MIL-HDBK-978-NASA
1 MARCH 1988

SUPERSEDING
MIL-HDBK-978-A (NASA)
15 MARCH 1984

MILITARY HANDBOOK

NASA PARTS APPLICATION HANDBOOK

(VOLUME 2 OF 5)
DIODES, TRANSISTORS, MICROWAVE DEVICES

FSC 59GP

AMSC N/A
DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

MIL-HDBK-978B (NASA)

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

WASHINGTON, D.C. 20546

NASA Parts Application Handbook

1. This handbook is approved for use by all elements of the National Aeronautics and Space Administration and is available for use by all departments and agencies of the Department of Defense.

2. Beneficial comments (recommendations, additions, deletions) and any pertinent data which may be of use in improving this document should be addressed to: Manager, NASA Parts Project Office, Goddard Space Flight Center, Greenbelt, Maryland 20771.

FOREWORD

This handbook provides a technological baseline for parts used throughout NASA programs. The information included will improve the utilization of the NASA Standard Electrical, Electronic, and Electromechanical (EEE) Parts List (MIL-STD-975) and provide technical information to improve the selection of parts and their application, and failure analysis on all NASA projects. This handbook consists of five volumes and includes information on all parts presently included in MIL-STD-975.

This handbook (Revision B) succeeds the initial release. Revision A was not released. The content in Revision B has been extensively changed from that in the initial release.

CONTENTS

GENERAL CONTENTS

VOLUMES 1 through 5

VOLUME 1

Section 1. Introduction
2. Capacitors
3. Resistors and Thermistors

VOLUME 2

4. Diodes
5. Transistors
6. Microwave Devices

VOLUME 3

7. Microcircuits

VOLUME 4

8. Crystals
9. Filters
10. Transformers and Inductors
11. Delay Lines
12. Motors

VOLUME 5

13. Connectors, Power
14. Connectors, Radio Frequency
15. Protective Devices
16. Switches
17. Relays
18. Wire and Cable

CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

CONTENTS

VOLUME CONTENTS

VOLUME 2

			Page
Paragraph	4.	DIODES	4-1
	4.1	General	4-1
	4.1.1	Introduction	4-1
	4.1.2	General definitions and abbreviations	4-1
	4.1.3	NASA standard parts	4-6
	4.1.4	General device characteristics	4-6
	4.1.5	General parameter information	4-10
	4.1.6	General reliability considerations	4-16
	4.2	Switching	4-35
	4.2.1	Introduction	4-35
	4.2.2	Usual applications	4-35
	4.2.3	Physical construction	4-37
	4.2.4	Military designation	4-40
	4.2.5	Electrical characteristics	4-41
	4.2.6	Environmental considerations	4-43
	4.2.7	Reliability considerations	4-44
	4.3	Rectifiers and power diodes	4-45
	4.3.1	Introduction	4-45
	4.3.2	Usual applications	4-45
	4.3.3	Physical construction	4-47
	4.3.4	Military designation	4-50
	4.3.5	Electrical characteristics	4-52
	4.3.6	Environmental considerations	4-53
	4.3.7	Reliability considerations	4-53
	4.4	Voltage regulator	4-55
	4.4.1	Introduction	4-55
	4.4.2	Usual applications	4-55
	4.4.3	Physical construction	4-62
	4.4.4	Military designation	4-66
	4.4.5	Electrical characteristics	4-66
	4.4.6	Environmental considerations	4-74
	4.4.7	Reliability considerations	4-74
	4.5	Voltage reference	4-76
	4.5.1	Introduction	4-76
	4.5.2	Usual applications	4-76
	4.5.3	Physical construction	4-77
	4.5.4	Military designation	4-81
	4.5.5	Electrical characteristics	4-82
	4.5.6	Environmental considerations	4-89
	4.5.7	Reliability considerations	4-89

CONTENTS

VOLUME CONTENTS

VOLUME 2

			Page
Paragraph	4.6	Current regulator- - - - - - - - - - - - - - - - -	4-91
	4.6.1	Introduction - - - - - - - - - - - - - - - - - - -	4-91
	4.6.2	Usual applications - - - - - - - - - - - - - - - -	4-91
	4.6.3	Physical construction- - - - - - - - - - - - - - -	4-93
	4.6.4	Military designation - - - - - - - - - - - - - - -	4-96
	4.6.5	Electrical characteristics - - - - - - - - - - - -	4-96
	4.6.6	Environmental considerations - - - - - - - - - - -	4-96
	4.6.7	Reliability considerations - - - - - - - - - - - -	4-96
	4.7	Voltage-variable capacitor - - - - - - - - - - - -	4-98
	4.7.1	Introduction - - - - - - - - - - - - - - - - - - -	4-98
	4.7.2	Usual applications - - - - - - - - - - - - - - - -	4-98
	4.7.3	Physical construction- - - - - - - - - - - - - - -	4-99
	4.7.4	Military designation - - - - - - - - - - - - - - -	4-102
	4.7.5	Electrical characteristics - - - - - - - - - - - -	4-102
	4.7.6	Environmental considerations - - - - - - - - - - -	4-104
	4.7.7	Reliability considerations - - - - - - - - - - - -	4-104
	4.8	Bipolar transient voltage suppressors (TVSs) - -	4-105
	4.8.1	Introduction - - - - - - - - - - - - - - - - - - -	4-105
	4.8.2	Usual applications - - - - - - - - - - - - - - - -	4-105
	4.8.3	Physical construction- - - - - - - - - - - - - - -	4-109
	4.8.4	Military designation - - - - - - - - - - - - - - -	4-112
	4.8.5	Electrical characteristics - - - - - - - - - - - -	4-113
	4.8.6	Environmental considerations - - - - - - - - - - -	4-114
	4.8.7	Reliability considerations - - - - - - - - - - - -	4-114
	4.9	Monolithic and multiple arrays - - - - - - - - - -	4-115
	4.9.1	Introduction - - - - - - - - - - - - - - - - - - -	4-115
	4.9.2	Usual applications - - - - - - - - - - - - - - - -	4-115
	4.9.3	Physical construction- - - - - - - - - - - - - - -	4-117
	4.9.4	Military designation - - - - - - - - - - - - - - -	4-121
	4.9.5	Electrical characteristics - - - - - - - - - - - -	4-121
	4.9.6	Environmental considerations - - - - - - - - - - -	4-123
	4.9.7	Reliability considerations - - - - - - - - - - - -	4-123
	4.10	Silicon controlled rectifiers (SCR)- - - - - - -	4-124
	4.10.1	Introduction - - - - - - - - - - - - - - - - - - -	4-124
	4.10.2	Usual applications - - - - - - - - - - - - - - - -	4-125
	4.10.3	Physical construction- - - - - - - - - - - - - - -	4-129
	4.10.4	Military designation - - - - - - - - - - - - - - -	4-131
	4.10.5	Electrical characteristics - - - - - - - - - - - -	4-131
	4.10.6	Environmental considerations - - - - - - - - - - -	4-138
	4.10.7	Reliability considerations - - - - - - - - - - - -	4-138

CONTENTS

VOLUME CONTENTS

VOLUME 2

			Page
Paragraph	4.11	Photodiodes	4-140
	4.11.1	Introduction	4-140
	4.11.2	Usual applications	4-142
	4.11.3	Physical construction	4-145
	4.11.4	Military designation	4-147
	4.11.5	Electrical characteristics	4-147
	4.11.6	Environmental considerations	4-160
	4.11.7	Reliability considerations	4-160
	5.	TRANSISTORS	5-1
	5.1	General	5-1
	5.1.1	Introduction	5-1
	5.1.2	General definitions	5-1
	5.1.3	NASA standard parts	5-4
	5.1.4	General device characteristics	5-4
	5.1.5	General parameter information	5-5
	5.1.6	General reliability considerations	5-5
	5.2	Low-power	5-21
	5.2.1	Introduction	5-21
	5.2.2	Usual applications	5-23
	5.2.3	Physical construction	5-41
	5.2.4	Military designation	5-43
	5.2.5	Electrical characteristics	5-44
	5.2.6	Environmental considerations	5-59
	5.2.7	Reliability considerations	5-59
	5.3	High-power	5-60
	5.3.1	Introduction	5-60
	5.3.2	Usual applications	5-64
	5.3.3	Physical construction	5-74
	5.3.4	Military designation	5-76
	5.3.5	Electrical characteristics	5-76
	5.3.6	Environmental considerations	5-97
	5.3.7	Reliability considerations	5-97
	5.4	Field effect transistors (FET)	5-98
	5.4.1	Introduction	5-98
	5.4.2	Usual applications	5-102
	5.4.3	Physical construction	5-105
	5.4.4	Military designation	5-107
	5.4.5	Electrical characteristics	5-107
	5.4.6	Environmental considerations	5-111
	5.4.7	Reliability considerations	5-111

CONTENTS

VOLUME CONTENTS

VOLUME 2

			Page
Paragraph	5.5	Optocouplers	5-112
	5.5.1	Introduction	5-112
	5.5.2	Usual applications	5-113
	5.5.3	Physical construction, packaging configurations, and device combinations	5-117
	5.5.4	Military designation	5-119
	5.5.5	Electrical characteristics	5-120
	5.5.6	Environmental considerations	5-123
	5.5.7	Reliability considerations	5-124
	6.	MICROWAVE DEVICES	6-1
	6.1	General	6-1
	6.1.1	Introduction	6-1
	6.1.2	General definitions	6-2
	6.1.3	NASA standard parts	6-3
	6.1.4	General device characteristics	6-3
	6.1.5	General parameter information	6-3
	6.1.6	General reliability considerations	6-4
	6.2	Microwave diodes	6-5
	6.2.1	Introduction	6-5
	6.2.2	Usual applications	6-6
	6.2.3	Physical construction	6-16
	6.2.4	Military designation	6-17
	6.2.5	Electrical characteristics	6-18
	6.2.6	Environmental considerations	6-25
	6.2.7	Reliability considerations	6-27
	6.3	Microwave transistors	6-29
	6.3.1	Introduction	6-29
	6.3.2	Usual applications	6-42
	6.3.3	Physical construction	6-51
	6.3.4	Military designation	6-64
	6.3.5	Electrical characteristics	6-64
	6.3.6	Environmental considerations	6-75
	6.3.7	Reliability considerations	6-75
	6.4	Microwave attenuators and loads	6-85
	6.4.1	Introduction	6-85
	6.4.2	Usual applications	6-85
	6.4.3	Physical construction	6-86
	6.4.4	Military designations	6-86
	6.4.5	Electrical characteristics	6-88
	6.4.6	Environmental considerations	6-88
	6.4.7	Reliability considerations	6-88

CONTENTS

VOLUME CONTENTS

VOLUME 2

			Page
Paragraph	6.5	Directional couplers	6-90
	6.5.1	Introduction	6-90
	6.5.2	Usual applications	6-90
	6.5.3	Physical construction	6-90
	6.5.4	Military designations	6-91
	6.5.5	Electrical characteristics	6-91
	6.5.6	Environmental considerations	6-93
	6.5.7	Reliability considerations	6-93
	6.6	Flanges and waveguide assemblies	6-94
	6.6.1	Introduction	6-94
	6.6.2	Usual applications	6-94
	6.6.3	Physical construction	6-95
	6.6.4	Military designations	6-95
	6.6.5	Electrical characteristics	6-96
	6.6.6	Environmental considerations	6-96
	6.6.7	Reliability considerations	6-96
	6.7	Radio frequency power dividers and combiners	6-97
	6.7.1	Introduction	6-97
	6.7.2	Usual applications	6-97
	6.7.3	Physical construction	6-107
	6.7.4	Military designation	6-114
	6.7.5	Electrical characteristics	6-114
	6.7.6	Environmental considerations	6-116
	6.7.7	Reliability considerations	6-117
	6.8	Electrical mixers	6-118
	6.8.1	Introduction	6-118
	6.8.2	Usual applications	6-120
	6.8.3	Physical construction	6-127
	6.8.4	Military designations	6-128
	6.8.5	Electrical characteristics	6-128
	6.8.6	Environmental considerations	6-141
	6.8.7	Reliability considerations	6-141
	6.9	Circulators and Isolators	6-145
	6.9.1	Introduction	6-145
	6.9.2	Usual applications	6-145
	6.9.3	Physical construction	6-145
	6.9.4	Military designation	6-146
	6.9.5	Electrical characteristics	6-146
	6.9.6	Environmental considerations	6-146
	6.9.7	Reliability considerations	6-148

CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4. DIODES

4.1 General.

4.1.1 Introduction. This section contains information on the various types of semiconductor diodes used in electronic equipment. Each type, although having essentially the same general construction, is specifically modified in die size, pellet metallization, doping, passivation, and physical configuration to accomplish a specific function. This includes small signal detection, rectification, switching, voltage reference, voltage regulation, variable capacitance, high voltage-high current rectification, transient suppression function, digital logic applications using diode arrays, and photo detection. The active portion of a diode is a semiconductor pn junction. The pn junctions are formed in various kinds of semiconductors by several techniques. The principal processes are either diffusion of a p- or an n-type impurity into a base material of the opposite type or by alloying a p-type metal into an n-type base. Another process uses a metal-semiconductor barrier type of junction (rather than the diffused type pn junction) to make a schottky diode.

Applicable military specifications.

MIL-S-19500	Semiconductor Devices, General Specification for
MIL-STD-750	Test Methods For Semiconductor Devices
MIL-STD-975 (NASA)	NASA Standard Electrical, Electronic, and Electromechanical (EEE) Parts List
MIL-STD-217	Reliability Prediction of Electronic Equipment
DOD-HDBK-263	Electrostatic Discharge Control Handbook for Protection of Electrical and Electronic Parts, Assemblies and Equipment (Excluding Electrically Initiated Explosive Devices)

4.1.2 General definitions and abbreviations. This list of definitions is presented as an aid for the interpretation and understanding of the specific diode sections.

4.1 DIODES, GENERAL

4.1.2.1 General definitions.

Anode. The electrode from which the forward current flows within the device.

Bias, forward. The bias which tends to produce current flow in the forward direction (p-type semiconductor region at a positive potential relative to the n-type region).

Bias, reverse. The bias which tends to produce current flow in the reverse direction (n-type semiconductor region at a positive potential relative to the p-type region).

Blocking. The state of a semiconductor device or junction which eventually prevents the flow of current.

Capacitance. The capacitance of a semiconductor device is the capacitance measured at designated terminals under specified conditions of bias and frequency.

Cathode. The electrode to which the forward current flows within the device.

Circuit, open. A circuit shall be considered as open circuited if halving the magnitude of the terminating impedance does not produce a change in the parameter being measured greater than the specified accuracy of the measurement.

Circuit, short. A circuit shall be considered short circuited if doubling the magnitude of the terminating impedance does not produce a change in the parameter being measured greater than the specified accuracy of the measurement.

Coefficient, temperature. The ratio of the change in a parameter to the change in temperature.

Coefficient, voltage-temperature. The change in voltage measured under specified conditions over a specified range of ambient or case temperatures. It is expressed as percent change per degree celsius.

Current, forward. The current flowing through the diode in the direction of lower resistance to the flow of steady direct current.

Current, reverse (leakage). The current flowing through the diode in the direction of higher resistance to steady direct current when a specified reverse voltage is applied.

Current, surge. The maximum current pulse which can be carried by the semiconductor diode for the length of time, repetition frequency, and waveform at the temperature specified.

4.1 DIODES, GENERAL

Diode, monolithic and multiple array. Consists of several diodes fabricated in a single monolithic chip. Monolithic arrays allow diode interconnection to form a desired circuit, whereas multiple arrays are restricted to a given circuit configuration such as a common anode circuit or a common cathode circuit.

Diode, current-regulator. A diode which limits current to an essentially constant value over a specified voltage range.

Diode, photo. A diode that is responsive to radiant energy.

Diode, semiconductor. A semiconductor device having two terminals and exhibiting a nonlinear voltage-current characteristic.

Diode, transient voltage suppressor. Transient voltage suppressors are characterized by two zener diodes oriented back to back and are capable of high voltage transient suppression.

Diode, tuning. A varactor diode used for rf tuning including functions such as automatic frequency control (afc) and automatic fine tuning (aft).

Diode, varactor. A two terminal semiconductor device in which use is made of the property that its capacitance varies with the applied voltage.

Diode, voltage-reference. A diode which is normally biased to operate in the breakdown region of its voltage-current characteristic, and which develops across its terminals a reference voltage of specified accuracy when biased to operate throughout a specified current and temperature range.

Diode, voltage-regulator. A diode which is normally biased to operate in the breakdown region of its voltage-current characteristic and which develops across its terminals an essentially constant voltage throughout a specified current range.

Efficiency, conversion. The ratio of the product of the average values of the direct voltage and direct current output to the total alternating current power input.

Efficiency, rectification. The ratio, expressed as a percentage, of the dc load voltage to the peak ac input voltage in a half-wave rectifier circuit with a resistive load.

Equilibrium, thermal. Thermal equilibrium is considered to have been reached when doubling the test time interval does not produce a change, due to thermal effects, in the parameter being measured that is greater than the specified accuracy of the measurement.

Impedance, dynamic. The ratio of the change in voltage to the corresponding change in current under the specified test conditions.

4.1 DIODES, GENERAL

Impedance, small-signal breakdown. The ratio of the small signal ac voltage to the ac current in the breakdown voltage region of the V-I characteristic under the specified test conditions.

Impedance, small-signal forward. The ratio of the small-signal ac voltage to the ac current in the forward region of the V-I characteristic under the specified conditions.

Junction, semiconductor. A region of transition between semiconductor regions of different electrical properties (e.g., p^+-n, n^+-p semiconductors) or between a metal and a semiconductor.

Matched pair. A pair of diodes identical in outline dimensions and with matched electrical characteristics. The two diodes may both be forward polarity, one forward and one reverse polarity, or both reverse polarity.

Metallurgical bond. A metallurgical bond occurs when two or more materials (metal or semiconductor) are brought into contact under temperature and pressure to form a eutectic melt, melt solution, or solid diffusion of the materials. The bond solidifies to form a regrowth or recrystallization region which contains material from both segments of the bond. A good bond is mechanically cohesive and able to withstand a predefined level of tensile stress.

Noise figure. At a selected input frequency, the noise figure is the ratio of the total noise power per unit bandwidth (at a corresponding output frequency) delivered to the output termination to the portion thereof contributed at the input frequency by the input termination, whose noise temperature is standard (293 ±5 K) at all frequencies.

Package type. A category in which all packages have the same case outline, configuration, materials (including bonding, wire, or ribbon and die attach) piece parts (excluding preforms which differ only in size), and assembly processes.

Rating. The nominal value of any electrical, thermal, mechanical, or environmental quantity assigned to define the operating conditions under which a component, machine, apparatus, electronic device, etc. is expected to give satisfactory service.

Rating, absolute maximum. The values specified on data or specification sheets for "maximum ratings," or "absolute maximum ratings" are based on the "absolute system" and unless otherwise required for a specific test method are not to be exceeded under any application or test conditions. These ratings are limiting values beyond which the serviceability of any individual semiconductor device may be impaired. Unless otherwise specified, the voltage, current, and power ratings are based on continuous dc power conditions at free air ambient temperature of 25 °C. For pulsed or other conditions or operation of a similar nature, the current, voltage, and power dissipation ratings are a function of time and duty cycle.

4.1 DIODES, GENERAL

In order not to exceed absolute ratings, the equipment designer has the responsibility of determining an average design value for each rating below the absolute value of that rating by a safety factor, so that the absolute values will never be exceeded under any usual conditions of supply-voltage variation, load variation, or manufacturing variation in the equipment itself.

Regulation, breakdown-voltage. The change in breakdown voltage between two specified values of reverse current.

Rectifiers, silicon controlled. A bistable semiconductor device that comprises three or more junctions and can be switched between conducting and nonconducting states.

Resistance, thermal. Thermal resistance is the temperature rise, per unit power dissipation, of a junction above the temperature of a stated external reference point under conditions of thermal equilibrium.

Small signal. A signal shall be considered small if doubling its magnitude does not produce a change in the measured parameter greater than the specified accuracy of the measurement.

Source, constant-current. A current source is considered constant if halving the generator impedance does not produce a change in the measured parameter greater than the required precision of the measurement.

Source, constant-voltage. A voltage source is considered constant if doubling the generator impedance does not produce a change in the measured parameter greater than the required precision of the measurement.

Surge Current. See Current, Surge.

Temperature, Ambient. The air temperature measured below a semiconductor device, in an environment of substantially uniform temperature, cooled only by natural air convection and not materially affected by reflective and radiant surfaces.

Time, delay. The delay time of a pulse is the time interval from a point at which the leading edge of the input pulse has risen to 10 percent of its maximum amplitude to a point at which the leading edge of the output pulse has risen to 10 percent of its maximum amplitude.

Time, recovery. See "Time, recovery, reverse" and/or "Time, recovery, forward."

Time, recovery, forward. The time required for the current or voltage to reach a specified condition after instantaneous switching from zero or a specified reverse voltage to a specified forward biased condition.

Time, recovery, reverse. The time required for the current or voltage to reach a specified condition after instantaneous switching from a specified forward current condition to a specified reversed bias condition.

4.1 DIODES, GENERAL

Time, rise. The rise time of a pulse is the time duration in which the amplitude of the leading edge of the pulse is increasing from 10 to 90 percent of the maximum amplitude.

Time, turn-on. Turn-on time is equal to delay time plus the rise time of the output pulse.

Voltage, breakdown. The breakdown voltage is the maximum instantaneous voltage that can be applied across a junction in the reverse direction without an external means of limiting the current. It is also the instantaneous value of reverse voltage at which a transition commences from a region of high small-signal impedance to a region of substantially lower small-signal impedance.

Voltage, forward. The voltage drop resulting from the flow of forward current through the semiconductor diode.

Voltage, reverse. The voltage drop resulting from the flow of reverse current through the semiconductor diode.

4.1.2.2 Abbreviations. Refer to MIL-STD-19500, Appendix B for abbreviations and symbols generally accepted by the electronic industry.

4.1.3 NASA standard parts. See subsection 1.1 Introduction for a complete description of the NASA Standard Parts Program. In addition to this handbook, the principal elements of this program include MIL-STD-975.

MIL-STD-975 is a standard parts list for NASA equipment; Section 4 contains a summary of standard diodes.

4.1.4 General device characteristics. The general device characteristics of switching, rectifier and power, voltage regulator, voltage reference, current regulator, voltage variable capacitor diodes, transient voltage suppressors, multiple and monolithic diode arrays, silicon controlled rectifiers, and photo diodes are briefly discussed in this section.

4.1.4.1 Switching diodes. Switching diodes are normally made of silicon or germanium. The silicon switching diode has a higher forward voltage drop than the germanium diode, but the silicon device has lower leakage currents and higher maximum junction temperatures. Silicon devices exhibit normal rectifying characteristics and have very fast reverse recovery times.

The fast reverse recovery time is accomplished either through the use of gold or platinum doping or by high energy (13 MeV) electron irradiation of the lightly doped region of the device, which greatly reduces the minority carrier lifetime in both that region and at the junction.

4.1.4.2 Rectifier and power diodes. This section includes three general types of rectifiers: power diodes, fast switching power rectifiers, and power Schottky barrier rectifiers.

a. Power diodes. These devices are general purpose low voltage rectifiers available in both axial leaded and stud mounted packages. The dc output reverse working voltages vary from 50 to 1000 V, with the dc output currents varying from 1 to 125 A (depending upon the type of package used). These devices are used extensively in ac power supplies for rectification such as half-wave and full-wave rectifiers, half-wave voltage doublers and full-wave voltage doublers.

b. Fast switching power rectifiers. These devices have reverse working voltages varying from 50 to 1000 V and dc output currents of 1 to 3000 A. They have fast recovery times, typically about 150 ns. This type of rectifier is suitable for use in such circuit applications as inverters, choppers, low rf interference, free-wheeling rectifiers, and dc power supplies.

c. Schottky barrier rectifiers. These devices make use of the rectification effect of a metal to a silicon semiconductor barrier. These devices have better forward conductivity characteristics than conventional pn-junction rectifiers (because the forward voltage is about 0.25 V) but higher reverse leakage currents. However, the lower forward voltage results in higher rectification efficiency, which is somewhat offset by a greater reverse power dissipation. Power Schottky diodes are used primarily in low voltage applications where pn-junction rectifiers would result in excessive power dissipation due to forward losses and/or commutation losses. Because the forward conduction is mainly by majority carriers, the usual recovery characteristics caused by minority carrier storage in pn junctions is completely absent. Schottky diodes can, therefore, switch very rapidly. For this reason they are ideally suited to rectification of high-frequency ac power for switching regulators and converters.

   In designing with Schottky power diodes, the reverse power dissipation is a significant factor (unlike pn-junction rectifiers). Therefore, the thermal design of the equipment must be sufficiently conservative to prevent thermal runaway and destruction of Schottky rectifiers.

4.1.4.3 Voltage regulators. The voltage regulator is a silicon junction diode possessing a very high back resistance up to its critical reverse breakdown, or zener, voltage. At this point, the back resistance drops to a very small value. In this region, the current increases very rapidly, whereas the voltage drop across the diode remains almost constant. Voltage regulators (commonly referred to as zener diodes), when biased in the reverse direction, can be used as voltage regulators, or reference elements because the reverse voltage at V_Z will remain essentially constant for a wide range of current. The power dissipation capabilities are usually in the range of 0.4 to 20 W. The temperature coefficient ranges from typically -2mV/°C for a V_Z = 3.0 V to +175 mc/°C for a V_Z = 200 V.

4.1 DIODES, GENERAL

4.1.4.4 Voltage references diodes. The silicon voltage reference diode consists of two or more silicon junctions encapsulated into a single package but designed to exhibit a relatively fixed voltage and a fixed current with a very low temperature coefficient (such as 0.0005%/°C for a V_Z = 6.35 V). The effective voltage time stability initial-to-peak varies from 10 to 50 ppm/1000 hours for these devices.

The voltage reference diode can be used in any circuit that requires a stable reference that is insensitive to shock vibration or position. Their inherent stability allows them to be used in circuits requiring an extremely high degree of voltage time stability.

4.1.4.5 Current regulators. The current regulator diode is basically a field effect transistor that has its gate and source connected. This particular device is useful in such applications as over-current protection, transistor biasing, linear ramp or stairstep generators, differential amplifiers, and precision reference voltage sources. The 1N5285 through 1N5314 family series covers a current range from 200 µA to 4.7 mA in 32 different current steps.

4.1.4.6 Voltage variable capacitance diodes. The voltage variable capacitance diode is a silicon pn-junction diode designed for use as a voltage variable capacitor. The capacitance varies essentially as $1/\sqrt{V}$ as the voltage across its terminals is varied. This type of device maintains constant characteristics over a wide temperature range but is less temperature sensitive when used at higher operating voltages. Therefore, if capacitance variations must be minimized, it is advisable to operate the device at higher bias voltages.

4.1.4.7 Transient voltage suppressors. Silicon transient voltage suppressors are used in applications where large voltage transients can permanently damage voltage-sensitive components.

Transient voltage suppressor (TVS) devices have a high current surge capability, extremely fast response time and a low impedance. Because of the unpredictable nature of transients, impedance is not specified as a parametric value. However, a minimum voltage at low current conditions (BV) and a maximum clamping voltage (V_C) at a maximum peak pulse current are specified in the manufacturer's performance specification.

TVS devices are designed to absorb a peak pulse power of 500, 1500 or 15,000 W dissipation for 1 ms. The response time of the clamping action of the TVS device is theoretically instantaneous (1×10^{-12} s): Therefore, they can protect integrated circuits, MOS and MSI integrated circuits, hybrids, and other voltage-sensitive semiconductors and components.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

TVS devices can be used in series or parallel to increase the peak power rating. TVS devices have proven to be effective in airborne avionics and controls, mobile communication equipment, computer power supplies and in many other applications where inductors and switching transients are present.

4.1.4.8 Multiple arrays and monolithic arrays. Diode arrays consist of several diodes which have been fabricated in an integrated circuit form. The diode array is not a unique device with characteristics that differ from discrete diodes; it is simply a packaging technique that permits a significant reduction in the size of electrical systems.

Monolithic arrays allow interconnection of diodes to form a desired circuit, whereas multiple arrays are restricted to a given circuit configuration, such as a common anode or common cathode circuit.

4.1.4.9 Silicon-controlled rectifiers (thyristors). The silicon-controlled rectifier (SCR) is a semiconductor device that has characteristics similar to those of a thyratron gas tube. The device can be switched between states by a current or polarized voltage pulse. Unlike conventional transistors, the device lends itself to use as a high current, high voltage rectifier or a static latch (limited to microseconds), or a sensitive high gain amplifier control.

The silicon-controlled rectifier is well adapted for use as a latching switch or high-power gain amplifier. The SCR can be turned on by a momentary application of control current applied to the control gate, whereas tubes or transistors require a continual signal. The turn-on time is about 1 μs, and turn-off time is about 10 to 20 μs. This latching action can be controlled by signals of only a few microwatts but can switch up to 200 V. With associated transistor-triggering circuits, unusual current gain on the order of 10^9 can be obtained. Because of this advantage there are many applications to which the SCR can be adapted.

4.1.4.10 Photodiodes. Photodiodes are used as light detectors in a variety of applications, such as card and tape readers, pattern and character recognition, shaft encoders, position sensors, and counters.

Photodetectors may be subdivided into four categories: conventional pn junction photodiodes, PIN photodiodes, avalanche photodiodes, and Schottky barrier photodiodes.

Conventional pn junction photodiodes are used in the applications outlined above in the photoconductive mode. The PIN photodiode and the avalanche photodiode are used as visible or infrared detectors in fiber optic system applications usually with a suitable optical source such as a laser diode for long distances and a LED device for shorter distances (1 km).

The Schottky photodiodes are recommended when a high blue response (less than 500 mm) or larger areas (greater than 1 cm^2) are required in the system application.

4.1 DIODES, GENERAL

The main parameters of the photodiode are response time, responsivity, spectral response, noise, and dark current. The peak wavelength of the spectral response curve (R_A vs λ) at a given temperature will largely dictate the circuit application requirements. The peak wavelength is a strong function of the semiconductor material band gap and the doping used to make the junction.

4.1.5 General parameter information. The basic function of the semiconductor diode and general parameter information is presented in the following discussion. For parameter information on transient voltage suppressors (see subsection 4.8), diodes arrays (4.9), silicon-controlled rectifiers (4.10), and photodetectors (4.11), refer to the appropriate individual paragraphs.

4.1.5.1 The pn junction. A pn junction can be formed only by a chemical process within a single crystal. If separate p- and n-type crystals were joined mechanically, a polycrystalline semiconductor would be the subsequent result. This type device will not furnish rectification.

Accompanying the formation of the pn junction is a region known as the depletion zone. This zone is so called because within it there is an absence of holes and excess electrons. This depletion zone is also referred to as the space charge zone because the acceptor and donor ions are fixed and charged electrically. This space charge forms a barrier to current flow.

When an external battery is applied with the positive output attached to the p-layer, the junction is in the forward-biased state. The external voltage source causes the holes in the p-region to move from the positive source potential to the negative potential. An opposite action occurs in the n-material. These two actions cause the depletion zone to shrink, thereby causing less resistance to majority carrier current flow.

When the applied battery or source potential is placed with the negative potential attached to the p-layer, the junction is in the reverse-biased state. The holes in the p-region are drawn toward the negative terminal, while the electrons in the n-region are drawn to the positive terminal. This resultant action causes the depletion zone to increase in thickness, subsequently creating a large resistance for majority carrier flow.

The reverse current tends to maintain a relatively constant value at all voltages up to a voltage called junction breakdown voltage. In this voltage region, current conduction across the junction interface increases rapidly and the diode is often destroyed by heating. There are two causes of voltage breakdown in semiconductor diodes: avalanche breakdown and zener breakdown.

Avalanche voltage breakdown can be thought of as an electrical multiplication process.

In the avalanche process, a free electron acquires enough energy from the applied voltage (which exists across the very narrow junction interface) to accelerate it sufficiently such that when it collides with a fixed electron, it knocks it free. These electrons are again accelerated until each undergoes a second collision, resulting in further electron multiplication. The higher the applied voltage, the more rapid the electron multiplication. The voltage across the junction does not increase substantially, because the energy of the avalanche electron is limited by the critical impact velocity.

Under zener breakdown, the actual breakdown of the semiconductor is initiated through the direct rupture of the covalent bonds due to an exceptionally strong electric field which is developed across the junction. Zener breakdown is more prevalent in a very narrow junction where the field intensity becomes very high. In wider junctions, avalanche breakdown is more prevalent, because the impressed voltage is not confined to such a narrow region. In diodes, the junction will recover when the magnitude of the reverse voltage is reduced below the breakdown voltage provided the diode junction has not been damaged by excessive temperatures while operating in the breakdown region.

4.1.5.2 Diode voltage-current relationship. The ideal pn junction follows the voltage-current relationship as predicted by the simple first-order theory as developed by Schottky. This relationship is expressed by the following equations:

$$I_F = I_{Sat}\left(e^{qV/nkT} - 1\right)$$

or

$$V = \frac{nkT}{q} \ln\left(1 + \frac{I_F}{I_{Sat}}\right)$$

where

I_F = Forward junction current
I_{Sat} = Reverse saturation junction current (due to bulk silicon only)
k = Boltzman's constant
T = Absolute temperature (degrees Kelvin)
q = Electronic charge
V = Voltage across the junction
n = Ideality factor, varies from 1.0 to 2.0, depending upon device design features.

A plot of the typical silicon diode characteristic is shown in Figure 1. This shows that for any reverse voltage, in excess of a few tenths of a volt, a small reverse current is produced which remains essentially constant. When a forward voltage is applied, the forward current increases exponentially.

4.1 DIODES, GENERAL

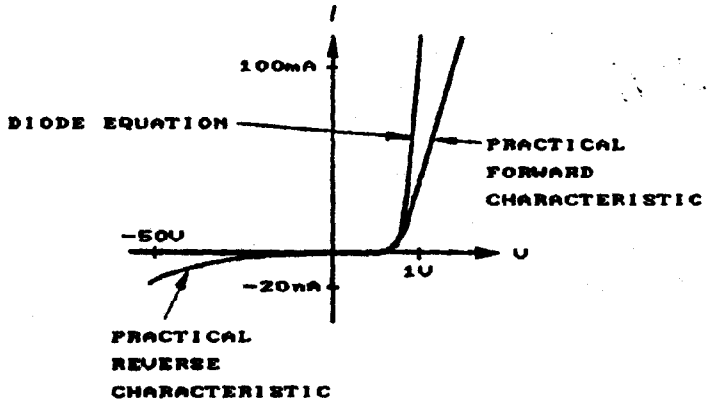


FIGURE 1. Typical silicon diode characteristics.

An ideal rectifier diode has a very low forward resistance and a very high reverse resistance. To create a diode with a low forward drop and a high reverse voltage capability, it is necessary for the semiconductor layer on one side of the junction to be highly doped with impurities (low resistivity) and the opposite layer to have a low doping level (high resistivity). If the p-region is more heavily doped with impurities than the n-region, it then has a greater number of current carriers and becomes the anode.

4.1.5.3 The pn junction turn-off theory. The characteristics of a diode under turn-off conditions may be explained by using a turn-off test circuit such as the one shown in Figure 2, which is a diode in series with a resistor placed across some fixed power supply. The polarity of the power supply may be reversed with the aid of a double-pole double-throw switch. If, initially, the switch is connected so that the diode is biased in the forward direction, the induced voltage causes a steady-state current flow I_F in the diode and load as shown in Figure 2. Consequently, $I_F = (V-V_D)/R_L$.

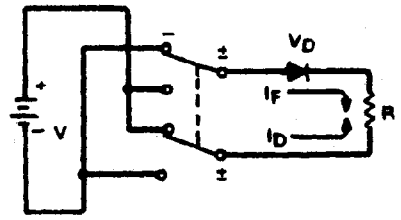


FIGURE 2. Diode turn-off time test circuit.

4.1.5.4 Reverse Transient. The reverse transient occurs when a diode is switched from a forward-conducting state to a reverse-biased condition. The reverse impedance of a diode is very high. However, this high impedance condition does not appear instantaneously when the diode is reverse-biased. A typical reverse recovery waveform is given in Figure 3.

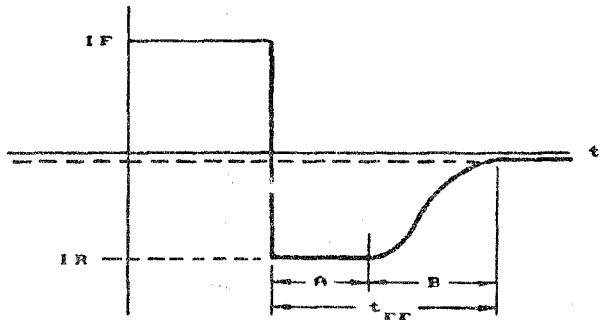


FIGURE 3. Typical reverse recovery waveform.

The delay in the appearance of the high reverse impedance is the result of two factors: minority carrier storage and junction capacitance. As a first approximation, region A can be attributed to minority carrier storage and region B to the junction capacitance.

4.1.5.5 Minority carrier storage. When the junction is forward biased, an excess of minority carriers builds up on either side of the junction region. From pn junction theory, the density of holes in the n region in the vicinity of the junction is given by:

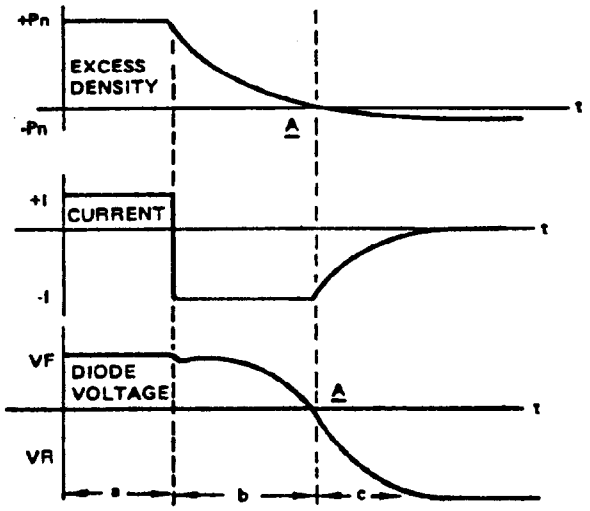
$$P_{(x = o)} = P_n e^{qV/KT}$$

From the definition of excess density,

$$P_{(excess)} = P_{(x=o)} - P_n(e^{qV/KT})$$

Therefore, the excess density is a function of the voltage across the junction of the device. When the diode is forward-biased, the term, $_e qV/KT$, is much greater than 1 because q/KT at room temperature is approximately 40. When the voltage across the device drops to zero, all excess carriers will have been swept from the junction region as shown in Figure 4 (point A). However, the current at point "A" does not instantaneously drop to zero because of the junction capacitance of the device.

4.1 DIODES, GENERAL



a = Forward bias
b = Minority carrier storage
c = Charging of the Junction Capacitance

FIGURE 4. Switching characteristics.

4.1.5.6 Storage time. When the pn junction is biased in the reverse direction for the majority carriers, it is biased in the forward direction for the minority carriers. The application of forward bias will, due to carrier action, increase the minority carrier density in both the anode and cathode sides of the junction. The minority carrier density in the highly doped p layer will be greater than in the n layer. In order to stop the I_F by reversing the external source voltage, it is necessary for the minority carriers to return to the junction. This travel of minority carriers takes a finite time. The diode will remain forward-biased until the minority carrier density at the junction becomes less than the equilibrium value. The finite time that it takes to return to equilibrium is known as storage time (t_s).

4.1.5.7 Junction capacitance. The voltage across the junction of the device creates a space charge on either side of the junction. This space charge is equivalent to a capacitor with a cross-sectional area and plate separation equal to that of the junction, and having a permittivity equal to that of the semiconductor material. The junction capacitance is a nonlinear function of the junction voltage, increasing with forward bias and decreasing as reverse bias increases. This capacitance is given by:

$$C = \frac{K}{(V + \phi)^n}$$

where: K = semiconductor doping constant
V = applied voltage
ϕ = built-in junction potential
n = diode constant (n = 1/3 for graded junction) (diffused) and n = 1/2 for abrupt junction (step).

A typical curve of junction capacitance versus voltage is shown in Figure 5.

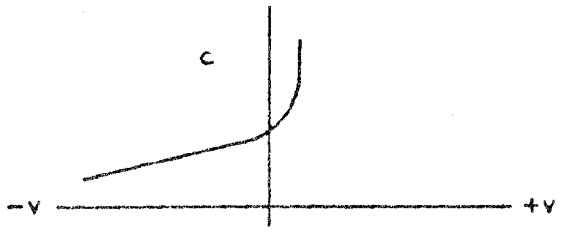


FIGURE 5. Typical junction capacitance vs voltage.

As stated previously, the recovery region C shown in Figure 4 is attributed to the junction capacitance. This capacitance in series with circuit resistance determines the effective time constant of the reverse voltage across the device.

Several factors are responsible for the magnitude and duration of the reverse recovery transient. The magnitude of the transient is primarily controlled by the external circuit in which the diode is used. Typically, this magnitude is given by VR/RL (VR = reverse voltage, RL - limiting resistance). The duration of the transient is dependent upon the forward current (magnitude and duration), reverse current, junction capacitance, series resistance, and temperature.

4.1.5.8 Metal barrier Schottky junction. A Schottky barrier diode consists of a metal semiconductor junction formed between aluminum, gold, silver or platinum metallization, and lightly doped n-type silicon or gallium arsenide with an N/N^+ epitaxial structure. Refer to Figure 6A for a typical structure.

In both materials, the electron is the majority carrier. In the metal, the level of minority carriers (holes) is insignificant. When the materials are joined the electrons in the n-type silicon semiconductor material immediately flow into the adjoining metal, establishing a heavy flow of majority carriers. Because the injected carriers have a very high kinetic energy level compared with the electrons of the metal, they are commonly called "hot carriers." In the conventional pn junction minority carriers are injected into the adjoining region. Here, the electrons are injected into a region of the same electron plurality. Schottky diodes are unique, in that conduction is entirely by majority carriers. The heavy flow of electrons into the metal creates a region near the junction surface depleted of carriers in the silicon material--much

4.1 DIODES, GENERAL

like the depletion region in the pn junction diode. The additional carriers in the metal establish a negative wall in the metal at the boundary between the two materials. The net result is a "surface barrier" between the two materials, preventing further current flow. That is, any electrons (negatively charged) in silicon material face a carrier-free region and a negative wall at the surface of the metal.

The application of a forward bias as shown in Figure 6 will reduce the strength of the negative barrier through the attraction of the applied positive potential for electrons from this region. The result is a return to the heavy flow of electrons across the boundary, the magnitude of which is controlled by the level of the applied bias potential. The barrier at the junction for a Schottky diode is less than that of the pn junction device in both the forward- and reverse-bias regions. The result is a higher current at the same applied bias in the forward- and reverse-bias regions. This is a desirable effect in the forward-bias region but highly undesirable in the reverse-bias region.

As the forward current flow is by the injection of electrons into the metal which already contains a high density of free electrons, the storage effect which limits the switching speed of pn junction diodes, does not occur in Schottky diodes. In addition, by making the junction area (surface metallization) small, the depletion capacitance due to space charge can be kept small. The overall small capacitance provides an extremely fast switching speed of typically few picoseconds.

Due to the high electron concentration in the metal, the drift of the electrons across the junction into the semiconductor is considerably higher than in the pn junction case, resulting in reverse leakage current at least an order of magnitude greater than that for a corresponding silicon pn junction (Figure 6).

4.1.6 General reliability considerations. The ultimate goal of a circuit designer is to produce circuits, which when assembled into a system, will enable the equipment to perform its intended function with less than a specified percentage of down time due to equipment malfunction. To do this, the designer must have a knowledge of all facets of reliability that contribute to system reliability, some of which follow:

a. The relationship of component reliability to system reliability

b. The causes of component failure

c. How reliability is measured

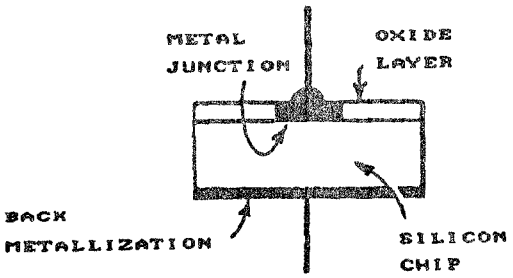
d. The various methods of specifying reliability assurance

e. The factors involved in selecting components

f. The effect of circuit design upon overall system reliability.

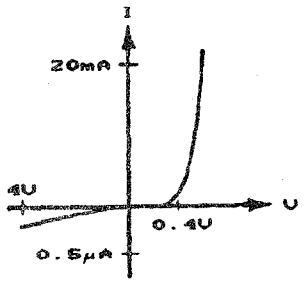
Reliability is usually expressed as failure rate in percent per thousand hours. Because failure rate is the reciprocal of MTBF, it is also possible to express diode failure rates in terms of MTBF, but such an expression can be extremely misleading. For example, failure rates on the order of 1.0 to 0.1% per thousand hours of life testing at maximum rated conditions are available from the semiconductor industry today. A failure rate of 0.1%/1000 hours is equivalent to an MTBF of 1,000,000 hours (over 114 years).

This is a rather meaningless figure for a number of reasons, some of which will be discussed in the following paragraphs.



A. Typical structure

B. Circuit symbol type



C. I-V characteristics type

FIGURE 6. Schottky-barrier diode.

4.1 DIODES, GENERAL

4.1.6.1 Achieving diode reliability. Three major factors contribute to diode reliability:

a. Good basic device design and material design as it relates to packaging

b. Good manufacturing processes

c. Quality and reliability control.

Only when all three factors are optimized will diode reliability be at a maximum.

4.1.6.2 Causes of failure. A general knowledge of the causes of semiconductor device failure is essential to an understanding of diode reliability. Diode failure mechanisms can be broadly classified as follows:

a. Surface defects

b. Mechanical defects

c. Bulk defects

d. Wire/bond defects

e. Contamination defects.

4.1.6.2.1 Surface defects. Most diode failures are related to the condition of the semiconductor surface. A poor surface condition may be caused either by imperfections within the encapsulated diode, or by failure of the package (which causes the semiconductor surface to be subject to the external environment), or a combination of the two. During diode fabrication, every precaution is taken to assure stability of semiconductor surfaces. This is particularly true for the fabrication steps just prior to encapsulation.

Such techniques as (1) the encapsulation of the diodes in an inert atmosphere (such as nitrogen) to reduce the possibility of chemical reaction with the semiconductor surface, (2) the use of getters which absorb moisture to maintain low partial vapor pressure within the package, and (3) the use of surface passivation for silicon devices to form a chemically bonded film for surface deactivation are all designed to stabilize or isolate the semiconductor surface from the surrounding environment.

Stresses, which cause a change in the state of the semiconductor surface during diode life, are another potential source of diode failure. Among the factors which can introduce mechanisms to change the state of the diode surface are the following:

a. Entrapment of moisture or other contaminants within the diode during encapsulation

b. Loss of the hermetic seal due to improper encapsulation; i.e., leaks present at the time the diode is manufactured or occurring during subsequent diode life.

One of the mechanisms of failure is the creation of conductive shunt paths, which can be above or below the oxide surface of passivated silicon diodes.

The surface passivation of silicon diodes by the growth of silicon dioxide or nitride, which are chemically grown on the surface, affords a greater degree of surface protection than has previously been available. However, unless properly designed and manufactured, even this class of diodes may have surface instability problems. Among the causes of these problems are contaminants sealed beneath the passivated surface, pin holes in the passivated film, and ionized conductive paths on the surface of the passivated film. As a result, hermetic encapsulation is still desirable, even for passivated diodes when maximum reliability is required.

Surface defects are most often detected by reverse current (I_R) instability over periods of life stressing. Fabrication techniques are not identical for all device types, and these differences can create different levels of I_R between the device types. The magnitude of I_R, therefore, becomes significant only when compared with the mean I_R for that device. (See Table I).

4.1.6.3 Mechanical defects. The mechanical defects which can occur in diffused diode include:

a. Poor bonding of die to header

b. Poor lead-to-die contact

c. Lack of hermetic seal.

Poor contact of the die to the header may increase the thermal resistance of the diode, resulting in high junction temperatures during high-power operation. Poor lead-to-die contacts may cause hot spots, but this is of secondary importance for relatively low level applications (See Table II). The effects of poor package sealing on surface stability have been reviewed in previous paragraphs.

4.1 DIODES, GENERAL

TABLE I. Failure mechanism analysis surface defects

Failure Mechanism	Description/Cause	Detection Method	Method to Minimize or Eliminate Cause
Ion migration through and across the oxide	Inversion of the n region near the anode contact, surface contamination and lack of an oxide sealant	High temperature, reverse bias, reverse current deltas	Improved control of surface cleanliness, application of phosphosilicate glass over the thermal oxide
Hole trapping in the oxide	Donor states in the oxide become positively charged resulting in an accumulation layer and premature breakdown	High temperature, forward bias, reverse current, deltas	Improved control of surface cleanliness, application of phosphosilicate glass over the oxide
Holes in oxide	Pinholes in the oxide result in shorting paths from anode to cathode	Electrical test	Improved process control

4.1.6.4 Bulk defects. Bulk defects in diodes are generally a less frequent cause of poor reliability than surface or mechanical defects. Bulk defects are often difficult to detect by in-process controls during the diode fabrication process, although they are usually detected at the final electrical test.

Included in this classification of defects are crystal imperfections and undesired impurities. Crystal defects can cause nonuniform diffusion, resulting in high current concentrations and hot spots. Undesired impurities can result in uneven voltage gradients. These uneven voltage gradients can cause, in a worst case, failure due to punch through. A second class of bulk defects results from diffusion of impurities and metal contacts into the bulk material at normal operating temperatures. This problem is generally minimized in a well designed and fabricated diode. (See Table III.)

4.1.6.5 Wirebond/interconnect defects. Diodes come in a variety of packages, some of which are similar to a transistor TO-5 or a microcircuit dual-in-line package. For these types, the wirebond failures will be similar to the transistor wirebond failures.

Gold-aluminum systems exhibit failures caused by excessive intermetallic formations or voids under the bond. Aluminum-aluminum systems exhibit failures in the wire adjacent to the bond area due to excessive pressure during bonding which causes the wire to crack.

Established, verified, and controlled bond schedules are a requirement for reliable bonds. The part manufacturer should have in-line control to assure that bond quality is being maintained.

TABLE II. Failure mechanism analysis mechanical defects

Failure Mechanism	Description/Cause	Detection Method	Method to Minimize or Eliminate Cause
A. Double slug construction			
Loss of hermetic seal	Fracture of glass case	Visual, dye penetrant	Handling techniques, mounting techniques, 100% visual inspection
	Seal fracture due to insufficient sealing area, low heat, short time, contamination	Visual	100% visual, better controls on preseal and seal operations
Poor solderability	Contamination of plating or base metal during plating	Hot oil dip of solder coated leads on sample basis, visual	Control of cleaning and plating operation, 100% visual
Intermittent	Variation of conductivity during temperature cycling, lack of control of assembly process	Monitor V_f during temperature cycling	Process Control, use monolithic double slug construction

4.1 DIODES, GENERAL

TABLE II. Failure mechanism analysis mechanical defects (Continued)

Failure Mechanism	Description/Cause	Detection Method	Method to Minimize or Eliminate Cause
B. Whisker (DO-7 type) construction			
Loss of hermetic seal	Fracture of the glass to seal	Hermetic seal test	Handling techniques, mounting techniques
Die lifted	Excessive voids in the eutectic bond	X-ray, thermal impedence test	Vacuum during eutectic bonding
Poor solder-ability	Contamination of plating or base material prior to or during plating	Hot oil dip of solder coated leads on a sample basis, visual inspection	Control of cleaning and plating operations
	Oxidation or contamination of the plated metal due to handling	Visual	Application of a protective coating that is compatible with normal fluxing and soldering operations
C. Whisker construction			
Loss of hermetic seal	Fracture of glass case	Visual, seal test	Handling techniques, mounting techniques
	Seal fracture due to insufficient sealing area, low heat, short time, contamination	Visual, seal test	100% visual, better controls on preseal and seal operations
Poor solder-ability	Contamination of plating or base metal during plating	Hot oil dip of solder coated leads on a sample basis	Control of cleaning and plating operation, 100% visual

TABLE II. Failure mechanism analysis mechanical defects (Continued)

Failure Mechanism	Description/Cause	Detection Method	Method to Minimize or Eliminate Cause
D. Voidless monolithic construction			
Loss of hermetic seal	Fracture of glass case	Visual, dye penetrant, electrical	Handling techniques, mounting techniques
External lead separation	Poor braze	Visual, pull test	Single element metallurgical bond, better control of brazing operation, vacuum brazing
Poor solderability	Contamination of plating or base metal prior to or during plating	Hot oil dip of solder coated leads on sample basis, visual inspection	Control of cleaning and plating operation, use of single element leads, 100% visual

TABLE III. Failure mechanism analysis bulk defects

Failure Mechanism	Description/Cause	Detection Method	Method to Minimize or Eliminate Cause
Metal precipitation	Gold precipitates during manufacturing operations such as glass sealing, resulting in soft breakdown characteristics	Power cycling with electrical test (breakdown voltage)	Not known, possibly related to dislocations in the silicon
Dislocations	Imperfections in the silicon that provide nucleation centers for metal precipitates that result in soft breakdown characteristics	Power cycling with an electrical test (breakdown voltage)	Monitor wafer processing using X-ray topography technique
Cracked die	Scribe crack propagates into the junction	Internal visual Electrical test	Ultrasonic or laser scribing

4.1 DIODES, GENERAL

Most of the S or C bend type of interconnect defects are easily detected by a visual inspection of the completed package prior to painting. These are the misplaced, excessive bend, or touching die type of defect. A more difficult type of defect to detect is insufficient pressure against the die contact or mechanical contact only.

Power stud diodes use a variety of anode/cathode interconnect techniques to effect low anode resistance ohmic contact. Some use an ultrasonic bond (aluminum-to-aluminum), some use an alloy type, and others use solder. The solder type of interconnect is desirable because of a thermal fatigue problem. The alloy or eutectic type has proven to be an acceptable method because it has been shown to be free from metal thermal fatigue problems. The ultrasonic bond is a reliable technique if properly established and controlled. Some parts may either have an alloy and an ultrasonic bond or ultrasonic bond and electrical weld.

The prime concerns should be that the process has been established, verified, is controlled, and uses materials that are metallurgically compatible as manufactured.

Some double-slug diodes depend on the glass case to provide the compression to maintain electrical contact between the die and the leads. This can result in parameter variation during temperature cycling. After a stabilization period, the parameters are within specification limits. The change occurs during the temperature change because of a difference in thermal coefficient of expansion between the glass and the heat sink materials. Also, a part can operate in an intermittent fashion, operating properly only when axial compression is applied. In these cases the glass case is not providing enough compression to hold the leads in contact with the die.

Some double-slug diodes use a construction technique that forms a monolithic structure of the heat sinks and the die. This is achieved by plating the die and the heat sinks with compatible materials and elevating them to a high temperature (900 °C) to form a bond. The bond interface is across the total die surface.

Methods of external lead attachment to the diode heat sinks vary. One method is to form the lead attachment directly to the heat sink using only the heat sink plating and the lead material to form a eutectic bond. This results in a bond that is as strong as the tensile strength of the wire. The bond is formed at about 700 °C and is relatively void free. Another method is to use a third metal as the eutectic. This method has a tendency to cause voids similar to those encountered in eutectic die mounting. A great many voids could result in a weak lead connection that could fail in shock or vibration during service.

4.1.6.6 Contamination defects. Contamination discussed in this paragraph refers to a material in the package, loose or attached, that is not part of the design. Examples are weld splatter, silicon chips, solder balls, and pieces of internal lead wire.

Sealing of metal packages by welding causes weld splatter. These particles of weld splatter may or may not be located so as to be detected during electrical testing. Those that are not may relocate during shock or vibration in usage and cause shorts.

Silicon chips can result from diamond scribing. Scribing may cause cracks in the silicon that will cause a chip of silicon to separate from the die. The propagation rate is dependent on time, temperature, and environment exposure. The chip could cause an electrical short or mechanical damage during shock or vibration in usage.

Some diodes are assembled using solder to attach the die to the header. During the soldering process, solder balls may be formed. This introduces a potential failure mechanism of the same type as weld splatter. The balls may break loose during environmental stress and cause shorts.

Diodes assembled with internal lead wires are exposed to potential failure mechanisms similar to transistors and integrated circuits. Wire particles and pieces may remain in the package to cause shorts or mechanical damage. These kinds of diodes lend themselves to internal visual inspections prior to capping, to screen out many of the defects.

4.1.6.7 Failure analysis. Although complete diode failure analysis is quite complex, preliminary analysis can prove very helpful in improving reliability.

If electrical testing shows the diode to be inoperative, the failure may be mechanical in nature, and the device should be X-rayed in an attempt to see the cause of failure before the diode is cut open. Opening a diode case should be the last operation in failure analysis because additional damage may be done which may mask the original cause of failure. After the diode is opened, the cause of mechanical failure will usually become apparent under microscopic examination.

If, when tested, the diode shows little or no deviation from the specification, it is well to observe its characteristics on a curve tracer where any irregularity in characteristic curves will be apparent. The diode should be tapped, while its characteristics are being observed, to detect any intermittent condition.

If the diode shows excessive leakage, the case should be thoroughly washed to remove any conductive paths that have formed externally. If the diode is from the low leakage series, it is necessary to assure that the body paint is still intact.

The investigation may be carried further by increasing and decreasing the diode temperature to the limits of the diode rating while observing the device characteristics on a curve tracer for irregularities.

4.1 DIODES, GENERAL

With the possible addition of a leak detection test, this is probably as far as failure analysis can be practically conducted outside of a semiconductor laboratory. Even for this preliminary analysis, thoroughly trained personnel and complete facilities are necessary.

4.1.6.8 Failure rate as a function of time. An idealized curve of component failure rate versus time is shown in Figure 7. Several features of this familiar "bathtub" curve are important in any consideration of diode reliability. The first portion of this curve indicates a high failure rate and then a steadily decreasing failure rate during the screening portion of diode life. This region is often referred to as "the infant mortality region." The portion of this curve which shows a decreasing failure rate for diodes has been demonstrated. These early life failures are generally classified as a result of poor workmanship.

The failure rate during the very early life of a diode depends upon a number of factors. Among these are the actual zero time in the life of a diode, the definition of failure, and the inherent reliability of the diode. Actually, the life of a diode begins when the encapsulating process is completed. On high-reliability diode lines a period of stressing at elevated temperature is often standard operating procedure to stabilize the diode's characteristics. The time and the stress applied during this stabilization process will affect the early life failure rate, and thus will significantly affect the shape of the curve.

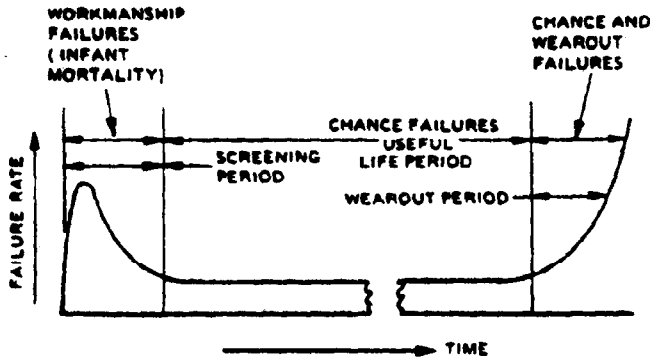


FIGURE 7. Failure rate as a function of time.

In any discussion of failure rate, the criteria used to define a failure will affect the failure rate for any given period of time. For example, a diode which has a certain amount of instability of characteristics early in life can exhibit different failure rates depending upon the relationship of initial limits and limits after a specified period of time. When tested to a life test specification, which defines a failure as a device exceeding the initial electrical parameter limits, these diodes will have a higher early life failure rate than they would have had if tested to a specification with life test limits relaxed from initial limits. If the diode parameters continue to drift

with time, even the relaxed life test limits would be exceeded and the total number of failures would be the same, regardless of the specified limits. However, if the diodes should stabilize after a short period of time, as is often the case, then the failure rate would be less to the relaxed life test points than to the tighter limits.

Figure 7 shows that after the initial high and decreasing failure rate period, which can be attributed to workmanship faults not detected during the manufacturing process, a period of relatively constant failure rate at a low level commences. This is the period of random failures.

The final portion of Figure 7 shows an increasing failure rate indicated as "wearout." This portion of the failure rate versus time curve is extremely difficult to define and will vary depending on diode method of fabrication and stress applied. This increasing failure rate can be introduced by such mechanisms as thermal fatigue of the solders between the silicon die and the mount due to repeated cycling of junction temperature while the case is at more or less a fixed temperature, by glass hermetic seal failure due to environmental cycling, by fatigue of internal construction due to mechanical stress, or by bulk defects. Little data is available from either diode life tests or system field tests to permit an accurate picture of this portion of the failure rate versus time curve. Contrary to the early life failure which may be characterized as workmanship faults, the failures which occur in the wearout period are believed to be a result of the basic design limitations of a diode.

The fact that failure rate is not constant with time throughout diode life dictates that any statement of failure rate must refer to the time period considered. Normally, failure rates are based upon the first 1,000 hours of diode life tests unless otherwise stated. This changing failure rate during diode life is a reason for not using MTBF as a measure of reliability on an individual diode basis.

4.1.6.9 Environmental considerations. All MIL-STD-975 diodes require vigorous screening and conformance testing in accordance with MIL-S-19500 and its applicable slash sheet. These tests are done to assure that the devices can withstand certain levels of environmental conditions such as pressure, vibration, moisture, temperature cycling, mechanical stress, and, if required, radiation. The use of these devices should take precautions so that these levels are not exceeded in the design of a system.

4.1.6.9.1 Electrostatic discharge. Because the damage caused by electrostatic discharges (ESD) is not always detectable, the user must take precaution when the user must take precaution when handling ESD sensitive devices. Refer to the Electrostatic Discharge Control Handbook, MIL-HDBK-263, for guidance.

4.1.6.10 Screening procedures. Because many early-life diode failures are the result of manufacturing flaws, it is possible to develop screening procedures to improve diode reliability. Actually, most diode manufacturers use screening procedures as a regular part of the diode fabrication process. The effectiveness of any screen procedure must be carefully verified for the particular type of diode under consideration.

4.1 DIODES, GENERAL

Screening tests are applied to uncover manufacturing defects and potential early-life failures. Depending on the manufacturing process used, screening tests such as power-applied burn-in, temperature cycling, and high temperature storage may be used.

All diodes are measured for significant electrical characteristics to detect devices with abnormalities, which may cause poor reliability. Most bulk and surface defects are detected at electrical characteristics screening.

Additional screening processes may be used to improve reliability, but unless properly selected, they may have the opposite effect of actually reducing diode life. For example, extreme mechanical stresses may not only destroy weak units, but may weaken good units.

Two conclusions, basic to diode reliability, may be drawn from the curve of failure rate versus time. These are as follows:

a. Relatively short term life tests (e.g., 1,000 hours) are sufficient to assure diode reliability for long time use.

b. Diode quality can be enhanced by the use of screening procedures to eliminate workmanship failures.

Item b above is best exemplified by the TXV process indicated in Figure 8 and the JANS process as found in Figure 9 in accordance with MIL-S-19500, General Specification for Semiconductor Devices.

The highest level of military product assurance is JANS (grade 1). The screens and process controls are of the latest technology. The product assurance level of JANS has been established for devices that are intended for space applications.

4.1.6.11 Power derating. The objective of power derating is to hold the worst-case junction temperature to a value lower than the normal permissible rating. The typical diode specification for thermal derating expresses the change in junction temperature with power for the worst case of the devices. The actual temperature rise per unit of power will be considerably less, but is not a value which can be readily determined for each unit. The user should refer to MIL-STD-975 for derating factor guidelines.

4.1.6.11.1 Junction temperature derating. Junction temperature derating requires that the ambient or case temperature for the part not be exceeded in the application.

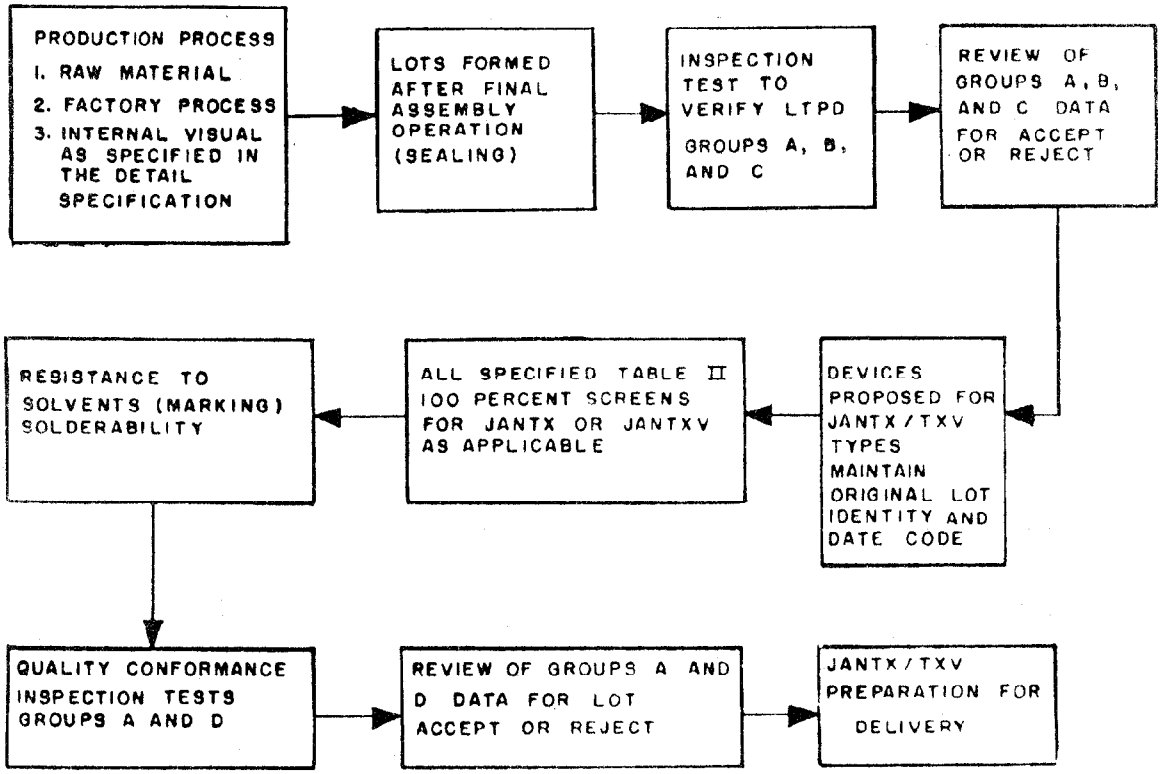


FIGURE 8. Order of procedure diagram for JANTXV devices.

The ambient temperature for a device that does not have some means of thermal connection to a mounting surface, includes the temperature rise due to the device, adjacent devices, and any heating effect which can be encountered in service. For space applications, the junction temperature should be limited to +125 °C maximum.

4.1.6.11.2 Thermal resistivity to air. The thermal resistivity to air is expressed in °C/W (or mW) or its reciprocal derating factor, which is usually expressed in mW/°C. These terms are based on the following assumptions:

a. The device is mounted in air at normal atmospheric pressures

b. Air is free to circulate by normal convection

c. The device is attached by its leads to an infinite heat sink at a controlled distance from the device

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.1 DIODES, GENERAL

d. The ability of the device to dissipate heat will be decreased by:

1. Longer leads
2. Restricted air circulation
3. Radiant heat reflection from adjacent areas.

e. The ability of the device to dissipate heat will be increased by:

1. Increase in air flow
2. Increase in air pressure
3. Decrease in the lead length
4. Surrounding the device with potting compounds having a thermal conductivity greater than air.

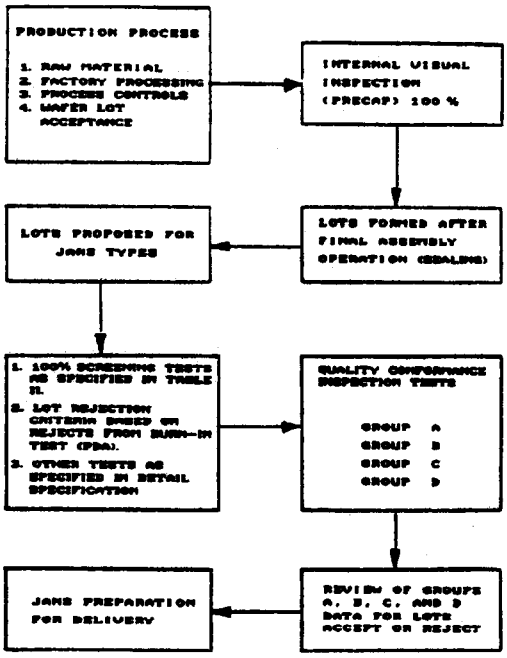


FIGURE 9. Order of procedure diagram for JANS devices.

4.1.6.11.3 Maximum allowable power for a device rated to air. This is established by controlling the thermal conductance of the medium surrounding the device under operational conditions so as to equal or exceed the normal conductance of air at one foot per second velocity.

The power may be determined from the following equations:

$$P_w(max) = \frac{T_{jo} - T_a}{\theta_{JA}}$$

where

$P_w(max)$ = Maximum allowable derated power
T_{jo} = Derated junction temperature from derating table
T_a = The temperature immediately surrounding the semiconductor in the most severe operating condition with all adjacent heat producing devices in operation
θ_{JA} = Manufacturer's thermal resistivity rating.

Where thermal conductance is expressed as a derating factor or 1/thermal resistivity, the above equation becomes:

$$P_w = (T_{jo} - T_a)\ (\text{Derating factor in air})$$

4.1.6.11.4 Maximum power for device rated to case temperature. This is established as follows:

a. Devices which have an established means for making thermal connection to a heat sink are rated for a case temperature instead of air temperature. The case temperature can be controlled by the derated power. The thermal resistivity, junction to case (θ_{JC}), or the thermal derating factor, is used to determine acceptable power for a maximum junction temperature when the case temperature is controlled.

b. Maximum power is determined from the case temperature of the device measured under the most severe operating conditions. The equation to be used is:

$$P_w(max) = \frac{T_{jo} - T_c}{\theta_{JC}}$$

where

θ_{JC} = Thermal resistance of the device rated to case
T_c = Measured case temperature
T_{jo} = Derated junction temperature.

4.1 DIODES, GENERAL

4.1.6.12 Voltage derating. The voltage rating of a device can vary with temperature, frequency, or bias condition. The rated voltage is the voltage compensated for all factors determined from the manufacturer's data sheet. The reliability derating consists of the application of a percentage figure to the voltage determined from all factors of the rating. For space applications voltages should be derated to 50 percent of the limits specified on the device data sheets. The user should refer to MIL-STD-975 for derating factor guidelines.

Instantaneous peak voltage derating. This derating is the most important and least understood derating and is required to protect the device against the high voltage transient spike of voltages which can occur as a result of magnetic energy stored in inductors, transformers, or relay coils. Transient spikes also can result from momentary unstable conditions which cause high amplitude oscillation during switching turn on or turn off.

Transient spike or oscillating conditions due to the discharge of leakage or static electricity will cause minute breakdown of surfaces within the body of the semiconductor. The minute breakdown may not cause failure but can cause a substantial increase in the probability of failure during service.

Circuit clamping by zener diodes, transient suppressors, or resistive voltage dividers, and current limiting by means of limiting resistors are required to maintain instantaneous transient voltages within the allowable limits.

Continuous peak voltage is the voltage at the peak of any signal or continuous condition which is a normal part of the design conditions. A continuous peak voltage is the highest voltage which can be observed on an oscilloscope under any normal operating condition.

Design maximum voltage is the highest average voltage. This is essentially the dc voltage as read by a dc meter. The ac signals can be superimposed on the dc voltage allowing a higher peak voltage providing the continuous peak voltage is not exceeded.

4.1.6.13 Selecting diodes. A number of factors must be considered in the choice of a diode type when reliability is of prime importance. These are the following:

a. The reliability of the device under consideration should be proven. New diode types with better electrical characteristics are constantly being announced. There is too often a tendency on the part of circuit designers to select these devices because of their high performance capabilities. It must be remembered that it takes time to adequately prove the reliability of a diode, and that generally the reliability of newer types of diodes has not been verified to the extent of older types.

b. The diode should have been in production long enough for any problems that may adversely affect reliability to have been eliminated. Early in the production phase of a diode type, major emphasis is often given to process improvement to optimize electrical characteristics. As the production process and yields improve, reliability will generally also improve.

c. The diode type under consideration should be a major portion of the manufacturer's yield. A characteristic of the semiconductor industry has been that a number of diode types of varying electrical characteristics are simultaneously produced on the same line. As manufacturing experience is gained, the process can be adjusted to optimize production of the most desired types. However, it is often true that a diode type which represents a small percentage of the yield of a production line may have some abnormality which will make its reliability different from the majority of the line output.

4.1.6.14 Circuit design considerations. The selection of the most reliable diode to perform the required function is basic, but is only the first step in assuring reliable circuit operation. Several circuit design considerations to assure reliable diode performance follow:

a. When possible, circuit performance should be based upon the most stable diode parameters.

b. Realistic limits for component variations due to tolerance, temperature, and time should be used. Wider limits must be allowed for characteristics which are less stable than those which show good stability with life.

c. Circuit design dependent upon diode characteristics that are uncontrolled can lead to poor reliability and should be avoided. If circuit performance is dependent upon unspecified diode characteristics and thus not controlled, there is no assurance that subsequent production will have consistent characteristics.

d. The use of derated operating conditions can be a factor to secure reliable circuit performance. The conditions to be derated and the amount of derating must be carefully determined to insure reliable circuit operation and still maintain required performance. Circuit performance and reliability, in this sense, compromise each other.

e. The environment which the diode, circuit, and system encounters during assembly, testing, and use must be controlled to assure maximum reliability.

4.1 DIODES, GENERAL

4.1.6.15 Radiation considerations. To ensure that a circuit functions properly in space applications, the design engineer must consider the effects of radiation exposure. Survivability requirements for systems that will encounter a radiation environment add to the challenge. The design engineer must know how radiation affects his circuit and components. When semiconductor devices are exposed to radiation environments, changes occur in their rated electrical parameters. The magnitude of the changes is a function of such things as the type of radiation, neutron or gamma rays, and time or duration. Generally, permanent damage is associated with displacement effects resulting from neutron radiation, and transient effects are the result of ionization from gamma ray radiation. The user should refer to MIL-STD-19500 for the four levels (M, D, R, and H) of radiation hardness assurance (RHA) and requirements. The letters M, D, R, and H following the S or TXV portion of the JAN prefix indicate the level of RHA. Parts without an RHA prefix either have not been tested for RHA or have not passed the testing.

Diodes exposed to radiation will display changes in the following electrical characteristics:

a. Breakdown voltage increases

b. Leakage current increases

c. Forward voltage increases

d. Reverse recovery time.

Each of these changes in the electrical characteristics of a diode is a result of radiation-induced changes in the semiconductor material itself. The increase in breakdown voltage can be an improvement, except for zeners, of the electrical characteristics over the value prior to radiation. This change is the result of an increase in the bulk resistivity of the base region of the diode. These increases in leakage current and forward voltage are undesirable electrical characteristic changes. When designing for the use of zener diodes, the design engineer must take into consideration the fact that the zener voltage will change. Zener diodes with a Vz > 10 V will have an increase in breakdown voltage as a function of neutron fluences, whereas those with a Vz < 10 V will have a decrease in breakdown voltage. These increases or decreases are caused by the two mechanisms involved in the breakdown. The low reverse breakdown voltage devices exhibit true zener internal field emission breakdown whereas the higher voltage devices break down as a result of the avalanche process. The amount of change or effect on a diode device is also dependent on the manufacturer's processes and the technology used. This varies from vendor to vendor and device to device. It is recommended that, for critical applications, the manufacturer be contacted and the specifics be discussed or resolved. Consideration should also be given to the type and amount of material shielding required for a given application.

4.2 Switching

4.2.1 Introduction. Switching diodes are normally made of silicon or germanium. The silicon diode has a higher forward voltage drop than the germanium diode; however, the silicon device has lower leakage currents and higher maximum operating junction temperatures. Silicon devices exhibit normal rectifying characteristics and perform the rectifying function very rapidly. They are classified according to the general groups below:

a. Ultrahigh speed switching types such as the 1N4148-1 with speeds of 1.0 to 50 ns

b. Medium speed types in the range of 50 to 150 ns (1N5417)

c. Standard switching types of 150 ns and up (1N645-1).

4.2.2 Usual applications. Switching diodes are useful in many types of computer circuits. They can perform a variety of functions in a particular circuit; specifically, clamping, saturation prevention, logic, and blocking. Care should be taken to make the forward current pulse as short in duration as possible. The magnitude of the forward current should be as small as the circuit application will allow. Some consideration also should be given to providing the maximum attainable reverse current pulse while holding the impedance of the external circuit to a minimum. Finally, the temperature at which the device operates should be as low as possible. Examples of some switching diode applications follow.

4.2.2.1 Clamping circuit. A clamping circuit is used to limit voltage pulses passing through the circuit. A simple clamping circuit is shown in Figure 10. When the magnitude of the input waveform is less than the threshold voltage, the diodes do not conduct. As the input signal becomes larger than the threshold voltage, the diodes begin to conduct. Diode X1 clamps the voltage during the positive cycle of the input, and diode X2 clamps during the negative cycle.

The resistance of the diodes and the series resistor form a voltage divider circuit which attenuates the signal. As the input signal and diode current increases, the diode resistance decreases, varying the voltage divider ratio to keep the output voltage near the threshold voltage. This variation in output voltage is called "soft limiting."

4.2.2.2 Logic circuit. Diodes may be used to implement Boolean logic functions. Of these, the OR (+) function is one of the simplest logic functions. In Figure 11 the function Y = A + B is implemented. The circuit will produce a positive output pulse across common-load resistor R3 if there is a positive pulse at any one of the inputs; that is if A or B is energized. The reason is that the input pulse forward biases the diode, and the resulting pulse current develops a pulse voltage drop across R3.

The resistance of input resistors R1 and R2 and output resistor R3 will depend on the diode characteristics and the characteristics of the driver and load circuits.

4.2 DIODES, SWITCHING

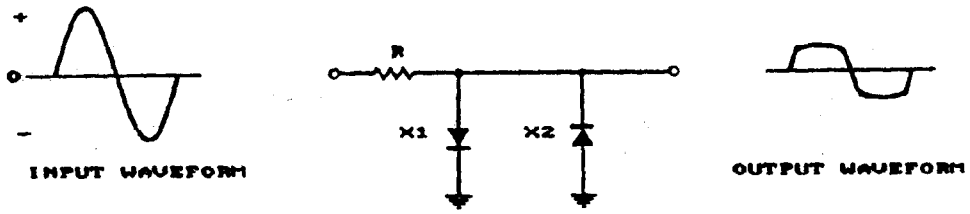


FIGURE 10. Clamping circuit.

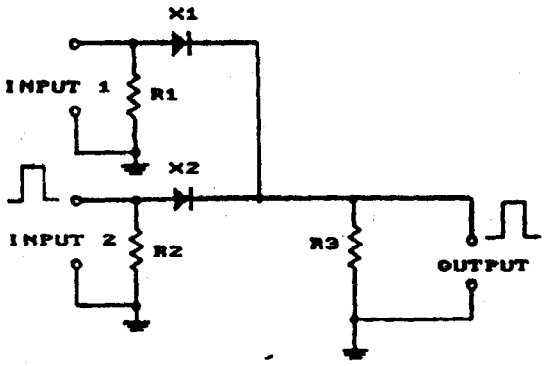


FIGURE 11. Logic circuit.

4.2.2.3 Saturation prevention circuit. When a transistor is used in a switching application, the transistor is driven from cutoff to saturation, the inherent storage time of the transistor limits the time required for it to turn off. This problem can be minimized by using a diode to keep the transistor in the operating region, thus preventing it from going into saturation. In Figure 12, the diode is used to reduce the turnoff time of the transistor. The transistor base voltage, 0.6 V for silicon transistors, is the anode voltage for the diode. If the diode is silicon, it will start to conduct when the voltage drop reaches about 0.35 V. Any additional reduction in collector voltage will turn the diode on harder and prevent the transistor from going into hard saturation.

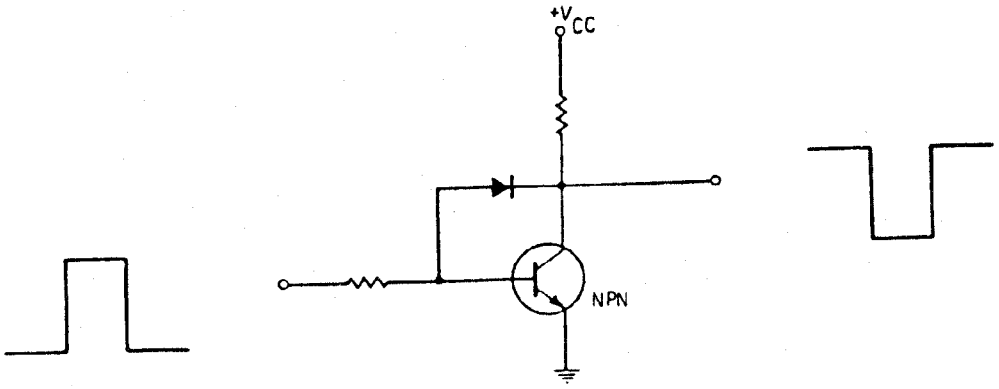


FIGURE 12. Saturation prevention circuit.

4.2.3 Physical construction. The axial leaded switching diode generally utilizes a passivated diffused or planar process die in a glass-to-metal or double-slug metallurgical bond construction. The two basic die processes are shown in Figure 13 and 14.

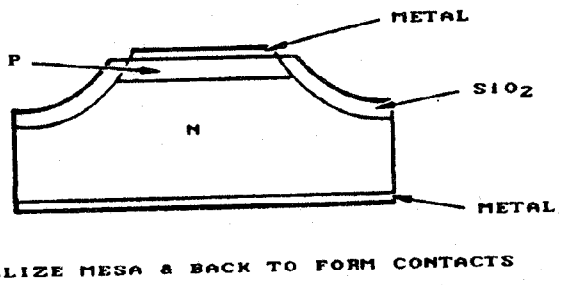


FIGURE 13. Passivated diffused process.

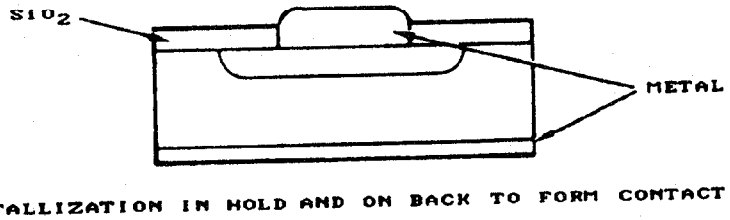


FIGURE 14. Planar process.

4.2.3.1 Back contact. Illustrated in Figure 15 below are three back contact techniques used in constructing axial leaded switching diodes. Although the alloyed back contact is the strongest, the high-temperature alloyed back (associated with the double-slug construction) is most commonly used.

4.2 DIODES, SWITCHING

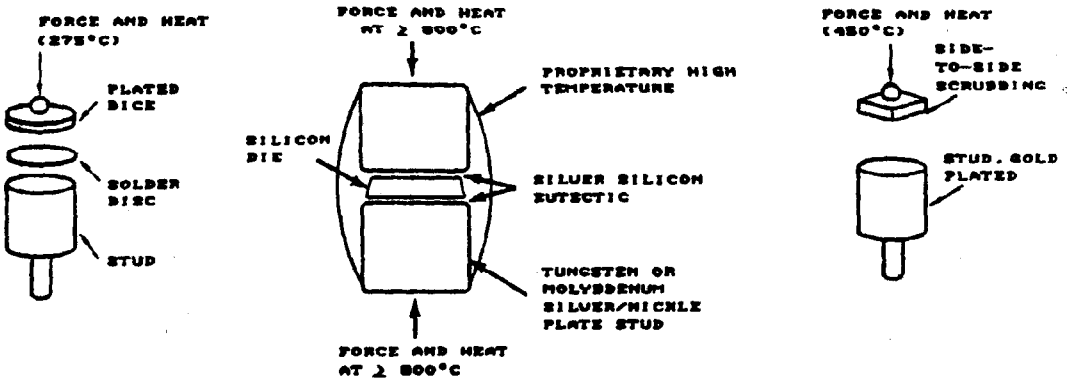


FIGURE 15. High-temperature alloyed back contact.

The solder contact has good mechanical strength and thermal conductivity. The silicon-silver alloy contact has good mechanical strength and thermal conductivity. The silicon chip is metallurgically brazed to plated molybdenum or tungsten slugs at temperatures greater than 700 °C. The silicon-gold alloy contact is the strongest and most thermally conductive of the above techniques. Because of the high melting temperature of the silicon-gold eutectic alloy (377 °C), the contact will withstand high temperatures during overload.

Molybdenum, tungsten, and dumet are usually used for studs in glass packages because their thermal expansion coefficients match those of certain glasses, allowing sealing of these glasses to them. Copper and copper alloys are usually used in devices which handle appreciable power because they have much better thermal conductivity than molybdenum and dumet.

4.2.3.2 Front contact. Electrical contact is made to the die in three ways as shown in Figure 16. Although three contact techniques are mentioned, the stud or lead contact associated with the double-slug construction is the most common and reliable.

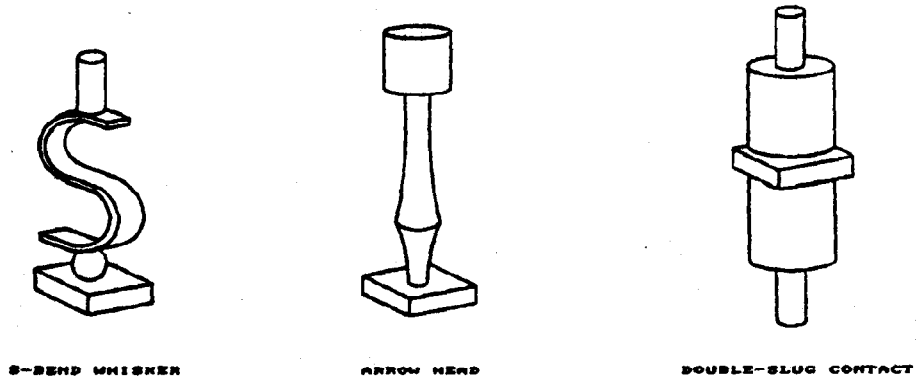


FIGURE 16. Various styles of electrical contact.

The whisker is held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker. The arrowhead contact is made using solder between the wire and the top of the die. The stud contact is usually soldered to the top of a diffused die, making good electrical and mechanical contact but not limiting the stress on the die.

The whisker contact has good shock and vibration tolerance. The arrowhead contact is somewhat superior due to its lower mass due to imbedding of the end of the contact in the metal on top of the die. The whisker structures provide good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion and contraction of the package.

The double-slug contact transmits more stress to the die than the whisker structure, but is used in spite of this because it provides good heat transfer, thus raising the power rating of the package.

4.2.3.3 Seals. Generally, two sealing techniques are used: metal-to-glass or double-slug construction. The package in Figure 17 uses glass-to-metal seals. These glass-to-metal seals are formed prior to final sealing and the final seal occurs by fusing glass-to-glass at the point noted in the figure.

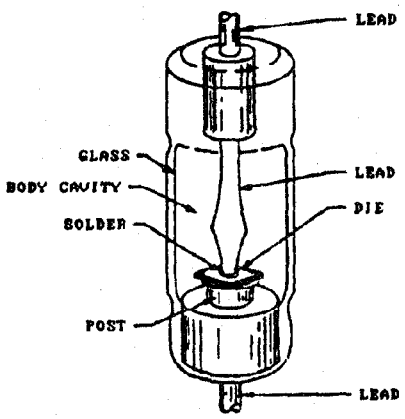


FIGURE 17. Glass-to-metal construction.

The double-slug package in Figure 18 is made by metallurgically bonding both sides of the die to either molybdenum or tungsten slugs to which a glass sleeve or glass slurry is fused for final sealing. The final step is a lead-brazing operation. This double-slug construction is most commonly used for producing axial leaded switching diodes for military applications.

4.2 DIODES, SWITCHING

Figure 19 shows a typical switching diode package.

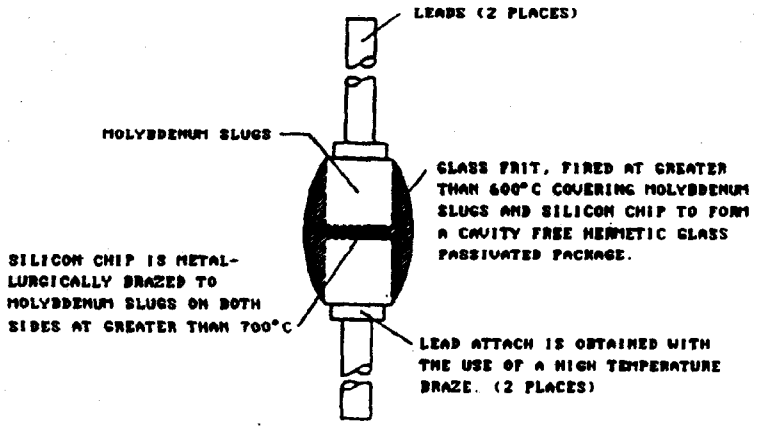


FIGURE 18. Typical double-slug construction.

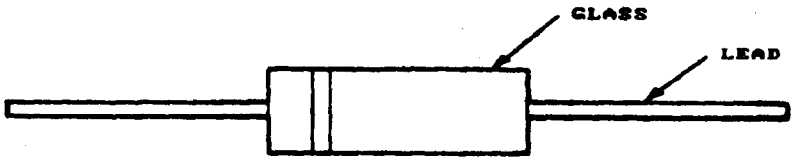


FIGURE 19. Outline drawing of a typical switching diode package.

4.2.4 Military designation. The military designation for diodes is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.2.5 Electrical characteristics. Table IV shows electrical ratings for various switching diodes. The values are for general reference only. Specific values should be found in the applicable MIL-S-19500 reference sheets.

TABLE IV. Typical electrical ratings for switching diodes

Device Type	V_R (V)	V_F (V)	I_F (mA)	Time trr maximum (ns)	19500 Slash Sheet
1N645-1	225	1.0	400	-	240
1N647-1	400	1.0	400	-	240
1N649-1	600	1.0	400	-	240
1N4148-1	75	1.0	10	5	116
1N5712	16	1.0	35	-	445

NOTE: This table is not to be used for part selection. Use MIL-STD-975 for that purpose.

4.2.5.1 Junction capacitance. In the p-n semiconductor switching diode, there are two capacitive effects to be considered. Both types of capacitance are present in the forward- and reverse-bias regions, but one so outweighs the other in each region that only that one is considered in each region. In the reverse-bias region there is the transition- or depletion-region capacitance (C_T), while in the forward-bias region there is the diffusion (C_D) or storage capacitance.

The basic equation for the capacitance of a parallel plate capacitor is defined by $C = \epsilon A/d$, where ϵ is the permittivity of the dielectric (insulator) between the plates of area A separated by a distance d. In the reverse-bias region there is a depletion region (free of carriers) that behaves essentially like an insulator between the layers of opposite charge. Because the depletion region will increase with increased reverse-bias potential, the resulting transition capacitance will decrease, as shown in Figure 20. The fact that the capacitance is dependent on the applied reverse-bias potential has application in a number of electronic systems.

4.2 DIODES, SWITCHING

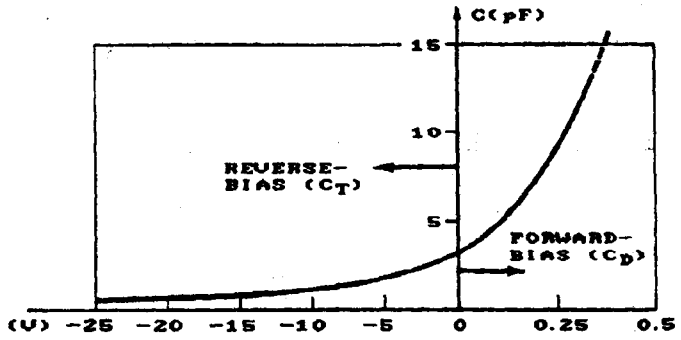


FIGURE 20. Transition and diffusion capacitance vs applied bias for a silicon diode.

Although the effect described above will also be present in the forward-bias region, it is overshadowed by a capacitance effect directly dependent on the rate at which charge is injected into the regions just outside the depletion region.

In other words, the capacitance effect is directly dependent on the resulting current of the diode. Increased levels of current will result in increased levels of diffusion capacitance. However, increased levels of current result in reduced levels of associated resistance, and the resulting time constant ($\tau = RC$), which is very important in high-speed applications, does not become excessive.

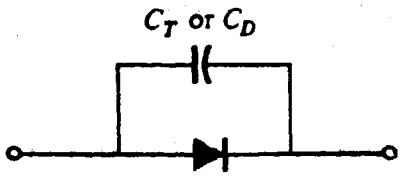


FIGURE 21. Including the effect of the transition or diffusion capacitance on the semiconductor diode.

The capacitive effects described above are represented by a capacitor in parallel with the ideal diode, as shown in Figure 21. For low- or midfrequency applications (except in the power area), however, the capacitor is normally not included in the diode symbol.

4.2.5.2 Reverse recovery time. In the forward-bias state, it has been shown in an earlier section that there are a large number of electrons from the n-type material progressing through the p-type material and a large number of holes in the n-type--a requirement for conduction. The electrons in the p-type and of holes progressing through the n-type material establish a large number of minority carriers in each material. If the applied voltage should be reversed to establish a reverse-bias situation, ideally the diode will change instantaneously from the conduction state to the nonconduction state. However, because of the large number of minority carriers in each material, the diode will simply reverse as shown in Figure 22.

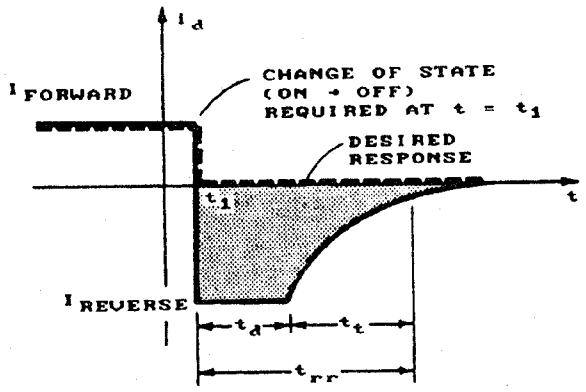


FIGURE 22. Defining the reverse recovery time.

Thus, the diode will stay at this measurable level for the period of time t_d (storage time) required for the minority carriers to return to their majority-carrier state in the opposite material. Eventually, when this storage phase has passed, the current will reduce in level to that associated with the non-conduction state. This second period of time is denoted by t_t (transition interval). The reverse recovery time is the sum of these two intervals: $t_{rr} = t_d + t_t$. Naturally, it is an important consideration in high-speed switching applications. Most industrially available switching diodes have a t_{rr} in the range of a few nanoseconds to 1 µs.

4.2.6 Environmental considerations. Typical environmental conditions and screening tests that switching diodes are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations, and 4.1.6.10 Screening procedures in subsection 4.1 Diodes, General. For the specific device selected consult the applicable MIL-S-19500 reference slash sheets.

4.2 DIODES, SWITCHING

4.2.7 Reliability considerations.

4.2.7.1 Failure modes. Predominant failure modes of the switching diode are found in the reverse-current and forward-voltage performance parameters. The failure mechanisms such as (1) mechanical stress, (2) thermal stress, and (3) operational life tests, account for the more significant incidences of failure, particularly the variable frequency exposure and longterm life tests. As noted previously, the TXV process with attendant V_F and I_R parameters should indicate any device that is beginning to show gradual deterioration. Other reliability considerations in device design are presented in the recommended approach to general diode construction, given in paragraph 4.1.6 Diodes, General.

4.2.7.2 Derating. History has shown that the largest single cause of diode failure is operation above allowable levels of electrical and thermal stress. Accordingly, it is imperative that derating of parts be performed to increase the reliability of electronic systems. Users should refer to MIL-STD-975 for derating factor guidelines. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress for a semiconductor. It will result in electrical parameter drift, and a general degradation of the electrical and mechanical characteristics of the device.

4.3 Rectifiers and power diodes.

4.3.1 Introduction. This section includes three general types of rectifiers: power diodes, fast switching power rectifiers, and power Schottky rectifiers. This general group does not include all available rectifiers but includes the most commonly used types.

4.3.1.1 Power diodes. The devices in this series are general purpose low-voltage rectifiers available in both axial-leaded and stud-mounted packages. Individual dc output voltages range from 50 to 1000 V. Maximum dc output currents are commonly supplied in values ranging from 1 to 6 amperes for axial-leaded packages and as high as 125 A for stud-mounted packages.

4.3.1.2 Fast switching power rectifiers. These rectifiers, with reverse voltages ranging from 50 to 1000 V and dc output currents of 1 to 300 A, have fast recovery times that are typically around 150 ns. Maximum dc output current ratings range from 1 to 6 amperes for axial-leaded packages and up to 300 A for studmounted packages and other high current packages.

4.3.1.3 Power Schottky rectifiers. These rectifiers are lower voltage devices with reverse voltages ranging from 15 V to 100 V and output currents up to 200 A. Schottky rectifiers come in axial-leaded, dual TO-3, and stud-mounted packages. Because they are majority carrier devices, their reverse recovery time is insignificant, which makes them excellent devices for high frequency applications.

4.3.2 Usual applications.

4.3.2.1 Power diodes. These devices are ideally suited for applications requiring rectification in the 1 to 125 amperes region, where low power loss and minimum physical size are important. In addition to breakdown voltage, and forward voltage, forward-surge current, and peak-reverse power, are of importance when selecting a power diode. A high-junction temperature and extreme low forward voltage drop and thermal impedance permit high current operation with minimum space requirements. Some types are available with negative polarity for use in bridge circuits or circuits that need a negative heat sink in half-wave and center-tap applications. Consideration should be given to absolute maximum ratings and derating characteristics.

4.3.2.2 Fast switching power rectifiers. This type of rectifier exhibits fast recovery time, typically below 150 ns, and is suitable for use in such circuit applications as inverters, choppers, low rf interference, free-wheeling rectifiers, and dc power supplies. In most of these applications, dissipation of heat is important, and therefore circuit design should take into consideration the specific device derating characteristics. See paragraph 4.3.7.

Consideration should also be given to the performance of the reverse recovery time, t_{rr}, at elevated temperatures because most of the existing MIL-S-19500 rectifier specifications specify t_{rr} at ambient temperature only.

4.3 DIODES, RECTIFIERS AND POWER DIODES

Evaluation of fast-switching power rectifiers on t_{rr} at high temperatures has shown a large variance in the percent change of t_{rr} from ambient to elevated temperatures among manufacturers. Because higher t_{rr} at elevated temperatures increases power dissipation losses and can possibly cause a timing problem in certain applications, the designer should evaluate the performance of t_{rr} in the circuit at the operating temperature of the application.

A typical reverse recovery time characteristic of a rectifier is shown in Figure 23.

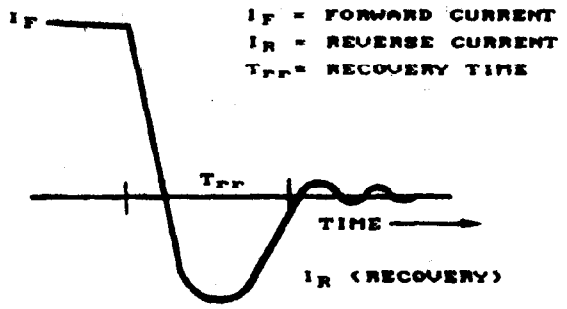


FIGURE 23. Reverse recovery time of a typical fast switching power rectifier.

4.3.2.3 Power Schottky rectifiers. The Schottky rectifier has a metal barrier type junction rather than the diffused pn-junction of standard power diodes. The most commonly used barrier metals or alloys are chromium, platinum, nickel platinum, molybdenum, and tungsten. The chromium barrier provides low-forward voltage with a very high-leakage current. However, the tungsten barrier provides low-leakage current with high-forward voltage. Because efficiency is a major consideration, the nickel platinum barrier provides the best choice due to its low-forward drop with a minimum of leakage current. The Schottky rectifier has two characteristics that make it useful as a rectifier in switching power supply circuits: (1) the low-forward voltage drop, 0.3 V, results in considerably less heat dissipation, compared with pn-junction rectifiers; therefore, it has a higher operating efficiency in the circuit (see Figure 24), and (2) the fast recovery time is also a significant advantage. The reverse recovery time of the power Schottky is generally less than 10 ns, compared with 150 ns for a standard fast recovery pn-junction diode.

The Schottky rectifier has some disadvantages which should also be considered. The most significant disadvantage is the reverse bias electrical characteristic. The peak reverse voltage is lower than that of the pn-junction diode, which limits its use to low voltage applications, and the reverse leakage current is higher than the pn types. The second disadvantage is that Schottky rectifiers have a lower operating temperature.

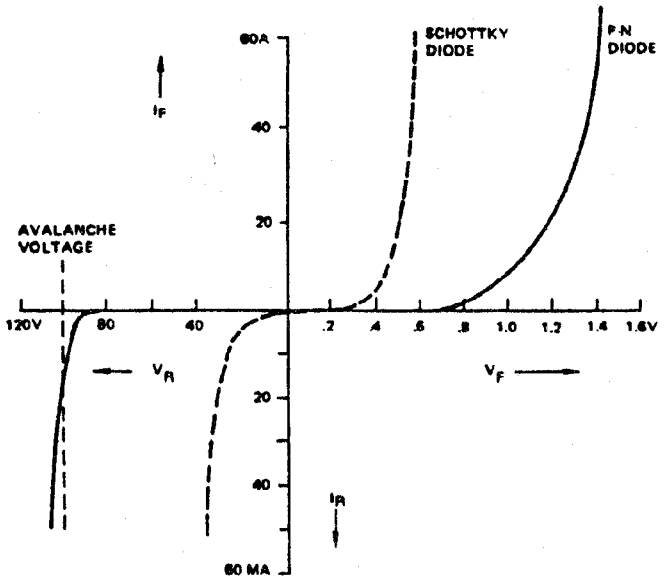


FIGURE 24. PN junction and Schottky rectifier I/V characteristics.

4.3.3 Physical construction. The rectifier and power rectifier are generally made using a passivated diffused or planar process whereas the power Schottky rectifier is made with a metal barrier Schottky die. Rectifiers commonly are of glass-to-metal or double-slug construction whereas high-power rectifiers and high-power Schottkies use a stud-mounted package. The three basic die processes are shown in Figures 25, 26, and 27.

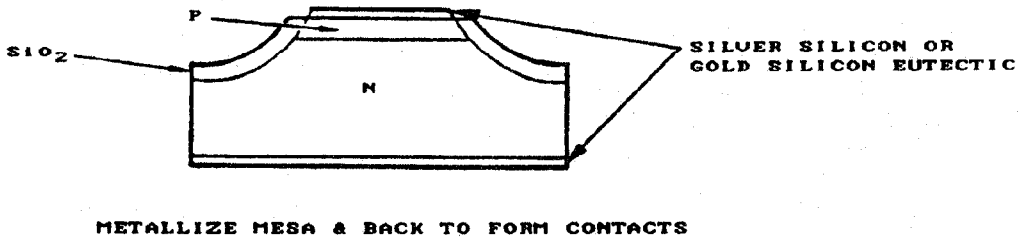


FIGURE 25. Passivated diffused process.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.3 DIODES, RECTIFIERS AND POWER DIODES

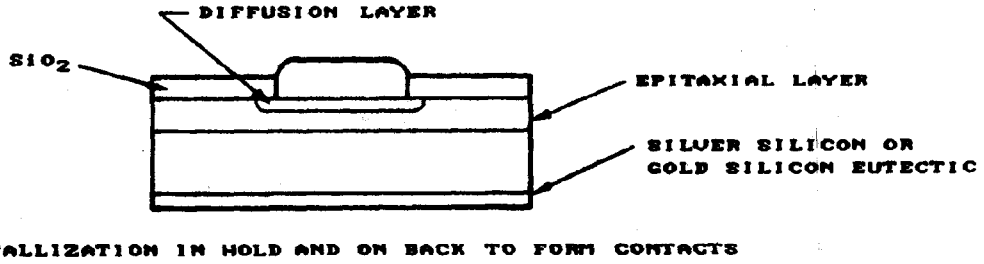


FIGURE 26. Planar process.

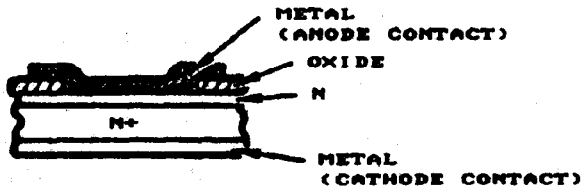
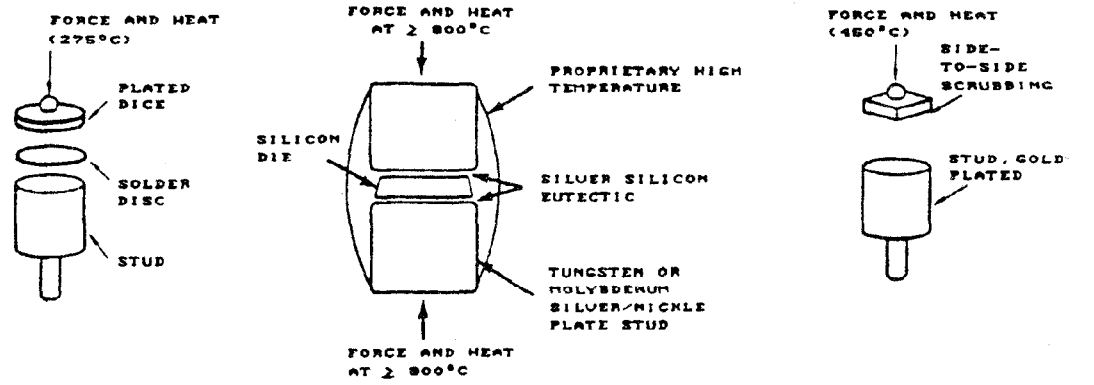


FIGURE 27. Metal barrier Schottky junction.

4.3.3.1 Back contacts. Illustrated in Figure 28 below are three back contact techniques used in constructing rectifier, power, and Schottky diodes. Although the alloyed back contact is the strongest contact, rectifier and power diodes generally made with high temperature alloy and solder back contacts, respectively, for high reliability applications.

The solder contact has good mechanical strength and thermal conductivity. The silicon-silver alloy contact, which is generally associated with a high temperature alloy, has good mechanical strength and thermal conductivity. The silicon chip is metallurgically brazed to plated molybdenum or tungsten slugs at temperatures exceeding 700 °C. The silicon-gold alloy contact is the strongest and most thermally conductive of the above techniques. Because of the high melting temperature of the silicon-gold eutectic alloy (377 °C), the contact will withstand high temperatures during overload.

