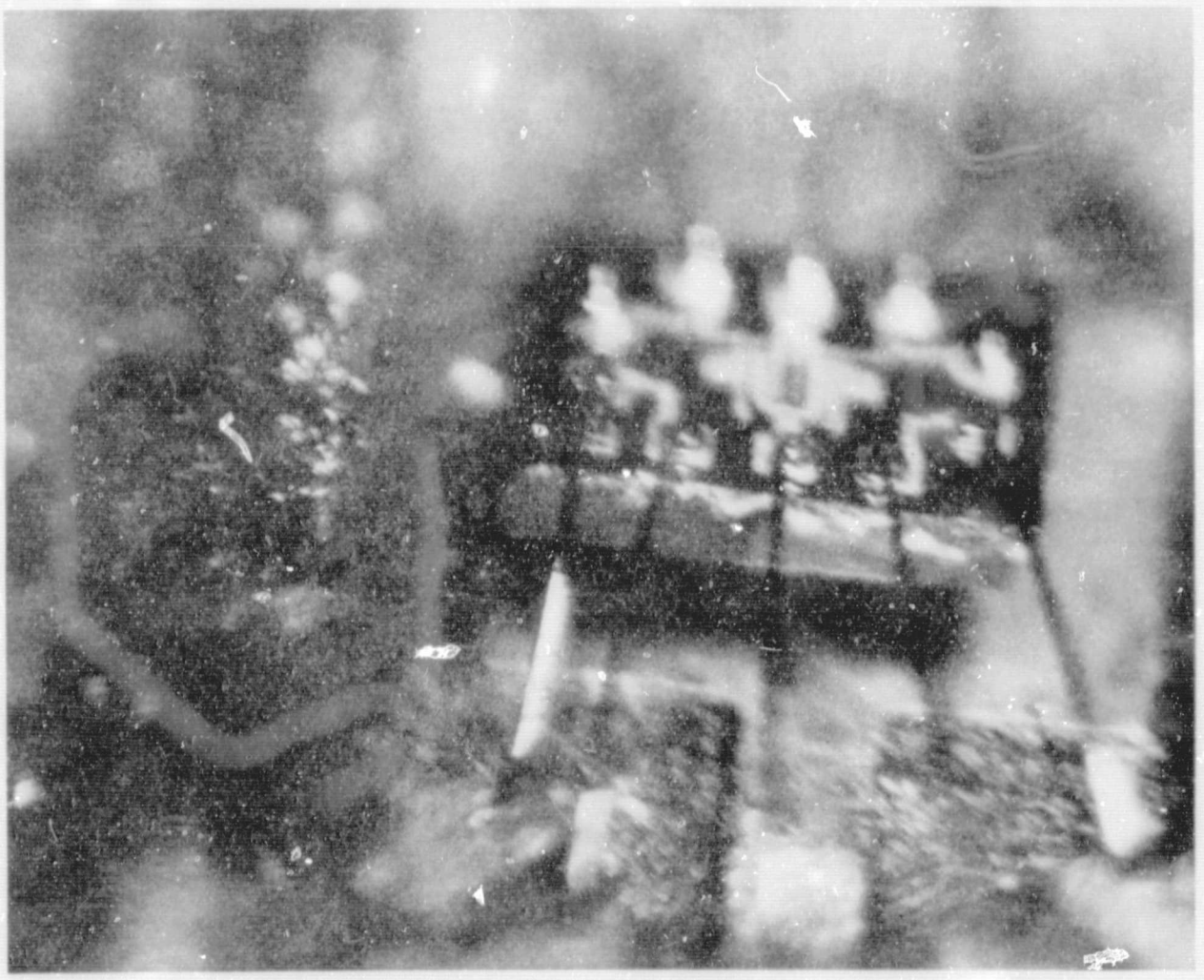
It should be noted that almost half of the failures fall in the category of wire bonding. The section on wire bonding in the NASA STD XX-2 (section 6.7) is the largest section of the Standard. This one operation, the bonding of lead wires, is the most difficult operation to get under control by the parts manufacturer.

Contamination (11%) ranks second in cause of failure, according to this tabulation, but is quite close to the next group of failures (8%), and probably in other studies may be slightly changed in rank. However the severity of this reliability hazard should not be minimized, especially in equipment used in Missile and Space applications. Loose conducting particles of length greater than the shortest

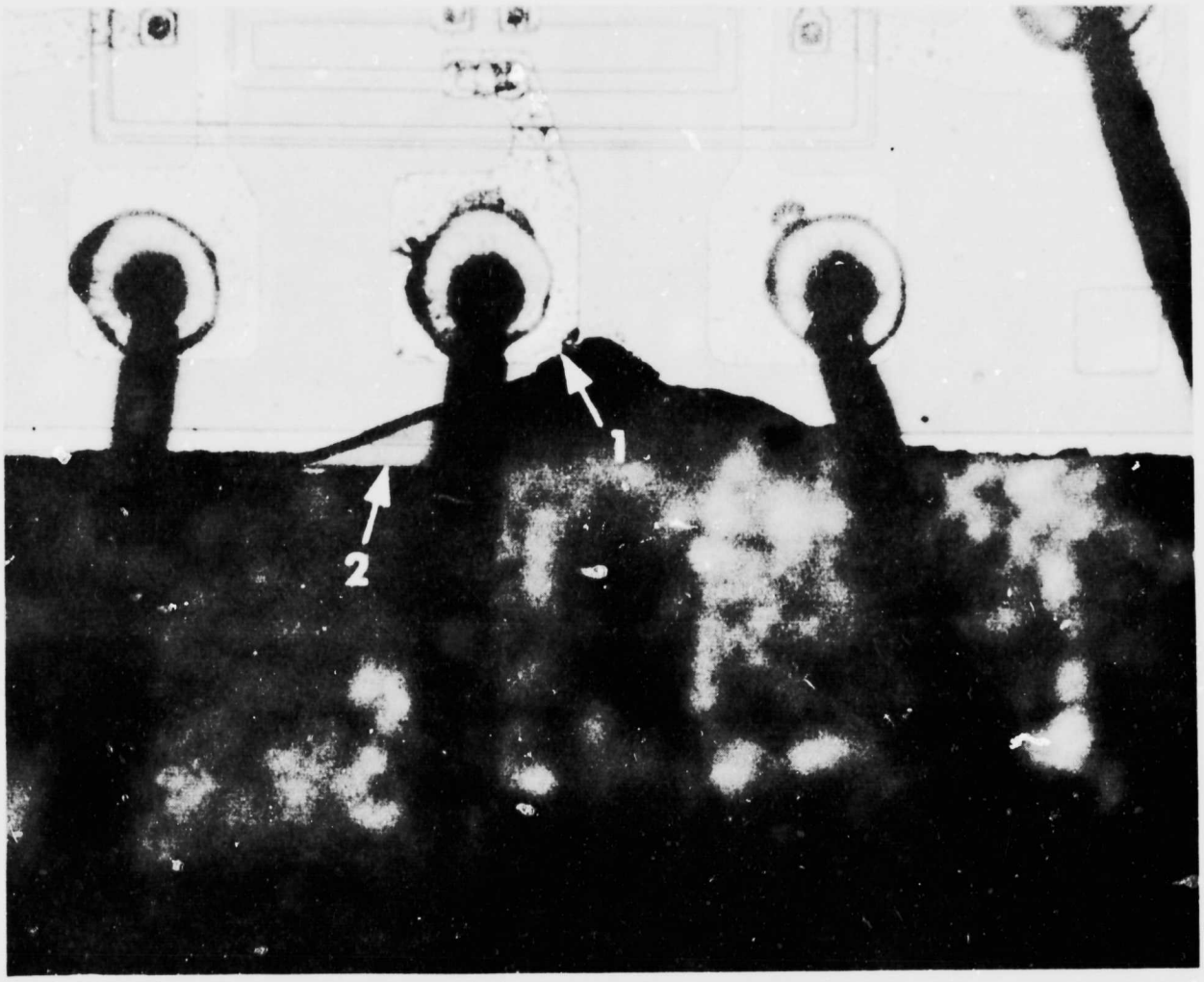
distance between two metallization paths provide a severe shorting hazard. There is also the possibility that small particles may hook up in chains through electrostatic attraction to form larger conducting paths. That is why in the NASA STD XX-2 loose particles of any size or number are cause for rejection of the part. Determining whether or not a particle is fixed or loose in the package is very difficult and is still the subject of investigation. A good rule to apply here is "When in doubt, throw it out".