Molybdenum, tungsten, and dumet are usually used for studs in glass packages because their thermal expansion coefficients match those of certain glasses for sealing. Copper and copper alloys are usually used in devices that handle appreciable power because they have much better thermal conductivity than molybdenum, tungsten, or dumet.



Solder back contact | High temperature alloyed back contact | Alloyed back contact

FIGURE 28. Back contact formation techniques.

4.3.3.2 Front contacts. Electrical contact is made to the die in three ways as shown in Figure 29. Although three contact techniques are mentioned, the stud construction is the most common and reliable.

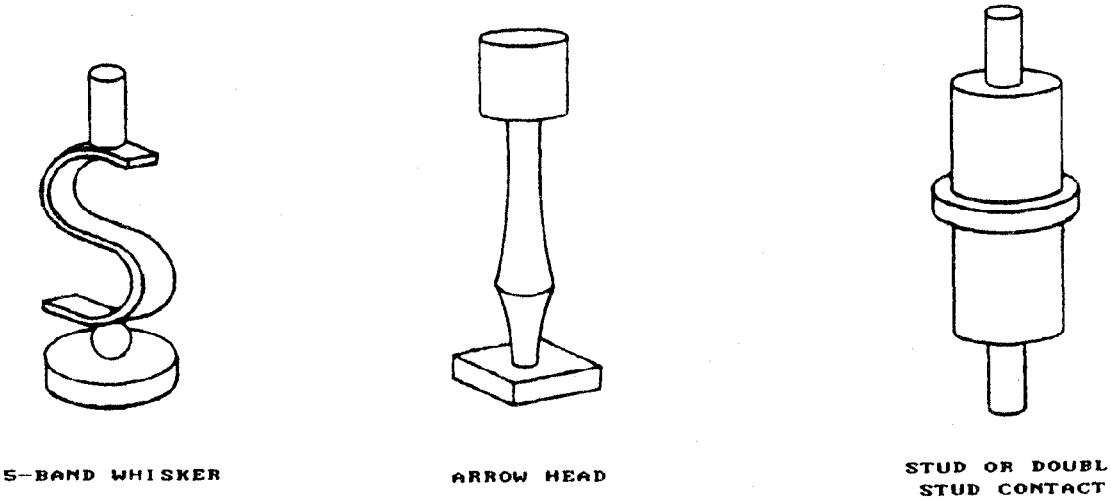


FIGURE 29. Three various styles of electrical contact.

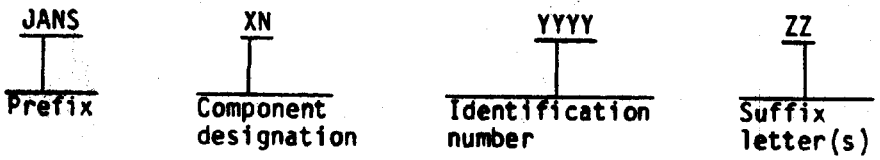
The whisker is held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker. The arrowhead contact is made using solder between the end of the wire and the top of the die. The stud contact is usually soldered to the top of a diffused die which makes good electrical and mechanical contact but does not limit the stress on the die.

The whisker contact has good shock and vibration tolerance. The arrowhead contact is somewhat superior due to the imbedding of the end of the contact in the metal on top of the die which prevents sliding of the contact. Both whisker structures provide good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion and contraction of the package.

The stud and double-slug contact transmits more stress to the die than the whisker structure but is used in spite of this because it provides good heat transfer, thus raising the power rating of the package.

4.3.3.3 Seals. Generally, three sealing techniques are used: metal-to-glass, double-slug, and stud mounted. The package (A) in Figure 30 uses glass-to-metal seals. These glass-to-metal seals are formed prior to final sealing. The final seal occurs by fusing glass-to-glass at the point noted in the figure. The double-slug package (B) in Figure 30 is made by metallurgically bonding both sides of the die to either molybdenum or tungsten slugs to which a glass sleeve or glass slurry is fused for final sealing. The final step is a lead-brazing operation. This type of construction is most often associated with rectifier diodes. The stud-mounted package (C) assembly sequence illustrated in Figure 30 has solder-back and band-front contacts. This construction is most often associated with power rectifiers and power Schottky rectifiers. Figures 31 32 show a typical double-slug rectifier and power diode package, respectively.

4.3.4 Military designation. Military designation for diodes is formulated as follows:

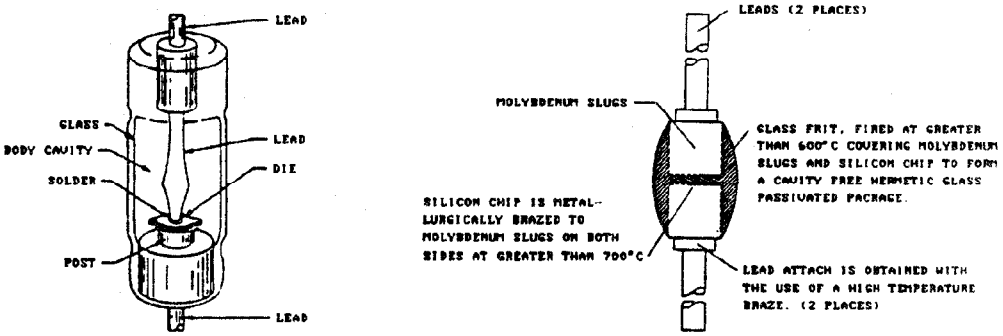


The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

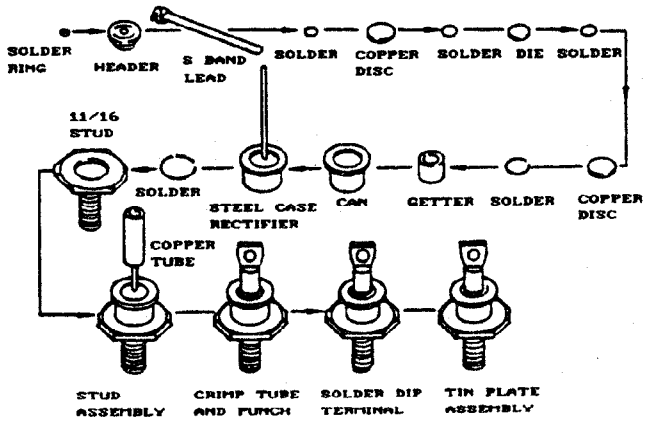
The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.



A. Glass-to-metal

B. Double-slug



C. Stud

FIGURE 30. Package sealing techniques.

4.3 DIODES, RECTIFIERS AND POWER DIODES

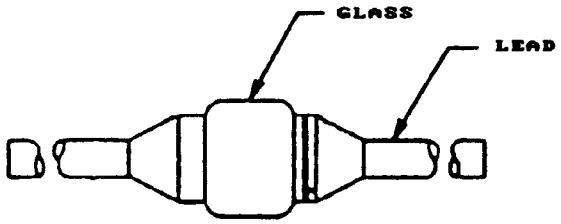


FIGURE 31. Outer drawing of a typical double-slug rectifier.

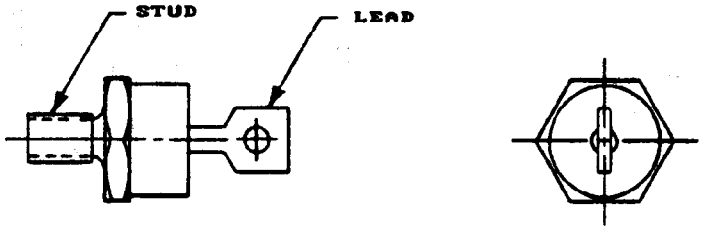


FIGURE 32. Outer drawing of a typical power diode.

4.3.5 Electrical characteristics.

4.3.5.1 Power diodes and switching rectifiers. Table V illustrates the maximum ratings of typical power diodes and switching rectifiers. Average output (rectified) current is referred to as I_o and surge current refers to the maximum current pulse that can be carried by the device for the length of time, repetition frequency, and waveform specified. Basic characteristics and parameter information on diodes and rectifiers are discussed in the "Diodes, General" section.

TABLE V. Power diode and switching rectifier maximum ratings

Device Type	V_R	Io T_A = 55 °C	If (Surge) 1/120 Sec.	MIL-S-19500 Slash Sheet
1N5415	50V	3A	80A	411
1N5416	100V	3A	80A	411
1N5417	200V	3A	80A	411
1N5418	400V	3A	80A	411
1N5419	500V	3A	80A	411
1N5420	600V	3A	80A	411
1N5615	200V	1A	25A	429
1N5617	400V	1A	25A	429
1N5619	600V	1A	25A	429
1N5621	800V	1A	25A	429
1N5623	1000V	1A	25A	429

NOTE: This table is not to be used for part selection, use MIL-STD-975.

The operating characteristics shown above are for general reference only. For any given type of device, refer to applicable MIL-S-19500 slash sheet.

4.3.5.2 Power Schottky rectifiers. For electrical characteristics on Schottky rectifiers refer to applicable MIL-S-19500 slash sheet.

4.3.6 Environmental considerations. Typical environmental conditions and screens which rectifier diodes are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 Screening procedures in subsection 4.1, Diodes, general. For the specific device selected, consult the applicable MIL-S-19500 reference slash sheet.

4.3.7 Reliability considerations.

4.3.7.1 Power diodes and fast switching rectifiers. Fast switching power rectifiers, as with other silicon diodes of this type, may exhibit localized mechanical weakness because of a large junction area, due to either device die and material design including ohmic contact metallurgy and package design, or defects and flaws that result from the fabrication process itself.

The former type of weakness would include the use of lead-rich solder for either or both anode and cathode solder ohmic contact regions. The resulting recrystallization (plastic deformation) of these regions results in progressively higher $R_{\theta JC}$ (thermal impedance) values, leading to higher junction temperatures that will permanently degrade and eventually destroy the device. The use of the standard gold-silicon eutectic alloy or the gold-germanium eutectic alloy will eliminate this failure phenomenon. If a lead-rich solder is used for a device contact, a power cycling test under applicable system conditions should be used to determine the thermal cycle capability of the device.

The purpose of high temperature reverse bias and power burn-in tests, as specified in the JANTXV or JANS specifications, is to eliminate defective devices caused by surface ionic contamination of the primary passivation, large shifts in the reverse leakage current, microcracks extending into the junction area, and poor metallurgical bonding in the die attach area.

4.3.7.2 Power Schottky rectifier. Although power Schottky diodes and rectifiers are extremely reliable when used properly in the application, catastrophic failures can occur if incorrect test methods are used.

A mistake commonly made is to measure reverse leakage current by applying reverse voltage until a specified current is reached. This can result in inadvertently applying excess reverse voltage to the rectifier which causes dielectric failure of the barrier.

The test must be made by applying the specified reverse voltage and measuring the resulting reverse current. The rectifier must not be stressed beyond the peak reverse voltage specified on the applicable electrical specification.

4.3 DIODES, RECTIFIERS AND POWER DIODES

Another mistake which causes an increase in high-leakage currents is to measure the leakage at an elevated ambient temperature by means of a dc reverse bias.

Due to its high-leakage current, the diode is actually heated during the test which causes the leakage current to increase still further. This can result in a thermal runaway condition and destruction of the diode. The solution is to use a reverse voltage pulsed at a low-duty cycle, about 1 percent, to minimize the power dissipation.

Schottky barrier diodes and rectifiers also have a dv/dt failure mode. If the test fixture is allowed to deliver the specified reverse voltage at a low-source impedance, some diodes may fail when inserted one after the other into the test socket. These failures occur because the reverse voltage application rate exceeds the limit of 700 V/μsec. However, by using a 1000-ohm series resistor, the junction capacitance of the diode under test is sufficient to limit the rate of applied voltage to about 20 V/μsec.

The question of whether to use insulating hardware or to mount the Schottky devices on an electrically isolated heat sink frequently arises. In terms of overall efficiency, the isolated rectifier with a grounded heat sink is better. However, if excessive junction temperature is a problem, due to high-ambient conditions or inadequate heat sinking, the isolated heat sink is the best approach.

4.3.7.3 Derating. History has shown that the largest single cause of diode failure is operation above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be performed to increase the reliability of systems. Users should refer to MIL-STD-975 for derating factor guidelines. In general, derating for voltage, current, and surge current, should be 50 percent of the value published on the vendor data sheet. The junction temperature should be limited to +125 °C maximum. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress a semiconductor device could be subjected to. It will result in electrical parameter drift, and a general degradation of the electrical and mechanical characteristics of the device.

4.4 Voltage regulator.

4.4.1 Introduction. Technological advances in the manufacture of silicon junction diodes have made them extremely valuable components for use in accurate voltage regulator reference designs. The silicon junction diode is a semiconductor device possessing a very high reverse bias resistance up to its critical reverse breakdown, or zener, voltage. At this point, the back resistance drops to a very small value. In this region, the current will increase very rapidly, while the voltage drop across the diode remains almost constant. Figure 33 illustrates that over a wide range of current, an essentially constant voltage will be maintained. Zener diodes, therefore, when biased in the reverse direction, can be used as a voltage regulator, or reference element.

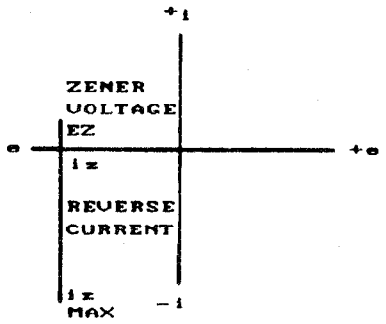


FIGURE 33. Idealized reverse breakdown characteristic.

The silicon diode regulator possesses definite advantages over other types of reference elements. It has a longer life expectancy because of mechanical ruggedness and does not suffer from deterioration under storage. There is essentially no aging during its operating life, unlike other regulating devices. Small size and light weight make its use especially useful in airborne or portable equipment. Moreover, the silicon regulator or combination thereof, can be supplied in values at any desired voltage and can operate over a wide range of current whereas other regulators are restricted to specific voltages and very limited current ranges.

4.4.2 Usual applications. Beyond breakdown, the characteristics of the silicon zener diode are almost identical to the gas voltage-regulator tube and may be considered to be the equivalent to the semiconductor. The zener diode can be used in exactly the same manner to provide a constant voltage output.

The simple shunt regulator, Figure 34, is a circuit in which a shunt element draws variable current through a resistor that is also in series with the load.

4.4 DIODES, VOLTAGE REGULATOR

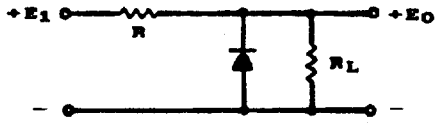


FIGURE 34. Shunt voltage regulator.

The current through resistor R is dependent upon the load requirements. As the load increases or decreases, the zener shunt element will draw more or less current. The net result is a constant output voltage across R_L.

4.4.2.1 Thermally induced resistance. Resistor R should be selected so that the current in the diode does not exceed I_Z max, or exceed the maximum power dissipation rating if the load R_L were removed. In practice, I_Z is chosen as approximately 20 percent of I_Z max, and as a shunt regulator, it will absorb current variations between the I_Z and I_Z max. limits. Referring to Figure 34, it can be seen that as I_Z increases, dynamic resistance decreases.

The greater the permissible current through the shunt diode, the larger the ripple reduction will be and the better the voltage regulation. However, as the diode current is increased, the junction temperature will rise to the point where the dynamic resistance will increase due to a thermally induced resistance in the element.

This thermally induced resistance will tend to limit regulation at high currents, whereas the dynamic resistance will dominate at lower currents. Therefore, when using a diode, a compromise operating point must be selected. The alternative would be to use a regulator of higher power dissipation capabilities possessing lower thermal resistance.

The maximum permissible diode current is limited by the temperature rise of the junction and by the heat sink provided to dissipate the heat. Thus, the use of a zener diode as a voltage regulator element is limited only by its rated current handling capabilities. Zener regulators are available over a wide range of power dissipations.

4.4.2.2 Multiple junctions. In many instances, the circuit design engineer requires regulation at higher voltages. Three alternatives are available. A single junction unit rated for the desired voltage can be used but it suffers from a high positive temperature coefficient and high dynamic resistance.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.4.2.2.1 Series arrangements. A second alternative to achieve optimum performance is to use a number of lower voltage units in series. The resultant temperature coefficient, dynamic resistance, and thermal resistance will be much reduced for the series combination.

An example of how a high voltage multiple junction assembly compares with a single high voltage unit can be shown when six 5-V zeners are connected in series. A single 30-V zener diode would possess a temperature coefficient of close to +0.1%/°C and a dynamic resistance of 60 Ω, and might handle only 30 mA for a particular style diode. The series combination, on the other hand, would have essentially zero coefficient, resistance of only 6 Ω, and the ability to handle over 200 mA of reverse current.

Regulation of a high voltage multiple junction assembly will be superior to a single high voltage area in all respects at any given current. It is also possible to obtain a much closer tolerance by such a series combination. If similar units are used, the series regulator assembly will also have higher current ratings as a result of its increased heat dissipation capability. The higher wattage rating is directly proportional to the number of similar series units used.

4.4.2.2.2 Packaged series assemblies. The third alternative is the use of packaged series assemblies. Here, six selected zener diodes are assembled in series to provide voltage regulators from 24 through 160 V. The total dissipation of the package is 5 W without heat sink and it allows the designer to place the assembly in a convenient location on the chassis without having to provide for mounting of individual diodes.

The power dissipation per diode in such a series combination will be inversely proportional to the number of units used; in this case, six. For a given current, there will be a significant lowering in the operating junction temperatures.

As the junction temperature is lowered, life expectancy of each diode will increase.

It is possible, therefore, to achieve an increase in overall reliability despite the fact that more individual components are involved. In general, multiple junction operation is preferred in all ways.

4.4.2.3 Double anode zener diodes. In the 6- to 8-V region, it is possible, to compensate for the reverse positive temperature coefficient by taking advantage of the negative coefficient of another diode operating in the forward direction. Two such diodes, when connected in series with reversed polarities make possible a stability of ±1 percent or better over the temperature range of -55 to +100 °C. Optimum compensation and stability will occur at currents of approximately 10 mA.

4.4 DIODES, VOLTAGE REGULATOR

These double anode zener diodes can also be used to provide a means of simple ac regulation because the diodes are constructed in a symmetrical manner; the reverse voltages of both diodes are matched to within ±5 percent.

Typical applications for such devices include: calibration source for oscilloscopes, ac limiting in servo control systems, and spike clipping. The basic ac regulator circuit using double anode zener diodes is shown in Figure 35. The idealized characteristics of a double anode-type zener diode is shown in Figure 36.

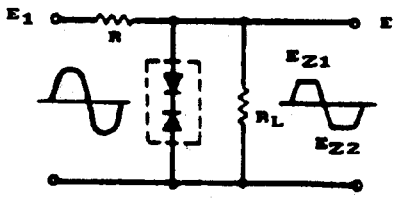


FIGURE 35. Basic ac regulator.

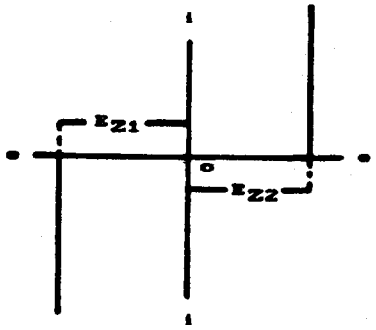


FIGURE 36. Idealized characteristics of double anode type diode.

4.4.2.4 Zener diode clamp for inductive spikes in MOSFETS. For inductive loads, as is the case in switching regulators, overvoltage transients can be produced when the device is switched off, even though the dc supply voltage for the drain circuit is well below the V_{DS} rating of the MOSFET. The faster the MOSFET is switched, the higher the overvoltage will be. If the device cannot absorb this extra inductive energy produced by the overvoltage spikes ($1/2\ LI^2$ or VIt) a catastrophic failure (drain-source short) may occur.

Inductance is always present to some extent in a practical circuit, and therefore, there is always danger of inducing overvoltage transients when switching off. Usually the main inductive component of the load will be clamped. Stray circuit inductance still exists, however, and overvoltage transients will still be produced as a result.

The first approach to this problem is to minimize stray circuit inductance, by means of careful attention to circuit layout, to the point that whatever residual inductance is left in the circuit can be tolerated. However, if this cannot be entirely accomplished, and for added safety against spikes by the remaining stray circuit inductance, a zener diode should be connected physically as close as possible to the drain and source terminals, as shown in Figure 37.

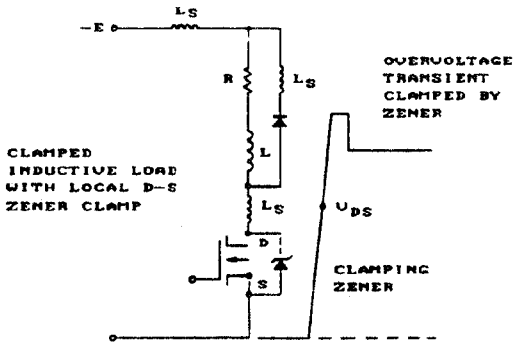


FIGURE 37. Zener voltage regulator.

4.4.2.5 Bias method for a relay amplifier. Figure 38 illustrates a method of bias for a relay amplifier that can be held in the off condition until a predetermined input level has been reached. The zener diode provides a bias for the amplifier that is close to cut-off, and current through the JFET will be insufficient to energize the relay. When the gate is made more positive, it can be seen that the bias will remain constant even though the drain current increases.

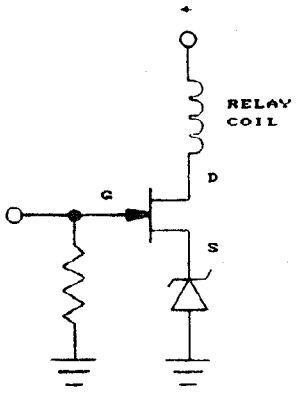


FIGURE 38. Relay amplifier bias supply.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.4.2.6 Zener diodes for fixed biasing purposes. The zener diode may be considered as an equivalent constant voltage source when used for biasing purposes. Figure 39 illustrates how a form of fixed bias for two JFETs in a single-ended push-pull servo amplifier may be provided. The result is a simple, highly efficient circuit with a minimum of components.

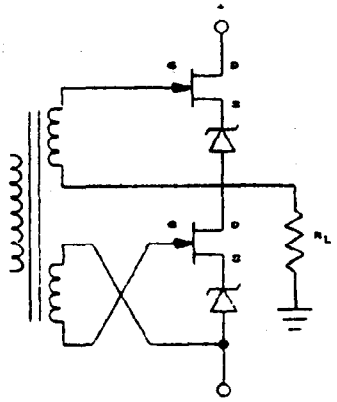


FIGURE 39. Zener diode as fixed bias.

4.4.2.7 Zener diodes as coupling devices. A silicon zener diode may be used as a coupling device between two amplifier stages in much the same manner as a capacitor; it is illustrated in Figure 40.

The dc level will be reduced only by an amount equal to the zener voltage drop. The zener element permits frequency response down to dc due to its extremely low resistance. Although the diode may be replaced by a resistor to achieve the required level change, there could be a loss in signal due to the voltage-divider action of the resistors.

4.4.2.8 Reference element application. Fortunately, the zener diode is basically a low voltage device. This fact makes it particularly attractive in the design of transistorized equipment. Of particular interest is its application as the reference element in a series-regulated power supply as shown in Figure 41.

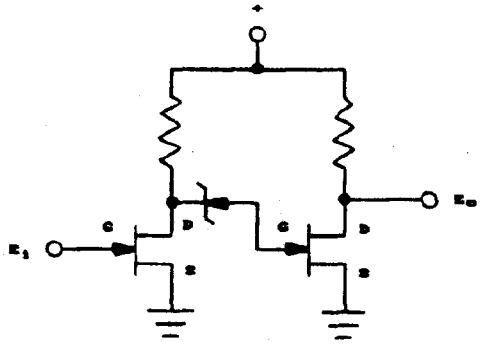


FIGURE 40. Zener diodes as dc coupling devices.

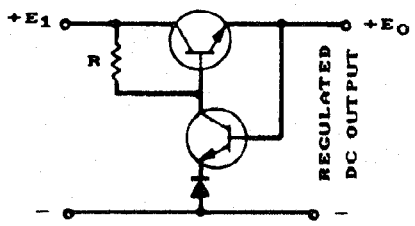


FIGURE 41. Zener reference element in transistor power supply.

4.4.2.9 Temperature sensing devices. The apparent disadvantage of a zener diode's temperature coefficient may be put to a useful purpose in the form of a temperature-sensing device. A bridge composed of two resistors and two similar diodes, Figure 42, can be constructed so as to indicate a temperature level when one of the diodes is held at a reference temperature and the other is subjected to the varying environment. A typical 10-V zener diode has a temperature coefficient of +0.07%/°C which corresponds to 7 mV/°C change. The sensing element will, therefore, indicate an imbalance of 0.7 V when undergoing a 100 °C temperature change. The output can be read directly or fed to a recording galvanometer for permanent records.

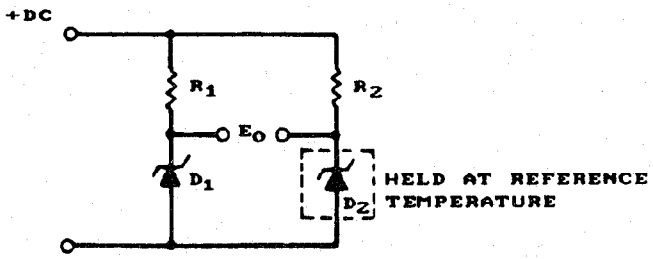


FIGURE 42. Temperature-sensitive bridge.

4.4.2.10 Selective signaling circuit. A precision selective signaling circuit can be made to operate a series of relays in a sequential manner corresponding to the value of applied voltage. In Figure 43, as the input voltage increases, the zener diode with the lower voltage will turn on first and will remain on, thereby energizing the appropriate relay coil. Subsequently, as the input voltage reaches the voltage level of each individual zener diode they too will energize the appropriate relay.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.4 DIODES, VOLTAGE REGULATOR

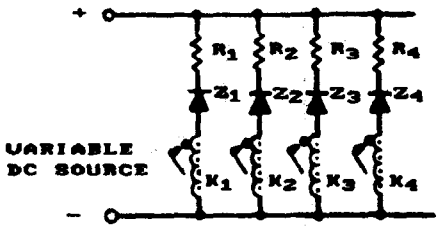


FIGURE 43. Selective relay signaling circuit.

The principal advantage over other means is that the pull-in of each relay is virtually independent of coil characteristics and is dependent only on the sharp reverse characteristic of the zener diode. The exact voltage value for relay operation can be set by preselection of the proper zener diode, rather than the often impossible task of choosing relays of dissimilar characteristics.

4.4.3 Physical construction. The voltage regulator diode generally uses a passivated diffused, planar or diffused planar junction or epitaxial substrate process die in a glass-to-metal or double-slug package. The three basic die processes are shown in Figures 44, 45, and 46.

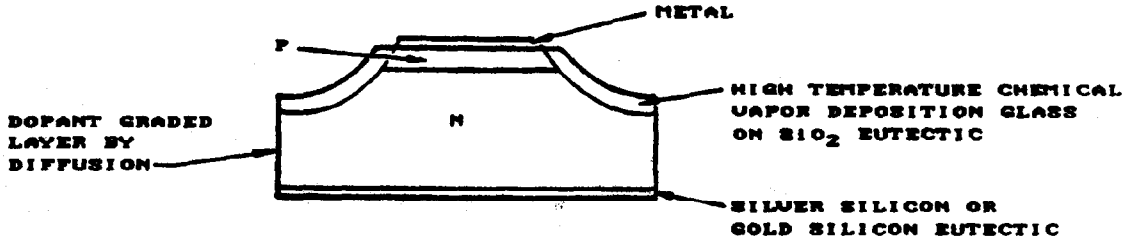


FIGURE 44. Passivated diffused process.

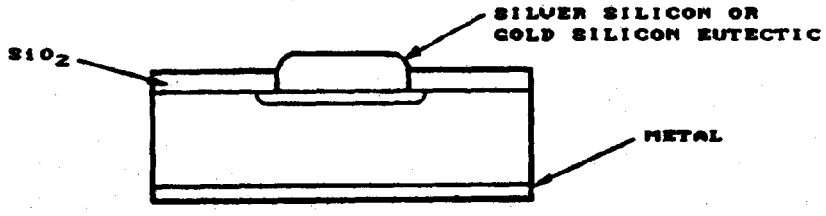


FIGURE 45. Planar process.

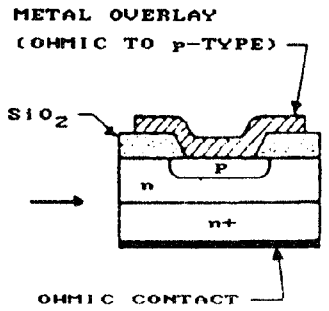


FIGURE 46. Diffused planar junction on epitaxial substrate.

4.4.3.1 Back contacts. Illustrated in Figure 47 are three back contact techniques used to construct voltage regulator diodes. Although the alloyed back contact is the strongest contact, voltage regulator diodes generally use high temperature alloy back contacts, associated with double-slug construction, for high reliability applications.

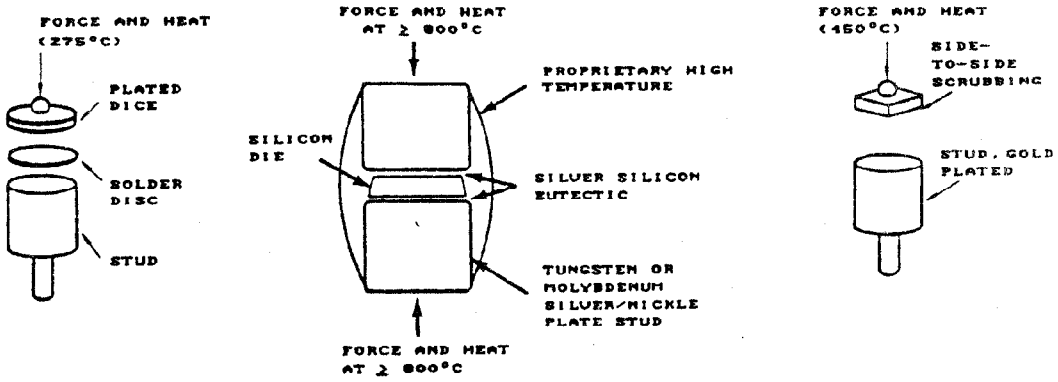


FIGURE 47. Back contact formation techniques.

The solder contact has good mechanical strength and thermal conductivity. The silicon-silver alloy contact, which is generally associated with the high temperature alloy, has good mechanical strength and thermal conductivity. The silicon chip is metallurgically brazed to plated molybdenum or tungsten slugs at temperatures exceeding 700 °C.

4.4 DIODES, VOLTAGE REGULATOR

The silicon-gold alloy contact is the strongest and most thermally conductive of the above techniques. Because of the high melting temperature of the silicon-gold eutectic alloy (377 °C), the contact will withstand high temperatures during overload.

Molybdenum, tungsten, and dumet are usually used for studs in glass packages because their thermal expansion coefficients match those of certain glasses, allowing sealing of these glasses to them. Copper and copper alloys are usually used in devices which handle appreciable power because they have much better thermal conductivity than molybdenum, tungsten, and dumet.

4.4.3.2 Front contact. Electrical contact is made to the die in three ways as shown in Figure 48.

FIGURE 48. Various styles of electrical contact.

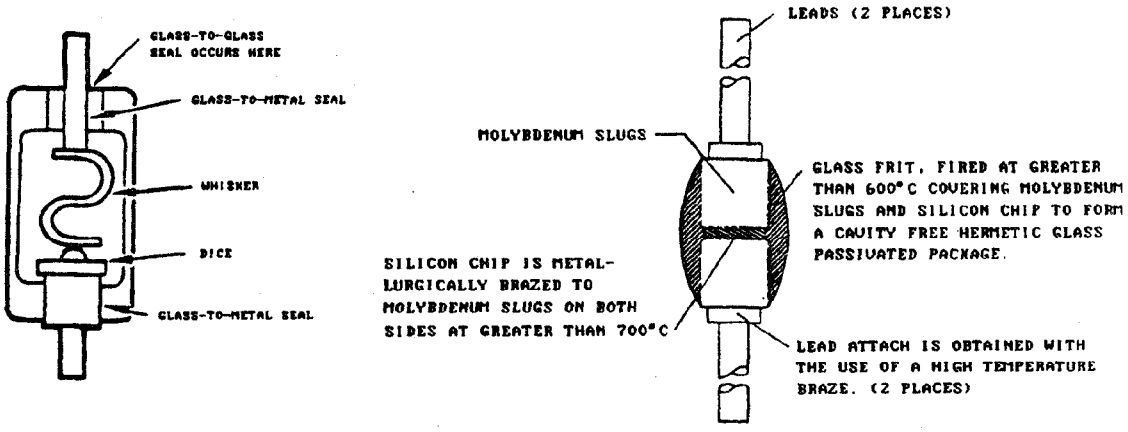
The S-bend and wire whisker are held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker. The stud contact is usually soldered to the top of a diffused die which provides good electrical and mechanical contact, but does not limit the stress on the die.

The whisker contacts have good shock and vibration tolerance. The wire whisker, however, is somewhat superior in this respect due to its lower mass and the imbedding of the whisker in the metal on top of the die which prevents sliding of the contact. Both whisker structures provide good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion and contraction of the package.

The double-slug stud contact transmits more stress to the die than the whisker structure, but is used despite this because it provides good heat transfer, thus raising the power rating of the package.

4.4.3.3 Seals. Generally, two sealing techniques are used: glass-to-metal and double-slug construction. The glass-to-metal package, in Figure 49A, uses only glass-to-metal seals. These glass-to-metal seals are formed prior to final sealing and the final seal occurs by fusing glass-to-glass at the point noted in the figure. The double-slug package in Figure 49B is made by metallurgically bonding (>700 °C) both sides of the die to either molybdenum or tungsten slugs. To this a glass sleeve or glass slurry is fused for final sealing. The final step is a lead brazing operation. This type of construction is most often associated with voltage regulator diodes.

Figure 50 shows a typical voltage regulator diode package.



A. Glass-to-metal seal B. Glass seal

FIGURE 49. Typical construction of sealing techniques.

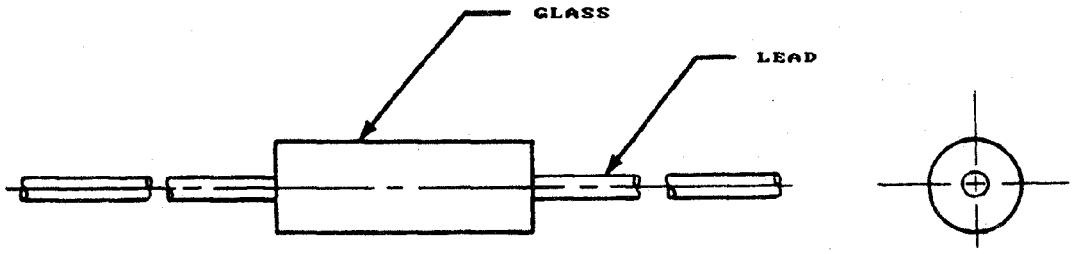
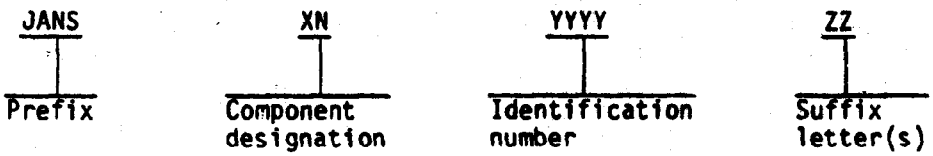


FIGURE 50. Outline drawing of voltage regulator diode.

4.4 DIODES, VOLTAGE REGULATOR

4.4.4 Military designation. The military designation for diodes is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.4.5 Electrical characteristics.

4.4.5.1 Voltage-ampere characteristics. The volt-ampere characteristics for a typical zener diode, given in Figure 51, show that it can conduct current in both directions. The forward current, I_F, is a function of the forward voltage, V_F, and is often useful for biasing applications.

The normal operating mode of the regulating diode is a function of its reverse characteristics. In this region, V_Z is the nominal zener voltage which, in various types, may range from 2.4 to 200 V in selected 5, 10 or 20 percent groupings. The zener knee is represented by a sharp break from virtual non-conductance to conductance at the nominal zener voltage. A small zener current I_{ZK}, is the minimum current required to establish the characteristics over the knee and essentially on the flat zener plateau. This minimum regulator current varies among the various types of diodes and is given in the specification sheets. The maximum zener current I_{ZM}, is limited by the power dissipation of the diode and is a function of the junction temperature. Values of I_{ZM} are also given in the specifications.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

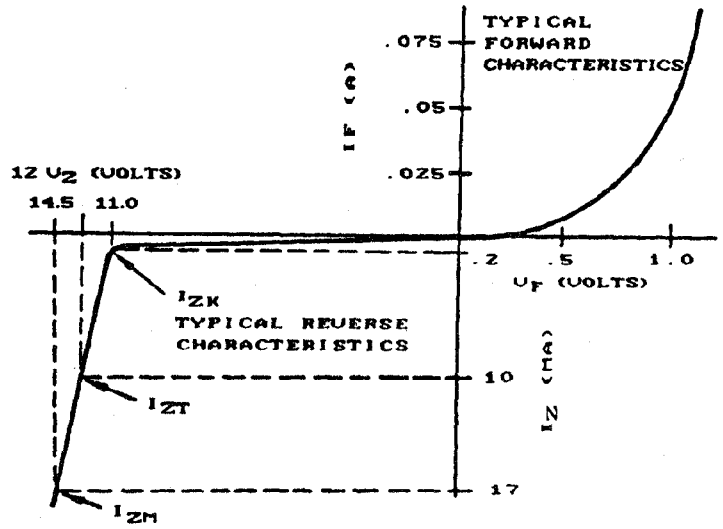


FIGURE 51. Typical forward and reverse characteristics for a 12 volt device.

V_Z is essentially constant between the limits of I_{ZK} and I_{ZM}. The plateau has, however, a small positive slope in which the precise value of V_Z will change as a function of I_Z. This change arises from the volt-ampere characteristics zener impedance, A, for fast changes I_Z, and the thermal effect due to junction temperature changes with d-c levels of I_Z. The latter will be discussed under temperature effects. The maximum zener impedance is usually specified at two selected test points for each type device. The first point is near the knee of the zener plateau and the second is near the midrange of the usable zener current excursion.

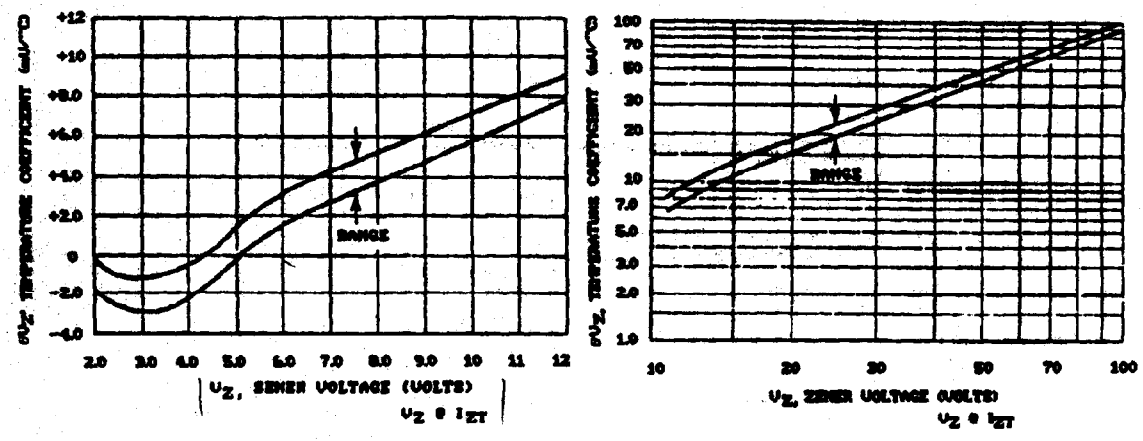
The temperature coefficient of the zener regulator diode for $V_Z \leq 5$ V via zener breakdown is negative and is somewhat dependent on the bias current as shown in graph C of Figure 52. The temperature coefficient of $V_Z = 5.0$ V may be somewhat positive or somewhat negative depending upon the bias amount as shown in graph C of Figure 52. The temperature coefficient for $V_Z > 6.5$ V is positive and is relatively independent of bias current levels for devices with V_Z up to 100 V and beyond. The variation in temperature coefficient (TC) is shown in graph A and graph B of Figure 52.

Temperature coefficient, often abbreviated as TC or represented by the symbol K_T, may be defined in either of the two following ways:

$$K_T = \frac{\Delta V_B(mV)}{\Delta T(°C)} \quad mV/°C,$$

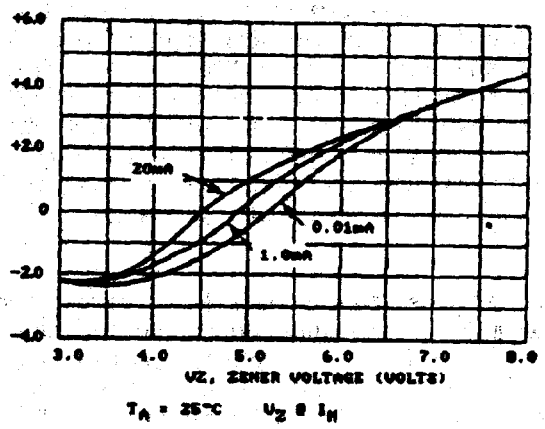
$$K_T^* = \frac{(\Delta V_B)/V_B}{100(\Delta T)} \quad \%/°C.$$

4.4 DIODES, VOLTAGE REGULATOR



A. Range for units to 12 V.

B. Range for units 12 to 100 V.



Below 3 and above 8 V, changes in zener current do not affect temperature coefficients

C. Effect of zener current

FIGURE 52. Zener diode temperature coefficient graphs.

Some designers prefer working with the %/°C figure especially when working with power supplies in which the output voltage is usually a constant multiplied by the V_B of the voltage regulator (VR) diode used as a reference. The TC of the output voltage due to the VR diode is equal to the value K_T* of the voltage regulator diode if expressed in %/°C.

Other designers are more concerned about the actual voltage change and prefer to express the value of K_T in mV/°C. To compute the total expected change in V_B caused by a temperature change, the value of V_T expressed in mV/°C is multiplied by the temperature difference. The result is in millivolts.

There may be a need to occasionally change back and forth between K_T as expressed in mV/°C and K_T* expressed in %/°C. The following equations make this a simple task:

$$K_T^* = \frac{K_T}{10\ V_B}$$

$$K_T = 10\ V_B K_T^*.$$

VR diodes which are in the mid-voltage range and exhibit effects of both zener and avalanche breakdown may have either a positive or negative temperature coefficient, depending on which effect is predominant. At low current levels, the zener effect is strongest and the TC is negative as mentioned previously.

At rather high current levels the avalanche effect becomes more evident and the VR diode takes on a positive TC. Because both zener and avalanche breakdown modes are actually at work simultaneously, with the mixture controlled by the relative current level, there is a variation in the value of TC as the current is changed.

At some specific current, the negative TC of the zener effect is exactly equal to the positive TC of the avalanche effect and the net TC is zero. In Figure 53 this condition of zero TC occurs at the point in which all three of the characteristic curves intersect.

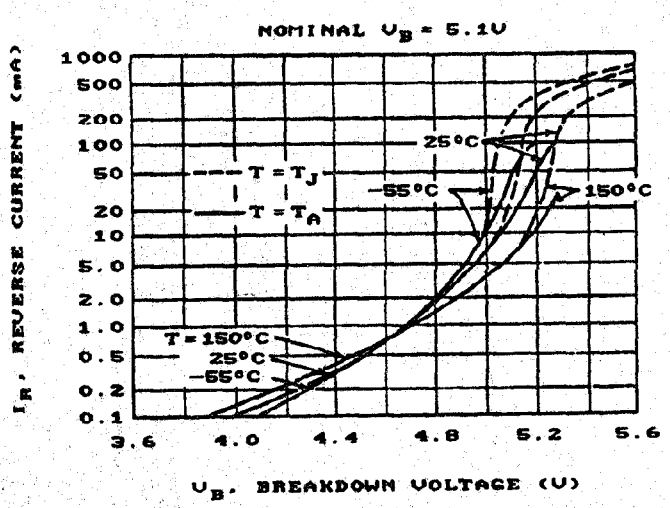


FIGURE 53. Reverse breakdown characteristics for a 1N5231 voltage regulator both zener and avalanche breakdown effects are exhibited.

4.4 DIODES, VOLTAGE REGULATOR

4.4.5.2 Zero temperature coefficient. The zero temperature coefficient characteristic is not limited to diodes of exactly a 5-V breakdown only, but can be found at various operating current points in regulators in the voltage range from 4.5 to 6.5 V. Below 4.5 V, the high current value necessary to achieve the condition of zero coefficient is prohibitive and approaches the maximum I_Z of the device. Above 6 V, the intersection of zero coefficient has reached the zero reverse current line.

In certain applications, it may be necessary to place the diode in a temperature-stable environment. A small oven system can control the temperature within about a degree. Due to the increased cost, larger size, and additional circuit complexity, this must be considered an inefficient method of achieving stability.

On the other hand, there are applications which demand a reference at a discrete voltage, usually high, where the use of an oven may be the only solution. Where operation at an arbitrarily lower voltage value is possible, either of the following alternatives should be considered:

a. A series of temperature-compensated regulator diodes is available. For example, the 1N827 with a V_Z = 5.9 to 6.5 V, has a temperature coefficient of only ±0.001%/°C or 0.062 mV/°C at I_Z = 7.5 mA, over a temperature range of -55 to +100 °C.

b. For optimum performance, a number of lower voltage units can be connected in series. The resultant temperature coefficient, dynamic resistance, and thermal resistance will be considerably reduced by using a series configuration; see paragraph 4.4.2.2.1 for a discussion of this approach. A single junction unit rated for the desired voltage exhibits a high positive temperature coefficient and a high dynamic resistance.

4.4.5.3 Maximum current limits. Any silicon zener regulator has a maximum limit of current range over which it can operate. This upper limit is established by the heat dissipation capability of the regulator. The maximum current which can flow through the diode is limited in practice by the junction temperature, which must not rise above some critical value. Heat generated internally at the crystal junction, together with the ambient temperature, contributes to the determination of the junction temperature. Refer to paragraph 4.4.7.3 for recommended reliability derating criteria.

Various power classes of regulators ranging from several hundred milliwatts to 100 W. are in production.

It should be noted that although I_Z is the recommended operating point, usually 20 percent of I_Z maximum, any current beyond the zener-breakdown curvature may be arbitrarily selected. If operation near the knee of the curve is desired, diodes exhibiting a V_Z of above approximately 7 V should be chosen.

4.4.5.4 Zener capacitance. Zener diodes are basically pn junctions operated in the reverse direction; therefore, they display a capacitance that decreases with increasing reverse voltage because the effective width of the pn junction is increased by the removal of charges as reverse voltage is increased. This decrease in capacitance continues until the zener breakdown region is entered; very little further capacitance change takes place, owing to the now fixed voltage across the junction. The value of this capacitance is a function of the material resistivity (ρ) (amount of doping--which determines V_Z nominal), the diameter (D) of junction or die size (determines amount of power dissipation), the voltage across the junction, V_R, and some constant, K. This relationship can be expressed as:

$$C_d = \frac{KA}{(\rho V_R)^n}$$

A = Effective Junction Area
C_d = Junction Capacitance
n = -0.5 for Alloy Junction
-0.33 for Graded Junction

After the junction enters the zener region, capacitance remains relatively fixed and the ac resistance then decreases with increasing zener current. Typically this value is between 10 and 10000 picofarad.

4.4.5.5 High frequency and switching considerations. At frequencies above 100 kHz, and switching speeds less than 10 µs, shunt capacitance of zener diodes begins to seriously affect their usefulness. In any application where the signal is recurrent, the shunt capacitance limitations can be overcome by operating a fast diode in series with the zener. Upon application of a signal, the fast diode conducts in the forward direction, charging the shunt zener capacitance to the level where the zener conducts and limits the peak. When the signal swings to the opposite direction, the fast diode becomes backbiased and prevents fast discharge of shunt capacitance. The fast diode remains backbiased when the signal rises and exceeds the charge level of the capacitor. When the signal exceeds this level, the fast diode conducts as does the zener.

Thus, between successive cycles or pulses, the charge in the shunt capacitor holds off the fast diode, preventing capacitive loading of the signal until zener breakdown is reached.

4.4.5.6 Noise density considerations. The breakdown voltage of the junction in a voltage regulator diode does not occur simultaneously across the entire area but it will take step-like functions. This contributes to a ragged characteristic in the knee region which is shown in an exaggerated form in Figure 54. When the current reaches a certain value, the entire junction takes part in the breakdown process and no further steps in breakdown current occur.

4.4 DIODES, VOLTAGE REGULATOR

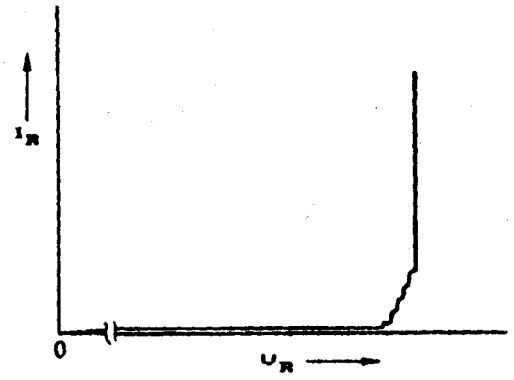


FIGURE 54. Exaggerated breakdown characteristics in the knee region.

If the voltage regulator diode is biased in the region of the knee, even very slight variation in current will cause jumps in the breakdown voltage which will appear as a noise voltage.

A small part of this noise is due to the internal resistance associated with the device. A larger part of zener noise is a result of the zener breakdown phenomenon and is called microplasma noise. This microplasma noise is generally considered white noise with equal amplitude for all frequencies from about zero cycles to approximately 200,000 cycles. To eliminate the higher frequency component of noise a small shunting capacitor can be used. The lower frequency noise generally must be tolerated, because a capacitor required to eliminate the lower frequencies would degrade the regulation properties of the zener device in many applications.

Most manufacturers rate the low power zener diode series, such as the 1N5221 and 1N4115 families, with a maximum noise density at 250 μA. The rating of microvolts rms, root-mean-square, per square root cycle enables the calculation of the maximum noise density, volts per square root bandwidth.

Noise density decreases as zener current increases. This can be seen by the graph in Figure 55, where a typical noise density is plotted as a function of zener current for some typical low noise zener diodes, 1N4115 and 1N4124.

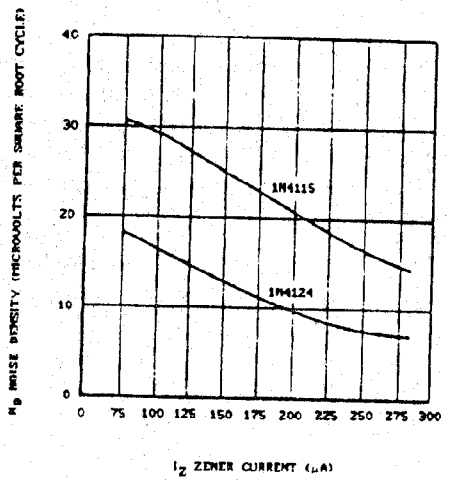
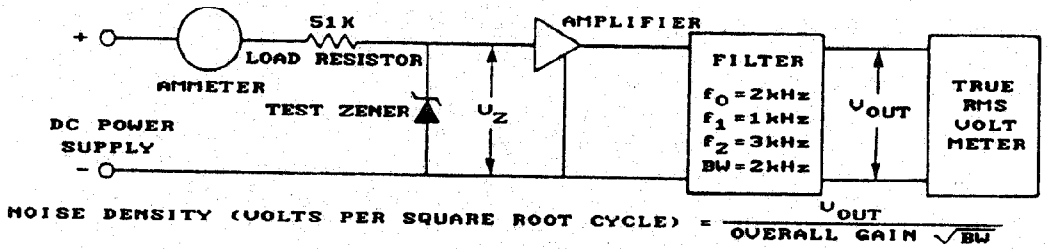


FIGURE 55. Typical noise density vs zener current.

The junction temperature will also change the zener noise levels. Thus the noise density rating must indicate bandwidth, current level, and temperature.

The circuit given in Figure 56 is used to measure noise density. The input voltage and load resistance is high so that the zener diode is driven from a constant current source. The amplifier noise must be negligible compared with the zener device under test. The filter bandpass is known so that the noise density in volts rms per square root cycle can be calculated.



$$\text{NOISE DENSITY (VOLTS PER SQUARE ROOT CYCLE)} = \frac{U_{OUT}}{\text{OVERALL GAIN } \sqrt{BW}}$$

WHERE: BW = FILTER BANDWIDTH (CYCLES)

U_{OUT} = OUTPUT NOISE (VOLTS RMS)

FIGURE 56. Noise density measurement method.

4.4 DIODES, VOLTAGE REGULATOR

The very low-voltage zener regulator diodes operate primarily on field emission or zener breakdown have very little noise. When avalanche breakdown becomes more dominant (V_Z > 6 volts), the noise density likewise increases as shown in Figure 57. In addition, noise levels in zener regulator diodes are greatly process-dependent and rather wide variations are probable among different manufacturers of the same type.

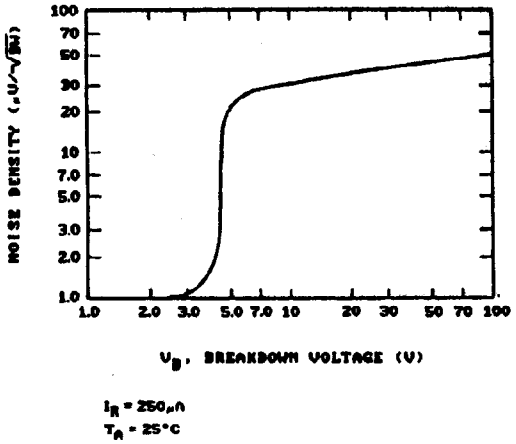


FIGURE 57. Typical noise density vs breakdown voltage.

4.4.6 Environmental considerations. Typical environmental conditions and screens which voltage regulator diodes are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 Screening procedures in subsection 4.1 Diodes-general. For the specific device selected consult the MIL-S-19500 reference slash sheet.

4.4.7 Reliability considerations.

4.4.7.1 Failure mechanisms and data. In extensive tests of zener voltage regulator diodes, the two most critical parameters were zener impedance and reverse current. Zener impedance is a measure of the slope of the diode's voltage-current characteristics in the breakdown or voltage-regulation region, and reverse current (leakage) is a measure of the diode's characteristics in the blocking or nonconducting direction at stated conditions of voltage and temperature. Long term life tests have yielded a significant number of failures in the reverse current and zener impedance parameters. Failure mechanisms of zener diodes are similar to those of other diode types discussed in paragraph 4.1.6 General reliability considerations, subsection 4.1 Diodes-general.

4.4.7.2 Screening. As with previous recommendations for other types of diodes, the voltage regulator procured as a JANS device offers the best protection against the above mentioned problems. It should be noted that the two characteristics which are usually measured to verify zener diode quality are the zener diode breakdown voltage (V_Z) and reverse current (I_R). Measurement of these characteristics during burn-in screening verifies the reliability of the diode. The magnitude of reverse current is largely dependent upon surface conditions, die size and thickness, and junction area of the semiconductor. Therefore, delta measurements during burn-in for I_R are a useful criterion of diode reliability.

JANTXV Zener diodes are not subjected to high temperature reverse bias (HTRB). This could allow some zener diodes to escape with surface contamination. JANS zener diodes are subjected to HTRB at 80 percent of nominal V_Z for parts with V_Z greater than 10 V. For ports with V_Z less than 10 V the test is not necessary.

4.4.7.3 Derating. History has shown that the largest single cause of diode failure is operating above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be performed to enhance the reliability of systems. Users should refer to MIL-STD-975 for derating factor guidelines. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress a semiconductor could be subjected to. It will result in electrical parameter drift and general degradation of the electrical and mechanical characteristics of the device.

4.5 DIODES, VOLTAGE REFERENCE

4.5 Voltage reference.

4.5.1 Introduction. A silicon voltage-reference diode is an assembly of two or more silicon junctions encapsulated in a single package. This device exhibits an extremely good voltage regulating ability over wide variations in ambient temperature for long periods of time.

A term often confused with the voltage reference diode is the zener diode. A zener diode is designed to operate under reverse bias for purposes of voltage regulation.

4.5.2 Usual applications. Application of the reference diode should take into consideration the following points: (1) current through the device should not vary widely because the diode then loses much of its value as a precise voltage reference; (2) anticipated operating temperature ranges should govern choice of current selected for device application. For control of this all important constant current over widely varying temperatures, the circuits of Figure 58 are suggested.

Circuit A of Figure 58 is usable only if a source of well-regulated high voltage is available. It must have sufficient reserve current available to supply the additional 10 mA required by the reference element. This power supply should be almost impervious to temperature effects. Variations in the value of R will, of course, result in undesirable current shifts.

In cases where no such separate regulated source is available, or where isolation is required, an independent supply may be designed using two series-connected reference diodes as preregulators, as shown in circuit B of Figure 58. Because these are also highly stable devices themselves, the end reference element is assured of an extremely constant current over a wide range of temperatures. The chief disadvantage is the cost of the additional reference devices.

Less costly zener diodes may be used as the preregulator instead of the reference diode as diagrammed in circuit C of Figure 58. The three 5.6-V diodes operating at their zero temperature coefficient current point provide excellent stability.

A fourth method involves only the use of a single zener diode in the range of 12 to 20 V as shown in circuit D of Figure 58. Such a diode, however, will suffer from an inherently positive temperature coefficient; i.e., its voltage will rise with temperature, which increases the current through the reference element. This disadvantage can be overcome by compensating with resistor R, which also has a positive coefficient. As the voltage across the zener regulator increases with temperature, the resistor will tend to increase in resistance so as to maintain a relatively constant current through the reference element.

In all the circuits shown in Figure 58, resistor R is selected to provide the 10 mA reference current, or less as discussed earlier.

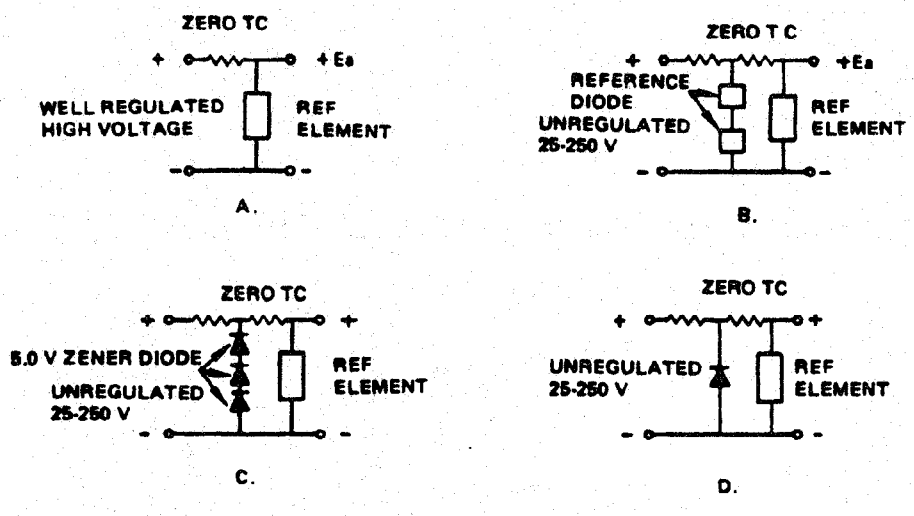


Figure 58. Typical Bias Current Systems.

4.5.3 Physical construction. The voltage reference diode is generally made with passivated diffused, planar or diffused planar junction on an epitaxial substrate process die in a single glass-to-metal or double-slug package. The multiple junctions are assembled back-to-back to achieve the desired temperature coefficient (TC). One junction operates in a reverse-bias condition and exhibits a positive TC; other junction(s) operate in the forward-biased condition and exhibit a negative TC. The net result is a near-perfect cancellation of voltage drift versus temperature change, ΔBV. The three basic die processes are shown in Figures 59, 60 and 61 respectively.

4.5 DIODES, VOLTAGE REFERENCE

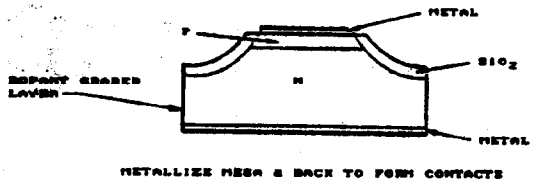


FIGURE 59. Passivated diffused process.

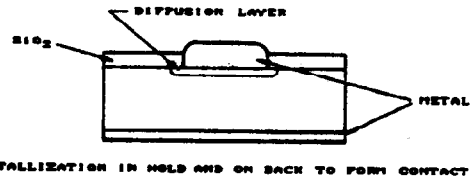


FIGURE 60. Planar process.

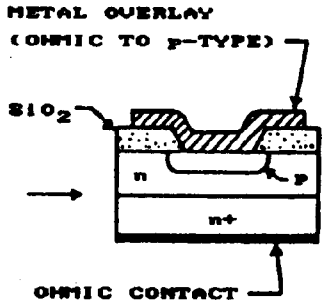


FIGURE 61. Diffused planar junction on an epitaxial substrate.

4.5.3.1 Back contacts. Illustrated in Figure 62 below are three back contact techniques used in constructing voltage reference diodes. Although the alloyed back contact is the strongest contact, voltage reference diodes generally use high temperature alloy back contacts, associated with double-slug construction, for high reliability applications.

The solder contact has good mechanical strength and thermal conductivity. The silicon-silver alloy contact which is generally associated with the high temperature alloy has good mechanical strength and thermal conductivity. The silicon chip is metallurgically brazed to plated molybdenum or tungsten slugs at temperatures exceeding 700 °C.

The silicon-gold alloy contact is the strongest and most thermally conductive of the above techniques. Because of the high eutectic temperature of silicon-gold eutectic alloy, 377 °C, the contact will withstand high temperatures during overload.

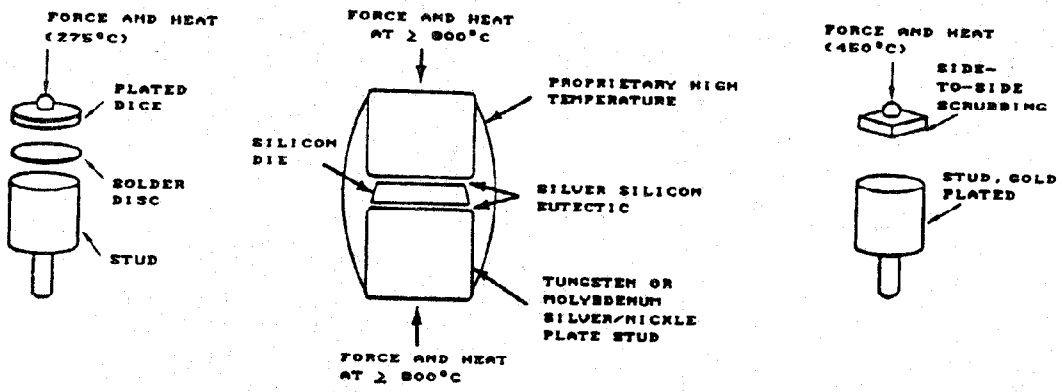


FIGURE 62. Back contact formation techniques.

Molybdenum, tungsten, and dumet are usually used for studs in glass packages because their thermal expansion coefficients match those of certain glasses, allowing sealing of these glasses to them. Copper and copper alloys are usually used in devices which handle appreciable power because they have much better thermal conductivity than molybdenum, tungsten and dumet.

4.5.3.2 Front contacts. Electrical contact is made to the die in three ways as shown in Figure 63.

The C bend and S bend whiskers are held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker. The stud contact is usually soldered to the top of a diffused die which makes good electrical and mechanical contact but does not limit the stress on the die.

The whisker contacts have good shock and vibration tolerance. Both whisker structures provide good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion and contraction of the package.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.5 DIODES, VOLTAGE REFERENCE

The double-slug contact transmits more stress to the die than the whisker structure, but is used in spite of this because it provides good heat transfer, thus raising the power rating of the package.

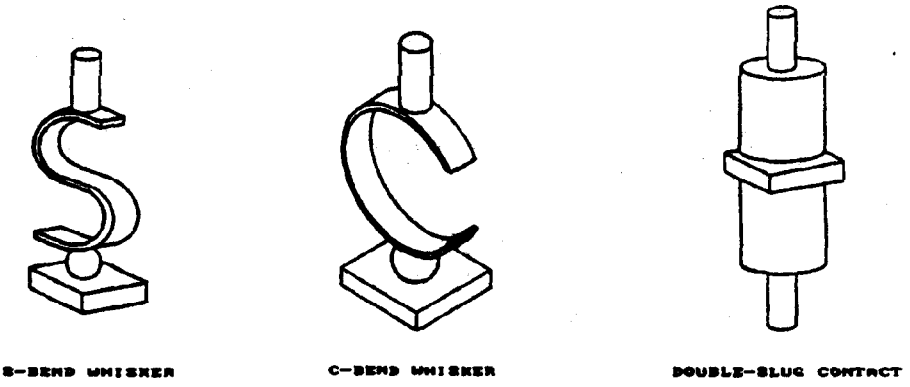


FIGURE 63. Various styles of electrical contact.

4.5.3.3 Seals. Generally, two sealing techniques are used: metal-to-glass and double-slug construction. The glass-to-metal package in Figure 64A uses only glass-to-metal seals. These glass-to-metal seals are formed prior to final sealing; the final seal occurs by fusing glass-to-glass at the point noted in the figure.

The double-slug package of Figure 64B is made by metallurgically bonding, > 700 °C, both sides of the die to either molybdenum or tungsten slugs to which a glass sleeve or slurry is fused for final sealing. The final step is a lead brazing operation. This type of construction is most often associated with voltage reference diodes.

Figure 65 shows a typical voltage reference diode package.

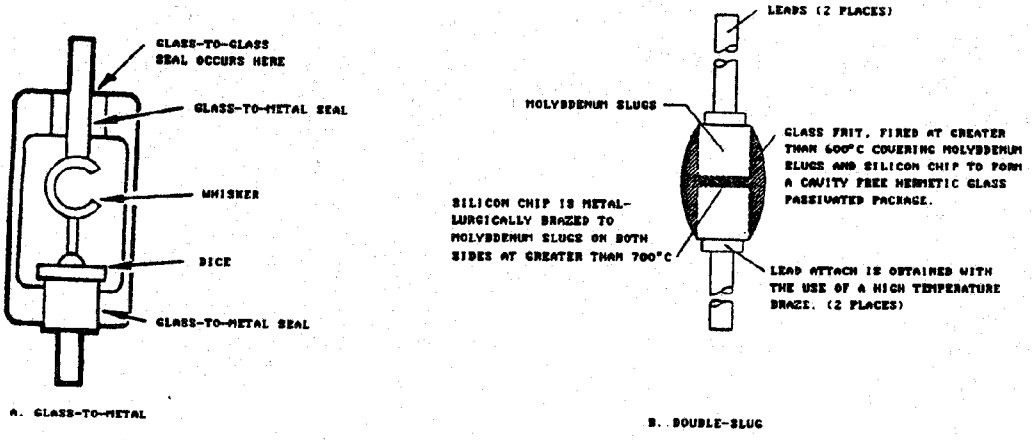


FIGURE 64. Various sealing techniques.

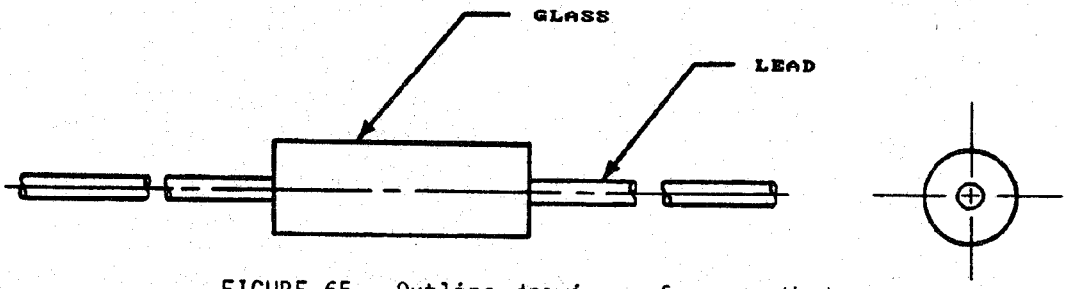
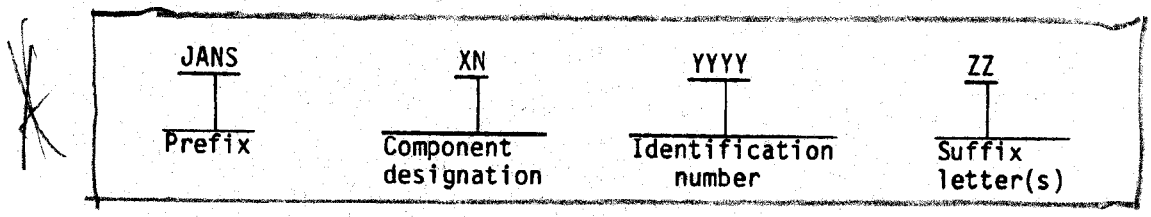


FIGURE 65. Outline drawing reference diode.

4.5.4 Military designation. The military designation for diodes is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable is placed after the prefix. See MIL-M-19500 for details.

4.5 DIODES, VOLTAGE REFERENCE

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.5.5 Electrical characteristics.

4.5.5.1 Stability. In the reference diode, the stability, often called absolute or time stability, is defined as the change in zener voltage over a period of time (operating or shelf life). This change is due to some characteristic of the device itself, rather than external conditions such as changes in operating temperature or current. This time stability is often confused with temperature stability.

Figure 66 is a stability curve for an 8.4-V reference diode conducting 10 mA current with TC = ±0.001%/°C. It can be seen that the reverse current can vary ±1 mA without exceeding the TC requirement. However, if the reverse current varies, the resultant variation in voltage drop, as shown in Figure 67, must be considered. If the element has a drop of 8.4 V at 10.0 mA, the drop at 9 and 11 mA will be 8.389 V and 8.411 V, respectively. This is a net change of ±11 mV and represents a change of about 1 mV per 100 mA deviation from the design center of 10 mA. Referring again to Figure 66, it can be seen that if, in addition to the normal allowable drift with temperature, a simultaneous change in reverse current from 10 mA to 9 mA occurs, an additional voltage variation corresponding to -11 mV will result. At 100 °C, the voltage across the reference element could possibly have varied from its original value at +25 °C by as much as -17.3 mV, 11 plus 6.3 mV. This represents a stability factor of only about ±0.003%/°C; three times more than the specification allows. Therefore, if the current through the element is allowed to vary during operation, this otherwise stable device loses much of its value as a precise voltage reference.

To investigate what happens to the stability of this reference element at reverse currents other than those specified, measurements were taken at a number of discrete current values from 8 mA through 12 mA. From Figure 68, it can be seen that current values above and below the 9 to 11 mA range result in the element exhibiting thermal stability characteristics that no longer conform to the ±0.001%/°C requirements. Note that at 12 mA the stability is poor at both high and low temperatures. Operation at 8 mA shows excellent low temperature characteristics, but at high elevated temperatures the drift is excessively negative. However, if a particular application calls for good voltage stability only over a limited temperature range (i.e., -25 °C through

+75 °C) it would seem advantageous to limit the current through the element to 9 mA, 10 percent lower than its specified rating. For laboratory use, and in industrial equipment, where the temperature seldom exceeds +55 °C, a stability of at least ±0.0005%/°C could be expected. If operation at only very low temperatures is contemplated, a current of even less than 8 mA would prove optimum. The choice of current should be set by the expected operating environmental conditions.

Figure 69 illustrates thermal voltage stability versus reverse operating currents, and shows how anticipated operating temperature ranges should govern the choice of current.

4.5.5.2 Forward Characteristics. Excellent stabilization of transistor collector current for variations in both supply voltage and temperature can be obtained by the use of a compensating diode operating in the forward direction in the bias network of amplifier or oscillator circuits. Figures 70 and 71 show the forward characteristics of a transistor and a compensating diode, respectively. The operating point is represented on the diode characteristics by the dashed horizontal line. The diode current at this point determines a bias voltage which establishes the transistor idling current. This bias voltage shifts with varying temperature in the same direction and magnitude as the transistor characteristic, and thus provides an idling current that is essentially independent of temperature.

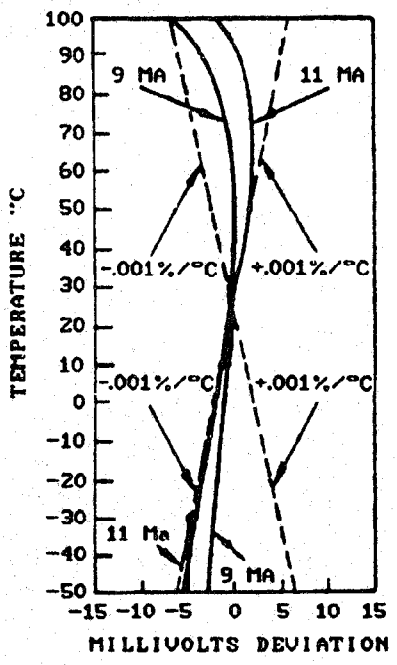


FIGURE 66. Typical stability factor at 9 and 11 mA reverse current.

4.5 DIODES, VOLTAGE REFERENCE

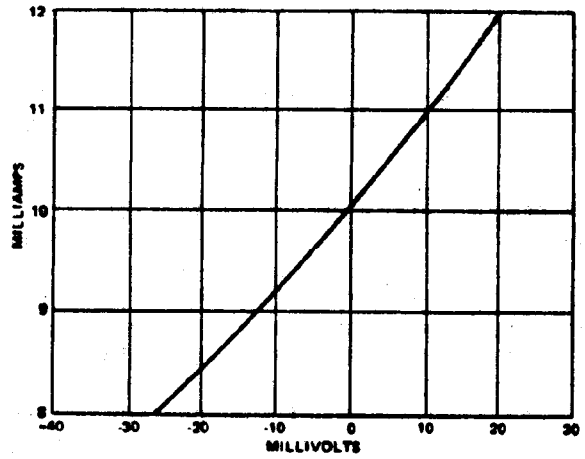


FIGURE 67. Typical voltage drop variation from a 10 mA reference point.

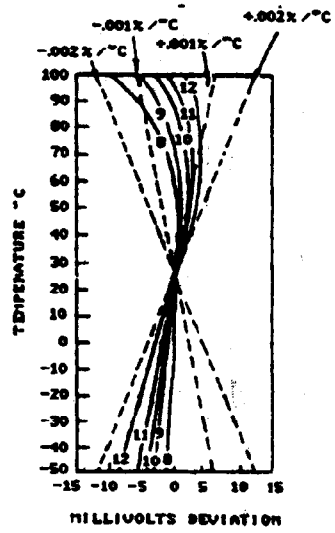


FIGURE 68. Typical stability at 8 through 12 mA.

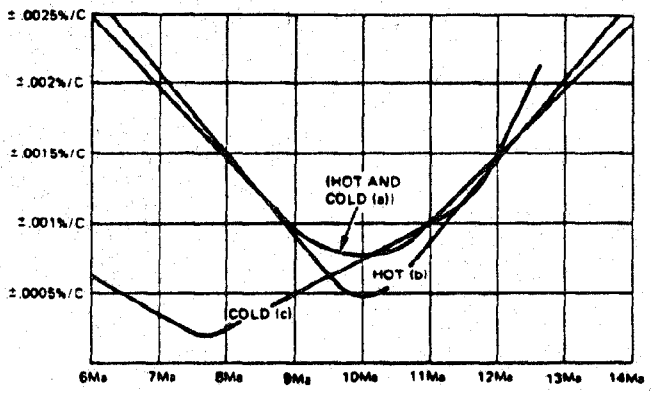


FIGURE 69. Typical thermal voltage stability vs reverse current.

The use of a compensating diode also reduces the variation in transistor idling current as a result of supply-voltage variations. Because the diode current changes in proportion with the supply voltage, the bias voltage to the transistor changes in the same proportion and idling-current changes are minimized.

4.5.5.3 Temperature effects. Figure 72 shows the effect of temperature on voltage reference diode characteristics. The forward characteristic does not vary significantly with reverse breakdown (zener voltage) rating. A change in ambient temperature from 25 to 100 °C produces a shift in the forward curve in the direction of lower voltage.

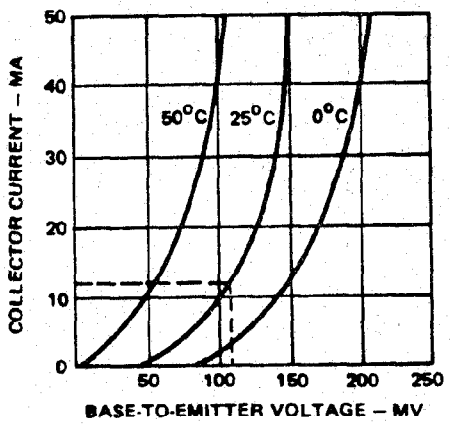


FIGURE 70. Transfer characteristics of transistor.

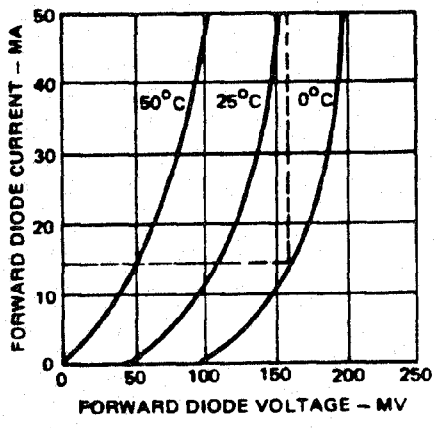


FIGURE 71. Forward characteristics of compensating diode.

4.5 DIODES, VOLTAGE REFERENCE

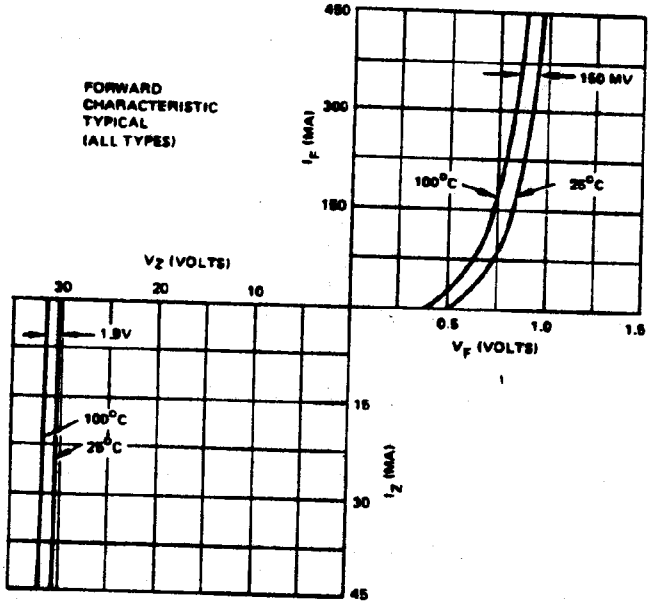


FIGURE 72. Effects of temperature on voltage reference diode characteristics.

4.5.5.4 Temperature compensation. Various combinations of forward biased junctions and reverse biased junctions may be arranged to achieve compensation. From Figure 73, it can be seen that if the absolute value of voltage change (ΔV) is the same for both the forward biased diode and the zener diode where temperature has gone from 25 °C to 100 °C, the total voltage across the combination will be the same at both temperatures since one ΔV is negative and the other positive. Furthermore, if the rate of increase or decrease is the same throughout the temperature change, the voltage will remain constant. The nonlinearity associated with the voltage temperature characteristics is a result of this rate of change not being a perfect match, i.e.,

$$V_{REF} = V_Z + \Delta V_Z + V_D - \Delta V_D$$

Figure 74 shows the voltage temperature characteristics of a typical voltage reference diode. It can be seen that the voltage drops slightly with increasing temperature.

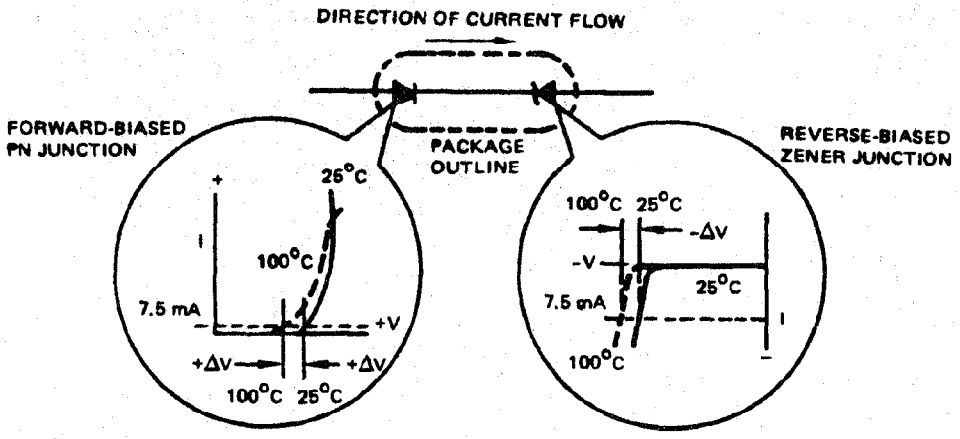


FIGURE 73. Principle of temperature compensation.

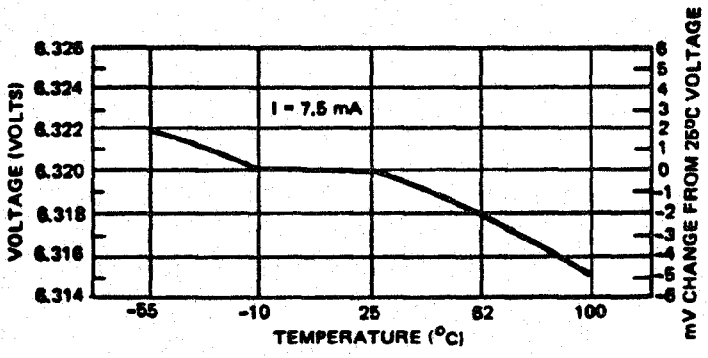


FIGURE 74. Voltage vs temperature for a typical voltage reference diode (1N827).

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.5 DIODES, VOLTAGE REFERENCE

Figure 75 illustrates that there is a significant change in the temperature coefficient of a unit depending on how much above or below the test current the device is operated. If the three curves intersect at "A," operation at I_A results in the least amount of voltage deviation due to temperature from the +25 °C voltage. At I_B and I_C there are greater excursions, $\Delta V_B + \Delta V_C$, from the 25 °C voltage as temperature increases or decreases.

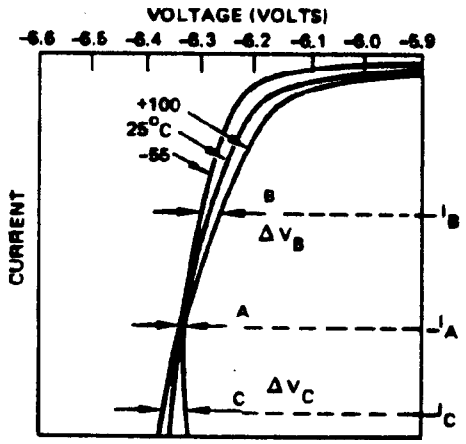


FIGURE 75. Voltage-ampere curves showing crossover at A.

To summarize, the following are significant considerations for reference diode applications:

a. Current regulation is probably the most critical consideration when using temperature compensated units.

b. Zener impedance is defined as the slope of the V-I curve at the test point corresponding to the test current. Impedance changes with temperature, but the variation is usually small and it can be assumed that the amount of current regulation needed at +25 °C will be the same for other temperatures.

c. Standard voltage reference diodes typically exhibit a time stability of 200-500 ppm per 1000 hours.

4.5.5.5 Electrical ratings. Table VI gives the electrical ratings for several typical voltage regulator diodes.

The information is for general reference only. For the specific device selected, consult the applicable MIL-S-19500 specification.

TABLE VI. Electrical ratings

Device Type No.***	Reference Voltage Range			Dynamic Impedance		Voltage Temperature Stability
	BV (Min. Volts)	BV (Max. Volts)	@ Iz (mA)	Z Ohms	@ Iz (mA)	BV 1/
1N827-1*	5.90	6.50	7.5	15	7.5	0.009
1N829-1*	5.90	6.50	7.5	15	7.5	0.005
1N937B-1**	8.55	9.45	7.5	20	7.5	0.037
1N938B-1**	8.55	9.45	7.5	20	7.5	0.018
1N939B-1**	8.55	9.45	7.5	20	7.5	0.009
1N940B-1**	8.55	9.45	7.5	20	7.5	0.0037

* From MIL-S-19500/159, -1 means metallurgically bonded, ±5 percent V_Z tolerance.
** From MIL-S-19500/156, -1 means metallurgically bonded ±5 percent V_Z tolerance.
*** All devices tabulated have military outline DO7.

1/ BV is measured in volts per span of temperature. The particular span of temperature varies for different devices and may be found in the applicable MIL-S-19500 slash sheet.

NOTE: This table is not to be used for part selection; use MIL-STD-975 for that purpose.

4.5.6 Environmental considerations. Typical environmental conditions and screens which voltage reference diodes are capable of withstanding are not substantially different from that given in paragraphs 4.1.6.9 Environmental considerations, and 4.1.6.10 Screening procedures, in subsection 4.1 Diodes-general. For the specific device selected, consult the applicable MIL-S-19500 reference slash sheet.

4.5.7 Reliability considerations. There are many failure modes associated with voltage reference diodes. The majority of failures are due to zener voltage and short and open circuits. The mechanical and thermal stress tests identify most of the conditions of these failures. As with other diode families, it is suggested that the JANTXV and JANS process would eliminate all or most of these failure types.

4.5.7.1 Failure mechanisms. Because the failure mechanisms for this type of diode are similar to those of other diode types, paragraph 4.1.6 Reliability considerations should be consulted for a discussion of diode failure mechanisms.

4.5 DIODES, VOLTAGE REFERENCE

4.5.7.2 Derating. History has shown that the largest single cause of diode failure is operation above allowable thermal and electrical stress levels. Accordingly, it is imperative that derating of parts be performed to increase the reliability of military systems. Users should refer to MIL-STD-975 for derating factor guidelines. In general, derating for zener current should be limited to no more than $I_Z = I_Z$ nom + .5 (I_Z max. - IZ nom) of the values published on the vendor data sheet. The junction temperature should be limited to +125 °C maximum. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress a semiconductor device could be subjected to. It will result in early end of life, electrical parameter drift, and a general degradation of the device's electrical and mechanical characteristics.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.6 Current regulator.

4.6.1 Introduction. The current-regulator diode is basically a field-effect transistor that has its gate and source terminals connected together. It is called either a constant-current diode or field-effect current regulator diode. It presents a constant current independent of the terminal voltage over a wide operating range and exhibits a very high circuit impedance.

4.6.2 Usual applications. The current-regulator diode is useful in such applications as over-current protection, transistor biasing, linear ramp or stairstep generators, differential amplifiers, and precision reference voltage sources. It is always used in the forward bias direction between the pinch-off voltage and the breakdown voltage. The symbol and polarity of a current regulator diode are shown in Figure 76.

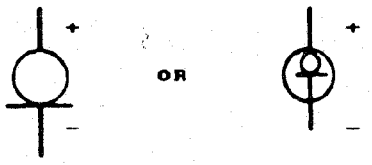


FIGURE 76. Symbol and polarity of current regulator diode.

Devices are available that cover the current range from 220 µA to 4.7 mA in 32 different current steps. The family series of devices that cover this range are the 1N5285 through 1N5314 diodes. The current-regulator diodes in this series have a tolerance of plus or minus 10 percent and are available as JANTXV versions in a DO-7 package.

To use the current-regulator diode as a constant-current bias source in a differential amplifier circuit, the device would be connected as shown in Figure 77.

The current-regulator diode replaces the transistor, zener diode, and resistors that would normally be used in a conventional circuit. To use the current-regulator diode as a low-voltage reference source, it would be connected in the circuit as shown in Figure 78.

The constant-current regulator diode drives a known resistor value that produces an output reference voltage whose value is determined by Ohm's law. This results in a precision millivolt reference source when it is desirable to generate a stable and accurate voltage reference at voltages lower than those offered by zener diodes.

4.6 DIODES, CURRENT REGULATOR

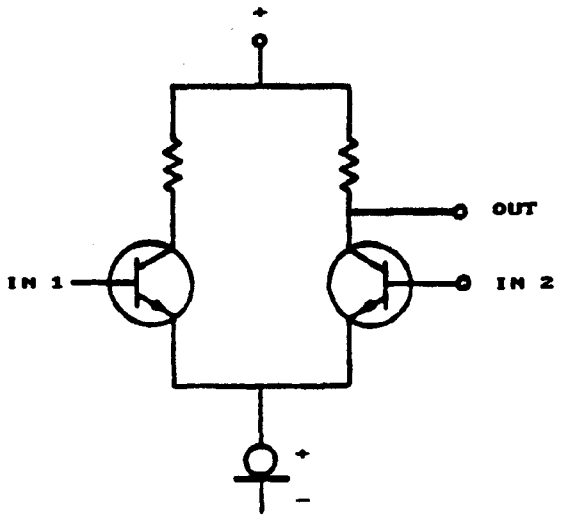


FIGURE 77. Constant current bias source in a differential amplifier circuit.

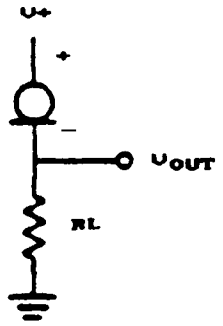


FIGURE 78. Circuit showing current regulator diode as a low-voltage reference source.

The current-regulator diode can also be combined with a zener diode to obtain better voltage and temperature performance than a zener diode could provide by itself. See Figure 79 for the circuit connection. This combination may be used when an accurate temperature-stable reference voltage is required.

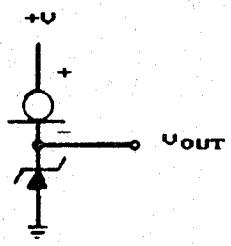


FIGURE 79. Current regulator diode combined with a zener diode.

4.6.3 Physical construction. The current-regulator diode generally has a planar process die in a glass-to-metal package. The basic die process is shown in Figure 80.

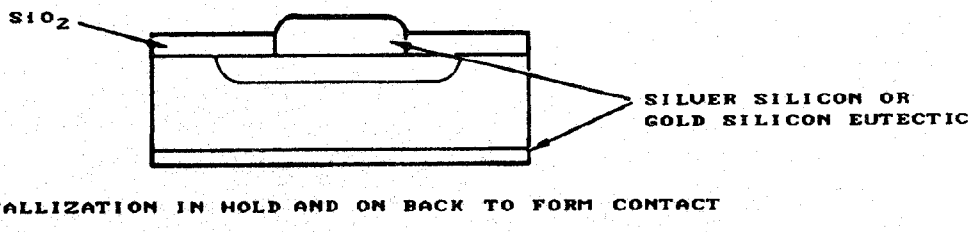


FIGURE 80. Planar process.

4.6.3.1 Back contacts. Illustrated in Figure 81 is the solder back contact used in constructing current regulator diodes.

The solder contact has good mechanical strength and thermal conductivity. The stud material is generally molybdenum, tungsten, or dumet because their coefficients of expansion closely match those of glass.

4.6 DIODES, CURRENT REGULATOR

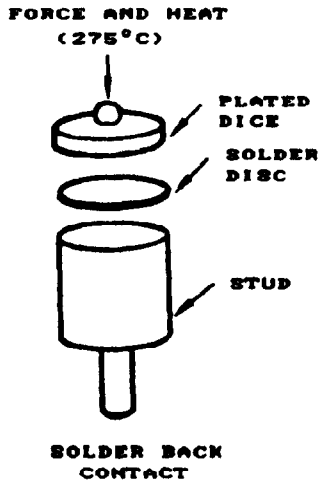


FIGURE 81. Solder back contact.

4.6.3.2 Front contacts. Electrical contact is made to the die in two ways as shown in Figure 82.

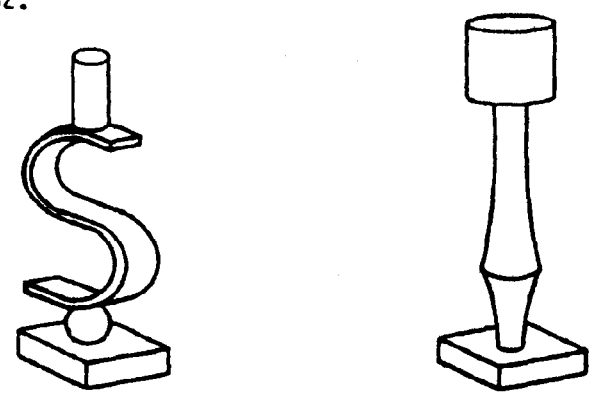


FIGURE 82. Two styles of electrical contact.

The whisker is held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker. The arrowhead contact is made using solder between the wire and the top of the die.

The whisker contact has good shock and vibration tolerance. The arrowhead contact is somewhat superior due to imbedding of the end of the contact in the metal on top of the die. The whisker structure provides good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion as well as contraction of the package.

4.6.3.3 Seals. Figure 83 shows the sealing technique for glass-to-metal seals.

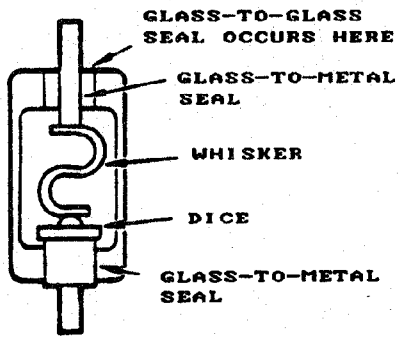


FIGURE 83. Typical construction of glass-to-metal seal technique.

The sealing technique generally used for a current-regulator diode is a glass-to-metal seal. The glass-to-metal seals in the above package are formed prior to final sealing, which is a glass-to-glass seal.

Figure 84 shows a typical current regulator diode package.

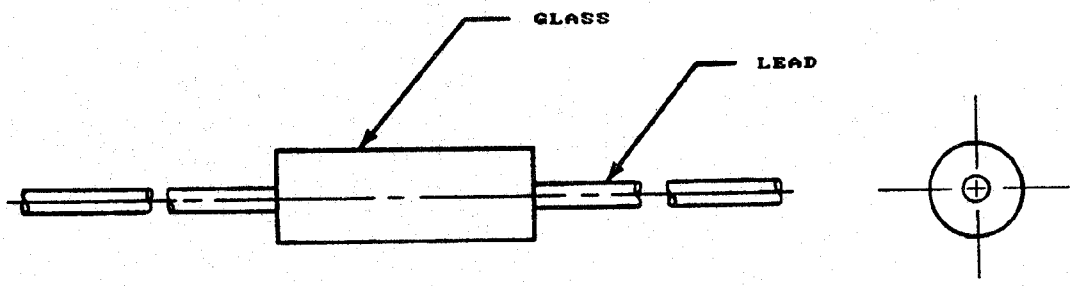
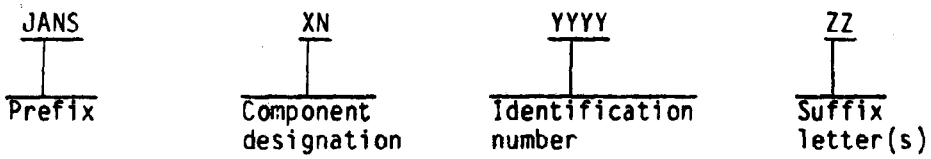


FIGURE 84. Outline drawing of a typical current-regulator diode (DO-7 package).

4.6 DIODES, CURRENT REGULATOR

4.6.4 Military designation. The military designation for diodes is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order by registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.6.5 Electrical characteristics. The current-regulator diode is basically an FET with its gate and source terminals connected. An FET connected in this way acts as a constant-current device over a wide voltage range. This range is from the pinch-off voltage to the forward breakdown voltage point. In this region, the power supply or circuit voltage can vary widely with very little change, if any, in the current the device allows to pass. Therefore, the device is biased in the forward direction between the pinch-off voltage and the forward breakdown voltage. Any circuit voltage within this range will result in a constant drain current. The ideal constant-current diode will have a very low pinch-off voltage, a high breakdown voltage, and as high a dynamic impedance between these limits as possible. The current-regulator diode should not be used in the reverse bias direction.

4.6.6 Environmental considerations. Typical environmental conditions and screening tests which current regulator diodes are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 Screening procedures in subsection 4.1 Diodes, general. For the specific device selected, consult the applicable MIL-S-19500 reference slash sheet.

4.6.7 Reliability considerations.

4.6.7.1 Failure modes. The major electrical failure modes for current-regulator diodes are a shift in the regulator current (I_P) or a failure of the device to

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

regulate. Break-down voltage (BV) is not a major electrical failure mode with current-regulator diodes because of the high voltage. Because of the above catastrophic failure modes, the devices are more likely to have shorts than opens.

Because the failure mechanisms for this type of diode are similar to those of other diode types, paragraph 4.1.6 Reliability considerations should be consulted for problems related to diode failure mechanisms.

4.6.7.2 Derating. History has shown that the largest single cause of failure for diodes is operation above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be performed done to increase the reliability of systems. Users should refer to MIL-STD-975 for derating factor guidelines.

The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress a semiconductor could be subjected to. It will result in electrical parameter drift, and a general degradation of the electrical and mechanical characteristics of the device.

4.7 DIODES, VOLTAGE VARIABLE CAPACITOR

4.7. Voltage-variable capacitor.

4.7.1 Introduction. Voltage-variable capacitance diodes are not included in MIL-STD-975, however, they are included in this handbook to provide a technical understanding of this part.

The voltage-variable capacitance diode is a silicon pn-junction diode designed for use as a voltage-variable capacitor. This device provides a means of changing circuit capacitance through bias voltage control. The capacitance varies essentially as $1/\sqrt{V}$ as the voltage across its terminals is varied. This device maintains constant characteristics over a wide temperature range but is less temperature sensitive when used at higher operating voltages. Therefore, if capacitance variation must be minimized, it is advisable to operate at higher bias voltages.

With voltage-variable capacitance diodes, it is possible to construct very low noise amplifiers. In principle, the noise figure of a parametric amplifier is equal to 0 dB if the amplifier is considered to be ideal; i.e., one that generates no internal noise. This is not possible in a conventional amplifier because of thermal and shot noise inherent in semiconductor devices. Parametric amplifiers more closely approach the ideal and are thus useful in the front end of receivers because their low noise increases the detection range by several factors.

4.7.2 Usual applications. The voltage-variable capacitance diode is useful from low audio to ultra high frequencies and from -65 °C to +150 °C. It is useful in automatic frequency control, voltage tuning, frequency modulation, and harmonic multiplication. An example of a high-power harmonic multiplier application is provided in Figure 85.

In the circuit, an input of 50 W at 50 MHz is converted to an output of 32.5 W at 150 MHz. This is accomplished by the varactor diode which builds up a waveform with flattened positive peaks because its capacitance is high at positive and low-negative bias; also, the negative peaks are exaggerated because the capacitance is low at high reverse bias. This distorted wave contains strong harmonics which can be picked up by the output tuned circuit, which in this case is tuned to the third harmonic.

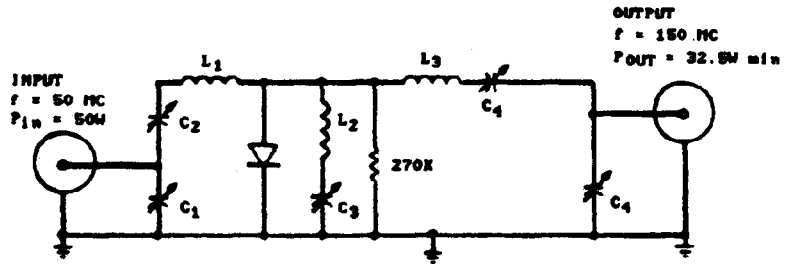


FIGURE 85. 50 to 150 MHz tripler test circuit.

In the design of harmonic multipliers, lumped circuit techniques are useful up to 450 MHz with little performance degradation provided coil and capacitor "Q" values of 200 to 300 are maintained.

Above 450 MHz, coaxial, stripline, or helical coil resonators are recommended. Component values are not particularly critical; however, excessive inductance or insufficient coupling can cause low efficiency, and insufficient inductance or excessive coupling can cause poor filtering. Simple experimentation with well constructed and shielded breadboards is generally sufficient for circuit optimization. Note that an adequate tuning range must be provided to insure input match over normal voltage variable capacitance diode variations, and that spurious signals between stages should be kept below 30 dB by suitable filter circuits.

In typical applications, doubling efficiency is 5 percent greater than that for tripling, and, quadrupling efficiency 5 percent less than that for tripling.

4.7.3 Physical construction. The voltage-variable capacitance diode is generally made with a passivated diffused process die in a glass-to-metal package. The basic die process is shown in Figure 86.

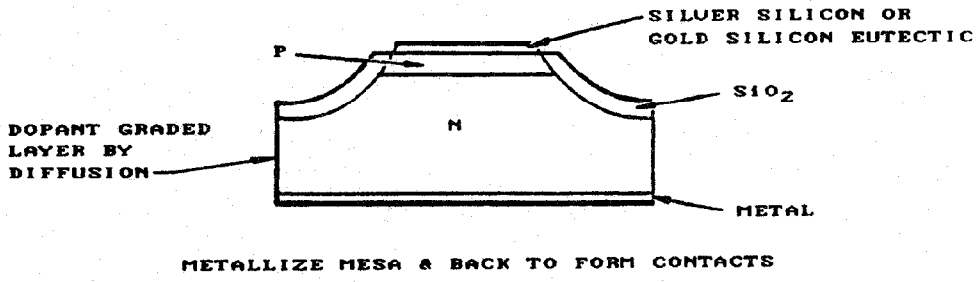


FIGURE 86. Passivated diffused process.

The following general description of construction characteristics is typical of most voltage-variable capacitance diodes. The high power varicap incorporates the high Q varicap crystals in a high power package. The alloyed abrupt junction gives the user a device which closely approaches the theoretical formula $C = 1/\sqrt{V}$ for the capacitance versus voltage curve. The capacitance nonlinearity of the high power varicap can be defined by the following formula:

$$C_V = (C - C_{pkg}) \cdot \left(\frac{4 + V_O}{V + V_O}\right)^n + C_{pkg}$$

where

voltage stock sensitivity (n) = 0.48; package capacitance = 1.5 pF, C = capacitance at -4 reverse voltage at -4 V barrier potential (V_O) = 0.7 @ 25 °C (TC of $V_O \cong 2$ mV/°C).

4.7 DIODES, VOLTAGE VARIABLE CAPACITOR

The high-power varicap has broad area contacts which are metallurgically bonded at 320 °C. These positive contacts and surface passivation on both sides of the crystal give the high-power varicap its increased power handling capabilities and high reliability characteristics.

As previously noted, the above typical physical characteristics are for reference only. For details of the specific device selected, consult the applicable MIL-S-19500 specification sheet.

4.7.3.1 Back contact. Illustrated in Figure 87 is the solder back contact technique used in the construction of voltage-variable capacitor diodes.

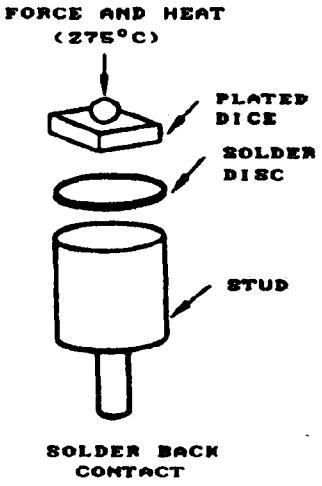


FIGURE 87. Back contact formation techniques.

The solder contact has good mechanical strength and thermal conductivity. The stud material is generally molybdenum, tungsten, or dumet because their coefficients of expansion closely match those of glass.

4.7.3.2 Front contact. Electrical contact is made to the die in two ways as shown in Figure 88.

The S bend and wire whiskers are held under compression in contact with the top of the die. In this way, electrical contact is made to the die while the force on it is limited by the spring constant of the whisker.

The whisker contacts have good shock and vibration tolerance. The wire whisker is somewhat superior in this respect due to its lower mass and the imbedding of the whisker in the metal on top of the die, which prevents sliding of the

contact. Both whisker structures provide good mechanical isolation of the die from mechanical strain induced by lead flexure and thermal expansion and contraction of the package.

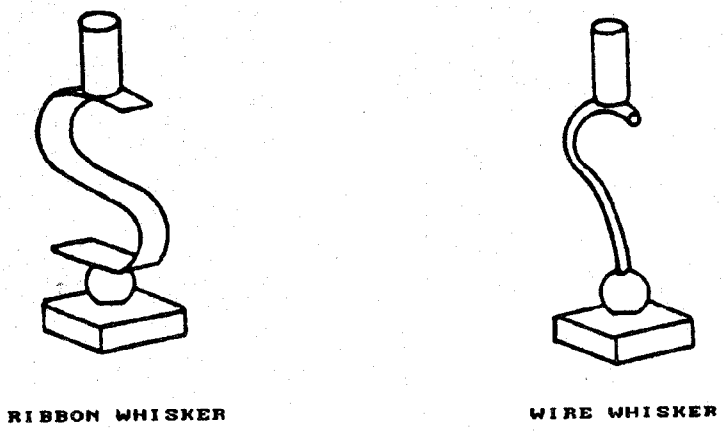


FIGURE 88. Two styles of electrical contact.

4.7.3.3 Seals. Figure 89 shows the sealing technique for glass-to-metal seals.

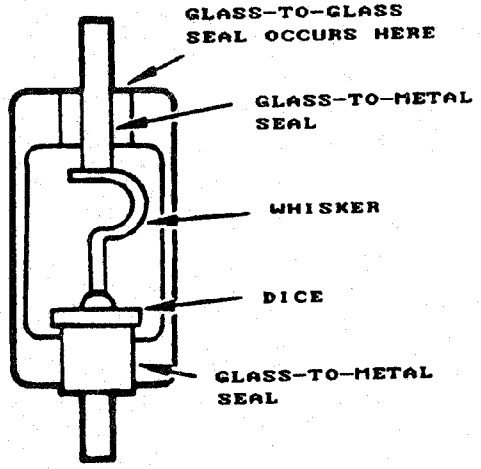


FIGURE 89. Glass-to-metal seal technique.

The sealing technique generally used for a voltage variable capacitor diode is a glass-to-metal seal. The glass-to-metal seals in the above package are formed prior to final sealing, which is a glass-to-glass seal.

4.7 DIODES, VOLTAGE VARIABLE CAPACITOR

Figure 90 below shows a typical voltage-variable capacitor diode package.

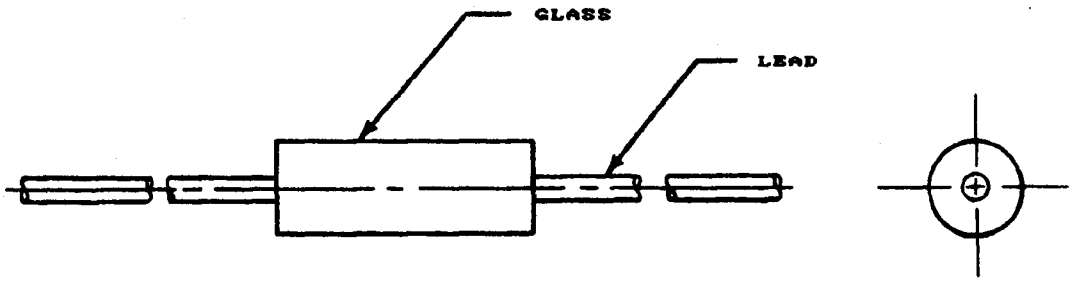


FIGURE 90. A typical voltage-variable capacitor diode package.

4.7.4 Military designation. The military designation for diodes is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification component	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.7.5 Electrical characteristics. Table VII illustrates typical electrical characteristics of voltage-variable capacitors (60-V series).

The values listed in Table VII are typical and for general reference only. For the specific device selected, consult the applicable MIL-S-19500 specification sheet. Diodes are listed by 1N numbers. Refer to subsection 4.1 Diodes, general for definition of electrical terms.

In considering the values of Table VII, reverse current, reverse voltage, capacitance and "Q" (quality factor) are significant parameters in determining performance of the voltage-variable capacitance diode. Of these parameters, the quality factor or Q is perhaps the most important, hence the following discussion is presented to further describe its significance.

Table VII. Voltage-variable capacitance diodes

Device Type	Cap. at V_R pF	Vdc	Cap. Ratio from Min.	I_R (μA)	P_T (max)	Q (Min)
1N5139A	6.8	4	2.7	10	400	350
1N5140A	10.0	4	2.8	10	400	300
1N5141A	12.0	4	2.8	10	400	300
1N5142A	15.0	4	2.8	10	400	300
1N5143A	18.0	4	2.8	10	400	250
1N5144A	22.0	4	3.2	10	400	200
1N5145A	27.0	4	3.2	10	400	200
1N5146A	33.0	4	3.2	10	400	200
1N5147A	39.0	4	3.2	10	400	200
1N5148A	47.0	4	3.2	10	400	200

NOTE: These parts are not included in MIL-STD-975. This table is not intended to be a part selection list.

The equivalent circuit of a reverse biased diode in the microwave bands can be represented as shown in Figure 91.

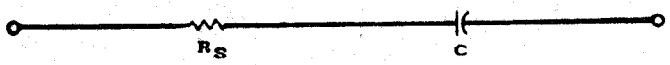


Figure 91. Equivalent circuit.

Where R_S is the series resistance or the spreading resistance of the diode due to the bulk resistivity of the semiconductor material; C is the transition capacitance of the reverse biased junction and is a function of the bias voltage.

Because a small dissipated power in the diode compared with the stored energy is desired, the usual concept of the figure of merit, Q, can be used to characterize the quality of the voltage-variable capacitance diode. By definition:

$$Q = 2\pi \frac{\text{(Average energy stored per cycle)}}{\text{(Average energy lost per cycle)}} \equiv \frac{\text{Reactive power}}{\text{Active power}}$$

4.7 DIODES, VOLTAGE VARIABLE CAPACITOR

For a series RC circuit this becomes:

$$Q = \frac{1}{2\pi f RC}$$

where f is the frequency at which the Q is measured.

A few notes pertinent to this formula when it is applied to characterize voltage-variable capacitance diodes are as follows:

a. It is a standard practice to measure microwave capacitance diodes at 1 MHz either at the voltage breakdown or at zero bias.

b. The $Q \cong 1/f$ simple relationship is not followed entirely by these diodes. The deviation increases as the frequency decreases. This fact makes it difficult to extrapolate the Q to other frequencies.

c. It must be made clear whether or not the influence of cartridge capacitance is included in the quoted Q of a diode. The inclusion of the cartridge capacitance increases the Q, thereby giving false information about the quality of the junction capacitance.

The Q of the best quality voltage-variable capacitance diodes when measured under standard conditions (i.e., 1 MHz, at V_{BR}, junction capacitance only) are between 10 and 16.

C_{min} means the minimum junction capacitance; i.e., the capacitance measured at the breakdown voltage.

4.7.6 Environmental considerations. Typical environmental conditions and screening tests which voltage variable capacitance diodes can withstand are not substantially different from those given in paragraphs 4.1.6.9 Environmental conditions and 4.1.6.10 Screening procedures in subsection 4.1 Diodes, general. For specific devices selected consult the applicable MIL-S-19500 reference slash sheets.

4.7.7 Reliability considerations.

4.7.7.1 Failure modes. Recent development of the JANTXV varicap specification for the IN5139A through 5148A has done much to reduce the incidence of failures. Particular attention should be given to the merits of a reverse bias bake at elevated temperatures. This test screens out diodes with leakage paths on junction surface areas which contribute greatly to many operational life failures.

4.7.7.2 Derating. History has shown that the largest single cause of failure of diodes is operation above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be performed to increase the reliability of the system. Users should refer to MIL-STD-975 for specific derating factor guidelines on diodes.

4.8 Bipolar transient voltage suppressors (TVSs).

4.8.1 Introduction. Transient voltage suppressors (TVSs) consist of two zener diodes oriented back-to-back and sealed in a package that provides adequate thermal dissipation. They exhibit extremely fast response times, low series resistance (R_{ON}), and very high surge voltage handling capabilities. Unlike a zener diode, whose function is voltage regulation, a transient voltage suppressor is designed to protect high voltage sensitive circuits by suppressing transient voltages.

4.8.2 Usual applications. Transients are a major concern of designers. They result from a variety of reasons. The most common are normal switching operations, power supply switching, and circuit disturbances caused by load switching, magnetic coupling and voltage spikes. Voltage transients are a major cause of component failures. Random voltage transient spikes can damage voltage sensitive devices such as ICs, hybrids and MOS devices. Because TVSs exhibit a fast response time and low clamping factor, they can protect these voltage sensitive devices.

TVSs have been used effectively in mobile communication equipment, computer power supplies, airborne avionics, and other applications where inductive and switching transients are present.

When choosing a TVS for a particular application, the following important factors should be considered:

a. The maximum clamping voltage (V_C) should be determined in order to provide adequate protection for a circuit or component. Once V_C has been determined, it will be used to calculate the power for worst case designs for a given application.

b. The TVS selected should exhibit a reverse stand-off voltage (V_R) equal to or greater than the circuit operating voltage (maximum ac or dc peak voltage with tolerances).

c. To select the appropriate TVS one must also determine the transient pulse power (P_P). This can be accomplished by using the simple definition; transient pulse power (P_P) equals the peak pulse current (I_{PP}) multiplied by the clamping voltage (V_C).

$$P_P = V_C \times I_{PP}$$

4.8 DIODES, BIPOLAR TRANSIENT VOLTAGE SUPPRESSORS (TVSs)

4.8.2.1 Microprocessor system TVS applications. The TVS is placed on the signal and input power lines to prevent system failures caused by the effect of switching power supplies, ac power surges and transients, such as electrostatic discharges as illustrated in Figure 92. A TVS across the signal line to ground will prevent transients from entering the data and control buses. TVSs shunted across the power lines ensure a transient-free operating voltage.

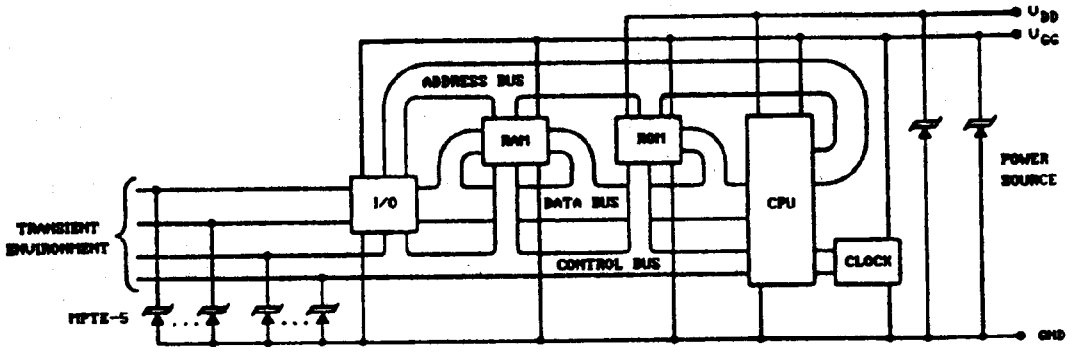


FIGURE 92. Microprocessor system TVS application.

4.8.2.2 DC line TVS applications. A TVS in the output of a voltage regulator can replace many components used as protection circuits such as the crowbar circuit illustrated in Figure 93. It may also be used to protect the bypass transistor from voltage spikes across the collector to emitter terminals.

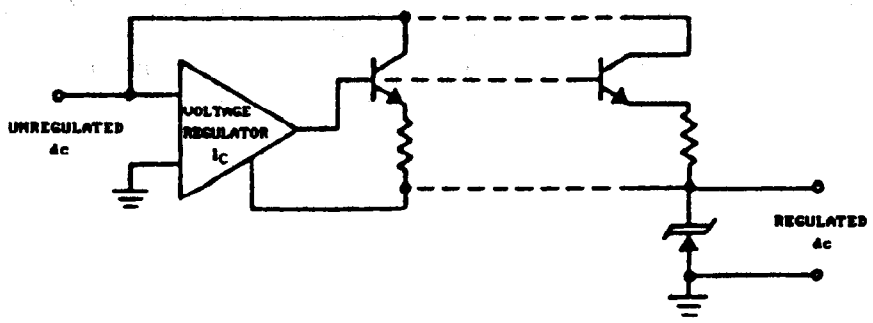


FIGURE 93. DC line TVS application.

4.8.2.3 Relay and solenoid TVS applications. The coil inductance of a solenoid or relay can release energy that can damage contacts or drive transistors. A TVS used as shown in Figure 94 would provide adequate protection.

The proper TVS can be selected by determining peak pulse power (Pp) and pulse time (t_p). Knowing the values of V_{CC}, L, and R_L the following equations can be used to determine Pp and tp.

$$I_0 = \frac{V_{CC}}{R_L}$$

$$Pp = Ip \times V_C$$

$$t_1 = \frac{V_{CC}/R_L}{V_C/L}$$

$$t_p = \frac{t_1}{2}$$

Figure 95 shows peak pulse power, Pp versus pulse duration, tp for a given TVS device.

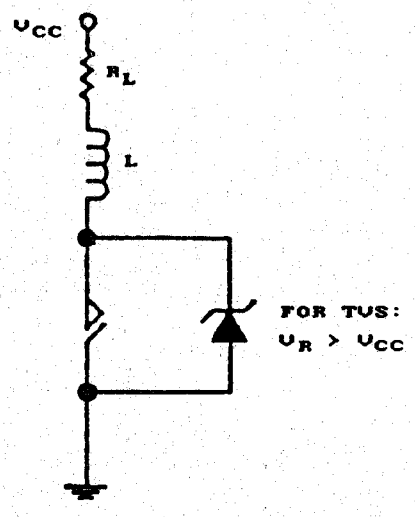
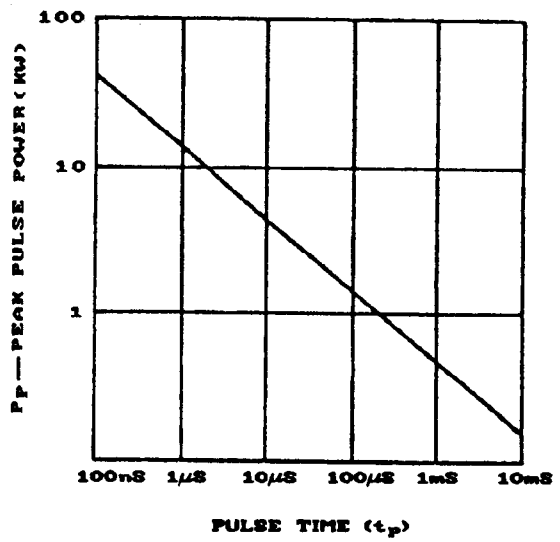


FIGURE 94. Coil and contacts, dc.



Max. duty cycle = 0.1%.

FIGURE 95. Peak pulse power vs pulse duration.

The applications described herein are but a few of the many possible uses of the transient voltage suppressor.

4.8.2.4 Protecting switching power supplies. A designer needs to protect against three types of transients: load transients, line transients, and internally generated transients. Because transients have high energy levels they cause improper operation and component failure. Protecting a switching power supply is an example of utilizing TVS devices. The following are typical components that need to be protected in a switching power supply:

a. High voltage switching transistors

b. Rectifiers

c. Output rectifiers

d. Control circuitry.

Figure 96 shows a typical switching power supply with TVS devices used for protection in voltage sensitive areas.

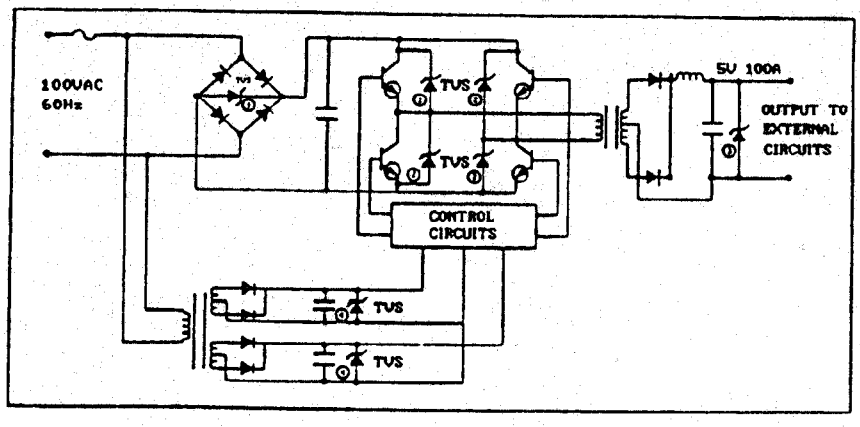


FIGURE 96. Typical switching power supply.

4.8.3 Physical construction. The transient voltage suppressor diode is generally made using two passivated diffused, planar or diffused planer junction on an epitaxial substrate process die placed back-to-back in a single glass-to-metal or double-slug package. The three basic die processes are passivated diffused, planar, and diffused planar junction on epitaxial substrate, as shown in Figures 97, 98, and 99, respectively.

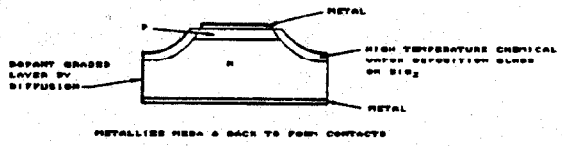


FIGURE 97. Passivated diffused process.

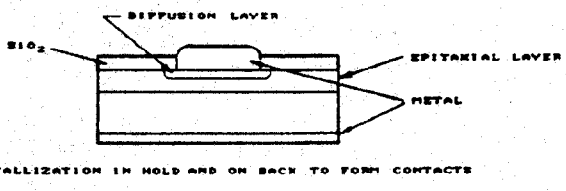


FIGURE 98. Planar process.

4.8 DIODES, BIPOLAR TRANSIENT VOLTAGE SUPPRESSORS (TVSs)

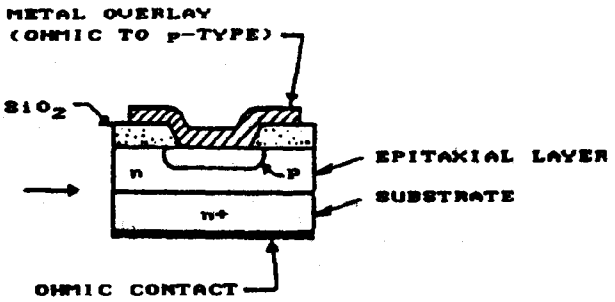


FIGURE 99. Diffused planar junction on epitaxial substrate.

4.8.3.1 Back contact. Illustrated in Figure 100 below is the back contact technique used in constructing transient voltage suppressor (TVS) diodes. TVS diodes generally use high temperature alloy back contacts, associated with double-slug construction.

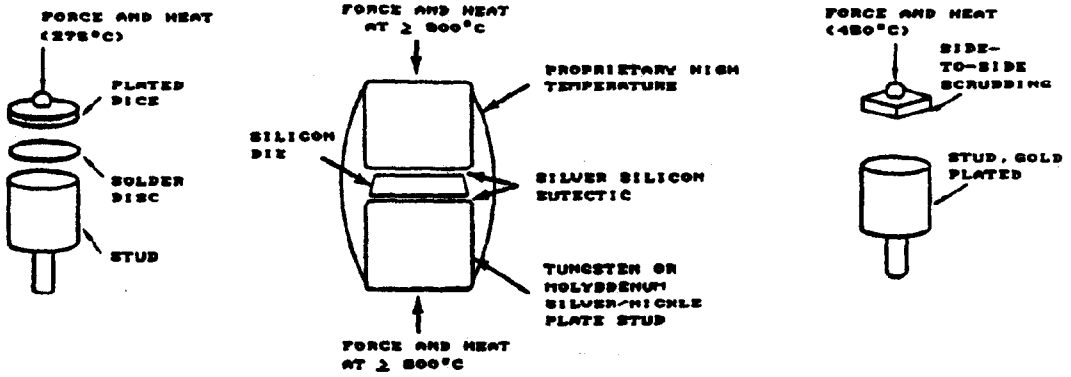
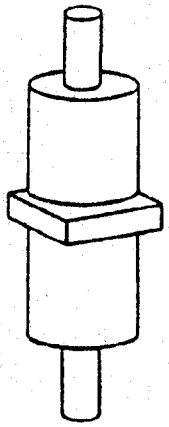


FIGURE 100. Back contact formation techniques.

The silicon-silver alloy contact which is generally associated with the high temperature alloy has good mechanical strength and thermal conductivity. The silicon chip is metallurgically brazed to plated molybdenum or tungsten slugs at temperatures greater than 700 °C.

Molybdenum, tungsten, and dumet are used for studs in glass packages because their thermal expansion coefficients match those of certain glasses, allowing sealing of these glasses to them.

4.8.3.2 Front contacts. Electrical contact is made to the die as shown in Figure 101.



DOUBLE-SLUG CONTACT

FIGURE 101. Various styles of electrical contact.

The double-slug contact provides good heat transfer, thus raising the power rating of the package.

4.8.3.3 Seals. The double-slug package in Figure 102 is made by metallurgical bonding (700 °C) both sides of the die to either molybdenum or tungsten slugs to which a glass sleeve or glass slurry is fused for final sealing. The final step is a lead brazing operation. This type of construction is most often associated with transient voltage suppressor diodes.

4.8 DIODES, BIPOLAR TRANSIENT VOLTAGE SUPPRESSORS (TVSs)

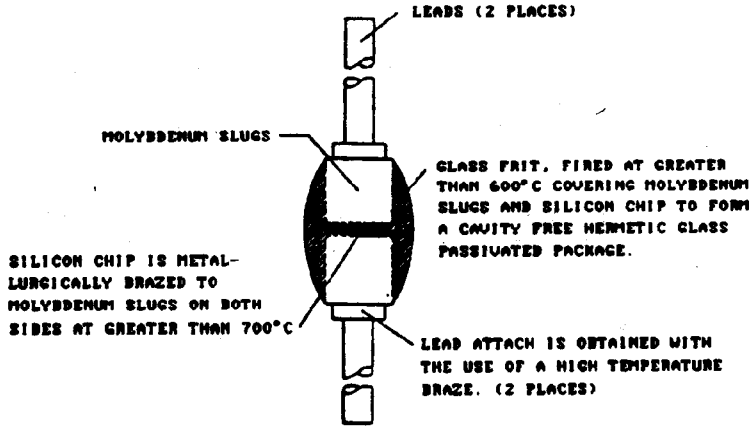


FIGURE 102. Sealing techniques.

Figure 103 shows a typical transient voltage suppressor package.

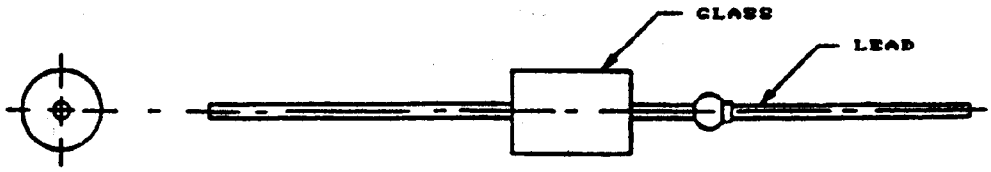


FIGURE 103. Typical transient voltage suppressor (DO-13 package).

4.8.4 Military designation. Military designation for diodes is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R) or any other letter to indicate a modified version.

4.8.5 Electrical characteristics.

a. Voltage-ampere characteristics. The volt-ampere characteristic for a typical transient voltage suppressor, TVS, given in Figure 104 shows that the device can conduct current in both directions. Characterized by its two zener die placed back to back, the TVS is useful for bi-directional transient suppression.

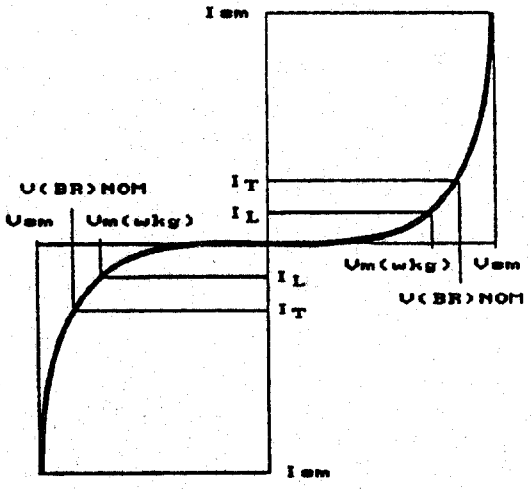


FIGURE 104. Typical characteristic curve for bi-directional transient suppressor.

b. Working peak voltage (VM, WKG) or stand-off voltage (V_R). This parameter is the maximum permissible dc working voltage. It is the highest reverse voltage at which the TVS will be nonconducting.

c. Maximum peak surge voltage (VSM) or maximum clamping voltage (V_{Cmax}). This is the maximum voltage drop across the TVS while it is subjected to the peak pulse current, usually for 1 ms.

d. Minimum breakdown voltage (BVmin). The breakdown voltage is the reverse voltage at which the TVS conducts. This is the point where the TVS becomes a low impedance path for the transient.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.8 DIODES, BIPOLAR TRANSIENT VOLTAGE SUPPRESSORS (TVSs)

e. Test current (I_T). This current is the zener current at which the nominal breakdown voltage is measured.

f. Maximum leakage current (I_L). This current is the current leakage measured at the maximum dc working voltage (VM or VR).

g. Maximum peak surge current (ISM). This current is the maximum permissible surge current.

4.8.6 Environmental considerations. Typical environmental conditions and screening tests which transient voltage suppressors are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 Screening procedures in subsection 4.1 Diodes, general. For the specific device selected, consult the MIL-S-19500 reference slash sheet.

4.8.7 Reliability considerations.

4.8.7.1 Failure mechanisms and data. In extensive tests of transient voltage suppressors, the three most critical parameters were peak pulse power, peak surge voltage and reverse leakage current. Because of its inherent use as a transient suppressor, the TVS must sustain high voltage pulse power. Opens or shorts caused by cracking of the die are main-failure modes if peak reverse power rating is exceeded. Long term life tests have yielded a significant number of failures in these areas. Failure mechanisms of TVSs are similar to those of other diode types discussed in paragraph 4.1.6 General reliability considerations in subsection 4.1 Diodes, general.

4.8.7.2 Screening. As with previous recommendations for other diode types, the transient voltage suppressor procured as a JANTXV or JANS device offers the best protection against the above problems. The two characteristics which are usually measured to verify TVS quality are the breakdown voltage (BV) and reverse leakage current (I_R). Measurement of these characteristics during burn-in screening will verify the reliability of the diode. The magnitude of reverse leakage current is largely dependent upon surface conditions, die size, thickness, and junction area of the semiconductor. Delta measurements of reverse leakage during burn-in is a useful criterion in determining the long term reliability of the device.

4.8.7.3 Derating. History has shown that the largest single cause of failure of TVSs is operation above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be performed to enhance reliability. Users should refer to MIL-STD-975 for specific derating factor guidelines. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress a semiconductor device could be subjected to. It will result in electrical parameter drift and general degradation of the electrical and mechanical characteristics of the device.

4.9 Monolithic and multiple arrays.

4.9.1 Introduction. Diode arrays consist of several diode circuits fabricated in an integrated circuit form. The diode array is not a unique device with characteristics that differ from discrete diodes. It is simply a packaging technique that permits a significant reduction in the size of electrical systems. The arrays are available in either 10- or 14-ceramic flat packs or ceramic 14- and 16-lead dual-in-line packages, for high density packaging and reliability.

Although not complex in operation, these arrays are widely used in applications requiring moderate current, fast switching, or digital logic operations.

The military designation, electrical characteristics, environmental conditioning and reliability of diode arrays are similar to those of their discrete diode counterpart.

4.9.2 Usual applications. Digital logic circuits commonly use diode arrays that can be used to construct logic gates. Figure 105 is an example of such a circuit; it is referred to as an AND gate.

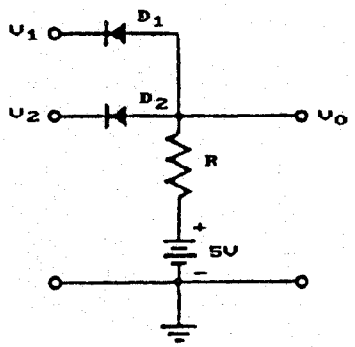


FIGURE 105. Diode logic gate, AND gate.

In the operation of an AND gate, all the signals are binary digital signals and have only one of two possible states (low and high). In Figure 105, the output voltage (V_0) becomes high only if V_1 and V_2 are high. However, if either of these voltages are zero, the corresponding diode is forward-biased, and V_0 becomes short-circuited by the diode and becomes zero (low state).

4.9 DIODES, MONOLITHIC AND MULTIPLE ARRAYS

Diode arrays are also used to construct read only memories (ROMs) where the data to be stored is in the form of a table. These are called "look-up tables." See Table VIII, which shows an e^{-x} data table using 4-bit words.

The e^{-x} look-up table can be implemented by using a 4- to 16-line demultiplexer such as a 74154 silicon microcircuit. This device responds to a 4-bit binary input by making its appropriate output low. To convert the 74154 into a ROM, diode arrays are connected as shown in Figure 106. When a given output line goes low and a diode is connected from V_{CC} to the output line of the 74154, the output bit goes low because the diode conducts. On the other hand, if a diode is not connected from V_{CC} to the 74154 output line, that output bit remains high. The diode array forms the ROM portion of the circuit.

TABLE VIII. e^{-x} data table

x	Binary	e^{-x}	Binary	Decimal Equivalent
0.0	0000	1.00	1111	15
0.1	0001	0.90	1110	14
0.2	0010	0.82	1100	12
0.3	0011	0.74	1011	11
0.4	0100	0.67	1010	10
0.5	0101	0.61	1001	9
0.6	0110	0.55	1000	8
0.7	0111	0.50	1000	8
0.8	1000	0.45	0111	7
0.9	1001	0.41	0110	6
1.0	1010	0.37	0110	6
1.1	1011	0.33	0101	5
1.2	1100	0.30	0101	5
1.3	1101	0.27	0100	4
1.4	1110	0.25	0100	4
1.5	1111	0.22	0011	3

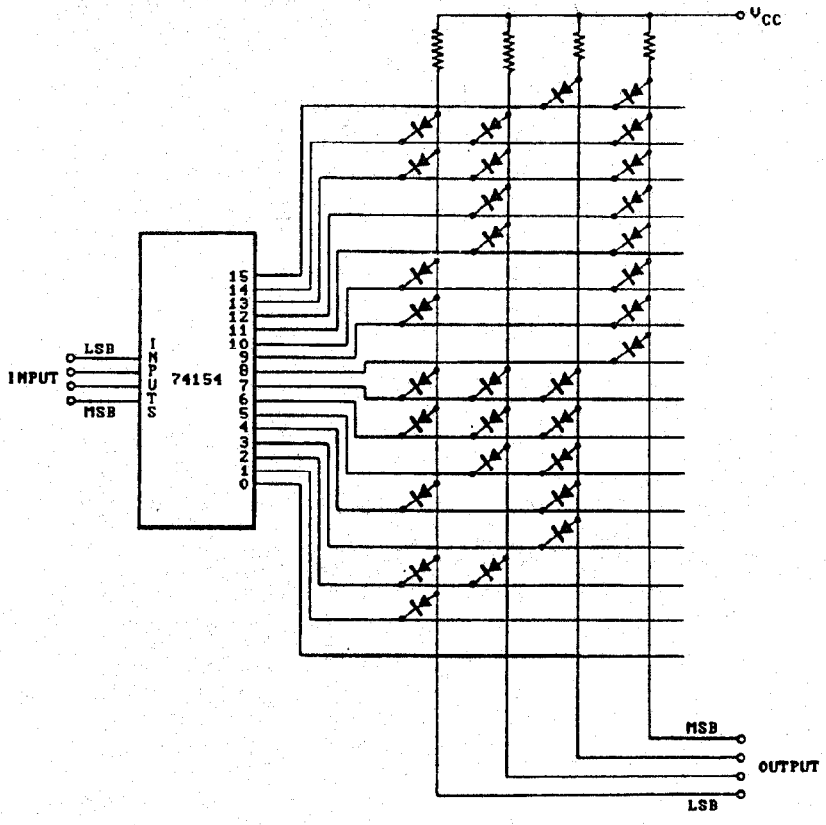
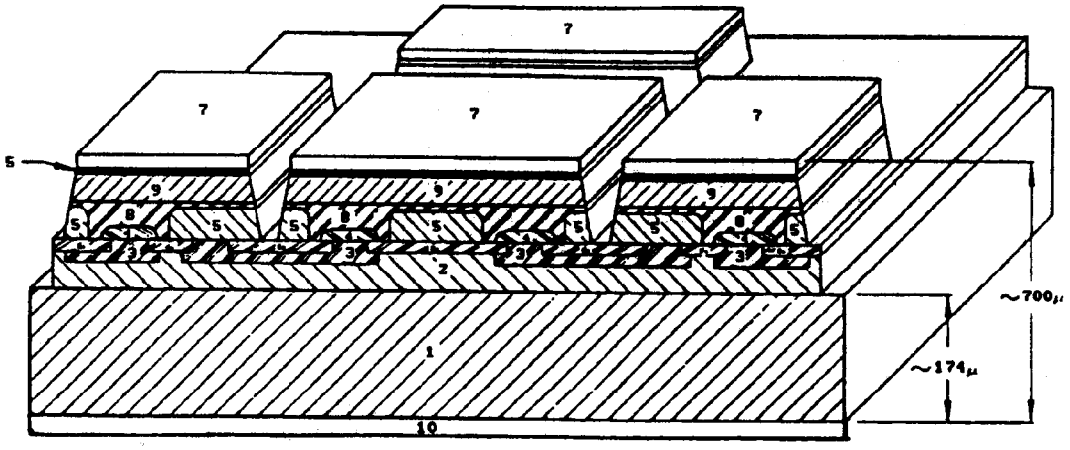


FIGURE 106. A ROM synthesized from a 74154 demultiplexer and the appropriate diode connections as a look-up for exponentiation (see Table XI).

4.9.3 Physical construction. All air-isolated monolithic planar or multiple diode arrays use the same fabrication process to facilitate the manufacture of diode arrays in integrated circuit form. The elimination of parasitic capacitance (which is a very important consideration in overall switching speed) is accomplished through air isolation of the diode circuits. Air isolation is obtained by selective chemical etching of the whole device wafer from the backside of the silicon wafer as shown in Figure 107.

4.9 DIODES, MONOLITHIC AND MULTIPLE ARRAYS

The process used by the manufacturer is a very general one, capable of fabricating a wide range of diode arrays. The process supplies a hermetically sealed structure at the chip level, but of integrated circuit size as compared with discrete diodes with similar electrical characteristics. The excellent hermeticity is the result of using silicon nitride over the device structure in conjunction with a unique fused glass sandwich construction. A physical diagram representing a typical diode array is shown in Figure 107, with Arabic numerals identifying the materials used in its construction.



1. Support wafer (silicon backing wafer)
2. Glass dielectric layer
3. Aluminum Interconnects (3 μ thick)
4. P^+ Junction
5. N^{++} Cathode contact (channel stops)
6. Passivation
7. Aluminum bonding pads (6-μ thick)
8. N type epitaxial layer
9. N^+ substrate
10. Gold film backing for mounting

FIGURE 107. Typical construction of an air isolated monolithic or multiple diode array.

In the process two silicon wafers are used, one containing the diffused junction configuration desired for the device, and the other the mechanical support during hot processing. The passive wafer contains no device structures, nor is it called upon to provide any electrical contact function. All contacts to the devices are provided from the "top" side of the finished chip, which is the "bottom" side of the active device wafer. The two wafers are sandwiched and sealed together using a fused glass frit between the wafers. Prior to sandwich sealing, the diffused side of the device wafer is first metallized with aluminum to provide the necessary interconnects between junctions, and it is subsequently coated with a glass powder slurry. The second mechanical support wafer is then applied to the device wafer and the two are sealed together under moderate compressive force by heating both wafers to an elevated temperature at which the glass frit fuses to form a glass seal between the two wafers. The assembled "sandwich" is then etched using an aluminum metal masking pattern to produce localized mesas. The mesas are inverted relative to the original diffusion orientation. These metallized mesa tops provide bonding pads for termination of the chip. A nitric acid rich, oxidizing, silicon-etch solution containing hydrogen fluoride (HF) is used for this purpose. The oxidizing nitric acid-rich etchant insures no attack on the aluminum back contact mask pattern. The mesa etching is then stopped when it "bottoms out" at the silicon nitride layer.

In this way, the manufacturer is able to create a wide range of diode device arrays using this process without resorting to material other than an N/N^+ epitaxial device wafer. Therefore, a great variety of diode array structures may be produced by judicious selection of mesa etch masking and back contact configurations. The flexibility of the process in producing common anode and common cathode diode configurations is demonstrated in the generalized schematic device cross section shown in Figure 107.

For cathode contacts, an ultrasonic bond is made using a 1.5 mil aluminum wire connecting the chip bond pad areas (item 7 in Figure 107) to the lead posts inside the package. For anode contacts, a similar aluminum-to-aluminum ultrasonic bond is made to the aluminum bonding pads on the glass dielectric (item 2, Figure 107) which are exposed after etching. Pad configuration is not shown.

The diode arrays come in various hermetically sealed ceramic flat packages as illustrated in Figures 108 and 109.

4.9 DIODES, MONOLITHIC AND MULTIPLE ARRAYS

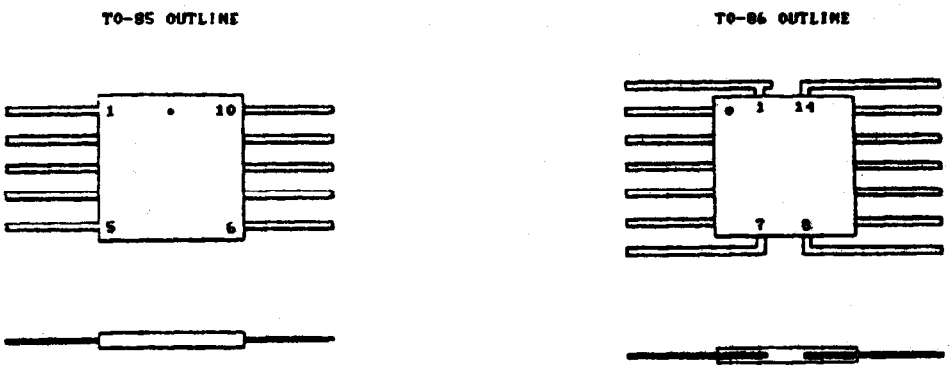


FIGURE 108. Package configurations for multiple arrays.

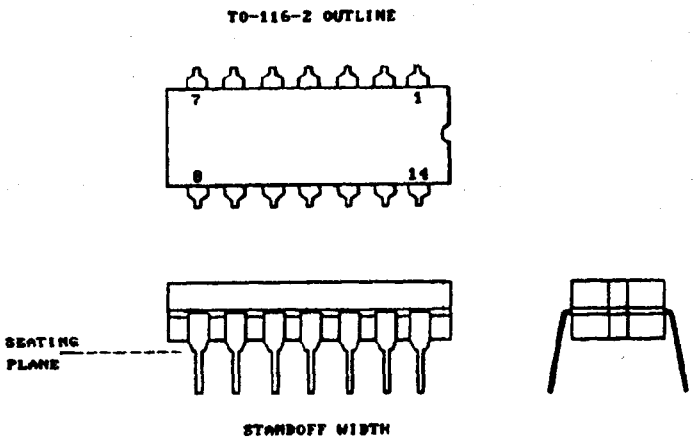


FIGURE 109. Package configuration for monolithic array.

4.9.4 Military designation. Military designation for diodes is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

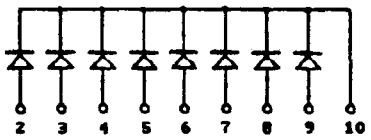
4.9.5 Electrical characteristics.

4.9.5.1 Multiple arrays. The basic diode arrays are the 1N5768, 1N5770, 1N5772 and 1N5774. These arrays consist of four different diode circuit configurations

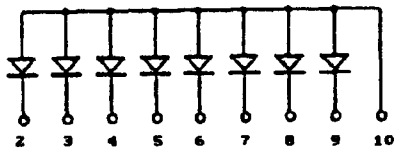
which come in either a TO-85 or TO-86 hermetically sealed ceramic flat package, as outlined in Figure 110.

Basic device characteristics and parameter information on diode arrays are similar to those discussed in 4.1 section, paragraph 4.1.5 General parameter information in subsection 4.1 Diodes, general.

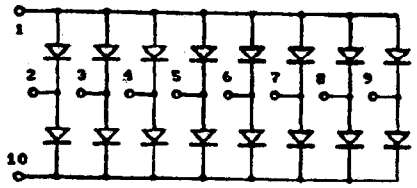
4.9 DIODES, MONOLITHIC AND MULTIPLE ARRAYS



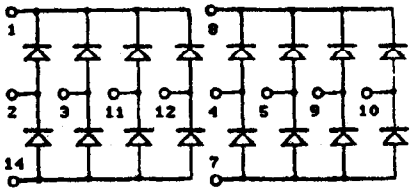
PACKAGE OUTLINE TO-85



PACKAGE OUTLINE TO-85



PACKAGE OUTLINE TO-85

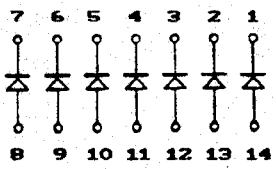


PACKAGE OUTLINE TO-86

FIGURE 110. Multiple diode array circuit configurations.

4.9.5.2 Monolithic arrays. The 1N6100 and 1N6101 devices are functionally identical but come in different package configurations as shown in Figure 111. Using the same manufacturing process as outlined under paragraph 4.9.3 Physical construction, the diode array can be wired for the application at hand while conserving valuable space on the printed circuit board. Also, the reverse recovery time (max) is 5 ns and not 20 ns as for the 1N5768, 1N5770, 1N5772 and 1N5774 series.

Basic device characteristics and parameter information on diode arrays are similar to those discussed in subsection 4.1, Diode-general and paragraph 4.1.5.1.



PACKAGE OUTLINES

TO-86	1N6100
TO-116-2	1N6101

FIGURE 111. Monolithic diode array.

4.9.6 Environmental considerations. Typical environmental test conditions and screens which switching diode arrays are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental consideration and 4.1.6.10 Screening procedure, is in subsection 4.1, Diodes-general.

4.9.7 Reliability considerations.

4.9.7.1 Failure modes. The statements of paragraph 4.2.7.1 Reliability considerations for switching diodes are applicable to switching diode arrays.

Considering the material design and construction of the diode array (see paragraph 4.9.3 Physical construction), the entire active portion of the device is buried in the glass dielectric layer and is inaccessible by any contaminant once the device is properly wire bonded and packaged.

4.9.7.2 Derating. The derating for voltage, current and forward surge current should be 50 percent of the value published in the vendor data sheet.

Users should refer to MIL-STD-975 for derating factor guidelines.

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

4.10 Silicon controlled rectifiers (SCR)

4.10.1 Introduction. The silicon-controlled rectifier (SCR) belongs to the thyristor family. It is a semiconductor device that can be switched between two states, from off to on, or from on to off, by a current or polarized voltage pulse. Unlike conventional transistors, the device lends itself to use as a high-current high-voltage rectifier, a static latch limited to microseconds, and a sensitive high-gain amplifier control. As shown in Quadrant I of Figure 112, under forward-bias conditions (anode positive with respect to the cathode) the SCR has two states. At low values of forward-bias, the SCR exhibits a very high impedance; in the forward blocking or off state, a small forward current, called the forward off-state current, flows through the device. As the forward-bias is increased, however, a voltage point is reached at which the forward current increases rapidly, and the SCR switches to the on state. This value of voltage is called the breakover voltage. When the SCR is in the on state, the forward current is limited primarily by the impedance of the external circuit.

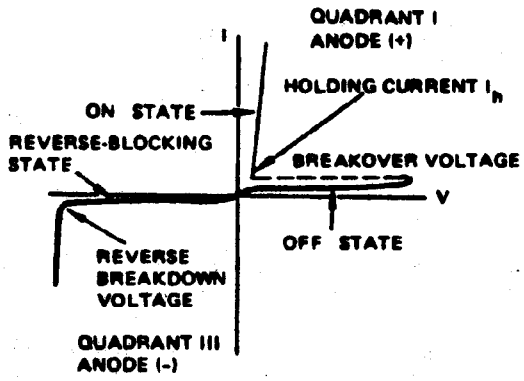


FIGURE 112. Characteristics of a silicon controlled rectifier.

Under reverse bias (anode negative with respect to cathode) as shown in quadrant III of Figure 112, the SCR exhibits a very high internal impedance, and only a small amount of current, called the reverse blocking current, flows through the device. The current remains very small and the device remains "off" unless the reverse voltage exceeds the reverse breakdown voltage limitation. At this point, the reverse current increases rapidly and the SCR undergoes thermal runaway, a condition that normally causes irreversible damage. The value of the reverse breakdown voltage generally is greater than the forward breakover voltage for most types of SCRs.

The breakover voltage of the SCR can be controlled or varied by application of a pulse at the gate electrode. Figure 113 demonstrates that as the amplitude of the pulse is increased, the breakover voltage decreases until the curve

closely resembles that of a rectifier. In normal operation, the SCR is operated with critical values well below the breakover voltage and is made to switch on by gate signals of sufficient magnitude to assure that the device is switched to the "on" state at the desired instant.

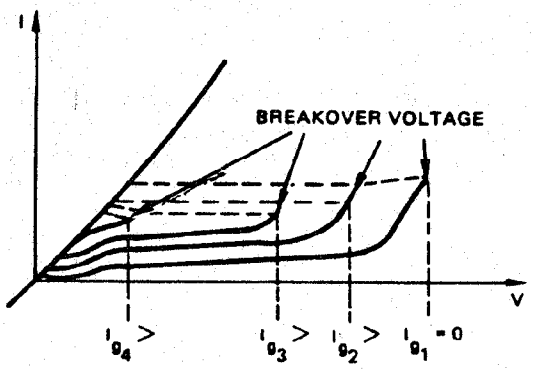


FIGURE 113. Curves showing the forward-voltage characteristics of an SCR for different values of gate current.

After the device is switched on (triggered) by the gate signal, the current through the device is independent of the gate voltage or current. The SCR remains in the on state until the principal current is reduced to a level below that required to sustain conduction.

Basic SCR types are offered with current ratings extending from 0.5 to 2000 A rms with voltage ratings spanning the range of 25 to 3000 V peak. Many specialized SCR types also are available, including high speed inverter SCRs, SCRs for use over very wide temperature ranges, light-activated SCRs, very high voltage SCRs, and SCRs tested to very rigid quality levels for high reliability applications. All graphs and charts are for typical SCR devices and not for a specific device. Reference should be made to the component-detailed specification for particular characteristic parameters.

4.10.2 Usual applications. The silicon-controlled rectifier is well adapted for use as a latching switch or high power gain amplifier. The SCR can be turned on by a momentary application of control current applied to the control gate, while tubes or transistors require a continual on signal. The turn-on time is about 1 µs, and turn-off time about 10 to 20 µs. This latching action can be controlled by signals of only a few microwatts and can switch up to 200 V. With associated transistor triggering circuits, unusual current gain of many billions can be obtained. Because of this advantage there are many applications to which the SCR can be adapted. A few of these applications follow.

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

4.10.2.1 The dc static switch. Figure 114 illustrates a static SCR switch for use in a dc circuit. When a low power signal is applied to the gate of SCR_1, this SCR is triggered and voltage is applied to the load. The right hand plate of C charges positively with respect to the left hand plate through R_1.

When SCR_2 is triggered, capacitor C is connected across SCR_1, so that this SCR is momentarily reverse biased between anode and cathode. This reverse voltage turns SCR_1 off and interrupts the load current, provided that the gate signal is not applied simultaneously to both gates.

SCR_1 should be selected so that the maximum load current is within its rating. SCR_2 need conduct only momentarily during the turn-off action; it can be smaller in rating than SCR_1. The minimum value of commutating capacitance C can be determined by the following equations:

for resistive load: $C = 1.5t_{off}I/E$ μF

for inductive load: $C = t_{off}I/E$ μF

where

t_{off} = turn-off time of SCR in microseconds

I = maximum load current, including possible overloads, in amperes at time of commutation

E = minimum dc supply voltage

The resistance of R_1 should be ten to one hundred times less than the minimum effective value of the forward blocking resistance of SCR_2. This latter value can be derived from the published leakage current curves for the SCR under consideration.

In some cases, a mechanical switch may be substituted for SCR_2 to turn off SCR_1 when it (the switch) is momentarily closed. Many other useful variations of this basic dc static switch can be devised.

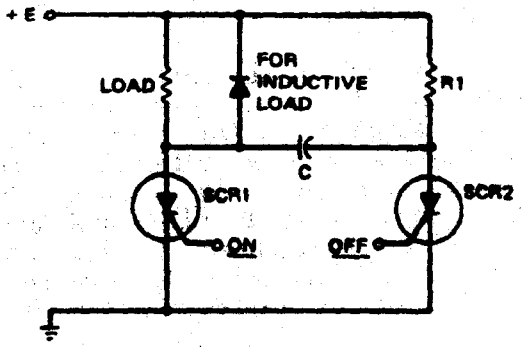
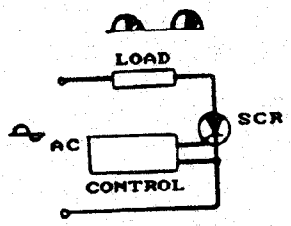


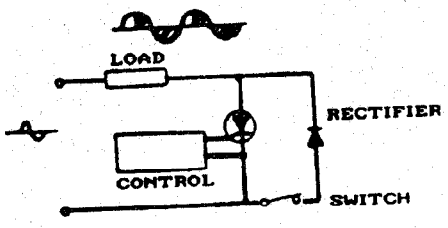
FIGURE 114. The dc static switch.

4.10.2.2 Principle of phase control. Phase Control is the process of rapid on-off switching which connects an ac supply to a load for a controlled fraction of each cycle. This is a highly efficient means of controlling the average power to loads such as lamps, heaters, motors, and dc supplies. Control is accomplished by governing the phase angle of the ac wave at which the SCR is triggered. The SCR will then conduct for the remainder of that half-cycle.

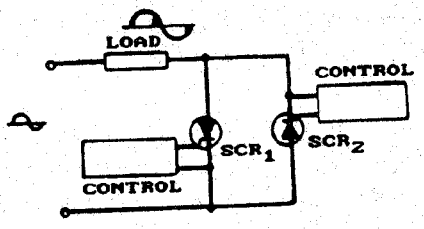
There are many forms of phase control possible with the SCR as shown in Figure 115. The simplest form is the half-wave control of Figure 115A, which uses one SCR for control of current flow in one direction only.



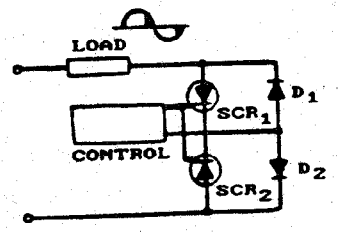
A. Controlled half-wave



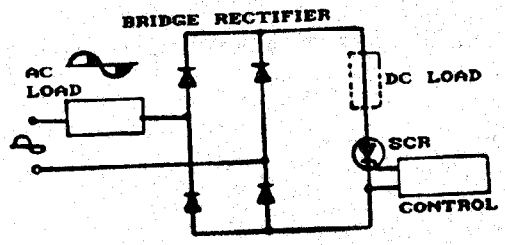
B. Controlled half plus fixed half-wave



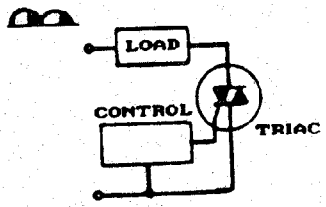
C. Controlled full wave



D. Controlled full wave



E. Controlled full wave for ac or dc



F. Full wave control with triac

FIGURE 115. Basic forms of ac phase control.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

This circuit is used for loads which require power control from zero to one-half of full-wave maximum and permit or require dc. The addition of one rectifier, as shown in Figure 115B, provides a fixed half-cycle of power which shifts the power control range to half-power minimum and full-power maximum but with a strong dc component. The use of two SCRs, shown Figure 115C, provides control from zero to full-power and requires isolated gate signals, either as two control circuits or pulse-transformer coupling from a single control. Equal triggering angles of the two SCRs produce a symmetrical output wave with no dc component.

Reversible half-wave dc output is obtained by controlling the symmetry of the triggering angle.

An alternate form of full-wave control is shown in Figure 115D. This circuit has the advantage of a common cathode and gate connection for the two SCRs. Although the two rectifiers prevent reverse voltage from appearing across the SCRs, they reduce circuit efficiency by their added power loss during conduction.

The most flexible circuit, shown Figure 115E, uses one SCR inside a bridge rectifier and may be used for control of either ac or full-wave rectified dc. Losses in the rectifiers, however, make this the least efficient circuit form and commutation is sometimes a problem. On the other hand, using one SCR on both halves of the ac wave is a more efficient utilization of SCR capacity, hence the choice of circuit form is based on economic factors as well as performance requirements.

4.10.2.3 Inverter configurations. Rectifier circuits occur in several configurations such as half-wave, full-wave, and bridge. Inverter circuits may be grouped in an analogous manner. Figure 116 shows the different types of configurations. Methods of triggering and commutation have been left out for clarity.

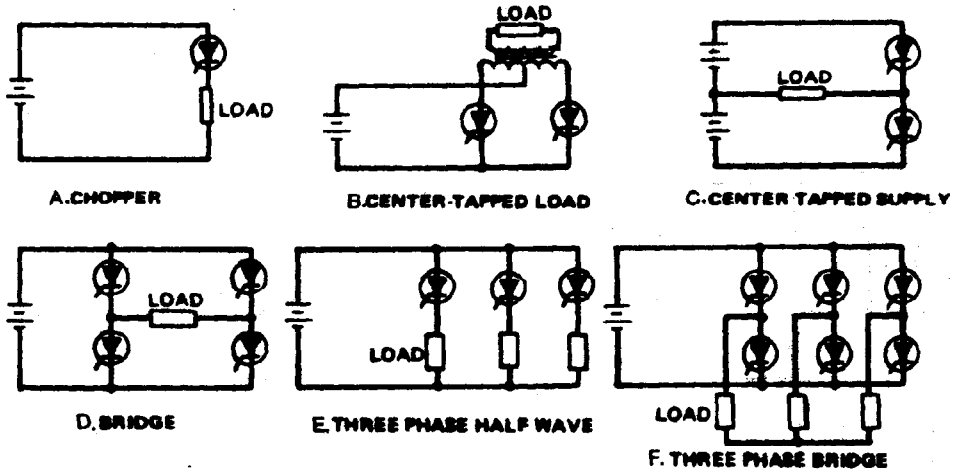


FIGURE 116. Inverter configurations.

4.10.2.4 Pulse modulator switches. The conventional pulse modulator circuit using SCRs as switches operates as a Class A inverter, as shown in Figure 117.

The current pulses through the SCR are narrow, ranging typically from 0.1 to 10 µs base width. The repetition rate is from several hundred to several thousand cycles per second.

Turn-off time and dv/dt are usually not critical characteristics of SCRs for this application. The most critical static characteristic is blocking voltage and the most critical dynamic characteristic is a high value of di/dt. The stress of high values of di/dt may be alleviated by using a saturating reactor in series with the SCR.

To minimize jitter and the delay and rise time, the gate of the SCR should be driven as hard as the ratings permit. The rise time of the gate pulse should be much less than 1 µs.

High frequency SCRs make ideal pulse modulator switches.

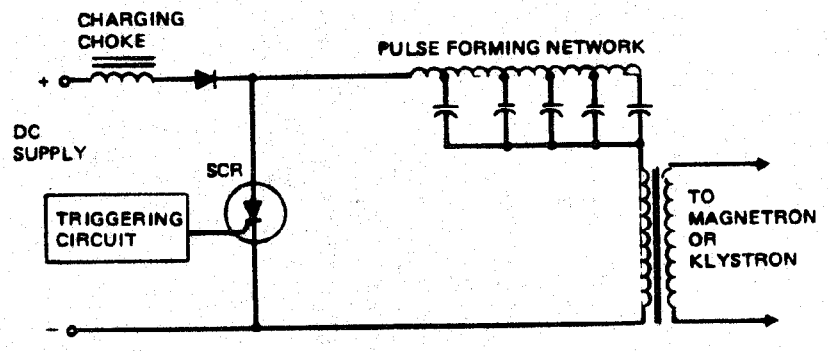


FIGURE 117. Basic pulse modulator circuit.

4.10.3 Physical construction. An SCR is a four-layer p-n-p-n device that has three electrodes, a cathode, anode, and control gate. See Figures 118, 119 and 120 for junction diagrams, typical cross-section views, and typical DO-5 package outline drawing, respectively.

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

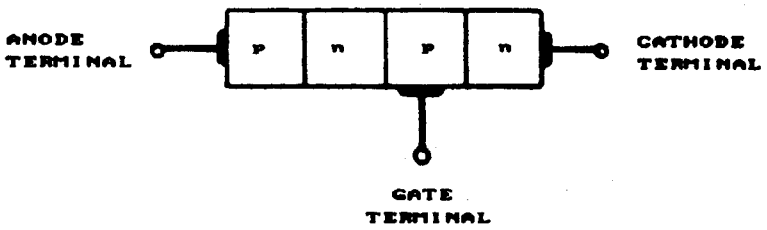


FIGURE 118. Typical p-n-p-n junction diagram.

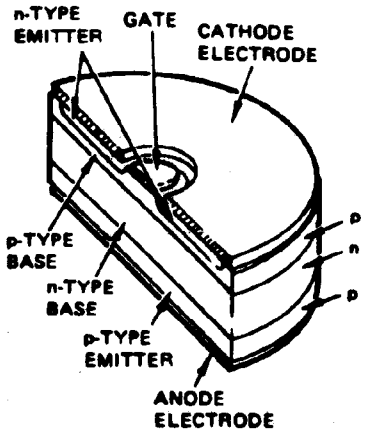


FIGURE 119. Cross-section of a typical SCR pellet.

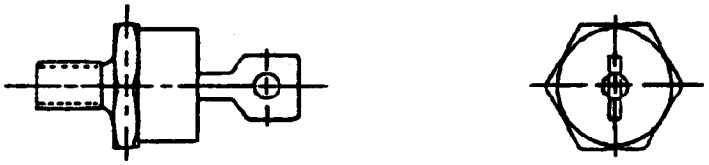
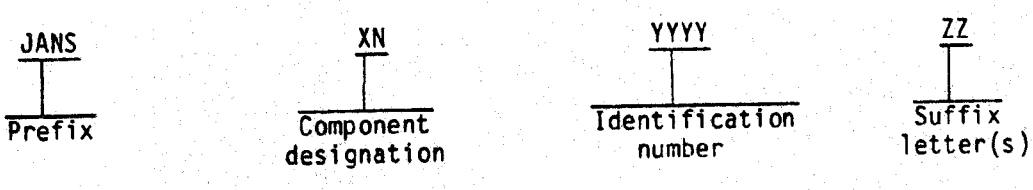


FIGURE 120. Dimensions of a typical SCR (DO-5 package).

4.10.4 Military designation. The military designation for diodes is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, it applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 1N for diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

4.10.5 Electrical characteristics

4.10.5.1 Surge and I^2t ratings, nonrecurrent. In the event that a type of overload or short circuit can be classified as nonrecurrent, the rated junction temperature can be exceeded for a brief instant, thereby allowing additional overcurrent rating. Ratings for this type of nonrecurrent duty are given by the surge current curve and by the I^2t rating.

Figure 121 shows the maximum allowable nonrecurrent surge current at rated load conditions for a high current SCR. The junction temperature is assumed to be at its maximum rated value immediately prior to surge; it is therefore apparent that the junction temperature will exceed its rated value for a short time.

The data shown in this curve are values of peak rectified sinusoidal waveforms (60 Hz) in a half-wave circuit. The one-cycle point, therefore, gives an allowable nonrecurrent half sine wave of 0.00834-second duration (half-period of 60 Hz frequency) of a peak amplitude of 150 A. The 20-cycle point shows that 20 rectified half-sine waves are permissible, separated by equal off times and each of an equal amplitude of 80 A.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

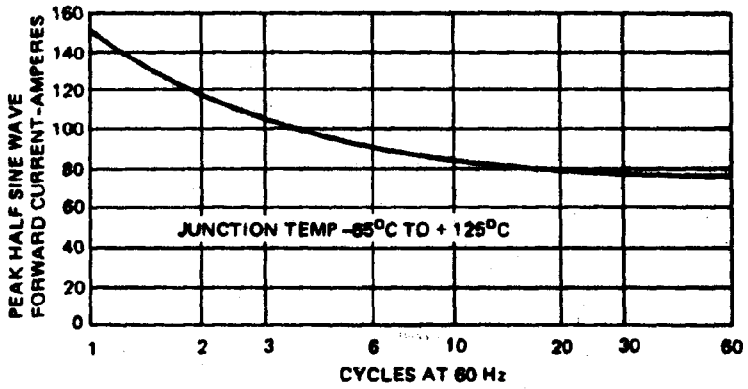


FIGURE 121. Maximum allowable nonrecurrent peak surge forward current at rated load conditions.

I^2t ratings apply for nonrecurrent overloads shorter than one half cycle. For such times, the SCR behaves essentially like a resistance with a fixed thermal capacity and negligible power dissipating means, and it displays a current capability which can be expressed as a constant, I^2t, where I is the rms value of current over an interval, t. The I^2t rating of the SCR is given in the specifications. This rating assumes that the SCR is already in the conducting state. If the SCR is turned on into a fault, the current-time relationships (di/dt) during the turn-on interval must be within the device's switching capabilities.

4.10.5.2 Holding and latching current. As with the solenoid of an electromechanical relay, an SCR requires a certain minimum anode current to maintain it in the closed or conducting state. If the anode current drops below this minimum level, designated as the holding current, the SCR reverts to the forward blocking or open state. The holding current for a typical SCR has a negative temperature coefficient; that is, as its junction temperature rises, its holding current requirement decreases.

A somewhat higher value of anode current than the holding current is required for the SCR to initially "pickup." If this higher value of anode latching current is not reached, the SCR will revert to the blocking state as soon as the gate signal is removed. After this initial pickup action, however, the anode current may be reduced to the holding current level. Where circuit inductance limits the rate of rise of anode current and thereby prevents the SCR from switching solidly into the conducting state, it may be necessary to make alterations in the circuit.

A meaningful test for the combined effects of holding and latching current is shown in Figure 122. The SCR under test is triggered by a specified gate signal. The test circuit allows the SCR to latch into conduction at a current level I_{F1}. The test circuit then reduces the current to a continuously variable level I_{F2}. The current I_{F2} at which the SCR reverts to the off state is the desired value of holding current.

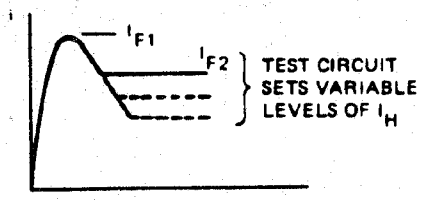


FIGURE 122. Holding current test waveform.

4.10.5.3 Rate of rise of forward voltage (dv/dt). A high rate of rise of forward (anode-to-cathode) voltage may cause an SCR to switch into the on, or low impedance, forward conducting state. In the interest of circuit reliability, it is important to characterize the device with respect to its dv/dt withstand capability.

SCRs are characterized with respect to dv/dt withstand capability in two ways:

a. The so-called static dv/dt withstand capability. This specification covers the case of initially energizing the circuit or operating the device from an anode voltage source that has superimposed fast rise-time transients. Such transients may arise from the operation of circuit switching devices or result from other SCRs' operation of adjacent circuits.

b. The maximum allowable rate of reapplication of forward blocking voltage while the SCR is regaining its rated forward blocking voltage (V_{FXM}) following the device turn-off time (t_{off}) under stated circuit and temperature conditions. In this context dv/dt is an important part of the overall SCR turn-off characterization.

Figure 123 shows the typical static dv/dt as a function of temperature for a medium current SCR with its gate open. The rate of rise of anode voltage shown on the ordinate is the slope of a straight line starting at zero anode voltage and extending through the one-time constant (τ) point on an exponentially rising voltage. The upper right hand portion of the figure illustrates this definition.

The definition shown in Figure 123 has become accepted as a standard by the industry. Some specification sheets give the time constant under specified conditions rather than by a curve as in Figure 123. It should be noted that: $\tau = 0.632 \times$ rated SCR voltage (V_{FXM})dv/dt.

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

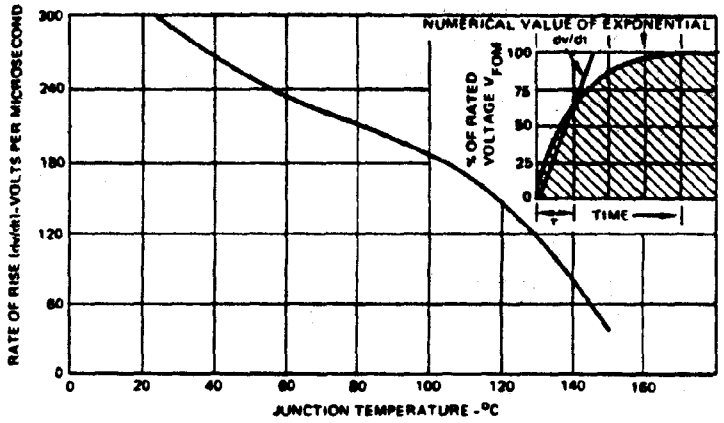


FIGURE 123. Rate of rise (dv/dt) of forward voltage that will not turn on SCR.

The initial dv/dt withstand capability will be recognized as being greater than the value defined. In terms of specified minimum time constant it is

$$\left.\frac{dv}{dt}\right|_{t = 0+} = \frac{\text{rated SCR voltage } (V_{FXM})}{\tau}$$

In terms of specified maximum dv/dt capability, the allowable initial dv/dt withstand capability is

$$\left.\frac{dv}{dt}\right|_{t = 0+} = (1/0.632)dv/dt = 1.58\ dv/dt$$

The shaded area shown in the insert of Figure 123 represents the area of dv/dt values that will not trigger the SCR. These data enable the circuit designer to tailor his circuitry in such a manner that reliable circuit operation is assured. Because a high circuit-imposed dv/dt effectively reduces $V_{(BR)FX}$ (the actual anode voltage at which the particular device being observed switches into the on state) under given temperature conditions, a higher voltage classification unit will allow a higher rate of rise of forward voltage for a given peak circuit voltage.

As an illustration of this, consider a SCR with V_{FOM} = 200 volts at T_j = 130 °C operating at a peak circuit voltage of E = 150 volts, or (150/200) x 100 = 75 percent of rated voltage. Figure 123 shows that a dv/dt = 120 V/µs can be applied to the one time constant (τ) point of an exponentially rising voltage. In this example, τ_1 = (0.632 x 200)/120 = 1.06 µs. Accordingly, from Figure 123, the time in which the anode of the SCR may reach 150 V, or 75 percent of rated voltage is t_1 = 1.5 x 1.06 ≅ 1.6 µs.

If a C38D type is selected with V_{FOM} = 400 V, under the same conditions it may be found by using the ratio of rated voltages that $t_2 = \tau_1$ 400/200 = 2 τ_1 = 2 x 1.06 = 2.12 µs to reach 63.2 percent of 400 or 253 V. Because in this case a maximum circuit voltage of 150 V or 150/400 = 37.5 percent of the rated 400 V V_{FOM} is the design requirement, it is seen from the insert of Figure 123 that t_2 = 0.5 τ_2 = 0.5 (2.12) = 1.06 µs.

By selecting a 400-volt device the time to reach the operating circuit voltage in this example can be reduced to t_2/t_1 = 1.06/1.6 (100) = 67 percent.

Reverse biasing of the gate with respect to the cathode may increase dv/dt beyond that shown in Figure 123.

The circuit shown in Figure 124 can be used to suppress excessive rate of rise of anode voltage. The time constant of the load resistance, R_L, in ohms and capacitor, C, in microfarads should be selected so that

$$\tau \leq R_L C \ \mu s$$

where

τ = minimum time constant of exponential forward voltage rise specified for SCR

Referring to Figure 124, SCR discharges capacitor, C, via resistor, R_d. R_d should be selected on the basis of limiting the peak capacitor discharge current E/R_d during the SCR turn-on interval to a value within the device capability. For best results, the circuit of Figure 124 should be wired and placed in close proximity to the SCR in order to minimize inductive effects. Also, the capacitor should have good high-frequency characteristics.

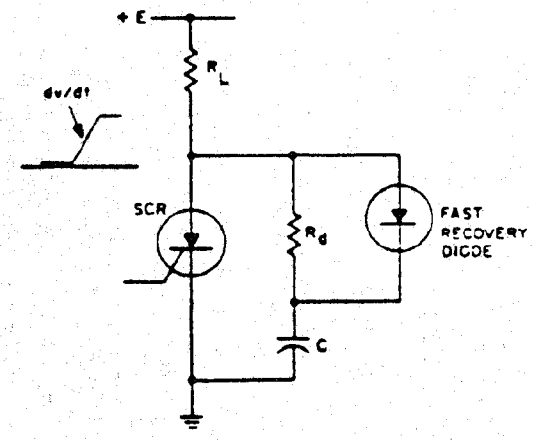


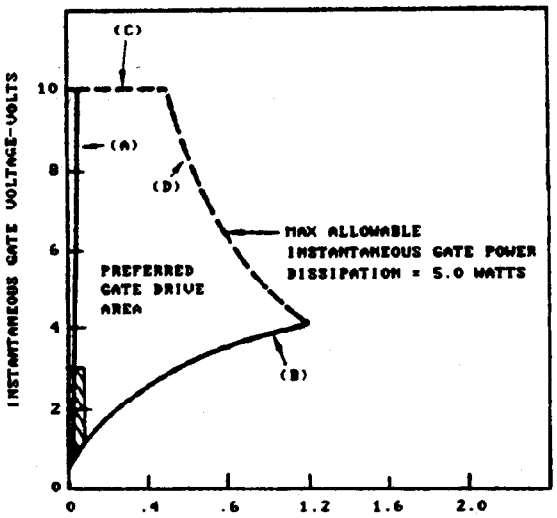
FIGURE 124. Rate of rise of anode voltage (dv/dt) suppression circuit.

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

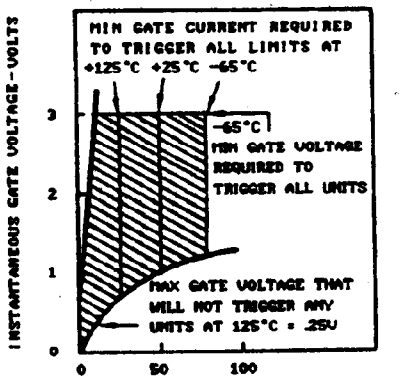
4.10.5.4 Triggering specifications for the dc gate. The typical dc gate triggering characteristics of an SCR are presented in the form of a graph illustrated in Figure 125A. The graph shows gate-to-cathode voltage as a function of positive gate current (flow from gate to cathode) between limit lines (A) and (B) for all SCRs of the type indicated. These data apply to a zero-anode-current condition (anode open).

The basic function of the triggering circuit is to simultaneously supply the gate current to trigger I_{GT} and its associated gate voltage to trigger V_{GT}. The shaded area shown in Figure 125B contains all the possible trigger points (I_{GT}, V_{GT}) of typical SCRs. The trigger circuit must therefore provide a signal (I_{GT}, V_{GT}) outside the shaded area to reliably trigger these SCRs.

The area of trigger circuit-SCR gate operation is indicated as the "preferred gate drive area." It is bounded by the shaded area in Figure 125B which represents the locus of all specified triggering points (I_{GT}, V_{GT}), the limit lines A and B, line C representing rated peak allowable forward gate voltage V_{GF}, and line D representing rated peak power dissipation P_{GM}, Figure 125A. Some SCRs may also have a rated peak gate current, I_{GFM}, which would appear as a vertical line joining curves B and D, Figure 125A.



125A.



125B.

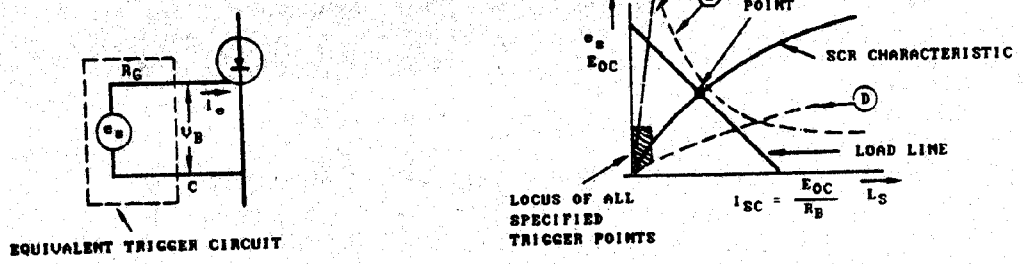
Note: Junction temperature -65 to +125 °C

FIGURE 125. Characteristics of dc gate triggering.

Figure 125B shows the detail of the locus of all specified trigger points and the temperature dependence of the minimum gate current to trigger I_{GTmin}. The lower the junction temperature, the more gate drive is required for triggering. Some specifications may also show the effect of forward anode voltage on trigger sensitivity. Increased anode voltage, particularly with small SCRs, tends to reduce the gate drive requirement. Also shown is the small positive value of gate voltage below which no SCR of the particular type will trigger. The reverse quadrant of the gate characteristic is usually specified in terms of maximum voltage and power ratings.

4.10.5.5 Load lines. The trigger circuit load line is the diagonal line connecting the properly derated values of current and voltage in Figure 125B and must intersect the individual SCR gate characteristic in the region indicated as preferred gate drive area. The intersection, or maximum operating point, should also be located as close to the maximum applicable (peak, average, etc.) gate power dissipation curve as possible. Gate current rise times should be in the order of several amperes per microsecond to minimize anode turn-on time, particularly when switching into high currents. This results in minimum turn-on anode switching dissipation and minimum jitter.

Construction of a load line is a convenient means of placing the maximum operating point of the trigger circuit-SCR gate combination into the preferred triggering area. Figure 126A illustrates a basic trigger circuit of source voltage (e_s) and internal resistance (R_G) driving an SCR gate. Figure 126B shows the placement of the maximum operating point well into the preferred trigger area close to the rated dissipation curve. The load line is constructed by connecting a straight line between the trigger circuit open circuit voltage (E_{OC}) entered on the ordinate and the trigger circuit short circuit current, $I_{SC} = E_{OC}/R_G$, entered on the abscissa.



A. Gate circuit

B. Load line superimposed on gate trigger characteristic

FIGURE 126. Gate circuit and construction of load line.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.10 DIODES, SILICON CONTROLLED RECTIFIERS (SCR)

If the trigger circuit source voltage, e_s (t), is a function of time, the load line sweeps across the graph, starting as a point at the origin and reaching its maximum position at the peak trigger circuit output voltage.

The applicable gate power curve is selected on the basis of whether the average or peak allowable gate power dissipation is the limiting factor. For example, if a dc trigger is used, the average maximum allowable gate dissipation (0.5 W) must not be exceeded. If a trigger pulse is used, the peak gate power curve is applicable (the 5-Watt peak power curve labeled D in Figure 125). For intermediate gate trigger waveforms, the limiting allowable gate power dissipation curve is determined by the duty cycle of the trigger signal according to the equation:

average allowable gate power = peak gate drive pwv x pulse width x pulse repetition rate.

4.10.6 Environmental considerations. Typical environmental conditions and screens which silicon control rectifiers are capable of withstanding are not substantially different from those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 Screening procedures in subsection 4.1, Diodes, general. For the specific device selected, consult the applicable MIL-S-19500 reference sheet.

4.10.7 Reliability considerations. Refer to the subsection 4.11 Diode, general, paragraph 4.1.6 Reliability section for a complete list of failure mechanisms. Some of the more pertinent failure mechanisms and how they relate to SCRs are explained below.

4.10.7.1 Structural flaws. Structural flaws are generally considered to be the result of weak parts, discrepancies in fabrication or inadequate mechanical design. Various in-process tests performed on the device, such as forward voltage drop at high current density levels and thermal resistance measurement, provide effective means for the monitoring of controls against such flaws.

The modes of failure generally associated with mechanical flaws for an SCR are excessive on-voltage drop, failure to turn on when properly triggered, and open circuit between the anode and cathode terminals. These types of failure mechanisms are relatively rare.

4.10.7.2 Encapsulation flaws. Encapsulation flaws are deficiencies in the hermetic seal that will allow undesirable atmospheric impurities to reach the semiconductor element. Foreign atmospheres, such as oxygen and moisture, can react in such a way as to permanently alter the surface characteristics of the silicon.

A change in surface conductivity is evidenced by a gradual increase of the forward and reverse blocking current characteristics. Because the SCR is a current-actuated device, it will lose its capacity to block rated voltage if blocking current degrades beyond some critical point. This type of mechanism

may eventually result in catastrophic failure. The rate of degradation is dependent mostly on the size of the leak and the level of stress, particularly thermal stress, that is applied.

4.10.7.3 Internal contaminants. The inclusion of a source of ionizable material inside the sealed package can result in semiconductor surface inversion, layer formation. This leads to permanent electrical instability in the device.

The occurrence of such defects can cause discrepancies in the turn-on and turn-off stages in an SCR. For example, a variation can occur in the voltage at which the SCR is turned on. On the other hand, a defect such as the one mentioned above, could cause a device not to be turned off completely.

In certain cases, it is possible, using external circuitry (negative gate or resistor biasing), to suppress the effects of such contaminants to the point of rendering the device still usable in the circuit.

4.10.7.4 Material electrical flaws. In certain uses, device failure can result from defective junction formation in the semiconductor. Crystal dislocations at a semiconductor p-n function can result in the formation of diffusion pipes or diffusion spikes which could cause emitter-collector shorts in the given device. These pipes or spikes become very important as the depletion width of the lightly doped region decreases. The probability of occurrence of oxidation-induced stacking faults and other defects can be reduced by careful processing of the semiconductor device.

4.10.7.5 Metal diffusion. Impure metal atoms lodged in a semiconductor device can, under certain circumstances, diffuse into the p-n junction, thus causing device degradation.

The probability and extent of device degradation and their effects on device lifetime are all functions of the temperature and the diffusing metal atom species.

Lithium, copper, silver gold, zinc, nickel, and iron are examples of metals that diffuse rapidly in silicon.

4.10.7.6 Derating. History has shown that the largest single cause of SCR failure is operating above allowable levels of electrical and thermal stress. Accordingly, it is imperative that derating of parts be performed to enhance the reliability of electronic systems. Users should refer to MIL-STD-975 for derating factor guidelines. The design and use of a part should always be within the thermal and electrical stresses defined. High temperature operation is the most destructive stress for a semiconductor. It will result in early electrical parameter drift, and a general degradation of electrical and mechanical characteristics of the device.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.11 DIODES, PHOTODIODES

4.11 Photodiodes.

4.11.1 Introduction. One of the most basic of all the optoelectronic solid-state devices is the photodiode. The photodiode may be used in the photoconductive mode or the photovoltaic mode. In the photovoltaic mode, the device is used to convert radiant energy to a generated voltage under zero bias conditions. In this mode, the output voltage generated across its electrodes will vary with the intensity of the light striking the thin top P^+ or N^+ layer. In the photoconductive mode, the device is reverse biased and the photocurrent generated (I_L) will vary linearly with the intensity of the light striking the top surface of the pn junction.

Figure 127 illustrates the equivalent circuit which is applicable for both modes of operation and defines the various terms used.

Figure 128 illustrates with separate circuit diagrams the photoconductive and photovoltaic modes of operation for photodiodes in general, whether the basic semiconductor material used is silicon, germanium, or some III-V semiconductor alloys.

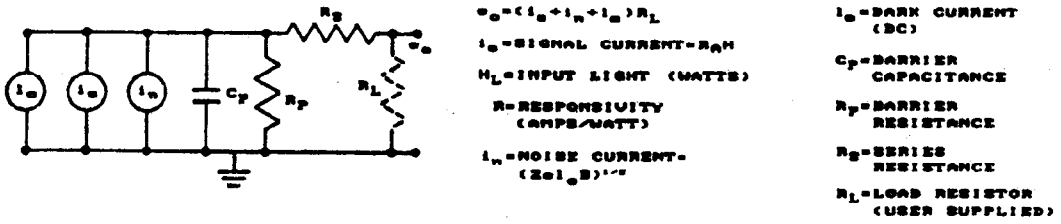


FIGURE 127. Equivalent circuit for photodiodes (all models).

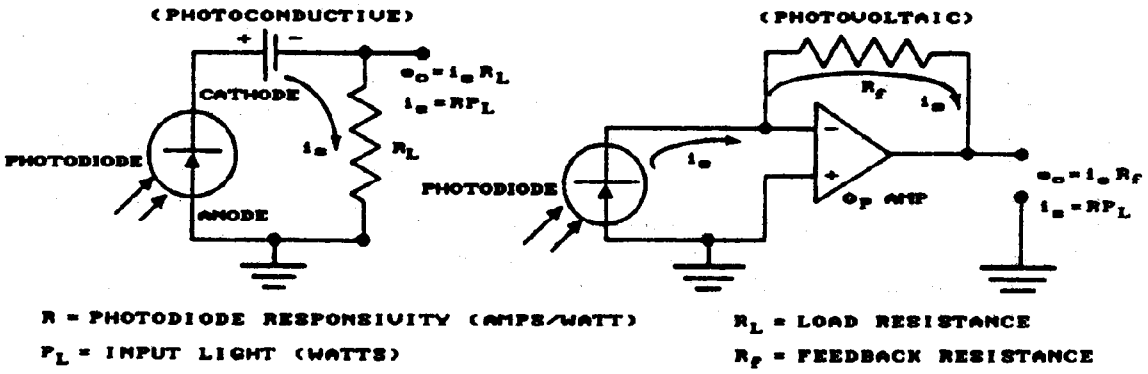


FIGURE 128. Typical hookup circuits for various photodiode modes.

The pn junction photodiode has a slow response time, being as slow as 30 µs. Responsivity can be good, however, depending somewhat on the construction of the device. Responsivity of the pn junction photodiode can be in the neighborhood of 0.3 to 0.7 A/W (amperes/watt) at the wavelength of peak sensitivity. The spectral response of most silicon pn junction photodiodes is from about 500 to 1000 nm. Again, this can be adjusted by varying the type of semiconductor material used. The pn junction photodiodes are also characterized by a high thermal noise level and high dark current, which make them useful only for short distance communication. In practice, most conventional pn junction photodiodes are used more for industrial sensing than for communication purposes.

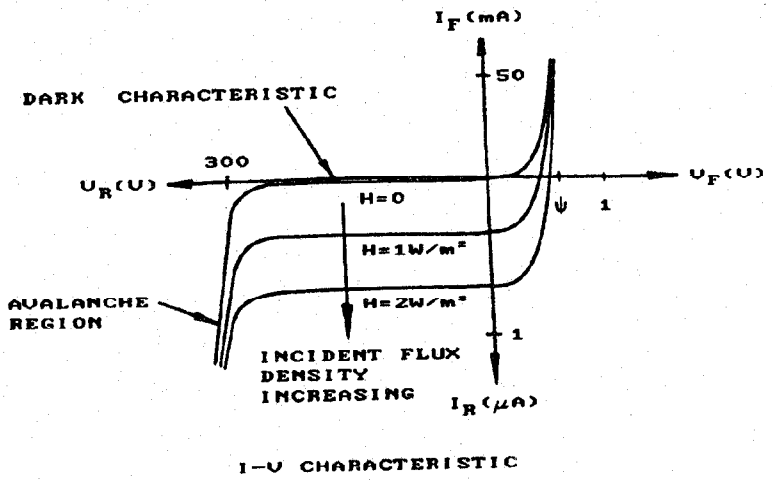


FIGURE 129. Typical photodiode I-V diagram.

Another way of describing a photodiode is with the familiar current-voltage, or I-V, diagram. Figure 129 is the I-V diagram for a typical photodiode. The arrow points to the voltage at which the photodiode is reverse biased. When the photodiode is illuminated, the I-V curve drops down to a more negative value, but the reverse voltage (X) remains the same. As the light intensity continues to increase, the I-V curve drops even further, as shown in the figure. In other words, the photo current increases with more light, though the reverse voltage is constant. However, the current cannot increase indefinitely because of the physical limitations of the photodiode. A point will be reached (saturation) at which the device will no longer respond to increases in light intensity.

An advantage of using photodiodes is that the photodiode reacts to changes in light intensity much more quickly than the photoconductor. This is because the charge carriers are accelerated by the depletion region's barrier field whereas the photoconductor's charge carriers move only by the slow thermal diffusion process. However, most conventional pn junction photodiodes are still not fast enough for most fiber optic work. This is because the depletion region in the photodiode is so small that a good portion of the light is absorbed by the semiconductor material outside of the depletion region.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.11 DIODES, PHOTODIODES

4.11.2 Usual applications. Photodiode applications include cared and tape readers, pattern and character recognition, shaft encoders, position sensors, and counters.

Both the PIN photodiode and the avalanche photodiode (APD) are very useful as photodetectors in fiber optics system applications.

4.11.2.1 Schmitt detector. In this particular circuit application, as shown in Figure 130, PIN photodiodes are superior to silicon npn phototransistors as signal detectors because when the frequency of the optical signal reached about 2 KHz, the output from the detector became very distorted whereas there was no signal distortion with the photodiode.

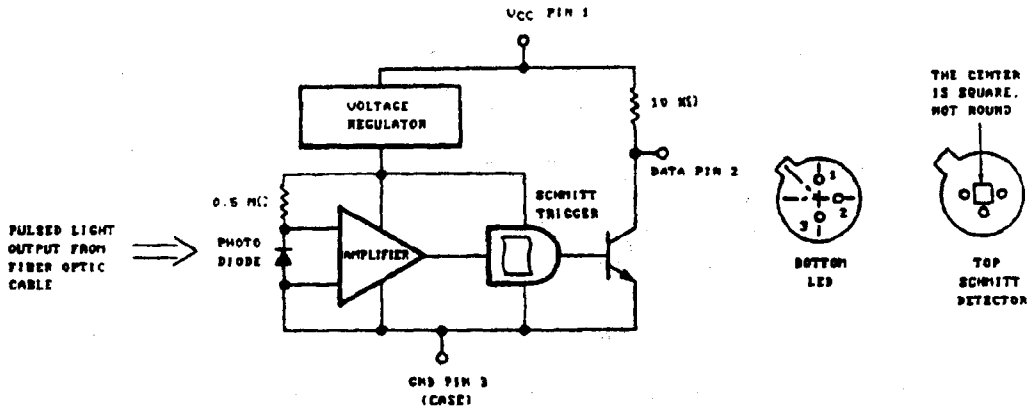


FIGURE 130. Schmitt detector circuit.

In addition, photodiodes, especially the PIN and APD types, are used extensively in fiber optic communication systems as detectors or as the key element in the receiver system as illustrated in the figures that follow.

It should be pointed out that there are three types of fiber optic communication systems; namely, simplex, half duplex, and full duplex, depending upon the direction in which their optical data may flow.

Reference to data direction applies to two terminals in a fiber optic system (transmitter and receiver). Between two terminals, data can travel one way, from one terminal to the other; two ways, from one terminal to the other and back again; and two-ways simultaneously, from one terminal to the other and bad again at the same time. In other words, one-way, two-ways, and two-ways simultaneously are referenced to as simplex, half duplex, and full duplex, respectively.

4.11.2.2 Simplex systems. Simplex systems are aptly named for they are the simplest kind of communication link possible. Figure 131 is a diagram of a typical simplex system. Note that the link consists of one transmitter connected to one receiver by one fiber. Data can travel in only one direction in this type of link, making it useful only for remote control or reception of telemetry. Unfortunately, this simple and inexpensive system has no means of feedback. That means that the sender has no way of knowing whether or not communication has been received correctly; therefore, there is no way to know when errors have occurred in transmission. Because of this inherent problem, simplex systems are used only in noncritical applications. However, the user may benefit from communications expense if these drawbacks are tolerable.

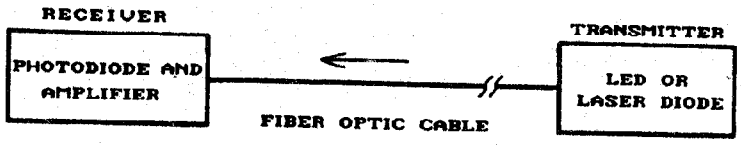


FIGURE 131. Simplex fiber optics communication system.

4.11.2.3 Half duplex systems. Figure 132 is a diagram of a typical half duplex communication system. It consists of two transmitter-receiver pairs connected by a single fiber optic cable. In this system, communication can travel in both directions, but not at the same time.

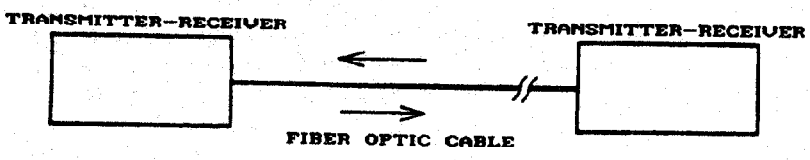


FIGURE 132. Half duplex fiber optic communication system.

To accomplish two-way communication on one line, the system must have an emitter and a photodetector coupled to each end of the fiber. This means having a combination LED/photodiode at each end of the system or, as shown in Figure 133, a way of splitting the transmission path at each end. Both of these arrangements are compromises in quality because (1) an LED/photodiode combination cannot work at peak efficiency as both a generator and detector of light, and (2) splitting the line results in smaller signal levels at the receiver. This means that the length of the line must be reduced to maintain the minimum acceptable signal level at the receiver.

4.11 DIODES, PHOTODIODES

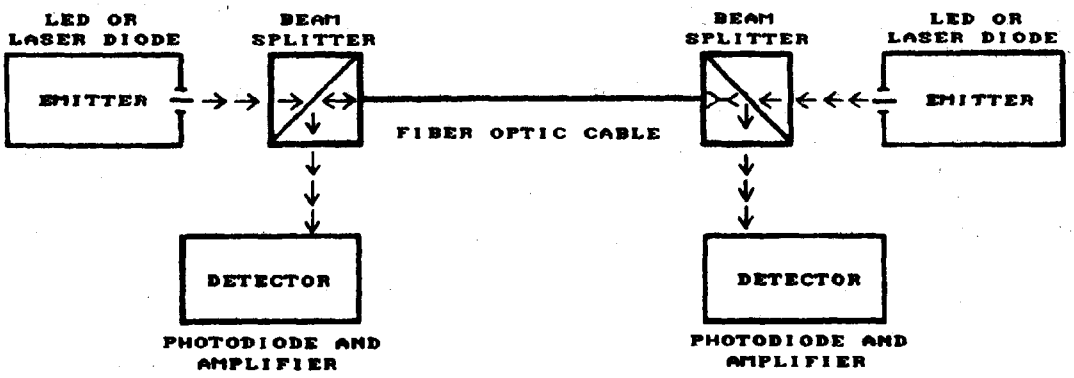


FIGURE 133. Half duplex system with split transmission path.

Special LED/photodiodes and beam-splitting connectors make half duplex systems more expensive and complicated than simplex systems. In terms of the number of lines used, half duplex systems are efficient because they only need one line. But in terms of the time spent waiting for the other party to stop sending so that you can send, the half duplex is inefficient. Also, because beam splitters reduce signal strength, half duplex systems are limited in the distance they can cover without a repeater. For these reasons, half duplex systems are also not widely used.

4.11.2.4 Full duplex systems. Figure 134 is a diagram of a typical full duplex system. It consists of two transmitters and two receivers connected by two fiber optic cables. Each cable carries communication in one direction only. In effect, a full duplex system is nothing more than two simplex systems side by side, operating in opposite directions.

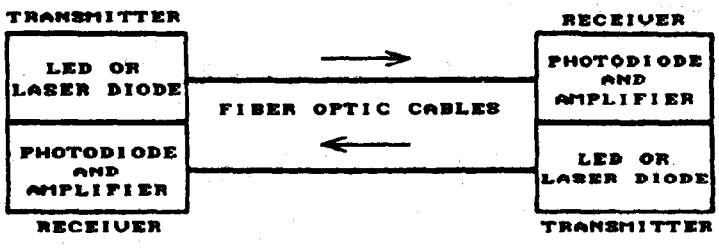


FIGURE 134. Full duplex fiber optic communication system.

The disadvantage of this arrangement is that two fibers are not needed instead of one. However, the advantage of simultaneous communication often outweighs this disadvantage. Another advantage is that the system can be optimized for maximum length by using separate LEDs and photodetectors instead of LED/photodiode devices.

It is possible to operate a full duplex on one line by means of wavelength multiplexing. However, this technique is not currently used outside of the laboratory, although it may be used in future applications.

4.11.3 Physical construction.

4.11.3.1 Conventional pn junction photodiode and PIN photodiode. Figure 135 is a typical cross sectional drawing of a conventional photodiode. There have been many processes by which pn junctions have been formed. The technique that has become the most popular and is most used by industry is the planar process. In this process silicon is used as the starting semiconductor material and is followed by gas-phase doping of the solid silicon at high temperatures (1000 - 1200 °C). By diffusion, the dopant gases penetrate the solid surface of the silicon. Diffusions of n-type and p-type impurities may be made onto a previously doped region of an n-type or p-type substrate wafer by appropriate masking off one side of the wafer after lapping off the previously diffused doped layer. This serves to provide manufacturing versatility.

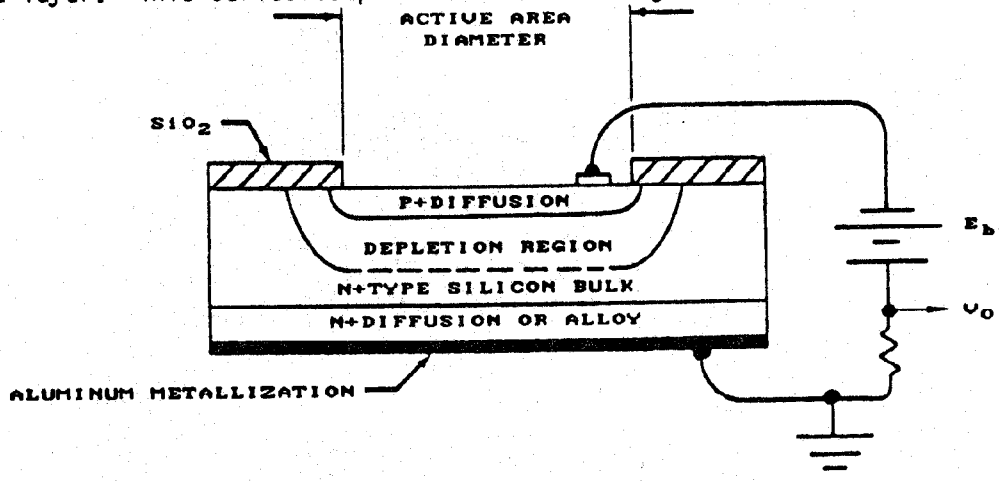


Figure 135. Cross section, planar diffused photodiode.

In the foregoing illustration, n-type bulk silicon is diffused on one side by n^+ dopant and on the opposite side by p^+ dopant. A depletion region between the n and p regions exists which is free of carriers while under appropriate reverse bias conditions. It is in this depletion area that the photon should be absorbed. The active area exposes the thin p^+ diffused region to the light beam. The light beam is in turn absorbed in the semiconductor material surface. When photons of the correct energy are absorbed into the semiconductor surface, electron-hole pairs are formed.

4.11 DIODES, PHOTODIODES

The PIN photodiode is similar to the one described in Figure 135 with one important difference. This device has an intrinsic (I) layer between its p and n regions and is commonly referred to as a PIN photodiode. The intrinsic layer is undoped and has a bulk silicon resistivity of typically 300 to 500 Ω-cm compared to conventional photodiodes which can have bulk silicon resistivities in the 5 to 20 Ω-cm range. Therefore, the depletion region in this case will extend further into the "I" region than it would in a more heavily doped semiconductor. In other words, adding the "I" layer results in a much wider depletion region for a given reverse voltage near the breakdown value. The wider depletion area makes the PIN photodiode respond better to the lower light frequencies (longer wavelengths).

4.11.3.2 Schottky barrier photodiode. The Schottky barrier photodiode is illustrated in Figure 136. The Schottky photodiode (often called a surface diode) differs from the conventional photodiode in the method by which the p-type material is formed. In the conventional photodiode the pn junction is developed by a chemical diffusion process. In the Schottky photodiode, a metal-semiconductor barrier is formed over a thin silicon epitaxial layer (2 to 5 μm) of several ohm-centimeters resistivity by evaporation of a thin gold metal film (10- to 20-nm thick).

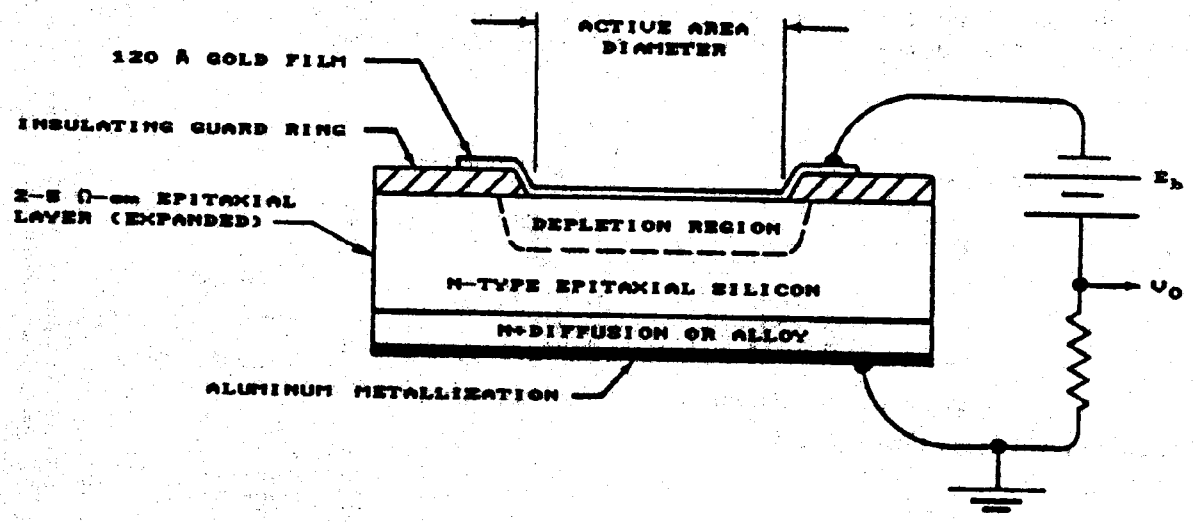


FIGURE 136. Cross section of Schottky barrier photodiode.

4.11.3.3 Avalanche photodiode. Figure 137 illustrates the general construction of an epitaxial silicon reach-through avalanche photodiode with n^+-p-π-p^+ structure (π refers to the intrinsic silicon of about 300 to 500 Ω-cm). The diameter of the light-sensitive, high-gain region is typically 100 μm.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

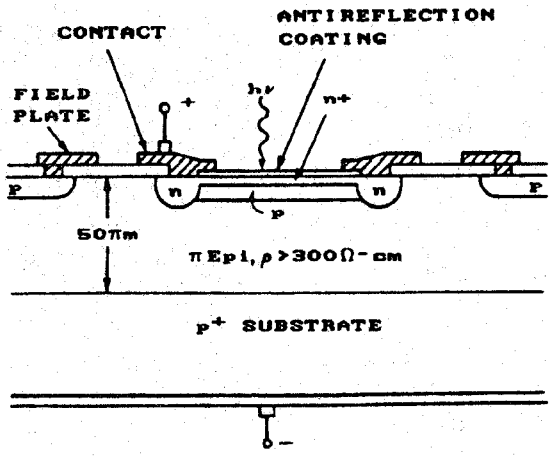


FIGURE 137. Cross section, avalanche/photodiode.

In the foregoing figure, the antireflection coating for silicon devices is usually a thermally evaporated silicon monoxide film of the appropriate thickness and refractive index. Its purpose is to increase quantum efficiency and thereby increase the responsivity of the device near its peak value.

Typical packages commonly used for photodiodes are JEDEC TO-5, TO-18, and TO-46.

4.11.4 Military designation. Currently, there are no JANS or JANTXV military device designations on DESC's QPL list (April 1, 1986) for photodiodes. For high reliability applications such as space, the user can generate a specification control drawing (SCD) describing a manufacturer's generic type screened in accordance with MIL-S-19500 to JANS or JANTXV requirements as applicable. It is anticipated that with increasing usage of photodiodes and other optoelectronic devices in high reliability applications, QPL registrations will follow.

4.11.5 Electrical characteristics. Short, medium, and long wavelengths of photon power are absorbed at different depths within the pn junctions. The depth of penetration is dependent on the photon wavelength or energy as illustrated in Figure 138 for higher resistivity silicon (>5 Ω-cm). Short wavelengths (higher photon energy), as expected, are absorbed near the surface. Long wavelengths (lower photon energy) may penetrate the entire die structure. To be most useful, the wavelengths should be absorbed in the depletion area.

The photon current is produced by electron-hole pairs being separated and drawn out in directions of more positive or negative sources whichever is the case. If the electron-hole pairs are generated outside the depletion area they will usually recombine and no photo current will be produced. The active p^+ diffused

4.11 DIODES, PHOTODIODES

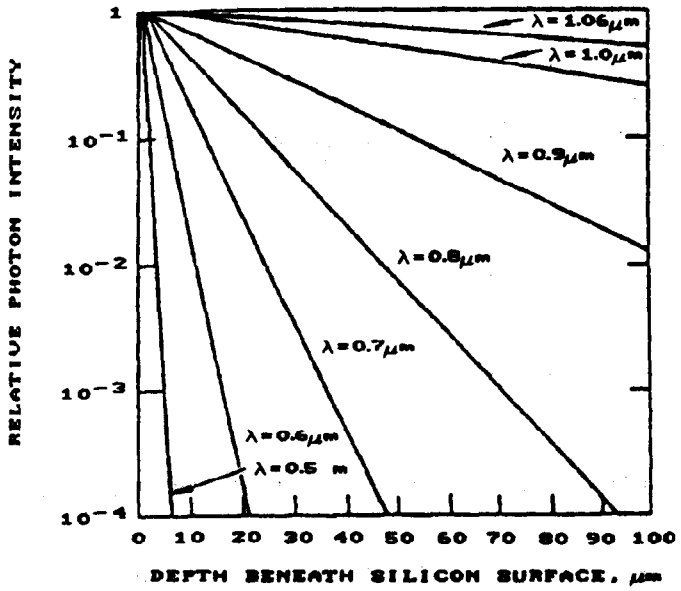


FIGURE 138. Decay of incident light intensity as a function of penetration depth into silicon substrate.

region should be extremely thin (less than 1 μm) to ensure maximum penetration. As in other reverse-biased pn junction diodes, the depletion width can be made larger by increasing the reverse bias and thereby reducing the junction capacitance.

In a practical photodiode, 100 photons will create between 30 and 95 electron-hole pairs, thus giving the detector a quantum efficiency ranging from 30 to 95 percent. To achieve a high quantum efficiency, the depletion layer must be thick enough to permit a large fraction of the incident light to be absorbed. However, the thicker the depletion layer, the longer it takes for the photo-generated carriers to drift across the reverse-biased junction. Because the carrier drift time determines the response speed of the photodiode, a compromise has to be made between response speed and quantum efficiency.

For very low level light signals, all of the previous photodiodes considered fall short in their performance. Their responsivity prevents them from being used in electronic systems covering large distances (over a kilometer) because the signal attenuation at that distance becomes excessive. A group of photo-diodes that can handle lower level signals is the avalanche-type (APD). With their internal amplification of light signals, they are well suited for long distance communication. These devices in general are more expensive than the corresponding PIN photodiodes.

Table IX is a performance specification table and Figure 139 is a spectral response figure for a typical PIN photodiode. The response time for this photodiode is about 1 ns, which is much faster than the 30 µs quoted earlier for the pn junction photodiode. Responsivity is about 0.5 A/W whereas dark current is about 1 nA at room temperature. In general, no figure for noise will be found in any vendor's performance specification. Instead, what will usually be given is the noise equivalent power (NEP), which is a measure of the light power required to produce a signal-to-noise ratio of one to one in the device. (Another name for NEP is sensitivity.)

Notice that the spectral response is not less than 40 percent of the maximum response for the wavelengths from 500 to 950 nm. This is typical for silicon photodevices.

Better response time than is possible with either a pn or PIN photodiode can be achieved through the use of the Schottky barrier photodiode.

The lower frequency photons contain less energy and tend to penetrate deeper into the diode's structure before producing electron-hole pairs and in many cases, do not produce pairs. The wider depletion region of the PIN photodiode increases the chance that pairs will be produced. The PIN photodiode is therefore more efficient over a wider range of light frequencies. The PIN device also has a lower internal capacitance due to the wide "I" region which acts like a wide dielectric between the p and n regions. This lower internal capacitance allows the device to respond faster to changes in light intensity. The wide depletion region also allows this device to provide a more linear change in reverse current for a given change in light intensity.

In the case illustrated, Figure 136, the active region has a very thin gold film that covers the metallized n-type or p-type silicon epitaxial layer. This film must be thin (200Å or 20 nm) to allow light penetration to the surface of the silicon epitaxial layer. The Schottky barrier photodiode has some advantages over the conventional photodiode. It operates well at wavelengths less than 500 nm and has a much simpler fabrication process. It should be pointed out that the Schottky photodiode does not operate well at high temperatures (greater than 70 °C) or with high light power (greater than 2 mW of radiant power). They have good response times, typically less than 25 ns.

These devices are recommended when high blue response (wavelengths less than 500 nm) or larger areas (greater than 1 cm^2) are required.

4.11 DIODES, PHOTODIODES

TABLE IX. Electrical characteristics of a PIN photodiode

Characteristic	Symbol	Min	Typ	Max	Unit
Dark current 1/ (V_R = 20 V, R_L = 1.0 Megohm)	I_D				nA
T_A = 25 °C		-	1.0	10	
T_A = 100 °C		-	14	-	
Reverse breakdown voltage (I_R = 10 µA)	$V_{(BR)R}$	100	200	-	V
Forward voltage (I_F = 50 mA)	V_F	-	-	1.1	V
Series resistance (I_F = 50 mA)	R_S	-	-	10	Ohms
Total capacitance (V_R = 20 V, f = 1.0 MHz)	C_T	-	-	4.0	pF
Responsivity (Figure 139)	R	0.4	0.5	-	µA/µW
Response time (V_R = 20 V, R_L = 50 Ω)	t_{on}	-	1.0	-	ns
	t_{off}	-	1.0	-	ns

1/ Measured under dark conditions. H = 0

For the avalanche photodiode shown in Figure 137, the guarding helps to ensure that the edge of the junction does not go into breakdown prematurely because of the high electric field strength at the edges. The guard ring is slightly doped to keep the electric field weak at the edge, forcing breakdown to occur in the bulk portion of the junction. This ensures a linear current reaction to the change in the light signal level.

The reach-through avalanche photodiode is composed of a high resistivity p-type material deposited as an epitaxial layer on a p^+ (heavily doped p-type) substrate. A p-type diffusion or ion implant is then made in the high-resistivity layer (200 to 500 Ω-cm) followed by the diffusion of an n^+ (heavily doped n-type) layer. For silicon, the dopants used to form these layers are normally boron and phosphorus, respectively. This configuration is referred to as a silicon p^+-π-p-n^+ "reach-through" structure.

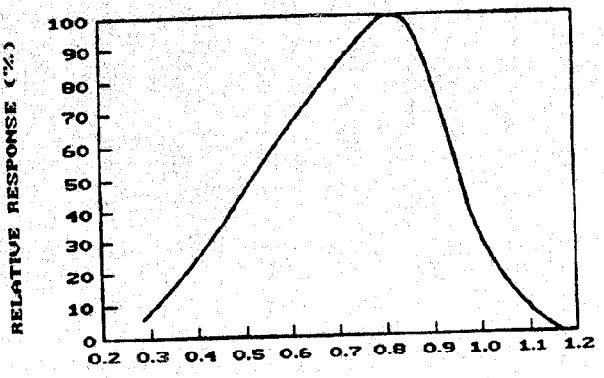


FIGURE 139. Spectral response.

The term reach-through arises from the photodiode operation. When a low reverse-bias voltage is applied, most of the potential drop is across the pn^+ junction. The depletion layer widens with increasing bias until a certain voltage is reached at which the peak electric field at the pn^+ junction is about 5 to 10 percent below that needed to cause avalanche breakdown. At this point, the depletion layer just "reaches through" to the nearly intrinsic region. The latter effect is illustrated in Figure 140.

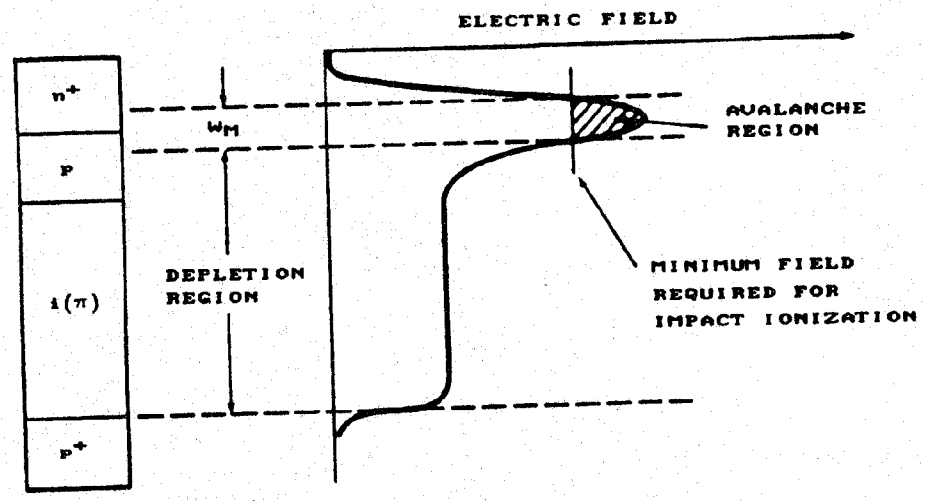


FIGURE 140. Reach-through avalanche photodiode structure and the electric fields in the depletion and multiplication regions.