The application of a strict Internal Visual Standard in the pre-cap stage of processing is mandatory, if parts are to perform according to the demands requested of them by our current Space Programs. As a further aid to the government inspector, the following section will discuss, in some detail, the application of the NASA STD XX-2, to a group of devices picked to give examples of some of the important reject categories.

4.4. Examination of devices. - The following photographs, many taken from actual device failures, are given to illustrate the types of rejectable defects in integrated circuits that the government inspector will observe in his inspection work. Each defect is discussed as to the type of defect, the probable cause of the defect, and the reliability hazard that the defect presents. In some cases line drawings have been included where they serve to provide a clearer understanding of the defect in question.



Defect: Die not firmly attached to package floor (para. 6.1.2).
Cause: Improper die attach operation. Specifically this could have been due to insufficient temperature, pressure, or motion at that operation; or it could have been due to insufficient plating on package floor.
Hazard: Die may change orientation causing shorts or putting added stress on bonds. In addition lack of adequate heat sinking of die may affect operation.



Defect #1: Chip-out having less than 1 mil (.001 inch) between the edge of the chip-out and the nearest diffused junction or metallization (para. 6.2.3). In this case, the chip-out actually touches the metallization.

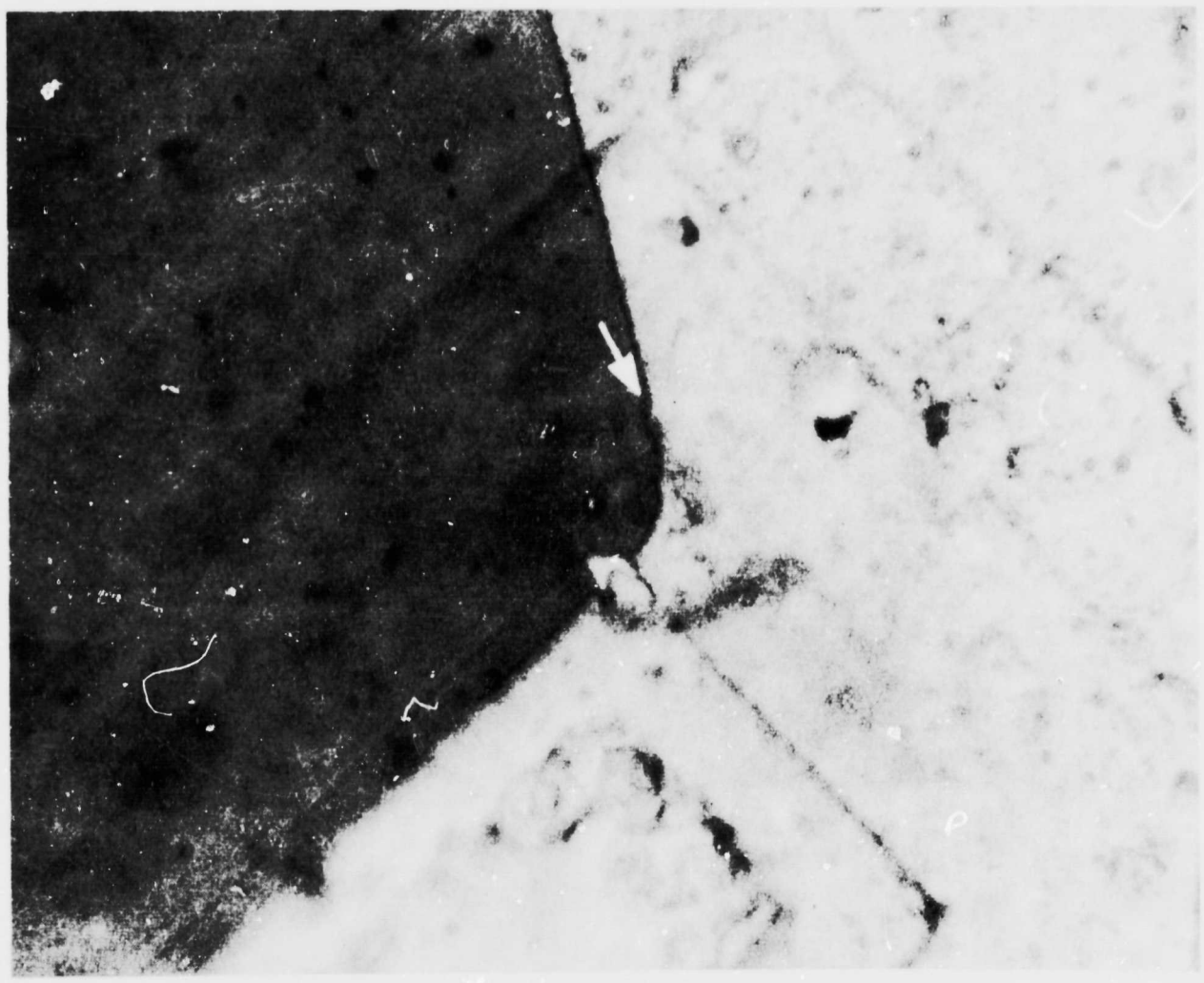
Cause: Imperfect scribing and dicing or mishandling of die after dicing.

Hazard: The metallization can short to the underlying semiconductor at the edge.

Defect #2: Severed chip of silicon still in place (para. 6.2.4).

Cause: Imperfect scribing and dicing or mishandling of die after dicing.

Hazard: The chip can break free and cause mechanical damage and/or cause a short.



Defect: Oxide flaw extending under metallization (para. 6.5.5.1).
Cause: Probably due to flaw in photoresist mask used during an oxide etching step. This mask flaw resulted in the etching of oxide in an undesired area.
Hazard: The defect could result in shorting of the metallization to the underlying semiconductor at a point where little or no oxide is present.

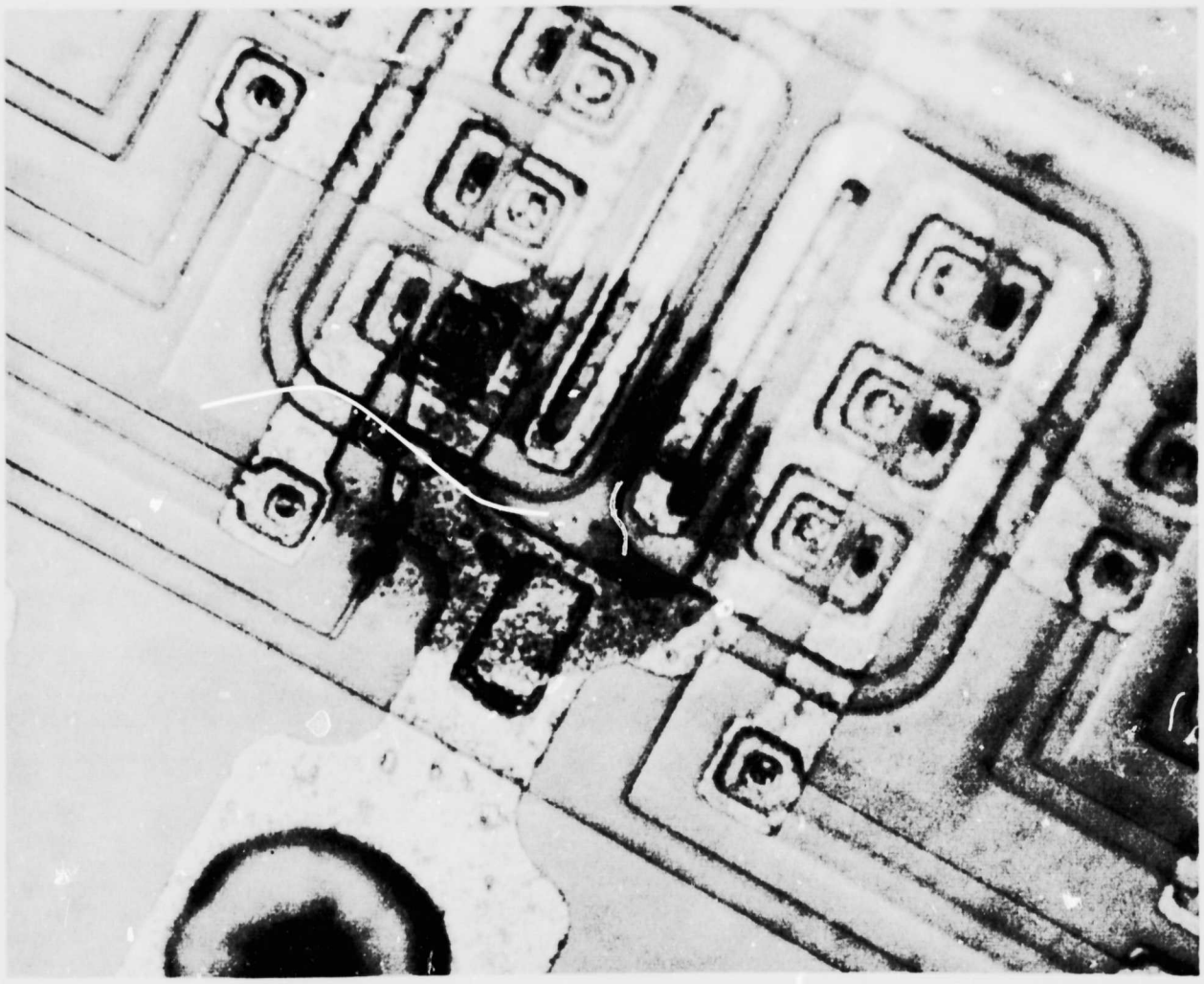
Defect #1: Oxide defect under metallization (para. 6.5.5.1).
Cause: Probably due to flaw in photoresist mask used during an oxide etching step. This mask flaw resulted in the etching of oxide in an undesired area.
Hazard: The defect could result in shorting of the metallization to the underlying semiconductor at a point where little or no oxide is present.
Defect #2: Less than 75% of contact window covered (para. 6.3.2).
Cause: This defect is due to misalignment of metallization.
Hazard: Decreased contact area between metal and semiconductor increases chances of an open circuit developing.
Defect #3: Metallization not intended to cover a contact window approaches too close to that window (para. 6.3.4).
Cause: This defect is due to misalignment of metallization.
Hazard: The defect increases the probability of shorting.