4.11 DIODES, PHOTODIODES

In normal usage the device is operated in the fully depleted mode. Light enters the device through the thin p^+ region and is absorbed in the π, intrinsic material, which acts as the collection region for the photogenerated carriers. Upon being absorbed the photon gives up its energy to create electron-hole pairs, which are then separated by the electric field in the π region. The photogenerated electrons drift-through the π region to the pn^+ junction where a high electric field exists (as shown in Figure 140). It is in this high-field region that carrier multiplication takes place.

In other words, electron-hole pairs generated by light signals generate many more electrons when the photodiode is in the avalanche state, so that a small amount of light generates a large change in photocurrent.

Table X gives typical performance values of an avalanche photodiode. The response time, listed as the rise time, is about 2 ns, which is not very much different from the PIN photodiode discussed earlier. Note that the responsivity in an avalanche photodiode can be as high as 75 A/W, which is many times greater than the responsivity of a PIN photodiode. Its noise equivalent power is high though, being 100 times greater than a PIN's NEP. At high-level optical signals, the shot noise component of the total noise makes the avalanche photodiode almost unusable.

TABLE X. Electrical characteristics of an avalanche photodiode

Electrical Characteristics at T_A = 22 °C At the dc Reverse Operating Voltage V_R Supplied with the Device	C30954E Light Spot Diameter 0.25 mm (0.01 in.) Typ.	Units
Breakdown voltage, V_{BR}	375	V
Temperature coefficient of V_R for constant gain	2.2	V/°C
Gain	120	
Responsivity		
At 900 nm	75	A/W
At 1060 nm	36	A/W
At 1150 nm	5	A/W
Quantum efficiency		
At 900 nm	85	%
At 1060 nm	36	%
At 1150 nm	5	%
Total dark current, I_d	5×10^{-8}	A
Noise current i_n f = 10 kHz, Δf = 1.0 Hz	1×10^{-12}	$A/Hz^{1/2}$
Capacitance, C_d	2	pF
Series resistance	-	Ω
Rise time, t_r		
R_L = 50 Ω, λ =900 nm, 10% to 90% points	2	ns
Fall time		
R_L = 50 Ω, λ =900 nm, 90% to 10% points	2	ns

4.11 DIODES, PHOTODIODES

4.11.5.1 Device performance considerations. Photodetectors performance is based on five major criteria.

a. Response time

b. Responsivity and quantum efficiency

c. Spectral response

d. Noise

e. Dark current.

In the following paragraphs, each of these criteria will be examined in more detail.

4.11.5.1.1 Response time and rise time. Response time is the speed at which a photodetector can change voltage or conductivity in response to a change in light intensity. Therefore, light pulses that are much shorter than a photo detector's response time will not be detected. Thus, it is important to choose a detector with a response time less than the pulse width of any signal that may be encountered in the system application.

Another name for response time is rise time. This is the time it takes the detector's output to reach the level that corresponds to the light intensity of the input signal. Figure 141 illustrates the idea of rise time.

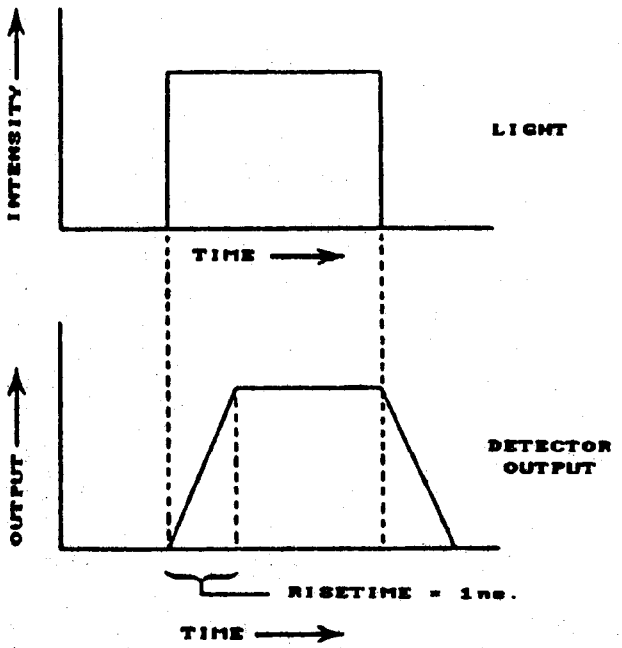


FIGURE 141. Rise time.

The upper half of the figure represents the input light signal. The lower half of the figure represents the response of the detectors. Notice that upon reception of the light signal, the detector does not immediately rise to its full output until about 1 ns later which is a very good response time.

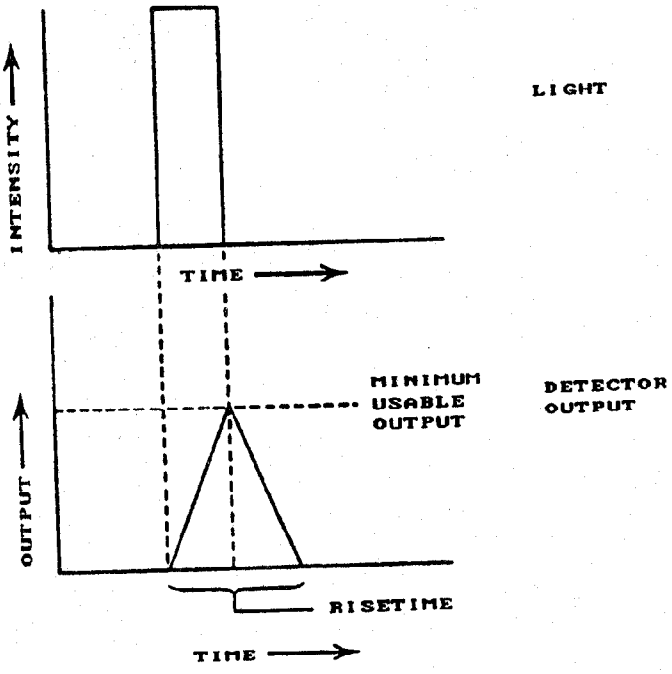


FIGURE 142. Insufficient rise time.

Figure 142 shows what happens when a narrow pulse is received by a very slow detector. Because the light pulse is narrower than the response time of the detector, the output of the detector never reaches its full value. If the output is too low, the receiver circuitry will not be able to distinguish it from noise. Therefore, the light signal pulse width and the detector response time must be compatible.

The response time is another way of measuring the amount of data a photo-detector can handle. If the response time is low, data pulses can be placed closer together and be made narrower. Because the data pulses can be made narrower in time, more pulses can be received per second than if the pulses were wider. Thus, a very short response time is a desirable feature in any photodetector.

4.11 DIODES, PHOTODIODES

4.11.5.1.2 Responsivity and quantum efficiency. The performance of a photodiode can also be expressed in terms of its responsivity (R_A). The responsivity of a photodiode is simply a measure of how much output (reverse) current is obtained from a given light input. It is expressed as a ratio of output current (or photocurrent) in microamperes (μA) to the input radiant incidence which is measured in milliwatts per square centimeter (mW/cm^2) as shown below:

$$R_A = \frac{I_L}{P0} = \frac{ng}{h\nu} + \frac{\mu A/cm^2}{mW/cm^2} \qquad \frac{\mu A}{mW} = \frac{A}{W}$$

I_L is defined as the photocurrent generated when a given value of optical power (P_0) strikes the photodiode surface.

The responsivity of a photodiode also varies with the wavelength of the radiant energy striking the device surface and reaches a peak valve when the quantum efficiency of the device is near its highest valve. A typical photodiode will have a maximum responsivity of 1.4 μA per milliwatt per square centimeter. The quantum efficiency, η, is defined on page 4-120 with the terms q, h and ν. In most photodiodes the quantum efficiency is independent of the power level falling on the detector at a given photon energy. Thus the responsivity is a linear function of the optical power. As shown above, the photocurrent I_L is directly proportional to the optical power (P_0) incident upon the photodetector, so that the responsivity R_A is constant at a given wavelength (for a given value of hν). Note, however, that the quantum efficiency is not a constant at all wavelengths because it varies according to the photon energy. Consequently, the responsivity is a function of the wavelength and of the photodiode semiconductor material because different materials have different band gap energies.

4.11.5.1.3 Spectral response. The foregoing graph shows the wavelength at which the detector is best suited to operate. At wavelengths away from the peak R_A value (vertical dashed lines), the amount of detectable signal falls off sharply, making communications more error-prone and inefficient. If a system is to be operated within a particular wavelength band, the spectral response error will reveal if the detector is suitable at those wavelengths. If the wavelengths have not been determined for the system, the spectral response curve will show what wavelengths are best for the detector.

Photodetectors, as can be seen in Figure 143, exhibit well defined, long wavelength thresholds (λc), the positions of which are determined by the magnitude of the band gap energy Eg or impurity activation-energy (E_A). The cutoff wavelength λc, corresponding to the given band gap energy is given by

$$\underset{(\mu m)}{\lambda c} = \frac{hc}{Eg} = \frac{1.24}{Eg(eV)}$$

where the product of Planck's constant, h, and the velocity of light, e, is 1.24; Eg is expressed in electron-volts and λc in micrometers (microns).

As shown in Figure 143, the cutoff wavelength (λc) is about 1.1 μm for silicon, 1.88 μm for germanium and 1.65 μm for $In_{0.53}Ga_{0.47}As$. For longer wavelengths, as can be observed from the above equation, the photon energy is not sufficient to optically excite an electron from the valence band to the conduction band.

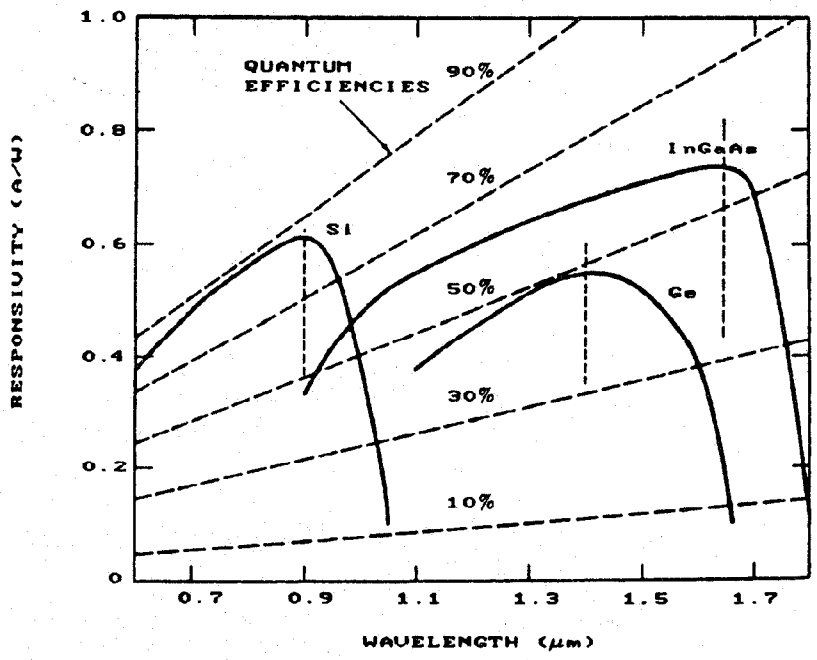


FIGURE 143. Spectral response curves (R_A vs λ) and quantum efficiency as a function of wavelength for PIN photodiodes for various semiconductor materials silicon, germanium and a III-V semiconductor alloy at 25 °C.

In Figure 144, the absorption coefficient α (measured in cm^{-1}) is plotted against the wavelength of incident radiation and the light penetration depth (μm) for various semiconductor materials used in photodiodes. At wavelengths longer than the band-edge wavelength (λc) (see Figure 144) the absorption decreases sharply. At the lower wavelength end, the photoresponse cuts off as a result of the very large α values at the shorter wavelengths. In this case the photon are absorbed very close to the photodetector surface where the recombination time of the generated electron-hole pairs is very short. The optically generated carriers thus recombine before they can be collected by the photodetector circuitry.

4.11 DIODES, PHOTODIODES

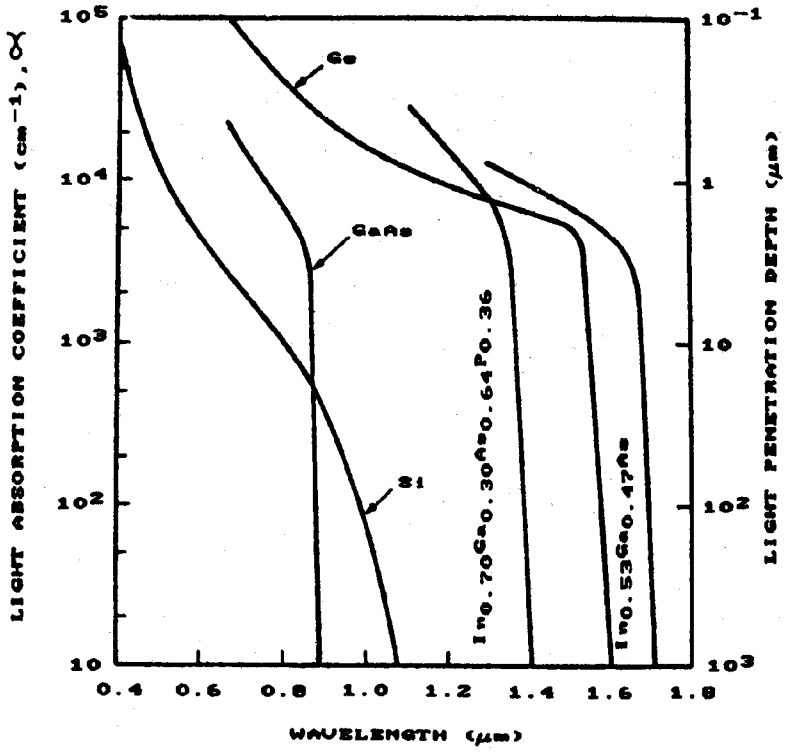


FIGURE 144. Absorption coefficient vs wavelength for Ge, Si, GaAs, InGaAsP, and InGaAs alloy at 25 °C.

A measurement similar to responsivity is quantum efficiency. Quantum efficiency tells what fraction of a single frequency light beam will be absorbed by electrons in a photodetector. In the ideal case, one electron-hole pair should be produced for each photon that strikes the diode surface, thus giving the ideal diode a quantum efficiency of 1 or 100 percent. However, the quantum efficiency of a typical photodiode will be lower than 100 percent and will vary with the wavelength of the incident radiant energy as shown in Figure 143.

The quantum efficiency, designated as η, is more quantitively defined as the number of electron-hole carrier pairs (per unit area per unit time) per incident-photon (per unit area/unit time) of given energy $h\nu$ and is given by

$$\eta = \frac{\text{Number of electron-hole pairs generated}}{\text{Number of incident photons}} = \frac{(I_L/q)}{(P_O/h\nu)}$$

Here I_L is the average photocurrent generated by a steady-state average optical power (P_O) incident on the photodetector, q is the charge of an electron, and h is Planck's constant.

4.11.5.1.4 Noise. Noise that is strong enough to mask a communications signal is called unwasted energy. In a photodetector, noise is made up of thermal noise and shot noise.

Thermal noise is caused by the thermal agitation of electrons in the detector. As the temperature rises electrons can absorb heat energy randomly, introducing a sporadic element into the detector's output. Usually this noise is so low (in the neighborhood of 2 nA) that it can be neglected, although it can be a problem in high temperature environments.

Thermal noise is not restricted to photodetectors. It can appear in other components as well. For instance, resistors in a receiver circuit can be a source of thermal noise. These extra thermal noise sources have to be considered when evaluating a system.

Shot noise is not a function of temperature, but of signal strength. As signal strength increases, that is, as the light intensity increases, the shot noise becomes greater. This is because the electrons are freed sporadically when light energy is present. Because of the numbers of electrons involved, the absorption appears to be even over time. But occasionally a large group of electrons are freed simultaneously, producing a burst of current, otherwise known as shot noise.

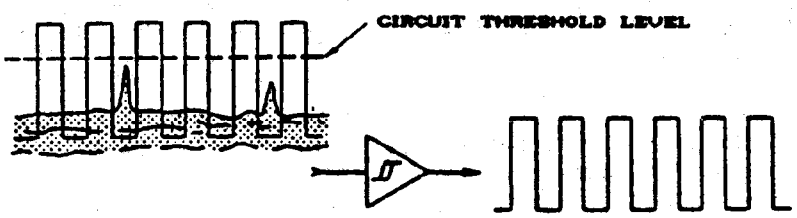
A measure of the effect of noise on a system is the signal-to-noise ratio (SNR). The SNR is the ratio of power of the detected signal and the underlying noise from all sources. It is commonly used as a measure of analog system performance.

Figure 145 illustrates the difference between a high and low SNR. The figure shows a signal and the background noise that is being applied to a pulse-restoring circuit. In Figure 145A, the SNR is high enough that the threshold detection level can be set well above the noise level. The output of the pulse-restoring circuit is then a faithful reproduction of the transmitted signal.

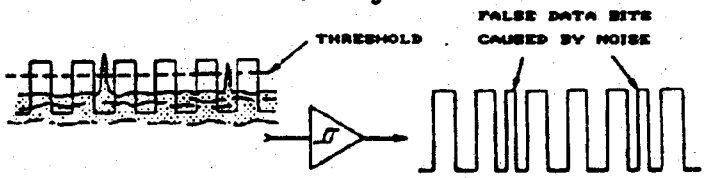
Figure 145B shows the same input signal greatly attenuated. To detect this signal, the threshold level of the pulse-restoring circuit must be set low. This makes it more likely that noise spikes will be detected and processed by the circuit, resulting in false data bits at the output. Therefore, to prevent loss of information, the detector must provide signals well above the noise level.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

4.11 DIODES, PHOTODIODES



A. High SNR



B. Low SNR

FIGURE 145. The effect of SNR on received pulse data.

4.11.5.1.5 Dark current. Dark current is the current output of a photodetector when no light is present at the input of the detector. It is caused by leakage current through the detector, which is a result of the detector's reverse biasing. Because the leakage current increases with temperature, the dark current also increases with temperature. Excess dark current can also be caused by incorrect reverse biasing.

Dark current is also very high in the avalanche photodiode, being about 10 to 100 times greater than a PIN's dark current. Another disadvantage is a high reverse voltage needed to place the photodiode into its avalanche operating region. This reverse voltage is usually in the range of several hundred volts, which is much more than the voltage required by PIN photodiodes.

The avalanche photodiode is seldom used for short distance communications because of the cost of the device and its support circuitry. In very long distance work, though, its cost can be justified because of its high responsivity.

4.11.6 Environmental considerations. Typical environmental conditions and screens of photodiodes are similar to those given in paragraphs 4.1.6.9 Environmental considerations and 4.1.6.10 of Screening procedures in subsection 4.1 Diodes, general.

4.11.7 Reliability considerations. Because the failure mechanisms for photodiodes are similar to those of other diode types, paragraph 4.1.7 General reliability considerations in subsection 4.1 Diodes, general should be consulted for a discussion of diode failure mechanisms.

4.11 DIODES, PHOTODIODES

Because planar diffused photodiodes are usually fabricated from bulk silicon material having a range of 5 to 25,000 Ω-cm for the base region, semiconductor photodiodes with a base resistivity lying in the higher end of the resistivity scale would be much more sensitive to inversion layer formation at the silicon-silicon dioxide (thermal oxide) interface than conventional diodes.

Therefore, the incorporation of a guard ring structure (channel stopper) is necessary to minimize the surface contribution to the dc reverse leakage current in order to contract the magnitude of the dark current under reverse bias because the dark current is a combination of surface leakage current and bulk leakage current. The guard ring will also help prevent premature edge voltage breakdown failures, especially for APD devices which operate at a high-reverse bias near the breakdown voltage of the device.

Figure 146 shows the total dark current as a function of the active junction area with and without the guard ring. Note that large photodiodes do not have as significant a dark current reduction with the guard because a higher percentage of the dark current is due to bulk leakage. Under operating conditions where the guard ring and active area are biased at the same potential, the surface leakage is shunted around the load resistor and flows through the guard ring. In this manner, the much lower bulk leakage becomes the limiting source of shot noise current through the load resistor. Because the shot noise current of the detector varies directly as the square root of the leakage current, the total noise performance of the detector is greatly improved by the addition of the guard ring.

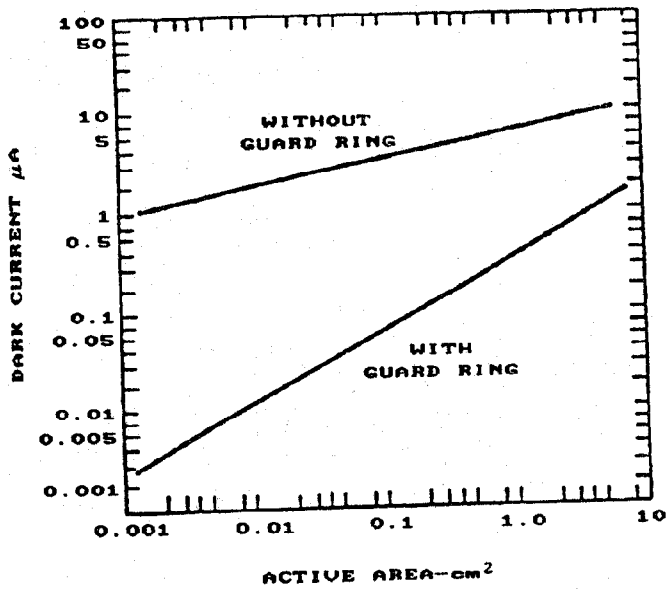


FIGURE 146. Dark current vs active junction area.

4.11 DIODES, PHOTODIODES

Surface leakage current is dependent on surface defects, cleanliness, bias voltage, and surface area. Bulk leakage current is dependent on the active area, the silicon resistivity and bias voltage. Both the surface leakage current and the bulk leakage current are temperature dependent. A good approximation for the temperature coefficient of dark current is that the dark current doubles for every 10 °C increase in operating temperature. Because dark current produces shot noise, it is evident that a reduction in dark current will improve the photo diode signal-to-noise ratio. The guard ring structure effectively accomplishes this by shunting surface leakage currents away from the load resistor as illustrated in Figure 147.

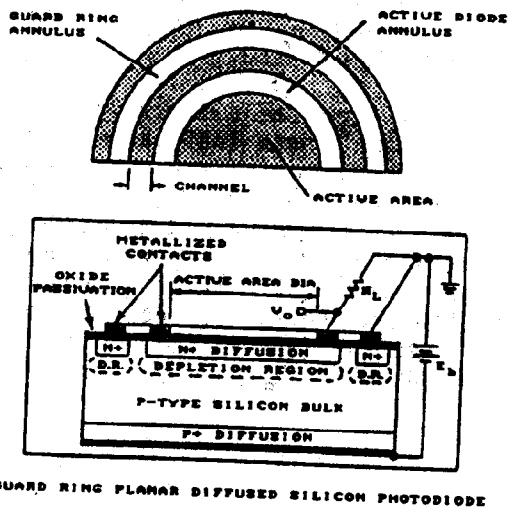


FIGURE 147. Guard ring planar diffused silicon photodiode.

5. TRANSISTORS

5.1 General.

5.1.1 Introduction. The transistor section is divided into five subsections: General; Low-Power Transistors which covers small-signal, switching, and chopper bipolar-transistors; High-Power Transistors which covers high-power bipolar-transistors; Field-Effect Transistors which covers low-power JFETs and MOSFETs devices; and Optocouplers. Radio frequency devices will be covered in Section 6.3, Microwave Transistors.

This general subsection will discuss the pertinent information that is common to all the subsections.

5.1.1.1 Applicable military specifications. There are many military documents that cover semiconductor devices. This section addresses itself to the primary documents.

MIL-Spec	Title
MIL-S-19500	General Specification for Semiconductor Devices
MIL-S-19500/---,	A group of military specifications used for procuring JAN, JANTX, JANTXV, and JANS transistors
MIL-STD-750	Test Methods for Semiconductor Devices
MIL-STD-975 (NASA)	NASA Standard Electrical, Electronics, and Electromechanical (EEE) Parts List
MIL-HDBK-217	Reliability Prediction of Electronic Equipment
DOD-HDBK-263	Electrostatic Discharge Control Handbook for Protection of Electrical and Electronic Parts, Assemblies and Equipment (Excluding Electrically Initiated Explosive Devices)

5.1.2 General definitions.

5.1.2.1 Abbreviations. Refer to MIL-S-19500 Appendix B for standard abbreviations used in this section. Only those not listed in MIL-S-19500 are listed herein.

5.1 TRANSISTORS, GENERAL

5.1.2.1.1 Chopper transistors.

I_{ECS} ..... Emitter cutoff current
$h_{FE(INV)}$ ... Static forward current transfer ratio (Inverted connection)
$r_{ec(on)}$ .... Small-signal emitter-collector on-state resistance
$V_{EC(ofs)}$ ... Offset voltage (Inverted connection)

5.1.2.2 Definitions. Definitions used in this handbook are the same as those found in MIL-S-19500 Appendix A.

5.1.2.3 Symbols. The following symbols are those generally accepted by electronic industries.

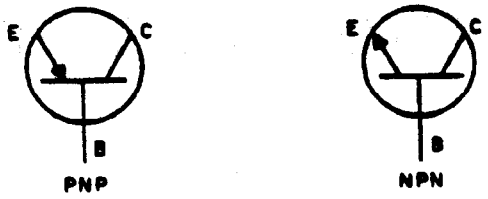
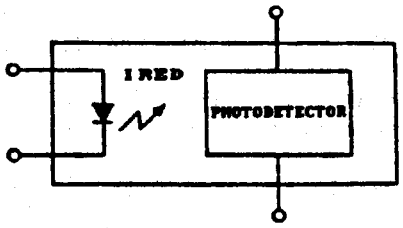
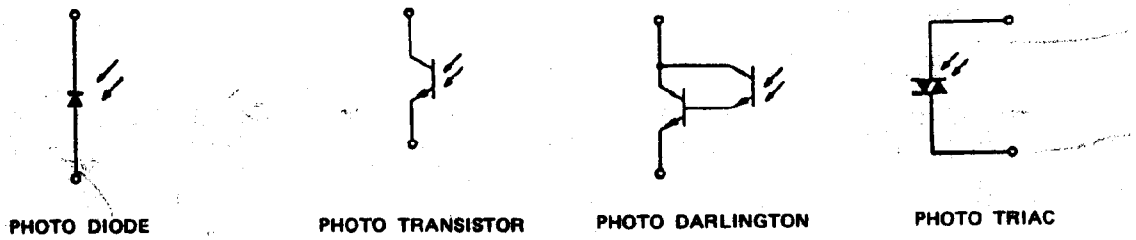


FIGURE 1. Bipolar transistors.



A. Basic optocoupler system



B. Symbol for photodetector

FIGURE 2. Optocouplers.

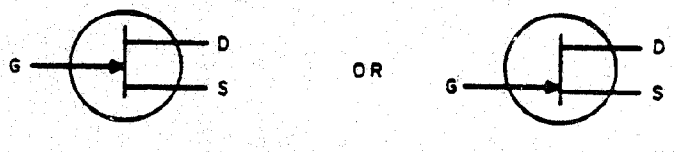


FIGURE 3. n-channel depletion mode JFET.

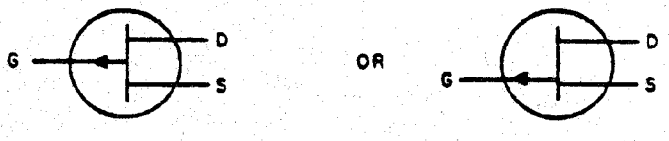
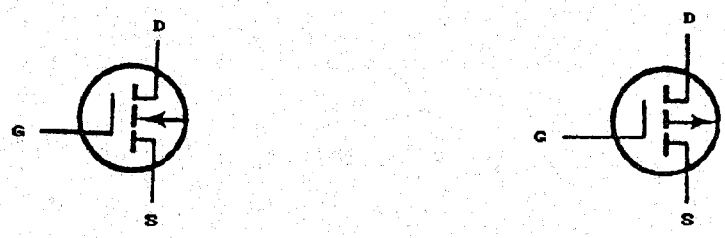


FIGURE 4. p-channel depletion-mode JFET.

n-Channel p-Channel

FIGURE 5. Depletion-mode MOSFET.



n-Channel p-Channel

FIGURE 6. Enhancement-mode MOSFET.

5.1 TRANSISTORS, GENERAL

5.1.3 NASA standard parts. See General subsection 1.2 for a complete description of the NASA Standard Parts Program. In addition to this manual, the other principal element of this program is MIL-STD-975(NASA).

MIL-STD-975(NASA) is a standard parts list for NASA equipment with Section 13 containing a summary of standard transistors.

5.1.4 General device characteristics. The following is a brief description of the four basic types of transistors covered in this section.

5.1.4.1 Low power and chopper transistors. Low power transistors (<2 W) include npn and pnp devices that are used both in small-signal and switching applications. The chopper transistor is also a low power bipolar transistor with special parameters that are useful in chopper applications.

5.1.4.1.1 Small-signal. A transistor is said to be operating in the small-signal mode when the bias is such that the largest ac signal to be amplified is small compared to the dc bias current and voltage. Transistors used this way are normally biased at currents between 0.1 mA and 10 mA and voltage between 2 and 20 V.

5.1.4.1.2 Switching. The high value of off resistance and the low value of on resistance associated with a transistor makes the device as valuable as diodes for switching applications. The transistor has one important advantage over the diode. Its state is easily controlled using the base lead because a relatively small current in the base can control a large current in the collector where as the diode can only be switched by altering its bias. This switching gain makes the transistor a more versatile device. Switching transistors generally have a frequency operating range of greater than 500 MHz.

5.1.4.1.3 Choppers. A chopper is a low-power bipolar transistor intended primarily for use in an inverted operating mode in electronic chopper circuits. Its main characteristic is its low offset voltage VEC(ofs).

5.1.4.2 High-power transistors. This category covers devices that have a power dissipation range of two watts or greater. These devices are normally used in applications where the assumptions of linear operation are not valid and where variation in collector voltage and current are a significant fraction of the total allowable range of operation.

5.1.4.3 Field effect transistor (FET). A field-effect transistor consists essentially of a current-carrying channel formed of semiconductor material whose conductivity is controlled by an externally-applied voltage. The current is carried by one type of charge carrier only, with electrons in channels formed of n-type semiconductor material and holes in channels of p-type material; therefore, the field-effect transistor is sometimes called a unipolar transistor. This is to distinguish it from the bipolar transistor, which is a junction transistor the operation of which depends upon both types of charge carriers.

5.1 TRANSISTORS, GENERAL

There are two types of FETs: the junction FET (JFET) and the insulated gate FET (IGFET). The most commonly used insulated-gate FET is the metal oxide semiconductor FET (MOSFET).

MOSFETs can be structured to operate in both the depletion and enhancement-mode while JFETs operate only in the depletion-mode. MOSFETs are common both in the depletion and enhancement-mode for low-power devices. Also, both JFETs and MOSFETs are available as n- or p-channel devices.

5.1.4.4 Optocoupler. An optocoupler is a signal coupling device usually consisting of an infrared emitting-diode and a photosensitive semiconductor enclosed in a housing that is sealed against outside light. This device is characterized by nearly perfect input-output isolation due to the extremely low capacitance between the light source and the photodetector.

5.1.5 General parameter information. Electrical parameters will be discussed in each of the individual subsections.

5.1.6 General reliability considerations.

5.1.6.1 Failure modes and mechanisms. There are four failure modes for transistors: parametric degradation, shorts, opens, and mechanical defects. Tables I through IV describe the failure mechanisms for these failure modes, the methods of detecting the failure-causing defect before the parts are accepted by the user, and methods for reducing the probability of obtaining a part with such defects or of causing defects in good parts. It should be pointed out, however, that a very small percentage of parts will elude detection, and in-service failures will be encountered even with the most rigorous screening and process control methods.

5.1.6.1.1 Parameter degradation. Parameter degradation is the degradation of the electrical characteristics of a part. It is the most common failure mode and the most costly to detect. Although it usually does not result in a catastrophic failure of operating hardware, it can seriously degrade hardware performance if sufficient design margin is not allowed. Furthermore, it is usually indicative of a part whose expected life has been severely shortened. The mechanisms contributing to a parameter degradation are shown in Table I.

5.1.6.1.1.1 Contamination mechanisms. Contamination and corrosion are the most prevalent causes of parameter degradation and can be categorized by the process in which they are introduced: diffusion, oxide growth, washes, and gas ambient sealed in the can. Contaminated doping materials and ambient atmospheres during the diffusion processes result in wafers with a low yield of acceptable transistor die, usually as a result of improper resistivities or channel forming ions at the silicon-silicon dioxide interface. Bad dice are normally removed by the manufacturer through electrical testing.

5.1 TRANSISTORS, GENERAL

a. Contamination detection methods. Elevated temperature bake-out for several hours in an inert atmosphere prior to capping is necessary to drive off surface contaminants. High temperature storage of sealed parts is effective in stimulating the effects of gas ambient and surface contaminants. The most effective method of detecting surface contaminants is through high temperature reverse bias (HTRB) testing. The temperature sufficient to activate the contaminating ion is usually 150 to 175 °C. Reverse bias of both junctions to achieve 10^6V/cm field strengths is desired as it moves the ions to localized areas so that they can be detected on post HTRB leakage current measurements. Higher temperatures permit shorter test times but should be carefully considered to avoid starting other time dependent mechanisms, such as the Kirkendall effect. Burn-in at the rated junction temperature will aggravate the effects of bulk contaminants as well as corrosive contaminants. Large increases in leakage or degradation of gain will identify devices which are contaminated.

b. Contamination minimization. Use of phosphosilicate glasses or silicon nitride passivations has been shown to be effective in gettering alkalis, thereby reducing their effects.

5.1.6.1.1.2 Bulk and oxide defect mechanisms. The next class of defects causing parameter degradation is somewhat related to contamination existing in the part. Contamination to some degree is inevitable so the more perfect the bulk and oxides, the less the effect of contaminants will be. Die faults and dislocations act as localization zones for mobilized gain failures.

Another phenomenon that is especially common to large area dice is inclusions of polycrystalline silicon, during the epitaxial growth, due to particulate contamination on the prepared silicon substrate surface. These propagate through the collector base junction and cause low breakdown voltages and high leakage currents.

Defective oxides having pin holes, cracks, misaligned oxide cuts, thin spots and other defects will contribute to several mechanisms. In bipolar transistors, mobile ions can migrate through discontinuous oxide to the silicon-silicon dioxide interface, depleting the base and thus degrading gain. In FETs and bipolar transistors, the discontinuities provide leakage paths from metal connections to active areas.

5.1 TRANSISTORS, GENERAL

TABLE I. Parametric degradation

Failure Mechanism	Description	Method of Detection	Method of Minimizing Defect
Surface contamination	Ionic material on surface of Si or SiO_2 providing leakage paths (Cl, Fl, etch) Alkali ions in or on oxides, creating channels (Na^+, CA^{++}) Precipitated dopants on surface of silicon during oxide growth	High temperature reverse bias testing (HTRB) with temperature sufficient to mobilize contaminant and field concentrations sufficient to localize it, monitor leakage, and gain variations before and after test	Proper cleaning and etching sequences and recycling of solutions Practically unavoidable Keeping oxide growth temperature below diffusion temperature
Contaminated gas ambient	Noninert gases in the capping station Nonhermetically sealed cans Evaporated contaminants from die surface	Fine and gross leak seal tests to detect nonhermetic seal, sample residual gas analysis to detect lot problem with sealed ambient	Sealed capping station supplied with pure, dry, N_2 gas at pressure greater than 1 atm High temperature bake-out prior to capping
Bulk defects	Die cracks in active areas providing leakage paths Inclusions, dislocations, stacking faults, and other discontinuities providing localization zones for impurities Poor geometry design leading to areas or current crowding and hot spots	Power cycling will aggravate cracks; power aging at high temperature to localize impurities in fault areas, increase leakage and lower gain and breakdown voltage; close monitoring of gain and leakage before and after aging to detect parts with unstable characteristics	Visual inspection of each die, rejecting cracked ones Virtually impossible to eliminate all dislocations Careful die layouts and thermal analysis using thermal microprobe on an operating device to locate hot spots in design
Oxide defects	Cracks, holes, and thin spots providing leakage paths to exposed silicon	HTRB with temperature sufficient to mobilize contaminant and field concentrations sufficient to localize then monitor leakage and gain variations before and after test	Gross defects detectable with a visual inspection after oxide growth, high temperature growth desired to form dense, uniform oxides
Die attachment	Voids in eutectic which results in increased thermal impedance	X-ray normal to die looking for voids under die, precap visual inspection looking for voids around die	Proper temperature, scrubbing time, materials and cleanliness prevents voids
Interconnect	Resistive contact of metallization to silicon due to oxide growth in the window	High temperature storage will aggravate these phenomena, forward diode drop at each junction should be monitored, watching for slight increases (50 - 100 mV)	Complete oxide cuts and proper alloy procedure Temperature should be sufficient to form a Si-Al alloy Use monometallic interconnects, preferably gold wire, gold metallization
Electrical stress	Electrostatic discharges rupturing low leakage junctions Excess power and current	Identify all devices with 1 nA leakage current or less as susceptible and protect accordingly	Short all leads together until part is installed, handle in electrostatically neutral environment Current limits, transient suppressed power supplies and adequate oscillation suppression for dc operational

5.1 TRANSISTORS, GENERAL

Other bulk defects which affect transistor operation, especially interdigitated power types, are misalignment of base and emitter diffusions or gradients in diffusion rates of the base and emitter that lead to irregular base widths and variations in emitter resistivity. These phenomena lead to localized secondary breakdown due to hot spot formation and punch-through under reverse bias. Die cracks which propagate into diffused areas reduce the breakdown voltage and increase the leakage current and ultimately produce shorted junctions.

a. Bulk and oxide defect detection methods. Generally most defective dice are removed by the manufacturer prior to part assembly through wafer-level electrical tests. Visual die inspection will also eliminate dice with gross defects.

   Thermal shock or power cycling will aggravate die cracks. HTRB and power burn-in at rated junction temperature with delta reject criteria on gain and leakage measurements will detect the majority of parts with time-dependent failures due to oxide and bulk defects.

b. Bulk and oxide defect minimization. Discovery of most bulk defects during die fabrication is expensive and tedious usually involving high magnification microscopy. Bulk defects seem to be tolerated by industry provided that yields are high enough; however, since transistor dice with some bulk defects are unavoidable at the present state of semiconductor processing technology, they pose a life-limiting failure mechanism for a small population of parts.

   Cleanliness and high temperature growth usually yield a very good grade oxide. A sample SEM inspection of the wafers will help detect oxide defects.

5.1.6.1.1.3 Die attach mechanisms. One of the most critical parameters of power devices and many small signal transistors is not electrical but thermal. Thermal resistance ($R\theta$) and junction-to-case ($R\theta_{JC}$), affects almost all other parameters because most parameters are dependent upon junction temperature. If this is too high, catastrophic part failure results. The primary interfaces affecting $R\theta_{JC}$ are the silicon-to-eutectic, eutectic-to-header, or insulator, and header- or insulator-to-stud. The most common defect affecting $R\theta_{JC}$ is voiding of the eutectic under the die leading to hot spots on the die and localized avalanching.

Voids can be caused by oxidized or contaminated silicon on the back of the die which precludes the formation of a good eutectic phase, intermediate phases formed in the eutectic die from diffusion or migration, inadequate gold-silicon preforms or platings, and poor die attach procedures resulting in the formation of undesirable intermetallics.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

a. Die attach defect detection methods. X-ray of the header prior to capping, with the X-ray axis through the die surface, could be used to detect voids. Pre-cap visual inspections will detect voids around the edges of die but not under the die. Power aging at maximum junction temperature and high temperature reverse voltage tests will detect some of these failures. A nondestructive secondary breakdown test, increasing voltage to an acceptable threshold at elevated temperature, would eliminate many parts with hot spots on the die which contribute to premature secondary breakdown.

b. Die attach defect minimization. Clean unoxidized surfaces and proper time-temperature controls for the attaching materials used are the only effective means for achieving proper die attachment. Gold-silicon eutectic materials afford the most reliable die solders, but sometimes even these may be unacceptable due to interaction with other part materials and instabilities with temperature and aging. Parts with soft solder die attachments should be carefully evaluated if power cycling requirements greater than 6000 cycles are required. Epoxies are not acceptable for use in high reliability applications.

5.1.6.1.1.4 Interconnect mechanisms. Interconnect flaws are another category of parameter degradation due to built-in defects. Interconnect wire degradations will be discussed later in the section dealing with opens, but their initial effect is an increase in contact resistance which increases the high current voltage drop in the forward direction.

Another phenomenon which increases contact resistance is the growth of silicon dioxide in the contact window; it is usually a result of incomplete etching of the silicon dioxide in the window.

a. Interconnect defect detection methods: High temperature storage will aggravate both phenomena. High current forward voltage drop measurements, especially emitter-to-base, will usually detect parts with resistive contacts.

b. Interconnect defect minimization. Proper cleaning of the base and emitter contact windows prior to metallization will help reduce interconnect defects.

5.1.6.1.1.5 Excess voltage, current, and power mechanisms. The last category of parameter degradation deals with the effects on structurally and electrically sound parts due to excess power, voltage, and current conditions. Whereas most overpower and current conditions result in shorts or opens, moderate excess power or current can lead to an increase in leakage and a decrease in gain as well as break-down due to formation of small channels or punch through sites in the junctions which are resistive in nature.

Degradation caused by overvoltage or current can be avoided through careful stress and thermal analysis of the part application and care in part and system level test conditioning.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.1 TRANSISTORS, GENERAL

5.1.6.1.2 Shorts. Electrical shorts and their related failure mechanisms are shown in Table II. Most shorts are a result of poor workmanship by the part manufacturer and are easily detected by a vigorous pre-cap visual inspection. Parts which have permanent electrical shorts are rejected the first time they are electrically tested. Intermittent shorts usually occur during part test or installation and are very difficult to detect after the part is capped. These failures can be catastrophic to operating hardware if proper circuit redundancy is not provided.

5.1.6.1.2.1 Contamination mechanisms. The leading cause of shorts is conductive contamination internal or external to the part. Aluminum flakes that accumulate on probe points during the wafer level test that drop on the die across metallization paths, large fragments of silicon contamination that are introduced at the wafer level, excess eutectic that accumulate over the sides of the die, fragments of eutectic either attached or loose, and fragments and silvers of steel wool from tweezer cleaning, are common forms of contamination introduced during die attachment. Excess pigtails and extra wires and balls from the wire cutting operation are types of conductive contaminates introduced during wire bonding. Improper header fabrication and plating results in flakes and slivers of Kovar or gold causing internal or external shorts. Improper storage and cleaning of die header assemblies and cans results in airborne contaminants, which are then sealed into the part. Weld slag balls can be introduced during the capping operation.

a. Contamination detection methods. Bound particles can often be loosened with high-impact shock test. Most free particles can be detected with an electrically monitored vibration test, but this is usually expensive. Several part-users employ a technique referred to as particle impact noise detection (PIND) testing. Acoustic noise generated by particle impingement on the sides of the device can be monitored during random vibration coupled with a mechanical shock pulse. X-ray can detect large particles, two mils or greater, that are opaque to X-rays; this includes gold and iron.

b. Contamination minimization. Conductive particles smaller than the smallest interconnect separation and nonconductive particles are not considered important. Ultrasonic cleaning of uncapped assemblies is one method of removing contaminants prior to capping. Glassivation over metallization is the best defense against conductive particles. A 100-percent pre-cap visual inspection would help weed out devices with contaminants.

TABLE II. Shorts

Failure Mechanism	Description	Method of Detection	Method of Minimizing Cause
Contamination	Aluminum flakes from probe points, silicon fragments, excess eutectic, excess pigtails, extra wires, weld slag balls	X-ray Acoustic particle detection 100% precap visual	Ultrasonic cleaning Glassivation
Assembly defects	Improper die orientation, excess wire loops, bent posts, tall posts, misaligned bonds	100% precap visual X-ray Constant acceleration	Proper device design, use of proper assembly techniques, and implementation of in-process quality control check points
Die defects	Masking alignment, improper diffusion depth, lifting of undercutting of photoresist discontinuities, cracks, poor die attach, poor wire bonding, holes or thin spots in the oxide, metallization smears	100% precap visual High temp storage HTRB	Glassivation Guard rings
Excess voltage current and power mechanisms	Melted metallization, discolored or melted die vaporized bond wires	100% precap visual	--

5.1 TRANSISTORS, GENERAL

5.1.6.1.2.2 Assembly defect mechanisms. Assembly defects such as improper die orientation that forces interconnect wires to cross, connection of the wrong contact to the external pins, excess wire loops, down bonding where inadequate clearance is provided between the edge of the die and the wire, misalignment of wire on bond pad, and bent posts or those that are too high so that they touch the side of the can, can cause shorts.

a. Assembly defect detection methods. A precap visual inspection eliminating parts with such defects is desired. X-ray will show crossed wires and improper post-to-cap clearance. Constant acceleration in all three axes will cause excess wire loops to touch other areas in the can. Constant acceleration in the Y1 direction will also tend to lift poor wire bonds off the bonding pads.

b. Assembly defect minimization. Proper device design, the use of proper assembly techniques, and implementation of in-process quality control check points will greatly reduce assembly defects.

5.1.6.1.2.3 Die defect mechanisms: Die defects causing shorts can be placed in three categories: bulk, oxide, and metallization. The bulk and oxide defects which cause parameter degradation, discussed previously, often produce shorts. Masking misalignment and improper diffusion depths can result in shorted junctions. Lifting and undercutting of the photoresist or mask defects and misalignment can result in bridged junctions or metallization. Dislocations, inclusions, and other discontinuities in the junctions can also cause shorts.

Cracks in the die as a result of scribing or improper die-attach and wire bonding provide channeling paths along exposed silicon surfaces. Cracked dice can also result from gross mismatches in the temperature coefficient from header to die in power devices. Cracks, holes, and thin spots in the oxide under metal paths frequently result in shorts to the active areas under the metal for expanded contact devices (wire bond at the die is made on the oxide rather than at the oxide window). The contaminants on the die surface or buried in the oxide often produce inversion layers and channels sufficient to short junctions. Smears of metallization due to improper handling during the die-attach or wirebonding operation may also short active areas.

a. Die defect detection methods. A good precap visual inspection will eliminate most gross die defects. High temperature storage and HTRB will aggravate oxide and bulk defects and mobilize contaminants to form shorting channels and inversion layers.

b. Die defect minimization. Glassivation of dice has been successful in reducing smearing of metal. Thick eutectics or temperature coefficient matching phases such as molybdenum tabs are desired in power devices to eliminate thermally induced cracks. Channel stoppers are desired around the outside of the die to keep channels from propagating to the edge of the die in single diffused epitaxial devices, and guard rings

or field plates around the base collector junction are often used to prevent channels in other geometries. Extension of base metallization over the base-collector junction line is a commonly used type of field plate.

5.1.6.1.2.4 Excess voltage, current, and power mechanisms. The last of the shorting mechanisms is electrical overstress. The most common is excess current causes aluminum from one metallization path to traverse across the surface of the silicon, underneath the oxide, to a metal path of different potential and forming what are known as white spears. Bulk breakdown due to localized thermal avalanche, punch through, and secondary breakdown is usually caused by excess temperature with bias applied. High frequency oscillations often occur during burn-in or circuit operation where degenerative feedback is not provided. Melted metallization, discolored or melted dice, and vaporized bond wires are all signs of misapplication of bias and signal or improper heat sinking.

5.1.6.1.3 Opens. The failure mechanisms producing opens are displayed in Table III. These mechanisms are frequently time-dependent and often require long periods to produce failures, and they can also be intermittent. These mechanisms pose a considerable question to the long-life performance of transistors. Open transistors are generally less catastrophic to hardware failure than shorted ones because they do not load power supplies, but they can still result in hardware failure due to loss of drive to succeeding stages. Redundancy is necessary where single-point transistor opens can cause catastrophic hardware failure.

5.1.6.1.3.1 Interconnect wire failure mechanisms. The most common cause of opens is interconnect wire failure. There are two types of metallization-interconnect wire systems used, gold wire-to-aluminum metallization and aluminum wire-to-aluminum metallizations. Each causes problems when not done properly.

The most common interconnect system in transistors is the gold wire-to-aluminum metallization interconnect. Insufficient time, temperature, or pressure during the formation of the bond results in weak or incomplete bonds.

A more significant mechanism is the Kirkendall effect, commonly called plague. Seemingly strong bonds grow weak with temperature and time due to intermetallic compound formation. Bond weakening is most severe in expanded contacts, but failures have also been observed in direct contacts.

The Kirkendall effect is the interdiffusion of two dissimilar metals at different rates producing voids at the metal interface. This phenomenon takes three forms, annular voids around the outer periphery of the ball, lateral voids along the entire surface of the ball, and depletion of the aluminum film.

5.1 TRANSISTORS, GENERAL

TABLE III. Opens

Failure Mechanism	Description	Method of Detection	Method of Minimizing Defects
Open interconnect wire at the die NOTE: The most common intermittent open-normal-open failure is due to die-wire bond failures	Gold-aluminum: Intermetallic formations (plague) Kirkendall voids under ball Excess pressure resulting in nicked wire exiting ball Depletion or peeling of Al film around and under ball Aluminum-aluminum: Microcracking at heel of bond due to poor tooling and excess bonding pressure Both: Insufficient time, temperature pressure, or energy leading to incomplete bonds Insufficient bond pad area on die	100% bond pull (non-destructive to an acceptable threshold level) to detect weak or poorly formed bonds. Threshold for Au-Al 2 grams min pull, Al-Al 0.7 gram min. Low duty power cycling will aggravate microcracking in Al-Al wedge bonds. High temperature storage or cycling followed by mechanical shock will aggravate plague and void failures	Use monometallic bonding system Oxygen in gas backfill seems to retard growth Visually inspect all bonds for improper ball dimensions and wire nicking Proper alloying (sintering) of Al film prevents peeling Proper tooling design and controlled bonding parameters Control over bond parameters Bond pad area must be greater than the final surface area of the bond (i.e., metallization visible around bond)
Open interconnect wire at the post and along the wire	Insufficient weld area at the post Burned or incomplete welds Si inclusions in wire or greater than 1% content (Al wire) Twists, nicks, crimps, kinks, or scratches in wires which reduce the wire diameter Corrosion on or in the wire from drawing or cleaning operations which corrodes metallization or otherwise interferes with bonding	Visual inspection of wires and bonds will detect gross defects, 100% bond pull (nondestructive to an acceptable threshold) will detect weak wires and bonds, thermal shock will aggravate incomplete or improperly formed post bonds, high temperature storage followed by thermal cycling will anneal and break poor quality aluminum wire	Use double bond at the post Proper weld or bonding schedule Close metallurgical surveillance of wires used, 1% Mg wire appears to be stronger Workmanship Clean all wires before use
Die lifted from header	Voids in eutectic die mount Failures of ceramic insulation to support die under mechanical stress (for dielectrically insulated parts) Cracked die due to thermal mismatched of die and header (power devices)	Constant acceleration at high g levels will expose these failures	Workmanship Mount ceramics directly to header, do not support from leads Use thick eutectics or intermediate materials to provide some stress relief
Metallization	Photolithography defects (stained or dirty masks, lifted or underexposed photoresists, etc.) resulting in incomplete metallization Improper annealing (sintering) resulting in peeling or lifting metallization Scratches opening metal paths Scratches reducing cross-sectional area of metal resulting in migration failures (expanded contact and interdigitated geometries) Insufficient design width and thickness resulting in excess current densities and burned metal or migration	Electrical test will detect permanently open parts, visual inspection of die will detect gross metal defects, power burn-in at reasonably high current levels will accelerate migration failures NOTE: Electromigration of Al is only a problem on devices with expanded contacts and is discussed more in the microcircuit section	Clean masks, uniform photoresist deposition exposure and removal will help Sintering temperature should be reasonably high Glassivation over metal to protect it Glassivation over metal to protect it Maximum current densities should be 5×10^4 A/cm^2 to minimize migration effects

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.1 TRANSISTORS, GENERAL

The next most widely used bonding system in transistors is aluminum wire, with 1 percent silicon, and aluminum metallization. The two failure mechanisms in this system are wire breakage due to silicon inclusion in the wire and microcracking at the heel of the bond due to necking of the bond by the bonding tool. Both failure mechanisms are time dependent and are a function of the thermal cycling, usually low rate, which causes mechanical movement and annealing of the wire.

Both gold:aluminum and aluminum:aluminum interconnect systems are prone to failure if the post bond does not provide sufficient wire-to-post surface contact. In addition, nicks and kinks in the wire, which reduce its designed diameter, could result in failure due to excessive current densities.

a. Interconnect wire failures, detection methods. In both gold:aluminum and aluminum:aluminum interconnect systems a stringent 100 percent precap visual inspection of the die and post bonds and the wire will detect most common interconnect defects. A sample bond wire pull test could also be performed to assure that a good bonding schedule is being maintained by the manufacturer.

b. Interconnect wire failures minimization. The majority of suppliers use gold:aluminum or aluminum:aluminum interconnect schemes. Double post bonds are recommended for either system to provide good adhesion to the post. Gold ball thermocompression bonds and ultrasonic aluminum bonds are some of the acceptable means of bonding the wire to the die. Wire quality and material content as well as metallization thickness and texture must be controlled by the manufacturer for reliable bonding. Bond location and adequate pad size also are important controlling factors in reliable interconnects.

5.1.6.1.3.2 Lifted die mechanisms. Excessive voids in the eutectic bond or undue mechanical stress can cause the die to lift off the header. In parts with dielectrically isolated collectors or gates, ceramic insulators which do not have adequate mechanical support to the header can break dice during mechanical stress and cause opens.

a. Lifted die detection methods. Parts which have marginal die attachments can usually be detected through X-ray of the die header assembly, when looking for voids. Power pulse $V_{CE}(SAT)$ and $R_{\theta JC}$ measurements will also detect such defects.

b. Lifted die minimization. Die and insulators should be firmly attached to the header rather than suspended from the leads. Ceramic temperature coefficients should be matched closely to silicon to preclude cracking due to thermal expansion.

5.1 TRANSISTORS, GENERAL

5.1.6.1.3.3 Metallization failure mechanisms. Metallization opens are usually a problem only in expanded contact devices since current is carried from the die to the wire bond by paths of metallization. Scratches due to improper die handling and missing metallization due to photolithography defects are commonly the sources of open metallization paths. Aluminum can also migrate from thin areas to create voids. This is a problem in rf and power devices where the metallization cross-sectional area and current requirements cause excessive current densities in the aluminum film. Improper alloying (sintering) of the aluminum film to the silicon and silicon dioxide surfaces results in peeling and lifting of the metallization. This also can be a problem in direct contact devices. Another mechanism is failure of the metallization to make contact with silicon due to an incomplete etch of the silicon dioxide at the window or growth of silicon dioxide at the interface of the silicon and aluminum. Molybdenum gold metallization system failures are sometimes caused by excessive undercutting of the molybdenum during etching or inadequate alloying of the molybdenum.

a. Metallization failure detection methods. A precap visual inspection will detect open or degraded metallization fingers in uncapped parts. Electrical testing will reveal those sealed parts having open metallization paths. Metal voiding due to migration can be discovered by a forward burn-in at rated current. Several integrated circuit users require a sample SEM inspection of each wafer to determine adequate metal coverage over oxide cuts.

b. Metallization failure minimization. Use of phosphosilicate glasses retards metal migration, except in gold metallization systems, and also protects the metal surfaces from scratches during die handling. High temperature storage will stimulate the growth of oxide at the contact window. The use of gold metallization significantly reduces metal migration effects because of the relative immobility of gold; however, while the gold is immobile, the moly underneath is not.

Derating maximum current limits to keep metallization current densities below 5×10^4 A/cm^2 will considerably reduce electromigration effects.

5.1.6.1.3.4 Excessive voltage, current, and power. Usually, overpower conditions will melt interconnect wires or vaporize metallization films due to excessive current density. Shorts in the die also cause open interconnects under normal power conditions.

5.1.6.1.4 Mechanical degradation. The last group of failure mechanisms is shown in Table IV and deals with failures that prevent proper installation of the part. While these failures are normally discovered during and just after part installation, they cause significant delay and occasionally go undetected until a considerable investment has been made in the using hardware. They are most easily detected prior to installation by an external visual inspection.

TABLE IV. Mechanical defects

Failure Mechanism	Description	Method of Detection	Method of Minimizing Defects
Unsolderable or unweldable leads	Finger greases or acids on leads	External visual inspection, dimensional inspection, solderability/weldability tests	Use of tweezers and gloves when handling parts
	Exposure of leads to corrosive solvents		Control of cleaning processes and storage of parts
	Exposure of part to O_2 and H_2O in high temperature environments		Use of dry N_2 environments during high temperature storage
	Poor plating		QC inspection by part manufacturer of plating integrity
	Improper metal content		QC inspection by part manufacturer of lead metals
	Improper dimensions		QC inspection by part manufacturer
	Bent, broken, nicked, scratched, or twisted leads due to improper handling or test equipment		QC inspection by part manufacturer
Degraded seals	Cracked glass due to mishandling	External visual inspection; hermetic seal tests, fine and gross, with maximum leak rate of 5×10^{-8} Atm/cc/sec	100% inspection of headers upon receipt by part manufacturer
	Bubbles or voids in glass		100% inspection of headers upon receipt by part manufacturer
	Improper can-to-header seal ring dimensions		100% inspection of headers upon receipt by part manufacturer
	Contaminated can-to-header seal ring		100% inspection of headers upon receipt and proper cleaning
	Improper can-to-header weld schedule		Regular certification of weld schedules and welders by part manufacturer
Degraded part package	Exposure to corrosive solvent or gases	External visual inspection, dimensional inspection	Store and test parts in dry inert environment
	Finger acids and greases		Handle parts with tweezers or gloves
	Improper dimensions and material content		QC inspection of headers and cans by manufacturer
	Deformation of part can or header due to excessive mechanical stress		Control of mechanical shock and acceleration fixtures and handling and packaging
	Galled or stripped threads and broken studs for stud-mounted packages	External visual inspection, thermal impedance measurement base-to-heatsink on installed parts	Use of properly derated stud torque values by part user; burn-in power parts in free air using free air ratings to achieve max temperature
	Poor finish and flatness on bottom of stud surface precluding proper thermal interface		QC inspection of headers and lot test of thermal impedance case to heat sink by part manufacturer
	Improper insulator resulting in poor mechanical support or high thermal impedance for collector or gate insulated devices	Thermal shock/high impact shock with electrical measurements prior to and post shock thermal impedance measurement junction-to-case	Use ceramic insulators which are solid (unperforated or slotted) that are eutectically mounted to part header rather than supported by lead parts
Mismarked or unmarked	Wrong part numbers placed on part	External visual inspection	Workmanship
	Unintelligible marking smears, blobs		Workmanship
	Tags applied to leads		No tags applied to leads, stamp part ID on package
	Part marking inks/paints soluble in cleaning solutions	Marking permanency test	Use of indelible inks/paints

5.1 TRANSISTORS, GENERAL

5.1.6.1.4.1 Degraded leads and packages. Corrosion of leads and cans results from any one of a number of poor practices. Handling of parts without hard protection and storage of parts in corrosive or oxidizing environments degrades leads preventing sound solder joints when the part is installed. Oxidized surfaces on the can where heat flow is required increase thermal impedance. Poor plating also contributes to solderability problems and reduces thermal conductivity. Improper material content and use of improper dimensions result in poor welds when contact resistance welding is used. Contamination can result in weakened joints. Scratched, bent, broken, and twisted leads make the part more difficult to install properly.

5.1.6.1.4.2 Studs. Galled or stripped threads, irregular or hollow surfaces on base of stud, and improper attachment of ceramic insulators, for dielectrically isolated devices, result in failures of stud mounted devices. The latter two usually cause a marked increase in thermal impedance of the device resulting in overheating and die failure. Damaged threads usually result in broken studs when proper torque is applied or in insufficient torque due to added mechanical resistance.

5.1.6.1.4.3 Seals. Cracks and bubbles around the external leads result in loss of hermeticity. The effects of a contaminated ambient have been discussed previously.

5.1.6.1.4.4 Part marking. Soluble inks, smeared or smudged marking, or absence of marking may cause the loss of traceability of screened qualified parts. The use of insoluble inks is mandatory. All part marking is to be stamped on the can. Do not use tags on leads or cans which may outgas into the hardware in which the parts are used or which would otherwise degrade the parts. An external visual inspection of all parts will eliminate such defects.

5.1.6.2 Transistor failure rate models. Failure rate models for these devices can be found in MIL-HDBK-217.

5.1.6.3 Environmental considerations. All MIL-STD-975 transistor devices require a vigorous screening and conformance testing per MIL-STD-19500 and its applicable slash sheet. These tests are performed to assure that these devices can withstand certain levels of environmental conditions; for example, pressure, vibration, moisture, temperature cycling, mechanical stress, and, if required, radiation. The user of these devices should take precautions so that these levels are not exceeded in the design of a system. Nevertheless, the following are some of the important environmental considerations.

5.1.6.3.1 Electrostatic discharge. One of the environmental factors which is not screenable is electrostatic discharge (ESD); therefore, the user must take special precautions when handling ESD sensitive devices; refer to DOD-HDBK-263 for guidance.

5.1.6.3.2 Radiation considerations. To insure that a transistor functions properly in space applications, the design engineer must consider the effects of radiation exposure. A designer of radiation resistant systems must know how radiation affects the circuits and components of the system. When semiconductor devices are exposed to radiation environments, changes occur in their rated electrical parameters. The magnitude of the changes is a function of such things as the type of radiation (neutron, proton, electron, gamma or heavy ions) and the duration of exposure.

Transistors exposed to radiation may display changes in the following electrical characteristics:

a. Breakdown voltage decreases

b. V_{CE} (sat) increases

c. Decrease in h_{FE}

d. Leakage current increases

e. Decrease in $V_{GS(th)}$ N-channel MOSFET

f. Increase in $V_{GS(th)}$ P-channel MOSFET

g. $V_{DS(on)}$ increases

Therefore, when designing with transistors, the design engineer must consider the fact that these parameters will change. The degree of parametric alteration and the long term effects are dependent upon the device technology and the manufacturing process. These vary from vendor to vendor and device to device. Power MOSFETs in particular are very susceptible to total dose induced radiation damage, and non-radiation hardened devices may fail at levels as low as 2 Krads (Si). Also, recent testing has shown that these devices exhibit catastrophic burn-out failure when exposed to energetic heavy ions. Since the cosmic rays present in natural space environment do contain some heavy ion flux, it is recommended that, for critical applications the specifics of each device type should be discussed with a radiation expert and/or manufacturer. For a more detailed discussion of radiation and radiation effects, see the section in microcircuits, paragraph 7.1.7.7.

5.1.6.4 Application considerations.

5.1.6.4.1 Thermal considerations. Temperature is one of the most critical reliability factors. For example, if the failure rate for a device is 0.01 failures/10^6 hours at a junction temperature of 75 °C, at 150 °C the failure rate would be 600 times that value, assuming a surface-inversion failure mechanism with an activation energy of 1.0 ev, which is a common failure mechanism for silicon transistors. Therefore, precautions must be taken to maintain a device junction temperature at the design operating level throughout its useful life. Some of the design precautions that can be taken are as follows:

5.1 TRANSISTORS, GENERAL

a. Design for the lowest feasible operating junction temperature.

b. Heat sink design and device mount down are crucial. For instance, a reduction in $R\theta_{JA}$ of 1.0 C/W for a power device dissipating 25 W would reduce the operating junction temperature by 25 °C.

c. Assure that the device mount down surface is flat or slightly convex to provide proper seating.

d. The device case temperature should be measured under actual operating ambient conditions to assure that the design operating junction temperature is not being exceeded.

$$T_J = T_C + R_{\theta JC} * P_d$$

e. Allow sufficient latitude in circuit design to accommodate some changes in device parameters with time.

5.1.6.4.2 Electrical considerations. To assure proper device operation throughout the lifetime of a system the following design criteria should be considered:

a. Allow for aging of electrical parameters.

b. Protect critical devices from transients by the use of snubbers, clamps, or other protective means.

c. Keep in mind device parameter variation with temperature when designing a circuit.

d. Beware of the unspecified parameter. It is possible that a JANTXV2N3714 from one vendor will function well in a circuit whereas the same device from another vendor will not, even though the device meets all specified parameter limits. When possible, devices from all possible sources should be tested under actual operating conditions.

5.1.6.4.3 Mechanical considerations. A mechanical reliability problem for hermetically sealed devices is handling. Bend leads can break the nonhermetic seal and can also break or loosen bond wires. Careless handling can cause particles to be released inside the package and that causes problems; therefore, care must be taken when handling hermetically sealed devices.

5.1.6.5 Screening. Both environmental and electrical screening are performed in accordance with the MIL-S-19500 slash sheet on all MIL-STD-975 devices. Users of these devices should become familiar with the test requirements to assure that they are not exceeded in the design of a system.

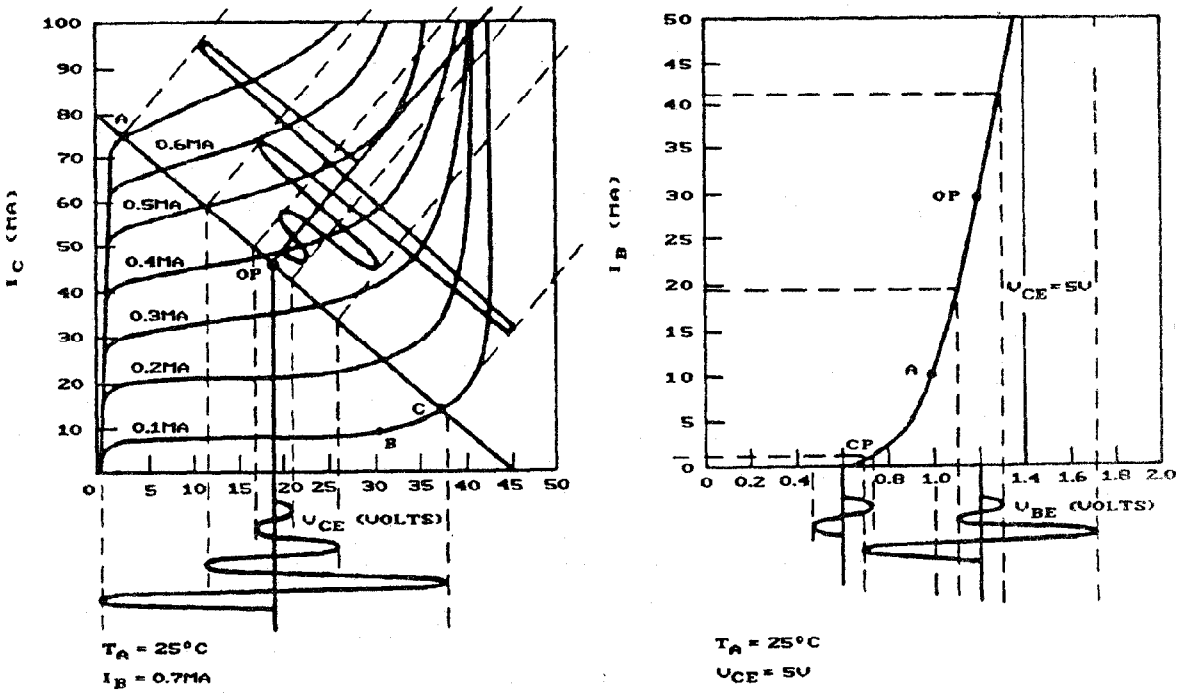
5.1.6.6 Derating factors. Derating factors for these semiconductor devices are in MIL-STD-975, and the user should refer to this document for derating factor guidelines.

5.2 Low-power.

5.2.1 Introduction. In recent years, integrated circuits have rapidly replaced many of the low-power functional circuits in which low-power discrete transistors were used. This is especially true in the case of digital and linear applications. The user should also refer to the microcircuit section of this handbook concerning these types of applications.

This low-power section will cover small-signal switching, and chopper transistors.

5.2.1.1 Small-signal. The transistor is a nonlinear active device. Figure 7 illustrates that, although slightly nonlinear throughout its range, the transistor's nonlinearities become very pronounced at the very low and high current and voltage levels below point A and above point B. If an ac signal is applied to the base of a transistor without dc bias, conduction would take place only during one half-cycle of the applied signal, and the amplified signal would be highly distorted. To avoid this problem, a dc bias operating point "OP" is chosen (see Figure 7). This bias moves the transistor's operation to the more linear portion of its characteristics. There the linearity, although not perfect, is acceptable and results in amplification with low signal-distortion.



Collector characteristic

Base characteristic

FIGURE 7. Typical V-I characteristics of a silicon planar transistor.

5.2 TRANSISTORS, LOW-POWER

The application of a dc bias is still not sufficient. A transistor could be biased exactly in the middle of its linear range and be operated at such large signal swings (see Figure 7) that the signal would encroach upon the nonlinear area and results in increased distortion. This is quite common in Class A audio output or driver stages.

In a great number of transistor applications, normal operating signal levels are small. Examples of such applications are the rf and most amplifier stages of radar, radio, and television receivers. Even after detection, as in audio or servo preamplifiers, the signal can be at a low-level.

In low-level stages, signal swings run from less than 1 μV to about 10 mV under normal operating conditions (for which these stages are generally designed). Therefore, it is important to analyze the transistor under conditions where the largest ac signal to be amplified is small compared with the dc bias current and voltage. The transistor is then said to be operating in the small-signal mode. Transistors used in this way are normally biased at currents between 0.1 and 10 mA and voltages between 2 and 10 V. An insufficient bias can cause distortion whereas an excessive bias exhibits unnecessarily increased power dissipation and a higher noise figure (the latter is primarily important in input stages). If the bias is sufficiently increased to make the stage operate in the high voltage nonlinear region, distortion will increase.

5.2.1.2 Switching. The most common usage for transistors is as a direct analog of the mechanical switch. That is, the transistor is made to have a high resistance (off-state) and a low resistance (on-state).

Figure 8 shows the on-state and off-state conditions. Point A occurs at a low current I_O and relatively high voltage implying a high resistance or an off condition. In a properly designed circuit, I_O will approach I_{CBO}. This current can be extremely small in a silicon planar transistor. Point B occurs at a relatively low voltage and high current point implying low resistance or an on condition.

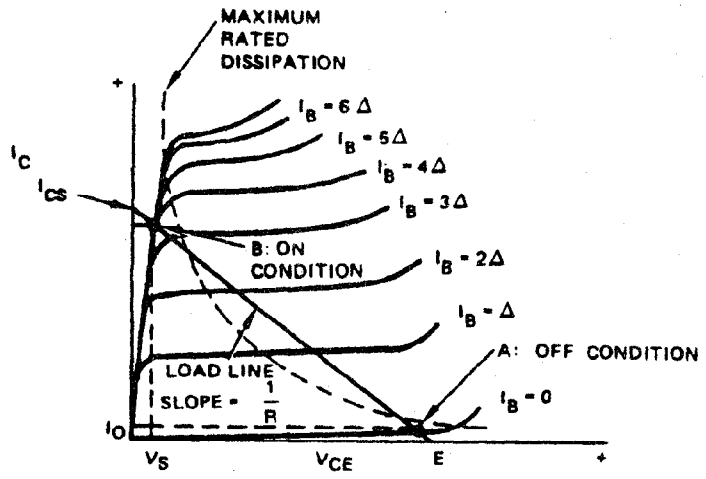


FIGURE 8. Generalized voltage-current characteristics for an npn transistors.

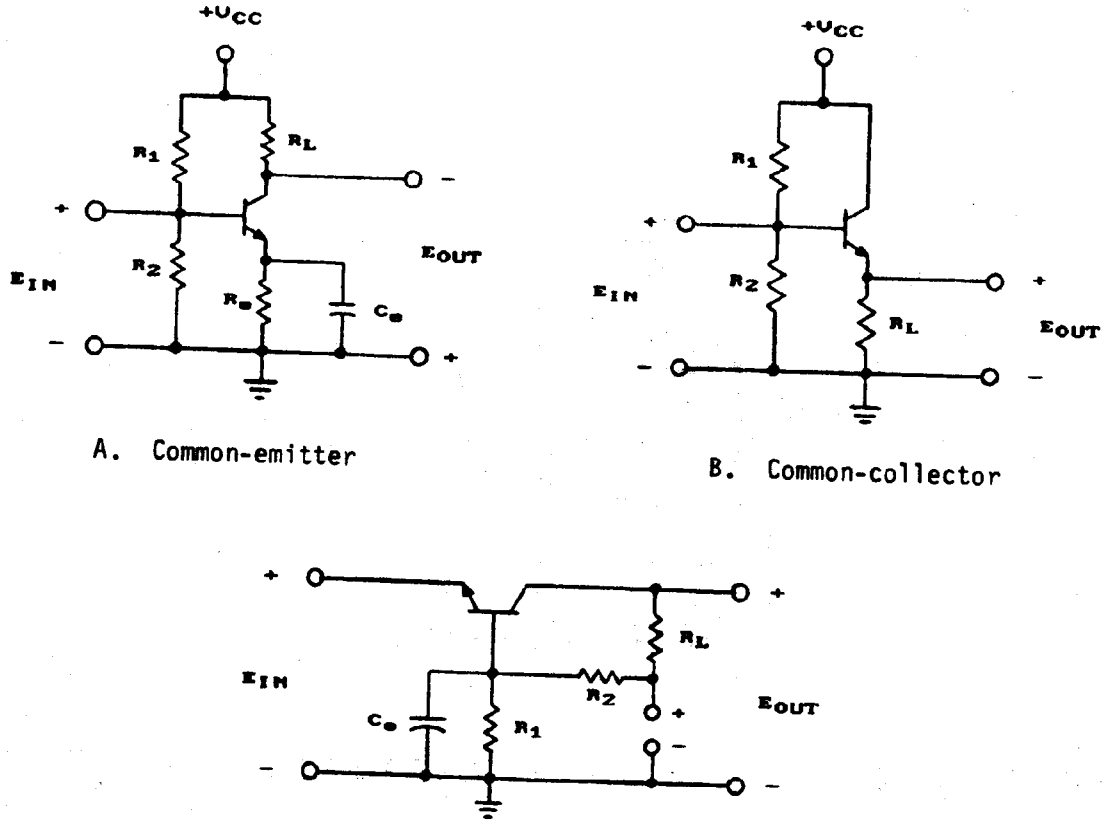
5.2.1.3 Choppers. A bipolar chopper transistor has similar characteristics (in the standard bias configuration) as a low power bipolar transistor; however, in the inverted configuration (where the collector is connected as the emitter and the emitter as the collector) the bipolar chopper transistor is designed to operate with a higher gain, a higher BV_{ECO}, and a lower Vec. These characteristics make the chopper transistor useful in both ac and dc chopper applications.

5.2.2 Usual applications. There are many applications in which the small-signal, switching, and chopper transistors can be used but only the basic applications are covered in this section.

5.2.2.1 Basic amplifiers. There are three basic types of biasing configurations. These are the common-emitter (CE), common-collector (CC) (also known as an emitter follower), and the common base (CB).

The basic common-emitter configuration, shown in Figure 9A, is the most versatile and is capable of giving both voltage and current gain. Its input and output impedance vary the least with variations in the input and output load impedance, and its input signal is inverted at the output.

5.2 TRANSISTORS, LOW-POWER



A. Common-emitter

B. Common-collector

C. Common-base

FIGURE 9. Basic biasing configurations.

The basic circuit of a common-collector is shown in Figure 9B. It has a current gain that approaches the CE configuration, but has a voltage gain less than one. Its main characteristics is its high input impedance. It has wide applications as a buffer stage between a high and low impedance source. There is no signal phase inversion when using this configuration.

The basic common-base circuit configuration is shown in Figure 9C. This configuration does not provide current gain but it provides voltage gain that approaches the CE configuration. Its low input impedance makes it useful for matching low impedance sources with high impedance sources. Its high cutoff frequency characteristic (f_{hfb}) also makes it useful for some high frequency applications. There is no signal phase inversion when using this configuration.

In the following paragraphs, a few of the basic amplifier circuits will be discussed using a simplified approach.

5.2.2.1.1 Single-stage amplifier. Figure 10 shows a typical single-stage amplifier using a 2N2222A npn transistor. With the resistance values shown, the bias conditions on the transistor are 1 mA collector current and 6 V from collector to emitter. At frequencies at which C_2 provides good bypassing, the input resistance is given by the formula

$$R_{in} = (1 + h_{fe})\, h_{ib}$$

For the 2N2222A, at a design center of 1 mA, the input resistance would be 51 times 20, or about 1000 Ω.

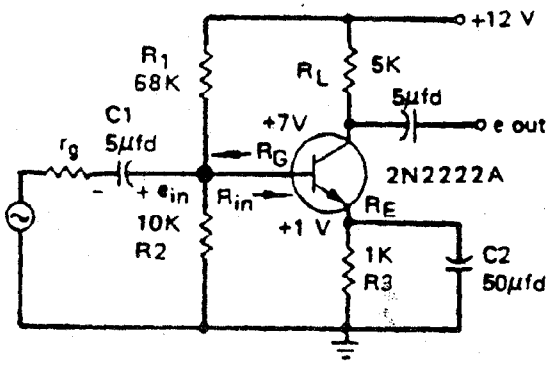


FIGURE 10. Single-stage amplifier.

The ac voltage gain e_{out}/e_{in} is approximately equal to R_L/h_{ib}. For the circuit shown, this would be 5000/20, or 250 (48 dB). The low frequency gain will drop 3 dB when the reactance of C_2 equals the parallel impedance of R_3 and R_E where

$$R_E \approx \frac{R_G + h_{ib}}{h_{fe} + 1}$$

and R_G is the parallel impedance of the bias network and generator r_g. Whereas the low frequency gain loss is mostly circuit dependent, the high frequency gain loss can be due to transistor characteristics.

5.2.2.1.2 Two-stage RC-coupled audio amplifier. The circuit of a two-stage RC-coupled amplifier is shown in Figure 11. The input impedance is the same as the single-stage amplifier and would be approximately 1000 Ω.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.2 TRANSISTORS, LOW-POWER

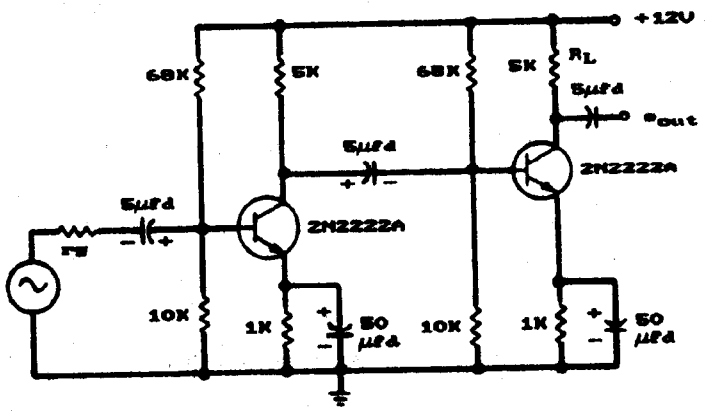


FIGURE 11. Two-stage RC-coupled audio amplifier.

The load resistance for the first stage is now the input impedance of the second stage in parallel with 5 KΩ. The approximate voltage gain for the two-stage circuit is given by the formula:

$$A_v \approx h_{fe} \frac{R_L}{h_{ib}}$$

5.2.2.1.3 Class B push-pull output stages. In the majority of applications, the output power is specified; therefore, a design will usually begin at this point. The circuit of a typical push-pull Class B output stage is shown in Figure 12.

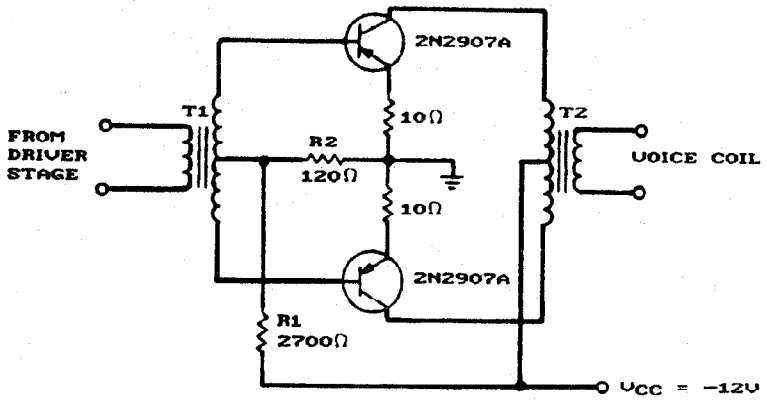


FIGURE 12. Class B push-pull output stage.

The voltage divider, R_1 and R_2 gives a slight forward bias (about 0.52 V) on the transistors to prevent crossover distortion. The 10-Ω resistors in the emitter leads stabilize the transistors to prevent thermal runaway. Typical collector characteristics with a load line are shown below in Figure 13.

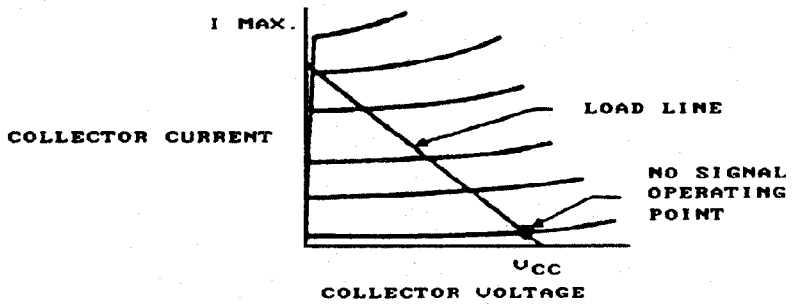


FIGURE 13. Typical collector characteristics and load line.

5.2 TRANSISTORS, LOW-POWER

The maximum ac output power, without clipping and using a push-pull stage, is given by:

$$P_o = \frac{I_{max}\ V_{CE}}{2}$$

where V_{CE} = collector to emitter voltage with no input signal.

Moreover, the load resistance R_L is equal to:

$$R_L = \frac{V_{CE}}{I_{max}}$$

And since the collector-to-collector impedance is four times the load resistance per collector, the output power (P_o) can be given by:

$$P_o = \frac{2V_{CE}^2}{R_{c-c}} \qquad (1A)$$

Where R_{c-c} = collector-to-collector load resistance.

Thus, for a specified output power and collector voltage, the collector-to-collector load resistance can be determined. For an output power in the order of 50 mW to 850 mW, the load impedance is so low that it is essentially a short circuit compared with the output impedance of the transistors. Thus, unlike small signal amplifiers, no attempt is made to match the output impedance of transistors in power output stages. The power gain is given by:

$$\text{Power gain} = \frac{P_{out}}{P_{in}} = \frac{I_o^2\ R_L}{I_{in}^2\ R_{in}}$$

Since I_o/I_{in} is equal to the current gain (β) for small load resistance, the power gain formula can be written as follows:

$$\text{Power gain} = \frac{\beta^2 R_{c-c}}{R_{b-b}} \qquad (1B)$$

where

R_{c-c} = collector-to-collector load resistance
R_{b-b} = base-to-base input resistance
β = grounded emitter current gain.

Since the load resistance is determined by the required maximum undistorted output power, the power gain can be written in terms of the maximum output power by combining equations (1A) and (1B) to give

$$\text{Power gain} = \frac{2\beta^2 V_{CE}^2}{R_{b-b}\ P_o}$$

5.2.2.1.4 Class A output stages. The Class A output stage is biased as shown on the collector characteristics in Figure 14.

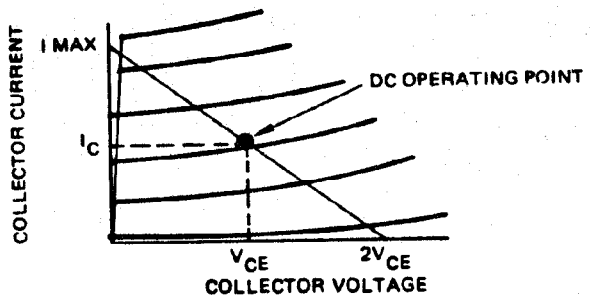


FIGURE 14. The dc operating point of class A audio amplifier.

The dc operating point is chosen so that the output signal can swing equally in the positive and negative direction. The maximum output power without clipping is equal to the following:

$$P_o = \frac{V_{CE}\ I_C}{2}$$

The load resistance is then given by:

$$R_L = \frac{V_{CE}}{I_C}$$

By combining these two equations, the load resistance can be expressed in terms of the collector voltage and power output is given by:

$$R_L = \frac{V_{CE}^2}{2\ P_o}$$

For output powers of 20 mW and above, the load resistance is very small compared with the transistor output impedance, and the current gain of the transistor is essentially the short circuit current gain β. Thus for a Class A output stage, the power gain is given by:

$$\text{Power gain} = \frac{\beta^2 R_L}{R_{in}} = \frac{\beta^2 V_{CE}^2}{2R_{in}P_o}$$

5.2 TRANSISTORS, LOW-POWER

5.2.2.2 Oscillators. An oscillator may be considered to be an amplifier that furnishes its own input signal from its output signal. Vital and fundamental to all oscillators is an amplifier. The process of oscillation simply involves a connection of output to input so that certain conditions are fulfilled. These conditions are called stability criteria, and different types of oscillators are classed primarily by the means employed to obey the basic stability criteria.

Figure 15 shows an amplifier with a voltage gain A whose output is connected to a second amplifier with a voltage gain B. The input of amplifier A is connected to the output of B.

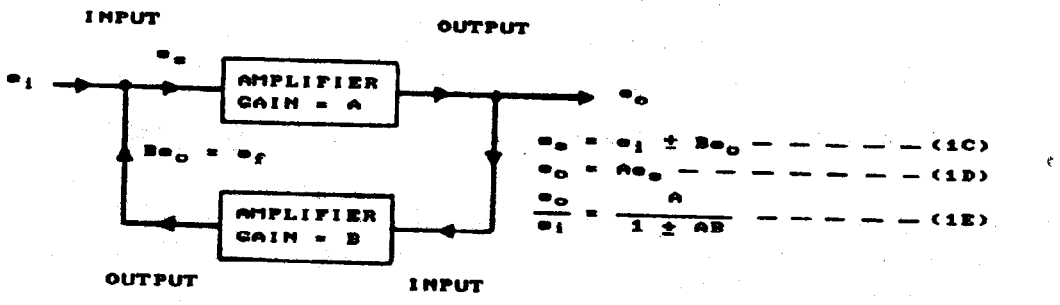


FIGURE 15. Stability criteria.

There is no need to know the amount of voltage, current, or power gain of either amplifier (whether it is greater or less than unity), or a need to know how the outputs are added. These equations are written in Figure 15. Equation (1E) is called a transfer function and describes how the output voltage/input voltage ratio behaves in terms of the gains of the amplifiers. In an oscillator, it might be expected that a small input signal would be amplified infinitely, or until some limit is imposed. This condition from Equation (1E) is possible if the product of AB = -1. This, in turn, would require one to choose the positive sign in Equation (1C). The block diagram shows that this positive sign stipulates that $e_f = Be_0$ "add" to e_i or is in phase with e_i. The ratio of e_0/e_i need not be infinite for oscillation to occur nor must the product of A and B equal one. The exact specification of the conditions for oscillation is most simply stated as $AB > 1$.

In some cases, the gain considered in writing the transfer function is most conveniently expressed as power or current gain. If power gain is considered, then the loop gain (net gain) must be greater than unity. This is called the Barkhausen criteria.

Four other common approaches are used for analytical determination of the stability of the system. These approaches (from direct examination of the system differential equation solutions) are: Routh's criterion, Nyquist's criterion, Bode's attenuation and the phase shift method, or a combination of these. In nearly all oscillators, the amplifier that we have labeled A has a voltage gain and current gain greater than unity (and consequently, a power gain greater than unity), and the amplifier that we have labeled B has gain less than unity. An active device, such as a transistor, furnishes the gain, and the resistors, capacitors, and inductors provide the loss and phase shift to insure the proper polarity and amplitude of the feedback.

The following paragraphs discuss the various types of oscillators and their applications.

5.2.2.2.1 Phase shift oscillators. Figure 16A depicts a simple versatile transistor amplifier. Its current gain is stabilized against transistor variation. The means used to stabilize both operating point and small signal current gain also allow it to be used over a collector voltage range of 2 to 24 V.

When this simple amplifier is combined with the phase shift network in Figure 16B, oscillation will occur at a frequency where there is a 360-degree total phase shift; 180 degrees of this 360 degrees is furnished by the grounded emitter amplifier, and 180 degrees is furnished by the high pass network. Figure 16C connects both together and provides a 5 KΩ pot for frequency adjustment. This pot adjusts frequency from about 200 to 400 Hz. Both h_{ie} and h_{ob} enter as terms in the expression for frequency, but the unusually low impedances chosen provide excellent temperature and voltage stability. Output is derived across the collector and is approximately equal to VCC. Frequency of oscillation is as follows:

$$f \approx \frac{1}{2\pi \sqrt{6R^2 C^2 + 4 R R_L C^2}}$$

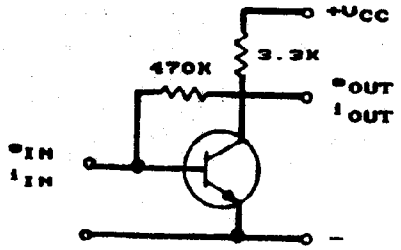
(This equation assumes $R > 10\ h_{ie}$ and $1/h_{oe} > 10\ R_L$)

The h_{fe} for sustained oscillation is as follows:

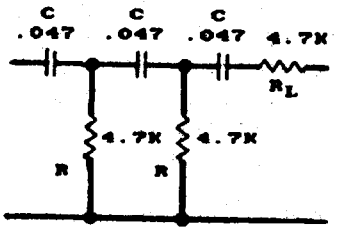
$$h_{fe} \approx 22 + \frac{30R}{R_L} + \frac{4R_L}{R_L}$$

5.2.2.2.2 Parallel-T oscillators. Figure 17 shows another RC phase shift oscillator using a parallel-T network. In this case, the simple amplifier is supplemented by an emitter follower to eliminate h_{ie} loading variations. Frequency stability of 0.2 percent is possible over the temperature range of -55 °C to +80 °C. In both phase shift oscillators, the effect of h_{oe} variations is swamped by the low (3.3 KΩ) collector load.

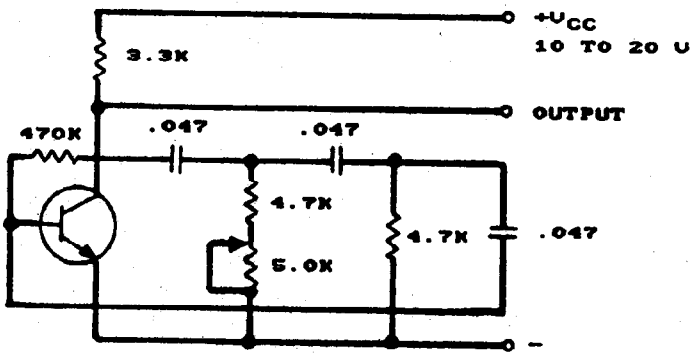
5.2 TRANSISTORS, LOW-POWER



A. Basic amplifier



B. Phase shift network



C. Phase shift oscillator with frequency adjust

FIGURE 16. Phase shift oscillator.

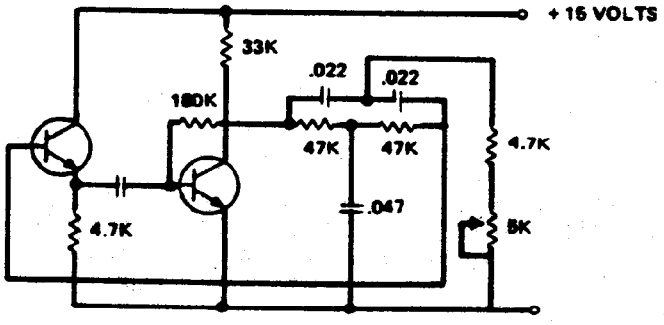


FIGURE 17. High stability parallel-T oscillator.

5.2.2.2.3 Resonant feedback oscillators. These are among the most prevalent and useful oscillators. For their resonators, they use either an inductance capacitance combination or a crystal electromechanical resonator. They are characterized by circuit simplicity, good power efficiency, and good stability. Figure 18A demonstrates how ac coupling between input and output of a transistor amplifier is accomplished. This circuit is classed as a tuned collector oscillator. L_1 and C, in the collector of transistor Q, form a parallel resonant circuit.

Furthermore, the mutual coupling between L_1 and L_2 (called a tickler winding) provides an input signal current whose direction and magnitude are set by the physical arrangement of L_1 relative to L_2, and by the direction and magnitude of the current through L_1 .

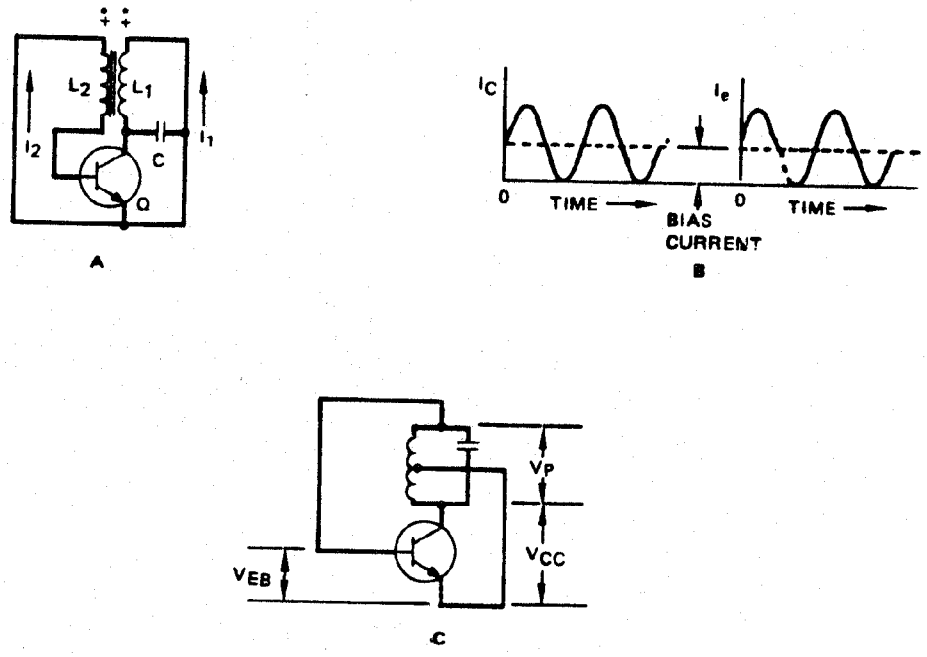


FIGURE 18. Resonant feedback oscillator.

The dots shown on the coils indicate winding start and phase coincidence so that an increase in collector current causes an increase in base current. This will provide the regenerative or positive feedback required to initiate oscillation. The core placed in the transformer is shown to indicate unity coupling between L_1 and L_2 for oscillator operation. The collector current and base current waveforms of Figure 18B show an offset which comes from a bias current which is not shown in this simple ac circuit. This offset current insures that

5.2 TRANSISTORS, LOW-POWER

the oscillator will start. Reversal of the increasing or decreasing base current is a limited cycle and comes about because the current gain of a transistor rolls off at both very high and very low collector currents. The exact points in I_C where reversal occurs are determined by loop gain and losses, and will always lie between transistor cutoff and saturation. The period of oscillation is very nearly set by the familiar relation between L_2 and C.

$$f_r = \frac{1}{\sqrt{2\pi\sqrt{L_2 C}}}$$

This assumes that the core coupling L_1 and L_2 is operated over a reasonably linear portion of its B-H characteristic. A more nearly exact expression, including the mutual inductance M, is

$$f_r = \frac{1}{2\pi\sqrt{C\,(L_2 + L_1 + 2M) - (L_2 L_1 - M)2\,h_{ob}/h_{ib}}}$$

In Figure 18C, an auto transformer can be substituted for the two winding transformers, and the emitter rather than the base may be allowed to float. This preserves the proper feedback polarity and is the basic circuit of the Hartley oscillator. Furthermore, it is a grounded-base oscillator and stability criteria are appropriately expressed in terms of grounded-base hybrid parameters. Analysis of this type of oscillator most frequently concerns itself with limit conditions. First, and fundamental, is the ability of the oscillator to start and continue oscillation. For this purpose small signal hybrid parameters may be used to establish the power gain of the circuit or the equivalent current or voltage gains. The Barkhausen criteria for loop unity gain forms the most convenient analytical approach. In terms of the mutual inductance M and the inductances $L_1 + L_2$, oscillation requires:

$$h_{fb} \approx \frac{L_2 + M}{L_1 + M}$$

In power applications, device ratings become important limit parameters to examine. The first point to consider is the power dissipation (P_d) of the device compared with the input power (P_i) needed in an attached load. For example, to furnish 1 W of output power (P_o) into a stipulated load, and by an additional winding or other means, provide a correct reflected load to the collector of the transistor, we can relate the P_d of the device to the desired P_o using the following equation:

$$n = \frac{P_o}{P_i} \times 100\% = \frac{P_o}{P_o + P_d} \times 100\% \quad \text{(neglecting circuit loss)}$$

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

where

R_L = required load V_{CC} = collector supply voltage
P_o = output power
P_i = input power
P_D = power dissipated in the transistor
n = efficiency as a percent

Because in most cases, oscillator efficiency will be greater than the theoretical Class A efficiency of 50 percent (relationship is in the biasing requirements needed to start and sustain oscillations); therefore, in our example, we can use 50 percent as the worst case efficiency for calculating the required P_d, $0.5 = 1/1 + P_d$ where $P_d = 1.5 - 0.5 = 1$ W.

The next concern is that of collector voltage swing. After selecting a supply voltage (V_{CC}), the maximum swing will be determined by the value of the load resistor (R_L).

where

$$R_L = \frac{V_{CC}^2}{2 P_o}$$

If $R_L = \infty$, then the worst case occurs, and the oscillator tank voltage swing is the largest. The peak voltage (V_p) appearing across the tank circuit is:

$$V_P \approx \frac{V_{CC}}{R_L} Q\omega L$$

$$Q = \frac{\omega L}{R_L} \quad \text{(assuming an unloaded Q of greater than 10)}$$

then

$$VP \approx \frac{V_{CC}\omega^2 L^2}{R_L^2} \quad \text{and the peak stress acros the transistor, then becomes}$$

$$V_{CE} \approx V_{CC} \frac{(1 + \omega^2 L^2)}{R_L^2}$$

The peak stress across the emitter base junction is related to V_{CE} through the transformer turns ratio. Therefore:

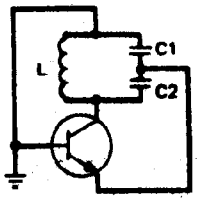
$$V_{EB} = \frac{V_{CC}\ \omega^2 L^2 N}{R_L^2} \quad \text{where N = turns ratio}$$

5.2 TRANSISTORS, LOW-POWER

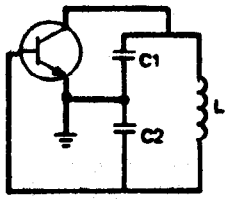
If the capacitor used to resonate the collector tuned circuit is split and used to form the feedback divider network, a Colpitts oscillator results. A simplified ac circuit for this configuration is shown in Figure 19A. Both C_1 and C_2 together set the effective capacity against which L resonates.

Figure 19B is identical to 19A except that the emitter rather than the base is shown at ground. One may relate nearly any LC resonant oscillator to the Hartley or Colpitts types, even though, at very high frequencies part of the circuit capacitive may be hidden as transistor capacitance.

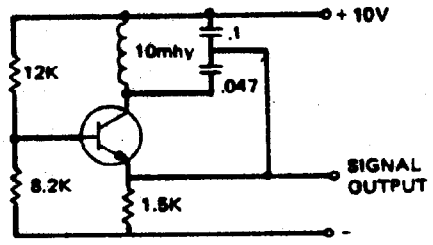
Figure 19C illustrates a practical 10 kHz Colpitts oscillator having a temperature drift rate of 0.035%/°C. This is the total drift rate and is determined by the temperature rate of incremental permeability of the coil core material.



A. Common-base Colpitts oscillator



B. Common Emitter Colpitts oscillator



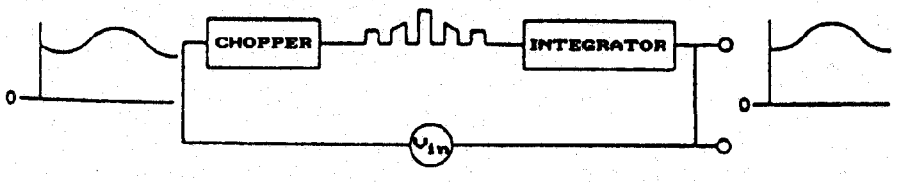
C. 10 kHz Colpitts oscillator

FIGURE 19. Colpitts oscillator.

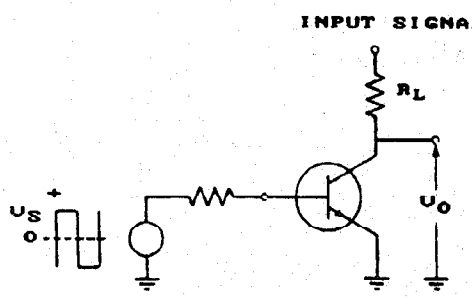
For the purpose of stability analysis, the Thevanin equivalent of the emitter resistor and the bias divider, together with transistor h_{ib} and the loaded voltage divider, are lumped to form the feedback loop.

If very high frequency stability is desired, the frequency determining network should be buffered from the amplifier. This provision cushions the frequency determining network from the inevitable changes induced from electrical environmental variation, but it also demands higher losses in coupling networks and therefore a higher amplifier gain to satisfy stability criteria. The alternative to this is a lower loss frequency determining network. The lower loss network is an alternative way of saying that a "high Q" is required where Q is defined as the energy stored per radian of angular period divided by the energy lost per radian of angular period. This Q definition is the usual figure of merit associated with resonant circuits but is equally applicable to many other networks having a transcendental solution.

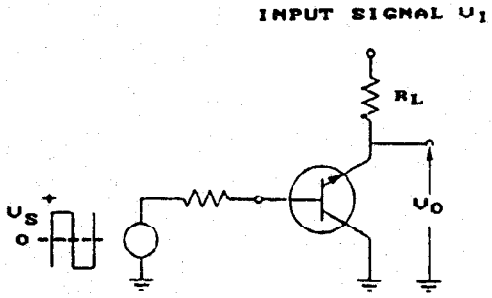
5.2.2.3 Chopper. A chopper circuit can be used to convert a slowly varying dc signal to a fixed frequency ac signal, but with the amplitude variations that match the amplitude variations of the original signal. The basic circuit is shown in Figure 20.



A. Block diagram showing action of a chopper circuit



B. Basic chopper circuit



C. Inverted chopper circuit

FIGURE 20. Basic chopper circuit.

5.2 TRANSISTORS, LOW-POWER

In this process a very small dc signal, after being converted to ac, can be amplified by a RC-coupled amplifier which eliminates the problem of thermal generated current (I_{CBO}) becoming mixed with the signal current. However, at these very small signal levels, a problem known as voltage offset causes an error in the output of a standard type bipolar transistor. This offset voltage can be greatly reduced by the use of a bipolar chopper transistor that is especially designed so that, when operated in the inverse mode (as shown in Figure 20C), the offset voltage, $V_{EC(ofs)}$, is in the microvolt range. Moreover, other inverse parameters such as I_{ECS}, $h_{FE(INV)}$, and $r_{ec(on)}$ are also specified so as to aid in the circuit design.

5.2.2.4 Switching applications. A proper choice of the I_C/I_B (force gain) ratio can minimize switching time for a given given transistor. Some circuit techniques are available that further improve switching speed. Two common techniques for this purpose, a collector-catcher circuit or the use of a speed-up capacitor in the base.

A simplified collector-catcher circuit is shown in Figure 21. Improvement in switching speed is obtained because the transistor is not allowed to go into saturation. Storage time, therefore, is drastically reduced. With no input pulse applied to the circuit, the transistor is initially biased off by $-V_{BB}$. A positive pulse of voltage turns the transistor on and the collector voltage begins to drop from $+V_{CC}$ toward $+V_R$. If V_R is greater than the sum of the voltage drop across CR and V_{CE} just out of saturation, then the collector is clamped at some value of voltage which maintains the transistor out of saturation.

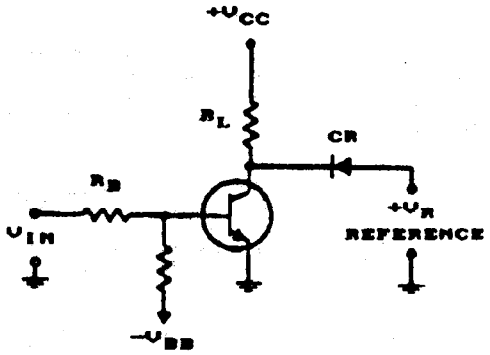


FIGURE 21. Collector-catcher circuit.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The difficulty with this circuit is that the maximum I_C is essentially beta-dependent and because beta varies with temperature, this circuit is very unstable. Burnout of the diode or transistor is possible.

More practical circuits are shown in Figures 22 and 23. In these circuits, the voltage V_B and the diode keep the transistor out of saturation. However, the feedback arrangement tends to keep I_C constant by automatically varying the base drive. For example, if beta increases, I_C tends to rise and the collector voltage begins to decrease. But if the base drive is decreased, the circuit is returned close to its original operating condition. The only disadvantage of this circuit is the requirements of an isolated power supply V_B.

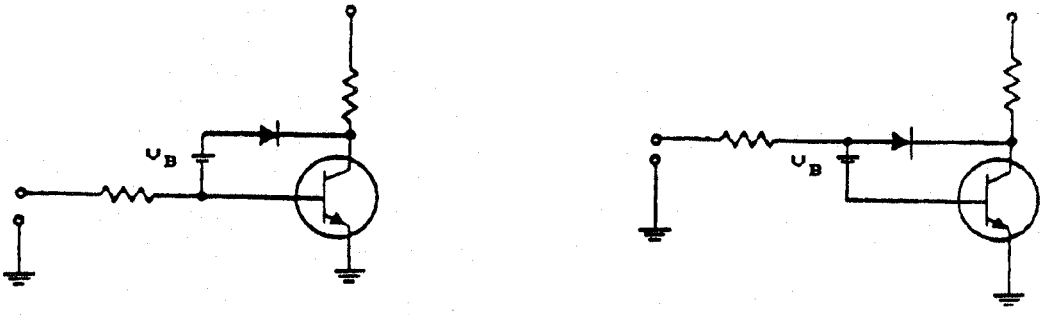


FIGURE 22. Collector-catcher circuits in which feedback is used.