Defect: Scratches (regardless of depth) cutting completely across a path (para. 6.6.6).
Cause: Careless handling. These could have been caused by tweezers used in handling either an entire slice or individual dice; or they could have been caused by two dice rubbing together after scribing and dicing; or any of many other causes.
Hazard: The metallization is reduced in thickness in the region of the scratch. This leads to the possibility of an open circuit occurring.

Defect: Discoloration of metallization and oxide (para. 6.6.10 and 6.9.5).

Cause: Incomplete cleaning off of processing chemicals at some stage.

Hazard: The presence of the discoloration indicates that chemicals may be present and that further chemical attack is possible. This could result in subsequent deterioration of oxide and/or metallization.

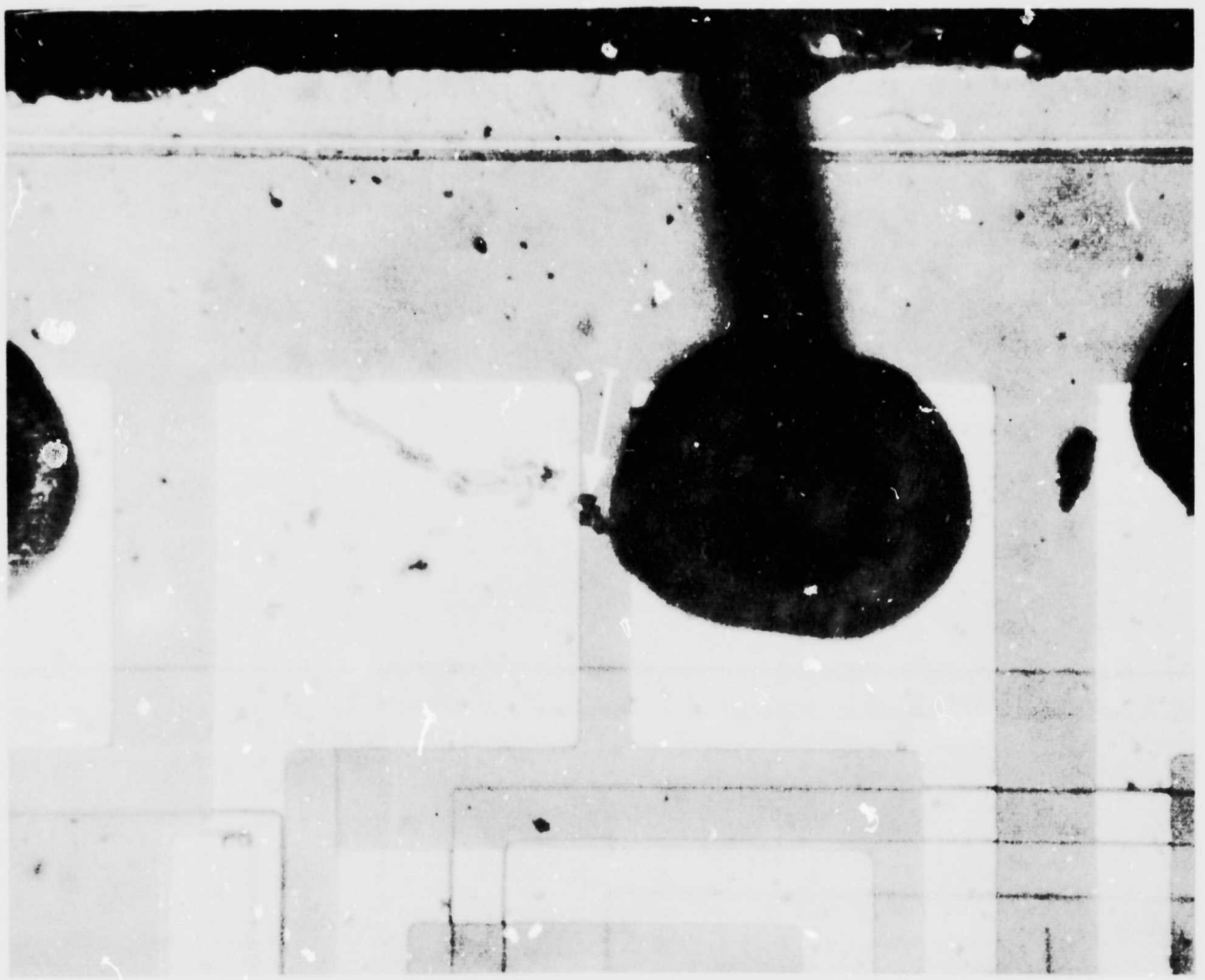


Defect: Discoloration of metallization and oxide (para. 6.6.10 and 6.9.5).

Cause: Incomplete cleaning off of processing chemicals at some stage.

Hazard: The presence of the discoloration indicates that chemicals may be present and that further chemical attack is possible. This could result in subsequent deterioration of oxide and/or metallization.

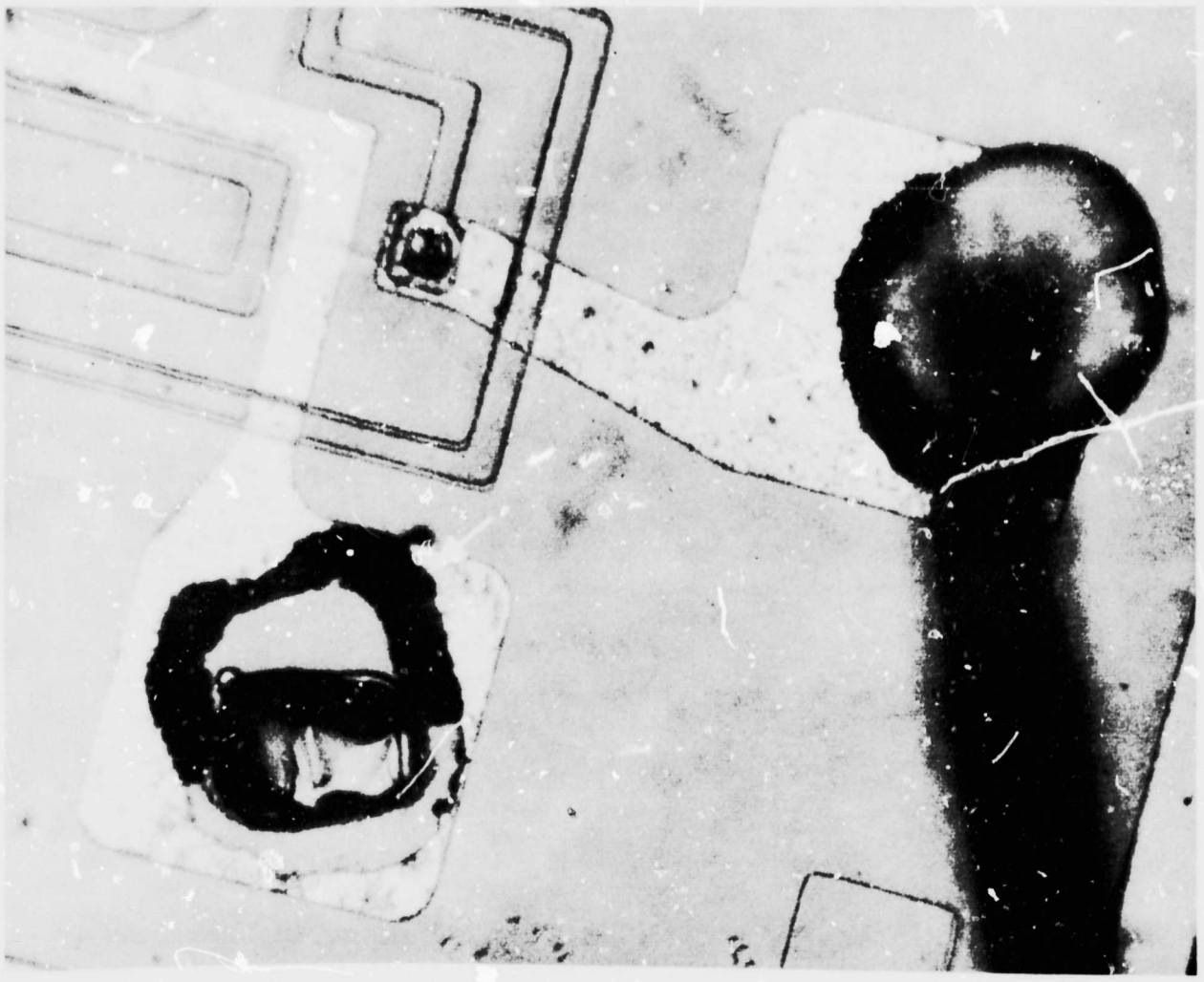
Defect: Voids in metallization leaving less than 75% of the minimum designed width at oxide step (para. 6.6.5.2).
Cause: In this case the cause was probably that the photoresist mask did not adhere sufficiently at the oxide step, thus allowing the metallization etch to penetrate under the mask and remove the metallization. The defect may also be caused or aggravated by thin metallization at the step due to "shadowing" effects during deposition of metallization.
Hazard: Restriction of current path width increases chances of an open circuit. (In this particular case an open developed during device operation when thin bridges of metal finally "burned out".)



Defect: Metallization smeared so that the distance between two adjacent metal areas is reduced to less than 50% of designed separation (para. 6.6.7). Also distance between a bond and metal other than that to which it is bonded is less than 1/2 mil (.0005 inch) (para. 6.7.7).

Cause: The smear appears to be due to a probe mark on the unused pad.

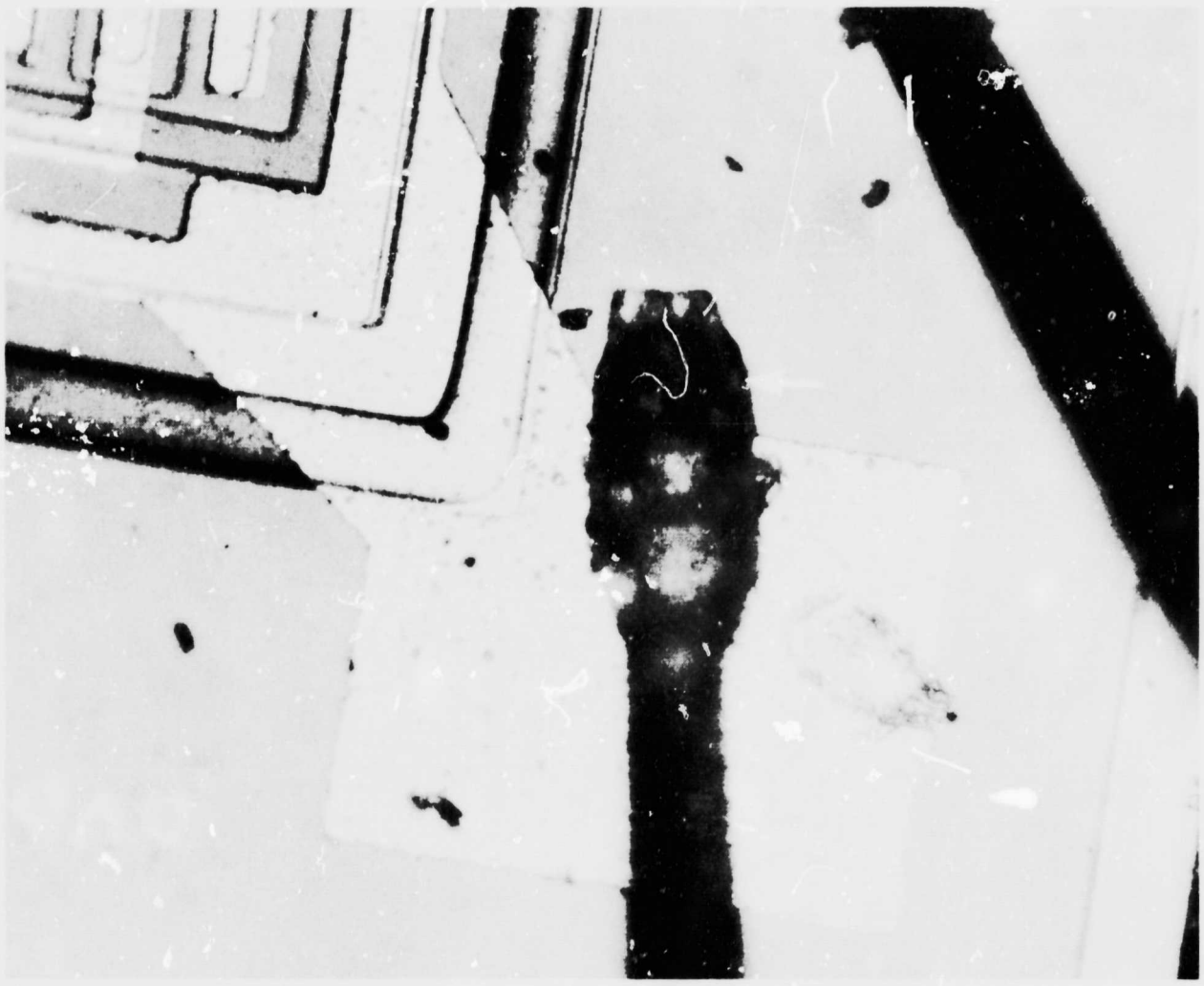
Hazard: A short could occur between the bond (or the pad to which it is attached) to the adjacent pad.



Defect: Intermetallic compounds projecting beyond the bond periphery when viewed from above (para. 6.7.3). In the accompanying photograph, a darkening of the metallization may be seen beyond the periphery of a gold ball bond. Another bond has been deliberately removed to show more of the intermetallic compound formation which is normally partially hidden by the bond.

Cause: These compounds always occur when certain dissimilar metals are alloyed. The amount of the formation and the rate of growth depend on many factors.

Hazard: The bond may become detached as a direct or indirect result of the intermetallic compound formation. Electrical conductivity between the bond and unaffected metal may be reduced because some phases of the intermetallic compound are poor conductors.