A practical circuit (also known as a Baker clamp) is shown in Figure 23. Operation is similar to that described above. The drop across CR_2 acts as the V_B supply. The diode pair CR_1 constitutes the feedback arrangement. The function of CR_3 is to keep the transistor turned off under low V_{BE} voltage to reduce delay time. C acts as a speed-up capacitor. Use of the proper speed-up capacitor effectively increases turn-on drive and turn-off drive without supplying large amounts of on base current. As a result, faster rise times and faster fall times are achieved without the penalty of long storage time.

5.2 TRANSISTORS, LOW-POWER

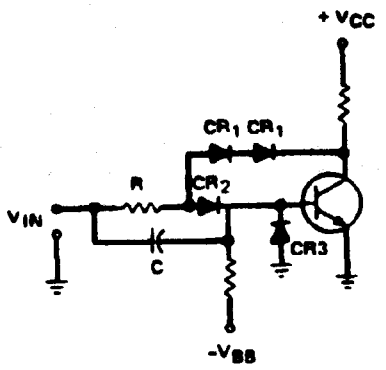


FIGURE 23. Practical collector-catcher circuit.

An example of a circuit that uses a speed-up capacitor is shown in Figure 23, and waveforms for this circuit are shown in Figure 25.

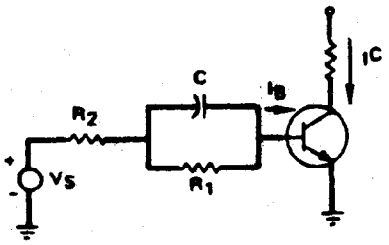
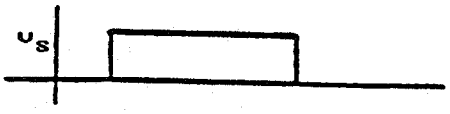
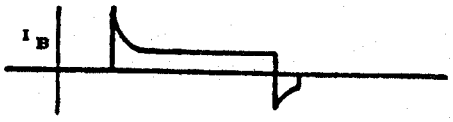


FIGURE 24. Circuit with speed-up capacitor.

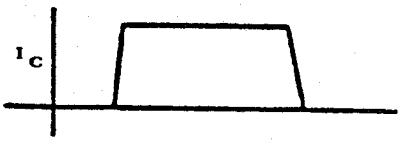
The optimum value of C for fastest response can be found experimentally. If V_S is large compared with V_{BE} and R_2 can be neglected, the charge stored on the capacitor while the transistor is on is V_C. This charge should equal the stored base charge for best response. Practical values for R_2, V_S, and V_{BE} will modify this relation.



A. Source voltage V_S



B. Base current I_B

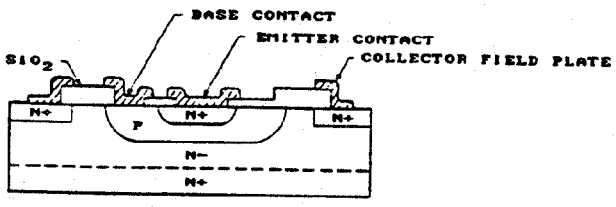


C. Collectors current I_C

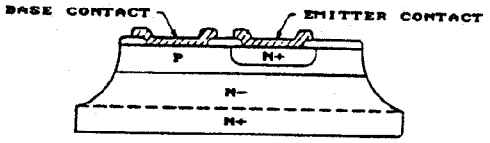
FIGURE 25. Waveforms for circuit shown in Figure 24.

5.2.3 Physical construction.

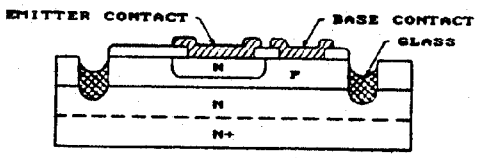
5.2.3.1 Die structures. The two basic structures of low-power and chopper transistors are the mesa (single and double) and planar types. These structures are shown in Figure 26.



A. Nonpassivated mesa



B. Glass passivated double mesa



C. Planar

FIGURE 26. Basic die structures.

5.2 TRANSISTORS, LOW-POWER

Planar devices offer the lower leakage current because both junctions are always protected by silicon dioxide or some other passivant. In a mesa structure, the base-collector junction might not always be protected by a primary passivant. If this is the case, some type of conformal coating may be used during the assembly operation as a secondary passivant.

There are three basic die processes that are used to produce the mesa and planar structures. These are the double diffused (epi-collector), the triple diffused, and the epi-base process. By using a combination of these processes and die structures, several different transistor die-types can be produced (with the exception of an epi-base planar). It is, therefore, possible that three JANTXV2N5667 devices, each from three different manufacturers, could contain three different die types. These three devices, although they may have passed all JANTXV2N5667 electrical parameters and other requirements, could each possess different characteristics. It is these differences that could lead to a problem known as the unspecified parameter; for example, while one manufacture's JANTXV device functions well in a circuit, the same type device from another manufacturer may not.

5.2.3.2 Package construction. Standard packages used for these devices are the TO-5, 18, 39, 46, 78, and 91. Figure 27 shows the basic construction for a TO-5.

Die bonds can be either soft solder or gold eutectic and both are reliable.

Wire bonding can be done with either gold or aluminum wire. The bonding techniques used can be thermocompression or ultrasonic weld to the die, and ultrasonic, thermocompression, or electrical weld to the leads. These bonds are reliable if proper process controls are implemented.

Sealing of the package is done in an inert atmosphere using electrical-weld techniques. It is possible that weld splashes might occur which form loose particles inside the package; therefore, MIL-STD-975 devices require that a PIND test and an X-ray be done to detect any defective devices.

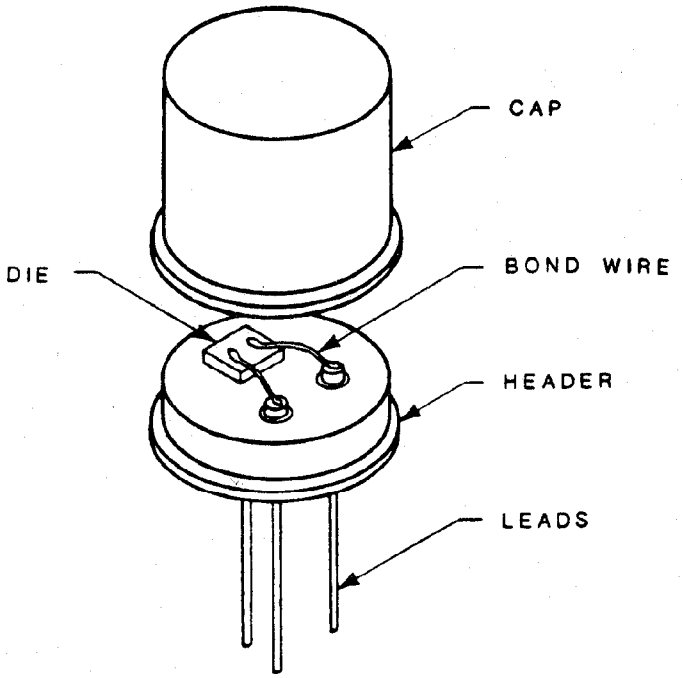


FIGURE 27. Typical TO-5 transistor construction.

5.2.4 Military designation. The military designation for transistors is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-15900 for details.

The component designation is 2N for transistors.

5.2 TRANSISTORS, LOW-POWER

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices, suffix M, longer or shorter terminal leads (suffix L or S), or any other letter to indicate a modified version.

5.2.5 Electrical characteristics. The electrical parameters of interest may be separated into static and transient groups. This is somewhat arbitrary because the same parameters may be in both groups.

5.2.5.1 Static parameters.

5.2.5.1.1 Collector cutoff current, I_{CBO}. I_{CBO} is defined as the dc collector current when the collector-base junction is reverse biased and the emitter is open. Its value is determined by the voltage applied as shown in Figure 28 and the temperature at which it is measured as is indicated in Figure 29. I_{CBO} essentially varies exponentially with temperature and above the "knee" of the voltage curve, and it tends to follow an exponential variation with voltage.

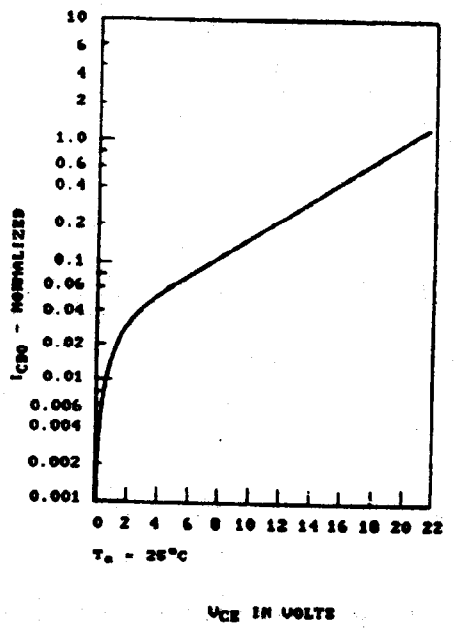


FIGURE 28. Behavior of I_{CBO} with voltage.

To eliminate the need to take voltage variation into account each time a circuit is designed, the I_{CBO} is usually specified at a voltage near the maximum rating of the transistor. It is then assumed that I_{CBO} is constant for voltages lower than this value. The temperature which determines the I_{CBO} value of a device is the junction temperature, not the ambient. If the basic measuring circuit is studied, it is seen that the power dissipated in the transistor is the product of I_{CBO} and V. Since the I_{CBO} and the power are very small, the junction temperature is essentially that of the ambient temperature. Many manufacturers label the leakage current versus temperature curves with the ambient rather than junction temperature.

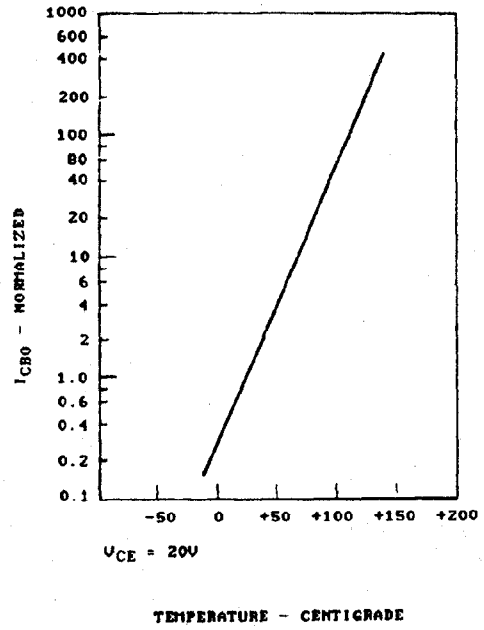


FIGURE 29. Behavior of I_{CBO} with temperature.

5.2.5.1.2 Collector cutoff current, I_{CEO}, I_{CER}. I_{CEO} is defined as the dc collector current when the collector-emitter junction is reverse biased and the base is open-circuited. I_{CEO} is important in the design of transistor circuits because it determines how close to a true off-state condition can be obtained. For example, the circuit in Figure 30 shows the base lead open rather than the emitter. Leakage current I_B (essentially I_{CBO}) flows across the reverse biased collector-base junction as before, but this current cannot return to the voltage source unless it flows across the base-emitter junction. Because polarities are such that this junction tends to be forward biased and this leakage current is essentially indistinguishable from a base-current supplied externally, the transistor will amplify the current to produce an additional current, (h_{fe} x I_{CBO}) in the collector. The net result is that there is a total collector current (I_{CEO}) that is equal to (1 + h_{fe}) x I_{CBO}.

5.2 TRANSISTORS, LOW-POWER

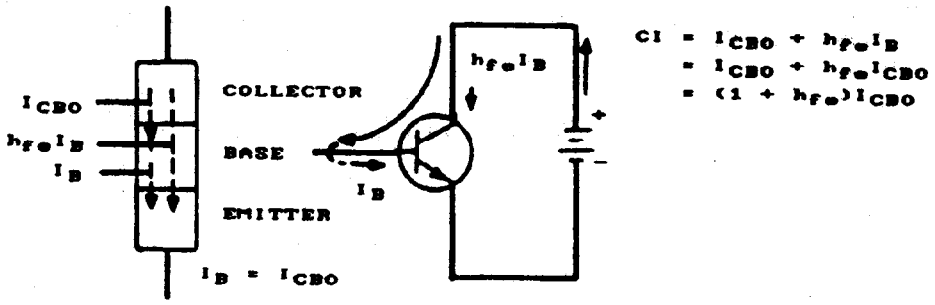


FIGURE 30. Effect of transistor gain on leakage current.

If a finite resistance is placed between the base and the emitter (R_{BE}) as shown in Figure 31, some of the I_{CBO} current will be shunted through this resistor. This shunted portion of the leakage current would not be amplified and therefore a collector current (I_{CER}) will flow. The I_{CER} would be much less than I_{CEO}, and as R_{BE} approached a zero value, I_{CER} would approach I_{CBO}. Conversely, if R_{BE} approaches an infinite value (usually 1 KΩ or more) I_{CER} would approach I_{CEO}.

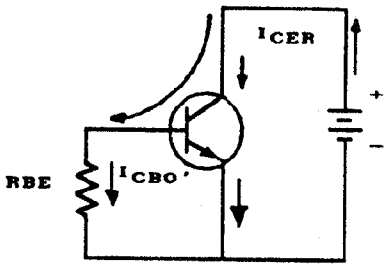
5.2.5.1.3 Emitter cutoff current (I_{EBO}). If the base-to-emitter of a transistor is reverse biased, there will be a leakage current (I_{EBO}) similar in every way to I_{CBO} except that it flows from emitter to base. Thus, to reverse bias a transistor, it is necessary to allow for I_{CBO} and I_{EBO} to flow out of the base lead. When I_{EBO} is not specified, it is usually assumed to be equal to I_{CBO}.

5.2.5.1.4 Current gain (h_{FE}). The direct current gain is of great interest in the design of transistor circuits and is defined as the following:

$$h_{FE} = \frac{I_C}{I_B}$$

where I_B and I_C are the absolute value of the base current and collector current respectively. The more commonly used parameter, h_{fe}, is essentially the ratio of a change in collector current for a small change in base current.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com



$I_{CER} = h_{FE}(I_{CBO} - I_{CBO'})$

FIGURE 31. Effect of base-emitter resistor on leakage.

The value of h_{FE} is usually measured at a voltage between collector and emitter; this value is close to the saturation voltage as this represents a minimum value. Normally, h_{FE} is not a very strong function of collector-emitter voltage outside of saturation. It is, however, a rather strong function of junction temperature and of collector current. Figure 32 is a set of typical curve for h_{FE} as a function of I_C. Each curve is associated with a different temperature.

The most important feature in Figure 32 is that over most of the current range, the gain decreases as the temperature decreases. This rule cannot be applied indiscriminately because the reverse begins to be true beyond 10 mA and 100°C. A second feature is that the gain has a definite maximum which may be quite broad at room temperature or rather sharp at 125°C. The collector current at which this maximum occurs is a function of the junction temperature. It follows that when selecting an operating point, the temperature range over which the circuit will be expected to operate should be considered.

It sometimes happens that a decreasing gain with increasing temperature is desirable. Magnetic cores, for example, often require less drive at high temperatures than at low. Generally, however, this characteristic cannot be controlled sufficiently well to be useful.

5.2 TRANSISTORS, LOW-POWER

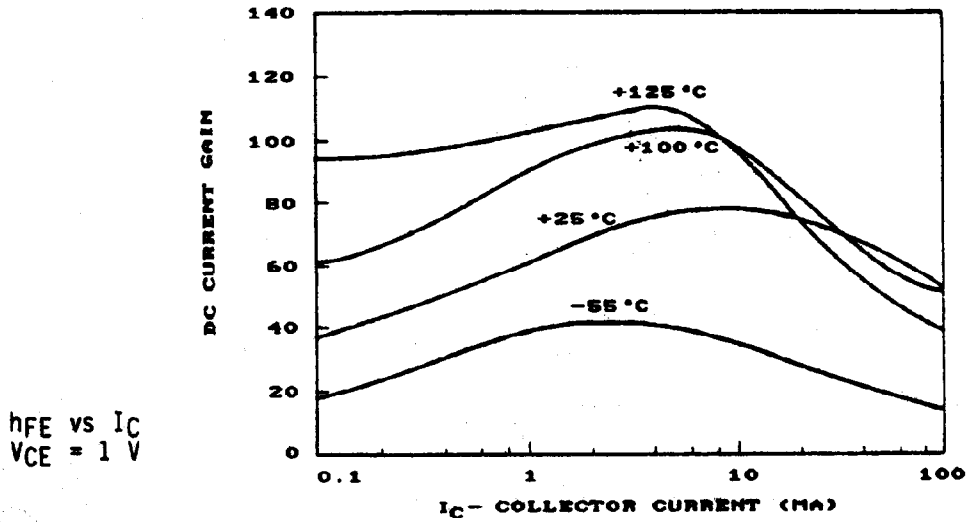


FIGURE 32. Variation of h_{FE} with temperature and current.

5.2.5.1.5 Collector saturation voltage [$V_{CE(sat)}$]. The collector saturation voltage is the parameter that effectively limits how closely the transistor approximates a closed switch. Figure 33 and 34 shows how this parameter varies with temperature, current ratio and collector current.

Figure 33 shows that when the temperature is held constant while the circuit current ratio is increased (or I_B decreased), the saturation voltage changes linearly.

In Figure 34 the curves indicate that temperature is not a particularly strong influence and that the saturation voltage increases with increasing temperature; however, this depends very much on the device being used. In some devices the saturation voltage is almost completely independent of temperature, while in others the temperature coefficient can be negative overall or part of the temperature range.

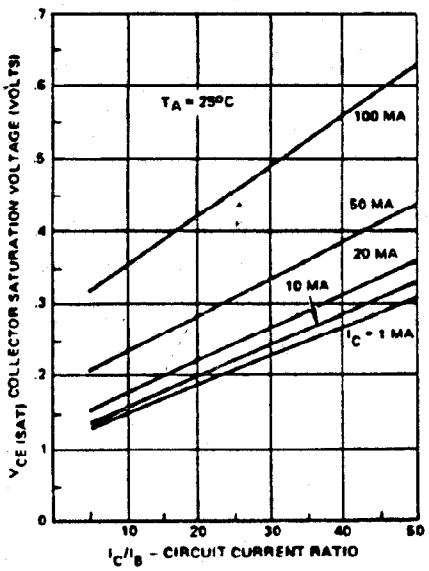


FIGURE 33. Variations in $V_{CE(sat)}$ with force gain and I_C.

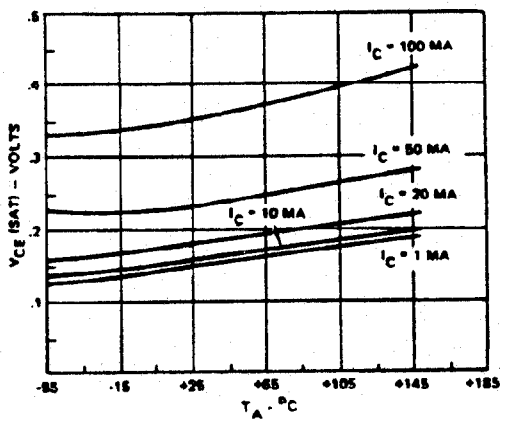


FIGURE 34. Variations in $V_{CE(sat)}$ with temperature and I_C.

5.2 TRANSISTORS, LOW-POWER

5.2.5.1.6 Base-emitter saturation voltage, $V_{BE(sat)}$. Two sets of $V_{BE(sat)}$ curves similar to the curves for collector saturation are shown in Figures 35 and 36. The characteristic feature is that the slopes of these curves are opposite to those shown in Figures 33 and 34. The temperature coefficient is negative and varies little over the entire range. Nominal values are about 2 mV change per degree centigrade.

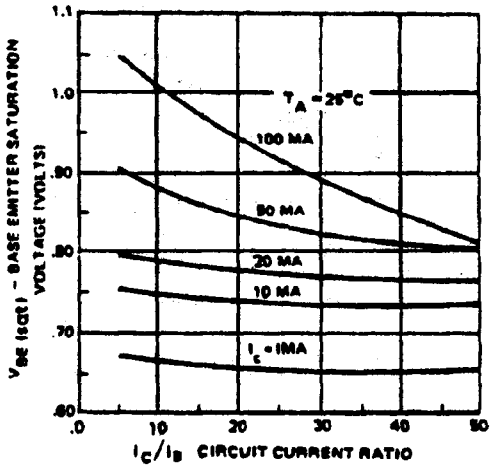


FIGURE 35. Variations in $V_{BE(sat)}$ with force gain and I_C.

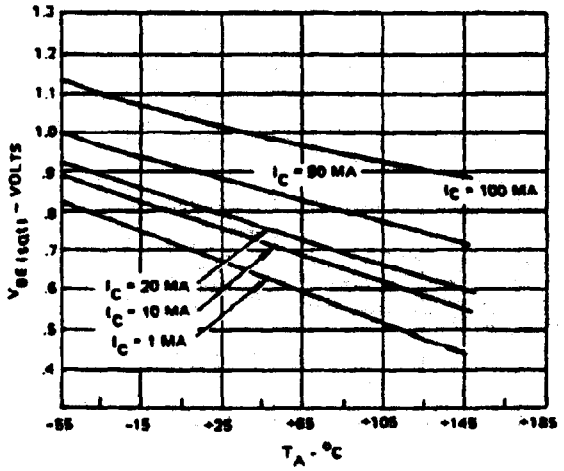


FIGURE 36. Variations in $V_{BE(sat)}$ with temperature and I_C.

5.2.5.2 Transient parameters of switching transistors. The factors which influence the transient response of the transistor are basically associated with the diffusion time of the carriers across the base region, the effect of capacitances due to the collector-base and base-emitter junctions, the associated parasitic capacitances between leads and from case to leads, and the operating conditions of the circuit. In predicting transient response, it is convenient to think of the turn-on delay time (t_d), the current rise time (t_r), the storage or turn-off delay time (t_s), and current fall time (t_f) as dependent variables whose values depends upon the operating conditions of the device as well as the capacitances and diffusion parameters. It follows that the calculations of the time intervals (t_d, t_r, t_s, and t_f) can be very complex.

Before considering the time intervals of a transistor and the effects of the collector-base and base-emitter capacitances and resistances on these intervals, the presence of these reactive components within the transistor is discussed.

There are two types of capacitances associated with any semiconductor junction: transition capacitance (C_T) and diffusion capacitance (C_D).

C_T is due to the high electric field in the depletion region caused by the voltage across the barrier. Hence C_T is voltage dependent. C_D is due to the current flowing through the depletion region. Hence C_D is current-dependent and therefore, the total junction capacitances is the sum of C_T and C_D.

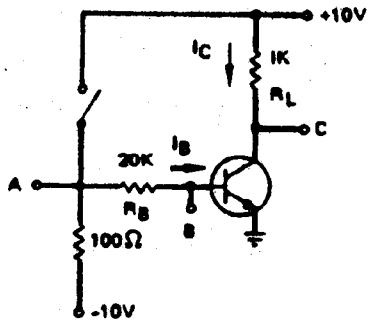
The collector capacitance (C_c) is made up primarily of the transition capacitance, (C_{TC} or C_{obo}), because the diffusion capacitance is small in a reverse biased junction. On the other hand, the emitter-base capacitance can be either diffusion capacitance (C_{DE}) or transition capacitance (C_{ibo}) if reversed biased.

The active portion of the base region of a transistor is not equipotential but exhibits an ohmic resistance to the flow of base current. This parasitic resistance is called the base-spreading resistance (r_b).

This base resistance is not purely resistive but takes on a distributed form (transmission line) in many transistor structures. However, to avoid complications, it will be assumed that r_b is resistive. As will be seen later, r_b is a very objectionable parameter because it contributes to the deterioration of transistor performance in many ways.

The time intervals (t_d, t_r, t_s, and t_f) mentioned above are defined in Figure 37. These definitions are commonly accepted for measurements and require no further explanation beyond that the 10 and 90 percent points of the collector waveform are taken as the points at which measurements are to be made. The collector waveform is the voltage from collector to emitter. For most calculations, we shall use collector current rather than collector voltage as the reference thereby avoiding some difficulty with what is meant by rise-time (t_r). In Figure 37D, the voltage is falling during t_r, because the current is increasing (or rising) during this interval.

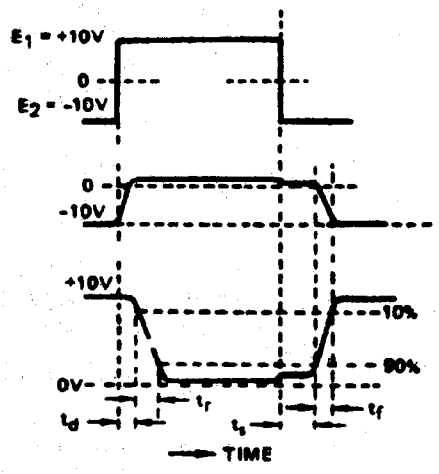
5.2 TRANSISTORS, LOW-POWER



$I_{B1} = I_{B2} = 0.5$ MA
$I_C = 10$ MA
$I_C/I_{B1} < h_{FE}$

$T_{ON} = t_d + t_r$
$T_{OFF} = t_s + t_f$

A. Typical circuit



B. WAVEFORM GENERATED AT NODE A BY SWITCH

C. WAVEFORM AT NODE B SHOWING FORWARD BIAS ON BASE DURING SATURATION RESPONSE TIMES

D. COLLECTOR WAVEFORM SHOWING STANDARD DEFINITIONS OF RESPONSE TIMES

B. Waveforms

FIGURE 37. Transient response.

5.2.5.2.1 Charge-control theory. To better comprehend switching parameters, it is necessary to have an understanding of the transistor as a charge-controlled device. A thorough understanding will also permit accurate comparisons among the switching abilities of different transistors as well as providing data to calculate switching speeds in a variety of circuits. The charge required to perform various switching functions can be readily deduced by examining the common hybrid-equivalent circuit of Figure 38, which is a simple but fairly accurate representation of a transistor.

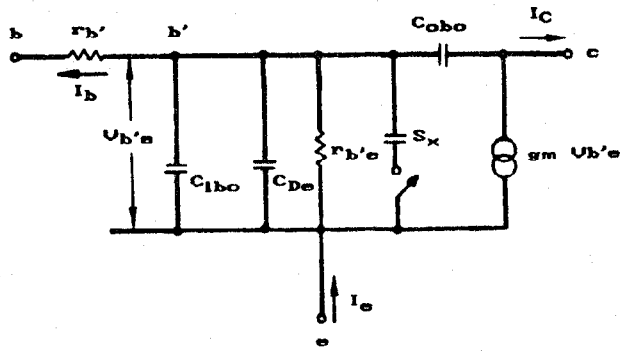


FIGURE 38. Hybrid-π transistor equivalent circuit.

The capacitances C_{ibo} and C_{obo} are the transition capacitances of the emitter-base and collector-base junctions respectively plus stray capacitance associated with the transistor package. Transition capacitance is somewhat analogous to a parallel plate capacitor in that charges of opposite polarity are separated by a narrow region depleted of charge. The emitter and collector transition capacitances are always present, and their capacitances increase somewhat as the reverse junction voltage decreases.

The diffusion capacitance C_{De} represents the charge involved in the flow of carriers across the base region, this is somewhat analogous to water flow in a pipe. The carrier flow is by diffusion; hence the name "diffusion capacitance." Because a transistor has a finite base width, the flow of carriers through it constitutes a base charge. The charge in the base will, therefore, be proportional to the emitter current, (I_e). Because a barely perceptible change in base-emitter voltage can change the emitter current of a transistor over wide ranges, a capacitor can be used to represent the base charge if it has a value proportional to emitter current. Therefore, the capacitance C_{De} has a value proportional to emitter current; C_{De} also depends somewhat upon collector voltage since the base width narrows as collector voltage increases.

5.2 TRANSISTORS, LOW-POWER

The remaining capacitance, S_X, and the switch represents the storage phenomenon. The switch closes and connects S_X to the circuit only when the transistor is saturated. The behavior of the store charge is nonlinear, being dependent upon values of the currents I_{b1} and I_C and upon temperature.

5.2.5.2.2 Delay time, t_d. The switching process, as the transistor is turned on from an off condition and then off from an on position, can be explained in terms of the capacitance. When the transistor is off, a reverse-bias voltage is present on both junctions, and only the capacitances C_{ibo} and C_{obo} are effective. The voltage on the base must be changed to forward-bias the base-emitter junction so that the transistor will turn on. The time required for the input base current to charge C_{ibo} and C_{obo} to a voltage level that forces the emitter to inject current is the delay time, t_d. The total change in charge on C_{ibo} and C_{obo} during the delay time interval is termed the off-bias charge, Q_{OB}.

5.2.5.2.3 Rise time, T_r. The rise-time interval commences as the emitter begins to inject current, causing the diffusion capacitance to enter into the picture. The base voltage must now change from its value at the threshold of injection to its final value. This voltage change requires change in charge upon C_{ibo} and C_{De}. In addition, the voltage across C_{obo} changes as the flow of collector current causes the collector voltage to drop from the cutoff level to the saturation voltage level. Thus, a charge is required by C_{obo}.

The time taken for the input current to supply the charge required by C_{ibo}, C_{De}, and C_{obo} is the rise-time interval. The total charge required by these capacitances to turn on a transistor is termed the active region charge, Q_A.

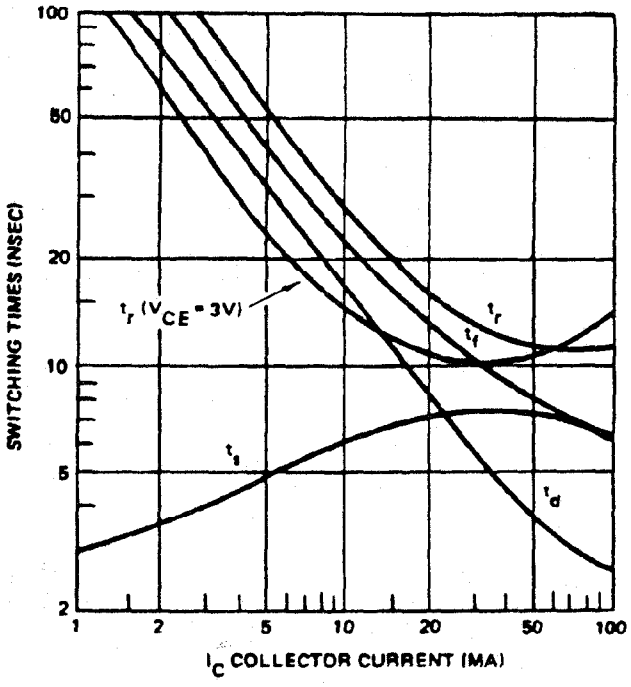
5.2.5.2.4 Storage time, t_s. To keep the transistor in the saturation region, a base current (I_{B1}) larger than that required to just maintain I_C is supplied. This condition results in an excess charge Q_X to be stored in the transistor. The storage excess charge is represented on the transistor equivalent circuit of Figure 38 by the inclusion of S_X and the switch.

Before the transistor can be turned off, the excess charge stored on S_X must be removed. It can be removed either through an external path or by internal recombination. Since the internal recombination time is long, faster switching is achieved by supplying a reverse-bias voltage to the transistor input circuit. The reverse bias permits a reverse current (I_{b2}) to flow and considerably shortens the time to deplete S_X of charge. The time for the excess charge to leave S_X is the storage time (t_s).

5.2.5.2.5 Fall time, t_f. At the conclusion of storage time, the fall-time interval commences. The charge which must be removed is the same as the rise time; that is, C_{ibo}, C_{De}, and C_{obo} must be charged from the voltages of the on condition to the voltages at the threshold of conduction. Often, however, some excess carriers stored in remote regions of the transistor did not have time to leave during the storage time interval and exit during fall time, producing a "tail" on the fall-time waveform.

The total charge which appears in the base circuit during turn-off is called the total control charge (Q_T). It represents the charge stored on all transistor capacitances and the store during the period of time the transistor was held on. It is normally the largest amount of charge involved in the switching process and, therefore, must be known as a function of all circuit variables (i.e., I_C, I_B, V_{CC}, and T_J) in order to design a circuit for any set of conditions.

5.2.5.2.6 Collector current effects on transient response. As illustrated in Figure 39, the diffusion capacitance charge becomes appreciable for this transistor at currents above 10 to 20 mA, depending upon collector voltage. At low currents, the rise time when V_{CC} = 3 V is faster than the rise time when V_{CC} = 10 V because the charge required by C_{ob} is less. However, at high currents, rise time at 10 V is faster than rise time at 3 V. This behavior occurs because at high currents, rise time is determined primarily by the diffusion capacitance, whose average value is smaller when switching from high collector voltages.



B_F = 10
V_{CE} = 10 V

FIGURE 39. Typical switching times.

5.2 TRANSISTORS, LOW-POWER

Fall time is less than rise time at all current levels because of the effect of recombination. Simply stated, the recombination current is subtracted from the turn-on current I_{b1} and added to the turn-off current I_{b2}. Therefore, the current available to charge Q_A during rise time is less than the current available to discharge Q_A during fall time.

Storage time behavior is very complex and is the most difficult transistor parameter to predict. Because of the space limitation in this manual, storage time will not be discussed. The t_s curve of Figure 39 is merely intended to serve as a guide.

5.2.5.3 Transistor frequency limitations.

5.2.5.3.1 Gain-bandwidth product (f_t). When operated at low frequencies in the common emitter configuration, the transistor exhibits a short-circuit gain ($R_L << r_{OUT}$) of h_{fe0}. This value may vary in small-signal transistors from a low of twenty to a high of several hundred. As the signal frequency is increased, the magnitude of h_{fe} decreases and its phase shift increases. This is because carriers when they are injected into the base-emitter junction will take a certain time to cross into the collector region. Figure 40 shows that there are three time constants limiting the speed of the injected carriers: the emitter time-constant $r_e C_e$, the collector time-constant $r_e C_{TC}$ (the collector is shorted to the emitter), and the base transit-time τ_B.

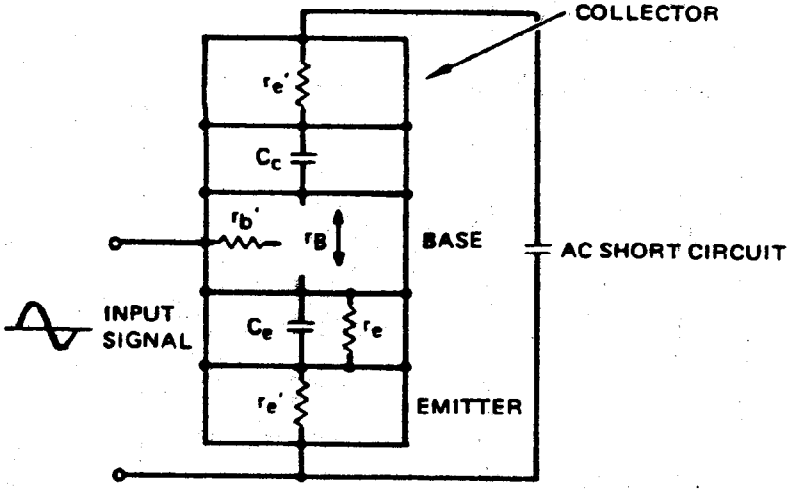


FIGURE 40. Various time-constants limiting the gain-bandwidth product of the transistor.

The sum of these three constants is defined as the gain-band-width product f_t, because it is the frequency at which h_{fe} falls to unity; it is given in the following equation:

$$f_t = \frac{1}{2\pi[\tau_B + r_e (C_e + C_c)]}$$

Even if C_e ($C_b'_e$) and C_c (C_{TC}) were made very small, the base transit-time would still limit the frequency response of the transistor. Therefore, high gain-bandwidth (f_t) transistors must be designed with extremely thin base regions or an accelerating field added into the base region. A well designed high-frequency transistor might have

$r_e C_b'_e$ of $25 \times 1 \times 10^{-12}$ = 0.025 ns

$r_e C_{TC}$ of $25 \times 0.4 \times 10^{-12}$ = 0.010 ns

τ_B of 125×10^{-12} = 0.125 ns

Total time for carriers to reach collector = 0.160 ns and the equation is as follows:

$$\text{Hence } f_t = \frac{1}{2\pi TC_{total}} = \frac{1}{628 \times 160 \times 10^{-12}} \approx 1.0 \text{ GHz}$$

A large collector bulk resistance will add a fourth time-constant of $r_c'C_{TC}$, which, if $r_c' = 100\ \Omega$ and $C_{TC} = 0.5$ pF will give the carriers another delay of $100 \times 0.5 \times 10^{-12} = 0.05$ ns, reducing f_t to 760 MHz. Therefore, a good high frequency transistor should also exhibit low extrinsic collector series resistance (r_c'). In epitaxial transistors, r_c' is small, and this fourth time-constant can be made negligibly small.

Figure 41 plots h_{fe}, h_{fb}, and maximum available power gain (MAG) versus frequency for a typical UHF transistor.

The first deduction that can be made is that there is an exact relationship between f_t and h_{feo} ($h_{feo} \times fh_{fe} = f_t$). Second, at a frequency five times this beta cutoff frequency (fh_{fe}) a dB/octave slope has been reached. Along this -6 dB/octave slope, the product of h_{fe} and its corresponding frequency is a constant, which is defined as the gain-bandwidth product (f_t).

5.2 TRANSISTORS, LOW-POWER

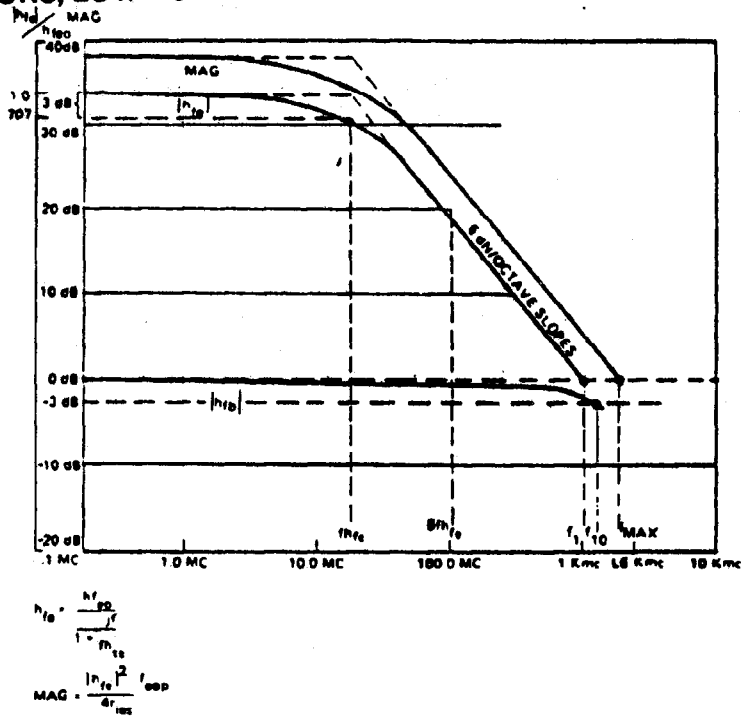


FIGURE 41. MAG, $|h_{fe}|$, $|h_{fb}|$ vs frequency of typical UHF transistor.

5.2.5.3.2 Alpha and beta cutoff frequencies. As previously stated, the beta-cutoff frequency (fh_{fe}) can be used as an aid to locate the proximity of the 6 dB/octave slope. Actually, modern transistor circuit analysis has done away with the formerly much used "alpha cutoff frequency" (fh_{fb}) and "beta cutoff frequency" (fh_{fe}). A small-signal transistor is primarily frequency-limited by its emitter, base, and collector time-constants, and may not be usable at fh_{fb} (due to feedback, there may not even be an fh_{fb}). Therefore, these two frequencies will simply be defined below.

a. Alpha cutoff frequency is the frequency at which the common base current gain α, falls to 0.707 of its low frequency value. In modern transistors fh_{fb} is usually somewhat above f_t.

b. Beta cutoff frequency is the frequency at which the common emitter current gain β, more recently identified as h_{fe}, falls to 0.707 of its low frequency value (h_{feo}).

5.2.5.3.3 Maximum frequency of oscillation. Maximum frequency of oscillation (f_{max}) is the frequency at which the maximum available unilateralized power gain (MAG) falls to unity. Considering the equation for MAG in Figure 41, it can be seen that as $|h_{fe}|$ drops to unity there still is a power gain given by $r_{oep}/4r_{ies}$ (impedance ratio to output to input impedance). Hence f_{max} will be generally higher than f_t. The f_{max} is sometimes also referred to as the $(\text{power gain})^{1/2}$ (bandwidth) product; stated as $\sqrt{PG}$ x BW which gives a 6 dB PG/octave slope, but should not be confused with f_t.

5.2.6 Environmental considerations. Refer to paragraph 5.1.6.3 Environmental considerations.

5.2.7 Reliability considerations.

5.2.7.1 Failure mechanisms. The failure mechanisms for low power transistors are discussed in detail with the stresses that cause these failures, under paragraph 5.1.6 Reliability considerations.

5.2.7.2 Derating. Refer to MIL-STD-975 for device derating guidelines.

5.3 TRANSISTORS, HIGH POWER

5.3 High-power.

5.3.1 Introduction. High-power transistors are used for controlling large amounts of power in various circuits by means of relatively low-power control signals. This category covers transistors that have a power dissipation of 2 W or more.

For high-power transistors, the design engineer will make frequent use of the common-emitter collector family of curves illustrated in Figure 42. This plot of the collector current versus collector-emitter voltage at various constant base current levels is divided into three regions; the limitations concerning these regions are also shown. The following describes the operation regions of all npn transistors (the polarities are reversed for a pnp transistor).

a. The active region is the region where V_{BE} is positive but is less than V_{CE}. Moreover, the collector current is a function of base current and consists of two areas: that of continuous operation and that of pulsed operation. These two operating areas are limited only by the maximum current and maximum voltage ratings of the device. In the pulsed operation area, pulse duration is also a limitation.

b. The saturation region is characterized by a positive V_{BE} greater than V_{CE} and is limited by the maximum base and collector currents only.

c. The cut off region, in which V_{BE} is negative, is limited only by the manufacturer's maximum voltage rating.

It is possible for a transistor to operate in any or all of these regions, depending on its mode of operation.

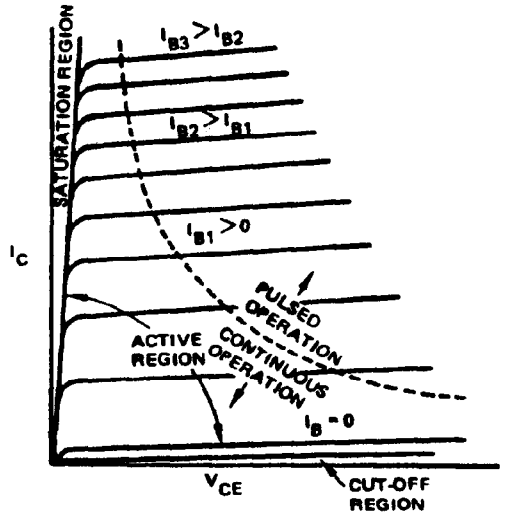
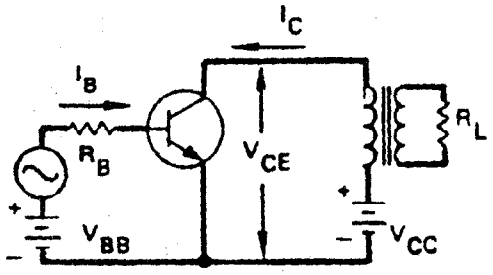


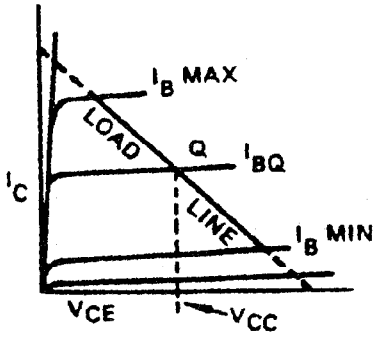
FIGURE 42. Collector family of curves for common emitter connection.

5.3.1.1 Operating modes. High-power transistors are devices for controlling large amounts of power in various circuits by means of relatively low-power signals. The operating mode of the transistor will depend upon the type of control desired. For example, the device may be used as a linear amplifier where the output signal should be a faithful reproduction of a lower power input signal. On the other hand, the device may be used as a highly efficient switch that simply opens and closes a circuit. The various modes of operation will be considered separately and related to the collector family of curves.

5.3.1.1.1 Class A. The Class A mode of operation is one in which collector current flows at all times and the load line lies within the active region. For example, the audio amplifier of Figure 43 operates in the Class A mode. The transistor is provided with a source of dc bias, I_{BQ}, by voltage V_{BB} and resistor R_B. The quiescent operating point Q is determined by the intersection of the dc load line (resistance of the transformer primary winding) and the collector curve corresponding to the quiescent base bias I_{BQ}. If an ac signal of base current is superimposed upon the quiescent bias current, such that I_B varies between the limits of I_{Bmax} and I_{Bmin}, the instantaneous operating point will traverse the ac load line as shown. Although for simplicity a resistive load is shown, in practice reactive loads are frequently encountered in which the load line is an ellipse rather than a straight line.



A. Class A amplifier



B. Load line

C. Collector current

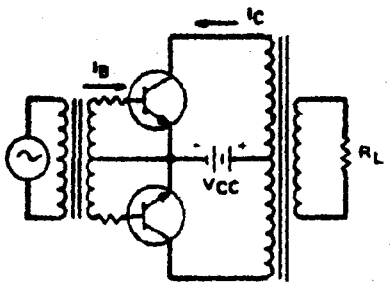
FIGURE 43. Example of class A mode.

5.3 TRANSISTORS, HIGH POWER

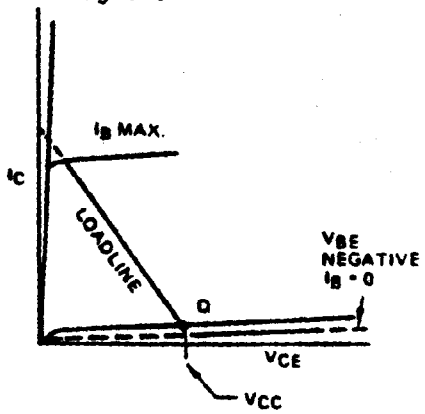
The Class A mode is not limited to audio amplifiers but includes all applications in which the collector current flows continuously.

5.3.1.1.2 Class B. The Class B mode is one in which the quiescent base bias is approximately equal to zero, and the corresponding collector current is also equal to zero. When the ac signal is applied, collector current flows for approximately one half of each cycle during the period when V_{BE} is positive.

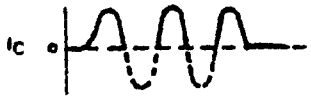
During the alternate half-cycles V_{BE} is negative, and the collector current is cut off. One common application of the Class B mode is in push-pull audio amplifiers as shown in Figure 44. Note that part of the load line is in the active region, and the remainder is in the cutoff region.



A. Class B amplifier



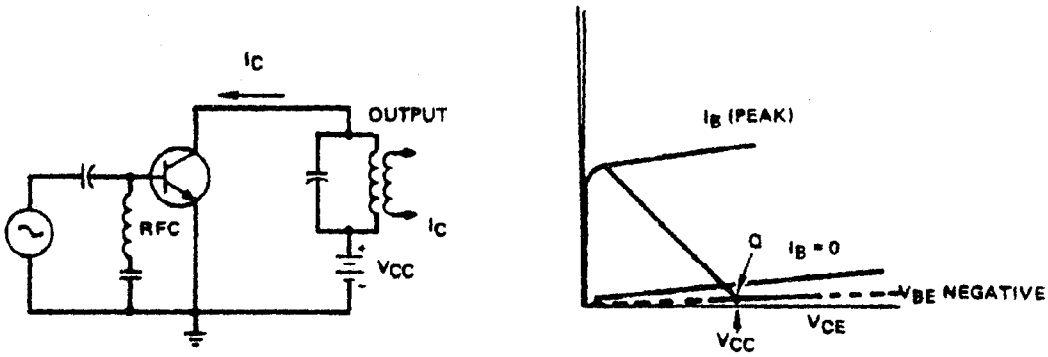
B. Load line



C. Collector current

FIGURE 44. Example of Class B mode.

5.3.1.1.3 Class C. The Class C mode is one in which the quiescent base-emitter bias is negative, and the corresponding collector current is completely cut off. When an ac base drive signal is applied, collector current flows for considerably less than one half of each cycle. A common application for the Class C mode is in rf power amplifiers as shown in Figure 45.



A. Class C amplifier

B. Load line

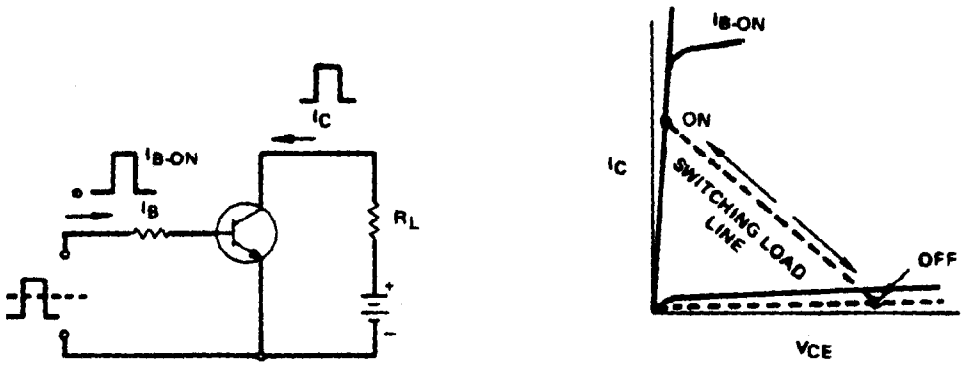
C. Collector current

FIGURE 45. Example of Class C mode.

5.3.1.1.4 Class D. The Class D mode is the term used to designate the operation of the transistor as a switch. For this type of operation it is necessary to consider three states: (1) The off state when the transistor is not conducting, (2) the on state when it is conducting, and (3) the "transition" state when it is changing from off to on or vice versa. In the off state, the base-emitter is reverse-biased (V_{BE} negative) so that the operating point is in the cut-off region. In the on state, the base current is raised to a level somewhat in excess of the value needed to maintain the collector current at the on value. This condition is sometimes described as overdriving the base. The operating point is in the saturation region. The value of V_{BE} under this condition is greater than that of V_{CE}; in other words, both the emitter-base junction and the collector-base junction are forward-biased and both the off and on states are low-power dissipation conditions for the transistor because in the former case the current is very low, and in the latter the voltage is low.

5.3 TRANSISTORS, HIGH POWER

In the Class D mode, the transition from off to on or vice versa is made very quickly by applying a positive or negative step of drive to the base. Therefore, the operating point does not remain in the transition state, which is in the active region, longer than a few microseconds. A simple switching circuit is shown in Figure 47.



A. Class D amplifier

B. Load line

FIGURE 46. Example of Class D mode.

5.3.2 Usual applications.

5.3.2.1 Transistor inverters. Transistor inverters have become an important means for converting dc to ac power in a wide variety of applications. They power complicated electronic systems of military, commercial, and space applications and are widely used as airborne power sources. In such applications, the transistor inverter offers advantages of high efficiency, absence of moving parts, low weight, small size, and high reliability.

5.3 TRANSISTORS, HIGH POWER

A typical inverter circuit is shown in Figure 47 and is comprised of two power transistors, a transformer with a square-loop core, and a source of dc voltage. The circuit operates as follows: assume transistor Q_1 is conducting and transistor Q_2 is not conducting. The supply voltage, V_{CC}, is connected to primary winding 2, which causes an opposing emf to be induced in winding 2 and in all other windings on the core. Feedback windings 1 and 4 are connected in the proper phase so as to bias bases of Q_1 and Q_2 in the off state. This condition will persist until the square-loop core material saturates. Magnetizing current in winding 2 then suddenly increases until the base drive of Q_1 is no longer sufficient to keep the transistor in saturation. Voltage across Q_1 increases and voltage across winding 2 decreases. The resulting induced voltages in windings 1 and 4 turn Q_1 off and Q_2 on. The supply voltage now is connected across winding 3, and the process repeats until the core saturates in the other direction, and another reversal occurs, and so on. The resulting output voltage is a square wave with the frequency

$$f = \frac{V}{4NAB_m} \times 10^8$$

where

V = applied voltage (volts)
N = number of turns in winding 2 or 3
A = core cross-sectional area (cm^2)
B_m = saturation flux density (gauss)

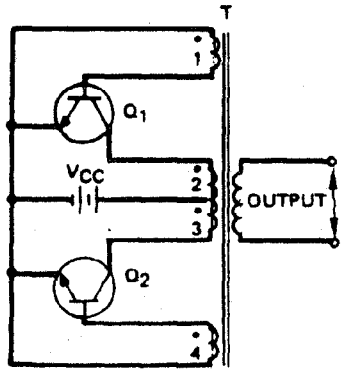
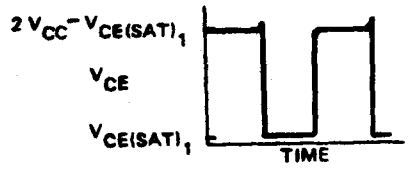


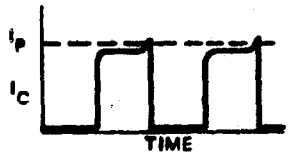
FIGURE 47. Basic inverter circuit.

5.3 TRANSISTORS, HIGH POWER

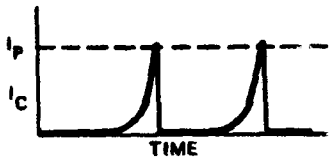
Typical voltage and current waveforms for one transformer inverter operation are shown in Figure 48. It can be seen from the collector-to-emitter voltage waveforms that each device is subjected in the off condition to a voltage approximately twice the supply voltage plus any induced voltage that may occur in the circuit due to such things as leakage inductance. Also significant is the fact that the same maximum collector current, i_p, is required for switching action whether this current is primarily reflected load current, as in Figure 48B, or totally magnetization current, as in Figure 48C. This will obviously limit efficiency at low output loads.



A. Collector-emitter voltage of transistor Q_1.



B. Full load collector current of transistor Q_1.



C. No load collector current of transistor Q_1.

FIGURE 48. Waveforms of basic inverter circuit.

Although the common-emitter connection is most useful, a common collector is sometimes used (when chassis ground polarity permits) because the transistors may be mounted without insulation on a chassis, giving better thermal contact.

At high-frequency and high-output power, the transformer requirements for saturated operation and efficiency present a difficult problem in the basic single-transformer inverter.

5.3.2.2 Saturable reactor controlled inverters. For power levels greater than approximately 25 W, several improvements can be made in the basic circuit of Figure 47. These improvements will provide a more economical circuit, improved transistor operation, and greater circuit flexibility. The modified circuit shown in Figure 49 is a conventional push-pull transformer circuit with feedback derived from additional windings on the transformer as in the previous circuit. A parallel feedback circuit with an additional transformer winding (N_C), voltage divider (P), and saturable reactor (SR) has been added. This modified circuit will oscillate and provide a square wave output voltage at

a frequency determined by the saturable core characteristics and the voltage applied to the winding on the core. The power output transformer is not driven into saturation; therefore, conventional lower-cost core materials may be used. Transistor operation is improved because the peak current, which the transistor must switch, is determined primarily by the load and not by the magnetizing current of a saturated output transformer.

Because the saturable reactor is required to pass only very short pulses of power, the saturation current is much smaller than that for the saturated power transformer of the basic inverter. With only one function to perform, the saturable reactor design can be improved to provide a faster transition from high to low impedance and, thereby, reduce the transistor switching time. The reduced current and shorter switching times significantly lower the transistor dissipation particularly at higher frequencies.

Greater design and circuit flexibility is achieved by permitting the operating frequency of the circuit to be controlled rather than fixed by the supply voltage as in the circuit of Figure 47. With the voltage divider (P), the voltage across the reactor can be varied to produce a variable output frequency.

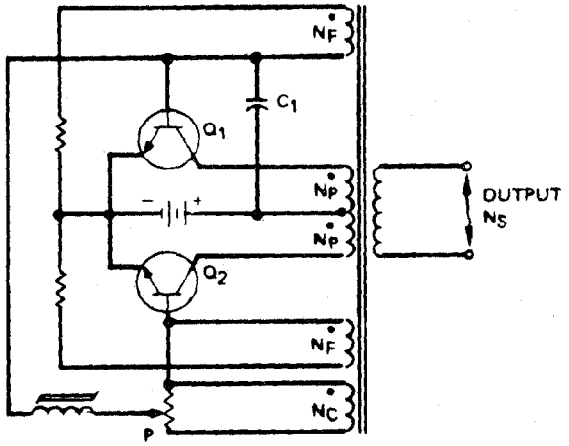


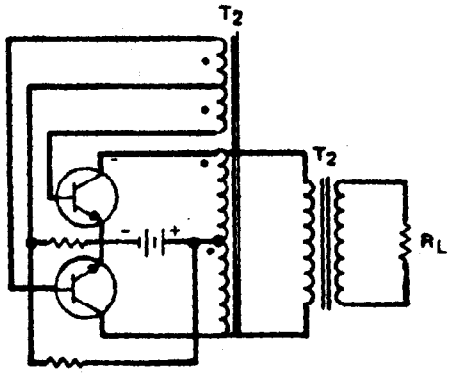
FIGURE 49. Saturable reactor-controlled inverter.

The design of this type of circuit is relatively simple. The output transformer must be selected to support the highest input voltage at the lowest desired frequency without saturating. This precaution will insure that the oscillation is always under the control of the frequency locking reactor circuit. The voltage supplied by the resistive divider to the saturable reactor should be at least equal to twice the voltage supplied by the feedback windings. The circuit will operate with lower voltages applied to the reactor. However, more reliable operation is achieved at the higher voltage. In calculating the turns required on the reactor, the total voltage appearing across the output terminals of the reactor as well as the voltage supplied by the resistive divider network must be considered.

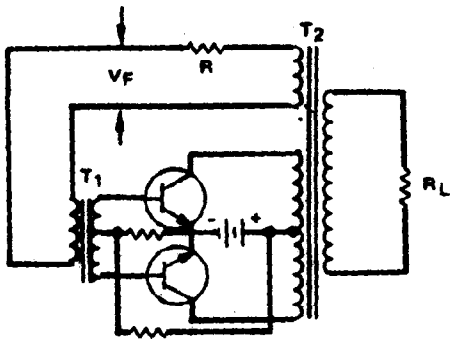
5.3 TRANSISTORS, HIGH POWER

Tests of the improved circuit indicate that the transistor switching time can be reduced to less than one-half that obtained with the basic circuit and that the dissipation occurring during the switching transient can be reduced to about one-fourth the value obtained with a saturated output transformer. The described circuit has been constructed in power ratings up to 500 W. Variable frequency converters have been built to operate over frequency ranges of 10 Hz to 1 KHz and up to frequencies of 20 kHz. These inverters are used to supply power to motors, electronic equipment, and a variety of other loads.

5.3.2.3 Two-transformer inverters. Other methods of obtaining high output power are the two-transformer inverters, shown in Figure 50. In these circuits, only the small driver transformers (T_1) saturate. This significantly reduces the magnetizing currents which the transistors must switch in the basic single-transformer inverter. The use of normal core material in the nonsaturating output transformer reduces transformer cost and increases efficiency.



A. Two Transformer inverter



B. Two transformer inverter with frequency control

FIGURE 50. Two-transformer inverter.

Frequency control may be accomplished as shown in Figure 50B where voltage (V_F), is regulated to provide constant frequency or varied to provide variable frequency. The circuits of Figure 51 are recommended to decrease transistor switching time and thereby reduce collector dissipation.

Fast switching is especially important at higher frequencies. Figure 51A shows capacitors in parallel with the base drive resistor of each transistor. This tends to promote faster switching by producing a spike of current limited only by the base-emitter junction impedance during turn-on. This capacitor also provides a low impedance path for the reverse junction current during turn-off. The capacitor discharges through the base drive resistor, which it shunts after the base-emitter junction is reverse biased. The capacitance must be greatly increased to provide any measurable improvement.

The circuit in Figure 51B uses cross-coupled capacitors to aid turn-on and turn-off. When Q_1 is on and Q_2 off, capacitor C_2 is charged to approximately 2 V_{CC} through the conducting base-emitter of Q_1. When Q_2 starts to turn-on, C_2 can discharge through Q_2 and apply reverse bias to the emitter-base junction of Q_1 to turn that transistor off. C_1 is aiding the turn-on of Q_2 as it charges to the off voltage level of Q_1.

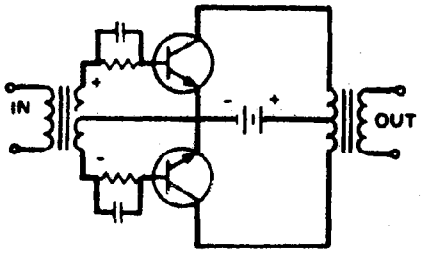
5.3.2.4 Audio amplifiers. The most important requirement of an audio power amplifier is to provide power gain over a wide band of frequencies with minimum distortion. This requirement is met by operating the transistor in the active region. The choice of circuit and operating mode will depend on the application.

In the Class A mode, the transistor is biased to some quiescent operating point with no signal applied. The ac signal then swings the operating point on either side of the quiescent point so that ideally the output collector current is a precise amplified reproduction of the input base current (see Figure 43 of this section). Because the quiescent operating point must allow the collector current to swing both positively and negatively, it is usually near the midpoint of the load line. Consequently, for the transformer-coupled Class A amplifier, the quiescent power dissipation in the transistor is equal to the input power as shown below.

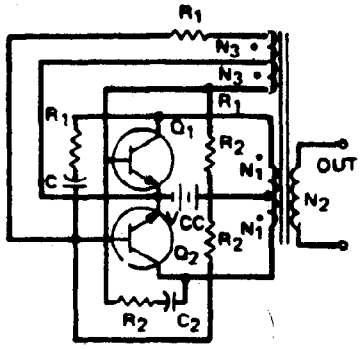
$$P_{IN} = V_{CC} \times I_Q$$

where V_{CC} is the supply voltage and I_Q is the quiescent current. Since the operating point swings both positively and negatively with the signal, the average input power remains constant in the Class A amplifier. The maximum transistor dissipation occurs at the zero signal or quiescent condition, and the maximum ideal efficiency, which occurs at maximum signal, is 50 percent. Because of the relatively high quiescent power dissipation and low efficiency, Class A amplifiers are usually limited to low-power levels.

5.3 TRANSISTORS, HIGH POWER



A. Inverter with speed-up capacitors



B. Inverter with cross-coupled feedback-capacitor

FIGURE 51. Inverter speed-up circuits.

In the Class B mode, the quiescent point bias is zero, so that the zero signal power dissipation in the transistor is zero. The ac signal swings the operating point alternately into the active region and into the cutoff region. Since conduction occurs for only 180 degrees of the ac cycle, it is necessary to operate the transistors in pairs in a push-pull circuit as shown in Figure 52. The two transistors are driven from a split-phase source (a centertapped transformer), and they conduct alternately.

In Class B operation, the maximum ideal efficiency is 78 percent. The maximum power output obtainable for Class B is five times the dissipation rating of the individual transistors. In contrast, the power obtainable in the Class A mode is only one-half the dissipation rating.

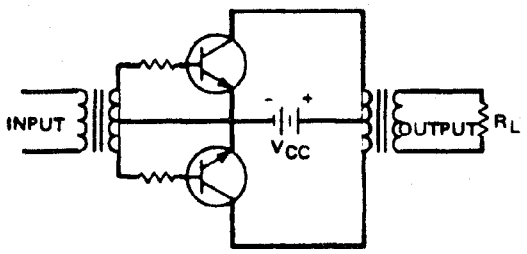


FIGURE 52. Class B amplifier.

The Class AB mode is often used for applications in which low distortion is a prime consideration. This mode is intermediate between Class A and Class B. It is necessary to overcome the so-called crossover distortion which results from the nonlinearity of the transistor input characteristics at low-current levels. Figure 53 shows how the Class B circuit of Figure 52 may be modified to accomplish this purpose. A resistor network, comprised of R_1 and R_2, is used to bias both transistors slightly into the active region for the quiescent condition. Thus, the conduction angle for each transistor is made slightly greater than 180 degrees, and there is a slight amount of overlapping near zero of the signal ac wave when both transistors are conducting. The quiescent point is chosen so that the base current is beyond the nonlinear knee region of the input characteristic. The transistor dissipation is still minimum at zero signal although it is no longer zero as with Class B. Also, the maximum efficiency is reduced somewhat from the Class B theoretical value. The Class C and Class D modes are not capable of linear operation.

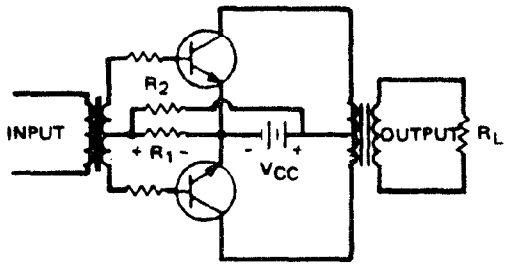


FIGURE 53. Class AB amplifier.

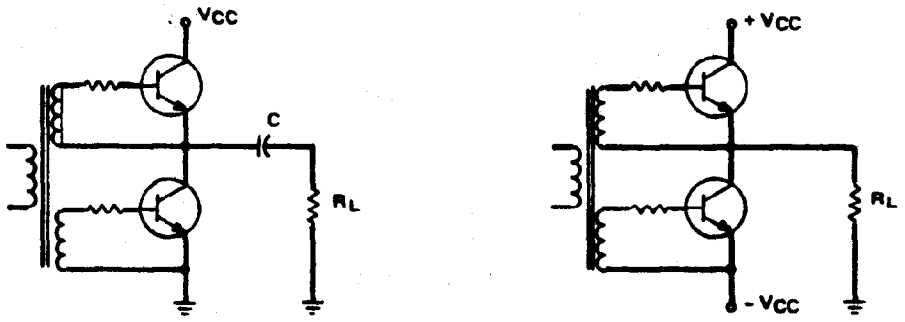
The following discussion is limited to the push-pull Class B (or AB) amplifier, which is the most useful type for high-power applications.

5.3 TRANSISTORS, HIGH POWER

A number of circuit configurations are available to the amplifier designer. The most familiar is the transformer-coupled circuit illustrated in Figures 52 and 53. This circuit is commonly used in applications where high-power levels and narrow-to-moderate bandwidths are required such as public address systems, sonar, and servo amplifiers. The output transformer allows great flexibility in matching the amplifier to a wide range of load impedances.

The phase shifts in transformers at the high and the low frequency ends of the amplifier bandwidth frequently lead to instability when the transformer is included in an inverse feedback loop. These problems have led to the development of a number of "transformerless" power amplifier circuits. One of these is the series transistor circuit shown in Figure 54. Two methods are shown for connecting the load. In Figure 54A, a coupling capacitor is used. In Figure 54B, the load is directly connected, which necessitates a dual power supply to set the input end of the load at ground potential. One disadvantage in this circuit is that it still requires 180-degree phase inversion for the base drive. This may be provided by a transformer with its attendant phase shift problems as shown in Figure 54, or by a split-load phase inverter as shown in Figure 55. The latter solution also may present problems such as difficulty of impedance matching a need for very large coupling capacitors.

These difficulties may be eliminated by the complementary circuit as shown in Figure 56. This circuit avoids the need for a transformer or a split-phase input. Advantage is taken of the availability of pnp as well as npn devices. The amplifier operates as a true push-pull amplifier because one transistor is driven on while the second is driven off and vice versa. The complementary circuit also allows the design of a direct-coupled amplifier.



A. Capacitor coupled load

B. Dual power supply

FIGURE 54. Series output amplifier.

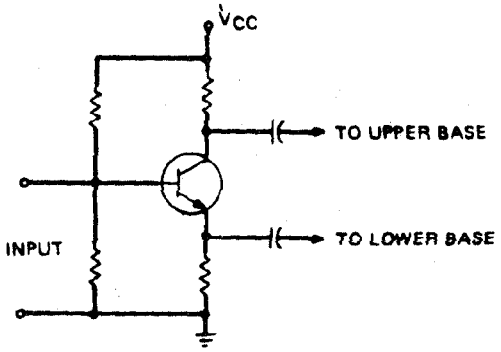


FIGURE 55. Split-load phase inverter.

A variation of the complementary circuit is the quasi-complementary circuit shown in Figure 57. This circuit allows the use of identical npn output transistors. The driver transistors give additional current gain, thus reducing the previous stage driving requirement.

The upper npn-npn pair of transistors is connected in the Darlington configuration and behaves as a high-gain npn device. The lower pnp-npn pair works as a high-gain pnp device. Analyzed in this way the circuit is seen to be the equivalent of that shown in Figure 56.

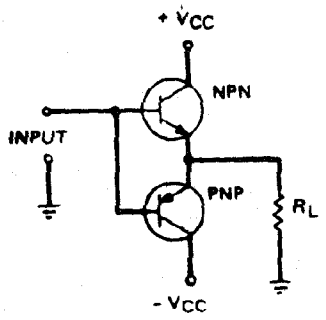


FIGURE 56. Complementary circuit.

5.3 TRANSISTORS, HIGH POWER

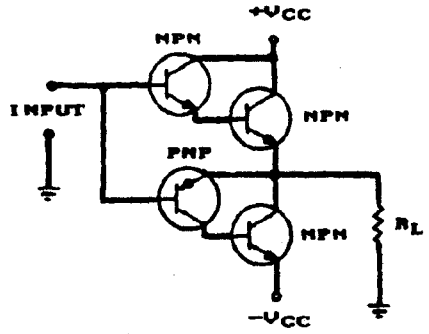
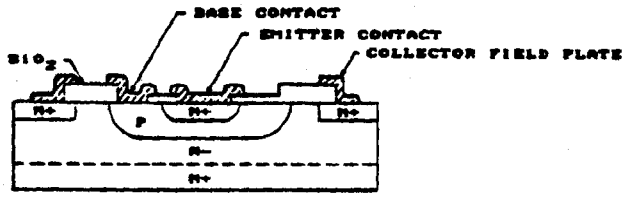


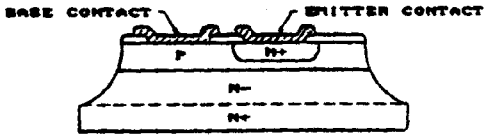
FIGURE 57. Quasi-complementary circuit.

5.3.3 Physical construction.

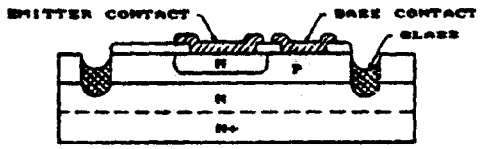
5.3.3.1 Die construction. Two basic types of structures are used in die fabrication for high-power transistors. These are the mesa (single and double) and planar structures. These structures are shown in Figure 58.



A. Planar



B. Nonpassivated mesa



C. Glass-passivated double mesa

FIGURE 58. Basic die structures.

5.3 TRANSISTORS, HIGH POWER

The planar structure offers the lowest leakage current because both the E-B and C-B junctions are always protected by ultra clean silicon dioxide or some other type of primary passivant. For spaceflight and critical aerospace applications semiconductor devices containing a conformal coating should not be used. Conformal coatings may have inherent problems when exposed to radiation and temperature extremes, such as outgassing, and physical and/or chemical alterations.

Moreover, three basic die processes are used to produce these two types of structures. These are the double diffused (epi-collector), triple diffused, and epi-base processes. It is, therefore, possible by using a combination of these processes and structures, to produce a variety of transistor types (with the exception of an epi-base planar). Consequently, three manufacturers supplying the same JANTXV or JANS device could provide devices containing dice with different structures and process combinations. Although these devices may pass all electrical requirements, they might possess different electrical characteristics. Such a situation can lead to a problem known as the unspecified parameter, that is, whereas one manufacturer's JANTXV devices functions well in an application, the same type from another manufacturer might not.

5.3.3.2 Packaging. Some of the standard packages used in the packaging of these devices are the TO-3, TO-66, TO-111 and MT-53. Figure 59 shows the basic construction of a TO-3 package.

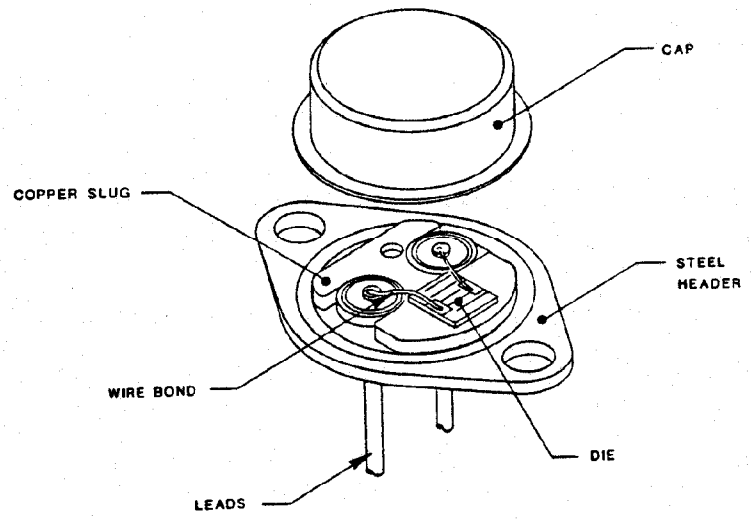


FIGURE 59. Typical TO-3 transistor construction.

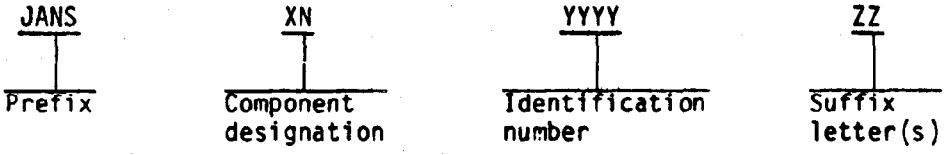
Die bonding methods used can be either soft solder or gold eutectic with moly interface. Both methods have been found to provide reliable bonds. However, if power cycling capability greater than 6000 cycles (at T_C deltas of 100 °C) are required, gold eutectic/molybdenum would be the best choice.

5.3 TRANSISTORS, HIGH POWER

Although there are various methods of making a die-to-lead interconnections the most popular wirebonding technique is ultrasonic bonding to the die and ultrasonic or electrical weld to the lead, with aluminum as the interconnecting wire material.

Sealing of the package is done in an inert atmosphere using electrical weld techniques. Due to this welding technique, it is possible that weld splashes might occur causing particles to be released inside the package. Nonetheless, all MIL-STD-975 devices require that a PIND and X-ray test be performed to remove any defective devices.

5.3.4 Military designation. The military designation for transistor is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 2N for transistors.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices, suffix M, shorter or longer terminal leads, suffix S or L, or any other letter to indicate a modified version.

5.3.5 Electrical characteristics. A knowledge of the capabilities and limitations of power transistors is essential for the economical and reliable application of these devices. Much of the required information can be obtained from the manufacturer's technical data sheets. However, the circuit applications engineer must have a clear understanding of the meanings of the various ratings and characteristics, how they were derived by the manufacturer and how they are related to his application in order to make optimal use of such information.

Some of the basic electrical characteristics have already been discussed in paragraph 5.2.5 in the low-power subsection; these are also applicable to high-power devices. The following paragraphs will emphasize some of the electrical characteristics of high-power devices and discuss thermal considerations.

5.3.5.1 Breakdown and leakage characteristics. I_{CBO} is the dc current that flows when the collector-base junction is reverse-biased by a specified voltage. It is temperature dependent; therefore, the temperature must be specified as one of the test conditions. A typical collector blocking characteristic is shown in Figure 60. As the voltage is increased, a value is finally reached beyond which the current increases very rapidly for small increments of voltage. This is the avalanche region. The breakdown voltage [$(V_{(BR)CBO})$] is that voltage in or near the avalanche region.

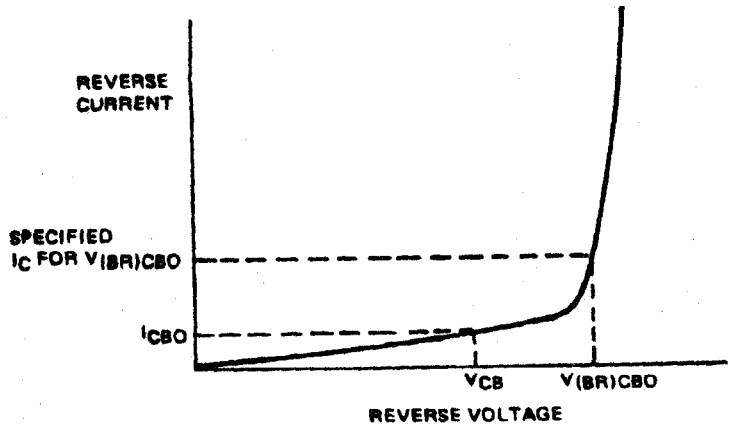


FIGURE 60. Collector junction blocking characteristics.

I_{EBO} is the dc current which flows when the emitter-base junction is reverse-biased by a specified voltage. The same comments and shape of characteristic curves apply to I_{EBO} as to I_{CBO}. However, in high-power transistors the emitter breakdown voltage [$(V_{(BR)EBO})$] is usually designed to be much less than $V_{(BR)CBO}$.

I_{CEV}, I_{CES}, I_{CER}, and I_{CEO} are the dc currents that flow when the collector and emitter terminals are connected to a specified voltage with polarity such that the collector-base junction is reverse-biased and the base terminal is returned to the emitter terminal through a circuit designated by the third letter in the subscript as follows:

V -- reverse base-emitter bias
S -- short circuit
R -- resistance
O -- open circuit.

The common emitter blocking currents are also temperature-dependent, particularly I_{CEO}. Figure 61 compares the relative leakage current levels for these four parameters. Also shown are the corresponding breakdown voltages.

5.3 TRANSISTORS, HIGH POWER

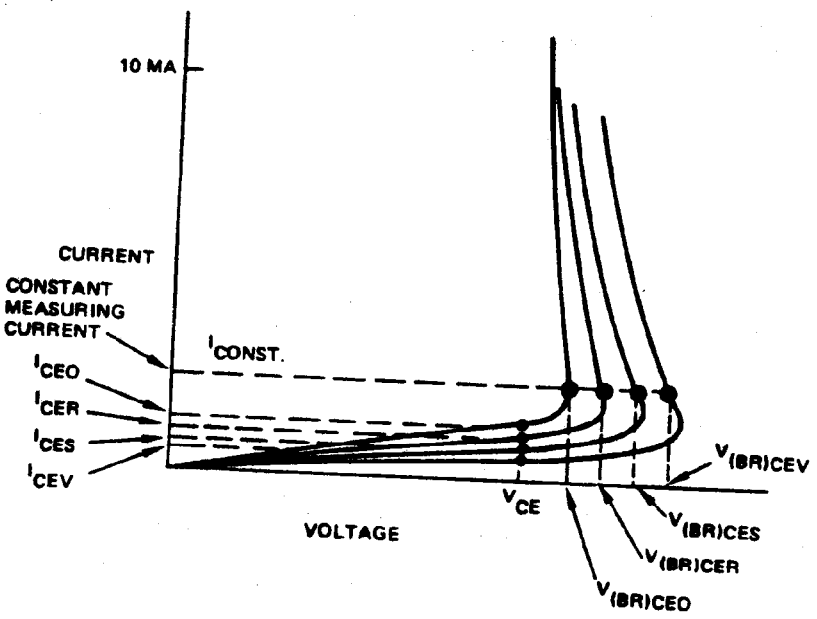


FIGURE 61. Common emitter blocking characteristics.

5.3.5.2 Sustaining voltages. Because of the negative resistance region, the breakdown voltages described above are still inadequate to specify completely the voltage capability of a transistor. Whereas the breakdown voltages are usually specified at a current of several milliamperes, the applications engineer may well be interested in the voltage limitations of full rated current.

The complete collector family of curves for a typical 30 ampere transistor is shown in Figure 62. It should be noted that the current scale of this family is greatly compressed as compared to that of Figure 61. Also, the shape of the blocking portion of the curves is somewhat exaggerated for the sake of clarity.

As shown in Figure 62, the V_{CEO}, V_{CER}, V_{CES}, and V_{CEV} curves, after passing through the breakdown points ($V_{(BR)CEO}$, etc.) and the negative resistance region, finally reach minimum values. This portion of the curves is known as the sustaining region because the voltage remains relatively constant over a wide range of collector currents. Also shown in Figure 62 are the curves for various constant values of base drive, I_{B1}, I_{B2}, etc., which constitute the

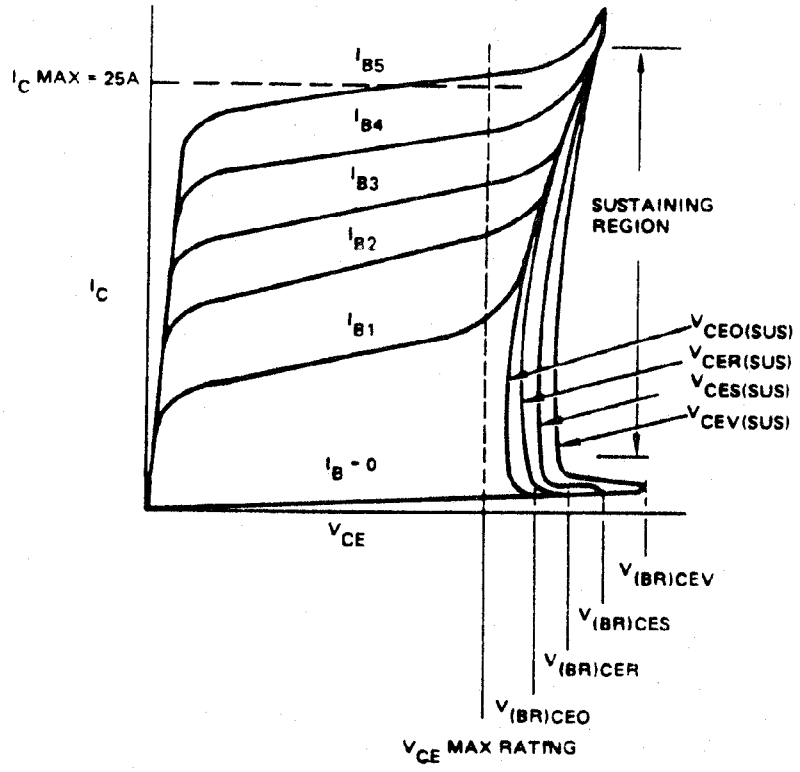


FIGURE 62. Transistor collector family of curves.

active region of the collector family. All of the forward-bias base curves finally bend over and become asymptotic to the open-base sustaining voltage curve.

The symbol for designating the sustaining voltage level is formed by adding (sus) to the suffix; e.g., $V_{CEO(sus)}$, although the nomenclature BV_{CEO} has been used in some literature. The sustaining voltages are specified at a current near the minimum voltage portion of the curves. For most high-power transistors, this current is in the range of several hundred milliamperes. The open base sustaining voltage [$(V_{CEO(sus)})$] is one of the most important parameters

5.3 TRANSISTORS, HIGH POWER

for a power transistor because it is the absolute upper voltage limit of the collector family in the active area.

5.3.5.3 DC current gain. The static value of the forward current transfer ratio (dc current gain in the common emitter configuration) is one of the most important parameters for transistors. Its value is required in low frequency applications involving large current swings and also in high frequency applications to establish biasing conditions. The common emitter dc current gain is defined as the ratio of dc collector current to dc base current and is represented by the symbol h_{FE}.

$$h_{FE} = \frac{I_C}{I_B}$$

The value of h_{FE} is not constant, but varies with collector voltage V_{CE}, collector current I_C, and junction temperature T_J. It is, therefore, necessary to specify V_{CE}, I_C, and T_J as test conditions in order to measure h_{FE}.

V_{CE} is usually specified at a value sufficiently above the saturation voltage so that the test operating point is beyond the region of maximum curvature shown in Figure 63.

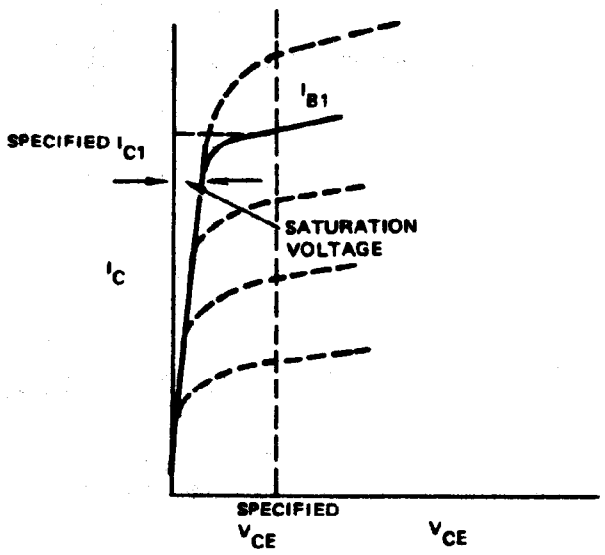
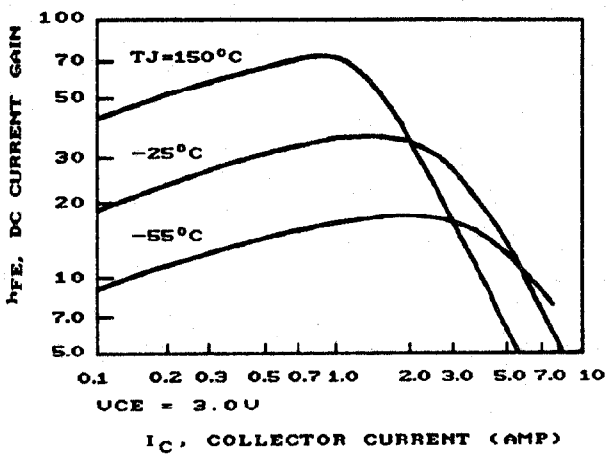


FIGURE 63. Test conditions for h_{FE}.

A typical curve for h_{FE} versus I_C at constant V_{CE} and three temperatures is given for high-power transistors in Figure 64.



T_J = 150 °C.

FIGURE 64. DC gain versus collector current.

5.3.5.4 AC current gain. The small-signal, short-circuit, forward current transfer ratio or ac current gain in the common emitter configuration is required in amplifier applications. It is defined as the ratio of ac collector current to ac base current and is represented by the symbol h_{fe}.

$$h_{fe} = \frac{di_c}{di_b} = \frac{I_c}{I_b}$$

at v_{ce} = constant.

If h_{fe} is plotted against frequency, the curve shown in Figure 65 results.

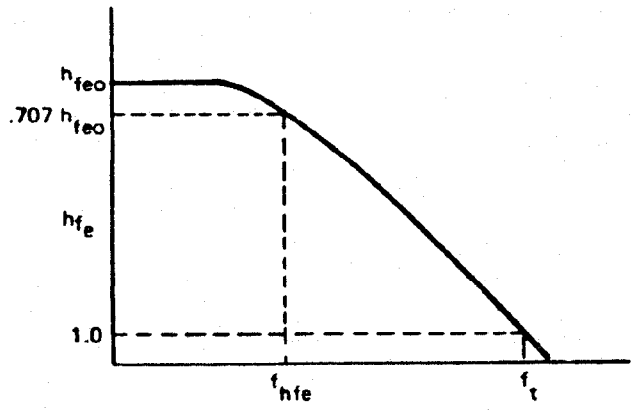


FIGURE 65. h_{fe} versus frequency.

5.3 TRANSISTORS, HIGH POWER

A number of important terms appearing on this curve should be noted:

a. h_{feo} is the low-frequency value of h_{fe}, usually measured at 1 KHz.

b. f_{hfe} (or β cutoff) is the frequency at which the ac gain has fallen to 0.707 h_{feo}.

c. f_t is the frequency at which the ac gain has fallen to 1.0.

At frequencies greater than 10 MHz, it is necessary to use high-frequency techniques to obtain ac gain measurements.

5.3.5.5 Thermal considerations.

5.3.5.5.1 Secondary breakdown. A power circuit design cannot be considered reliable unless the circuit has been checked to insure that the transistors will not undergo secondary breakdown. Operating within the power-temperature ratings and insuring against thermal runaway will not alone guarantee circuit reliability.

Figure 66 shows a sketch of the voltage-current characteristics of a transistor operating in the reverse breakdown mode. At low collector currents, the voltage across the device exceeds the open base breakdown rating. The peak of the curve and the negative resistance region is called either the first or normal breakdown and results from avalanche action in the transistor. However, as current in the avalanche mode is increased to higher values, a critical current (I_m) is reached at which the voltage across the device drops to a very low level. This behavior is called secondary breakdown.

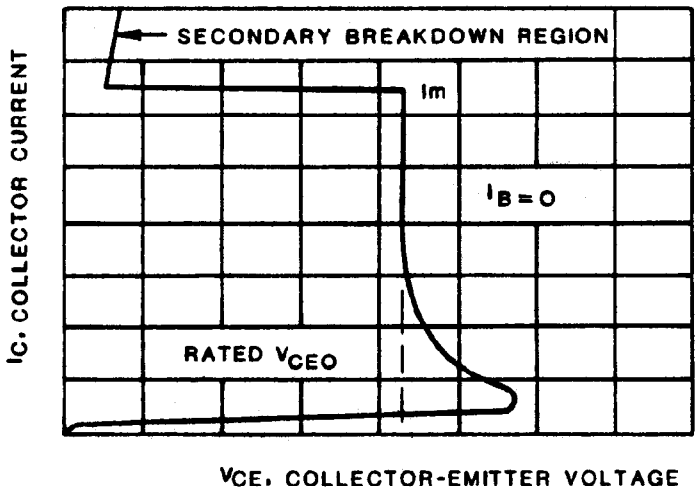


FIGURE 66. Manifestation of secondary breakdown in a transistor.

5.3 TRANSISTORS, HIGH POWER

Transistors need not be operating in avalanche breakdown in order to encounter secondary breakdown. Figure 67 shows a family of collector curves and the locus of critical or trigger currents at which the transistor enters secondary breakdown. As collector voltage is increased, maximum voltage occurs at lower currents and becomes extremely low as the emitter-base junction becomes reverse-biased. It has also been found that the amount of time a power pulse is applied at a particular operating point also determines whether or not secondary breakdown will occur. The observed behavior is a result of hot spots forming in the device as a result of nonuniform current density.

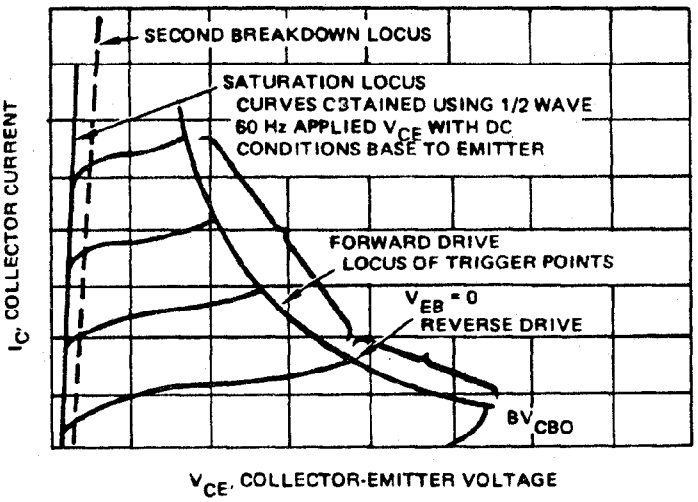


FIGURE 67. Locus of secondary breakdown trigger points.

ndary breakdown in transistors is triggered when the temperature of some part of the material achieves an extremely high value. It causes material to go into intrinsic conduction and become a very low resistance material. If the current is removed or held to very low values while the transistor is in secondary breakdown, no particular harm occurs and the semiconductor will act normally when it cools.

However, when the transistor enters secondary breakdown, a collector-emitter short results. The reason for this is that it is difficult to prevent the junction temperature from exceeding the melting point of the material. High pin-point temperatures in the transistor are produced because certain biasing conditions cause current to concentrate into a few very small areas. These small areas sustain the full collector current which normally, under ideal conditions, would be distributed evenly along the entire semiconductor active region. If the temperature at one of these hot spots becomes sufficiently high, secondary breakdown will occur. The resulting decrease in voltage across the device normally will result in an increase of current through the transistor because of the action of the other circuit elements. One small area must now

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.3 TRANSISTORS, HIGH POWER

sustain even more current than before, so the temperature can become extremely high. Extreme temperatures can rapidly change the characteristics of the device or melt the material. If enough melting occurs in the region of the high current path from collector to emitter, upon cooling a low resistance path or short circuit from collector to emitter will exist.

If it were possible to remove the current within a few microseconds after the onset of secondary breakdown, no particular damage to the transistor might result. Actually, under far less ideal conditions it has been possible to observe devices being switched in and out of secondary breakdown.

The current at which secondary breakdown occurs decreases markedly with increases in collector-emitter voltage when operating in the normal active region mode. The curves of Figure 68 illustrate this behavior and are typical for most types of semiconductors.

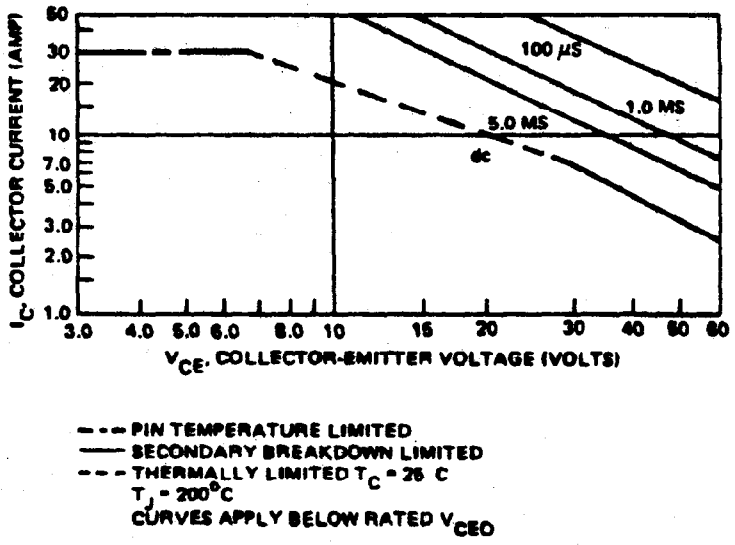


FIGURE 68. Example of an active region safe operating area.

The reason for collector voltage being an important variable is that as collector voltage is increased, the base width is reduced, thereby accentuating current crowding effects because the higher electric field caused by the shorter path reduces the current spreading or fan-out. Therefore, transistors with narrow base widths (i.e., higher f_t) will encounter secondary breakdown at lower power levels than lower frequency devices, other conditions being equal.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.3 TRANSISTORS, HIGH POWER

Because of the secondary breakdown problem, many vendors of power transistors provide safe operating area (SOA) charts as shown in Figure 68. The solid lines show secondary breakdown limitations whereas the dotted lines represent thermal limitations. For this transistor family, dc currents above 30 A cause excessive emitter pin temperatures (on low-level devices, I_C is limited by the bonding wire at low values of V_{CE}); therefore, operation above 30 Adc is not recommended. Above 6.5 V, the allowable dc current is junction temperature limited. If the case temperature is 25 °C, power dissipation must not exceed 200 W. The dc curve also shows that the dc power level must be lowered from the 200 W level as voltage increases above 30 V, if secondary breakdown is to be avoided. At case temperatures higher than 25°C, thermal resistance information must be taken into account to derate the power dissipation limit to keep the junction temperature (T_J) below its limits, $T_{J(max)}$. The secondary breakdown limitation curve is valid when $T_J < T_{J(max)}$. Therefore, the curve is not derated with increases in case temperature, because the curve is dependent on junction temperature only.

The various pulse curves given show no thermal limitation for currents to 50 A. The transistor power is limited solely by secondary breakdown. (At higher case temperatures the transient, thermal resistance must be used to determine if operation is within $T_{J(max)}$.) The given pulse curves are used if the duty cycle is 10 percent or less. For higher duty cycles, the curves will gradually degrade to the dc curve.

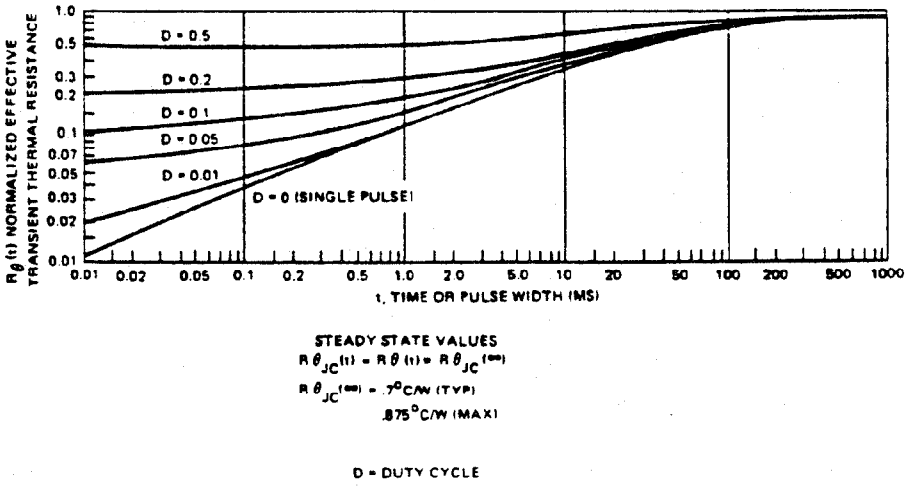


FIGURE 69. Example of thermal response data.

5.3 TRANSISTORS, HIGH POWER

It is common industry practice to rate power devices at a case temperature of 25 °C, even though it is very unlikely that the case would ever be held at 25 °C in a practical operating situation. It is instructive to find the effect on the safe area curve for this particular transistor if the case were at 125 °C.

In Figure 69, which shows the thermal response of a typical power transistor, the normalized effective thermal resistance, $R\theta(t)$, can be read for a single pulse (D=0) as 1.0 for dc, 0.23 for 5 ms, 0.13 for 1 ms, and 0.034 for 0.1 ms. The effective thermal resistance, $R\theta_{JC}(t)$, is found by multiplying these normalized values for the steady state value of 0.875 °C/W. The allowable power is found from

$$P_D = \frac{T_J - T_C}{R_{\theta JC}(t)}$$

where $T_J - T_C = 75°C$ in this example.

The power levels allowed are, respectively, 86, 370, 660, and 2500 W. These values can be used to construct a safe area curve based upon $T_C = 125$ °C as shown in Figure 70.

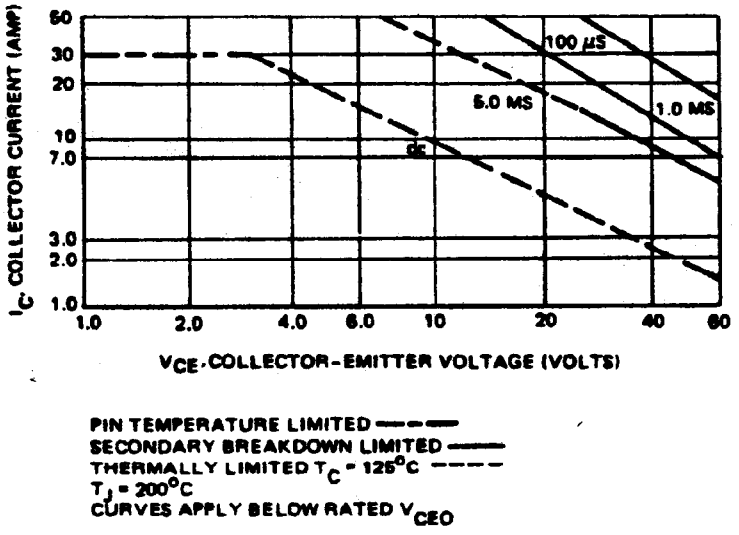


FIGURE 70. Constructed safe operating area data at $T_C = 125$ °C for transistor of Figures 67 and 68.

The secondary breakdown curves are unchanged. They are already based upon $T_J = T_{J(max)}$. The dc operation is entirely thermally limited and thermal and thermal limitations appear on the lower voltage ranges of the longer pulse times. However, at 100 µs no thermal limitations appear below 50 A. Such behavior is typical of most transistors; i.e., dc power becomes thermally limited at fairly low case temperatures, whereas under short pulse operation power dissipation is limited by secondary breakdown even at fairly high case temperatures.

The curves indicate the limits of power dissipation which may be sustained for a time interval as shown by the designated line; i.e., a 20-V, 30-A power pulse may be sustained for 1 ms without encountering secondary breakdown. In switching applications, the load line may traverse along one of the limit lines for the time indicated.

Figure 71 shows some of the circuits and load lines that might be encountered. Note that load lines can vary considerably in different applications because of the effects of transistor switching speed, circuit capacitance and inductance values, strays due to wiring, and transformer leakage inductance.

The first five circuits have typical inductive load lines. The most demanding is the transformer-coupled audio amplifier (see Figure 71A). Under normal operation, the load line is resistive and as long as it falls below the dc safe area line, no failures should occur. However, should the load be opened with signal applied, a reactive, nearly rectangular load line could result. It is important that the drive be restricted to keep the collector current within bounds and the whole load line below the dc safe operating area or failure could result.

The solenoid drivers (Figures 71B and C) cause relatively few problems. Peak currents are restrained by R_L and switch off is fast. The collector current remains fairly constant until the clamping level of the diode is reached, then the current transfers to the diode. The simplest and fairly reliable way to check for safe areas is to obtain the collector current waveform and make an equivalent rectangular model having the same peak amplitude and a width adjusted to have the same energy as the actual pulse. Assume collector voltage is constant at the clamp level (V_K)--it will generally rise to this level quickly--and note if this condition is within the safe area. For example, suppose the voltage rises from A to B in 1 µs and it takes 10 µs for the current to linearly decay from point B to point C. A 6 µs rectangular pulse of amplitude I_{max} would contain the same energy. Locate the point on the safe area at a current I_{max} and a voltage of V_K(point B) and note how much time is allowed there. If it is greater than the equivalent rectangular pulse, the transistor is safe. Generally, at times below 10 µs, all transistors can withstand their rated current and voltage simultaneously, so that this mode of operation seldom results in transistor failure.

5.3 TRANSISTORS, HIGH POWER

The power inverter (Figure 71D) and the ignition system (Figure 71E) can be treated in the same way as the relay drivers. The fact that no clamp exists in Figure 71D makes the circuit more sensitive to transients, but the critical point is still at the high power corner and the waveform modeling approach just described will work.

The switching motor control (Figure 71F) and the switching regulator (Figure 71G) differ from the previous circuits in that a large power transient occurs on both the switch-off interval (because of the inductor) and switch-on intervals (because of the diode recovery current); therefore, both intervals need to be checked. The same modeling principle applies, but both forward bias SOA (FBSOA) and reverse bias SOA (RBSOA) curves have to be scrutinized.

Should both voltage and current change at about the same speed, the foregoing approach may be too conservative. For a more exact analysis, it is necessary to plot the load line and note the time between certain intervals. A comparison to the SOA curve can then be made. Note that the pulse curves indicate that the transistor operating point can spend a given time at any one point on the curve or may travel along the curve for the same given time.

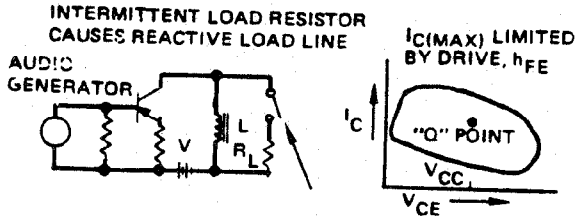
5.3.5.5.2 Power transistor cooling. Transistors with power ratings greater than 1 W are usually provided with a large, flat surface that can be clamped against a metal exchanger. The purpose of the heat exchanger is to transfer the heat to a larger surface from which it can then be dispersed by a cooling medium. The heat must pass through several thermal impedances as it flows from its source at the collector junction to the final cooling medium. These "thermal impedances" may be modeled as a series circuit, as shown in Figure 72.

The heat flowing through the thermal impedances will produce a succession of temperature drops, just as current flowing through electrical resistance will produce voltage drops. The value of the thermal impedance ($R\theta_{JC}$) between the collector junction and case is a function of the thermal design of the transistor. It is affected by the thermal conductivity of the various solders that bond the parts together. The $R\theta_{JC}$ values for transistors are given on their data sheets.

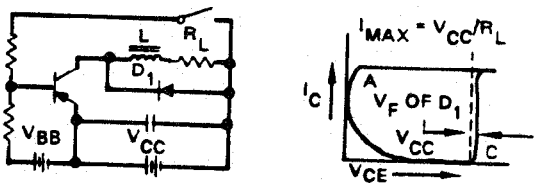
The thermal characteristics of the transistor are also specified indirectly on the data sheets in terms of thermal derating curves. These derating curves are often preferred over simple $R\theta_{JC}$ values because they include thermal transient effects and secondary breakdown limitations.

Although a maximum junction temperature is given in data sheets, this parameter cannot be measured directly. Consequently, most of the values and curves for transistor thermal analysis are based upon the case temperature (T_C) which can be measured by conventional thermocouple probes.

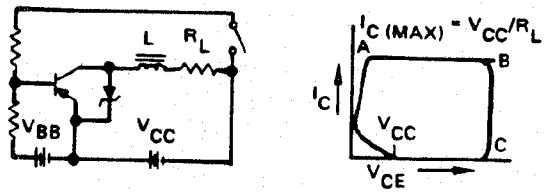
5.3 TRANSISTORS, HIGH POWER



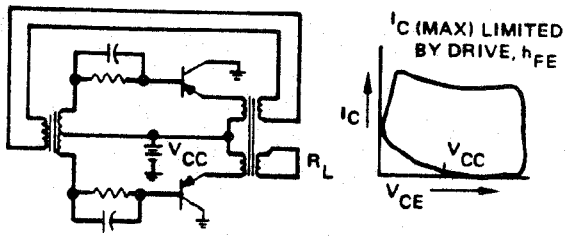
A. Class A audio amplifier



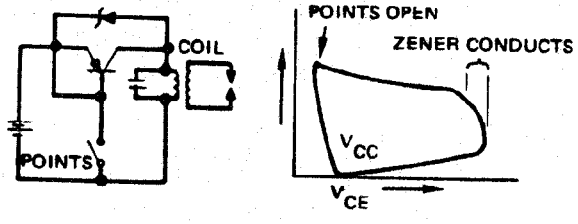
B. Solenoid driver with rectifier diode to suppress inductive kick



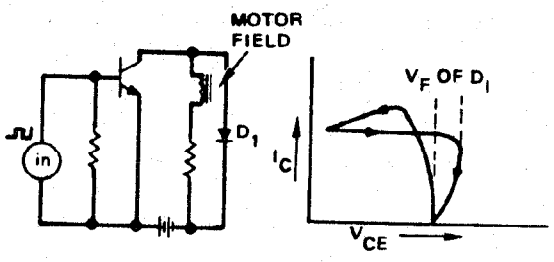
C. Solenoid driver with zener diode suppression



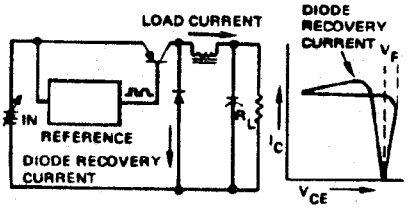
D. Power inverter



E. Ignition system



F. Switching type motor control



G. Switching type series regular

FIGURE 71. Examples of circuits.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.3 TRANSISTORS, HIGH POWER

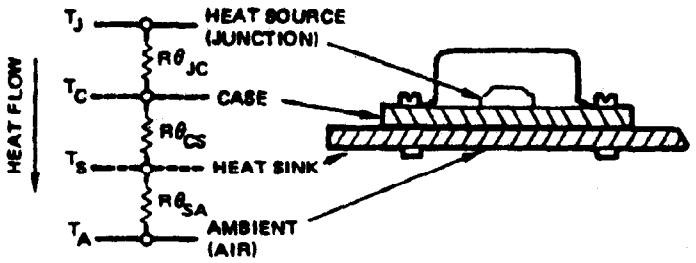


FIGURE 72. Thermal circuit.