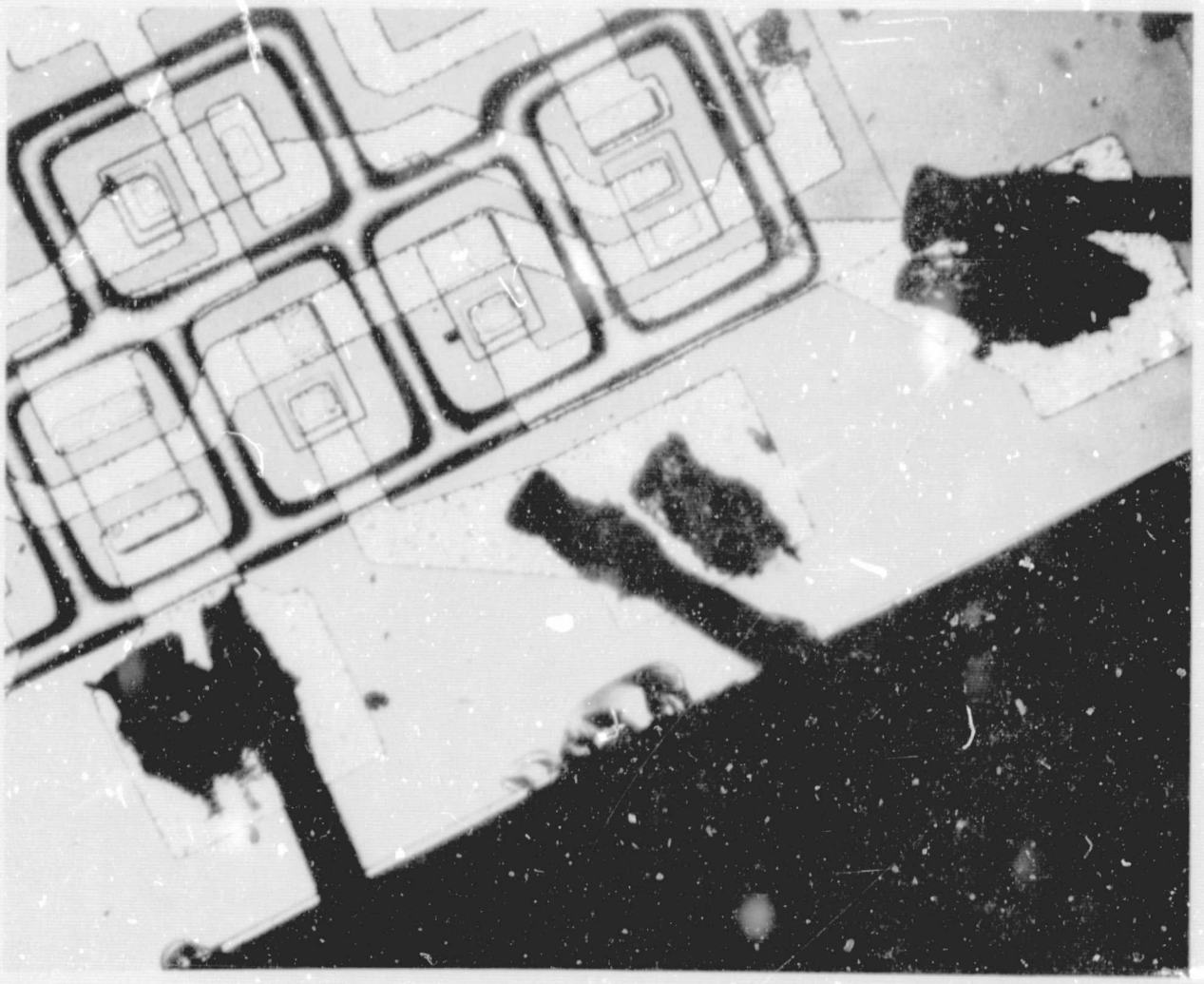


Defect: Less than 75% of compressed area of bond within bonding pad (para. 6.7.5).
Cause: Poor bond placement by bonding operator.
Hazard: Increases chances of an open developing at the bond.

Defect: Less than 1/2 mil (.005 inch) separation between bond and metal to which it is not bonded (para 6.7.7).

Cause: Poor bond placement aggravated by moderately long bond tail.

Hazard: Situation increases possibility of a short circuit.



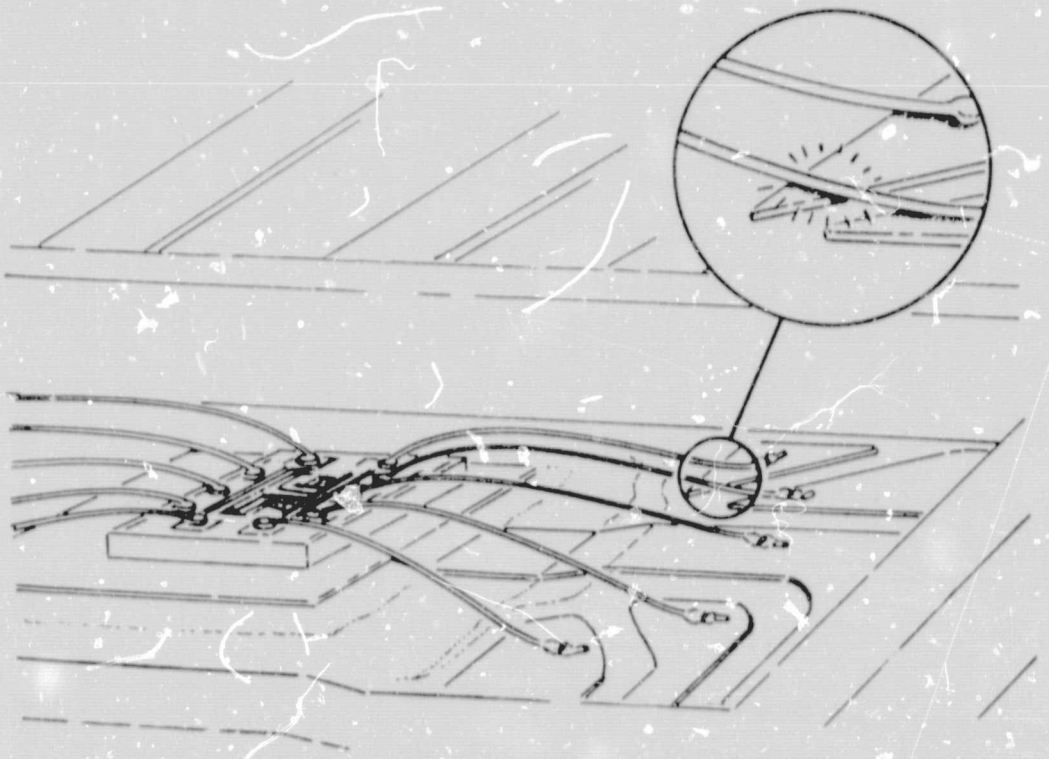
Defect: Evidence of rebonds on die (para. 6.7.11).
Cause: Original bond was defective.
Hazard: Aside from indicating that bonding process may have been out of control at time device was bonded, the possibility exists that unused bond fragments may break free causing presence of loose conducting particles.

Defect: Evidence of rebonding (para. 6.7.11).

Cause: Initial bond pulled off, removing metallization in the process.

Hazard: Missing or disturbed metallization due to a prior bond attempt may result in inadequate bonding in subsequent attempts.

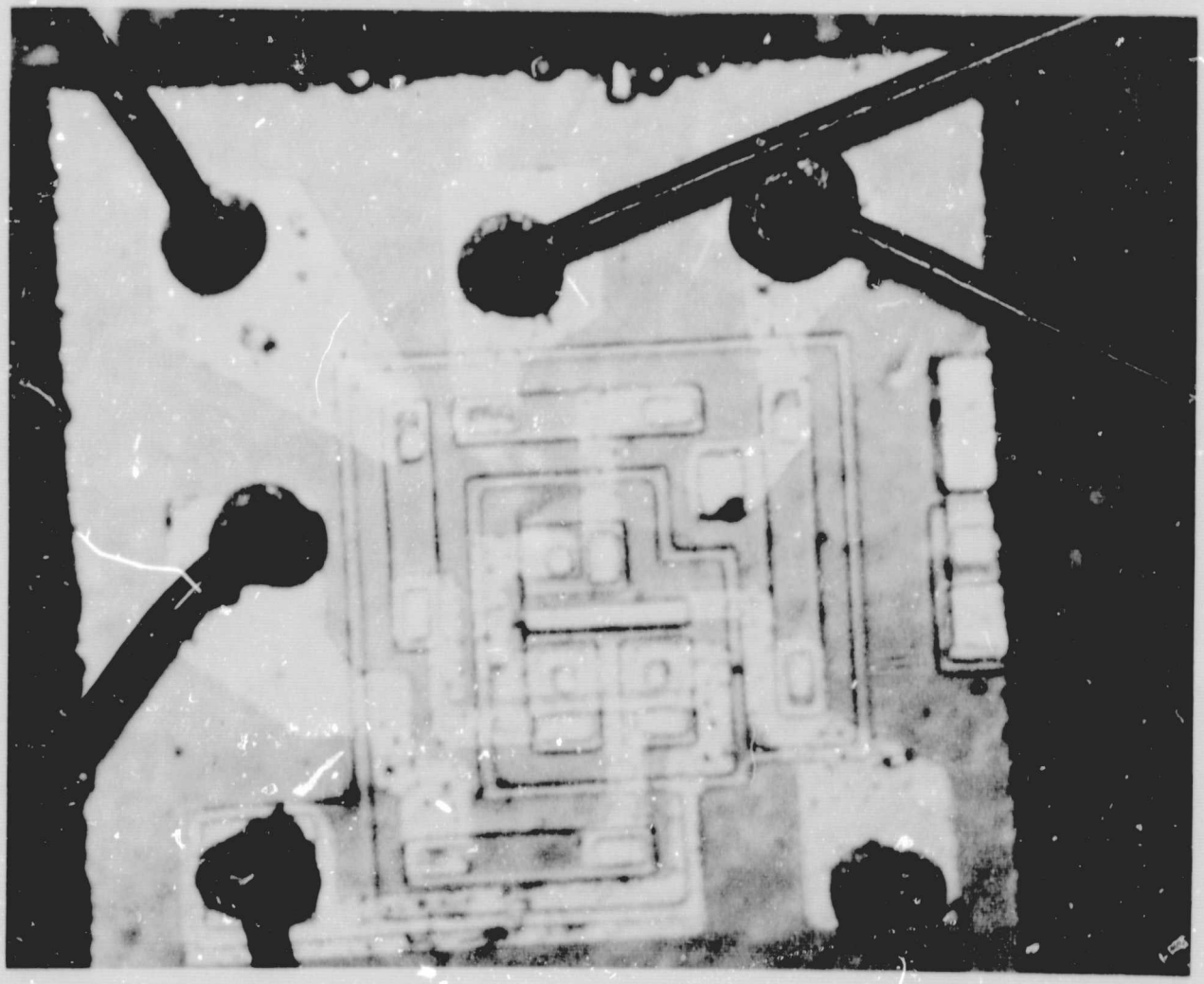
Defect: Lead wires passing over other metallization (para. 6.8.2).
Cause: In this case, the cause is in the design of the device.
Hazard: There is a possibility of shorting between the wire and the metallization over which it passes.



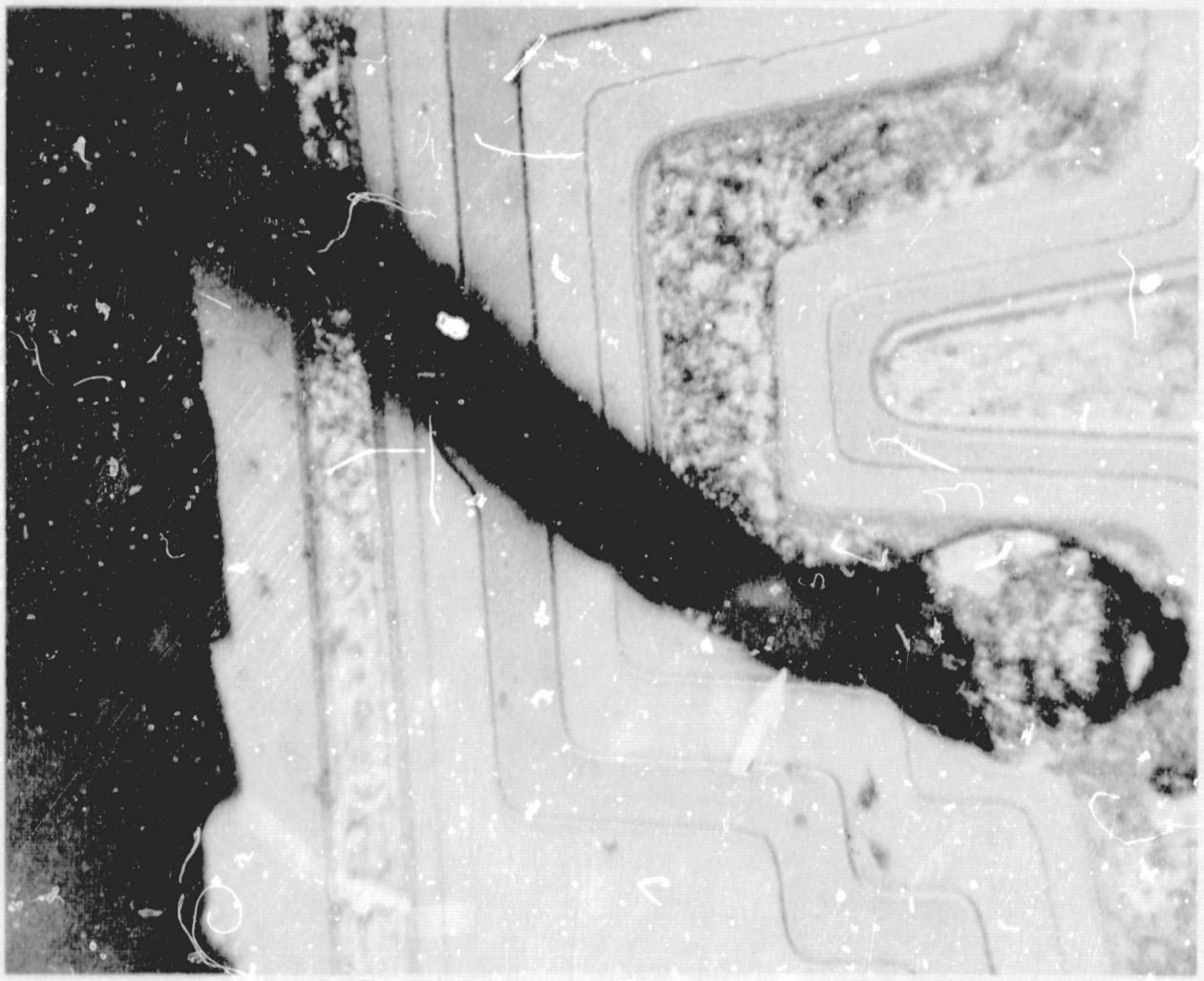
Defect: Wire crosses over other terminal lead (para. 6.8.2).

Cause: Poor placement of bond on terminal lead.

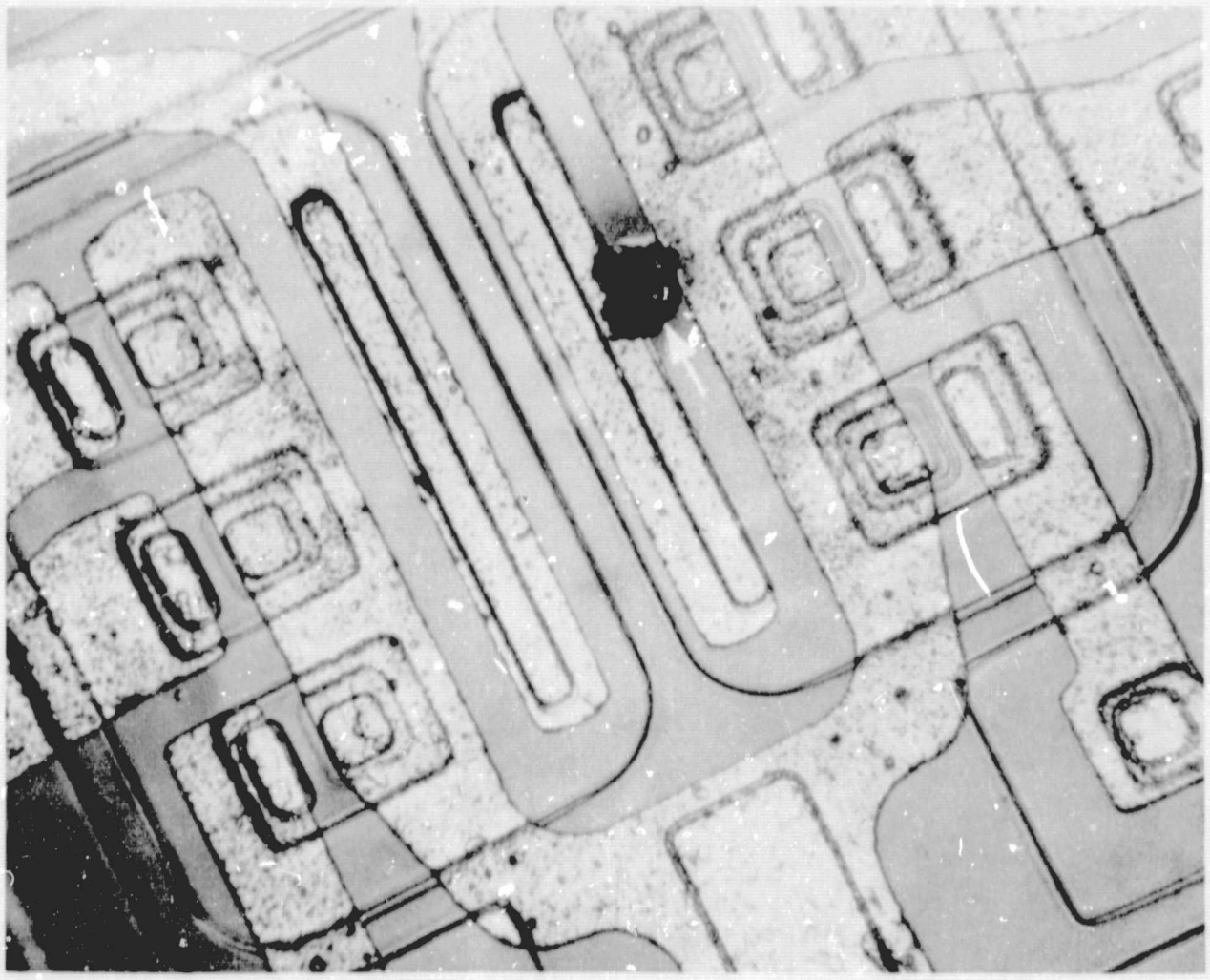
Hazard: Short circuit may occur.