Because of the importance of the case temperature in practical design problems, it is useful to note how T_C should be measured. The case temperature reference point for stud-mounted transistors is in the plane of the mounting surface, directly below the center of the collector junction. Access to this reference point is obtained by drilling an axial hole in the threaded stud as shown in Figure 73.

For case styles in which the mounting surface is clamped to the sink by bolts through a flange, such as the TO-3 case, the above procedure must be modified. The practical approach is to choose the nearest reference point to the ideal location. For the TO-3, the thermocouple should be placed into a very small hole drilled into the top surface of the mounting flange, at the end of the case closest to the transistor die (see Figure 74).

5.3.5.5.3 Mounting practices. The next thermal impedance in the heat removal circuit to be considered is $R\theta_{CS}$, the small but unavoidable thermal impedance at the interface between the transistor case and the heat sink (see Figure 72). In order to obtain the maximum rating from a power transistor, certain recommended heat sink mounting practices should be observed.

Regardless of the type of heat sink employed, the objective is to obtain the best possible thermal contact between the flat mounting surface of the transistor case and the sink. In order to avoid high thermal drop, a large mating surface area should be provided. Holes should be drilled or punched clean with no burr or ridges.

Careless deburring can seriously increase the thermal resistance and result in transistor overheating.

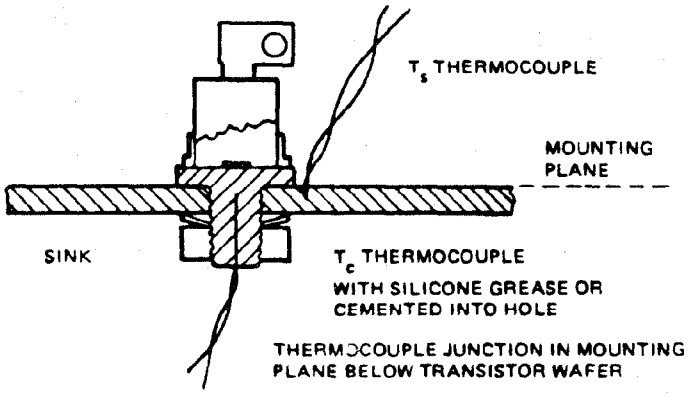


FIGURE 73. Thermocouple placement for stud-mounted case.

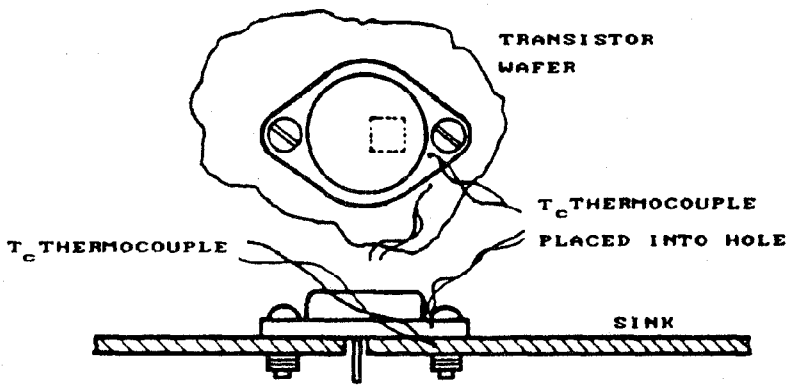


FIGURE 74. Thermocouple placement for TO-3 case.

However, even with the above precautions the surface will not be perfectly smooth. The two surfaces at the interface are actually in contact at only a limited number of points, with the remainder of the interface area consisting of a thin gap. Because air has very poor thermal conductivity, a marked reduction in $R\theta_{CS}$ can be obtained by filling the gap with an approved heat sink compound (HSC). Table V gives typical values of $R\theta_{CS}$ for various types of power transistors, both with and without HSC on the mounting interface. Care should be exercised to exclude foreign particles, particularly when using HSC, since this material can easily capture grit and dust during mounting.

5.3 TRANSISTORS, HIGH POWER

The nuts used on transistor studs should be carefully torqued down. Insufficient torque can lead to overheating, whereas excessive torque can strip threads, elongate studs and damage the transistor collector junction. It is important to remember that the studs are made from relatively soft copper to obtain high thermal conductivity. They cannot withstand the same torque as comparable-sized steel bolts. Torque wrenches should always be used when mounting power transistors and the specification limits for both maximum and minimum torque carefully observed. The effect of torque on the value of $R\theta_{CS}$ is illustrated in Figure 75.

Recommended contact force should be maintained throughout the operating temperature range and allowance made for the stress relaxation of components with time. The practice of screwing the transistor into a tapped hole should be avoided because temperature cycling can produce "thermal ratcheting" and gradual loosening of the device.

It is sometimes necessary to electrically insulate the transistor case from the sink by means of an insulating washer. Unfortunately, this washer contributes appreciably to the value of $R\theta_{CS}$, the magnitude of this contribution being shown in Table V. Therefore, it is preferable to mount the transistor directly on the sink and insulate the sink from the rest of the circuit wherever possible.

TABLE V. Typical values of $R\theta_{CS}$

Stud Threads	Thermal Impedance, °C/Watt			
	$R\theta_{JC}$	$R\theta_{CS}$ (Dry)	$R\theta_{CS}$ (HSC)	$R\theta_{CS}$ (Insulator-HSC)
1/4-28	0.75	0.7	0.3	1.1
5/16-24	0.7	0.5	0.2	0.7
5/16-24	0.5	0.5	0.2	0.7
1/2-20	0.45	0.3	0.15	0.5
TO-3	1.5	0.5	0.2	0.7

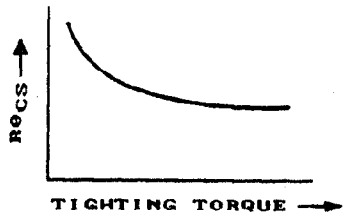


FIGURE 75. Relation between tightening torque and thermal impedance.

5.3.5.5.4 Heat sinks. The heat sink-to-ambient thermal impedance, $R\theta_{SA}$, is a function of the heat sink design. Copper and aluminum have the highest thermal conductivity (of the common metals) and are therefore usually preferred as heat sink materials.

An effective heat sink may take many forms. The simplest form is merely a flat, square plate of aluminum or copper 1/32- to 1/4-inch thick. This is mounted in the vertical plane to obtain the maximum "chimney effect." The term chimney effect, of course, relates to the natural convection type drafts set up as the air next to the surface of the heat sink is heated and rises, pulling more air past the heat sink surface. But even in these natural convection arrangements, it is possible to transfer heat away from the sink by radiation. This secondary transfer of heat by radiation can contribute appreciably to the total heat transfer, if the emissivity of the metal heat sink surface is increased by painting, oxidizing or anodizing (also see paragraph 5.3.5.5.5 Heat dissipation by radiation). Table VI shows some typical characteristics of simple, square plate heat sinks.

TABLE VI. Thermal characteristics of square plates

Sink Description	$R\theta_{SA}$, °C per Watt
5" x 5" x 1/8" Bright aluminum	4.3
5" x 5" x 1/8" Painted aluminum	2.5
7" x 7" x 1/8" Bright aluminum	2.7
7" x 7" x 1/8" Painted aluminum	1.6

5.3 TRANSISTORS, HIGH POWER

A further reduction in $R\theta_{SA}$ is obtained by using forced convection cooling. Figure 76 shows how increasing air velocity can markedly reduce the thermal impedance.

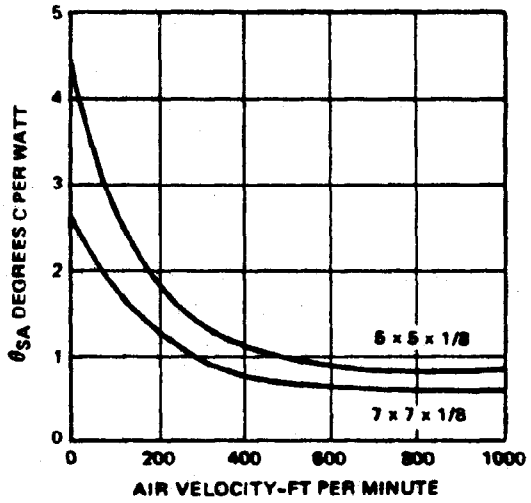


FIGURE 76. Sink-to-ambient thermal impedances for 1/8-inch thick square aluminum plates.

For applications where the thermal requirements are more critical, the simple flat plate sink may not be adequate. Die case and extruded aluminum sinks or assemblies of copper fins are available from many heat sink suppliers. They provide more efficient thermal paths and allow considerably more cooling capacity in a given space. The $R\theta_{SA}$ values for these improved heat sinks range from 3°C per watt to 0.3°C per watt for natural convection and as low as 0.1°C per watt for forced convection. Figure 77 shows how $R\theta_{SA}$ is related to overall volume for typical commercially available heat sinks. Additional information may be obtained from the heat sink manufacturers.

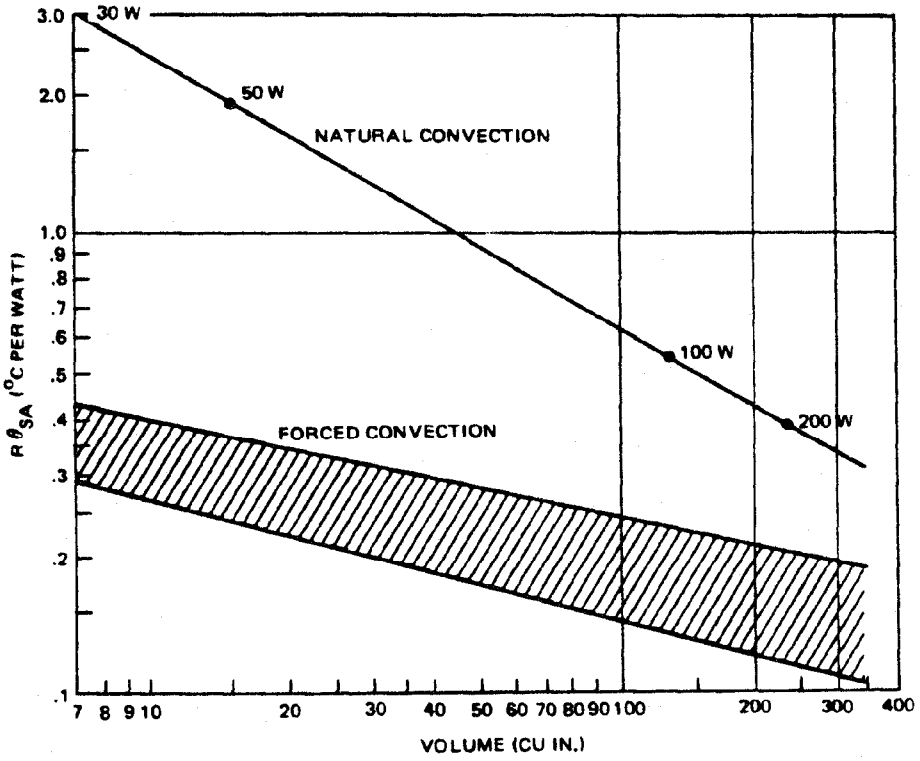


FIGURE 77. Thermal impedance, θ_{SA} versus volume (overall) of commercial heat sinks.

5.3.5.5.5 Heat dissipation by radiation. Radiation, the only mode of heat transfer that can take place across a vacuum, involves the transfer of photons from a hotter body, the emitter, to a cooler body, the absorber. An example is the heat one feels when standing next to a fireplace.

Opaque bodies absorb part and reflect the rest of the radiant energy which falls on the surface. The amount absorbed and reflected depends on the surface characteristics of the body, such as color and finish. Perfectly black bodies absorb all the radiant energy whereas perfectly shiny bodies reflect all energy. The radiation characteristics of a surface are characterized by a dimensionless quantity called the emissivity. A perfect absorber and emitter has an emissivity of unity whereas a perfect reflector has an emissivity of zero. Real bodies have emissivities between zero and one.

5.3 TRANSISTORS, HIGH POWER

The rate of heat transferred by radiation is low when the difference in temperatures between the emitting and absorbing bodies is small, say 20 °C or less, and the temperatures of the bodies are close to normal ambient temperature (25 °C). As shown in Figure 78, the thermal resistance decreases rapidly as the temperature difference increases, because the heat transfer rate is proportional to the quantity $T_E^4 - T_A^4$

where

T_E = the absolute temperature of the emitter
T_A = the absolute temperature of the absorber

The absolute temperature is measured relative to absolute zero; in the metric system, it is measured K.

Guidelines for minimizing the thermal resistance to radiation are:

a. High emissivity values for absorber and emitter

b. Good view of the absorber by the emitter

c. Larger absorber and emitter areas.

These guidelines are more important in high altitude (>70,000 ft) and a space environment where heat dissipation by natural or force convection is almost nonexistent.

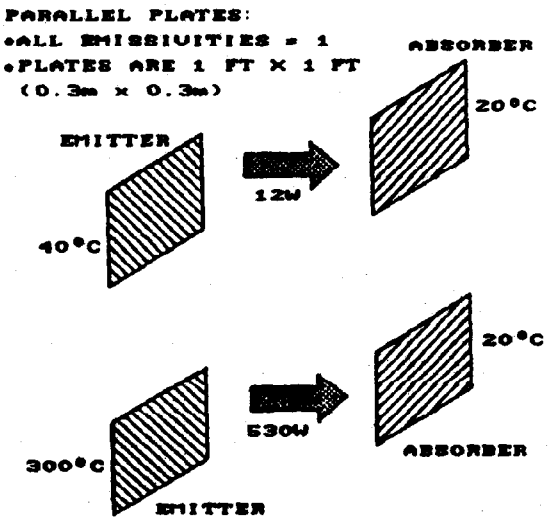


FIGURE 78. Rate of radiation heat transfer as a function of emitter and absorber temperatures.

5.3.6 Environmental considerations. Refer to paragraph 5.1.6.2 Environmental considerations.

5.3.7 Reliability considerations. Refer to paragraph 3.1.6 Reliability consideration, for a complete list of failure mechanisms and application considerations.

5.4 TRANSISTORS, FIELD EFFECT

5.4 Field effect transistors (FET).

5.4.1 Introduction. The field effect transistor is a semiconductor device in which the flow of charge carriers is controlled by a charge on the gate electrode. Unlike conventional transistors which are bipolar devices (i.e., performance depends on the interaction of two types of carriers, holes, and electrons, field-effect transistors are unipolar devices (i.e., operation is basically a function of only one type of charge carrier, holes in p-channel devices and electrons in n-channel devices).

There are two basic types of FET devices, the junction FET (JFET) and the metal-oxide-semiconductor FET (MOSFET). They can be structured to operate in either the depletion-mode or enhancement-mode; however, enhancement-mode JFETs are almost nonexistent. Both JFETs and MOSFETs are available as either p- or n-channel devices.

5.4.1.1 JFET operation. Before considering the operation of a JFET, a review of pn junction behavior will be valuable. Figure 79 illustrates a basic silicon pn junction.

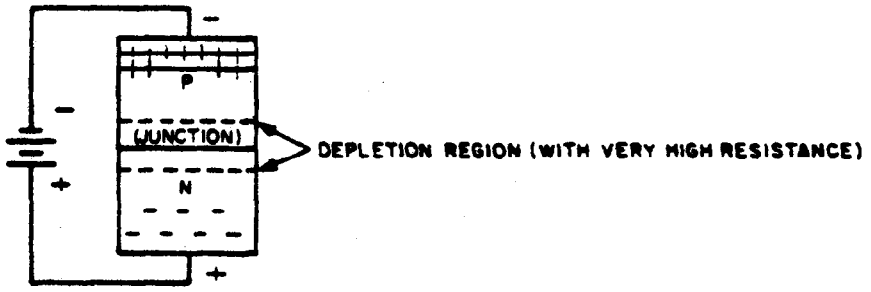


FIGURE 79. Basic silicon pn junction.

With the voltage connected as shown, the junction is reverse biased because the positive and negative carriers (holes and electrons) are attracted away from the junction. The empty area created at the junction is the key to JFET operation. This area is called the depletion region and contains neither holes nor free electrons in sufficient quantity for easy current flow; therefore, resistance is very high. Only a small leakage current flows.

The JFET can best be described as a bar of n-type silicon with two p-type regions diffused into it as shown in Figure 80.

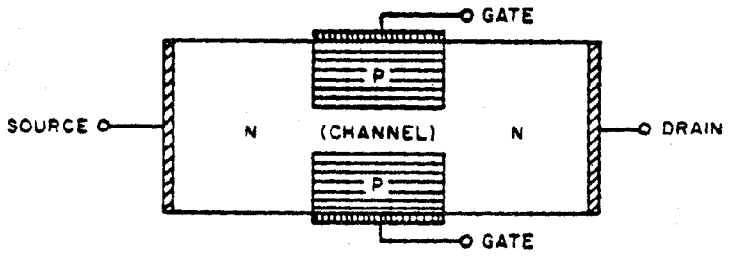


FIGURE 80. Physical configuration of a n-channel JFET.

The n-region between the two p-regions is called the channel. The source connection is so named because this is the source of the carriers which enter the device. The opposite connection is called the drain because this is where the carriers flow out. The gate terminals (which are normally connected together to one lead of the device or one external circuit connection) control the turn-on and turn-off.

In the configuration shown in Figure 81A, both junctions are reverse biased and a depletion region is created at the pn junction. If these regions extend across the channel and meet, a pinch-off condition is developed.

If a small positive voltage is applied on the drain as shown in Figure 81B, the two depletion regions are widened even more because of the reverse-bias voltage.

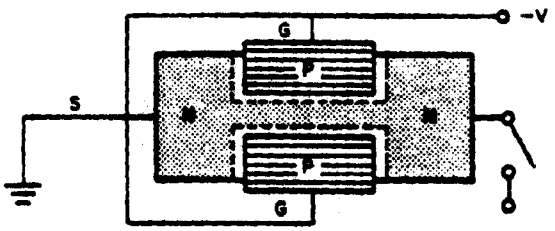
As reverse bias voltage is increased, the width of these depletion regions increases. If sufficient voltage is applied, the two depletion regions will extend completely through the channel and meet at one point, as shown in Figure 81C. In this condition, the channel for drain-source current (I_D) flow is pinched off.

The voltage value at which this condition occurs is a parameter normally specified on JFET data sheets as the gate-source cutoff voltage ($V_{GS(off)}$). Raising the voltage above $V_{GS(off)}$ will result in no appreciable decrease in I_D.

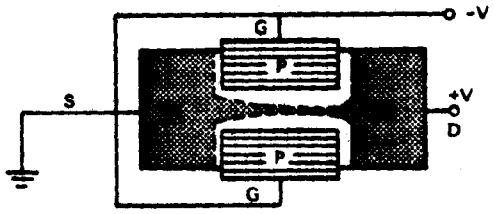
It can be seen from the above that a JFET transistor is a voltage-control device which can perform the complete switching function.

The operation of a p-channel JFET is the same, except that the polarities are reversed.

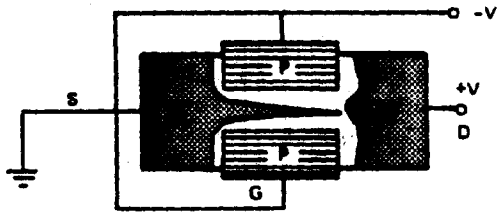
5.4 TRANSISTORS, FIELD EFFECT



A. Gate biased and drain unbiased



B. Gate and drain biased



C. Pinched off condition

FIGURE 81. Operation of an n-channel JFET.

5.4.1.2 MOSFET basic operation. The following paragraphs discuss the operation for both an enhancement-mode and a depletion-mode MOSFET device.

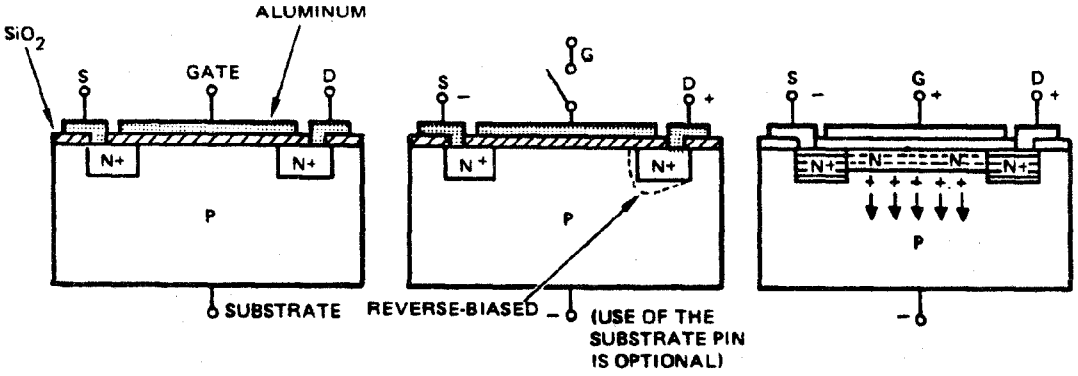
5.4.1.2.1 Enhancement-mode MOSFET. Figure 82A illustrates an unbiased enhancement-mode device.

An aluminum plate (gate) and the p-type substrate form equivalent plates of a capacitor, with the silicon dioxide acting as a dielectric. The other aluminum areas are used as bonding pads for the source and drain connections. If voltage is applied between the source and drain, only leakage current flows, because the pn junction will be reverse biased (gate open), as shown in Figure 82B.

In order to make the device conduct, a positive charge is applied to the gate. Because unlike charges attract, the available free electrons in the substrate are pulled up to the area under the silicon dioxide, as illustrated in Figure 82C. This concentration of electrons causes the normally p-type region to become n-type in a layer on the device between source and drain. This induced channel allows electrons to flow from source to drain. The gate-source voltage (V_{GS}) at which I_D barely starts to flow is called V_{GS} threshold (V_{GSth}). This configuration of a MOSFET, which uses the induced channel principle, is called an enhancement-mode MOSFET because the current flow is enhanced by the application of gate voltage.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The operation of a p-channel enhancement-mode device is the same except that polarities are reversed.



A. Unbiased B. Drain biased C. Gate and drain biased

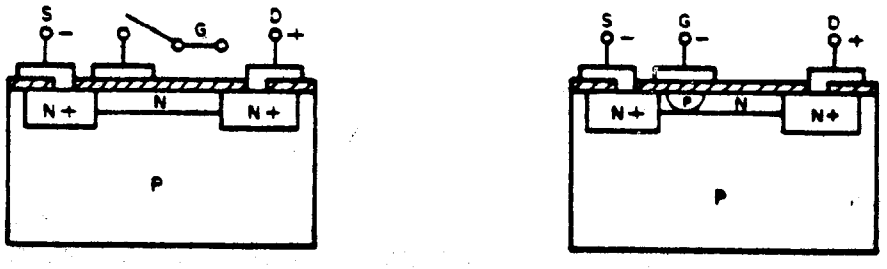
FIGURE 82. Operation of an n-channel enhancement-mode MOSFET.

5.4.1.2.2 Depletion-mode MOSFET. In this structure, a lightly doped n-channel is diffused into a p-material, as illustrated in Figure 82A. As a result, current will flow from source to drain with no voltage on the gate.

However, if a negative voltage is applied to the gate, the free electrons will be repelled out of the diffused channel, as illustrated in Figure 83B, creating a current-limiting condition similar to the no-gate-voltage condition of the enhancement-mode MOSFET described previously. The V_{GS} required to cutoff I_D is also called $V_{GS(off)}$ as in a JFET device. In fact, JFET and MOSFETs have very similar characteristics, because they are both depletion type devices.

Again the operation of a p-channel depletion-mode device is the same except that polarities are reversed.

5.4 TRANSISTORS, FIELD EFFECT



A. Drain biased

B. Gate and drain biased

FIGURE 83. Operation of an n-channel depletion-mode MOSFET.

5.4.2 Usual applications. There are three basic single-stage amplifier configurations for MOSFETs: common-source, common-drain, and common-gate. Each provides certain advantages in particular applications.

5.4.2.1 Common-source. The common-source arrangement shown in Figure 84 is most frequently used. This configuration provides a high input impedance, medium to high output impedance, and voltage gain greater than unity. The input signal is applied between gate and source, and the output signal is taken between drain and source. The voltage gain without feedback, A, for the common-source circuit may be determined as follows:

$$A = \frac{g_{fs}\, r_{os}\, R_L}{r_{os} + R_L}$$

where: g_{fs} is the gate-to-drain forward transconductance of the transistor

r_{os} is the common-source output resistance

R_L is the effective load resistance

The addition of an unbypassed source resistor to the circuit of Figure 84 produces negative voltage feedback proportional to the output current.

The voltage gain with feedback, A', for a common-source circuit is given by

$$A' = \frac{g_{fs}\, r_{os}\, R_L}{r_{os} + (g_{fs}\, r_{os} + 1)\, R_s + R_L}$$

where Rs is the total unbypassed source resistance in series with the source terminal.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The common-source output impedance with feedback, Z_o, is increased by the unbypassed source resistor as follows:

$$Z_o = r_{os} + (g_{fs}\, r_{os} + 1)\, Rs$$

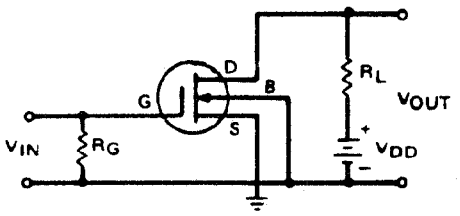


FIGURE 84. Basic common-source circuit for a MOSFET.

5.4.2.2 Common-drain. The common-drain arrangement, shown in Figure 85, is frequently referred to as a source-follower. In this configuration, the input impedance is higher than in the common-source configuration, the output impedance is low, there is no polarity reversal between input and output, the voltage gain is always less than unity, and distortion is low. The common-drain is used in applications which require reduced input-circuit capacitance, downward impedance transformation, or increased input-signal-handling capability.

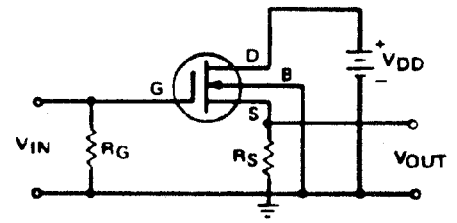


FIGURE 85. Basic common-drain circuit for a MOSFET.

5.4 TRANSISTORS, FIELD EFFECT

The gain A' of the common-drain circuit is given by

$$A' = \frac{Rs}{\frac{\mu + 1}{\mu} Rs + \frac{1}{g_{fs}}}$$

The amplification factor (μ) of a MOSFET is usually much greater than unity. The equation for the gain of a common-drain can be simplified as follows:

$$A' = \frac{g_{fs} R_s}{1 + g_{fs} R_s}$$

For example, if the gate-to-drain forward transconductance, g_{fs}, is 2000 microsiemens [2 x 10^{-3} siemen], and the unbypassed source resistance, R_s, is 500 ohms, the resulting stage gain, A', would be 0.5. If the same source resistance is used with a transistor having a transconductance of 10,000 microsiemens [1 x 10^{-2} siemen], the stage gain would increase to 0.83.

If the resistor, R_g, were returned to ground, as shown in Figure 85, the input resistance, R_1, of the source follower would equal R_g. However, if R_g were returned to the source terminal, the effective input resistance, R_1', would be given by:

$$R_1' = \frac{R_G}{1 - A'}$$

where A' is the voltage amplification of the stage with feedback.

If the load resistance and the effective input capacitance, C_1', of the common-drain are reduced by the inherent voltage feedback, the following equation holds:

$$C_1' = c_{gd} + (1-A')\, c_{gs}$$

where c_{gd} and c_{gs} are the intrinsic gate-to-drain and gate-to-source capacitances, respectively.

The effective output resistance, R_o', of the common-drain stage is given by the following equation:

$$R_o' = \frac{r_{os} R_s}{(g_{fs} r_{os} + 1) R_s + r_{os}}$$

where r_{os} is the transistor common-source output resistance
g_{fs} is the gate-to-drain forward transconductance
r_{os} is the common-source output resistance
R_s is the source resistance, and r_o' is the effective output resistance.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The common-drain output capacitance, C_o', may be expressed as follows:

$$C_o' = c_{ds} + c_{gs} \left(\frac{1 - A'}{A'} \right)$$

where c_{ds} and c_{gs} are the intrinsic drain-to-source and gate-to-source capacitances, respectively, of the MOS transistor.

5.4.2.3 Common-gate. The common-gate circuit shown in Figure 86 is used to transform a low input impedance to a high output impedance. The input impedance of this configuration has approximately the same value as the output impedance of the source-follower circuit. The common-gate circuit is also a desirable configuration for high-frequency applications, because its relatively low voltage gain makes neutralization unnecessary in most cases. The common-gate voltage gain, A, is given by

$$A = \frac{(g_{fs}\, r_{os} + 1)\, R_L}{(g_{fs}\, r_{os} + 1)\, R_G + r_{os} + R_l}$$

where Rg is the resistance of the input-signal source.

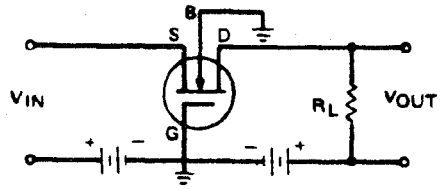


FIGURE 86. Basis common-gate circuit for a MOSFET.

5.4.3 Physical construction.

5.4.3.1 Die structure. Figure 87 illustrates the basic JFET structure, and Figure 88 shows the basic enhancement-mode MOSFET.

5.4 TRANSISTORS, FIELD EFFECT

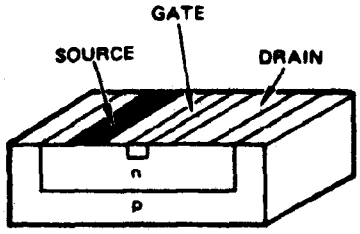


FIGURE 87. Structure of an n-channel JFET.

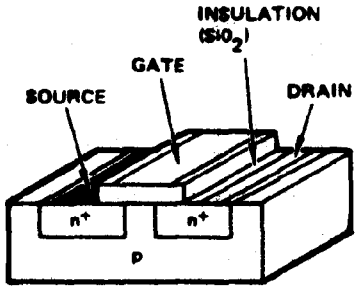


FIGURE 88. Structure of an n-channel enhancement-mode MOSFET.

5.4.3.2 Packaging. Some of the standard packages used for FET devices are the TO-5, TO-18, TO-39, and TO-72. Figure 89 shows the basic construction for a TO-5 device.

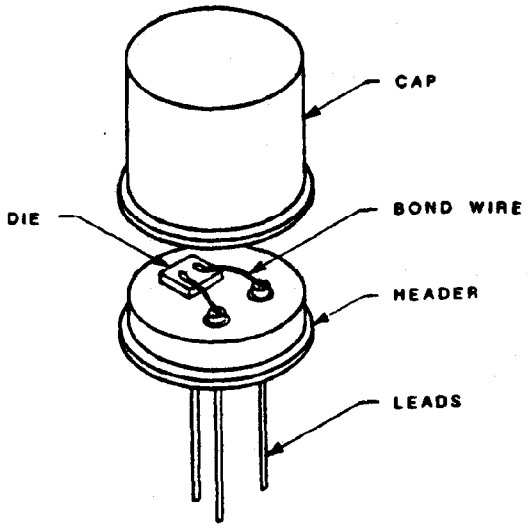


FIGURE 89. Typical TO-5 FET construction.

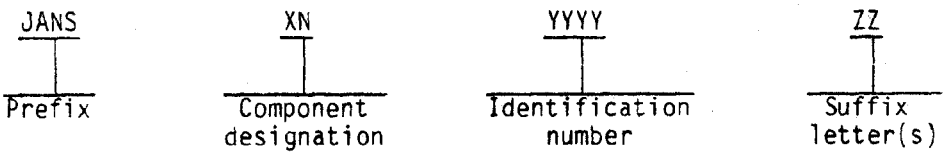
5.4 TRANSISTORS, FIELD EFFECT

Die bonding of these devices can be with either soft solder or gold eutectic. Both methods have been found to be reliable.

Wire bonding is done by using either aluminum or gold wires. The wire bonding methods most frequently used are thermocompression or ultrasonic bonding.

Sealing of the package is performed in an inert atmosphere by using electrical welding techniques; consequently, it is possible that weld splashes may occur and cause particles to be released inside the package. However, all MIL-STD-975 devices require 100 percent PIND testing and X-raying which would detect any defects.

5.4.4 Military designation. The military designation for transistors is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975.

The component designation is 2N for transistors.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices, suffix M, shorter or longer terminal leads, suffix S or L, or any other letter to indicate a modified version.

Radiation hardness assurance (RHA) designation code is placed after the prefix. See MIL-S-19500 for details.

5.4.5 Electrical characteristics. In both types of FETs, the operation is similar. Both rely on changing the charge on the gate electrode and the amount of charge in the conductive channel to change the drain current (see Figure 90).

The current-voltage relation of the FET is nearly the square law; it is

$$I_D = I_{DSS} \left[1 - \frac{V_{GS}}{V_{GS(off)}} \right]^2$$

5.4 TRANSISTORS, FIELD EFFECT

where

V_{GS} is the gate voltage

$V_{GS(off)}$ is the value of V_{GS} necessary to reduce the drain current to zero

I_{DSS} is the zero-gate-voltage drain current (saturation drain current)

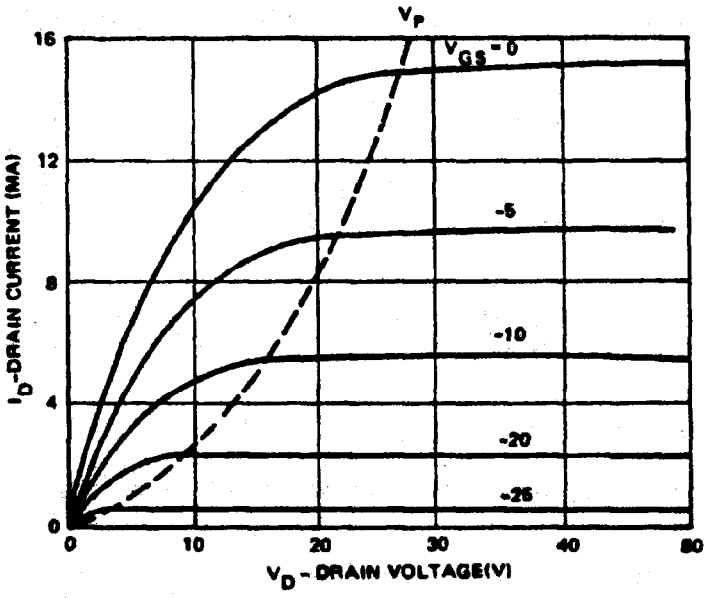


FIGURE 90. Typical output characteristics of a JFET.

The transconductance of the FET varies with gate voltage according to the relation:

$$\frac{\gamma I_D}{\gamma V_G} = g_m = g_{mo}\left[1 - \frac{V_{GS}}{V_{GS(off)}}\right]^{n-1}$$

5.4 TRANSISTORS, FIELD EFFECT

The basic current-voltage relationship for a depletion-mode MOSFET operating in the common-source configuration is shown in Figure 91. At low drain-to-source potentials and with the gate returned to source (V_{GS} = 0), the resistance of the channel varies linearly with voltage, as illustrated in region A-B. As the drain current increases beyond point B, the voltage drop in the channel produces a progressively greater voltage difference between the gate and points in the channel which are closer to the drain. As the potential difference between gate and channel increases, the channel is depleted of carriers (becomes constricted), therefore the drain current increases at a much slower rate with further increase in drain-to-source voltage, as shown in region B-C. An additional increase in drain-to-source voltage beyond point C produces no change in drain current until point D is reached. This condition leads to the description of region B-D as the "pinch-off" region. Beyond point D, the transistor enters the "breakdown" region and the drain current may increase dramatically. The upper curve in Figure 91 also applies to enhancement-mode MOSFETs, provided that the gate voltage V_{GS} is large enough to produce channel conduction.

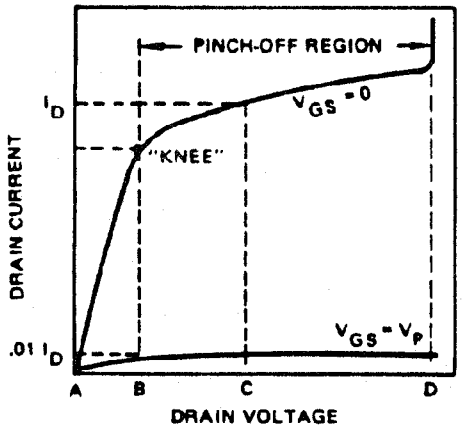


FIGURE 91. Basic current-voltage relationship for a depletion-mode MOSFET.

A MOSFET channel may achieve self pinch-off by the intrinsic IR drop alone, by a combination of intrinsic IR drop and an external voltage applied to the gate, or by an external gate voltage alone, which has the same magnitude as the self pinch-off IR drop V_p. In any case, channel pinch-off occurs when the sum of the intrinsic IR drop and the extrinsic gate voltage reaches V_p. The pinch-off voltage V_p is usually defined as the gate cutoff voltage $V_{GS}(off)$ that reduces the drain current between 0.1 and 1 percent of its zero-gate-voltage value at a specified drain-to-source voltage (this corresponds to the "knee" voltage, point B in Figure 91, of the zero-gate-voltage output characteristics).

5.4 TRANSISTORS, FIELD EFFECT

The pinch-off region between points B and D in Figure 91 is where MOS FET transistors are especially useful as high impedance voltage amplifiers. In the ohmic region between points A and B, the linear variation in channel resistance makes the device useful in voltage-controlled resistor applications, such as the chopper unit at the input of some dc amplifiers.

Typical output-characteristic curves for n-channel MOSFETs are shown in Figure 92. (For p-channel MOSFET, the polarity of the voltage and current are reversed.) In the pinch-off region, the dynamic output resistance, r_{os}, of the transistor may be approximated from the slope of the output-characteristic curve at any given set of conditions.

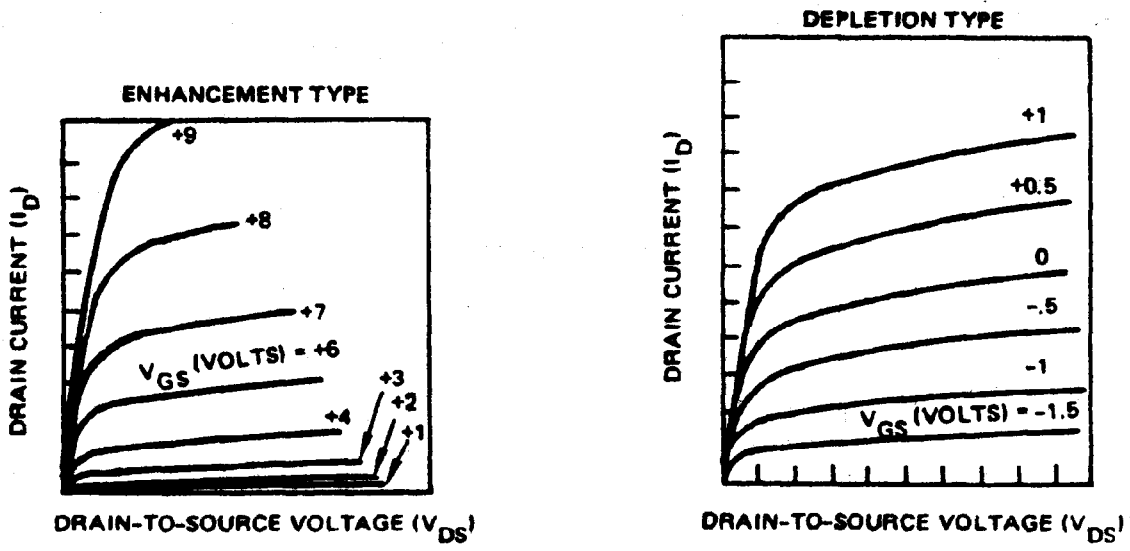
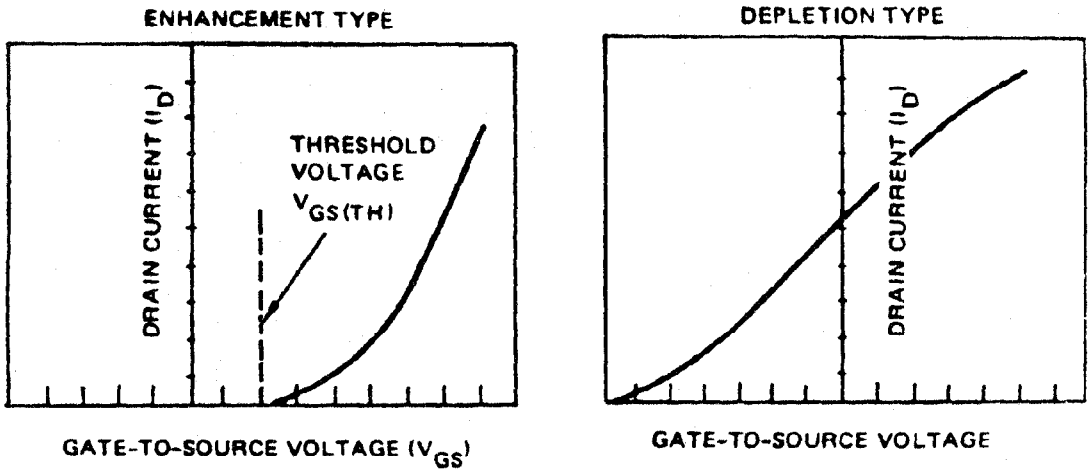


FIGURE 92. Typical output characteristics curves for a n-channel MOSFET.

Typical transfer characteristics for n-channel MOSFETs are shown in Figure 93 (polarities would be reversed for p-channel devices). The threshold voltage shown in Figure 93 is an important parameter for enhancement mode MOSFETs, because it provides a desirable region of noise immunity for switching applications.



A. Enhancement mode

B. Depletion mode

FIGURE 93. Typical transfer characteristics for a n-channel MOSFET.

5.4.6 Environmental considerations. Refer to paragraph 5.1.6.3, Environmental considerations in the General section.

5.4.7 Reliability considerations. The failure modes seen in FETs are also commonly seen in bipolar transistors. These failure modes are discussed in paragraph 5.1.6, "Reliability considerations" in the General section.

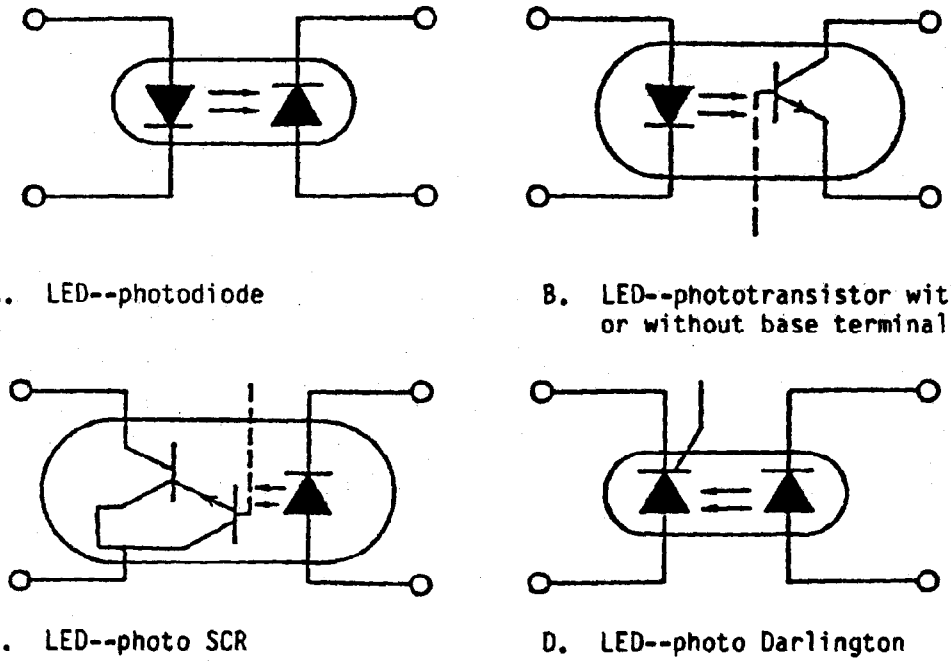
An additional defect may arise if caution is not exercised when applying voltage to the MOSFET device. Performance of MOS transistors depends on the relative perfection of the insulating layer between the control electrode (gate) and the active channel. If this layer is punctured by an inadvertent application of excess voltage to the external gate connection, the damage is irreversible. If the damaged area is relatively small, the additional leakage may not be noticed in most applications. However, greater damage may degrade the device to the leakage levels associated with JFET transistors. Therefore, it is very important that appropriate precautions be taken to ensure that the MOSFET gate-voltage ratings are not exceeded.

5.5 TRANSISTORS, OPTOCOUPLERS

5.5 Optocouplers.

5.5.1 Introduction. Optically coupled isolators, also called optocouplers, are used to isolate electrical systems from each other in an electronic circuit. Optocouplers allow very good circuit control with a high degree of electrical isolation between the input and output. These isolators are ideally suited for eliminating problems such as ground loop isolation, common mode noise rejection, and electromagnetic interference. These devices replace mechanical relays and pulse transformers. Optocouplers provide electrical isolation of potentially dangerous voltages in the equipment.

An optically coupled isolator consists of a light source coupled through an optically transparent insulation to a light detector and is housed in a light-excluding package. The light source may be an incandescent or neon lamp, or a light-emitting diode (LED). The transparent insulation may be air, glass or plastic. The detector may be a photoconductor, photodiode, phototransistor, photo silicon controlled rectifier (SCR), photo Darlington, or an integrated combination photodiode/amplifier. The discussion will be limited here to optocouplers having infrared-emitting diodes (IRED) coupled to a solid state photodetector. Diagrams of some basic optically coupled isolators are shown in Figure 94.



A. LED--photodiode

B. LED--phototransistor with or without base terminal

C. LED--photo SCR

D. LED--photo Darlington

FIGURE 94. Basic types of optically coupled isolators.

5.5 TRANSISTORS, OPTOCOUPLERS

Because optocouplers have no external optical properties, they are specified as electrical devices. Two functional characteristics that define an optically coupled isolator are how well they transfer information from input to output and how efficiently they maintain electrical isolation from input to output. Some important optocoupler parameters that determine these two characteristics are current transfer ratio (CTR), isolation resistance, isolation capacitance, and isolation voltage.

5.5.1.1 Current transfer ratio. The current transfer ratio, which is a measure of the optocoupler efficiency, is the ratio of the output current over the input current and is generally expressed as a percentage. It is dependent upon the radiative efficiency of the IRED, as well as the area, sensitivity, and amplifying gain of the detector.

5.5.1.2 Isolation resistance. This is the dc resistance between the input and the output of the optocoupler. Typically, the isolation resistance is very high and has a value of about 100 to 1000 GΩ (10^{11} to 10^{12} Ω).

5.5.1.3 Isolation capacitance. Since the optocoupler consists of two semiconductor materials (emitter and photodetector) separated by a transparent insulating material, this results in an isolation capacitance between the input and output. Typical values of isolation capacitance range from 0.5 to 3.0 pF.

5.5.1.4 Isolation voltage. One of the primary functions of an optically coupled isolator is to provide electrical isolation between the input and output circuits. The isolation rating refers to the maximum voltage difference that can be safely applied across the device without danger of insulation breakdown.

5.5.2 Usual applications. Optically coupled isolators are used to transmit information between electrically isolated electronic circuits. Typically, this isolation has been provided by relays, isolation transformers, line drivers, and receivers. However, if an optically coupled isolator is utilized, it will provide certain unique advantages over its mechanical counterparts.

a. Smaller size

b. Wide operating range

c. High gain

d. Faster operating speeds

e. High-voltage electrical isolation

f. Compatibility with circuits.

Some typical circuits and systems applications of optically coupled isolators follow:

5.5 TRANSISTORS, OPTOCOUPLERS

5.5.2.1 Line receivers. An optically coupled isolator can provide line isolation between two systems coupled by a transmission line. Figure 95 shows a typical interface system using TTL integrated circuitry coupled by a twisted pair line.

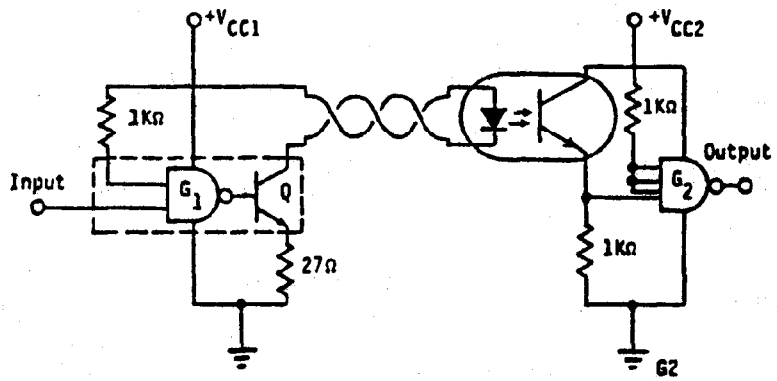


FIGURE 95. Typical interface system with TTL-integrated circuitry, twisted line pair, and optical coupler.

Gate G_1 and transistor Q constitute the input stage which drives the transmission line and emitter of the optically coupled isolator. At the receiving end of the line, the phototransistor is coupled to a fast switching gate (G_2) for fast pulse generation. In this system the optically coupled isolator provides isolation for both noise generated by electromagnetic interference and high voltage differences between the input and output.

An isolation standard transformer or relay can also accomplish this, but it would not be as fast as an optically coupled isolator. A standard line driver and receiver combination can eliminate noise and increase speed, but cannot provide isolation for high voltage differences between input and output. Digital input units containing optical isolators protect the microcomputer from standard as well as over voltage conditions.

5.5.2.2 Solid state relays. Optically coupled isolators can eliminate undesirable relay noise and spikes in relay circuits by providing a high degree of isolation between the input and output. A block diagram of a typical solid-state relay using an optically coupled isolator is shown in Figure 96.

The control stage consists of discrete transistors or integrated circuits whereas the output stage consists of high-power switching devices.

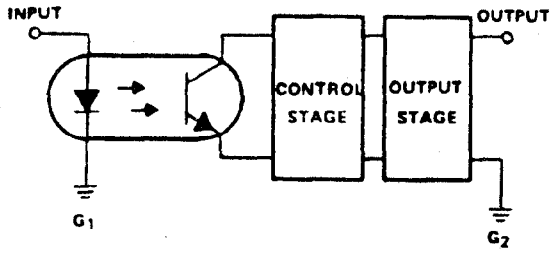


FIGURE 96. Typical solid state relay using an optically coupled isolator.

5.5.2.3 Digital logic interface. A very useful application of optically coupled isolators is interfacing between digital systems. An optically coupled isolator has output currents compatible with TTL inputs. It is beneficial to use this device when high voltage differences exist between systems. Other important applications of an optically coupled isolator are those involving 54/74 TTL and similar digital integrated circuit families. There is a wide variety of standard TTLs with different logic levels or logic conditions. A general interface circuit using an optocoupler is shown in Figure 97.

When the output of a logic circuit 1 (Gate 1) is low (V_{OL1}), the output of the optically coupled isolator is also low (V_{OL2}). Since V_{OL2} is the input to logic circuit 2 (Gate 2), it must be less than the maximum required low input voltage (V_{IL2}), to hold logic circuit 2 (Gate 2) in a stable state as shown below.

$$V_{OL2}\ (\text{isolator}) < V_{IL2}\ (\text{max})$$

Optically coupled isolator specifications should be evaluated to be level compatible with the TTL logic. The R_1 and R_2 values can be calculated using the following equations:

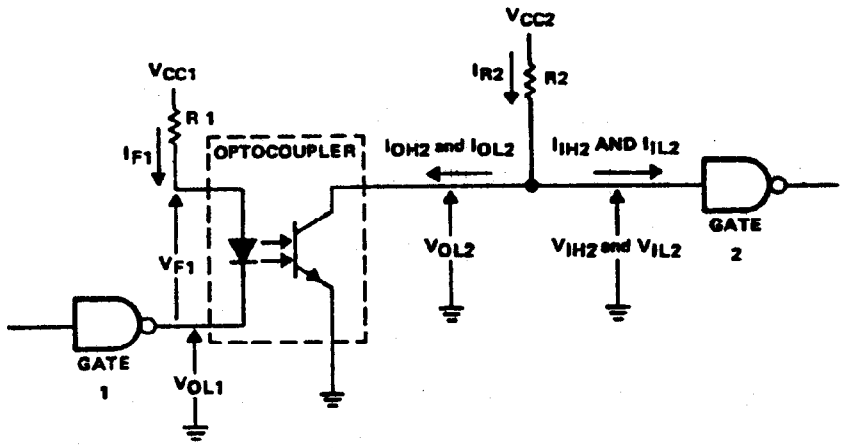
$$R_1 = \frac{V_{CC_1} - V_{F_1}(\text{typ}) - V_{OL_1}(\text{typ})}{I_{F_1}(\text{typ})}$$

$$R_2(\text{min}) = \frac{V_{CC_2}(\text{max}) - V_{OL_2}(\text{max})}{I_{OH_2}(\text{max}) + I_{IH_2}(\text{max})}$$

$$R_2(\text{max}) = \frac{V_{CC_2}(\text{min}) - V_{IH_2}(\text{min})}{I_{OH_2}(\text{max}) + I_{IH_2}(\text{max})}$$

R_2 is selected between the limits of R_2(min) and R_2(max).

5.5 TRANSISTORS, OPTOCOUPLERS



NOTE: V_{OL_2} = low-level output voltage of coupler when coupler is on
V_{IL_2} = low-level input voltage specified for Gate 2.

FIGURE 97. Optocoupler interface circuit.

5.5.2.4 Pulse amplifiers. Pulse amplification, as well as isolation, can be achieved by using an optically coupled isolator with a pulse amplifier. This circuit is shown in Figure 98, which illustrates an optical isolator used with a UA741 operational amplifier to amplify the pulse appearing at the anode of the IRED. The gain of this circuit is controlled by the feedback resistor R_F.

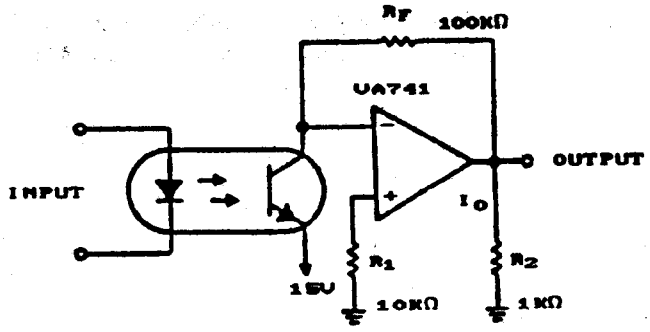


FIGURE 98. Isolated pulse amplifier using optically coupled isolator and operational amplifier.

5.5.3 Physical construction, packaging configurations, and device combinations. An optically coupled isolator contains both an IRED and a photodetector in the same package with a coupling medium of infrared-transmitting glass or a silicone rubber material. Three major packages are widely used:

a. Transistor can packages

b. Ceramic DIP packages.

The construction of each is described below.

5.5.3.1 Transistor package. One effective construction approach is shown in Figure 99. The phototransistor and the LED are separated by a piece of glass, which serves as a dielectric and light pipe. The close proximity of about 5 mils of the two active elements allows a high CTR, about 200 percent, and an isolation voltage of about 1.5 KV.

Exact positioning of the elements or of the case is not required because the phototransistor is designed to have a base area larger than the LED. This gives a bigger target for the LED during bonding and makes its placement less critical.

The package is assembled by attaching the phototransistor to a transistor-type header. A thin layer of clear silicone resin is applied and then covered by a thin piece of glass. The partial assembly is baked for about an hour to cure the epoxy resin. Another thin layer of clear silicone epoxy resin is applied to the glass and the LED is positioned in it. The assembly, with LED, is again baked for curing. At this time, the wire bonding operation takes place. Once the wire bonding is completed, the entire assembly is coated with the clear resin and cured. The last step is to attach the cover to the header. The overall case outline and internal connections are shown in Figure 100.

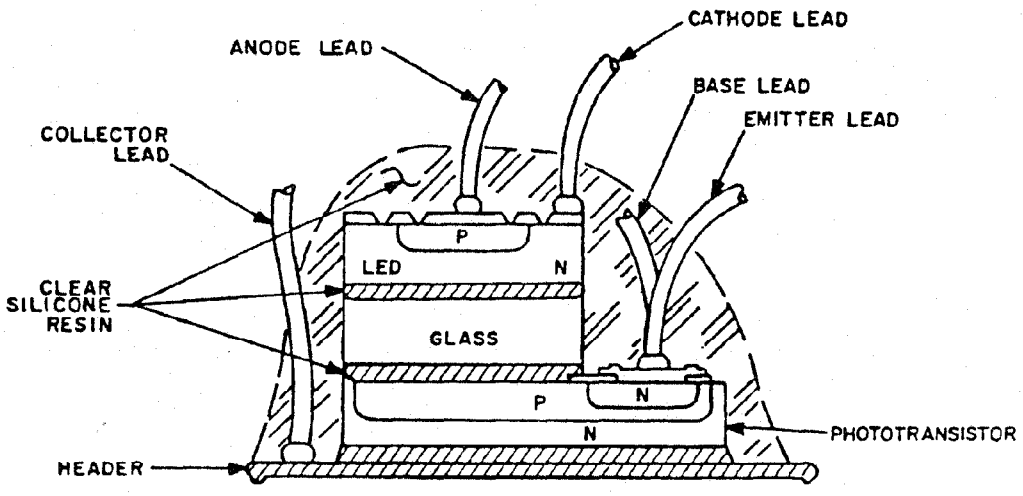


FIGURE 99. A typical phototransistor package.
SEE NOTE PAGE 5-118

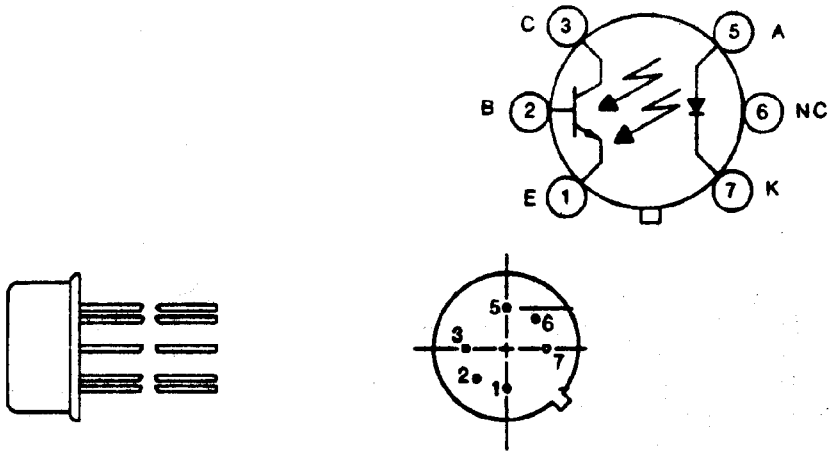


FIGURE 100. A typical transistor case outline.

5.5.3.2. High reliability ceramic package. Some of the high reliability optocouplers are packaged in a 16-pin ceramic package and hermetically sealed. Figure 101 shows the mechanical construction of a 16-pin hermetically sealed, optically coupled isolator.

Silicon detectors are die-attached and wire bonded inside a 16-pin ceramic package. GaAsp emitters (LEDs) are die-attached and wire bonded to a separate ceramic insert. Solder preforms are applied between the ceramic package and the insert and soldered in place. The assembly is then potted with transparent insulating material to improve the optical coupling and electrical insulation between the emitter and detector. Finally, a metal lid is attached to the package to insure a hermetic seal. The finished optically coupled isolator can withstand storage temperatures from -65 to +150 °C, operating temperature from -55 to +125 °C and 98 percent relative humidity at 65 °C without failure. Some manufacturers of optocouplers use an organic polymer (black-pigmented silicon rubber) to surround the light channels to isolate the channels in their multichannel devices to reduce cross talk. However, this imposes a limitation on constant acceleration testing to 5000 g, which might cause lateral movement of the pigment and reduce the current transfer ratio.

Note:

For those applications which are of a critical nature, opto devices containing opaque or transparent conformal materials should not be used. These coatings may have inherent problems when exposed to radiation and temperature extremes, such as outgassing and physical and/or chemical changes. Only hermetically sealed devices should be used.

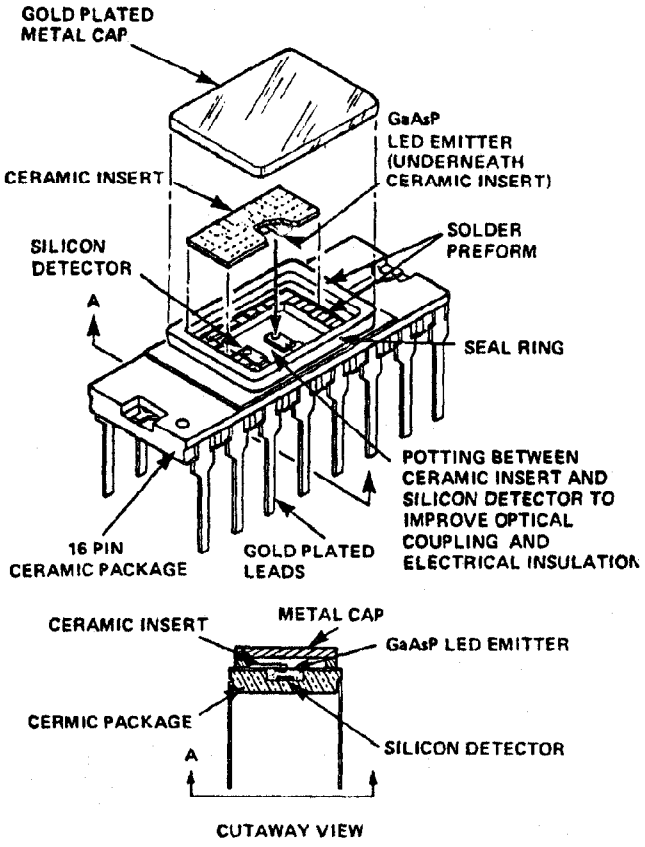


FIGURE 101. Mechanical construction of an optically coupled isolator in a hermetically sealed 16-pin ceramic package.

5.5.4 Military designation. The military designation for optocouplers is formulated as follows:

JANS	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

5.5 TRANSISTORS, OPTOCOUPLERS

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is 4N and 6N for optocouplers.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix M), longer or shorter polarity (suffix L or S), or any other letter to indicate a modified version.

5.5.5 Electrical characteristics. One of the most widely used general purpose optically coupled isolators uses a gallium arsenide light-emitting diode as the input and a silicon phototransistor as the output. It has a useful current transfer ratio (CTR) and is reasonably fast. Photo-Darlington outputs have higher CTR, but are slower. These two types are useful, but the collector-emitter breakdown voltage is limited to 25 to 50 V. By using a photo silicon-controlled rectifier, reverse voltages up to 400 V may be obtained. The speed of the photo SCR is between that of a phototransistor and a photo-Darlington. The fastest optocouplers use photodiodes in conjunction with a transistor amplifier or a logic gate to supply the output of the optically coupled isolator. Photodiodes are not used alone because they have very low CTR of about 0.1 percent.

Since the LED/phototransistor optocouplers are most widely used, these devices will be discussed in detail and a comparison of optocouplers using different output devices will be provided. The LED/phototransistor optocoupler consists of a GaAs LED and a silicon phototransistor.

5.5.5.1 LED characteristics. For most applications, the basic LED parameters I_F and V_F are sufficient to define the input. Figure 102 shows these forward characteristics, providing the necessary information to design an LED drive circuit.

Another important LED characteristic is the luminous intensity or the power output (P_O). At very low values of I_F, no radiative current predominates and, hence, very little or no light is emitted. As I_F increases, the radiative current begins to flow and increases faster than the non-radiative portion. This relationship allows a fairly linear relationship between I_F and luminous intensity for a range of forward currents. Eventually, the nonradiative current rate increases faster and the luminous intensity levels off as the radiative current virtually stops increasing.

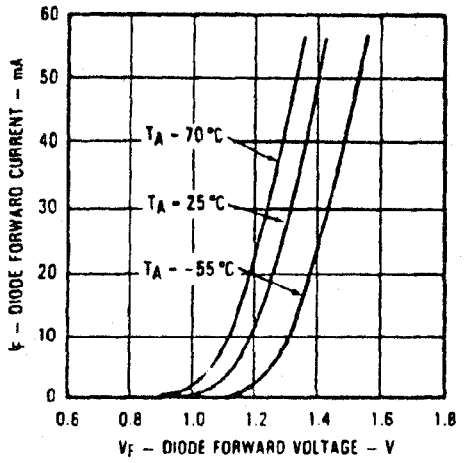


FIGURE 102. Diode forward current vs forward voltage.

Another important factor to consider is that the output of an LED decreases with time. Lifetime is often expressed as the time required for the output to fall to 50 percent of its original value. The typical lifetime may be 10,000 to 100,000 hours of operation. The LED degradation mechanisms will be discussed in paragraph 5.5.7, Reliability considerations.

Typical electrical characteristics of LEDs used in optically coupled isolators are forward voltage (V_F) at I_F = 10 mAdc, ranging from 0.8 to 1.5 V, and maximum reverse leakage current of about 100 µA at V_R = 2.0 Vdc).

Although the specifications do not call out any optical characteristics, the LED and the photodetector must have matching spectral characteristics to optimize the operation of the optically coupled isolators. Typically, when forward current (I_F) is passed through a GaAs LED, the optocoupler emits infrared radiation that peaks at a wavelength around 900 nanometers. The sensitivity of the photodetector should peak at the same wavelength for maximum signal coupling. The detector material (which is silicon) also peaks at this particular wavelength.

5.5.5.2 Phototransistor characteristics. The output of the optically coupled isolator is typically a phototransistor. The radiant energy from the LED falls on the base surface of the phototransistor. Phototransistors are designed to have a very large base region and, therefore, a very large collector base area with a small emitter. The incident energy, in the form of photons, creates electron-hole pairs in the base region and causes current to flow in the circuit.

5.5 TRANSISTORS, OPTOCOUPLERS

In most common test circuits the base is left open, the emitter is grounded, and a positive voltage is applied to the collector. However, the base lead is left available for unique circuit applications.

Typical phototransistor characteristics are listed in Table IX.

5.5.5.3 Coupled characteristics. Because the output stage of an optically coupled isolator is a phototransistor, the on-state parameters to be specified are $V_{CE(sat)}$, $I_{C(on)}$, and $I_{B(on)}$. Additional parameters that fully characterize the optically coupled isolator are current transfer ratio (CTR), isolation voltage, isolation resistance, and isolation capacitance. These additional characteristics are listed in Table VII.

TABLE VII. Typical isolator characteristics

Device Type	Current Transfer Ratio (CTR) (%)	Isolation Voltage V_{iso} (V)	Isolation Resistance (Ohms)	19500 Slash Sheet
4N22	25	1000	10"	486
4N23	60	1000	10"	486
4N24	100	1000	10"	486
4N47	50	1000	10"	548
4N48	100	1000	10"	548
4N49	200	1000	10"	548

NOTE: This table is not to be used for part selection; instead, use MIL-STD-975.

5.5.5.4 Switching characteristics. Switching characteristics are important for circuit applications and are usually specified by rise time (t_r) and fall time (t_f).

TABLE VIII. Switching characteristics

Device Type	Maximum Rise Time t_r (µs)	Maximum Fall time t_f (µs)	Comments
4N22,4N23,4N24	20	20	
4N47,4N48,4N49	25	25	Phototransistor mode
4N47,4N48,4N49	3	3	Photodiode mode

5.5.5.5 Optically coupled isolator characteristics. The current transfer ratio (CTR) and isolation voltage are the characteristics unique to an optically coupled isolator. The characteristics of 4N22, 4N23, 4N24, 4N47, and 4N48 are discussed in paragraph 5.5.5.3. All these devices are included in the qualified parts list of MIL-S-19500/486 and /548.

However, various types of optocouplers that meet special application requirements are available. Comparisons of the characteristics of various isolators are listed in Table IX.

TABLE IX. Comparison of different optically coupled isolators

Output Device 1/	CTR	Speed	C_{ISO}, R_{ISO}	Voltage Isolation	Package
Photo-transistor	10% to 100% min	1 - 10μs	1 - 3 pF; 10 - 100 GΩ	1 - 5 kV	DIP, metal can
Photo-Darlington	100% to 500% min	50 - 200μs	1 - 3 pF; 10 - 1000 GΩ	1 - 5 kV	DIP, metal can
Photo SCR	Approx. 20 mA trigger current	2 - 20μs	1 - 3 pF; 10 - 1000 GΩ	1 - 2.5 kV	DIP
Photodiode and transistor amplifier	7% - 400% min	0.5 - 60μs	--	1 - 2.5 kV	DIP
Photodiode and logic gate	400% - 600% min	t_{plh}, t_{phl} 50 - 100 ns	--	1 - 2.5 kV	DIP

1/ All the devices use LEDs for the input.

5.5.6 Environmental considerations. Typical environmental conditions and screens that the optically coupled isolators are capable of withstanding are not substantially different from those given in paragraph 5.1.6.3 of Screening procedure in subsection 5.1, Transistors, general. However, caution should be observed when selecting test conditions for constant acceleration and vibration, so that the dielectric resin material is not affected.

5.5 TRANSISTORS, OPTOCOUPLERS

5.5.7 Reliability considerations.

5.5.7.1 Failure modes. The optically coupled isolator differs from standard semiconductor devices in that it contains a GaAs/GaAsp LED chip and a silicon npn phototransistor chip, coupled with a light transmission medium. The major failure modes are CTR (current transfer ratio) degradation and the isolation voltage breakdown. The CTR degradation occurs due to light-emitting diode and phototransistor degradation. The isolation voltage breakdown is related to the light transmission medium and package-related defects.

5.5.7.2 Failure mechanisms.

5.5.7.2.1 Light-emitting diode (optocoupler input). Most current transfer degradation is due the reduction in the efficiency of the light-emitting diode within the optocoupler. The LED current consists of two components, a diffusion current and a space-charge recombination current.

$$I_F(V_F) = A_e^{qV_F/KT} + B_e^{qV_F/2KT}$$

where

A is the diffusion current component
B is the recombination current component
q is the electron charge
K is the Boltzman constant
T is temperature in K
V_F is the forward voltage

The diffusion current component is the radiative or light-emitting current and the recombination current is nonradiative. As the LED ages during use due to an increase in the value of B, the recombination current increases, then the radiative current will decrease for a fixed total LED current, causing a reduction in the light output of the LED. The specific reasons for the increase in the recombination current are not fully understood.

Dark-line defects are the dominant physical degradation mechanism of a light-emitting diode. Material or crystal imperfections act as growth sites for dislocation lines under forward bias conditions. The dislocation lines appear as dark regions giving rise to the term "dark line defects." As the dark line grows the light output progressively decreases.

Increasing the device temperature or operating current also accelerates the degradation rate. The rate depends on the device structure, assembly of the LED chip into the package and package thermal resistance. Impurities in the chip due to process contamination, exposed junction and metallization processes can also cause LED degradation.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

A typical CTR degradation with time is represented in Figure 103.

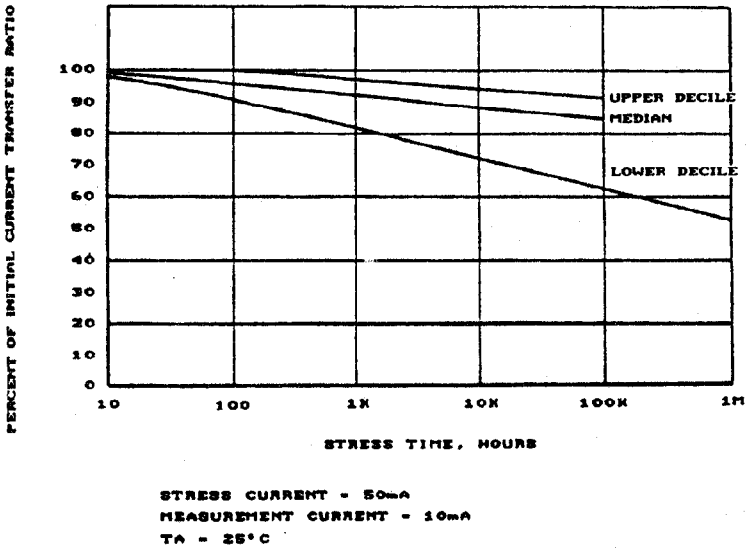


FIGURE 103. Current transfer ratio vs time.

In summary, the LED degradation depends on the following factors:

a. Quality of starting materials (wafer with few crystal dislocations and processed by the liquid phase epitaxial method).

b. Clean processing

c. LED structure or design

d. Metallization process

e. Operating temperature

f. Operating current.

5.5.7.2.2 Detector degradation. Although the detector has less influence than the light emitting diode, detector stability is very important. In a photo-transistor the failure modes are h_{FE} instability, increase in off state leakage current, and drop in breakdown voltages. The leakage instability is caused by the formation of the surface inversion channel in the base region of the

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

5.5 TRANSISTORS, OPTOCOUPLERS

phototransistor. The effect of this surface inversion can be remedied by subjecting these devices to a thermal anneal without bias. The surface inversion channel is considered to originate from ionic contamination in the silicon polymer material that covers the phototransistor. The large base area of the phototransistor is very sensitive to such contamination. The leakage current (I_{CEO}) on a phototransistor increases rapidly when subjected to high temperature reverse bias (collector-base), when such contamination is present.

Suitable design and processing of the phototransistor with appropriate passivation can be used to produce a stable and reliable detector.

5.5.7.2.3 Mechanical failure mechanisms. The failure modes are generally opens or shorts. Assembly-related failure mechanisms are:

a. Broken bond wire at dielectric interface

b. Lifted bond wire off the chip, due to a thermal expansion mismatch between the chip and the resin covering it

c. Improper bonding; bond wire shorting to the edge of the chip or package lead

d. Contamination carried with the package by the clear resin on the light pipe between the emitter and the detector. Optically coupled isolators are especially susceptible to this contamination. The effects of such contamination were previously discussed under detector degradation.

e. Dielectric instability. Optically coupled isolators transmit signals from input to output, while maintaining a high degree of isolation. Human safety and equipment protection are often critically dependent upon dielectric stability of the insulator under severe field conditions.

5.5.7.3 Derating. The largest single source of failure of these devices was proved to be operating above allowable levels of thermal and electrical stress. Accordingly, it is imperative that derating of parts be invoked to enhance the reliability of military systems. Users should refer to MIL-STD-975 for derating factor guidelines. The life of the LED in the optically coupled isolator can be extended by pulsed operation at low drive currents by a factor equal to the inverse of the duty cycle. High temperature operation is the most destructive stress a semiconductor can encounter. It will result in early end of life, electrical parameter drift, and general degradation of the device's electrical and mechanical characteristics.

6. MICROWAVE DEVICES

6.1 General.

6.1.1 Introduction. This section contains information on the various types of microwave components used in microwave communications and radar. Microwave components can be divided into three categories: active, passive, and hardware. The components range from dielectrically tuned transistor oscillators to hybrid power dividers and rectangular waveguides. The wide range of parts make general descriptions difficult, but there are some parameters and concepts that apply to all microwave devices. General microwave diode and transistor electrical parameters and reliability considerations are covered by the diode and transistor general sections. Microwave parts do not appear in NASA standard parts lists with the exception of some rf transistors and Schottky diodes. This constraint requires designers of high-reliability microwave systems to choose parts that meet all system requirements and are manufactured using high quality tests and controls. In most cases, this will require additional tests, process controls, and selected qualification tests because the products do not come from qualified suppliers. Some components are not mentioned in this section because they cannot be adequately treated in a handbook of this type. Examples are: assemblies such as antennas; advanced components such as HEMTs, GaAs MMICs, millimeter wave components, and subsystems; complex tubes, antenna feeds, and electromechanical switches.

Applicable military specifications are listed below.

Mil Spec	Title
MIL-A-3933	Attenuators, Fixed, General Specification for
MIL-A-22641	Adapters, Coaxial to Waveguide, General Specification for
MIL-A-24215	Attenuators, Variable (Coaxial and Waveguide), General Specification for
MIL-C-3928	Switches, (Coaxial), Radio Frequency Transmission Line, General Specification for
MIL-C-15370	Couplers, Directional (Coaxial Line or Waveguide), General Specification for
MIL-C-28790	Circulators, Radio Frequency, General Specification for
MIL-D-3954	Dummy Load, Electrical, Waveguide
MIL-D-39030	Dummy Loads, Electrical, Coaxial and Stripline

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.1 MICROWAVE DEVICES, GENERAL

Mil Spec	Title
MIL-F-3922	Flanges, Waveguide, General Purpose, General Specification for
MIL-F-39000	Flanges, Waveguide, Ridge, General Specification for
MIL-I-28791	Isolators, Radio Frequency, General Specification for
MIL-L-3890	Lines, Radio Frequency Transmission (Coaxial, Air Dielectric)
MIL-M-28837	Mixers, Stages, Radio Frequency, General Specification for
MIL-P-23971	Power Dividers/Power Combiners, and Power Divider/Combiners, General Specification for
MIL-S-55041	Switches, Waveguide, General Specification for
MIL-W-23068	Waveguides, Rigid, Circular
MIL-W-23351	Waveguides, Single Ridge and Double Ridge, General Specification for
MIL-STD-1352	Attenuators, Fixed and Variable, Selection of
MIL-STD-1358	Waveguides, Rectangular, Ridge and Circular, Selection of
MIL-STD-1637	Dummy Loads, Electrical, Waveguide, Coaxial, and Stripline, Selection of

6.1.2 General definitions. This list is presented as an aid for the interpretation and understanding of the specific microwave sections.

Attenuation. The reduction in amplitude or power of an electromagnetic wave as it propagates down a lossy transmission line.

Characteristics impedance, Z. The square root of the ratio of the series impedance to the shunt admittance of the equivalent circuit of a transmission line.

Insertion loss. A measure of the amount of loss a signal will suffer as it propagates through a component. In general, the measured value contains both the reflected loss and transmission loss.

Isolation. A measure of the reduction of power from the power at the input that will be expected at an isolated port of a component. In the case of a switch, isolation is the difference (in decibels) of power at the load with the switch off as compared to the switch on.

Reflection coefficient, Γ or ρ. The ratio of the voltages of the reflected signal to the incident signal when a signal is applied to the impedance terminating a transmission line. In general, the terminating impedance will not match the transmission line impedance and there will be reflected energy.

Voltage standing wave ratio (VSWR). The ratio of the maximum voltage to the minimum voltage of the standing wave produced by a mismatched termination on a transmission line.

6.1.3 NASA standard parts. Microwave parts do not appear in MIL-STD-975 with the exception of two Schottky diodes and several rf transistors.

6.1.4 General device characteristics. Device characteristics, such as materials, processes, package, and electrical characteristics, are covered in the individual microwave component subsections.

6.1.5 General parameter information. Details of electrical parameters for microwave diodes and transistors appear in this section as well as in the diode and transistor sections of this handbook. Details of electrical parameters for the remaining microwave components are presented in the section for the particular device. General electrical parameters include the following.

a. VSWR. VSWR is usually calculated from the return loss of a device. The return loss is the relation between the power returning along the line from a mismatched load to the power incident to a load. Thus, a return loss of 10 dB means that the reverse power traveling in a line from a mismatched load is 10 dB below the reference power incident on that mismatched load. For example, the Standing Wave Ratio (SWR) with a certain mismatched load is 2.0, which is equal to a return loss of about 9.5 dB (i.e., the reflected power is 9.5 dB below the incident power).

   That quantity for the same SWR can be confirmed as follows: The voltage vector of the backward-flowing power is 0.333 (if incident voltage were 1.0). Power returning down the line is then $(0.333)^2$ or 0.111. Note that 0.111 is approximately -9.5 dB. Mathematically, return loss is equal to $-20 \log_{10} \rho$ (where ρ is the voltage reflection coefficient), and VSWR is $\frac{1+|\rho|}{1-|\rho|}$

6.1 MICROWAVE DEVICES, GENERAL

b. Insertion loss. Expressed in decibels is 10 log $\frac{P_{out}}{P_{in}}$

Careful calibration of a test circuit can reduce the contribution of reflection loss to this number. The result will be the approximate transmission loss of a line with finite attenuation.

c. Isolation. Isolation is calculated the same as insertion loss in the case where the isolation of the output port from the input port is desired. The isolation of one output port from another output port is the difference in decibels of the power out of one of the first ports from the second port with power into the input. If more than one signal is present, the isolation of one particular signal power from the total power output may be calculated.

6.1.6 General reliability considerations. The general reliability considerations for transistors and diodes also apply to microwave transistors and diodes. Unique considerations for specific types are presented in this section. Many passive parts dissipate very little power and may be used at full rating over the full temperature of operation. Exceptions are loads, attenuators, and ferrite devices. Active components, such as switches, attenuators, and amplifiers, should be used in applications with proper consideration for derating and reliability.

Three major factors contribute to microwave component reliability.

a. Good basic device design and good mechanical design of packages and transmission medium

b. A good manufacturing process

c. Quality and reliability.

Only when all three factors are optimized will component reliability be at its maximum.

6.2 MICROWAVE DEVICES, DIODES

6.2 Microwave diodes.

6.2.1 Introduction. The only microwave diode that is included in MIL-STD-975 is one type of Schottky diode; however, microwave diodes are included in this handbook to provide a technical understanding.

Microwave diodes can be divided into three general groups which are determined by their general application. These applications are generation, detection, and control of microwave energy. Even though microwave devices utilized in these applications are similar in construction to other diodes, microwave diodes are sufficiently different to warrant a special discussion of their radio frequency (rf) properties.

Discussion of microwave diodes is facilitated by a review of the terms describing device characteristics. Following are several important parameters.

Conversion loss. The ratio of the available rf input power to the available intermediate frequency (IF) output power under the specified conditions.

Figure of merit. The measure of excellence of a video crystal in a video receiver.

Impedance at intermediate frequency (IF). The impedance presented at the output terminals of the mixer when the device is driven by the local oscillator under the specified conditions.

Load impedance. The input impedance of the diode load circuit at the converted frequency of the signal used for the measurement of conversion loss or overall noise figure.

Minority carrier lifetime. The measure of the time it takes to switch a microwave diode from totally on to totally off.

Mixer ratio frequency impedance. The impedance measured at the local oscillator terminals of a mixer under specified conditions.

Negative resistance. The dynamic slope of the current versus voltage (IV) curve when the microwave diode is tested in a suitable high frequency fixture.

Output noise ratio or noise temperature ratio. The ratio of the available noise-power output of the diode at the IF when driven by a local oscillator under the specified test conditions, to that of a resistor at standard temperature (293 ±5 K).

Overall noise figure. The ratio of the available signal-to-noise power ratio at the input to the signal-to-noise power ratio at the output of a network. This value is usually expressed in decibels.

6.2 MICROWAVE DEVICES, DIODES

Series resistance. A measure of the sum of the loss elements in the internal structure of the microwave diode such as lead and contact resistance as well as bulk semiconductor resistance.

Tangential sensitivity. Generally referred to as the signal power below a 1 mW reference level necessary to produce an output pulse whose amplitude is high enough to raise the noise change by an amount equal to the average noise level.

Threshold voltage. The minimum dc voltage applied to a Gunn diode to achieve threshold current. In a typical case, V_{TH} = 0.33 V operating.

Threshold current. The maximum dc current applied to a Gunn diode. Typically, I_{TH} = 1.5 I operating.

Transition time. A measure of the time for a microwave diode to switch from a reverse conduction state to a totally off state.

Video impedance. The impedance at the specified frequency presented at the output terminals of a semiconductor diode under the specified conditions.

6.2.2 Usual applications. As noted previously, microwave diodes can be divided into three general groups as determined by their application. The first group provides generation of microwave power. The rf frequency multiplication is best accomplished through use of the varactor device, whereas the bulk effect, avalanche, and tunnel diodes best provide power generation for amplifier applications. Noise power and oscillator power generation both use the avalanche diode. The second general group provides microwave power control. Attenuator, limiter, and switching applications are effectively controlled by the PIN diode, whereas tuning is best controlled by the varactor diode. Phase shifting may incorporate either type. The third general group covers the receiving or detecting of microwave power. The function of detection and mixing in receiving rf energy is best accomplished by the Schottky barrier diode, point contact, tunnel diode, or back diode. The rectification of microwave power is accomplished by use of the Schottky barrier diode.

6.2.2.1 Generation of rf power. Generation of rf power by frequency multiplication is often accomplished with the step recovery diode (SRD). The SRD is a type of varactor that provides efficient and versatile frequency multiplier performance. Its advantage is its charge storage and recovery characteristics. The SRD stores charges while biased in the forward direction by either part of a sinusoidal wave or by a steady bias current. It is turned off by either the negative portion of the sinusoid or by a negative pulse. This device is designed to conduct for a period of time in the reverse direction as minority carriers stored near the junction are depleted. When the stored minority carriers are depleted, an abrupt step in current that is rich in high-order harmonics occurs.

For discrete frequency multiplication the voltage pulse should be resonated (e.g., in a quarter wavelength transmission line) to generate an exponentially

decaying sine wave. To generate a comb spectrum, the voltage pulse is terminated in a resistive load (see Figure 1).

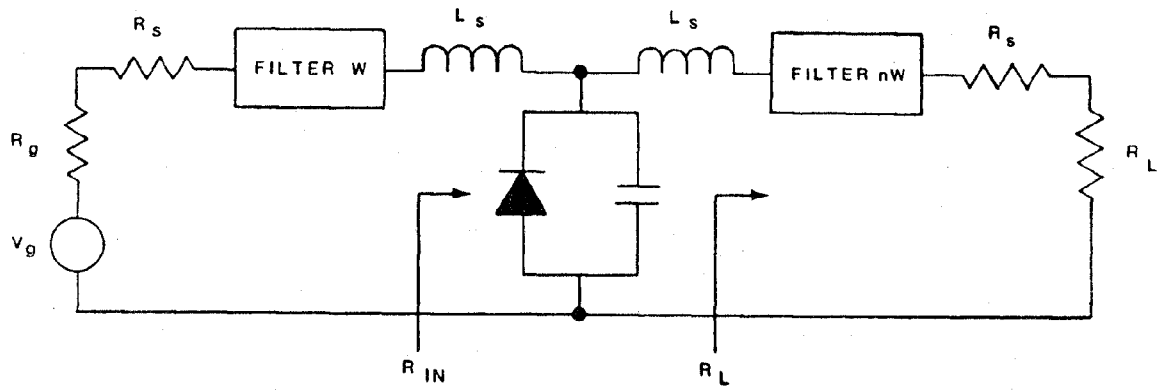


FIGURE 1. SRD multiplier circuit.

The ideal storage diode should have the following properties:

a. Infinite capacitance in the forward direction (zero incremental voltage drop)

b. Infinitely fast switching from reverse storage conduction to cutoff

c. Zero capacitance in reverse direction

d. Long minority carrier lifetime

e. No excessive parasitic elements; a small amount of series resistance is tolerable and a small amount of series inductance may be desirable to reduce the resistance-capacitance transition time.

A second application of the varactor for rf power generation is in parametric amplifiers and up converters in which the varactor becomes the active element that produces a negative resistance. The term parametric is derived from the achievement of amplification, oscillation, and up conversion or harmonic multiplication by choosing a lossless parameter in the system. Parametric amplifiers, when cooled to cryogenic temperatures, offer low-noise performance for low-noise receiver applications.

In the reflection parametric amplifier circuit a high-power, high-frequency pump signal is reactively mixed with the incoming low-frequency signal. The resulting lower sideband is supported in a resonant circuit whereas the upper sideband

6.2 MICROWAVE DEVICES, DIODES

signal is terminated in a high impedance. The lower sideband frequency mixes again with the pump signal providing an additional component to the low-frequency signal. This double reactive mixing process introduces a 180-degree phase shift between the incoming and outgoing signal frequency which results in the appearance of a negative resistance. The power is drawn from the pump source (see Figure 2).

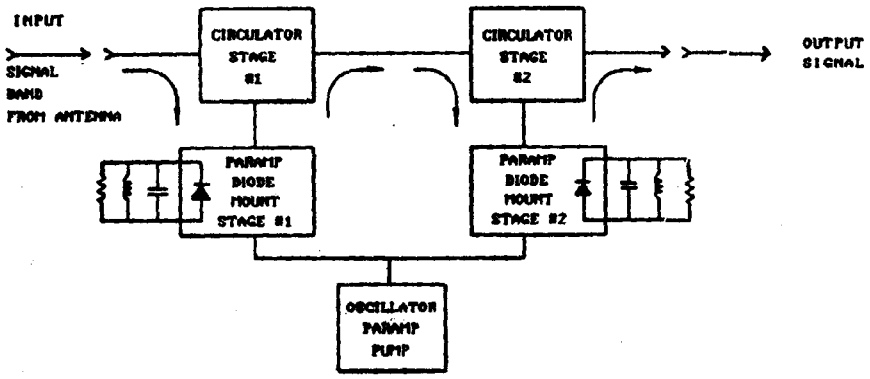


FIGURE 2. Reflection parametric amplifier.

Bulk-effect microwave devices (Gunn diodes) differ considerably from conventional pn junctions in that there is no discrete junction area. The bulk semiconductor material is usually gallium arsenide whereas conventional diodes are silicon. The n-type bulk material, when dc-biased, creates a charged region that travels from the negative terminal through the bulk material. In a microwave resonant circuit this charged region produces microwave power generation in oscillator or amplifier applications.

The major microwave applications of the Gunn device have been in low-power oscillators in the 6 to 20 GHz range with average output power to about 200 mW. Gunn devices have been used in applications exceeding 90 GHz. This device has about a 10-dB less noise figure and requires a lower bias voltage than the impact ionization avalanche time (IMPATT) diode. However, the Gunn device generates less power than the IMPATT.

The peak power output of a Gunn diode occurs at the transit time frequency which is inversely proportional to the length of the device and directly proportional to the electron drift velocity. This is in turn, proportional to the operating voltage. The power output of a Gunn oscillator increases rapidly beyond a threshold voltage until optimum transit time is obtained, then decreases with further

increased voltage. If possible, the diode should operate slightly below the power maximum (1 V) and never over it. Diodes become noisy when operated above the power-peak voltage.

Thus, the important characteristics of the device are the IV threshold voltage, turn-on voltage, peak power voltage, and turn-off voltage. Figure 3 shows the characteristic shape of the diode current versus voltage and microwave power versus voltage.

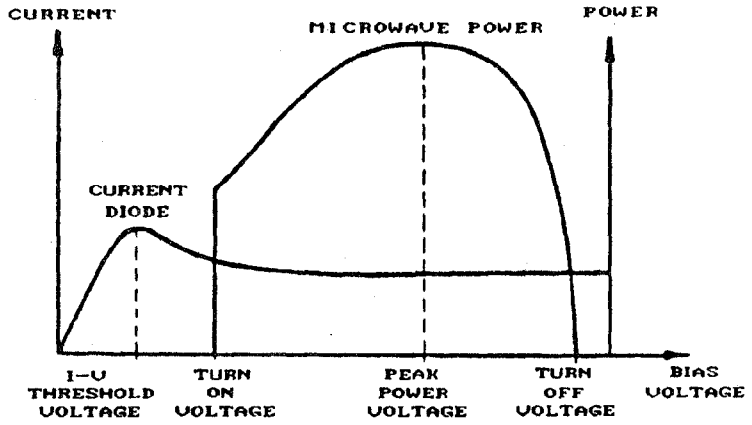


FIGURE 3. Gunn oscillator IV characteristic curve and power-voltage curve.

The frequency over which the device exhibits negative resistance will determine whether or not the device will oscillate in a given cavity. A reduced height cavity is often used to combat low temperature turn-on problems Figure 4 shows a Gunn oscillator circuit with varactor diode tuning.

A Read diode or IMPATT diode amplifier has the ability to produce more microwave power than any other semiconductor device at frequencies as high as 300 GHz. The IMPATT diode utilizes conventional pn, PIN, or Schottky junctions under external, reverse bias. The diode couples energy into a microwave field by providing charge carriers from avalanche breakdown of an n^+p region at one end of the diode to a drift region at precisely the time the applied ac field is starting to decrease. The clump of holes obtained from the avalanche region

6.2 MICROWAVE DEVICES, DIODES

begin traveling through the diode intrinsic, or drift, region under the force of the dc field. At this point the holes are traveling against the direction of the ac field; therefore, they give up energy to the ac field. Traveling through the drift region opposite the applied ac voltage produces a current with the opposite polarity. This implies a negative resistance has been developed. The drift region is made just long enough for the charge clumps to reach the end of the drift region as the ac voltage again starts going positive.

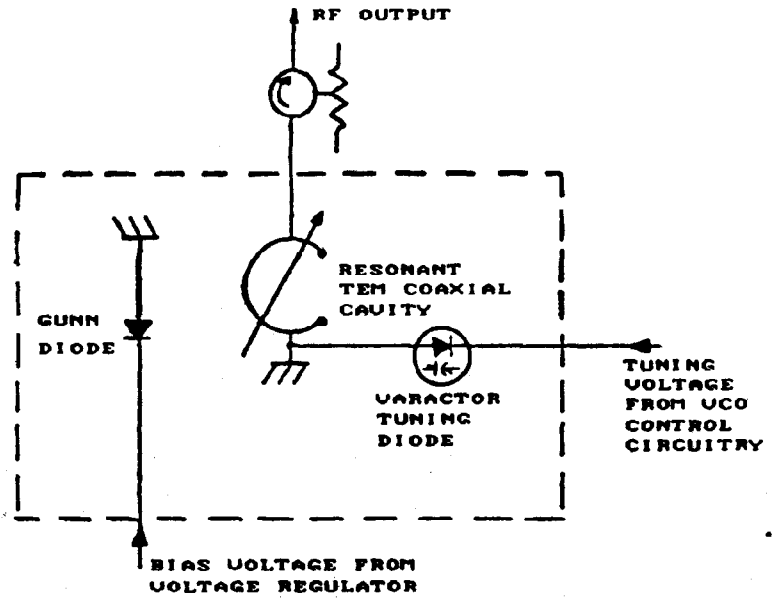


FIGURE 4. Gunn oscillator circuit with varactor diode tuning.

IMPATT diodes are often combined in such a fashion that the total output power is approximately equal to the sum of the powers of the individual diodes. Combiners can be 2-way or N-way. The 2-way combiners, such as rat-race couplers, hybrid couplers, or 2-way Wilkinson combiners, can be combined to form an N-way combiner where $N = 2n$ and n is an integer. This method is not efficient for $N > 4$. The N-way combiner combines diodes in a single structure. This structure may be either a nonresonant or a resonant type. In the nonresonant type, the power sources are isolated from one another so power combining can only occur if it is excited by an external signal. In the resonant type, the power sources are coupled together so the circuit can be used as an oscillator without the use of an external locking signal. Figure 5 shows one type of resonant combiner oscillator.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

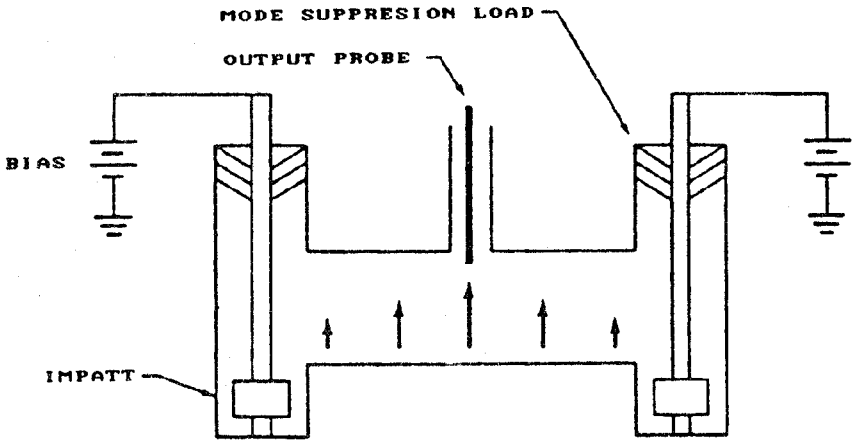


FIGURE 5. IMPATT combiner oscillator.

The trapped plasma avalanche triggered transit (TRAPATT) diode provides microwave energy from approximately 0.5 to 4 GHz, with efficiency approaching 50 percent. However, the TRAPATT diode requires a complicated circuit and gives a higher noise level than the IMPATT diode.

Tunnel diodes and back diodes have normal diode junctions except the pn-doping level is much higher than conventional devices. Electrically, tunnel diodes provide negative resistance at microwave frequencies and are useful as low-noise amplifiers due to their inherent low-noise figures. Back diodes differ from tunnel diodes in that they allow little current flow when the diode is reverse biased. Tunnel and back diodes can be used in detector and mixer applications.

6.2.2.2 Microwave power control. PIN diodes utilize an undoped semiconductor layer referred to as an intrinsic layer, located between the p- and n-doped layers. This intrinsic region provides low-junction capacitance that has little or no change due to reverse bias. Under forward biasing conditions, the intrinsic layer becomes a conductor. When the PIN device is placed in a microwave transmission application, the device has the effect of a shunt attenuator. In limiter applications, the device acts as a lossy switch.

The switching action results from variations of the I-layer resistance with applied dc signal. At zero or reverse bias, R_i is high and the diode acts as a fairly high capacitance at microwave frequencies. In the forward biased state, the I-layer resistance R_i is lowered due to conductivity modulation.

Conductivity modulation can be produced either by sufficiently high-level microwave power or by dc bias.

6.2 MICROWAVE DEVICES, DIODES

The breakdown voltage is approximately proportional to the I-layer width. The intrinsic resistance is proportional to the square of the I-layer width and inversely proportional to the minority carrier lifetime and dc bias current $R_{RF} = \frac{W^2}{\tau \cdot I_{DC}}$ (as shown in Figure 6).

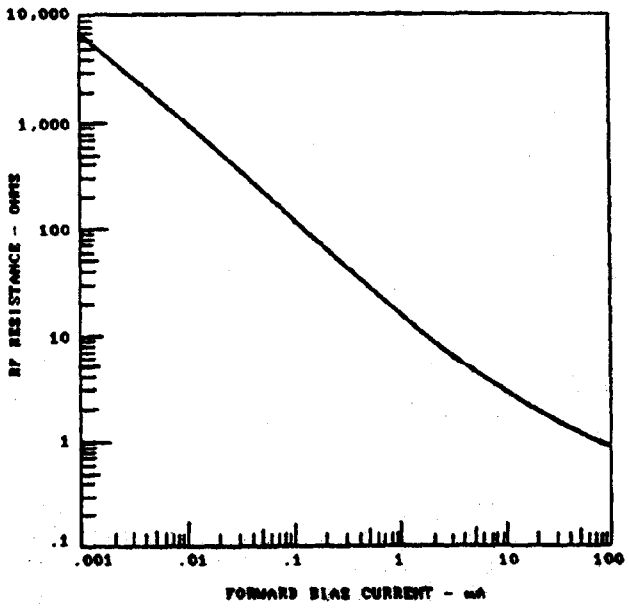


FIGURE 6. Typical PIN diode rf resistance vs forward bias current.

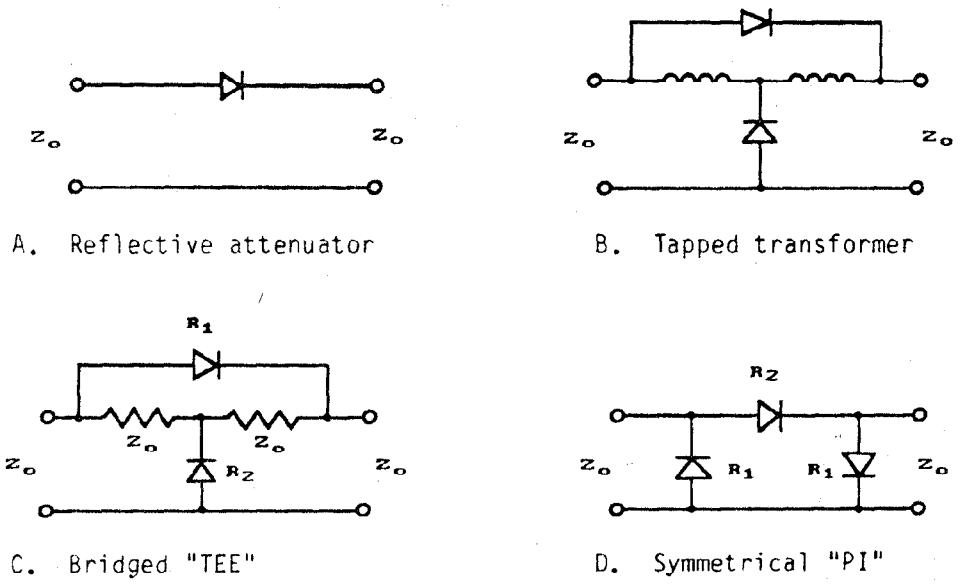
The PIN diode may be used in series or shunt configurations. The key electrical parameters are the insertion loss, isolation, and bandwidth of the particular application.

Four simple attenuator circuits are shown in Figure 7. The forward biasing is not shown. The first attenuator is the simplest and cheapest implementation. The second circuit has the highest isolation whereas the "TEE" and "Pi" circuits have the broadest bandwidth. A diode in shunt (not shown) could be implemented in waveguide. If n shunt diodes are spaced at quarter wavelength intervals, the overall attenuation can be increased by more than n times that of a single diode.

A series PIN switch and a shunt PIN switch are shown in Figures 8 and 9 respectively. When used as a switch, the residual attenuation that exists in the PIN diode when the switch is on is called insertion loss. The attenuation provided

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

when the switch is off is called isolation. If the diode is assumed to be a pure resistance at rf, the attenuation for each circuit is a function of the ratio of the circuit resistance to the diode resistance. As the bias on the diode is varied the load resistance as seen by the source also varies. Reflection from the mismatch is the primary source of attenuation, but some of the energy is also transmitted or dissipated in the PIN diode.



A. Reflective attenuator

B. Tapped transformer

C. Bridged "TEE"

D. Symmetrical "PI"

FIGURE 7. PIN diode attenuator circuits.

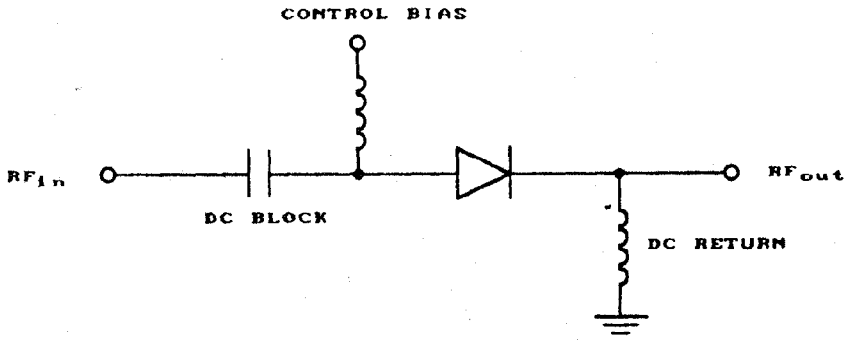


FIGURE 8. Series PIN reflective switch.

6.2 MICROWAVE DEVICES, DIODES

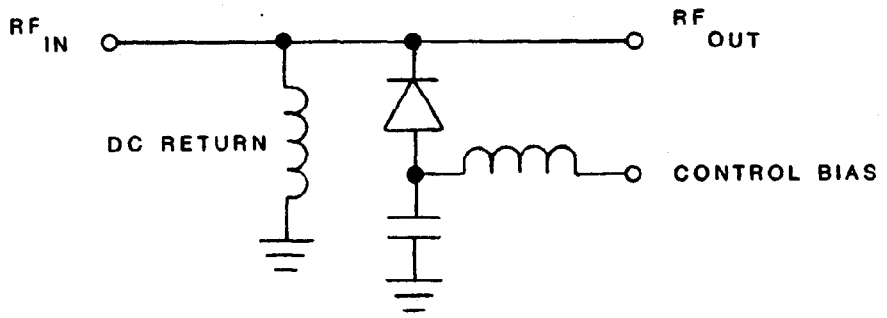


FIGURE 9. Shunt PIN reflective switch.

A packaged part, or even a wire-bonded chip component, has stray parasitic elements that must be tuned out at microwave frequencies by the addition of external reactances. Such circuits are generally referred to as resonant switches in which the PIN diode is used essentially to switch the circuit parameters from a parallel resonant condition to a series resonant condition. The high and low impedances produced by the parallel and series resonant circuits, respectively, constitute the on and off states of the switch.

A basic limiter utilizes the decrease in PIN diode resistance when incident rf power increases to reflect some of that power and maintain a nearly constant output power. The limiting action only comes into play when the input power reaches a threshold value. After that, the output power will remain constant for increasing input power until the diode resistance is reduced to its smallest value. The output will begin to increase if any additional power is input.

Figure 10 shows the output power versus the input power for a shunt-mounted 50 ohm limiter diode.

The power rating, breakdown voltage, intrinsic resistance, and speed of the diode determine the limiter characteristics. The turn-on characteristics of the PIN diode in the limiter can be improved by using a detector diode to add dc current to the PIN diode junction at the onset of rf power.

Varactor diodes have a modified pn junction constructed to maximize the decrease in junction capacitance with an increase in reverse voltage. Consequently, they are used as the voltage-controlled, tunable impedance in voltage-controlled oscillators (VCOs) and phase shifters.

Figure 11 shows an equivalent circuit of a VCO. The active device and cavity could be the Gunn diode oscillator discussed previously, a field effect transistor (FET), or crystal oscillator. Important parameters are capacitance versus voltage, capacitance ratio (a measure of the tuning range), and the Q (a measure of the efficiency).

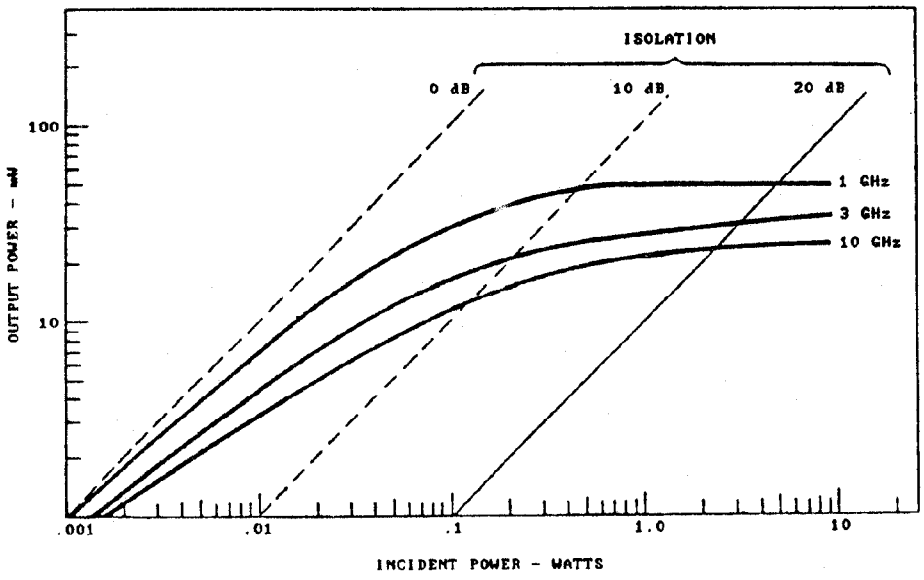


FIGURE 10. Output power vs input power for a 50-ohm limiter diode.

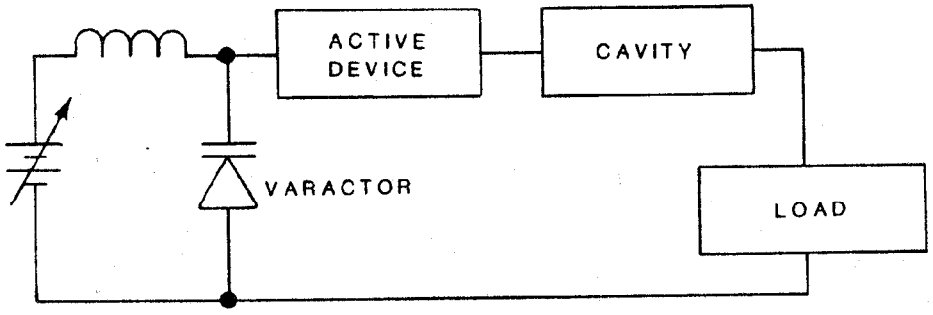


FIGURE 11. Equivalent circuit of a VCO.

6.2 MICROWAVE DEVICES, DIODES

6.2.2.3 Microwave detection. Schottky-barrier diodes differ from the normal silicon pn construction in that a metal layer deposited on the silicon forms the junction. As a result of this special layer, the device has a small junction capacitance and a nonlinear voltage-current relationship. This type of diode is useful as a microwave detector because of its low noise properties and efficient rectifying properties.

The point-contact diode is similar to the Schottky barrier diode except that the anode contact is accomplished by using a tungsten whisker to form a pressure contact to the silicon. A more detailed discussion of this oldest of the microwave semiconductors is provided in the last portion of this section.

Tunnel diodes have the lowest output resistance and lowest 1f noise characteristics due to the high doping levels of the back diode semiconductor wafer. This results in a faster video rise time and larger bandwidth. The tunneling region of the IV characteristic, where the detector operates under small signal conditions, is relatively independent of temperature. Figure 12 shows a zero-biased detector circuit.

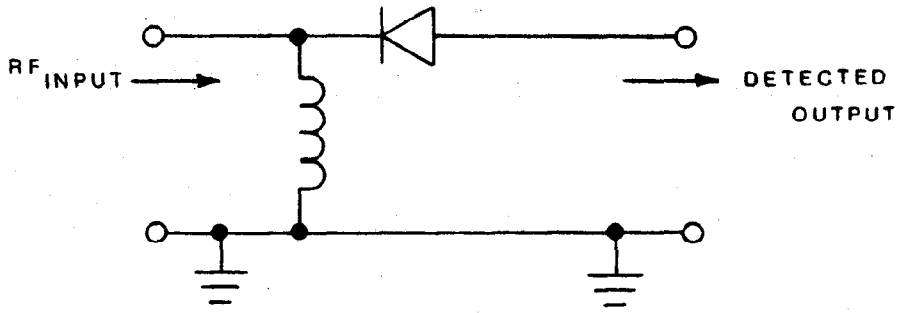


FIGURE 12. Detector circuit.

Schottky barrier diodes are more reliable when operated with a low impedance load. High-resistive loads are acceptable if the diode is in the square law region. A 50Ω load is best for fast pulse shape preservation. The input match and R-C time constant of the input circuitry will also affect the shape of a fast changing pulse. High-power detector loads should be less than 1,000Ω with 300 to 500Ω giving maximum power transfer.

6.2.3 Physical construction. To accomplish the various microwave functions previously discussed, microwave diodes are constructed in a different manner from

the normal pn junction and associated packaging. The following brief outline of the comparative internal junction construction (Figure 13) is presented as an aid to the discussion of electrical characteristics in paragraph 6.2.5 of this section. Details on beam lead devices are given in section 3.1, Transistors, general.

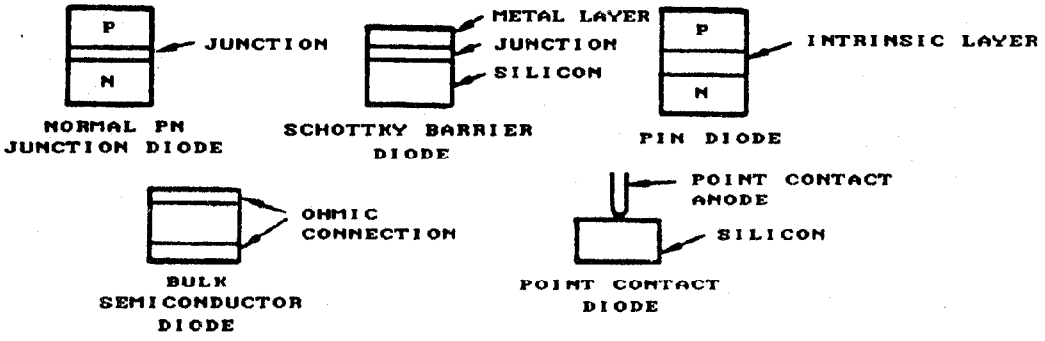


FIGURE 13. Microwave diode construction.

External packaging of microwave diodes varies according to the application. Figure 14 illustrates a few of the commonly used types.

6.2.4 Military designation. The military designation for microwave diodes is formulated as follows:

JANTXV	XN	YYYY	ZZ
Prefix	Component designation	Identification number	Suffix letter(s)

The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is IN for microwave diodes.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describe matched devices (suffix M), reverse polarity (suffix R), or any other letter to indicate a modified version.

6.2 MICROWAVE DEVICES, DIODES

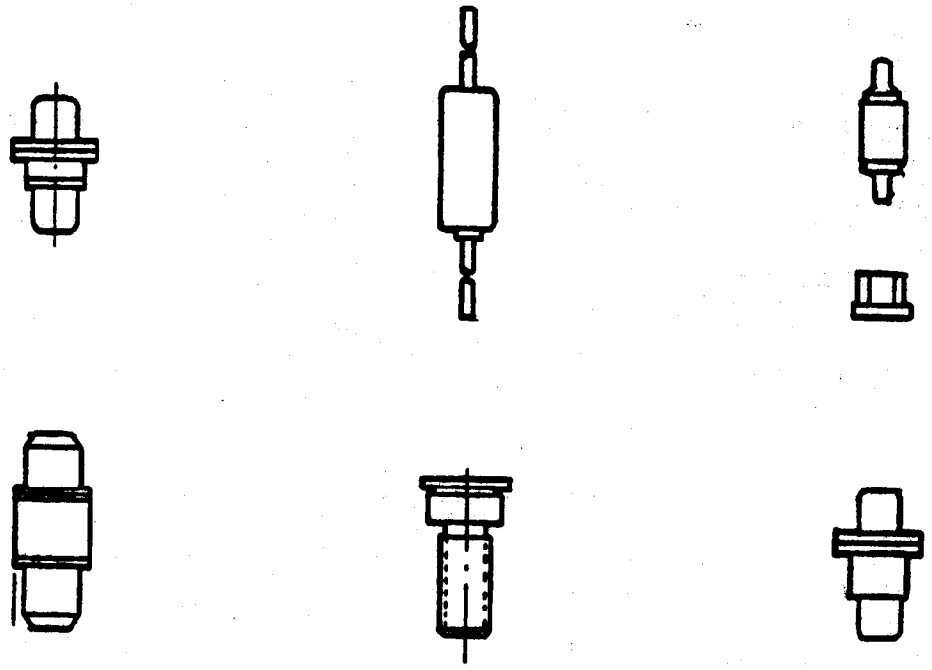


FIGURE 14. Microwave diode packages.

6.2.5 Electrical characteristics. Table I presents some typical detector and mixer diode electrical characteristics for JAN-type rf and microwave diodes. An examination of this oldest class of microwave diodes points out their applications.

6.2.5.1 Detection and mixing based on the dc EI curve. The curve for an ideal diode is shown in Figure 15. In this instance the forward current is infinite for the smallest conceivable amount of forward voltage. The reverse current is zero for any value of reverse voltage. The ideal diode would make a perfect detector or mixer. As a low-level detector operating at close to zero volts, the extreme nonlinearity of the characteristic curve over the range of some small plus-to-minus voltage would lead to perfect or lossless detection. As a mixer the dV/dI at some operating point above zero volts would once again result in perfect or lossless mixing. Because conditions are seldom ideal, consideration will be given to more practical cases.

6.2 MICROWAVE DEVICES, DIODES

TABLE I. Electrical ratings

Device Type No.	Test Freq. (MHz)	Operate Freq. (MHz)	Noise Figure (dB)	Conversion Loss (dB)	Output Noise Ratio	VSWR Max.	IF Impedance (ohms) Min.	IF Impedance (ohms) Max.	Burnout (Ergs)
1N830A	100	100	---						
1N25	1000	---	---	---	---	---	---	---	375
1N82A	1000	---	14.5	8	2.5	1.3	100	400	---
1N21WE	3060	---	7.0	---	---	---	---	---	---
1N21WEM	3060	---	7.0	5.5	1.5	1.3	350	450	5
1N21WEMR	3060	---	7.0	5.5	1.5	1.3	350	450	5
1N32	3295	1550-	---	5.5	1.5	1.3	350	450	5
		3060		---	---	---	---	---	375
1N31	9375	5200-	---						
		9375		---	---	---	---	---	175
1N23WE	9375	---	7.5						
1N23WEM	9375	---	7.5	6.0	1.4	1.3	335	465	2
1N23WEMR	9375	---	7.5	6.0	1.4	1.3	335	465	2
1N78	16000	---	---	6.0	1.4	1.3	335	465	2
1N78C	16000	---	9.5	7.5	2.5	---	325	625	1
1N26	23984	---	---	6.0	1.9	1.5	400	565	1
1N26B	23984	---	1.0	8.5	2.5	---	300	600	.1
1N53	34860	---	---	7.5	2.0	1.5	400	600	.1
				8.5	2.5	1.6	400	800	---

NOTE: These parts are not included in MIL-STD-975. This table is not intended to be a part selection list.

6.2 MICROWAVE DEVICES, DIODES

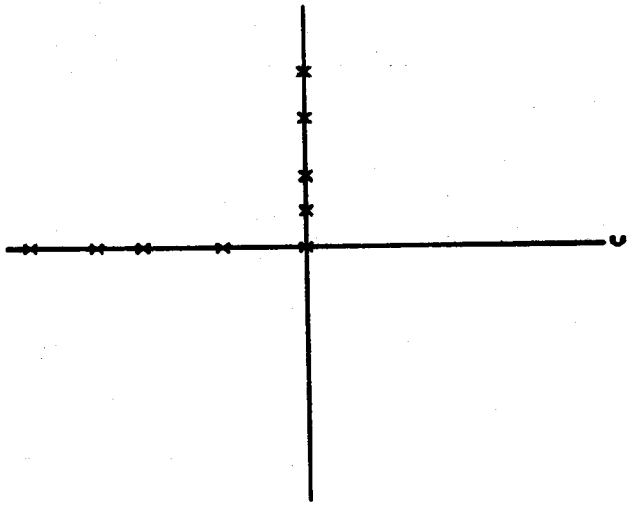


FIGURE 15. An ideal diode dc characteristic curve.

Figure 16 is a plot of the EI curve of the point-contact diode. Concerning detection, forward conduction starts at zero for the point-contact diode, and the ordinate, in this case, is 0-10 μA. This scale has been chosen because detection is being discussed. For the purpose of this discussion, the reverse current is assumed to be zero over the range of the signal input.

If a small rf voltage (40 mV P-P) is applied, it results in a detected output as shown in Figure 17.

As can be seen, the point-contact diode is detecting efficiently due to the fact that there is conduction over the positive half cycle of the rf input.

The mixer diode can be treated in somewhat the same manner. Mixers typically use Schottky barrier diodes or hot-carrier diodes. Figure 18 is a plot of the EI curve of a Schottky barrier diode with the ordinate, in this case, being from 0-10 mA. This scale has been chosen because the mixer diode is now being discussed and conduction in the range of milliamperes is important, not micro-amperes, as was the case for the detector. Since the local oscillator supplies the bias level, we see that the Schottky barrier diode will operate as an efficient mixer, for the purpose of this discussion, with a local oscillator bias of about 500 mV peak-to-peak and a signal of 40 mV peak-to-peak. It should be pointed out, however, that dc bias will also enable the Schottky barrier diode to operate at lower drive levels.

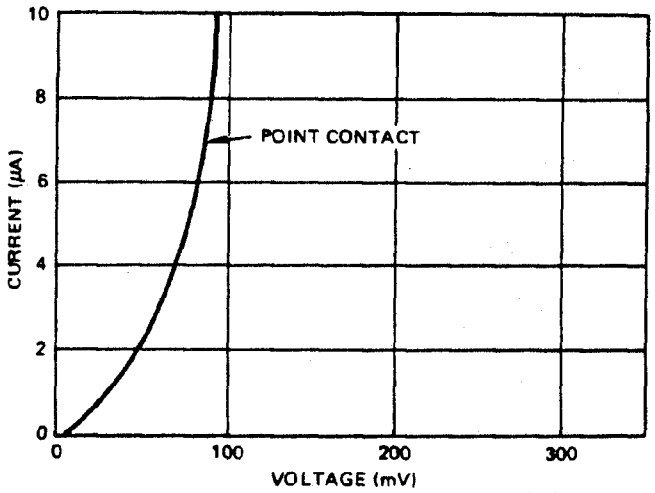


FIGURE 16. Forward direction dc characteristic EI curve.

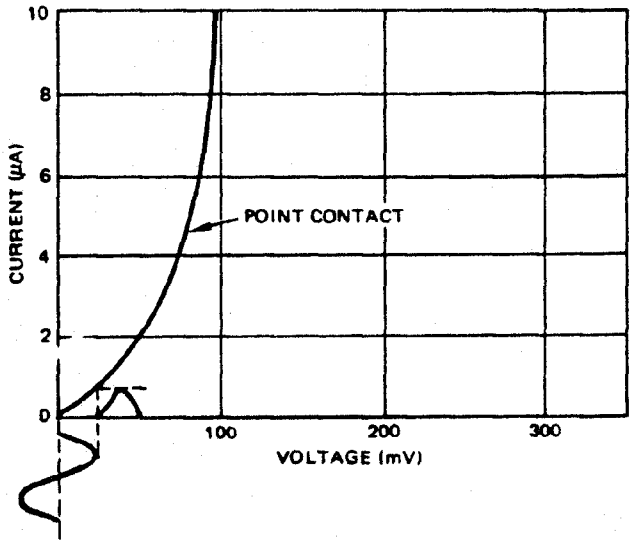


FIGURE 17. Video detection based on the dc characteristic curve without dc bias.

6.2 MICROWAVE DEVICES, DIODES

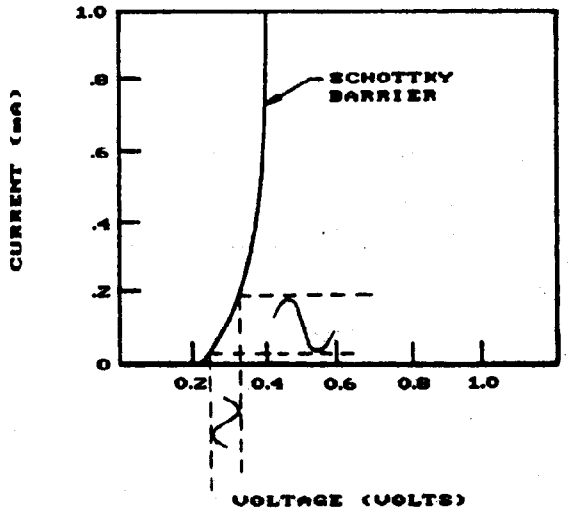


FIGURE 18. Mixing based on the dc characteristic curve.

Shown in the next several figures are the actual forward and reverse characteristics of point-contact diodes. Comparisons are made between typical S, X, and Ka-band types, and typical diode characteristics versus temperature.

Figure 19 shows the forward characteristic for the point contact diodes in the S-band, X-band, and Ka-band. The curves are similar.

Figure 20 is a plot of the reverse characteristic of the point contact diode. Point contact diodes have very low reverse voltage drops for any given reverse current. Although not generally specified, point contact diodes typically have 3-V peak-inverse voltage (piv) at 100 μA.

Figure 21 shows the forward characteristic of a point-contact diode versus temperature. At currents of 1 mA or more, the change versus temperature is much less than at currents of 100 mA or less. This is to be expected, because internal heating due to the I^2R drop (forward conduction) tends to swamp out any differences due to ambient temperature. This is not so at lower current values.

Figure 22 shows the reverse characteristics of the point-contact diode versus temperature. As can be seen, the change in piv at 150 °C is extreme, becoming one-third of its room temperature value for this diode.

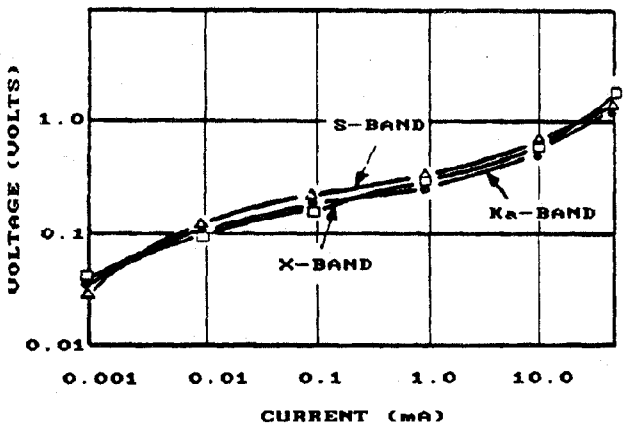


FIGURE 19. Typical forward dc characteristics at 25 °C for point-contact diodes.

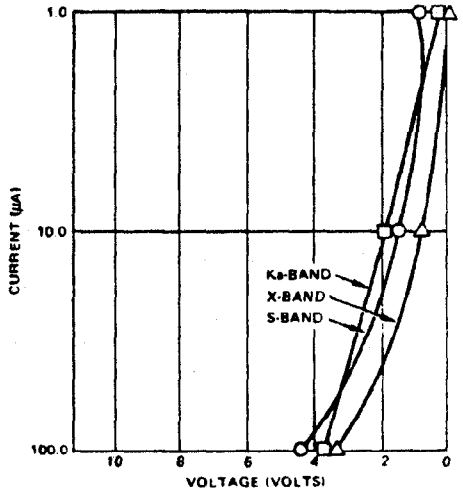


FIGURE 20. Typical reverse dc characteristics at 25 °C for point-contact diodes.

6.2 MICROWAVE DEVICES, DIODES

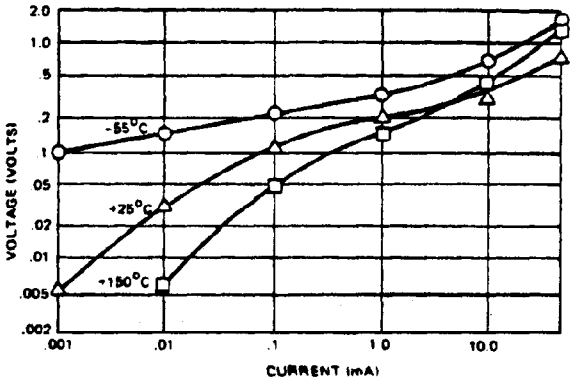


FIGURE 21. Typical forward dc characteristics vs temperature for point contact diodes.

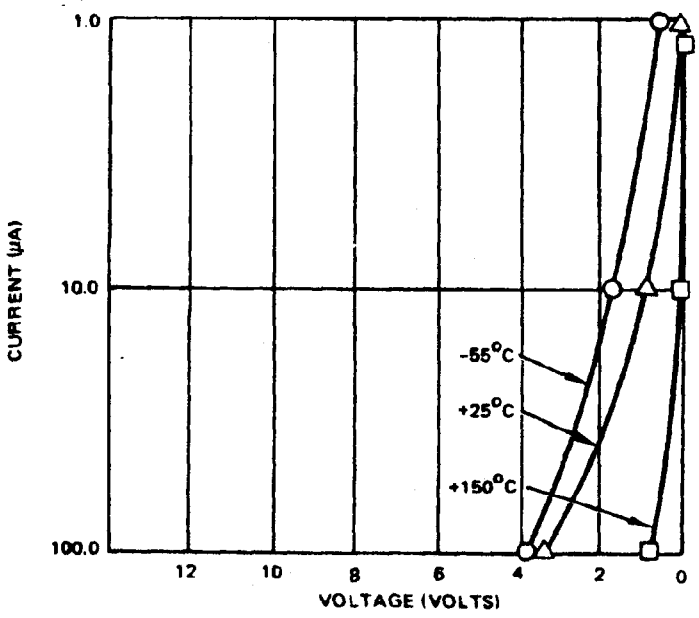


FIGURE 22. Typical reverse dc characteristics vs temperature for point contact diodes.

6.2 MICROWAVE DEVICES, DIODES

The hot-carrier diode or Schottky barrier diode has several advantages over the point-contact diode.

a. More rugged. The point-contact diode depends on a mechanical contact between a sharp metal whisker and a semiconductor material. Shock and vibration often change or destroy this contact.

b. Higher burnout. The uniformity and larger area of contact of the hot-carrier diode make it less susceptible to burnout from either cw power, peak power, or pulse energy. This is a common mode of failure for detectors and mixers.

c. More reproducible. Each hot-carrier diode is very nearly like every other hot-carrier diode in forward characteristics. This actual forward characteristic is closer to the ideal diode law characteristics than any other type of diode. Point-contact diodes, however, vary widely. See Figure 23 for a comparison of current voltage characteristics of a batch of point-contact diodes and a batch of hot-carrier diodes.

d. Lower series resistance. The lower series resistance of hot-carrier diodes leads to lower losses and better noise performance.

e. Higher off-to-on ratio. The ratio of reverse current (at a given reverse voltage) to forward current (at a given forward voltage) is much higher for the hot-carrier diode. This leads to better performance as a switch, rectifier, or detector.

f. Lower noise figure. If an average hot-carrier diode is compared with good selected-point contact the hot-carrier diode exhibits slightly better overall NF_o. However, compared with the average point-contact, the average hot-carrier diode is superior.

g. Low frequency (lf) noise. Figure 24 shows typical lf noise performance for a hot-carrier diode at several bias points and a good point-contact diode. The lf noise of the hot-carrier diode is significantly better.

6.2.6 Environmental considerations. Microwave diodes generally conform to the environmental specifications of conventional pn-junction diodes (see section 4.2, Diodes, switching). For more detailed requirements see applicable slash sheets of MIL-S-19500.

Typical environmental conditions and screens that microwave diodes are capable of withstanding generally conform to the requirements of MIL-S-19500 for JANTXV or JANS diodes.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.2 MICROWAVE DEVICES, DIODES

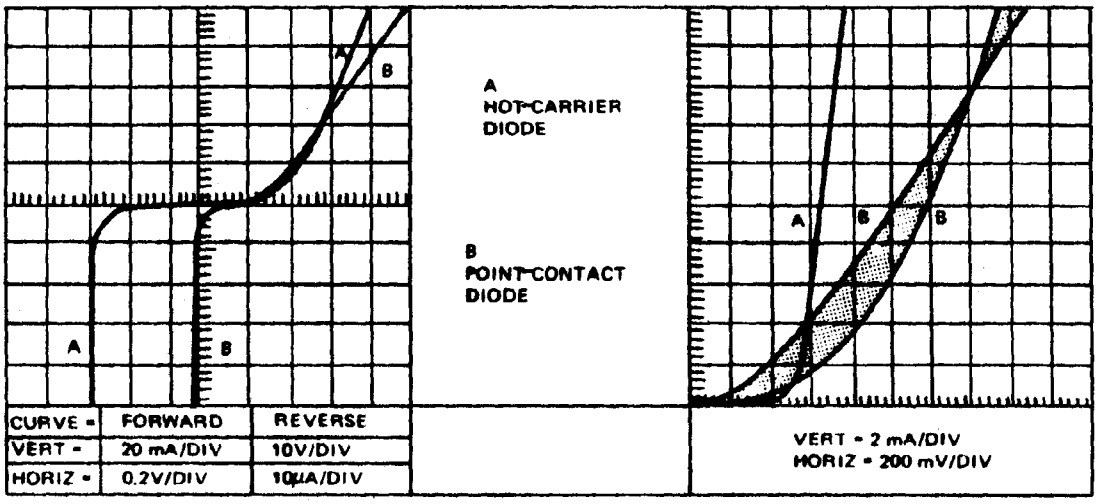


FIGURE 23. Current and voltage characteristics for hot-carrier and point-contact diodes.

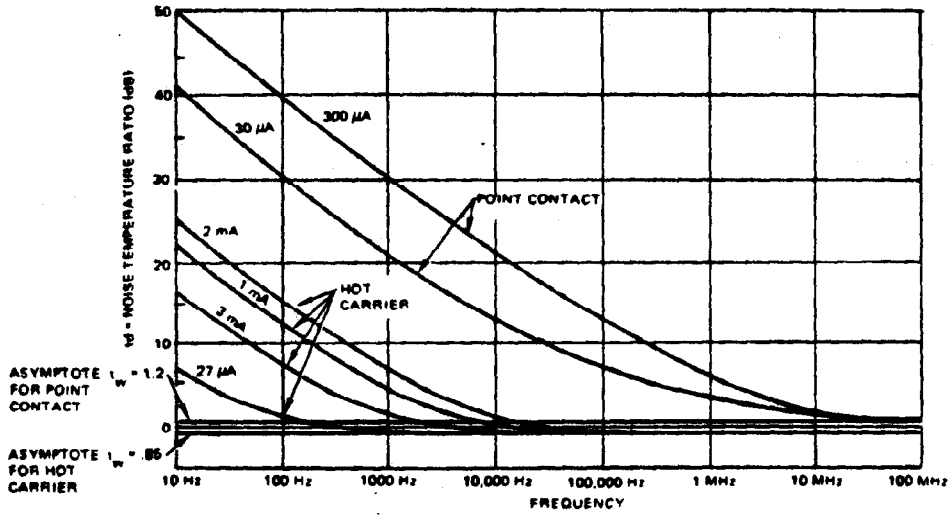


FIGURE 24. Hot-carrier and point-contact diodes--1f noise.

6.2.7 Reliability considerations.

6.2.7.1 Failure modes. The primary mode of failure for microwave diodes is burnout. Burnout is defined as a catastrophic failure, such as an open or short circuit, or a 3 dB degradation in electrical performance. Burnout may be due to excessive energy, either continuous or transient, which is dissipated in the diode. In some circuits where a diode acts as a detector over a wide dynamic range and maximum sensitivity is sought, the diode capability will be exceeded at the high-power levels. Transient energy is more difficult to protect against, and transient spike leakage power is damaging because the heat generated in the diode during the duration of the spike is not dissipated throughout the bulk material of the diode semiconductor. Proper circuit design and selection of correct device characteristics can eliminate the preceding failure modes.

Another reliability consideration in the use of microwave diodes is that the point contact can damage the junction, depending on the amount of pressure exerted on the contact or external shock-induced oscillation of the whisker.

6.2.7.2 Derating. This is the intentional reduction of electrical stresses on a device to reduce the operating temperature of the diode for the purpose of extending its life. History has shown that the largest single source of failure of diodes is due to operating above allowable levels of stress. Accordingly, it is imperative that derating of parts be invoked to enhance the reliability of military systems. Derating conditions specified in MIL-STD-975 for diodes apply to microwave diodes. In general, derating for voltage, current, and surge current should be 50 percent of the value published on the vendor data sheet. The junction temperature should be limited to 125 °C maximum. The design and use of a part should always be within the thermal and electrical stresses defined. High-temperature operation is the most destructive stress a semiconductor could be subjected to. High temperature will cause early destruction life, electrical parameter drift, and general degradation of the electrical and mechanical characteristics of the device.

6.2.7.3 Screening. Development of the JANTXV process for several microwave diodes has done much to reduce the incidence of the failures discussed in this section. A description of the JANTXV process is given in section 4.1, Diodes, general.

6.2.7.4 Static sensitivity. Schottky diodes and point-contact diodes are electrostatic-sensitive devices. The following procedures should be adhered to when handling high frequency devices:

a. The operator, as well as tweezers or any other pickup tool, must be grounded to the test or inspection station. This prevents the build-up of static charge, which can damage the diode if allowed to pass through it.

6.2 MICROWAVE DEVICES, DIODES

b. All test fixtures should be equipped with a short across the terminals that is disconnected after the diode is inserted.

For leakage measurements where a short across the terminals is not practical, a series resistor may be used. The series resistor (approximately 10 K) should be physically close to the test terminal.

This will prevent discharge through the diode of any charge built up on the capacitance that is present in the fixture, leads, and test equipment. It is also advisable to minimize this capacitance.

c. Spurious pulses generated by test equipment, contact bounce during switching, induced voltage in the leads, etc., must be eliminated.

d. When passing a diode from one operator to another, the receiving operator should grasp the lead held by the passing operator. If the opposite lead is taken, there is a possibility of passing static charge difference between the operators through the diode.

e. All soldering apparatus should be transformer isolated from the power line and should be free of leakage.

6.2.7.5 Radiation considerations. To insure that a circuit functions properly in space applications, the design engineer must consider the effects of radiation exposure. Requirements in military programs for systems that are survivable in a radiation environment adds to the challenge. The design engineer must know how radiation affects the circuit and components. When semiconductor devices are exposed to radiation environments, changes occur in their rated electrical parameters. The magnitude of the changes is a function of such things as the type of radiation, neutron or gamma ray, and time or duration. Generally, permanent damage is associated with displacement effects resulting from neutron irradiation, and transient effects are the result of ionization from gamma ray radiation.

Diodes exposed to radiation will display changes in the following electrical characteristics:

a. Breakdown voltage increases

b. Leakage current increases

c. Forward voltage increases.

These changes are a result of changes to the bulk material and are permanent.

Four radiation hardness assurance levels (RHA) are provided for diodes screened and tested to JANS requirements in accordance MIL-S-19500. These are designated by the letters M, D, R, and H.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 Microwave transistors.

6.3.1. Introduction. There are no microwave transistors in MIL-STD-975. They are included here for information only. However, a few rf transistors appear in MIL-STD-975.

The design of modern microwave amplifiers, oscillators, and occasionally multipliers and mixers incorporate a microwave transistor as the active, two-port element. The majority of these devices are either npn silicon bipolar or gallium arsenide metal electrode semiconductor FET (GaAs MESFET). A comparison of the relative gain, power, and noise figure is summarized in Table II. The GaAs MESFET has lower noise, higher gain, and higher power output than the silicon bipolar transistor. The MESFET can operate at higher frequencies than the bipolar transistor because the latter suffers from minority carrier storage effects. The MESFET is a majority carrier device and is not slowed down by this mechanism. The MESFET also has a higher input impedance and a negative temperature coefficient that prevents thermal runaway from occurring. The MESFET's main disadvantage is higher 1f noise, which can be significant for oscillator noise.

TABLE II. Comparison of microwave transistors

	Si Bipolar Transistors			GaAs MESFET			
	4 GHz	8 GHz	12 GHz	4 GHz	8 GHz	12 GHz	18 GHz
Gain typical (dB)	15	9	6	20	16	12	8
Power output (W)	6	2	0.25	25	8	4	1
NF min (dB)	2.5	4.5	8	1.0	1.8	2.2	2.5

The physical principles of microwave transistors are similar to those of low frequency transistors. Operation at rf and microwave frequencies puts more severe requirements on active area dimensions, process control, heat sinking, and packaging. The construction of most types is a planar, epitaxial process. Microwave transistors also include pnp bipolar, insulated gate (IG or MOS) FET and JFET; however, those types will not be covered in this section.

The first microwave transistors were fabricated using the interdigitated geometry. These were followed by the overlay metal matrix, sometimes called mesh or emitter grid, and the FET. Each of these technologies represents a distinct advancement of the state of the art. Also, each was developed with a certain end result in mind.

The following discussion explains the development, applications, advantages, and disadvantages of the four basic microwave transistor types: interdigitated, overlay, metal matrix, and GaAs FET. Also discussed are the refinements of each technology with respect to the materials, design optimization, and manufacturing methods.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS

6.3.1.1 Device geometries.

Interdigitated. Devices with interdigitated geometry are still in the vast majority. The primary reasons for this are relative ease of fabrication and the broad spectrum of applications. The surface topography is shown in Figure 25.

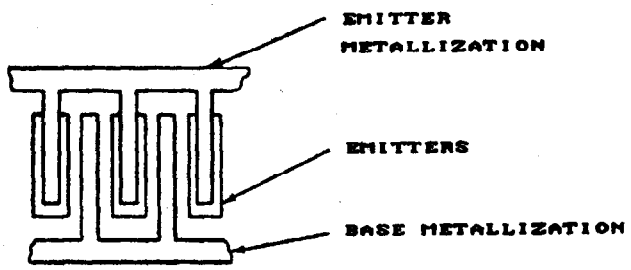


FIGURE 25. Surface geometry of an interdigitated transistor.

The greatest advantage of the interdigitated device is its inherently high ratio of emitter periphery to base area (P_e:A_b). This number is a figure of merit of transistors and is usually on the order of 3 to 5 cm^{-1}. The significance of this ratio is that it is an indicator of the device's gain-bandwidth product (f_t). Generally, the higher this ratio can be made, the higher the f_t of the device.

Through refinements of today's technology small signal interdigitated devices can achieve the following limits:

f_t = 8 GHz

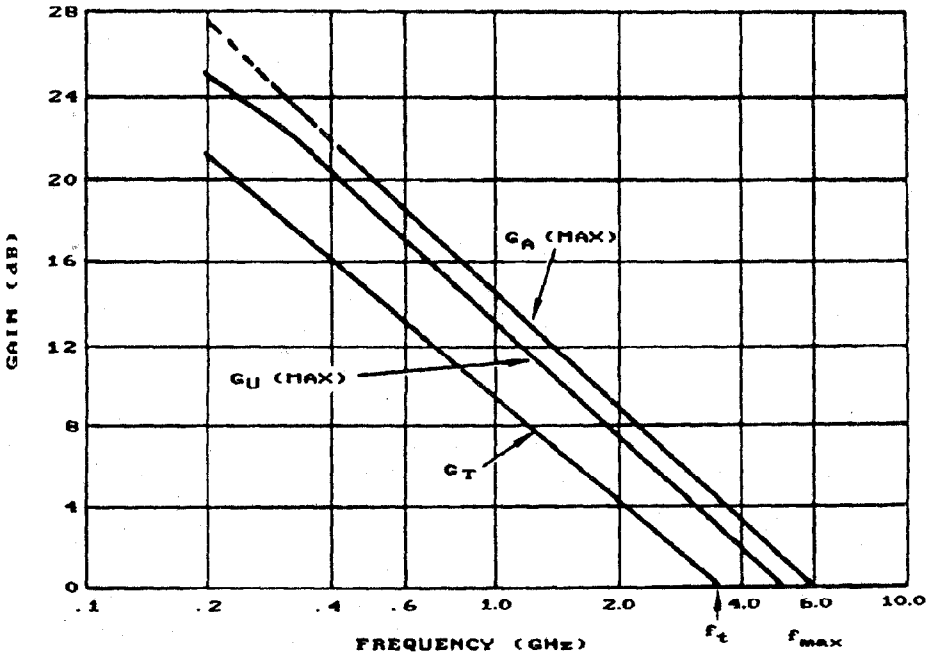
f_{max} = 25 GHz

The difference between f_t and f_{max} is as follows: f_t is the frequency where the transducer power gain $|S_{21}|^2$ (known as G_T) goes to 0 dB, while f_{max} is the frequency above which oscillation cannot occur (Figure 26).

These values were significantly improved by the development of different diffusion technologies. Both arsenic-diffused emitters and ion implantation will be discussed later in this section. Basically, the key to improving f_t lies in the ability to control the depth of the base diffusion, the final base width, and the uniformity of dopant concentration. These newer technologies both have significant advantages over the conventional phosphorus-doped emitter.

The noise figure is one of the most critical parameters of a small signal microwave amplifier. Key factors for keeping the noise figure to a minimum are minimizing the collector-base capacitance and careful consideration of

circuit design. Generally, for microwave devices, the noise figure is lower in the common emitter configuration than common base. Negative feedback through the collector-base capacitance tends to reduce the intrinsic noise figure. However, this effect will also reduce the forward gain G_T. Obviously, this trade-off must be considered in light of the application.



$G_{U(MAX)}$ = Unilateral gain
$G_{A(MAX)}$ = Max available gain

FIGURE 26. Definition of f_t and f_{max}.

A typical noise figure versus frequency and emitter current for an npn bipolar transistor in the common emitter configuration is shown in Figures 27 and 28, respectively.

6.3 MICROWAVE DEVICES, TRANSISTORS

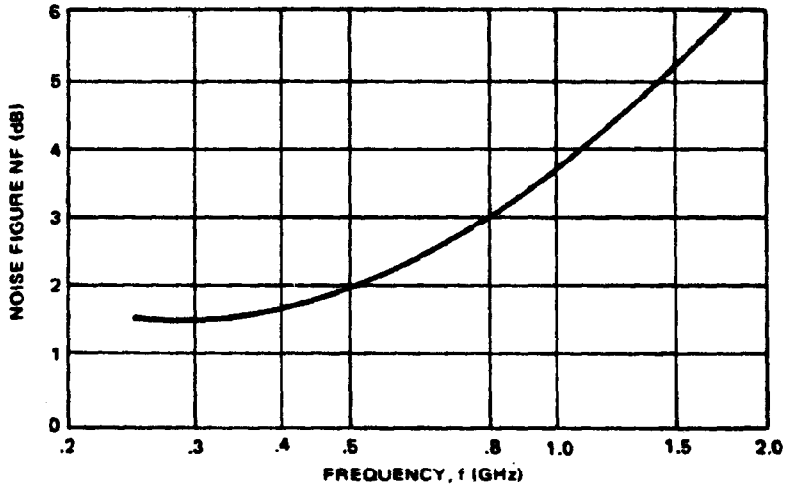


FIGURE 27. Noise figure vs frequency.

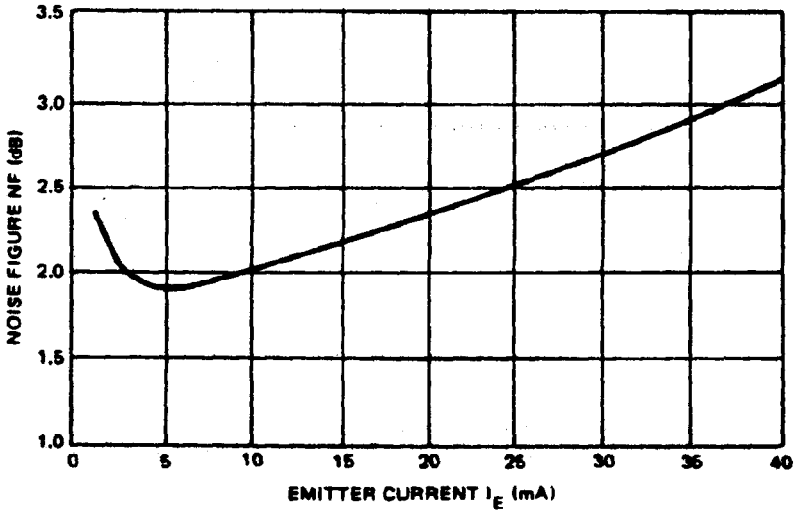


FIGURE 28. Noise figure vs emitter current.

Overlay. The intrinsic ability of the overlay device to withstand greater current densities at a time when aluminum was the only metal in use for contact metallizing gave the industry a significantly more reliable product than any other yet produced.

The overlay geometry is shown in Figure 29 and the construction is shown in Figure 30. The details of fabrication will be covered fully in the section on fabrication and assembly. However, several features merit mention now.

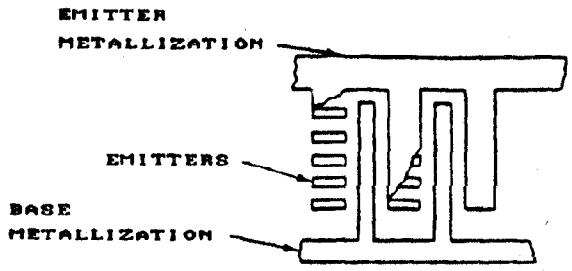


FIGURE 29. Overlay geometry.

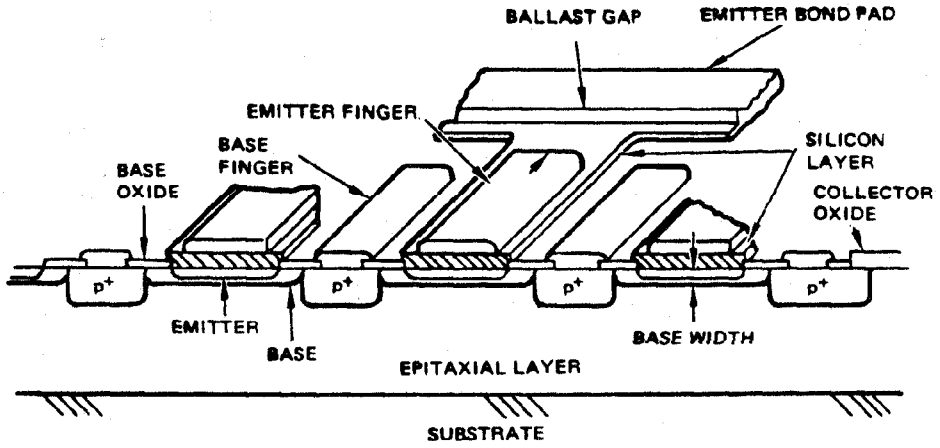


FIGURE 30. Typical construction of an overlay device.

The first is the layer of polycrystalline silicon under the emitter metallization. This technique was developed in an effort to hinder the spiking of aluminum through the device to the base and collector.

6.3 MICROWAVE DEVICES, TRANSISTORS

At first, the silicon was sputtered at high temperatures (>1000 °C). However, this caused further diffusion of the emitter. A method for depositing the silicon near room temperature was eventually discovered. This silicon barrier vastly improved the ruggedness of the overlay device.

The second feature was the development and use of ballast resistors on overlay devices. The purpose of the ballast resistor was to equalize the current in each emitter finger. (Note--in an overlay device, an emitter finger contacts many emitter sites as opposed to the fingers of an interdigitated device). Thus, more current would tend to flow in fingers closer to the bonding pad, causing potential overheating problems, whereas the ballast resistor design would evenly distribute the current before it could cause any thermal damage.

The material used for ballast resistors is most often nichrome, because it has higher resistivity than most contact metals. The formation of the ballast resistor is accomplished by depositing the nichrome in one long strip so that the fingers are at 90 degree angles to it. Doped silicon has also been used for this purpose.

The ballast technique is now used on practically every microwave power device regardless of geometry. Thermal effects in ballasted and unballasted devices will be discussed later in this section.

Metal matrix. The metal matrix geometry can best be described as a sibling to the overlay geometry. However, several distinct advantages are evident. This geometry affords the highest ratio of emitter periphery to base area of any other fabrication technology. Thus, the upper frequency range can be extended considerably. Power and current handling capability also remains high. The surface geometry of the metal matrix device is shown in Figure 31.

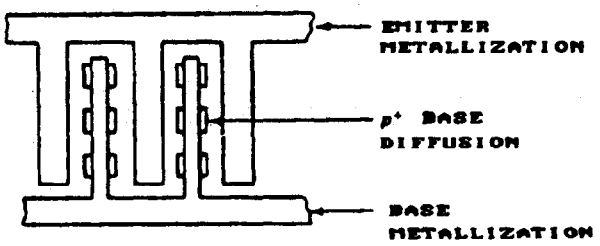


FIGURE 31. Metal matrix geometry.

The matrix structure achieves its high emitter periphery due to the fact that the emitter runs lengthwise across the chip as opposed to the interdigitated configuration.

The wide metal fingers enhance the current handling capability of the device. Even where aluminum metallization is being used, the increase in metallization cross sectional area significantly reduces the rate of aluminum migration. Where refractory metals are used, migration is virtually eliminated as a problem.

These advantages are somewhat offset by the extreme difficulty in processing the matrix transistor. The accuracy with which the regions are aligned is probably the most crucial factor, along with precise diffusion control.

6.3.1.2 GaAs Schottky barrier field effect transistor (FET). To date, the most successful microwave field effect transistors are the n-type GaAs metal semiconductor field effect transistors. Devices can be fabricated that meet the following specifications:

Unilateral power gain (G_U) > 25 dB at 1 GHz (See paragraph 6.3.5 for a definition of G_U.)

f_{max} > 100 GHz

Noise factor < 3dB at 4 GHz

Probably the greatest initial problem in the development of these devices was the difficulty in controlling the thickness of the epitaxial layer. Factors affecting this thickness include dopant concentration, source temperature, and substrate temperature. Experimentation with these variables resulted in the capability to reduce the growth rate drastically. However, other problems such as poor vapor etching arose at lower source temperatures.

The orientation of the substrate determined whether the epitaxy was <100> or <111>. Although the <111> structure has a more controllable growth rate, <111> contributed to low yields in the scribe and break operation. Therefore, the <100> structure is being used almost exclusively.

Performance of the GaAs FET is extremely dependent on geometry. The gain-bandwidth product, f_t, is effectively maximized by reducing the dimensions of the gate and gate-source/gate-drain spacing. Initial configurations included an "enclosed drain" mesa geometry and an interdigitated mesa geometry.

Dimensions associated with the gate were kept in the 2- to 4-micron region. Figure 32 shows f_{max} and G_U as a function of gate length for various values of pinch-off voltage, V_p. These predicted values, however, could not be obtained with the interdigitated geometry due to two major problems: 1) formation of parasitic transistors and 2) resistors in the fingers of the interdigitated geometry devices.

Redesign of the enclosed drain geometry was pursued. A more refined geometry evolved that essentially eliminated the more serious problems. A comparison of performance of the interdigitated and enclosed drain designs is shown in Figure 33.

6.3 MICROWAVE DEVICES, TRANSISTORS

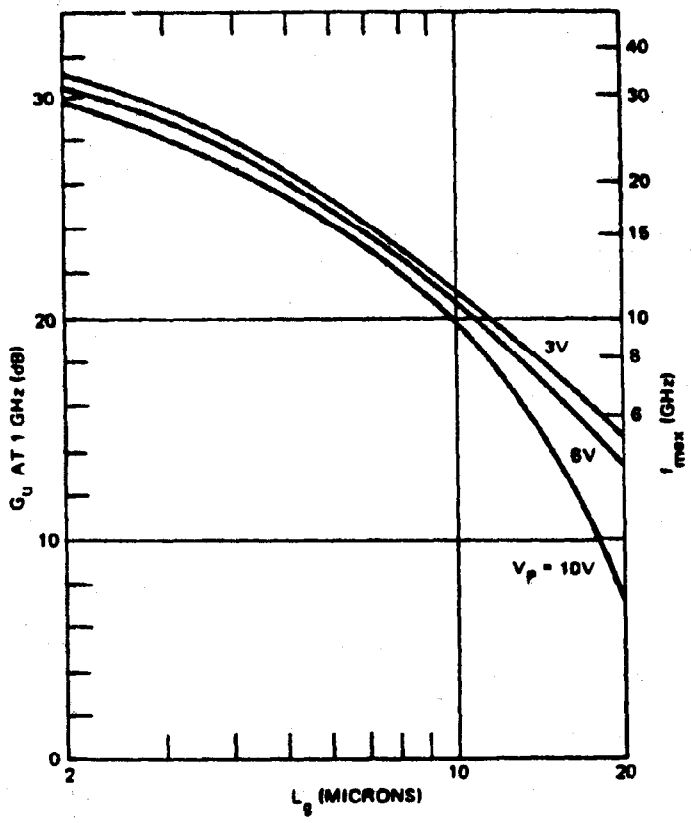
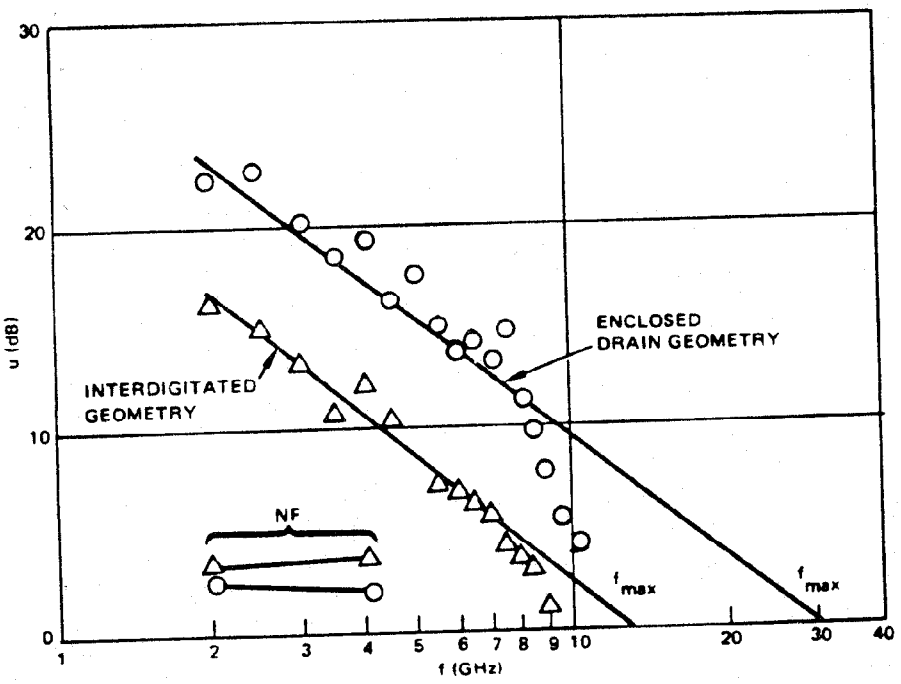


FIGURE 32. Performance prediction for the GaAs FET as a function of gate length.

6.3.1.3 Thermal considerations. Microwave power transistors, if they are to have a useful mean lifetime, must have an adequate thermal design. Thermal efficiency must be taken into account at dc levels and at the frequencies of operation.

From a dc standpoint, the most critical fabrication process is the attachment of the transistor chip to the header. The formation of the silicon-gold eutectic, so that 100 percent of the underside of the chip is wetted, ultimately determines the dc thermal capability of the device.

In many instances, tiny voids in the eutectic are created. These voids eliminate the heat transfer path that would be present if the eutectic had been properly formed. The rapidity of device degradation under dc load is proportional to the percentage of voided area under the chip.



V_{DS} = +6 V, I_{DS} = 40 mA

FIGURE 33. Actual performance of the two GaAs FET models.

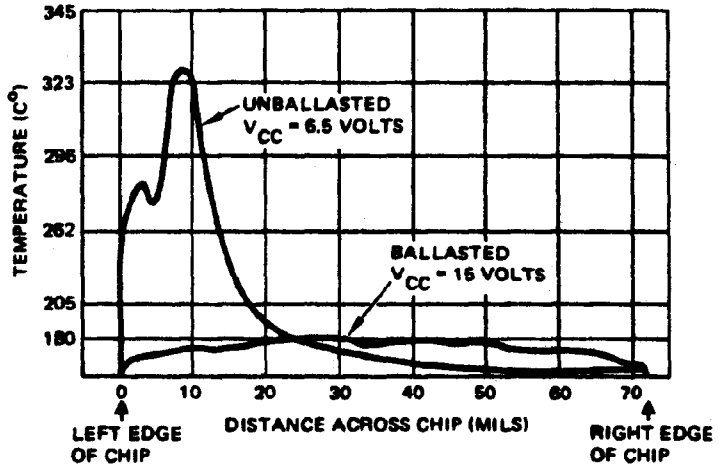
The reason for this is quite simple. The thermal impedance, θ_{JC}, of a void (hot spot) can be on the order of 20 to 50 °C/W. The measured θ_{JC} of the entire device is an average value, perhaps about 5.0 °C/W. The current density around the void raises the temperature in that vicinity much higher than the average junction temperature. This phenomenon is self-accelerating and is commonly known as thermal runaway.

The value of ballast resistors for a good thermal design has already been mentioned. This is applicable to any microwave transistor geometry. The equalizing of current in each metal finger is a great aid in slowing thermal runaway.

Thermal analysis of a transistor chip can be performed with great accuracy through the use of an infrared scanning microscope. A thermal profile, such as the one in Figure 34, can give a direct indication of the temperature anywhere on the surface of the chip. Figure 34 also shows the advantage of ballast

6.3 MICROWAVE DEVICES, TRANSISTORS

resistors in keeping the junction temperature at a significantly lower level than that of an unballasted device. The resulting increase in safe operating area can be seen in Figure 35.

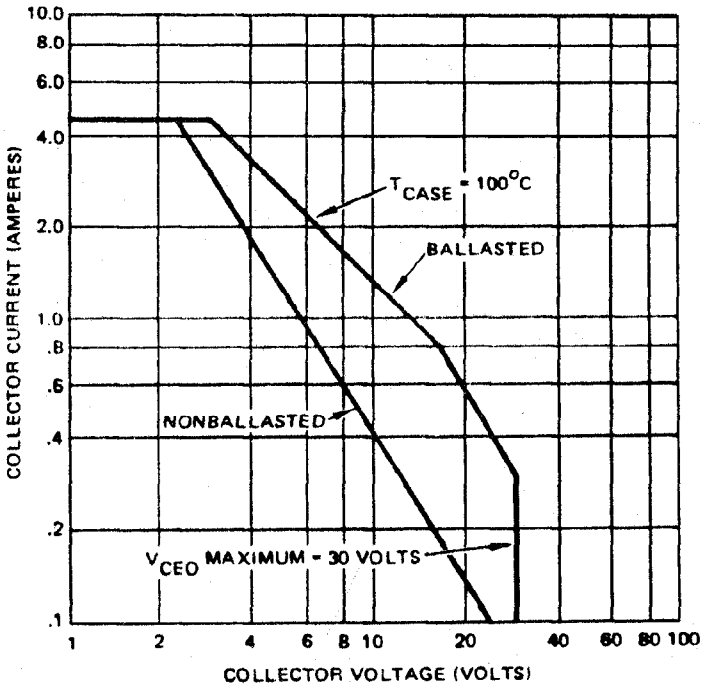


P_{DISS} = 13 W
T_{CASE} = 100 °C

FIGURE 34. Thermal profiles of a ballasted transistor vs an unballasted transistor.

It was previously mentioned that conventional θ_{JC} measurements yielded the average value for the device. The IR technique offers the unique ability to determine what can be called the hot spot, θ_{JC}. This means that the θ_{JC} is calculated from the temperature of the hottest spot on the chip. The advantage of this is a conservative specification of a device for a longer mean lifetime.

Adverse thermal effects at rf levels are generally attributed to the amount of mismatch (VSWR) the output of the transistor encounters. Figure 36 shows the variation in collector efficiency, hot spot temperature, and hot spot θ_{JC} over a range of frequencies for two values of output VSWR. Power reflected back into the transistor at high VSWRs causes excessive dissipation. This raises the average junction temperature as well as the hot spot temperature, causing an overall loss in collector efficiency.



$T_{HOT\ SPOT}$ = 200 °C maximum
I_C maximum = 4.5 A

FIGURE 35. Safe operating area ballasted vs unballasted.

6.3.1.4 Metallization systems. Since the beginning of the semiconductor age, the contact metallization for transistors has classically been aluminum. In general applications, virtually no shortcomings could be cited. Basic advantages are as follows:

a. Excellent conductivity

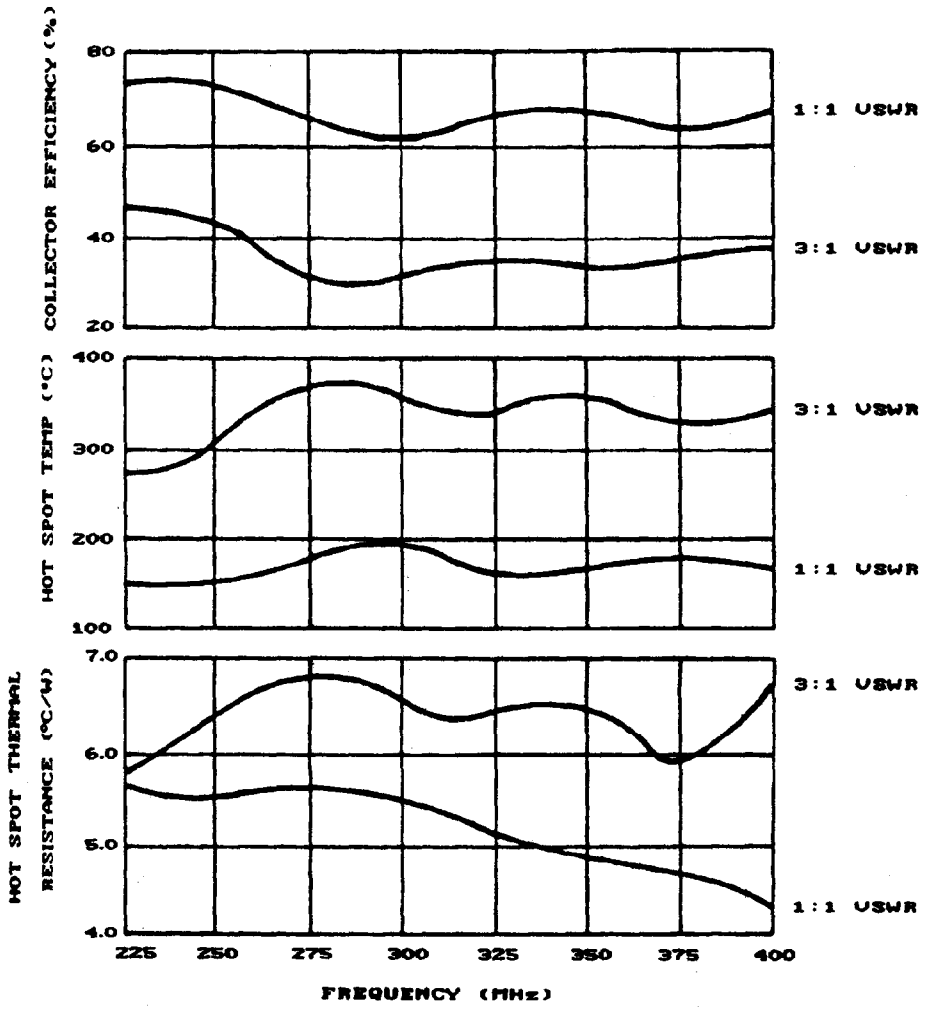
b. Easily sputtered to contact areas

c. Formed good eutectic with silicon

d. Easily bonded to

e. Adheres well to SiO_2.

6.3 MICROWAVE DEVICES, TRANSISTORS



V_{CC} = 28 V
T_{CASE} = 85 °C
P_{OUT} = 25 W at VSWR of 1:1

FIGURE 36. Hot spot thermal resistance, hot spot temperature and collector efficiency as a function of frequency.

The discovery of aluminum's propensity to migrate in masses into silicon under certain conditions caused a great deal of alarm in the industry.

Aluminum migration as a failure mechanism was not discovered until power devices in the 100 MHz region were in general usage. Because the migration is a function of current density and microwave devices use extremely fine geometries, the conditions for failure were perfect. Since then it has also been discovered that junction temperature and time also play an important role. J.R. Black, in 1969, related all of these factors by the equation

$$MTBF = \frac{1}{AJ^2 e^{-\emptyset/kT}}$$

where:

MTBF = Mean time between failure (hours)

A = a constant containing a factor equal to the cross sectional area of the metallization

J = current density (amperes/cm^2)

$\emptyset$ = activation energy (ev)

k = Boltzman's constant

T = metallization temperature (K)

This left only two possible choices: redesign the devices to lower the current density or redesign using a metal which has less of a propensity to migrate. In actuality, both of these alternatives were used.

The conclusion reached by J. R. Black and widely accepted in the rf semiconductor industry is the following:

For an unpassivated device with

$T_J = 200\ °C$
$A = 10^{-7}\ cm^2$

the current density must be kept under 1×10^5 amperes/cm^2 for a 2-year mean lifetime. For higher current densities, the maximum rated junction temperature must be reduced accordingly.

Development of other metal systems has taken place and many are in use. Some examples are:

a. Molybdenum-gold

b. Nichrome-tungsten-gold

c. Platinum silicide-tungsten-gold

d. Chromium-silver-gold.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS

All of these metals are heavy, or refractory. In essence, they are capable of resisting mass transport via momentum exchange due to their high atomic weight. The majority of products today employ these metallization systems.

In summing up the introduction to microwave transistors, it is clear that, of the many current and new technologies discussed, each has its relative merits and shortcomings. The metal matrix geometry is widely used in power applications due to its higher frequency limit. For small signal applications, the only feasible conventional geometry is the interdigitated geometry. Low-current requirements and high-frequency capability (f_t < 6 GHz) are the primary reasons. However, if the application requires amplifying small signals up to the frequency of x band and above with a low noise figure, the GaAs MESFET is the best possible three-terminal device.

Device failure due to metal problems can be considerably reduced through the use of refractory metals such as gold, tungsten, and chromium instead of aluminum. In addition, power devices should use emitter ballast resistors to equalize current sharing. The fact that metal failure mechanisms are self-accelerating make these requirements mandatory. These technologies are currently being used in high reliability microwave devices.

6.3.2 Usual applications. The primary use of microwave transistors is in strip transmission line amplifier and oscillator applications for low-noise receiver front ends, transmitters, and local oscillator operation. Figure 37 shows a microwave transistor-amplifier circuit with input and output matching networks. The unilateral power gain is the sum of the additional power gain (or loss) resulting from the input impedance matching network between the device and the source, the forward power gain of the device with the input and output terminated in matching loads, and the additional power gain (or loss) due to the impedance matching network between the output of the device and the load: $Gu = G_s + G_f + Ge$ in dB.

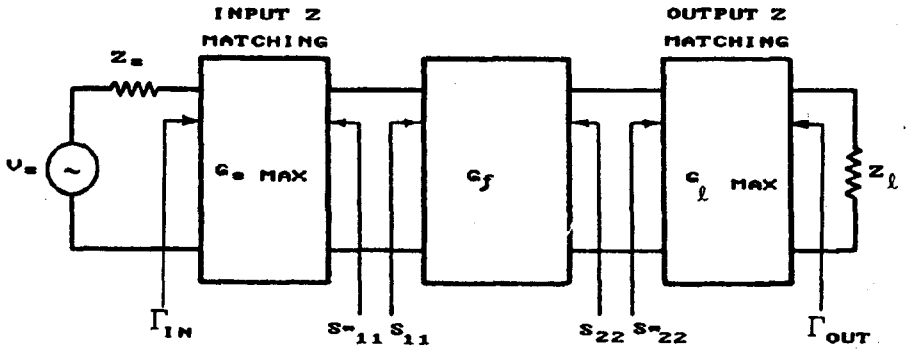


FIGURE 37. Amplifier circuit for maximum gain.

The maximum unilaterial gain will be obtained by choosing impedance matching networks so the reflection coefficients of the source and the load are equal to the complex conjugates of the input and output reflection coefficients of the device, respectively: $\Gamma_S = S_{11}^*$ and $\Gamma_\ell = S_{22}^*$, $Z_S = Z_O$, $Z_\ell = Z_O$.

If a transistor amplifier is to be unconditionally stable, the magnitudes of S_{11}, S_{22}, Γ_{in}, and Γ_{out} must be smaller than unity and the transistor's inherent stability factor K must be greater than unity and positive. K is computed from

$$K = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{12}S_{21}|} > 1$$

where $\Delta = S_{11}S_{22} - S_{12}S_{21}$

the input and output reflection coefficients are given by

$$\Gamma_{in} = S_{11} + \frac{S_{21}S_{12}\Gamma_\ell}{1 - S_{22}\Gamma_\ell}$$

and

$$\Gamma_{out} = S_{22} + \frac{S_{21}S_{12}\Gamma_S}{1 - S_{11}\Gamma_S}$$

Also, the boundary conditions for stability are given by

$$\left|\Gamma_{in}\right| = 1 = \left|S_{11} + \frac{S_{21}S_{12}\Gamma_\ell}{1 - S_{22}\Gamma_\ell}\right|$$

and

$$\left|\Gamma_{out}\right| = 1 = \left|S_{22} + \frac{S_{21}S_{12}\Gamma_S}{1 - S_{11}\Gamma_S}\right|$$

Substitution of real and imaginary values for the S parameters in the previous two equations and solving for Γ_ℓ and Γ_S yields

$$R_S \text{ (radius of } \Gamma_S \text{ circle)} = \frac{|S_{12}S_{21}|}{|S_{11}|^2 - |\Delta|^2}$$

$$C_S \text{ (center of } \Gamma_S \text{ circle)} = \frac{C_S^*}{|S_{11}|^2 - |\Delta|^2}$$

$$R_\ell \text{ (radius of } \Gamma_\ell \text{ circle)} = \frac{|S_{12}S_{21}|}{|S_{22}|^2 - |\Delta|^2}$$

$$C_\ell \text{ (center of } \Gamma_\ell \text{ circle)} = \frac{C_\ell^*}{|S_{22}|^2 - |\Delta|^2}$$

6.3 MICROWAVE DEVICES, TRANSISTORS

where $\Delta = S_{11}S_{22} - S_{12}S_{21}$

$C_s = S_{11} - \Delta S_{22}^*$

$C_\ell = S_{22} - \Delta S_{11}^*$

The reflection coefficient of the source impedance required to match the input of the transistor conjugately for maximum power gain is

$$\Gamma_{sm} = C_s^* \quad \frac{B_s \pm \left(B_s^2 - 4|C_s|^2\right)^{1/2}}{2|C_s|^2}$$

where $B_s = 1 + |S_{11}|^2 - |S_{22}| - |\Delta|^2$

The reflection coefficient of the load impedance required to match the output of the transistor conjugately for maximum power gain is

$$\Gamma_{\ell m} = C_\ell^* \quad \frac{B_\ell \pm \left(B_\ell^2 - 4|C_\ell|^2\right)^{1/2}}{2|C_\ell|^2}$$

where $B_\ell = 1 + |S_{22}|^2 - |S_{11}|^2 - |\Delta|^2$

If the computed values of B_s and B_ℓ are negative, the plus sign should be used in front of the radical in the equations for Γ_{sm} and $\Gamma_{\ell m}$. Conversely, if B_s and B_ℓ are positive, the negative sign should be used.

Stability circles can be plotted directly on a Smith chart. These circles separate the output or input planes into stable and potentially unstable regions. A stability circle plotted on the output plane indicates the values of all loads that provide negative real input impedance, thereby causing the circuit to oscillate. A similar circle can be plotted on the input plane which again indicates the values of all loads that provide negative real output impedance and causes oscillation. A negative real impedance produces a reflection coefficient that has a magnitude greater than unity. The regions of instability occur within the circles whose centers and radii are R_s, C_s, R_ℓ, C_ℓ.

By using an appropriate sign, only one answer is possible in either of the equations for Γ_{sm} or $\Gamma_{s\ell}$ and a value of less than unity is obtained. The maximum available power gain possible is expressed as

$$G_{max} = \frac{|S_{21}|}{|S_{12}|}\left|K \pm (K^2 - 1)^{1/2}\right|$$

Hence, maximum available power gain is obtained only if the microwave transistor amplifier is loaded with Γ_{sm} and $\Gamma_{\ell m}$ as reflection coefficients. The maximum frequency of oscillation is determined after the maximum available power gain is achieved.

It is possible to design more than one stage of amplification to achieve high gain and gain flatness over a specified bandwidth (Figure 38A). However, for multistage amplifiers where gains of 30 dB or greater are required, the use of such "single-ended" designs can lead to difficulty in producing the required gain flatness and distortion. This is because of the mismatches occurring between gain stages causing gain and phase variations. It is possible to minimize mismatches by designing interstage networks that will match the output impedance of the previous stage to the input impedance of the following stage and provide a positive sloped gain over some bandwidth to compensate the transistor's roll-off (Figure 38C) and give an overall flat gain (Figure 38E). Isolators can be added to the input, output, and between stages. Isolators can produce significant loss which will increase the noise figure.

A balanced amplifier shown in Figure 39 can operate with both high gain and wide bandwidth. Two FET amplifier units of the same performance are arranged between the output and input ports of two 3-dB, 90-degree hybrid dividers or two Wilkinson power dividers with 90-degree phase shifters.

The input and output VSWRs are predominately those of the hybrid dividers. If the individual amplifiers are not perfectly matched at certain frequencies, a signal in the zero degree arm of the divider will be reflected from the corresponding amplifier and a signal in the $\pi/2$ arm of the divider will be similarly reflected from its amplifier. The signals will be phased by 0 or $\pi/2$ again by traveling through the divider after reflection, and the signals will be fed to the load at the isolated port.

The total reflection and power transmission can be expressed as

$$S_{11} = 1/2\ (S_{11a} - S_{11b})$$

$$S_{22} = 1/2\ (S_{22a} - S_{22b})$$

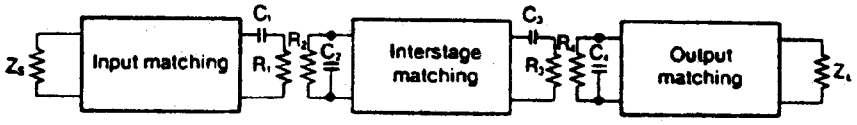
and

$$\text{Gain} = |S_{21}|^2 = 1/4|S_{21a} + S_{21b}|^2$$

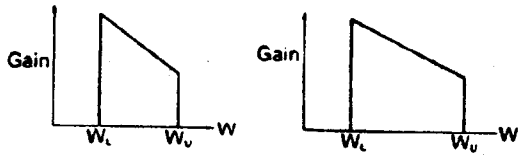
where a and b indicate the two GaAs MESFETs. The 1 and 2 refer to the input and output ports, respectively. The input and output reflections are reduced to one-half of the corresponding difference of reflection coefficients of the two MESFETs, and the total power gain is equal to the individual power gain if the two MESFETs are identical.

6.3 MICROWAVE DEVICES, TRANSISTORS

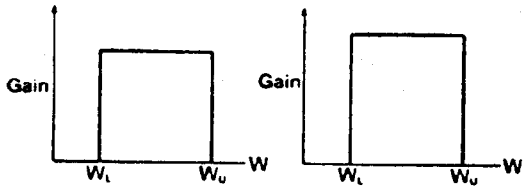
A. Amplifier schematic



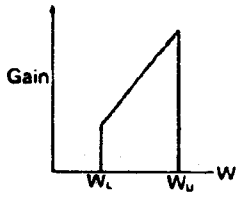
B. FET impedance models



C. Frequency response of transistors

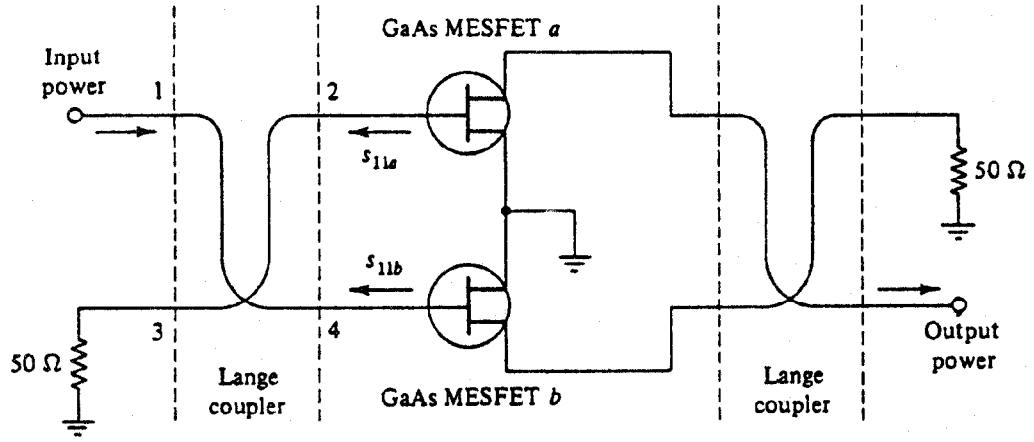


D. Frequency response of input and output matching networks

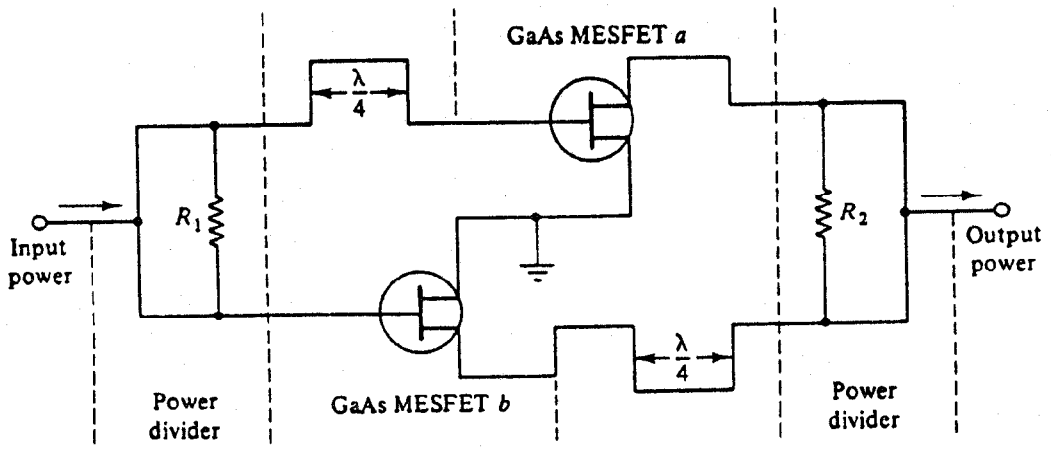


E. Frequency response of interstage

FIGURE 38. Multistage amplifier with input, output, and interstage matching network.



A. Balanced amplifier with Lange couplers



B. Balanced amplifier with Wilkinson power dividers and phase shifters

FIGURE 39. Two types of balanced amplifier circuits.

6.3 MICROWAVE DEVICES, TRANSISTORS

If one MESFET fails, the loss for the balanced amplifier would be only 6 dB. The balanced amplifier configuration is by far the most common design in modern microwave integrated circuits. Its advantages are:

a. Good input and output VSWRs

b. Good stability

c. High reliability

d. Low tuning work

e. Linear output power increases by 3 dB

f. Up to 20 dB power gain with 10 percent bandwidth at X band.

Both silicon bipolar transistors and GaAs MESFETs can be designed as an amplifier or an oscillator, depending on whether the stability factor K is greater or less than unity. Common-emitter or common-source configuration is preferred for amplifier design and, alternatively, common-base or common-gate configuration for oscillator design. The stability factor K is greater than unity for an amplifier, but the input and output reflection coefficients (Γ_{in} and Γ_{out}) are less than unity. For an oscillator, however, the stability factor K is less than unity, but the input and output reflection coefficients are equal to or greater than unity. In amplifier design, two matching networks are required to match the input and output ports of the device to give a maximum transducer power gain, a minimum noise figure, and low input and output reflection coefficients. In a small-signal oscillator design the same two-port (or one that is properly modified with a feedback path) can be designed to deliver nearly the same output power to the same 50-Ω load.

Two-port oscillator design may be summarized as follows:

a. Select transistor with sufficient gain and output power capability for the frequency of operation. This may be based on oscillator data sheets, amplifier performance, or S-parameter calculation.

b. Select a topology that gives $k < 1$ at the operating frequency. Add feedback if $k < 1$ has not been achieved.

c. Select an output load matching circuit that gives $|S_{11}'| > 1$ over the desired frequency range. In the simplest case, this could be a 50-Ω load.

$$|S_{11}'| = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L}$$

d. Resonate the input port with a lossless termination so that $\Gamma_G S'_{11} = 1$. The value of S'_{22} will be greater than unity with the input properly resonated.

$$S'_{22} = S_{22} + \frac{S_{12}S_{21}\Gamma_G}{1 - S_{11}\Gamma_G}$$

The transistor will oscillate with any of the six configurations given in Figure 40. In all cases, the transistor delivers power to a load and the input of the transistor. Practical considerations of circuit design and dc biasing will determine the best design. A bipolar transistor oscillator circuit is shown in Figure 41.

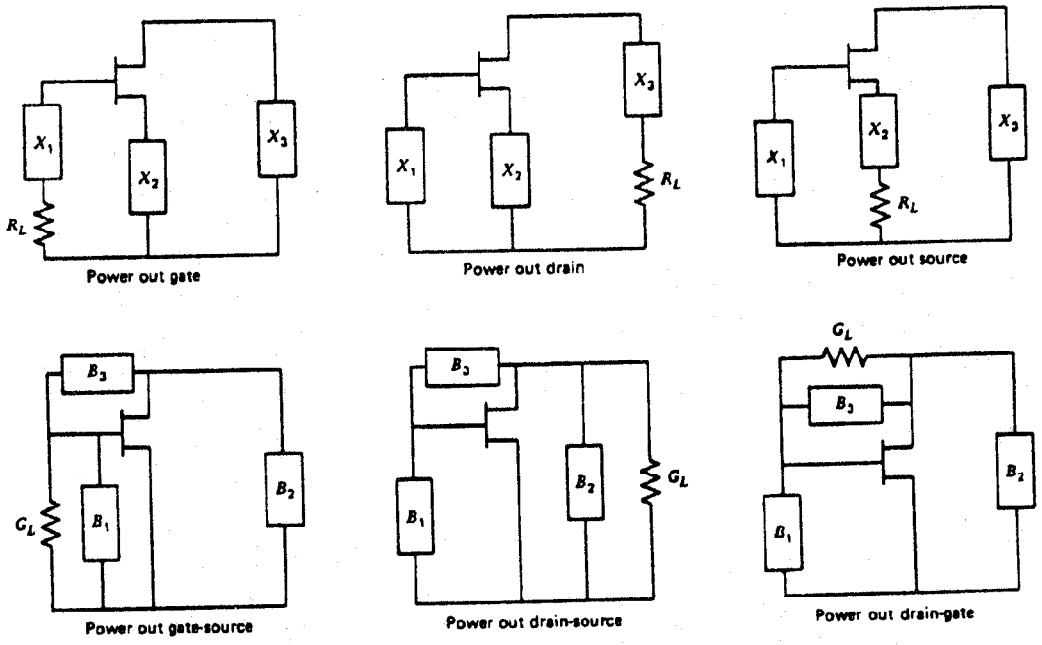


FIGURE 40. Six oscillator structures.

Microwave transistors can be used as mixers as shown in subsection 6.8. The non-linearities of bipolar transistors and FETs that produce mixing are used in frequency multiplier applications. The GaAs MESFET and, in particular, the dual-gate FET work exceptionally well. The major advantage over SRD multipliers, which is the greater than 100 percent efficiency, results in the replacement of complicated varactor diode chains and resultant pre- and post-amplification. Table III compares the performance of varactor diode, bipolar transistor, and FET multipliers.

6.3 MICROWAVE DEVICES, TRANSISTORS

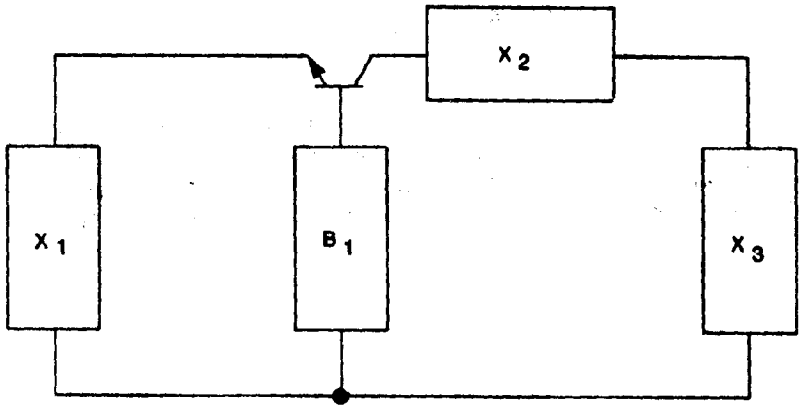


FIGURE 41. Bipolar transistor oscillator circuit.

TABLE III. Comparison of microwave frequency multipliers

Device Characteristic	Diodes	BJT	FET 1/
Bandwidth	Narrow	Medium	Wide
RF driving power	300 mW	100 mW	1 mW (10 mW)
Output maximum frequency	SRD: 18 GHz Varactor: 120 GHz	11 GHz	30 GHz
Isolation	Poor	Medium	Good
Idlers	Critical	Critical	Less Critical
Power handling (X-band)	1 to 4W	0.5W	1 to 4 W
Stability	Good	Poor	Excellent
Higher harmonic distortion	High	Low	Low
RF efficiency	-1.5 to -3 dB	-2 dB	Up to +10 dB (up to 3 dB)

1/ Figures in parentheses are results for a single-gate FET.

6.3.3 Physical construction.

6.3.3.1 General. Microwave transistor fabrication hinges most critically on the three items in the introduction of this subsection. It is the accuracy with which these processes are conducted which has the most profound effect on the electrical characteristics of the final device. These will be discussed one by one, paying particular attention to the problems involved in microwave transistor fabrication.

Masking. The function of a photo mask is to define an area on a silicon wafer which is to be subsequently etched. In concept the process is very simple.

a. The wafer is coated with approximately 4000 Å of photo resist. A layer this thin is achieved using centrifugal action.

b. The photo mask is placed over the wafer, and light passing through the clear areas of the mask hardens the photo resist by polymerization.

c. The pattern is further developed and hardened so it will withstand a chemical etch. The etching procedure removes the unwanted, unexposed regions of the wafer.

In practice, several difficulties must be settled. First, there is the problem of producing masks for the extremely fine geometries necessary for microwave transistors. The fingers of a small signal interdigitated device in the 1- to 3-GHz region are on the order of 1.0 to 2.5μ wide. In addition, the tolerances required for subsequent operations complicate matters even more. For instance, the alignment of emitters in an overlay device requires a registration accuracy of approximately 1μ.

To achieve these levels of accuracy, many refinements in equipment and technology were required. Today, photolithography and self alignment can produce reliable widths to submicron dimensions in association with ion implantation. Direct writing techniques can increase accuracy and further reduce the critical dimensions.

Diffusion and dopants. The key factor regulating the f_{max} of a microwave transistor is the emitter-collector transit distance or base width. The narrower this region, the higher the f_{max}. Typical base widths are on the order of 1,000 to 2,000 A. Ability to control diffusion depth from one run to the next is essential to the production of relatively homogeneous devices.

The dependence of the gain-bandwidth product, f_t, on the emitter-collector transit distance can best be shown by the equation:

$$T_{ec} = \frac{kT\,C_{te}}{qI_e} + \frac{W^2}{nD} + \frac{X}{2V} + C_c r_c$$

6.3 MICROWAVE DEVICES, TRANSISTORS

where

T_{ec} = emitter-collector transit time

$\frac{kT}{q}$ = constant, 26 mv at 300 K

I_e = emitter current

C_{te} = emitter transition capacitance

W = base width

n = doping profile constant

D = base carrier diffusion constant

X = collector space charge width

v = carrier velocity

C_c = collector-base capacitance

r_c = collector series resistance

This equation can also be written as

$$T_{ec} = T_c + T_b + T_{sc} + T_c$$

where

T_e = emitter charging time

T_b = base transit time

T_{sc} = collector space charge transit time

T_c = collector charging time.

It can be seen that each term in the equation has either distance or capacitance as the independent variable. Capacitance is extremely dependent on geometry and materials, whereas distance is controlled by the diffusion. Therefore, the key to producing devices with the highest possible f_t depends largely on the controls established for the diffusion process.

Diffusion, in practice, is an extremely delicate operation, especially when final base depths of only a few thousand angstroms are desired. Subsequent high temperature operations will cause the initially diffused base to be further diffused in both a downward and lateral direction.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

Emitter diffusion also has its problems. The phenomenon of "emitter push," that is, the deepening of the base diffusion directly under the emitter diffusion, destroys the f_{max} characteristic of the device.

Recently, many advances have been made in the technology and materials used in the diffusion process. The remainder of this section will be devoted to discussing them and their relative merits.

Ion implantation is a technique through which ions are driven into substances due to energy received from particle accelerators. Obvious advantages of this technology are the great degree of control over implantation depth, because the depth is determined by the energy level of the ions (Figure 42) and secondly, the ability to obtain extremely fine line widths. From Figure 42, another obvious advantage can be seen. That is the absence of lateral diffusion for an implanted junction.

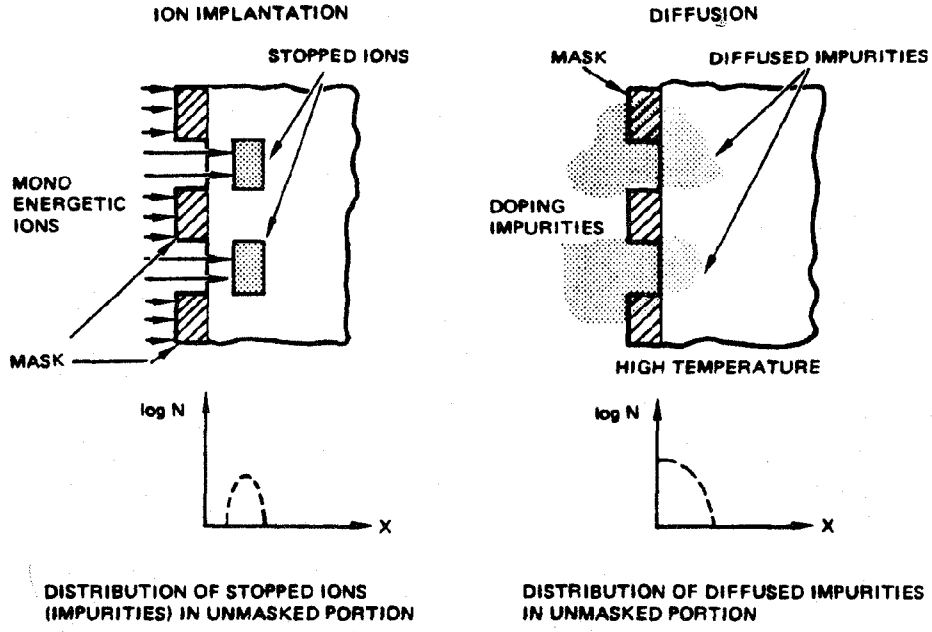


FIGURE 42. Impurity profiles--ion implantation vs diffusion.

The major drawback of ion implantation for use in microwave transistor fabrication is the possible damage of the lattice structure by radiation. Though ion implantation has definite possibilities, it seems that more favorable results are being obtained by using arsenic as an emitter dopant.

6.3 MICROWAVE DEVICES, TRANSISTORS

The classic n-type dopant used in transistor fabrication is phosphorus. It is easily obtained and relatively simple to use. With the development of microwave transistors its use continued until it was discovered that many problems in device performance were linked directly to phosphorus. Both lateral diffusion and the "emitter push" phenomenon were associated with the properties of phosphorus itself or the processing conditions which had to be used.

The next n-type element on the periodic table is arsenic. Its use was initially avoided due to inability to obtain a satisfactory diffusion process. After a successful process was developed and experimental devices fabricated, the results of testing were conclusive. The heavier arsenic atom exhibited an unprecedented superiority over phosphorus. The depth of diffusion was controllable to such an extent that ion implantation efforts were temporarily tabled.

A brief summary of microwave transistor types and their relative merits and shortcomings is shown in Table IV.

TABLE IV. Summary of the four microwave technologies

Emitter Periphery	Inter-digitated	Overlay	Metal Matrix	GaAs FET
Base area (frequency capability)	Higher	High	Highest	Highest
Metal cross-sectional area	Low	High	High	Low
Processing	Easy	Easy	Difficult	Difficult
Chip size	Small	Variable	Large	Small
Current/power handling capability	Low	High	High	Low
Available MIL parts	Several	Several	None	None
Availability (no. of manufacturers)	All	Several	One	Many

When viewed from the more or less ideal situation depicted on a flow chart, the fabrication of a microwave transistor seems relatively simple. Mask tolerances, diffusion depths, oxide thicknesses, and other variables are tightly controlled so that the outcome is, theoretically, a high yield lot of homogeneous devices.

However, real world forces can change the theoretical. Producing microwave devices on a repeatable basis is, in actuality, one of the most difficult tasks to accomplish. The product designer starts with a list of desired parameters, as follows:

P_d Power dissipation

f_t Gain bandwidth product

S_{21} Forward transmission coefficient

S_{12} Reverse transmission coefficient

S_{11} Forward reflection coefficient

S_{22} Reverse reflection coefficient

Noise figure.

From this the product designer must then decide the qualities of the required raw material. The primary considerations are the resistivity of the bulk silicon and the resistivity and thickness of the epitaxial layer that will be grown. Then, depending on the required limits on the electrical parameters, the designer calculates the optimum values for the following:

Diffusion depths

Dopant concentrations/profiles

Oxide thickness

Line widths/layout

After these are established, the designer generates a flow chart to insure process control.

In-process testing of the product at specified intervals is an integral part of this flow chart. In this manner, the product engineer can see how the established controls are affecting the parameters of the devices. (This testing is of prime importance because it affects percentage yield).

So the reader may develop a feeling for what is actually involved in processing a microwave transistor, the key steps will be described in detail. Because the interdigitated geometry is the simplest type and the most common, it will be used for illustrative purposes.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS

6.3.3.2 Interdigitated device fabrication. The fabrication of an npn microwave device begins with the preparation of a slice of n^+-type material. Initial slice thickness is approximately 15 mils and the resistivity is on the order of 0.01 Ω-cm. The crucial first step is the growth of a thin epitaxial layer of high resistivity n-type material, Figure 43A. Epitaxial growth is nominally 10 μM thick and has a resistivity of approximately 1Ω-cm. Exact values are determined by the desired parameters. For instance, devices with thinner epitaxial layers have higher values of f_t and lower collector-base breakdown voltages than devices with thicker epitaxial layers. True microwave devices, by our definition, are those devices that operate at 1 GHz or higher and must have this epitaxial growth. Without it obtaining the required values of f_t would be impossible.

On top of the epitaxial layer, silicon dioxide is grown to a thickness between 8000 and 12,000 Å, Figure 43B. This layer of oxide serves a dual function. First, the layer passivates the entire surface of the epitaxial layer. Secondly, in conjunction with photoresist, masking, exposure, and etching operations the layer allows for the definition of the base region, shown in Figure 43C. The opening of small windows through the oxide provides the means through which the p-type base dopant (usually boron) can be diffused into the epitaxial layer. For microwave devices, the initial depth of this diffusion is shallow, no more than a few thousand angstroms. The final depth of the base diffusion, which depends on the next process step, is instrumental in determining the f_t of the device.

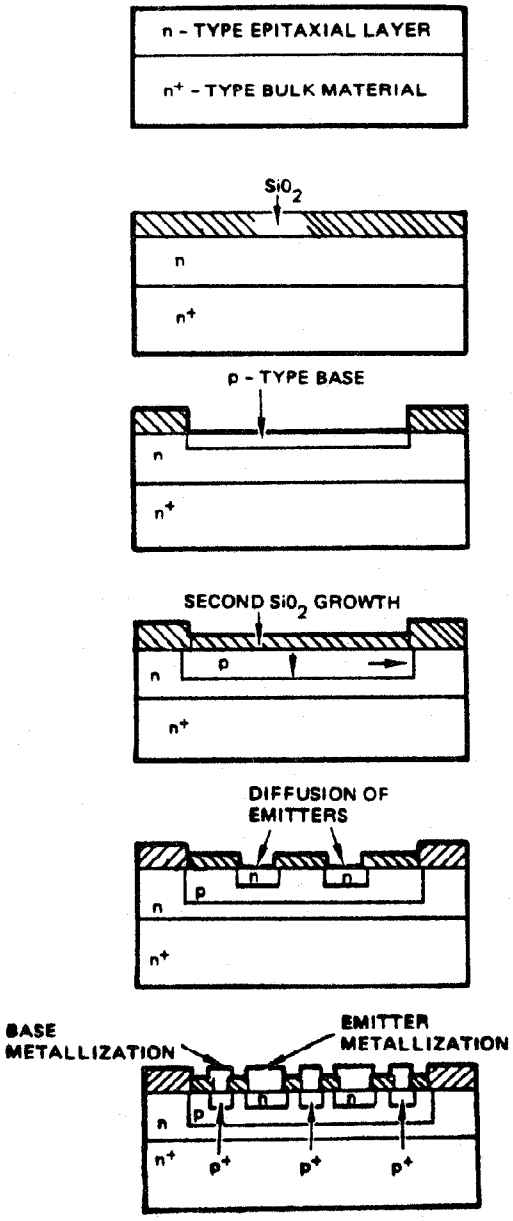
Following the diffusion of the base region, another layer of silicon dioxide is grown, Figure 43D. This oxide passivates the entire base region. For the fabrication of microwave devices, however, this process has one definite disadvantage. As the silicon dioxide is being grown, further diffusion of the base is occurring simultaneously. This is caused by the elevated temperatures (approximately 1,000 °C) required for oxide growth. At these temperatures thermal excitation moves the p-type impurity atoms in both a downward and lateral direction, as indicated by the two arrows in Figure 43D. The downward diffusion degrades the device's capability to attain the highest possible f_t because it effectively increases the distance through which charge carriers must travel from emitter to collector. Several techniques, such as plasma chemical vapor deposition, have been developed to grow silicon dioxide at reduced temperatures.

The emitter stripes and bonding pad are the next regions to be defined. Again, selective etching, following the usual photoresist, masking, and exposure cycle, creates windows through which the n-type emitter dopant (phosphorus or arsenic) can be diffused as shown in Figure 43E. Following the emitter diffusion, these emitter stripes also serve as the contact area for the emitter metallization. However, before the metallization process can take place, contact areas to the base region must be formed.

So, again, the procedure of photoresist, masking, exposure, and etching must be conducted to create the metallization contact areas for the base stripes and bonding pad. The base contact area is actually a p^+ region diffused into the p region.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

The final process, evaporation of metal (sputtering) deposits from 8,000 to 20,000 angstroms of metal to the contact areas, is shown in Figure 43F. For a more complete view of the device refer to Figure 44.



A

A high resistivity n-type epitaxial layer is grown on the n^+-type bulk material.

B

This is followed by the growth of a thick (1200 Å typ.) layer of silicon dioxide (SiO_2).

C

The base region is defined by etching away an area of oxide and then diffusing p-type material into the exposed n-type epitaxial layer.

D

The growth of another layer of SiO_2 simultaneously causes further diffusion of the base in both a lateral and downward direction (as indicated by arrows)

E

This oxide is etched away in selected locations so that the n-type emitter can be diffused.

F

The final operations are the etching of oxide over the base region, diffusion of p^+ base contact areas and sputtering of metallization.

FIGURE 43. Interdigitated device fabrication.

6.3 MICROWAVE DEVICES, TRANSISTORS

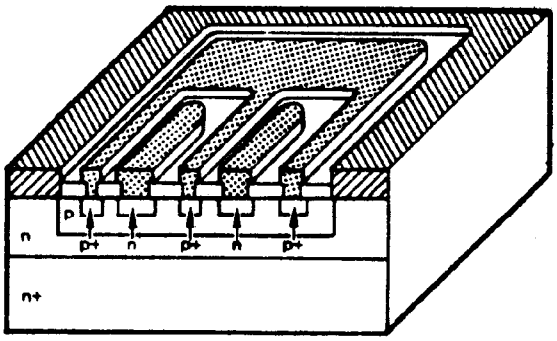


FIGURE 44. Cross-section of an interdigitated device.

The processes just described were done in an oversimplified manner. Though they are commonly used throughout the industry, the inherent difficulties of processing are always present. The processes should be examined in light of some of the necessary properties of a microwave transistor and the precision that must be maintained with respect to the following: epitaxial thickness, epitaxial resistivity, base diffusion depth, emitter diffusion depth, mask fabrication, masking alignment tolerances, emitter stripe width, metallization widths, and overall chip size.

The following variables that must be controlled throughout the entire fabrication operation should be considered: diffusion furnace temperatures, dopant concentration and flow rate, process time duration, etch rates, growth rates, and diffusion rates.

Each variable can shift the parameter distribution for a given production lot entirely out of specification. Hence, their control within precise limits is mandatory for product uniformity and repeatability.

6.3.3.3 Overlay device fabrication. The fabrication of the overlay transistor begins to differ from that of the interdigitated transistor during the diffusion of the p-type base dopant. The mask incorporated causes the dopant to be diffused in the shape of a grid.

Emitter sites are diffused precisely at the center of each square formed by the p-type base grid. Before the metallization process, the entire chip is passivated with polycrystalline silicon. Contact areas are then etched through the

passivation to the emitter sites and base grid lines. The metallization is then sputtered in the usual manner. A cross section of the overlay geometry is shown in Figure 45.

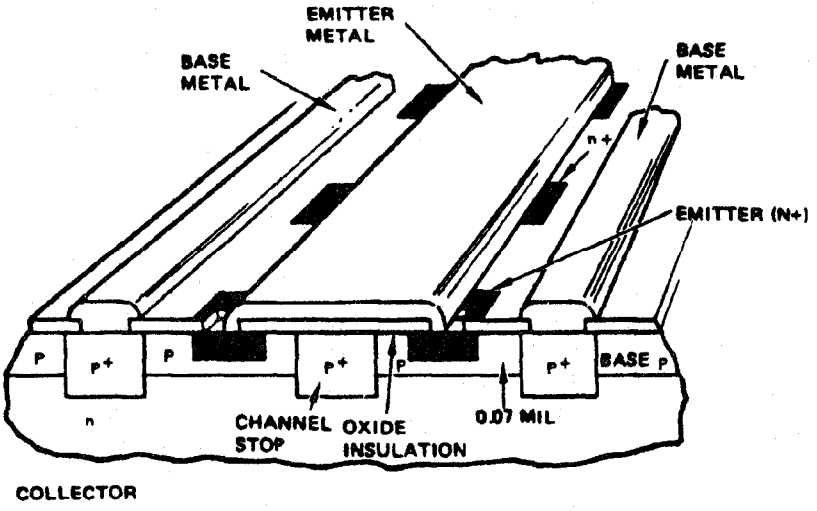


FIGURE 45. Overlay device fabrication.

6.3.3.4 Metal matrix device fabrication. Fabrication of the metal matrix device is probably the most difficult of the three geometries discussed. But because the inherent advantages outweigh the processing difficulties, many manufacturers are experimenting with this technology. After the problems are solved this geometry will be a major factor in the power microwave market.

The major problem is the accuracy required for alignment during photoresist and masking operations. Definition and location of emitter regions must be accomplished within 1 μ. Most electrical rejects during wafer probe are directly traceable to poorly aligned geometries.

The number of processing steps required for metal matrix device fabrication is much greater than required for either overlay or interdigitated. Therefore, initial manufacturing costs are higher. However, with time, these will ultimately be reduced.

Current fabrication refinements have yielded devices that can be used at nearly 3 GHz with 5 W output.

6.3 MICROWAVE DEVICES, TRANSISTORS

6.3.3.5 GaAs FET fabrication. The two techniques used for GaAs MESFET fabrication that have gained the most popularity are the self-aligned and the etched-channel technologies; both can exploit either optical or electron beam lithography to define the gate stripe, depending on the transistor's gate length. Electron beam lithography is used generally for gate lengths of less than 0.5 μ.

Self-aligned gate technology. Figure 46 shows the basic processing steps of the self-aligned gate technology. The first step is defining the isolation mesa by etching away the active n layer until the semi-insulating substrate is reached. The gate metal, (i.e., aluminum) is then evaporated over the active area as shown in Figure 46A.

Source and drain areas are defined in photoresist and the exposed gate metal is removed by etching. Over-etching is used to undercut the resist, as shown in Figure 46B, to allow the necessary space between gate and drain and gate and source. Gold ohmic contact metallization, usually In-Ge-Au or Au-Ge-Ni, is then evaporated (Figure 46C). The resist, which is protecting the gate stripe, is now covered with this ohmic metallization, but this is conveniently removed by "floating off" the gold by dissolving away the resist. Thus, the remaining thin gate is left situated between source and drain contacts as shown in Figure 46D.

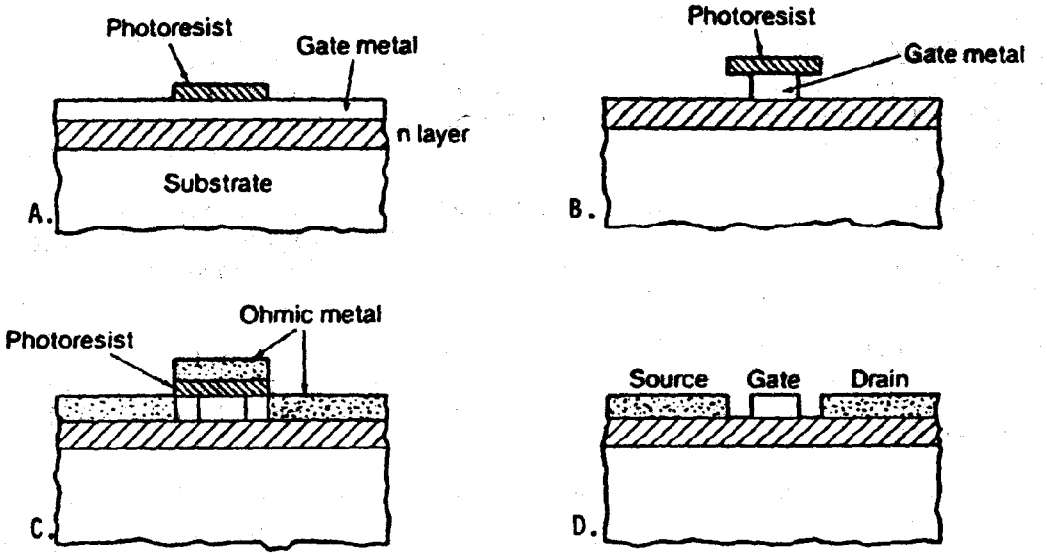


FIGURE 46. Processing steps of self-aligned gate technology.

Rather than define the active channel thickness by the thickness of the n-type epitaxial layer, a thicker layer is grown and the channel region under the gate is defined by etching. This removes the high tolerance in thickness required for the epitaxial layer when the channel region is not etched. Most companies use a preferential etch that gives a flat bottomed recess. Source and drain contacts are deposited first, as in Figure 47A, and the gate is defined in photoresist.

A channel is etched in the GaAs until a specific current is measured between the source and drain contacts. Gate metal is then evaporated and the excess metal (Figure 47C) removed by using the "float-off technique." This basic method works equally well with both photolithographic and electron beam resist exposure techniques. Gate lengths as low as 0.25 μ have been produced.

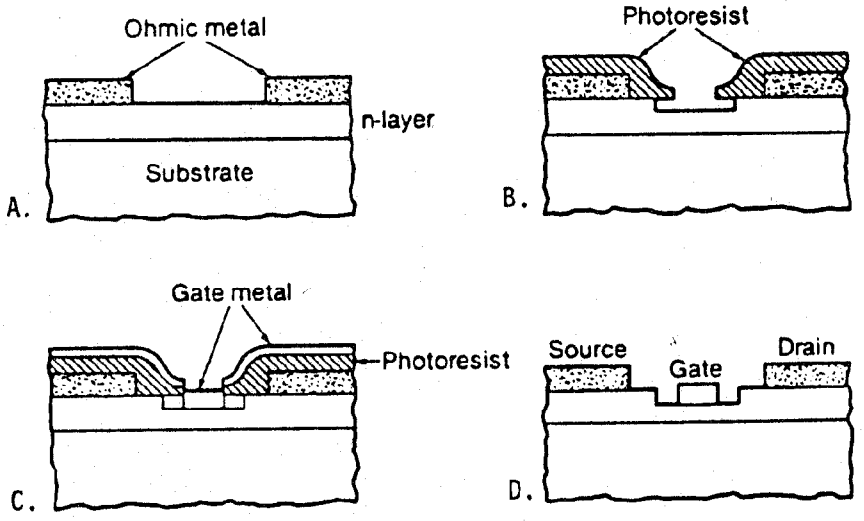


FIGURE 47. Processing steps of etched-channel technology.

6.3.3.6 Packaging of microwave transistors. Until microwave transistors were developed, packaging was simple. Ordinary switching transistors or low-frequency amplifiers could be sealed in epoxy for commercial use or hermetically sealed in metal cases for high reliability applications. In essence, the package did nothing to adversely affect the performance of the chip it contained.

Unfortunately, the same cannot be said about the microwave transistor. Conventional packages introduce stray capacitances and inductances which can cause such problems as parasitic oscillations, degradation of gain, poor broadband response, and impedance mismatch.

6.3 MICROWAVE DEVICES, TRANSISTORS

Therefore, it was necessary for new packaging concepts to be developed as fast as higher frequency devices were being produced. This task became somewhat easier with some of the breakthroughs in materials and in microwave techniques.

The most significant advances in microwave technology were the development of the stripline and microstripline techniques. These two methods, which were developed to achieve low loss transmission of microwave signals, proved to be extremely instrumental in the development of microwave transistor packages. Package development proceeded rapidly with the advent of the microstripline transistor package.

Overall package design and compatibility with the microstripline technique virtually eliminated performance problems associated with conventional packages and circuits. The major advantage that the microstripline technique offered was an excellent impedance matching network.

Consequently, the VSWR was kept quite close to 1:1. Package/circuit compatibility can be readily seen from the high collector efficiency and low reflected power.

A microwave transistor chip will always exhibit some degree of performance degradation when placed inside a package. Although refinements in packaging technology have reached a point where the actual degradation is minute, further refinements are always being sought. At microwave frequencies, designers strive for every 0.1 dB of gain they can possibly obtain while trying to keep the noise figure at a reasonable level. These are the two parameters that exhibit the greatest degradation due to interaction between package and chip.

For example at 2 GHz, the particular package in which a chip is mounted can account for a difference of 0.5 dB in gain.

Materials used at microwave frequencies have a profound effect on device performance. This point has already been established in the discussions on metallization systems and dopants. With the packaging of a microwave chip comes another critical decision in material selection.

Substrate material must exhibit such properties as good heat transfer capability, negligible effect on transistor performance, and ease of die attachment. Several materials perform well above 1 GHz. Among them are beryllium oxide, alumina, spinel, and sapphire. The latter is rarely used on discrete devices but is used more often in microwave integrated circuits.

Overall package technology has advanced to a point of successful high reliability application. Hermetic seal and ability to withstand rigid military environmental requirements are considered standard among the larger microwave packages.

However, a definite problem exists with certain extremely small packages. Structurally, they are incapable of surviving environmental tests such as vibration or centrifuge.

Recent developments in microwave integrated circuits and hybrids make use of another versatile package, the microstrip chip carrier. An example is shown in Figure 48. The chip is attached to the collector metallization, and wire bonds are made to the input and common, depending on the desired configuration, common emitter, or common base.