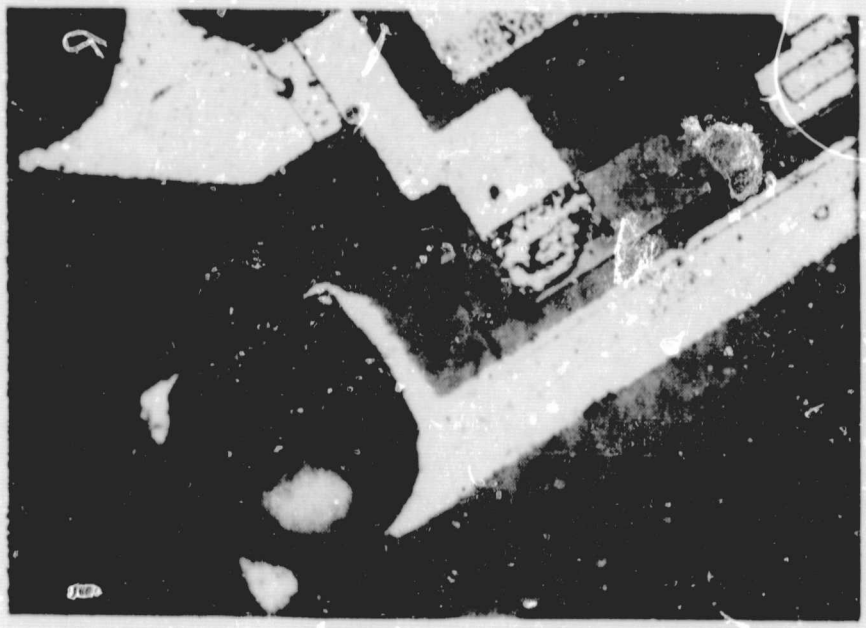
Defect: Bonding wire passes over adjacent bond (para. 6.8.2).
Cause: Improper die orientation resulted in defective wire routing.
Hazard: A short between wire and adjacent bond could develop.

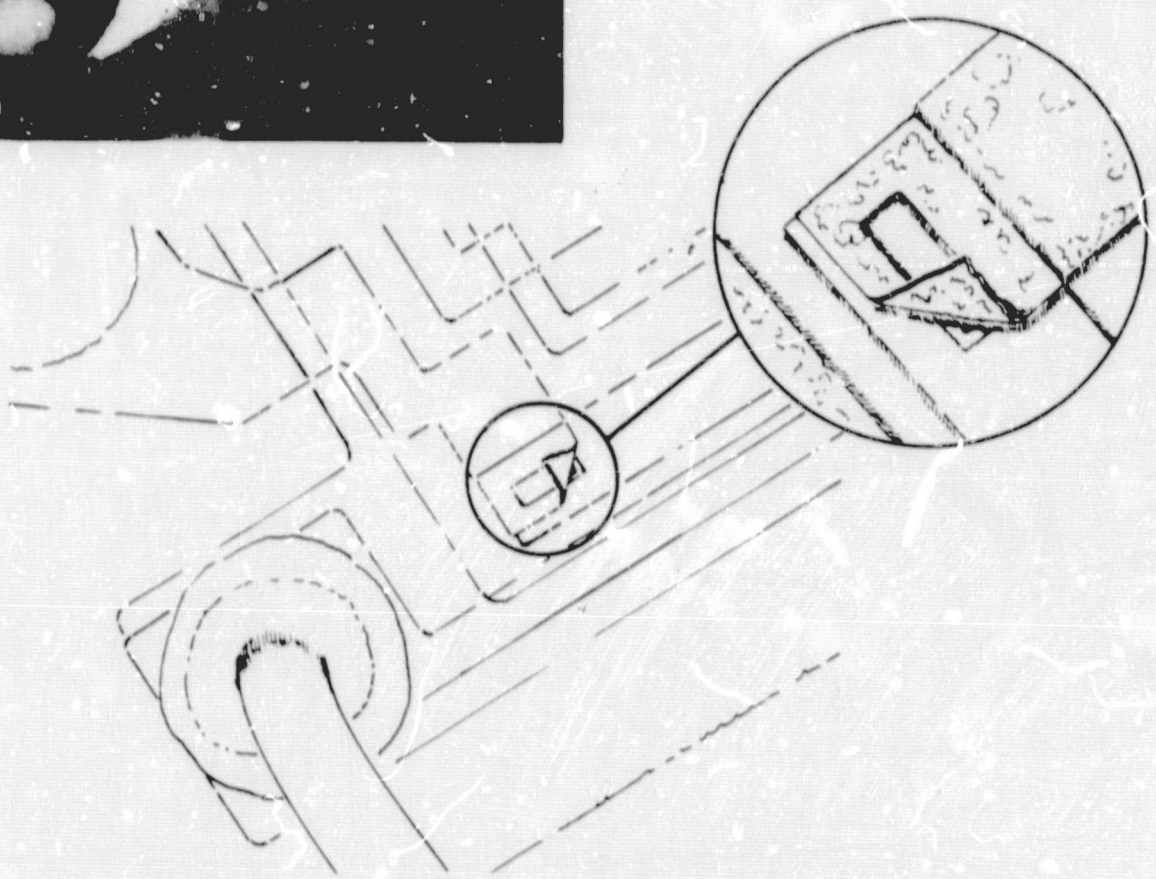


Defect: Horizontal angle between bond axis and the direction of the wire exceeds 15 degrees (para. 6.8.3).
Cause: Improper angle of bond placement in relation to position of bonding posts to which the other end of the wire is attached.
Hazard: The lateral pull at the "heel" of the bond can cause tearing which weakens the bond at that point.

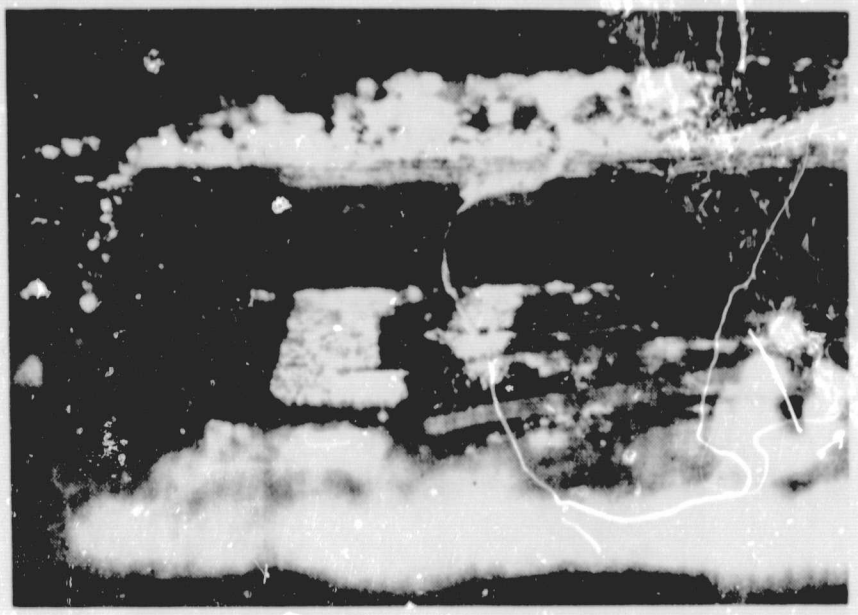


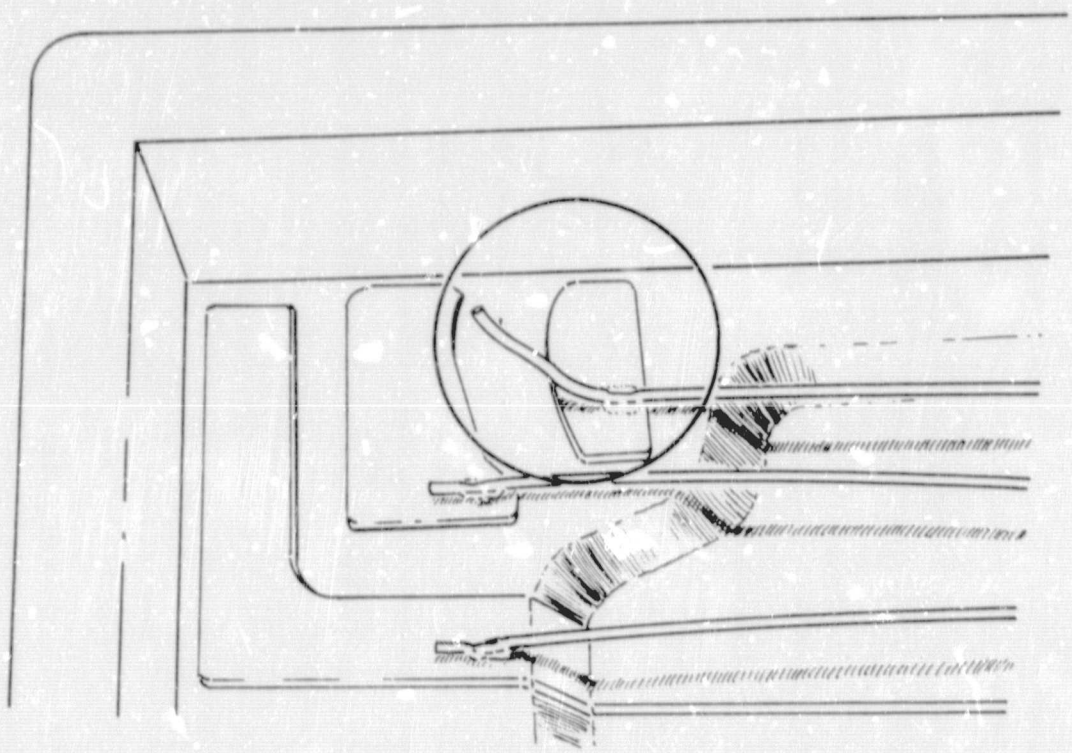
Defect: Foreign matter on die which is larger than the narrowest metallization separation (para. 6.9.2).
Cause: Improper cleaning and/or contamination control.
Hazard: The contamination creates the possibility of a short.





Defect: Metallization peeling from die surface (para. 6.6.2).
Cause: A poor metallization adhesion possibly due to inadequate alloying or sintering.
Hazard: Open circuits and/or creation of loose particles.

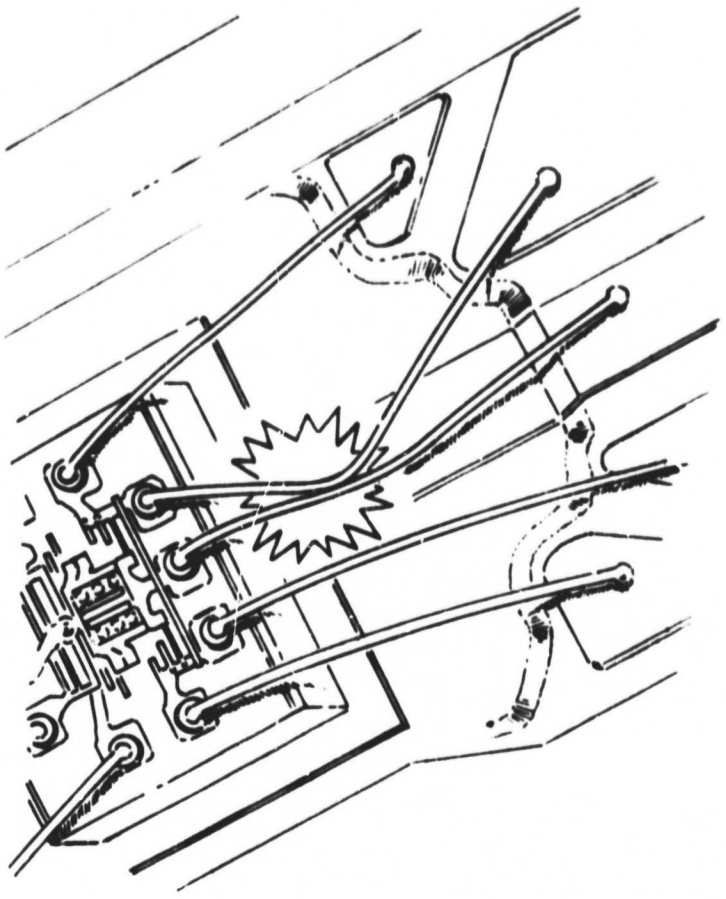




Defect: Remaining portion of bond "pigtail" which is more than 3 wire diameters in length. (para 6.8.12).

Cause: "Pigtail" not removed after bonding.

Hazard: Short and/or a potentionally loose particle.



Defect: Less than 2 wire diameters between adjacent wires (para 6.8.1).

Cause: Overlong wire and/or wire hit before lid sealing.

Hazard: Shorts.

BIBLIOGRAPHY

J. Lindmayer, C.Y. Wrigley; Fundamentals of Semiconductor Devices; D. Van Nostrand Co., Inc., Princeton, N.J.; 1965.

R.M. Warner, Jr., Editor; Integrated Circuits, Design Principles and Fabrication; McGraw-Hill Book Company, New York; 1965.

S. Schwartz, Editor; Integrated Circuit Technology; McGraw-Hill Book Company, New York; 1967.

S.L. Marshall, Editor; Microelectronic Technology; Boston Technical Publishers, Inc., Cambridge, Mass., 1967.

D.J. Manus; "Device Fabrication Technology", T.I. Quality and Reliability Semiconductor Seminar Papers (pamphlet published by Texas Instruments, Inc., Dallas, Texas).

A TRAINING PROGRAM OUTLINE FOR GOVERNMENT INSPECTORS

1. INTRODUCTION

In compliance with that portion of the RFP pertinent to this contract that reads, in part, "A discussion of training of inspectors for internal visual inspection shall be included in the standard." the following Training Program Outline has been constructed. It is understood that the "inspector" herein referenced is the inspector provided by the government.

It is recommended that, wherever possible, the place of training be separated from that location where the actual inspection of parts is to be undertaken, such as in a semiconductor manufacturing plant. It is important that the inspector be given an unbiased and objective view of the problems inherent in integrated circuit manufacturing.

2. TRAINING SCHEDULE

The following training schedule is proposed:

	DAYS
2.1 A brief course in semiconductor processes and fabrication techniques will be given. Topics for discussion will include:	2
2.1.1 Nature and properties of a semiconductor crystal.	
2.1.2 Methods used to fabricate P-N junctions.	
2.1.3 The Planar process.	
2.1.4 Fabrication of mono and bi-polar transistors.	
2.1.5 Metallization and bonding methods.	
2.1.6 Written test on above material.	
2.2 Instruction in the use of a microscope and accessories.	1
2.2.1 Low power microscopy ($\leq$ 80X).	
2.2.2 High power microscopy.	

	DAYS
2.2.3 Use of bright and dark field, polarized light, phase contrast, and filters.	
2.2.4 Photomicrography.	
2.2.5 Filar eyepiece.	
2.2.6 Test on microscopes and accessories.	
2.3 Viewing instructions.	2
2.3.1 The significance of the various colors seen when viewing an integrated circuit under direct light.	
2.3.2 Identification of circuit areas (transistors, resistors, isolation regions, buried layers).	
2.3.3 Identification of regions prone to having workmanship defects.	
2.4 Review of NASA STD XX-2.	5
2.4.1 Explanation of terms used.	
2.4.2 Discussion of reject criteria.	
2.4.3 Test.	
2.4.4 Viewing of actual examples of reject criteria as specified in the workmanship standard.	
2.4.5 Test of actual units, both good and bad.	