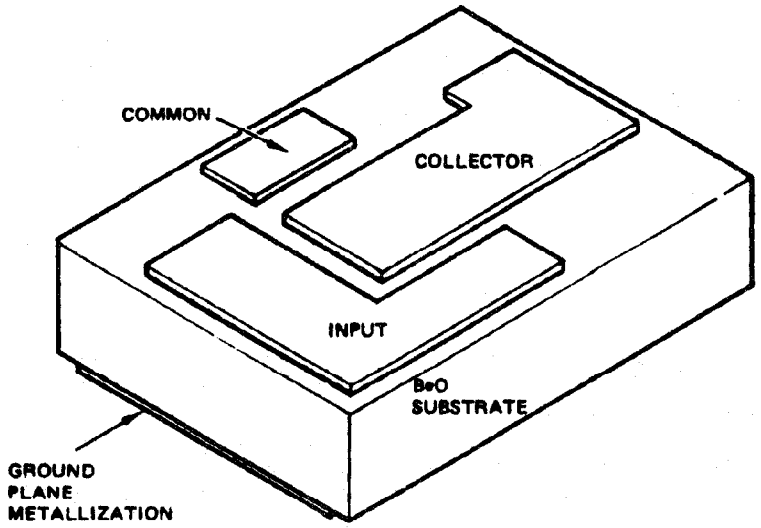


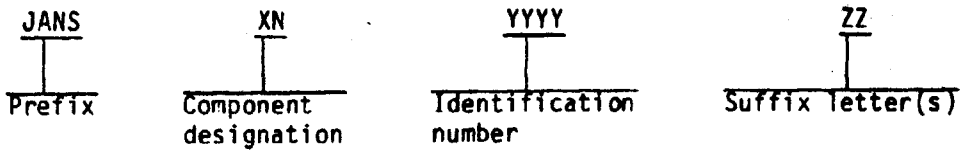
FIGURE 48. Microstrip chip carrier.

The advantage of the carrier over a chip for microcircuit purposes is simply described as follows. A chip-on-carrier is characterized after die attachment to the carrier. At microwave frequencies this is important because a device's electrical characteristics undergo a drastic change during the die attachment operation due to a piezoelectric effect. The extreme flexibility of the carrier makes it a desirable part in the microwave component repertoire.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS

6.3.4 Military designation. The military designation for microwave transistors is formulated as follows:



The prefix includes the level of product assurance. For NASA applications, the level of product assurance is restricted to JANTXV or JANS in accordance with MIL-STD-975. A radiation hardness assurance (RHA) code, if applicable, is placed after the prefix. See MIL-S-19500 for details.

The component designation is "2N" for microwave transistors.

The identification number is assigned in order of registration and has no other significance.

The suffix letter symbolically describes matched devices (suffix "M"), reverse polarity (suffix "R"), or any other letter to indicate a modified version.

6.3.5 Electrical characteristics. This section shall concern itself primarily with the basic concepts and considerations required for accurate rf measurements on microwave transistors. At frequencies greater than 1 GHz, the methods used for conventional transistors or h or y parameters yield no meaningful information. The following basic difficulties are evident:

a. H or y parameter measurements require open or short-circuited terminations and resonant lines. The resulting frequency dependency eliminates the advantage of swept frequency, broadband measurements.

b. Measurements must take into account the parasitic effects of the package of the transistor as well as the effect of transmission lines.

c. Bias networks that could cause the transistor to oscillate must not introduce reactive impedances into the circuit.

d. Small signal measurements are not easily adaptable to power devices.

6.3.5.1 Definition of S-parameters. Consider the two-port circuit in Figure 49. The response of this circuit, in s-parameter representation, establishes two independent variables:

$$a_1 = \frac{1}{2\,(Z_0)^{1/2}}\,(v_1 + z_0 i_1)$$

$$a_2 = \frac{1}{2\,(Z_0)^{1/2}}\,(v_2 + Z_0 i_2)$$

and two dependent variables:

$$b_1 = \frac{1}{2\,(Z_0)^{1/2}}\,(v_1 - Z_0 i_1)$$

$$b_2 = \frac{1}{2\,(Z_0)^{1/2}}\,(v_2 - Z_0 i_1)$$

where $Z_0 = Z_{01} = Z_{02}$ is the characteristic impedance of the transmission line.

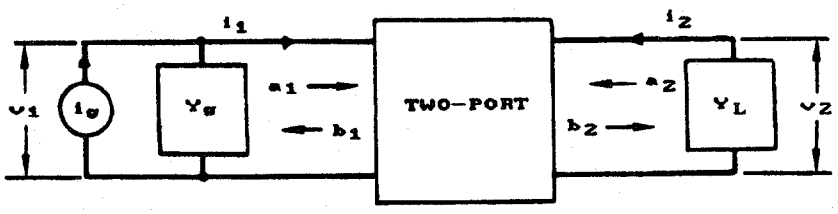


FIGURE 49. Two-port circuit.

From this the s-parameter matrix is defined as:

$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

6.3 MICROWAVE DEVICES, TRANSISTORS

or in matrix form:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

For simplicity, we will refer to a_1 and a_2 as incident voltage waves and b_1 and b_2 as reflected voltage waves. It can be seen that:

when $a_2 = 0$, $S_{11} = \frac{b_1}{a_1}$ and $S_{21} = \frac{b_2}{a_1}$

and

when $a_1 = 0$, $S_{12} = \frac{b_1}{a_2}$ and $S_{22} = \frac{b_2}{a_2}$

Typical block diagrams for measuring the forward and reverse transmission coefficients, S_{21} and S_{12}, and the forward and reverse reflection coefficients, S_{11} and S_{22}, are shown in Figures 50 and 51, respectively.

In Figure 50, a swept frequency source feeds a power divider. The outputs of the power divider feed both a reference channel and a test channel. A line stretcher is placed in the reference channel and the device to be characterized is placed in the test channel. The function of the line stretcher is to compensate for the extra electrical length of the transistor. A test set and analyzer then compares the two signals, and yields S_{21} for the forward direction and S_{12} for the reverse direction.

The reflection coefficients, S_{11} and S_{22}, are the most easily measured by the reflectometer technique shown in Figure 51. The swept frequency source feeds a directional coupler. The transistor is placed at the measurement port of the coupler and is terminated in a 50-Ω load. The incident wave is measured at the reference port of the coupler. The reflected wave is measured at the test port of the coupler. The analyzer again compares the ratios of the signals. Representative plots of S_{21}, S_{12}, S_{11}, and S_{22} can be seen in Figures 52 through 55.

6.3.5.2 Gain. While the measurements of the s-parameters of a transistor characterize it as far as transmission and reflection coefficients are concerned, these measurements do not yield the complete device performance picture. The measurement of gain, however, has several unique advantages. The value of gain is not affected by package parasitics. Furthermore, the location of a reference plane is of no consequence.

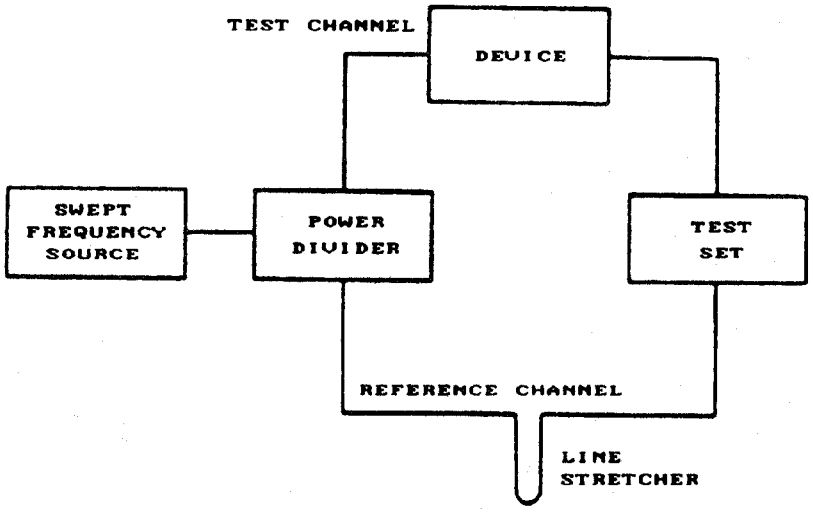


FIGURE 50. Block diagram for measurement of S_{12} and S_{21}.

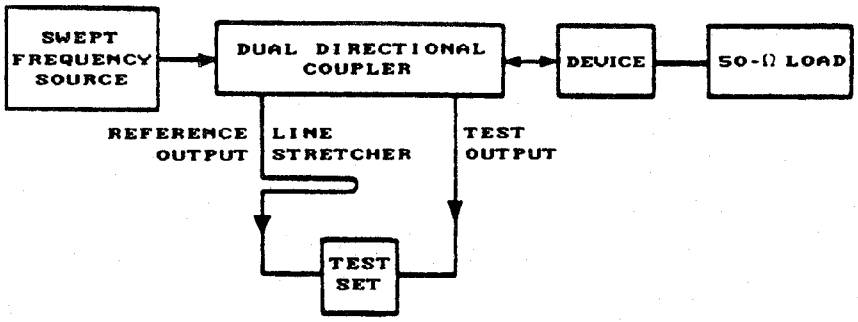


FIGURE 51. Block diagram for measurement of S_{11} and S_{22}.

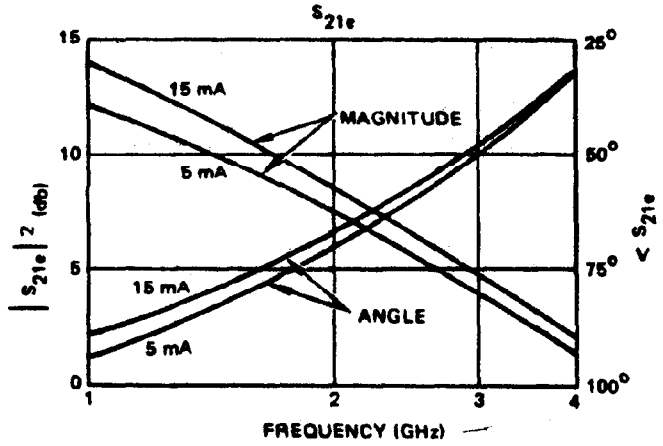
The gain can be measured under several different conditions. The definitions of these are as follows:

a. Maximum stable gain GMS is equal to

$$\left[|S_{21}|^2 / |S_{12}|^2 \right]^{1/2} \quad \text{or} \quad \left| \frac{S_{21}}{S_{12}} \right|$$

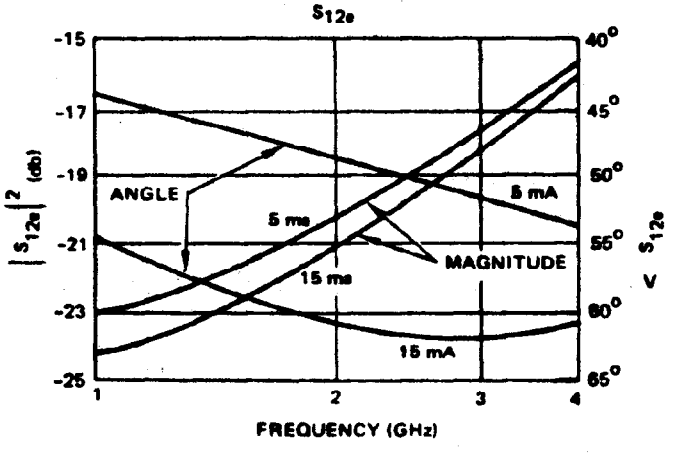
USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS



C_B = 7 V

FIGURE 52. S_{21} vs frequency.



V_{CB} = 7 V

FIGURE 53. S_{21} frequency.

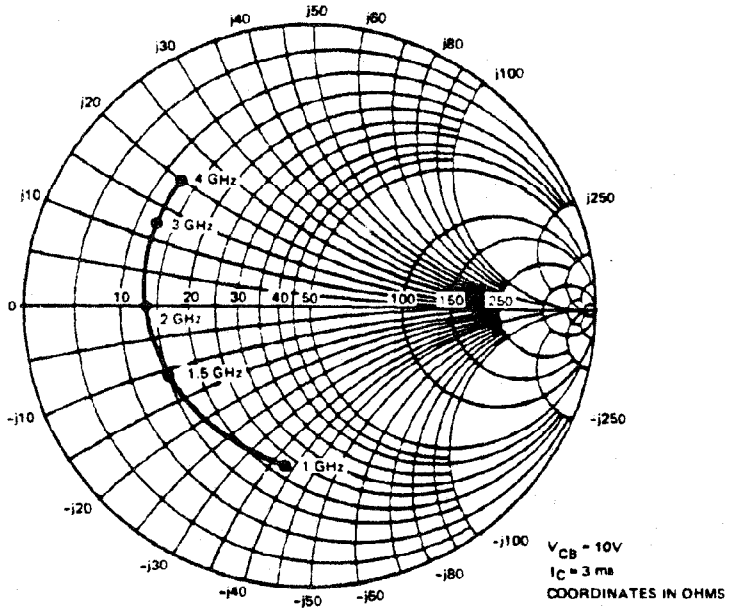


FIGURE 54. S_{11} vs frequency.

b. Maximum available (tuned) gain (G_{MAX}) is measured with the input and output conjugately matched. The only other condition is stability. The stability factor K must be greater than 1 (unconditional stability).

c. Unilateral gain (G_U) is the forward power gain of an amplifier whose reverse gain has been adjusted to zero.

d. Stability factor (K) determines the tendency of a transistor to oscillate. For $K > 1$, stability is guaranteed. For $K < 1$, oscillation can be induced by introducing reactive components in to the load.

Each of these four parameters can be calculated from the measured s-parameters. The circuit is tuned so the magnitude and phase of G_{MAX} are recorded. Reversal of ports A and B and repeating the procedure yields.

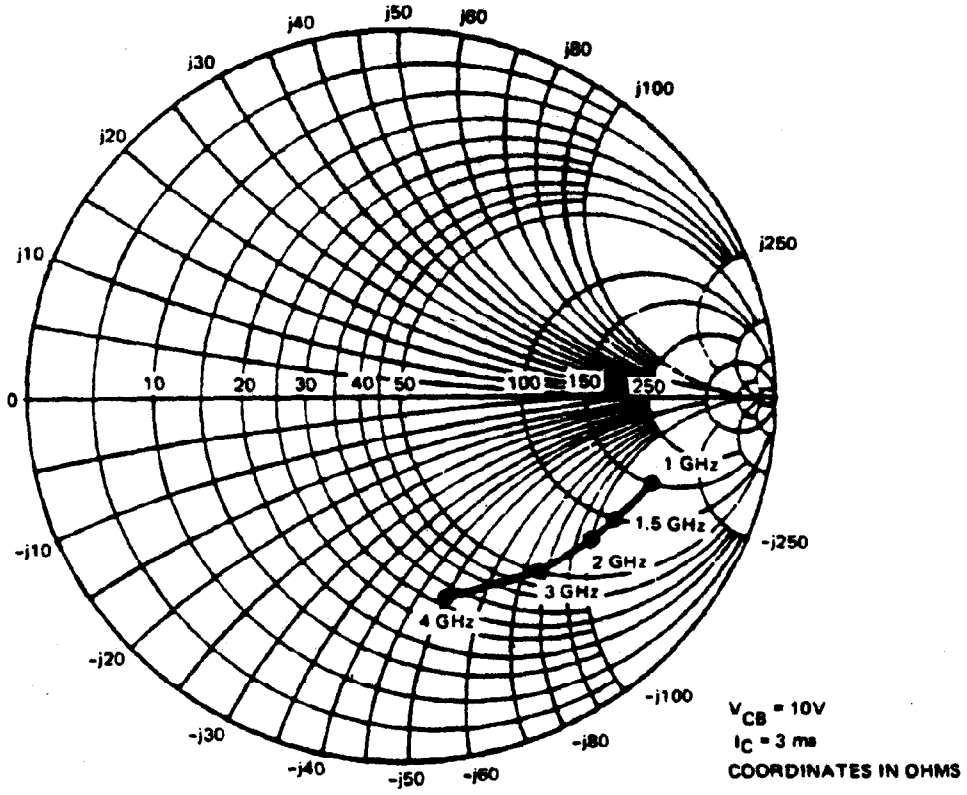
6.3 MICROWAVE DEVICES, TRANSISTORS

Then $G_{MS} = \left| |S_{21}|^2 \quad |S_{12}|^2 \right|^{1/2} = |S_{21}/S_{12}|$

$$K = \frac{1}{2}\left[\frac{G_{MS}}{G_{MAX}} + \frac{G_{MAX}}{G_{MS}}\right] \quad \text{FOR } K > 1$$

$$G_U = \frac{G_{MS} - (2\cos\theta) + (G_{MS} - 1)}{2(K - \cos\theta)}$$

(θ is the phase difference recorded during the measurement of $|S_{21}|^2$ and $|S_{12}|^2$.)



V_{CB} = 10 V
I_C = 3 mA
Coordinates in ohms.

FIGURE 55. S_{22} vs frequency.

In addition to the swept frequency measurement of s-parameters and the various fixed frequency gain measurements, one other family of curves is important. This is the G_{MAX} as a function of actual rf power output. A typical family of curves is shown in Figure 56. The flat areas of each curve represent regions of uniform gain over a wide range of power output.

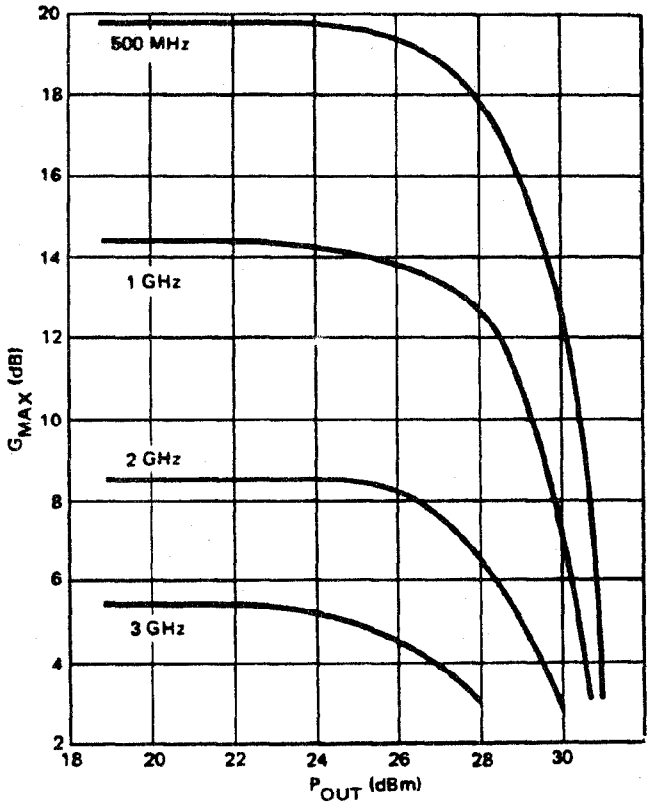


FIGURE 56. Gain vs power output.

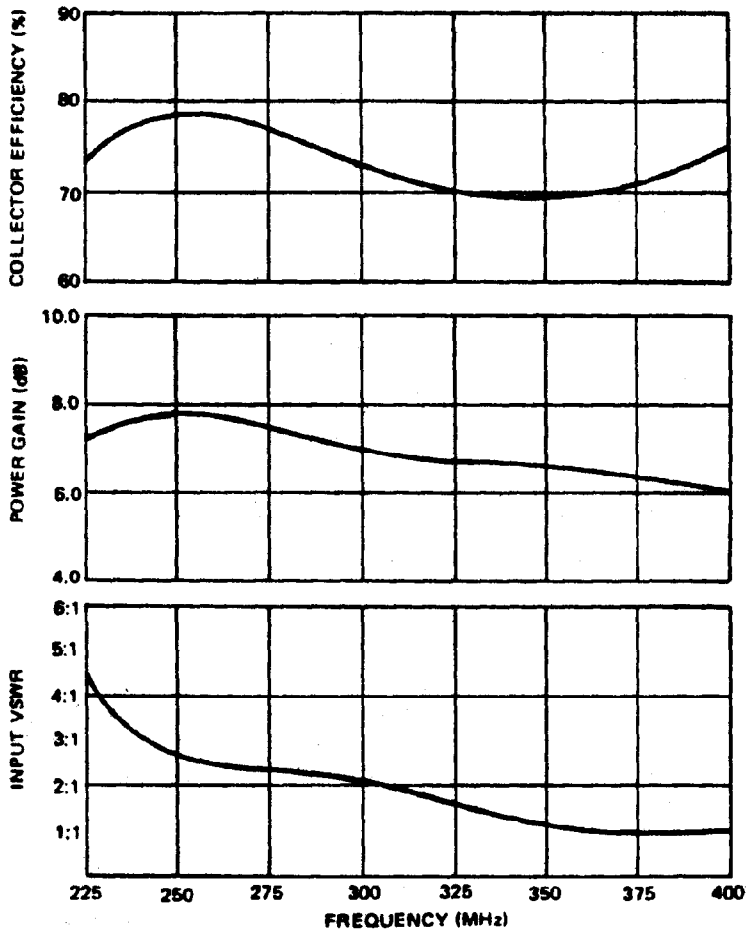
6.3.5.3 Noise figure. The measurement of noise figure is especially applicable to small signal devices. This parameter is a measure of the device's sensitivity to low level signals. It is a function of many variables, including frequency, source impedance, and current level.

Thus far, the discussion has been limited to the measurement of discrete devices. Of equal importance is the measurement of devices in a circuit. Because most microwave amplifiers are broadband, swept frequency measurements can be used to a great extent. Primary response characteristics include collector efficiency, input and output VSWR, power gain, and noise figure.

Collector efficiency is a term more commonly used for a power amplifier than a small signal amplifier. The major factor in determining efficiency is the class of operation. Class A amplifiers, because they draw current and dissipate heat 100 percent of the time, have the lowest efficiency, about 20 to 25 percent.

6.3 MICROWAVE DEVICES, TRANSISTORS

In Class C operation the transistor is biased so conduction occurs only when a signal of sufficient magnitude is fed into the input. The quiescent current is zero when the driving signal is not present. Therefore, because current flows for less than 50 percent of each cycle, the collector efficiency can be as high as 70 to 80 percent. See Figure 57 for a typical plot of efficiency in a Class C amplifier over a frequency range.



P_{OUT} = 30 W
V_{CC} = 28 V
T_{CASE} = 35 °C

FIGURE 57. Typical broadband response of a UHF power amplifier.

The amount of reflected power from the input of an amplifier is best demonstrated from the measurement of the VSWR. Several factors can affect the VSWR of a device in a given circuit. Most critical is the impedance matching both the input and output.

The input impedance of a power amplifier is generally low compared with the output impedance of the preceding driver. The real (resistive) component is inversely proportional to the area of the transistor on the order of 0.5 to 5 Ω. The imaginary (reactive) component of the input impedance is determined by the inductance of the package and the input capacitance of the device. At microwave frequencies, inductive reactance is dominant.

The output impedance is an extremely important factor in determining the amount of power gain of an amplifier stage. Although an optimum match is not always achieved due to other circuit considerations, the amount of mismatch adversely affects both gain and VSWR. Figure 57 depicts typical variations in gain and VSWR for an amplifier over a frequency range.

6.3.5.4 Amplifier parameters.

6.3.5.4.1 Intermodulation intercept. The intermodulation intercept is an expression of the low-level linearity of the amplifier. The intermodulation ratio (IMR) is the difference in decibels between the fundamental output signal level and the generated distortion product level. The relationship between intercept and intermodulation ratio is illustrated in Figure 58, which shows product output levels plotted versus the level of the fundamental output for two equal strength output signals at different frequencies. The upper line shows the fundamental output plotted against itself with a 1 dB to 1 dB slope. The second and third order products lie below the fundamentals and exhibit a 2:1 and 3:1 slope, respectively. The intercept point for either product is the intersection of the extensions of the product curve with the fundamental output.

The intercept point is determined by measuring the intermodulation ratio at a single output level and projecting along the appropriate product slope to the point of intersection with the fundamental. When the intercept point is known the intermodulation ratio can be determined by the reverse process. The second order IMR is equal to the difference in decibels between the second order intercept and the fundamental output level. The third order IMR is equal to twice the difference between the third order intercept and the fundamental output level. These are expressed as:

$$IP_2 = P_{out} + IMR_2$$

$$IP_3 = P_{out} + 1/2\ IMR_3$$

where P_{out} is the power level in units of decibel above a millwatt (dBm) of each of a pair of equal level fundamental output signals, IP_2 and IP_3 are the second and third order output intercepts in dBm, and IMR_2 and IMR_3 are the second and third order intermodulation ratios in decibels.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.3 MICROWAVE DEVICES, TRANSISTORS

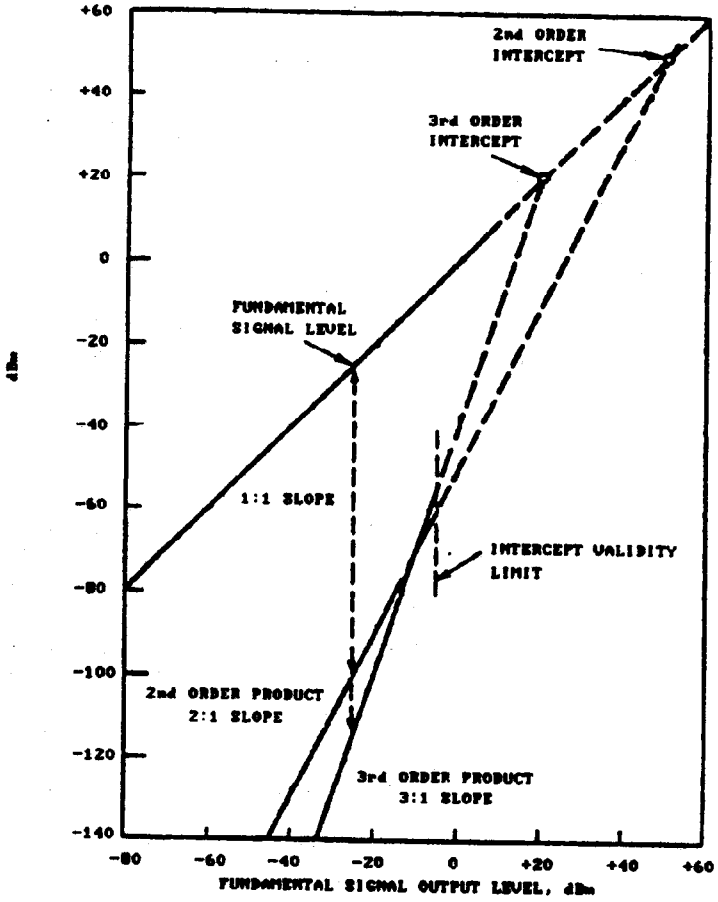


FIGURE 58. Intercept diagram.

The intermodulation intercept is a valid indicator of intermodulation performance only in the small signal operating range of the amplifier. Above some output level which is below the 1 dB compression point, the active device moves into large signal operation attended by signal strength dependent bias level shifts. At this point, the intermodulation products no longer follow the straight line output slopes, and the intercept description is no longer valid.

The intermodulation ratios are determined by measurement using a conventional spectrum analyzer. The measurement dynamic range is enhanced using appropriate cancellation techniques when required to accommodate high dynamic range amplifiers.

6.3.5.4.2 Compression. The 1 dB compression point is the output level at which the amplifier gain drops 1 dB below its small signal value. It is an indication of the signal level at which small signal conditions no longer apply. At this level the intermodulation intercepts no longer adequately predict the amplifier distortion behavior.

The measurement of the compression level is made using a frequency discriminating detector to detect the gain decrease at the fundamental frequency. This is preferable to using a total power detection measurement, since the fundamental frequency is probably of most interest to the user.

6.3.5.4.3 The most obvious conclusion that can be made regarding device selection is that circuit requirements should govern device selection, and not vice versa. Two seemingly identical devices will not necessarily perform comparably in a given circuit. Therefore, it is the entire network of transistor, package, parasitics, microstrip runs, and other components that must be considered as one fundamental unit.

The importance of S-parameter measurements cannot be overemphasized. S-parameter characterization gives the design engineer specific information at microwave frequencies on the performance of the device. Coupled with computer aided circuit design techniques, it enables a designer to optimize his circuit rapidly and accurately.

Complete S-parameter characterization is usually only meaningful for small signal devices. Power microwave devices are considered differently. The most important characteristic of a power device is its ability to operate in a mismatched load. Response curves for collector efficiency and junction temperature, as a function of frequency for various values of VSWR, can be considered to be of primary importance to a designer.

6.3.6 Environmental considerations. Typical environmental conditions and screens that microwave transistors are capable of withstanding can be found in MIL-S-19500. For the specific device selected, consult the applicable MIL-S-19500 reference sheet.

More information on environmental considerations can be found in section 3, Resistors, paragraph 3.8.7, Reliability considerations, and paragraph 3.1.7, General reliability considerations.

6.3.7 Reliability considerations.

6.3.7.1 Introduction. Until a few years ago there was a void in the transistor market area in the rf and microwave regions. This void is gradually being filled by the industry today using both old and new techniques: interdigitated, overlay, and metal matrix geometries. However, due to the lack of the volume in sales in this area, there has been only a minimum effort to fully understand the failure mechanisms associated with these devices and the necessary corrective action required to significantly reduce the failure mechanisms.

6.3 MICROWAVE DEVICES, TRANSISTORS

The dominant failure mechanisms of a device vary according to the peculiarities of the design, fabrication process, and application of the device. Following is a list of the major failure mechanisms for rf and microwave transistors.

a. Aluminum migration

b. Die attach failure

c. Metal-over-oxide-step coverage

d. Lead bond failures

e. Device inability to operate into a mismatched load.

6.3.7.2 Failure mechanisms for rf and microwave transistors.

6.3.7.2.1 Aluminum migration. This well documented phenomenon relates to the migration of aluminum in the presence of high temperature and high current densities. The transport of mass in metals when the metals are stressed at high current densities was recently recognized as a potential wear-out failure mechanism in semiconductor devices. On most semiconductor devices this failure mode is indicated by an electrical open due to voids or localized loss of conductor metal. In the case of a transistor, this normally begins on the emitter fingers. When one of the emitter fingers opens due to this migration, the adjacent fingers are required to carry this additional current. These fingers, now operating at an increased current, will accelerate their own failure rates, due to the increased current. Figure 59 illustrates this failure mechanism.

However, due to the shallow diffusion depth required on high frequency transistors, this failure mode can, on occasion, also be indicated by a shorted emitter-base junction. Figure 60 shows the cause of this shorted condition.

Studies have shown that aluminum migration is dependent on current density, surrounding material, and aluminum grain structure. Consequently, to reduce the possibility of aluminum migration, these phenomena must be closely scrutinized in the selection and application of rf and microwave transistors.

The current density problem is directly related to the cross-sectional area of the metallization; that is, the smaller the cross section, the smaller the MTBF. J.R. Black, in his paper on "Electro Migration Failure Modes in Aluminum Metallization for Semiconductor Devices," in the IEEE, Volume 57, No. 9, September 1969, used the following equation as the basis for calculating MTBF for large-grained films 6000 A thick, with the pre-exponential constant normalized to the conductor cross-section areas of $9.65 \times 10^{-8} cm^2$.

$$MTBF = \frac{(w)\ (t)\ \exp\ (\emptyset/kT)}{(5 \times 10^{-13})\ J^2}$$

w = conductor width in centimeters

t = conductor thickness in centimeters

Ø = an activation energy in electron volts, for large grained films 0.84

k = Boltzman's constant

T = film temperature in degrees Kelvin

J = current in amperes per square centimeter

It can readily be observed from this equation that the MTBF of the film is directly related to the cross-section (w and t) of the conductor.

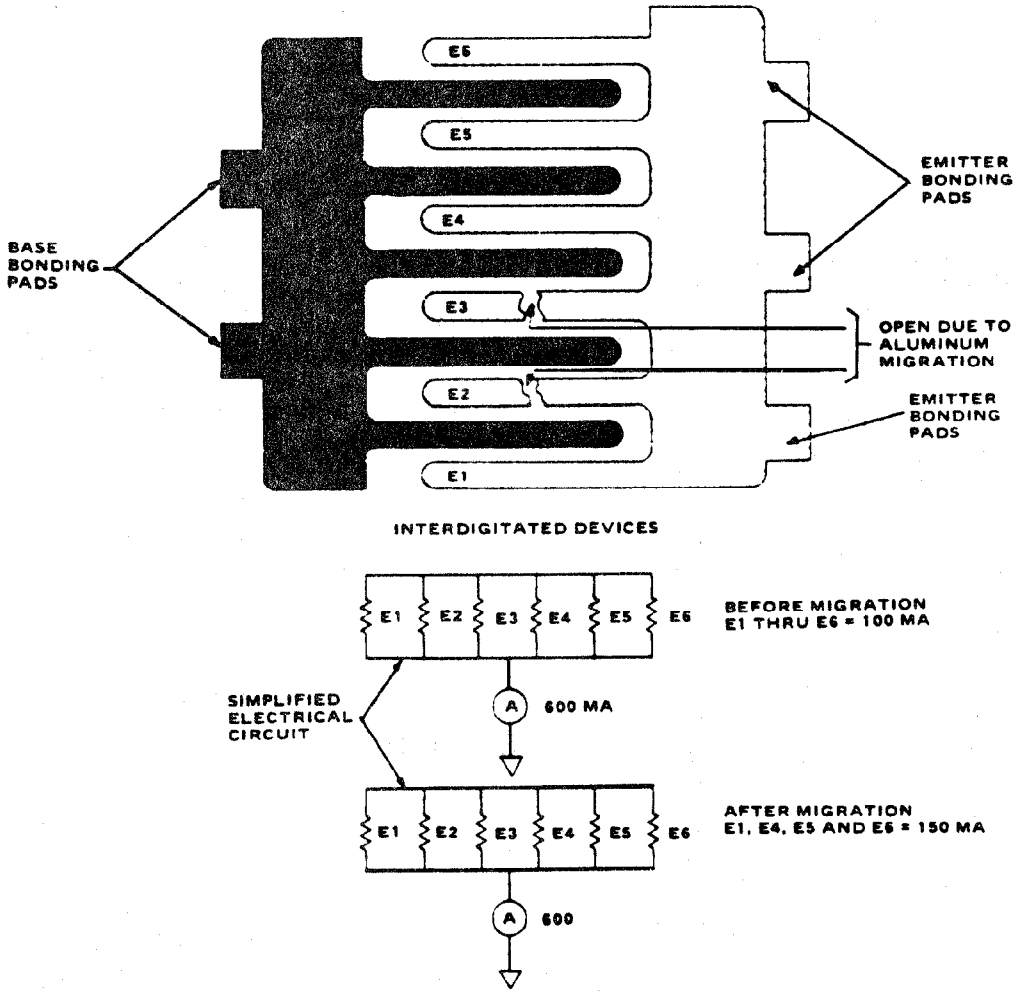


FIGURE 59. Current density, secondary problem.

6.3 MICROWAVE DEVICES, TRANSISTORS

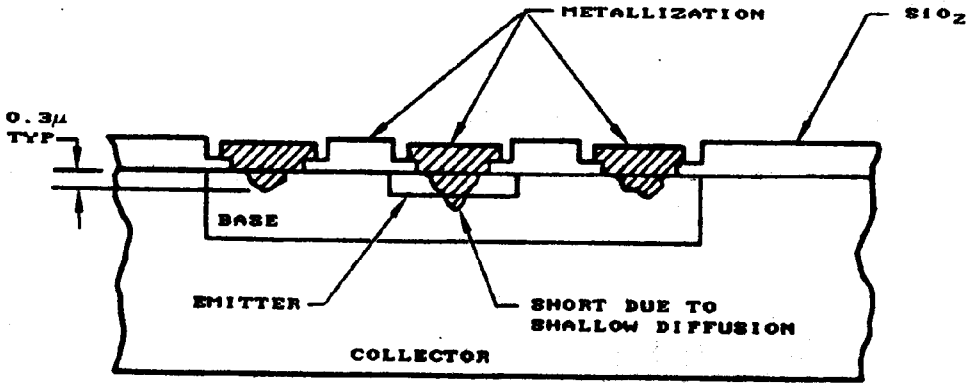


FIGURE 60. Short due to shallow diffusion.

6.3.7.2.2 Die attach failure. This defect is prevalent in any power transistor because of problems associated with attaching a die that has a large mass area; it is difficult to attach a die to the header without leaving any voids between the die and header. Figure 61 shows a typical cross section of a die bonded to the header. An X-ray photograph taken through the bottom of the header of a transistor package clearly shows the voids between the header and die. However, X-ray testing is potentially destructive and, therefore, cannot be used as a 100 percent screen.

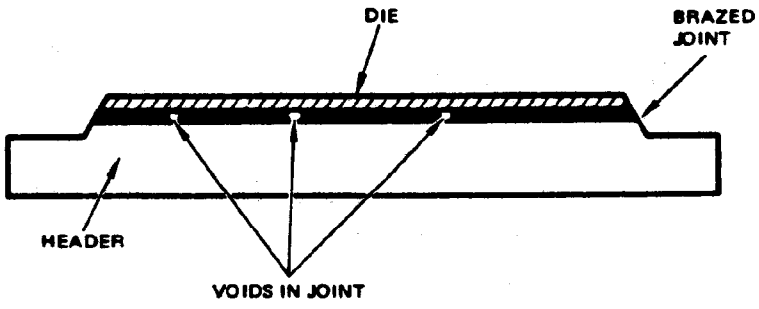
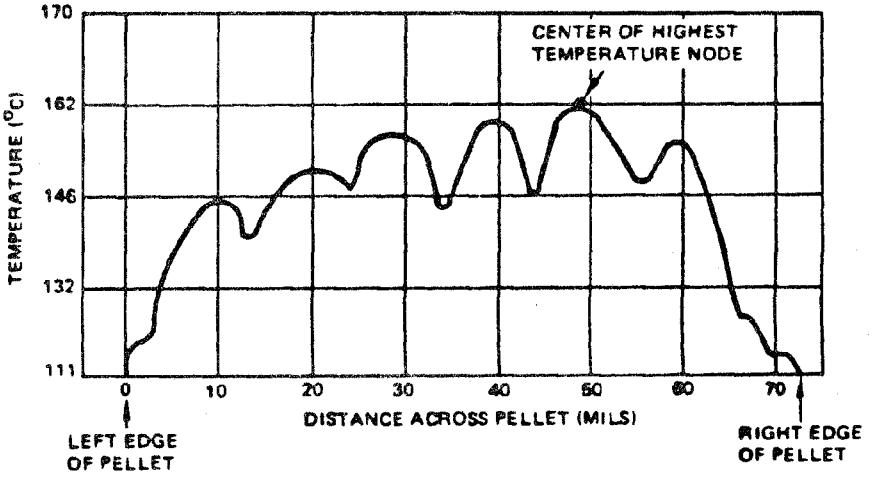


FIGURE 61. Cross-section of a die bonder to a header.

These voids cause hot spots on the surface of the die, due to current injection, which, in turn, can cause thermal runaway. The measurement frequently used to characterize the transistor thermal properties is thermal resistance, θ_{JC} or θ_{JA}. These characteristics, θ_{JC} and θ_{JA}, are normally calculated by techniques that use average die temperature to arrive at their values. The result is an average value of thermal resistance when, in actuality, the thermal profile across the die varies greatly. Figure 61 shows a typical thermal profile across a die in operation. If a void is beneath a section of the die, as indicated in Figure 61, this will cause surface hot spots that cause greater variations than those shown in Figure 62.



f = 400 MHz
P_O = 25 W
V_{CC} = 28 V
T_{CASE} = 85 °C
VSWR = 1.0

FIGURE 62. Thermal profile of a die during operation.

Two main approaches have been taken to overcome the problems associated with these voids. The first and most significant is employing current ballast resistors on the multiple emitter fingers (refer to Figure 30). With this structure the effects of a hot spot can be cancelled by the voltage drop developed across the ballast resistor, which feeds current to the emitter sites.

The use of ballast resistors has been successful in significantly reducing the failure rate due to hot spots on transistors operating below 100 MHz on Class A and B power amplifiers.

6.3 MICROWAVE DEVICES, TRANSISTORS

The second action being taken by the industry to reduce the effect of hot spots is improving the initial thermal resistance measurements. This improvement is being completed using infrared scanning techniques. This technique arrives at a worst case thermal resistance instead of an average value as the older method does. Therefore, if the vendors use these values for rating their devices, the users will be able to realistically design their circuits with these ratings.

Since infrared measurement is a sample test done only periodically for economic reasons, this measurement does not take into account the day-to-day variations of die attaching. Therefore, die attach failure remains an industry problem.

6.3.7.2.3 Metal-over-oxide-step coverage. Surface metal interconnects are subject to some of the more insidious failure modes. The more common symptom of metal problems is the total or intermittent open caused by fracturing or migration. It is now an established fact that aluminum migrates as the current density in the metal approaches 10^6 A per cm squared. Fracturing can be caused by thermal mismatch between the metallization and the silicon or silicon dioxide substrate. Thinning of the metallization at the metal-over-oxide step is caused by the method of metallizing the transistor.

In electron-beam evaporation, the source from which the metal is deposited is normal to the surface of the wafer. As a result the thickness of the metal over the oxide steps is far thinner than the metal deposited on the flat surface. It is much easier to measure the thickness of the metal on the flat surface. Therefore, all control of metallization revolves around the surface thickness. A typical cross section of a contact area is shown in Figure 63. It can be observed that at the metal-over-oxide step the metallization is only approximately 10 percent of the surface metal thickness. When the metal is caused to expand and contract by current heating or operating at thermal extremes, micro cracks can develop, causing an open or intermittent open.

Many approaches have been taken to improve the thickness of metal-over-oxide steps. Two of the approaches contour the oxide steps. They are: 1) tapering of the oxide steps, and 2) having a double step at each oxide step. Although tapering the oxide step by controlling of the etch rate is the better approach, it requires additional etching steps. If optimized properly, the results are worth the extra time and cost. Figure 64 shows a cross section of contact area with tapered oxide. As the chip complexity increases, especially with double metallization devices, this tapering of oxide becomes very important.

The other method, a double step, accomplishes the oxide height minimization by growing the oxide in two steps which require additional masking steps. This is also shown in Figure 63.

Also, the industry has been depositing the metal at more than one angle by electron-beam evaporation while universally rotating the wafers. This has helped to limit this problem. However, the metal can still migrate down into the contact area because current density causes the metal-over-oxide step to

become the weak area again. Therefore, most integrated circuit manufacturers have modified their design rules to allow a maximum current density of 10^2 A per cm squared. This rule is based on the assumption that the metal thickness over-oxide step is only 10 percent of the surface metal thickness. However, due to the miniature geometry required for high frequency transistors the current density runs in the range of 10^5 A per cm squared.

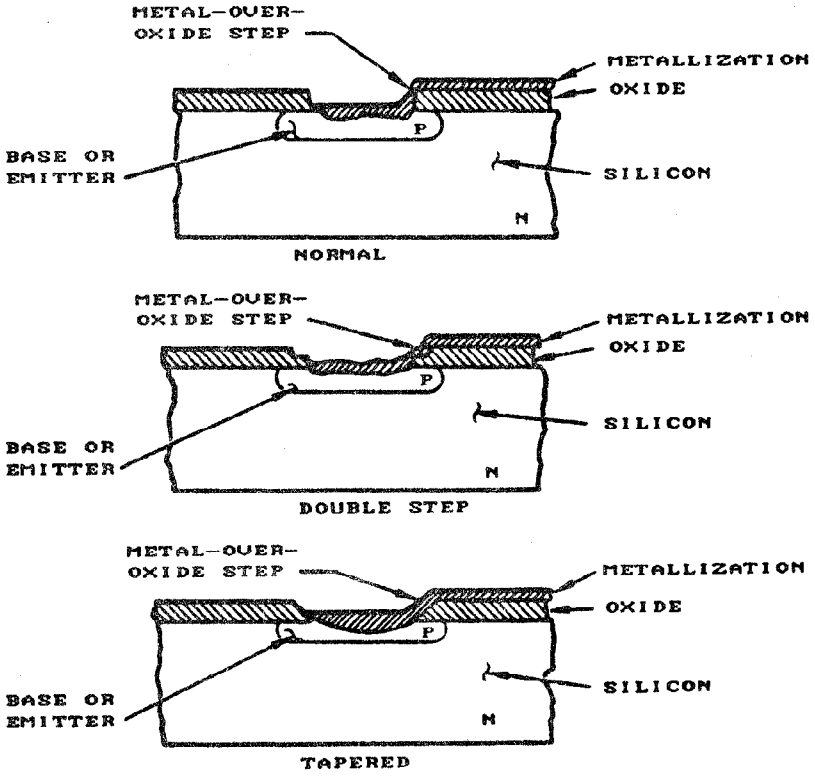


FIGURE 63. Cross-section of a contact area with tapered and double-stepped oxide.

The industry is presently developing metal systems other than aluminum for use on rf and microwave transistors. These metal systems are of refractory metal and, due to their high atomic weight, resist mass transport by way of momentum exchange.

Any other cause for the reduction of metal cross section can generate the same problems. Most common are photolithographic errors, scratches in the metal, and metal too thinly deposited. It is generally conceded that metal which opens due to thermal mismatch is caused by severely reduced cross section.

6.3 MICROWAVE DEVICES, TRANSISTORS

Therefore, it can be assumed that improvements that increase metal thickness over steps will solve both failure modes simultaneously. The only other significant failure mode in the surface metallization itself is lack of adherence to the substrate either on the surface or down in the oxide cutouts. Lifting of metal can be caused either by residual contaminants present during metallization or by severe overalloying of the metal after deposition. These types of failure modes should be screened out by normal 100 percent tests such as burn-in and temperature cycling.

6.3.7.2.4 Internal lead bond failures. Defective internal lead bonds traditionally have been a large cause of device failure. The majority of these devices had seen a 100 percent screening, which included constant acceleration and temperature cycling. Lead bond failure will be more prevalent in rf and microwave transistors due to: (1) the increase in the number of bonds within a package, (2) the smaller bonding pads, and (3) the smaller wire diameter required.

Generally, the industry has settled on aluminum-aluminum and gold-aluminum metal systems for transistors. Both methods are satisfactory, although both can fail under certain circumstances.

Gold thermal compression bonding is inherently a stronger bonding method than the aluminum ultrasonic bonding technique. However, the gold ball bond is more massive and, therefore, more sensitive to high-g shock. An ultrasonic bond has a narrow cross section at the heel of each bond and it is difficult, if not impossible, to examine the depth of the wire deformation (Figure 64). A ball bond inherently exhibits no deformation at one end of the bonding wire, but the bond made at the package end suffers the same problem as an ultrasonic bond.

During storage, there is no concern over "purple plague" or other bimetallic formations using a gold bond method. These formations occur rapidly at temperatures above 200 °C. Below that temperature the bimetallic formation rate is not easily measured. The bimetallic formation or "purple plague" is of course not a failure mode set in aluminum bonding. Hermetic seal leakage that allows moisture to enter the package is very likely to destroy either class of bond, since the aluminum metallization corrodes regardless of bond type.

Presently, the industry uses three types of screens to reduce this cause of device failure: 1) constant acceleration, 2) temperature cycling, and 3) power cycling. However, these screening techniques are only partially effective; therefore, this failure mode will continue to plague the industry.

6.3.7.2.5 Mismatched loads. In rf and microwave transistor Class C power amplifiers, failures are readily caused when the output circuit is detuned at full power. This can happen in practical applications. The resulting mismatch causes a high VSWR condition at the collector of the transistor, subjecting the transistor to instantaneous voltage peaks many times the supply voltage. These voltage peaks will cause avalanching, which takes place within the collector depletion region. Therefore, there is no "current steering" effect by the emitter ballast resistors to reduce the localized current densities in the

collector depletion region. Some devices are said to be capable of operating into an infinite VSWR; however close investigation of their specifications shows that this can only be done when the devices are derated greatly to accommodate the instantaneous voltage peaks.

The only form of corrective action to be approached by the industry is to use "collector ballasting" provided by a thick, undepleted collector layer. However, due to its effect on the electrical characteristics of the device, this approach is very limited.

6.3.7.2.6 Radiation effects. Radiation-induced damage in the form of permanent damage to the semiconductor crystal dominates other radiation-induced changes in microwave transistors. The type of damage produced by energetic electrons and gamma rays are called point defects because they are spread uniformly throughout the crystal. In this type of particle-atom collision an atom of semiconductor material will be displaced to an interstitial position, leaving behind a vacancy in the crystal lattice. The resulting vacancies and interstitials are mobile and will tend to unite with themselves or impurities in the crystal and form defect complexes. Different kinds of defect complexes are formed, depending on the temperature.

Neutron radiation can displace many more lattice atoms. As a result regions of localized damage occur. These regions are called defect clusters. Curves of the total number of defects per cubic centimeter of semiconductor per neutron per centimeter squared have been calculated. The effects of permanent radiation damage show on the dc, low frequency, and microwave properties of microwave transistors.

Low frequency radiation and dc effects have been thoroughly studied. The predominate effect is a decrease in current gain (β) due to recombination in the base region. For microwave transistors, the base width is sufficiently narrow that this effect can be neglected at neutron fluences up to 10^{17} n/cm^2. This value is moderated by degradation of current gain caused by excessive recombination in the base emitter space charge region and emitter bulk regions.

Permanent radiation effects on the microwave properties of small-signal bipolar transistors can be divided into two categories: 1) changes in the s-parameters, and 2) changes in noise performance.

The most significant changes are seen in the forward transfer gain (S_{21}). One study showed the f_t of a bipolar microwave transistor dropped from 4.5 GHz preradiation to 540 MHz postradiation. This was caused by a drastic increase in T_b due to a phenomenon similar to base pushout. The f_t increased from 540 MHz to 3 GHz when the collector bias was increased from 15 V to 30 V. Thus, the radiation tolerance of this type of transistor can be improved by operating it at a higher collector voltage and at lower collection currents.

6.3 MICROWAVE DEVICES, TRANSISTORS

Experiments with small signal bipolar microwave transistors in an amplifier circuit have shown that the noise figure is fairly insensitive to neutron radiation. This is because the noise mechanisms introduced by neutron damage only contribute excess noise at low frequencies. The increase in the noise figure that does occur can be attributed to the decreased current gain (βo).

The general military specification for semiconductor devices specifies four radiation hardness assurance levels (RHA) for JANTXV and JANS product assurance levels. Microwave transistors may specify RHA designation M, D, R, or H in accordance with the requirements of the military specification.

The major failure mechanisms on rf and microwave transistors, listed in order of importance are as follows:

a. Aluminum migration

b. Device inability to operate into a mismatch load

c. Die attach void failures

d. Internal lead bond failures

e. Metal-over-oxide-step coverage.

6.4 Microwave attenuators and loads.

6.4.1 Introduction. Attenuators and loads, which are known as dummy loads and terminations, are similar devices. A load terminates a line and dissipates the signal with low reflections. An attenuator lowers or attenuates the transmitted signal and has low reflections.

An attenuator is a device used to attenuate rf signals a given amount without introducing mismatch when operating within its specified frequency range. Frequency range is limited by the connection mechanism and the match of the dissipative element to the transmission media. Broadband devices are only achieved using resistive elements.

A load is a device used to match the transmission media and dissipate all signals presented at the load. A load used to terminate a line in place of additional circuitry is called a dummy load. A load that is designed into a circuit to terminate an unwanted signal path is called a termination.

6.4.2 Usual applications.

6.4.2.1 Attenuators. Attenuation, like insertion loss, is not usually desired in microwave design. In those cases where signal levels must be adjusted down to lower levels, attenuators or directional couplers can be used. Attenuators present low mismatch to the source, attenuate the signal a given amount, introduce little delay or phase shift, have relatively little change in attenuation, and present a low mismatch to the load. Attenuators may be coaxial, waveguide, or strip-transmission media.

Most present coaxial attenuators are designed to operate at 1 W or less from dc to 18 GHz, with some units available to 26 GHz. Lossy material attenuators can be made for slightly higher power but are bandwidth limited. High power attenuators use fins for air cooling or are liquid cooled. Variable attenuators are made using resistive or lossy elements, which are movable and coupled to the rf field. Such units have less than an octave bandwidth.

Waveguide attenuators are usually made of resistive material matched to the waveguide, with the resistive material mounting structure causing slight mismatch and small frequency variable phase shift. Variable waveguide attenuators can be made by changing the position of the resistive card in the waveguide. Precision variable attenuators are available for laboratory use only.

Attenuators are used to lower the power into amplifiers, detectors, and mixers to maintain operation in the linear region and reduce reflections at the input. Attenuators are used in the laboratory to substitute a specific amount of loss in a microwave circuit. Fixed attenuators (commonly called pads) ranging from 0 to 3 dB in 0.5 dB steps allow adjustment of signal levels in two paths to the same level. Attenuators can be used as matching devices where loss of power is not a concern. A 10 dB attenuator terminated in a short circuit returns a signal of -20 dB from the input signal that is equal to an input VSWR of 1.2:1.

6.4 MICROWAVE DEVICES, ATTENUATORS AND LOADS

6.4.2.2 Loads. Coaxial loads are available with various connectors and with frequency ranges as high as 18 GHz. Stripline loads and flange wall-mounted loads are available for use in stripline and microstrip applications. Power limitations are based on the dissipation element and its heat sinking. Ratings of 1 W are typical for most coaxial loads with some air cooled units available below 10 GHz at higher powers. Lossy elements are rarely used in coax except for a few applications at a few watts over limited frequency ranges. Typical applications are in loaded circulators and as terminations for hybrids and directional couplers.

Waveguide loads are predominately lossy material types and can have kilowatt ratings with liquid cooling. Stepped transition lossy material inserts are widely used in waveguide couplers. Waveguide loads are also used for circulators, hybrids, duplexers, and system self test circuits.

6.4.3 Physical construction

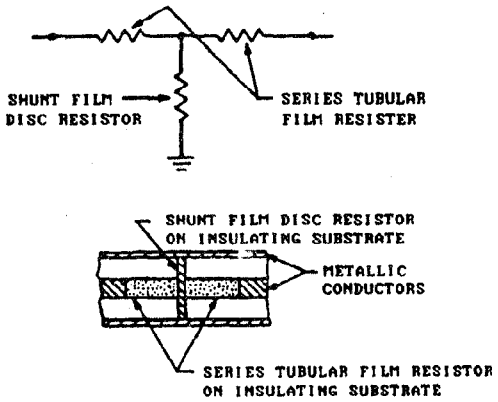
6.4.3.1 Attenuators. Coaxial attenuators construction has evolved from "Pi" and "T" types using rod and disk resistors to thin film resistor material on a substrate, Figure 64. Connector and transmission line matching into the resistive element allows VSWRs less than 1.3:1 at 18 GHz and attenuation flatness with frequency of less than 10 percent (typically ±0.50 dB) of attenuation value.

Variable attenuators are achieved by slot-line techniques with a lossy element movable in the slot. Frequency limitations are inherent in the length of the slot and the contour of the element.

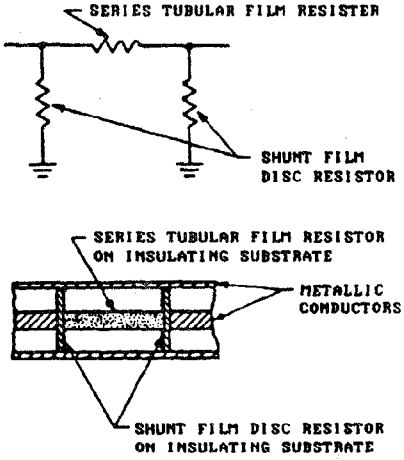
Power limitations are based on the heat capacity of the element, its thermal resistance, and adequate heat sinking.

6.4.3.2 Loads. Loads are made in a similar manner as attenuators since a terminated attenuator is a load. Most coaxial loads use rod-type resistors with one end soldered to the center conductor and the other end soldered into the outer conductor. Lossy material coaxial loads are shaped material placed at the end of a coaxial launcher. Waveguide loads are usually lossy material formed to match the waveguide size. The load material can be mechanically held by epoxy or screws, or both. Tapered lossy material may be introduced directly to strip transmission media to terminate a line.

6.4.4 Military designations. The military selection standard for attenuators is MIL-STD-1352. The military selection standard for loads is MIL-STD-1637. Military attenuators are found in MIL-A-3933 and MIL-A-24215. Military loads are found in MIL-D-3954 and MIL-D-39030.



A. Coaxial TEE section



B. Coaxial PI section

FIGURE 64. Construction of coaxial attenuators.

6.4 MICROWAVE DEVICES, ATTENUATORS AND LOADS

6.4.5 Electrical characteristics.

6.4.5.1 Attenuators. Attenuation is the primary parameter of an attenuator. The device is designed to have a nominal attenuation with a tolerance. A typical part may be specified as 30 dB ±/0.50 dB at center frequency. Frequency sensitivity of such a device could be 1.0 dB from dc to 18 GHz for coaxial attenuators and 0.50 dB for waveguide. VSWR is also an important parameter. For coaxial devices the VSWR mismatch is set by the interface (connector) and the transition into and out of the dissipative element. VSWR usually increases with frequency and a 1.1:1 VSWR at 100 MHz may increase to 1.3:1 at 18 GHz. In this range, the values of VSWR will have periodic peaks as the two primary sources of reflections (connector and element transition) add in phase. Waveguide devices do not have the pronounced variations of the coaxial attenuators due to limited frequency coverage.

Either coaxial or waveguide attenuators will have both attenuation and VSWR changes with either temperature or power, or both. This is due to the resistance change of the element with temperature. VSWR changes are primarily affected and attenuation changes are secondarily affected. Changes due to power in an application is really a misapplication of the part. Proper derating with temperature should be observed. Applications requiring high peak power with low average power should be tested over frequency and altitude to assure no arcing.

6.4.5.2 Loads. The primary parameter of a load is VSWR. The power rating of the load is also important. These parameters are inherent to the matching of the dissipative element to the transmission media and the resistance in the thermal path to the heat sink.

6.4.6 Environmental considerations. Environmental considerations for loads and attenuators are mostly mechanical in nature. In coaxial devices, the stresses on the load or attenuator element result from movement of the center conductor at the interface. The body of the element needs to be supported to prevent lateral movement. Heat removal from the element is also important. Temperature cycling and vibration fatigue are the most stringent environments. Elements made of nichrome should be avoided unless the nichrome is adequately protected from moisture (disappearing resistor phenomena).

6.4.7 Reliability considerations.

6.4.7.1 Failure modes. Attenuator and load failures are usually mechanical in nature. Separation of the element at its interface is a predominate failure mode. Temperature induced changes in the dissipative element may occur if the element is not adequately heat-sunk. In high-peak power applications, arcing may occur due to spacing, burrs, or as a result of power and altitude conditions. Some high power waveguide loads may corrode due to brazing salts left in the device during manufacturing.

6.4.7.2 Screening. Because the major failure modes are mechanical, temperature cycling is an effective screen to assure compatibility of materials. Electrical characterization of VSWR and attenuation (if appropriate) before and after such a screen can detect potential failures. Resistance measurements (center conductor to ground and to the other center conductor) are much more sensitive to changes than a VSWR test. Delta limits should be imposed. For resistor elements, delta limits should be representative of the capability of the resistive material.

Resistive element devices should undergo a dc burn-in at full rated power and temperature for 100 hours minimum (50 hours each end for attenuators). Delta limits of resistance change should again be imposed. When burn-in cannot be imposed, a high temperature storage for 100 hours at the maximum temperature of the device (above system use) should be done. High power testing by the manufacturer at rated conditions may not be possible and testing by the user must take its place. High-power testing at maximum rated temperature (both above system usage) should be done. A one-hour test under these conditions while observing for hot spotting of the element and body surface temperature monitoring can be effective.

6.5 MICROWAVE DEVICES, DIRECTIONAL COUPLERS

6.5 Directional couplers.

6.5.1 Introduction. Directional couplers (DCs) are devices that couple power via electric and magnetic fields from the primary signal path to a secondary signal path. Low frequency (signal processing) devices are coupled by transformer action. High frequency devices are coupled by proximity of transmission lines in either coaxial or waveguide. Coupling is the ratio of the power in the primary arm to the power in the secondary arm expressed in decibels. Directivity is the ability of the secondary arm to sense the direction of power flow in the primary arm.

A very useful form of the DC is the microwave hybrid. The hybrid is a four-port, symmetrical device that can be made from transformers at high frequency and in stripline, microstrip, coaxial or waveguide at microwave frequencies. Power entering one port is split between two of the ports with the fourth port having a high isolation from the input signal.

6.5.2 Usual applications. The most common DC is the hybrid. It is used in amplifier chains to combine or split power, in balanced mixer designs for push-pull operation, and in double-balanced mixers to provide image rejection. It is also used in duplexers to separate signals, and in bridges for precision phase measurements, or anywhere a 3 dB, symmetrical, four-port, high-isolation device is needed.

Other uses of DCs are as signal or pulse samplers. For example, a DC that samples reflected power from a load indicates return loss or VSWR. In these applications, the sampled incident or reflected signal level is set by the coupling ratio to the power in the primary arm. Standard coupling values are 3, 6, 10, 20, and 30 dB. Any coupling value from 3 dB up is achievable. If the application involves sampling a high-power signal the worst possible load VSWR must be considered to avoid sending excess power into the secondary line reverse power termination. DCs span the frequency range from 1.0 MHz to greater than 100 GHz with limited operating frequency range per device.

Signal processing directional couplers are used in the frequency range of 1 to 1,000 MHz. Above this frequency, coaxial and waveguide couplers are available. Coaxial models are typically limited to a bandwidth of an octave. Waveguide couplers are limited by their respective waveguide transmission characteristics. Directivity in coaxial units is typically 15 to 20 dB, except in special designs. Directivity in waveguide can range from 20 to 40 dB, depending on the design. Insertion loss of couplers is affected by the coupling value for low coupling values as shown in Figure 65.

6.5.3 Physical construction. All DCs are essentially four-port devices with one port of the secondary either terminated or unterminated. Signal processing DCs are typically made using toroidal transformers with chip resistors and capacitors used for tuning, balancing, or terminating one-port. Package configurations include TO cans, DIPs, flat packs, and stripline with coaxial connectors. Hybrids of 180 and 90 degrees are common.

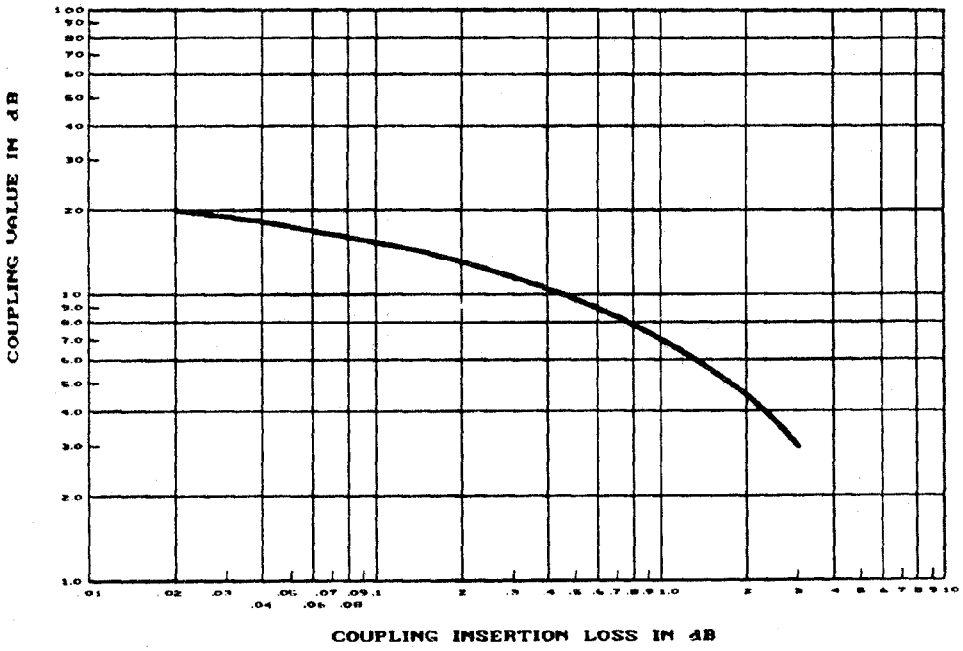


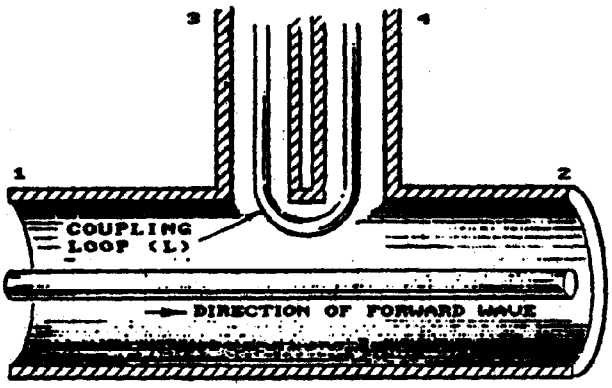
FIGURE 65. Coupling value vs main line insertion loss.

Coaxial units are constructed using coupled line in coaxial, stripline, microstrip, and slab line. Secondary arm terminations are internal or screw-on coaxial loads, Figure 66A. Waveguide units are made by joining two pieces of waveguide with appropriate coupling apertures in the joining walls (i.e., there are broad wall, side wall and cross guide couplers), Figure 66B. A waveguide resistive loop coupler is made with a coaxial line as the secondary arm. A coupling slot in the outer conductor is embedded in the broad wall of the waveguide and is rotated and brought into proximity with a coupling hole in the waveguide until the desired coupling coefficient is obtained.

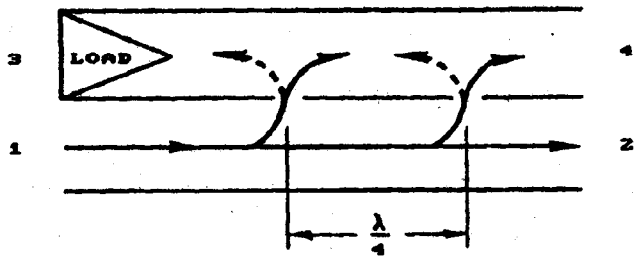
6.5.4 Military designations. The directional coupler selection standard is MIL-STD-1328. The military directional couplers are defined in MIL-C-15370.

6.5.5 Electrical characteristics. VSWR, insertion loss, coupling, coupling variation, and directivity are the primary parameters of a directional coupler. VSWR of the primary arm is measured with the device terminated in a matched load. VSWR of the secondary arm is measured with no load if the device is internally terminated. For DCs of less than 10 dB, the primary arm should be terminated in a matched load. Insertion loss of the primary arm is measured with a matched load on the secondary arm.

6.5 MICROWAVE DEVICES, DIRECTIONAL COUPLERS



A. Coaxial coupler



B. Waveguide coupler

FIGURE 66. Directional coupler construction.

Coupling is usually a measure of the ratio of the power into the primary arm in the forward direction to the power out of the secondary arm at center frequency under matched conditions. The coupling coefficient is a directional parameter, because the reverse of coupling value is directivity. Coupling variation is the change in coupling value across the frequency band. A calibration chart plotting actual coupling versus frequency is often provided so that it is possible to dispense with extremely constant coupling.

Directivity is the ratio of the power at the secondary arm output with power into the primary arm in the forward direction, to power at the secondary arm with the same power as in the first condition going into the primary arm in the reverse direction at center frequency. The directivity equals the coupling coefficient under perfect short circuit conditions. Because the actual directivity is affected by load VSWR in the reverse direction, devices should be tested under conditions of system usage to assure adequate directivity.

6.5.6 Environmental considerations. Environmental considerations for directional couplers are both mechanical- and temperature-related. Because coupling is achieved by bringing two transmission media close to each other, any severe mechanical shock or vibration, or differences in coefficients of expansion and contraction can cause variations in the coupling coefficient. Braze joints, flanges, and bodies of waveguide DCs must withstand humidity and salt corrosion. High power waveguide DCs may have to be pressure tight and able to withstand high power at high altitudes.

6.5.7 Reliability considerations.

6.5.7.1 Failure modes. The dominant failure modes of directional couplers tend to be mechanical in nature. These failures are caused by leaks, structural separation, or displacement under stress. As previously mentioned, there is a possibility of electrical burn-out from secondary line termination overload.

6.5.7.2 Screening. Probably the best single screen is repeated temperature cycling (10 cycles). A 100-percent seal test such as MIL-STD-202, Method 112, condition A, provides an endpoint measurement of mechanical integrity.

If the secondary line termination is a discrete resistor, a power burn-in may be helpful. As with attenuators, there is little purpose in attempting power burn-in with distributed or bulk terminating resistors. However, in both discrete, distributed, and bulk terminating resistors, there is an advantage in a temperature-cycling screen of both the resistor and the method used to attach it to the structure. High-power waveguide DCs may require a pressure test and high-power, high-temperature or high-altitude, 1-hour burn-in to screen from arcing.

As with attenuators, while monitored shock and vibration may be helpful, such a procedure may be more indicative of weakness in the monitoring equipment than the coupler under test.

6.6 MICROWAVE DEVICES, FLANGES AND WAVEGUIDE ASSEMBLIES

6.6 Flanges and waveguide assemblies.

6.6.1 Introduction. Microwave circuits are often configured in the waveguide transmission medium. Common usage ordinarily limits "waveguide" to some sort of metallic tubing and "waveguide assembly" to denote an assembly of waveguide sections. The hardware involved makes up the category "waveguide accessories" and includes flanges, adapters, and gaskets.

The waveguides discussed are air dielectric and may have cross sections that are coaxial, rectangular, rectangular with a re-entrant ridge in one or both of the wide sides, circular, or elliptical (see Figure 67). Waveguides may be further subdivided into rigid or nonrigid structures. Waveguide assemblies may be composed of one or more of these subdivisions.

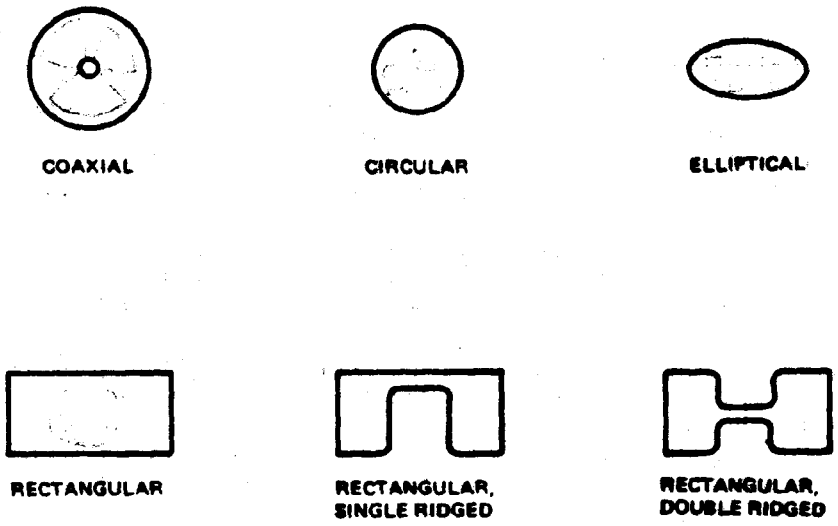


FIGURE 67. Waveguide cross-sections.

6.6.2 Usual applications. Waveguide flanges are the interface connections for waveguides. They are usually brazed onto waveguides and bolted or screwed together to make a waveguide interconnection. Two types of flanges generally used are cover and choke flanges. A choke to cover flange interface is used in pressurized systems. The choke flange is similar to a flat-face cover flange, but it has a recessed area for a gasket to maintain pressure and an rf choke area to prevent rf leakage.

6.6 MICROWAVE DEVICES, FLANGES AND WAVEGUIDE ASSEMBLIES

Waveguides have four standard shapes which offer the designer trade-offs in power handling capability, attenuation, and bandwidth. The rectangular waveguide, most commonly used, has the highest power capability. It has reasonably low attenuation, but limited bandwidth. Circular waveguide is usually used in rotary joints. It also has the advantage of having no upper frequency cutoff and is extremely wideband. Single ridge and double ridge waveguides offer wide bandwidth at low attenuation with limited power handling capability.

Waveguide assemblies are the building blocks of a waveguide system. Devices include bends (E and H), twists, and flexible waveguide--all with selected flanges. Also available are an assortment of devices to construct waveguide assemblies, including short slot hybrids, magic Ts, directional couplers, radiating devices, etc.

Adapters are waveguide transition devices which are used to make a transition to a different size or shape of waveguide or coaxial transmission line. Waveguide-to-waveguide transition must be several wavelengths long to match the complete bandwidth of the waveguide. Waveguide-to-coaxial adapters may have moderate VSWRs at the waveguide band edges.

Gaskets are designed to fit on waveguide flanges and prevent pressure leakage. Gaskets are typically silicone rubber and, when filled with silver-plated copper, prevent rf leakage.

6.6.3 Physical construction. Standard waveguides, flanges, assemblies (other than flexible) and adapters are constructed of either copper or aluminum alloys. Flexible waveguide assemblies are sometimes rubberized on the outside to reduce vibration fatigue and prevent pressure leakage. Fabricated waveguide parts and purchased waveguide parts should have finishes to prevent corrosion and reduce insertion loss. Military specifications describe in detail the materials and finishes used.

6.6.4 Military designations. Standard waveguide flanges may be selected from MIL-STD-1327. The general specifications for waveguide flanges are MIL-F-3922 and MIL-F-39000.

Rectangular, circular, and ridge waveguide selection is covered in MIL-STD-1358 and in MIL-W-85, MIL-W-23068, and MIL-W-23351.

Waveguide to coaxial adapters selection is covered in MIL-STD-1636 and MIL-A-22641.

Gaskets are usually supplied with flanges and are defined in MIL-F-3922 and MIL-F-39000.

6.6 MICROWAVE DEVICES, FLANGES AND WAVEGUIDE ASSEMBLIES

6.6.5 Electrical characteristics. The primary characteristics of waveguide hardware are VSWR and insertion loss. Secondary characteristics would include rf leakage or heating effects. Electrical characteristics of waveguide hardware depend primarily on mechanical attributes. VSWR is affected by interface mismatch due to tolerance build up in mounting, internal physical dimensions, surface roughness, and smoothness of transitions. Any of these mechanical faults causes discontinuities and results in power reflections. Surface finish, skin depths, and resistivity will affect insertion loss. Low insertion loss will keep heating effects to a minimum. The rf leakage occurs predominately at interfaces and depends on mechanical fit; it may be reduced with conductive gasketing. High-power applications may cause arcing or breakdown. Surface smoothness and finish may prevent arcing. Pressurization of waveguide increases the peak power rating of waveguide.

6.6.6 Environmental considerations. Metal corrosion of waveguide hardware may occur in the presence of moisture. Causes of the corrosion may be inadequate removal of brazing salts, improper or inadequate metal finishes, or contact of dissimilar metals. Metal porosity and inadequate seals can cause pressurization leakage. Metal fatigue due to usage or system vibration can occur in flexible waveguide. Thin wall waveguide may cause am and pm noise to appear on a conducted signal, due to high system vibration levels. High usage interconnections should use corrosion resistant, passivated steel inserts for durability.

6.6.7 Reliability considerations. In systems requiring pressurization, screening for air leakage by an air-bubble test or a dye penetrant test should be performed. Either thermal shock (10 cycles) or temperature cycling (25 cycles) is an excellent test for compatibility of materials and finishes. Physical appearance and delta VSWR or delta insertion loss measured in a standard fixture should be as acceptance criteria. Material and finish certification with periodic salt atmosphere testing should assure adequate finishes.

6.7 Radio frequency power dividers and combiners.

6.7.1 Introduction. Communications equipment has often been designed to have the signal of interest travel a single path. As equipment has become more varied and elaborate, designers have found advantage in having signals travel two or more paths. This causes the twin problems of dividing the signal into the desired parts and, after processing, combining the results.

Because most power dividers can also be used as power combiners, the problem is simplified. There are two pairs of conditions commonly found: (1) Signals are at a sufficiently high level so that dissipation will not cause critical degradation of noise level and low enough so that dissipation will not cause heating problems, or the reverse. (2) The division and combination is at zero (or 180) degrees, or in quadrature (90 degrees) phase relation. In each of the first alternatives, the devices may be resistive; in the second alternatives there must be at least some reactive elements.

The designer may also want to provide some minimum degree of isolation between loads (in a divider) or sources (in a combiner). The degree of necessary isolation and impedance match will determine which of several choices should be used.

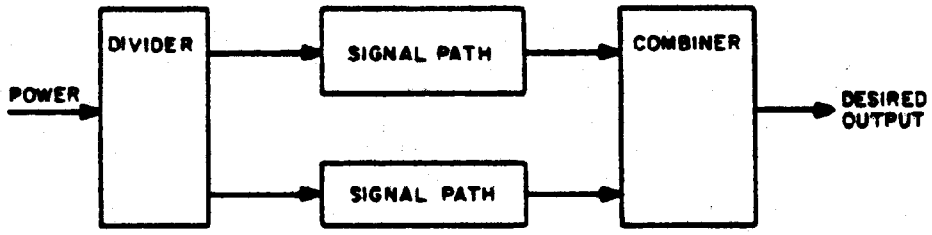
A two-way, in-phase (0 degrees) power divider is actually an internally terminated 180-degree hybrid. If the division and combination is at 180 degrees or in quadrature (90 degrees), the device is a hybrid.

6.7.2 Usual applications. The most common use of dividers/combiners is the division into two equal parts or the addition of two signals, (Figure 68A) One obvious solution, if more than two parts are desired or are to be combined (Figure 68B), is to "branch" divider/combiners, as in Figure 69. This method may have very persuasive advantages when the number of divisions/combinations is some power of 2 such as 4, 8, 16, or 32. The intermediate numbers, however, represent wasted power, for the unwanted paths should be terminated in dummy loads. Where there is no wasted power desired and the number of outputs and inputs is known, there are "n-way" dividers/combiners available. It is possible to combine 2-way and n-way devices, but for reasons related to the phase of the output or loss it may not be advisable.

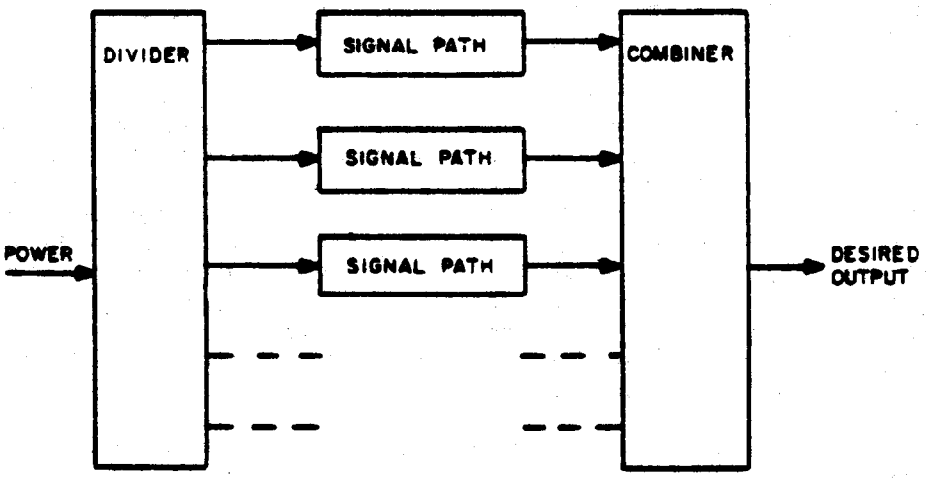
One obvious use of a divider is to proportion power to the various elements of an antenna array for transmission or, alternatively, to combine signals from the same elements for receiving.

Transistor amplification at high frequencies and microwaves is often accomplished with the use of power dividers and combiners. Signals can be amplified at lower power levels (if divided) and the power added from the several amplifiers (in the signal paths) in a combiner (see Figure 70). There is an additional advantage to this type of amplification if a form of divider/combiner is used that provides isolation between the various paths. One or more amplifiers can fail catastrophically without the array losing more than 1/n of the array gain/amplifier lost. The type of divider/combiner used in such application is commonly known as a "hybrid," and under some circumstances may provide 20 dB to 40 dB isolation.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS



A. 2-way power divider/combiner



B. N-way power divider/combiner

FIGURE 68. Use of 2-way and n-way power divider/combine.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

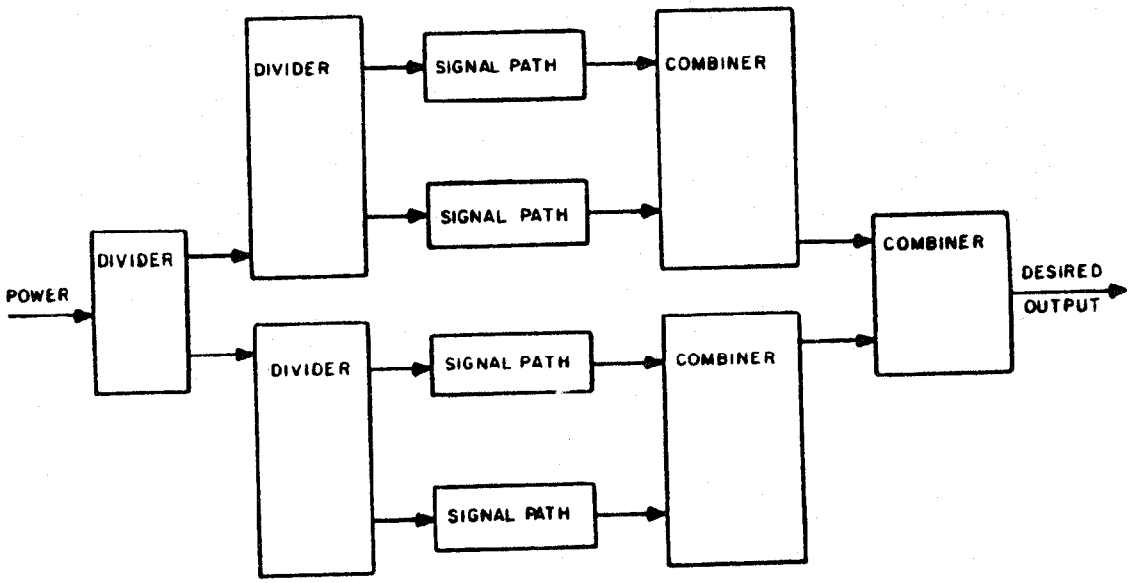


FIGURE 69. Use of 2-way dividers/combiners where $n = 2^x$.

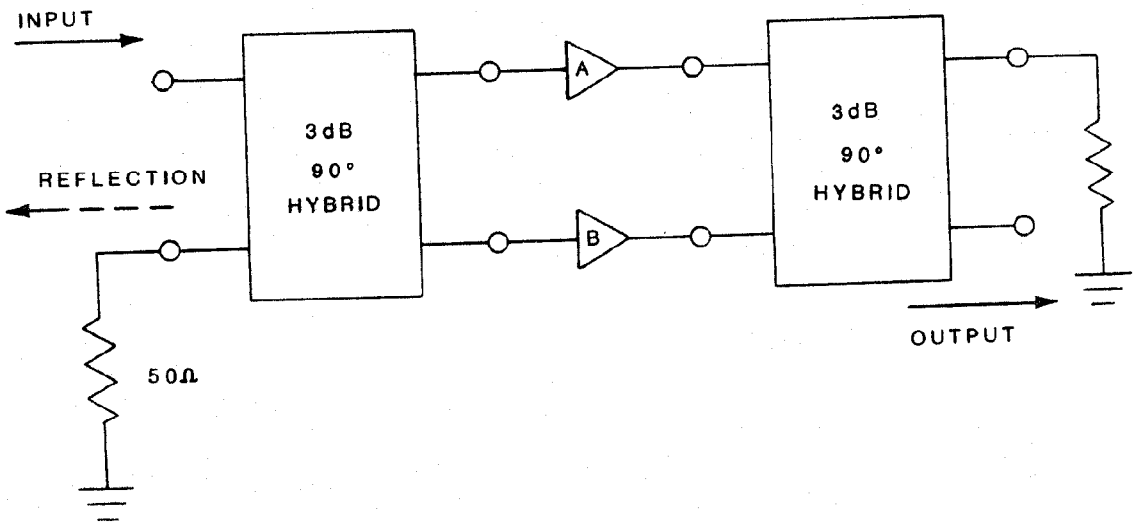


FIGURE 70. Balanced amplifier.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

A 180-degree hybrid is a reciprocal four-port device which provides two equal amplitude in-phase signals when fed from its sum port and two equal amplitude, 180-degree, out-of-phase signals when fed from its difference port. Opposite ports of the hybrid are isolated.

Utilizing the functional diagram of Figure 71, we can consider the application of signal at one or more of the ports of the hybrid. The cases that are important to consider are the following:

a. Operation as a power divider--one source operating at ports A, B, C or D

b. Operation as a power combiner--two sources operating at ports A and B, or C and D.

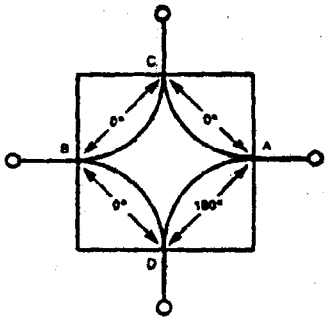


FIGURE 71. 180° Hybrid.

As a power divider, the hybrid will split equally the input signal and deliver one half the power to each load. Because all ports are considered to be at Z_0 impedance, the voltages at the outputs will be proportional to the square root of the output power and will be phase shifted by the amount indicated for that path of the hybrid, since $P_{out} = \frac{1}{\sqrt{2}} P_{in}$, $V_{out} = \frac{1}{\sqrt{2}} V_{in}$. For example, under matched conditions, a source voltage of $2E \cos \omega t$ will supply a voltage of $E \cos \omega t$ to the input of the hybrid. If an input signal at Port A of $E \cos \omega t$ is injected, the resultant output is:

At Port C $\frac{1}{\sqrt{2}} E \cos \omega t$

At Port D $\frac{1}{\sqrt{2}} E \cos (\omega t - 180°)$

No signal will appear at Port B. The various power divider relationships are summarized in Table V.

TABLE V. Power divider relationships for 180-degree hybrids

Input Signal	Input Port	Output Signals			
		Port A	Port B	Port C	Port D
$E \cos \omega t$	A	-	0	$\frac{1}{\sqrt{2}} E \cos \omega t$	$\frac{1}{\sqrt{2}} E \cos(\omega t + 180°)$
	B	0	-	$\frac{1}{\sqrt{2}} E \cos \omega t$	$\frac{1}{\sqrt{2}} E \cos \omega t$
	C	$\frac{1}{\sqrt{2}} E \cos \omega t$	$\frac{1}{\sqrt{2}} E \cos \omega t$	-	0
	D	$\frac{1}{\sqrt{2}} E \cos(\omega t + 180°)$	$\frac{1}{\sqrt{2}} E \cos \omega t$	0	-

When the hybrid is used as a combiner, it performs the vector addition of two signals. For example, in the general case of two equal amplitude, equal frequency signals of arbitrary phase applied to Ports A and B; $E \cos(\omega t + \alpha) + E \cos \omega t$, the resultant outputs C and D will be

$$\text{Resultant C} = \frac{E}{\sqrt{2}} [\cos \omega t + \cos(\omega t + \alpha)]$$

$$= \frac{E}{\sqrt{2}} [2 \cos 1/2 (2\omega t + \alpha) \cos 1/2 (-\alpha)]$$

$$= \sqrt{2} E \cos(-\alpha/2) \cos(\omega t + \alpha/2)$$

$$\text{Resultant D} = \frac{E}{\sqrt{2}} [\cos(\omega t + 180°) + \cos(\omega t + \alpha)]$$

$$= \frac{E}{\sqrt{2}} [2 \cos 1/2 (2 \omega t + \alpha + 180°) \cos 1/2 (180° - \alpha)]$$

$$= \sqrt{2} E \cos(90° - \alpha/2) \cos(\omega t + \alpha/2)$$

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

Table VI lists the relationships for various combinations of signals applied to ports A and B or C and D.

A 90-degree hybrid functions in much the same manner as a 180-degree hybrid. Equal amplitude outputs result when a signal is fed to one of the inputs. Opposite ports of the 90-degree hybrid are also isolated as in the 180-degree hybrid. Figure 72 shows the functional diagram of the quadrature hybrid. As can be seen from this diagram, a signal applied to any input will result in two quadrature, or 90-degree outputs. Ports A and B or C and D are isolated.

Table VII shows the two equal amplitude outputs when a matched signal E cos ωt is applied to each of the inputs. To use the hybrid as a combiner, apply two matched, equal amplitude signals to ports A and B, or C and D, with arbitrary phase. If, for example, the signal E coswt is applied to port A and E $\cos(\omega\tau + \alpha)$ is applied to port B, the resulting outputs are

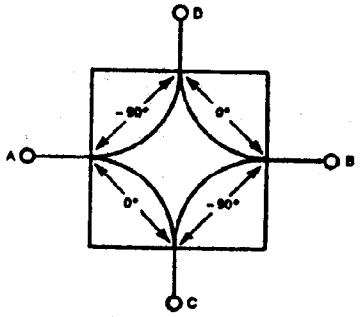


FIGURE 72. 90-degree hybrid.

$$\text{Resultant } C = \frac{E}{\sqrt{2}}[\cos \omega t + \cos (\omega t + \alpha - 90°)]$$

$$= \frac{E}{\sqrt{2}}[2 \cos 1/2 (2 \omega t + \alpha - 90°) \cos 1/2 (-\alpha + 90°)]$$

$$= \sqrt{2}\, E \cos (45° - \frac{\alpha}{2}) \cos (wt + \frac{\alpha}{2} - 45°)$$

TABLE VI. Power combiner relationship for 180-degree hybrids

Input Signal	Input Port	Output Signals			
		Port A	Port B	Port C	Port D
*E cos ωt	A	-	-	$\sqrt{2}$ E cos $\omega t(-\alpha/2)$	$\sqrt{2}$ E cos$(90°-\alpha/2)$
*E cos$(\omega t+\alpha)$	B			[cos$(\omega t+\alpha/2)$]	[cos$(\omega t+\alpha/2+90°)$]
E cos ωt	A	-	-	$\sqrt{2}$E cos ωt	0
E cos ωt	B				
E cos$(\omega t+180°)$	A	—	—	0	$\sqrt{2}$ E cos ωt
E cos ωt	B				
E cos ωt	C	0	$\sqrt{2}$ E cos ωt	-	-
E cos ωt	D				
E cos ωt	C	$\sqrt{2}$ E cos ωt	0	-	-
E cos$(\omega t+180°)$	D				
E cos $\omega_1 t$	A	-	-	$\frac{1}{\sqrt{2}}$ E (cos $\omega_1 t$ + cos $\omega_2 t$)	$\frac{1}{\sqrt{2}}$ E(cos $\omega_1 t$ + cos $\omega_2 t$
E cos $\omega_1 t$	B				

*See previous page for derivation of this example.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

$$\text{Resultant } D = \frac{E}{\sqrt{2}} [\cos(\omega t - 90°) + \cos(\omega t + \alpha)]$$

$$= \frac{E}{\sqrt{2}} [2 \cos 1/2 (2\omega t + \alpha - 90°) \cos 1/2 (-\alpha - 90°)]$$

$$= \sqrt{2}\, E \cos\left(-45° - \frac{\alpha}{2}\right) \cos\left(\omega t + \frac{\alpha}{2} - 45°\right)$$

$$\text{Phase of Resultant C} = \text{Phase of Resultant D} = \left(\omega t + \frac{\alpha}{2} - 45°\right)$$

TABLE VII. 90-degree hybrid as a power divider

Input Signal	Input Port	Output Signals			
		Port A	Port B	Port C	Port D
E cosωt	A		$\frac{E}{\sqrt{2}} \cos(\omega t - 90°)$	$\frac{E}{\sqrt{2}} \cos \omega t$	0
	B	$\frac{E}{\sqrt{2}} \cos(\omega t - 90°$		0	$\frac{E}{\sqrt{2}} \cos \omega t$
	C	$\frac{E}{\sqrt{2}} \cos \omega t$	0		$\frac{E}{\sqrt{2}} \cdot \cos(\omega t - 90°)$
	D	0	$\frac{E}{\sqrt{2}} \cos \omega t$	$\frac{E}{\sqrt{2}} \cos(\omega t - 90°)$	

The amplitudes of the resultant outputs at C and D vary, based on the phase of the inputs, whereas the phases of the outputs are always equal. This property is useful for receiving phase-modulated signals. The relationships for a 90-degree hybrid with signals applied to ports A and B or C and D are shown in Table VIII.

TABLE VIII. 90-degree hybrid as a power combiner

Input Signal	Input Port	Output Ports			
		Port A	Port D	Port C	Port B
$E \cos \omega t$	A		$\sqrt{2}\, E \cos(-45° - \alpha/2)$	$\sqrt{2}\, E \cos(-45° - \alpha/2)$	
$E \cos(\omega t + \alpha)$	B		$[\cos(\omega t + \alpha/2 - 45°)]$	$[\cos(\omega t + \alpha/2 - 45°)]$	
$E \cos \omega t$	A		0	$\sqrt{2}\, E \cos \omega t$	
$E \cos(\omega t + 90°)$	B				
$E \cos(\omega t + 90°)$	A		$\sqrt{2}\, E \cos \omega t$		
$E \cos \omega t$	B				
$E \cos \omega t$	D	0			$\sqrt{2}E \cos \omega t$
$E \cos(\omega t + 90°)$	C				
$E \cos \omega t$	D	$\sqrt{2}E \cos \omega t$		0	0
$E \cos(\omega t + 90°)$	C				
$E \cos \omega_1 t$	B		$\frac{E}{\sqrt{2}}\left[\cos \omega_1 t + \cos(\omega_2 t - 90°)\right]$	$\frac{E}{\sqrt{2}}\left[(\cos(\omega_1 t - 90°) + \cos \omega_2 t\right]$	
$E \cos \omega_2 t$	A				

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

The quadrature hybrid is often used to generate or assist in the detection of single sideband signals, in combination with mixers. One characteristic of mixers (actually algebraic multipliers) in heterodyning systems is the creation of sum and difference frequencies. A receiver is sensitive to an image frequency as well as a desired frequency. Processing a received signal through in-phase (I) and quadrature (Q) channels, however, allows addition of the desired signal components and reduces the image-frequency components -10 dB to -40 dB (see Figure 73). The reverse process can be used to generate a single-sideband signal.

The hybrid may be used as a directional coupler when one of the four ports of a 2-way hybrid is terminated in its design. The pair of ports isolated from each other may be used as the "input" and "reflected-power" ports of a directional coupler with the "load" connected to the combined port. Any reflection from the load will cause a proportional amount of power to appear at the reflected-power port. Power loss of 3 dB in the reactive hybrid or 6 dB in the resistive hybrid commonly eliminates consideration of use as a directional coupler, except for switching and instrumentation applications.

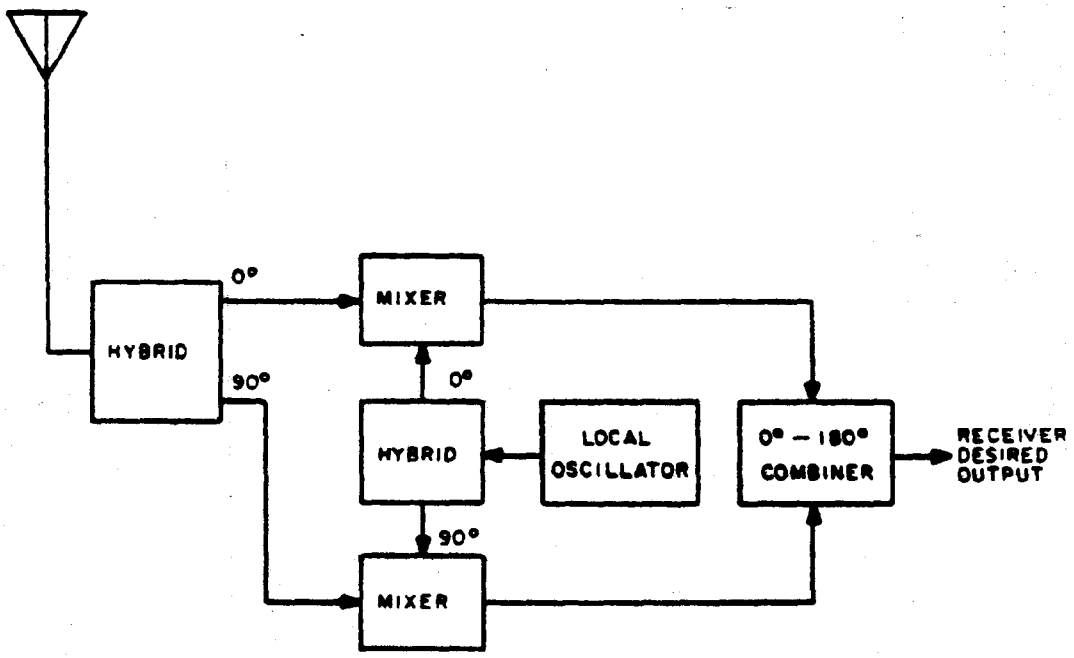


FIGURE 73. Phasing system of single-sideband reception.

The two common switching applications for the reactive hybrid are to provide "lossless" TR and ATR switching between transmitter and receiver by total reflection from open or short-circuit terminations, or provide a digitally-determined phase shift as in an electronically-steerable antenna array (see Figure 74).

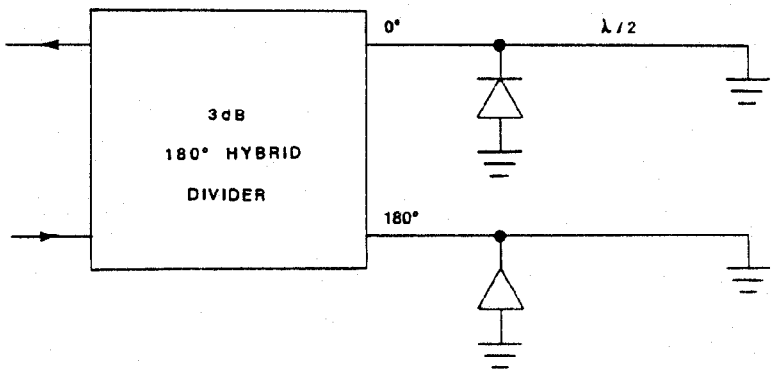


FIGURE 74. Single bit phase shifter.

The amplitude and phase of the reflection from a nonzero or noninfinity termination may be used to either discover what the termination is measurement, or to provide a phase shift (by reflection from a fixed or variable reactance or sliding short).

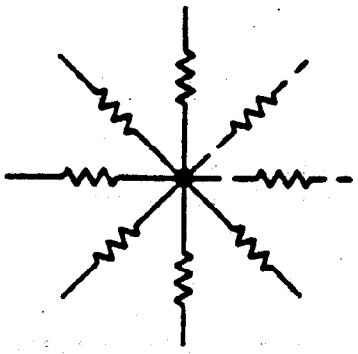
6.7.3 Physical construction. The most obvious physical features are controlled by the intended frequency range. At dc and relatively low frequencies, the parts making up the divider/combiner tend to be bulky with relatively open construction and wiring. When a high degree of isolation is required in the hybrids, balance of part values, wiring layout, and coupling between the windings of transformers become very important. A slight change of wire position may be the difference between 40 dB and 60 dB isolation. Shielding at low frequencies may be helpful, but even at low frequencies stray capacitance of wiring-to-shielding may cause problems.

As frequency rises, the parts tend to shrink and part types having less frequency sensitivity must be used. Film resistors replace composition and wire-wound types. Transformers become toroidal with bifilar (or 3-in-hand or even 4-in-hand) windings. Core material changes from iron to powdered iron to ferrite. Capacitors become low-inductance ceramic or mica. Connectors become coaxial or fitted with pins for insertion into the easily-duplicated wiring of a printed circuit board.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

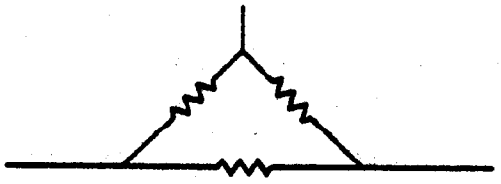
As frequency continues to rise above approximately 500 MHz, discrete parts give way to transmission lines. The transmission lines may be used singly to provide phase delay or reactance, or as coupled pairs to provide signal transfer. Center-tapped transformers are replaced with transmission-line baluns (balanced-to-unbalanced transforming networks). There is overlap in frequency of usage between coaxial transmission line and stripline, and then stripline to microstrip. Progressive increase of frequency or power mandates change to waveguide structures.

The simplest form of a divider/combiner is shown in Figure 75A. Any of the resistors may be used as the input port or ports with the remainder as output ports. The delta-circuit equivalent of the 3-port version of Figure 75A is shown in Figure 75B. The delta equivalent is rarely used with more than 3-ports, because for those cases, it requires more resistors than the star circuit and is physically more awkward. Some possible physical configurations are shown in Figure 76.



Each resistor = Z_0 (N-2)/N

A. Star network with n-ports



Each resistor = Z_0

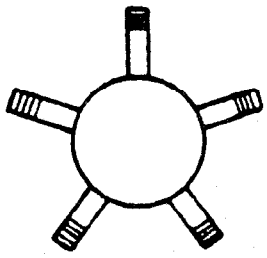
B. Delta network with ports

FIGURE 75. Resistive divider/combiners.

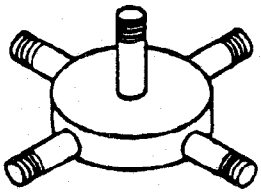
6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

A high frequency, 2-way, in-phase divider/combiner is shown in Figure 77. Transformer T2 is usually a bifilar-wound toroid and T1 provides a 2:1 impedance stepup to port 1. The resistor is an internal termination, and in dividing a port 1 input, it has zero required dissipation when the loads on ports 2 and 3 are equal to each other. When the loads on ports 2 and 3 are unequal, the resistor is called upon to dissipate power. The ghosted capacitor is often used to broaden the useful bandwidth of the device.

The most famous early waveguide hybrid is the Magic-T (Figure 78). Ports are made in straight waveguide so that one port receives the "electric" field and the other the "magnetic" field of any wave attempting passage. Because half the wave power is in each field, the power division is equal. Obversely, input to the E-plane port will result in 180-degree phase-difference output in the colinear ports, and input to the H-plane port will result in zero phase difference in the colinear port output. Because of mismatches at the junction, it is fairly common to provide an impedance match by such means as a post for H-plane compensation or an asymmetric diaphram for E-plane compensation. This matching necessarily narrows bandwidth.

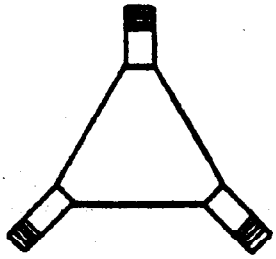


A. Five-port star (may have less or more branches)



B. Five-port star (may have less or more branches). It is a useful shape when one port is dedicated as input or output.

FIGURE 76. Some external configurations particularly adapted to resistive divider/combiner network.



C. Frequent shape of a 3-port delta

FIGURE 76. Some external configurations particularly adapted to resistive divider/combiner network (continued).

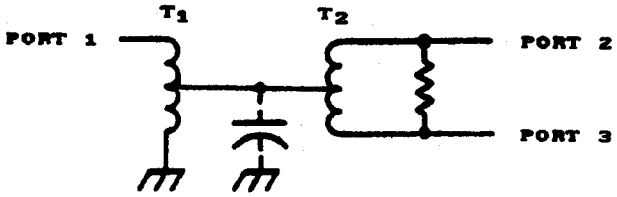


FIGURE 77. Simple zero-degree hybrid.

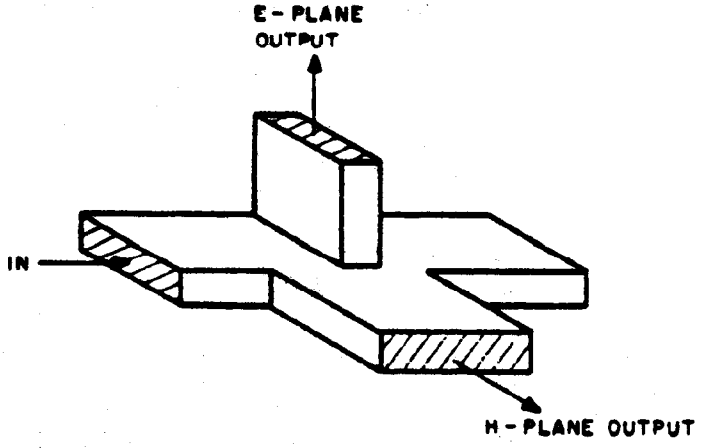


FIGURE 78. The "Magic-T."

A hybrid that yields similar results may be obtained with the "6/4" rat-race (Figure 79). This configuration may be constructed in waveguide, coaxial cable, or stripline transmission lines. Again, the impedance of the transmission line in the rat-race should be different from the port impedances to provide impedance match.

A similar appearing quadrature, shown in Figure 80A, is known as a pi circuit. For stripline applications it has generally been superseded with the quarter-wave coupled quadrature hybrid as shown in Figure 80B. Where increased bandwidth is desired, a 3/4-wave coupled hybrid can be used (Figure 80C). The lines of the hybrid circuit are usually one above the other, but may be adjacent and are bent back and forth to save board space at relatively low frequencies. An example of this type of hybrid is shown in Figure 81.

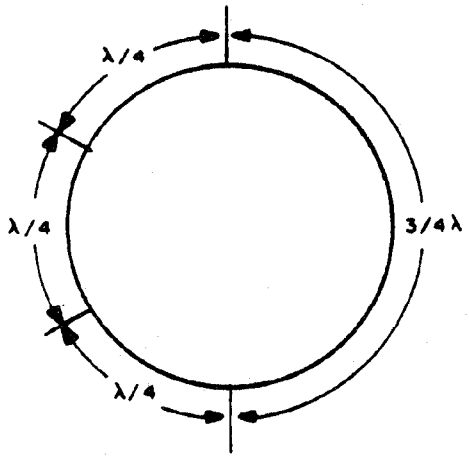
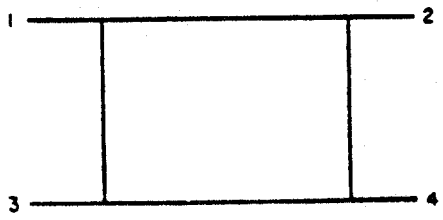


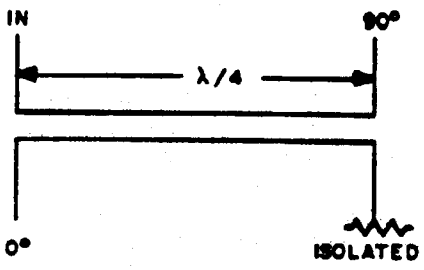
FIGURE 79. "6/4" rat race hybrid ring.



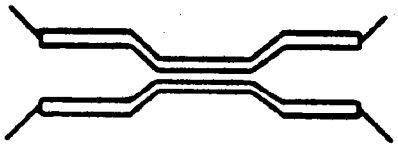
A. Branch line quadrature hybrid

FIGURE 80. Some forms of hybrid particularly adapted to stripline.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS



B. Quarter-wave coupled quadrature hybrid



C. Three-quarter-wave coupled hybrid

FIGURE 80. Some forms of hybrid particularly adapted to stripline (continued).

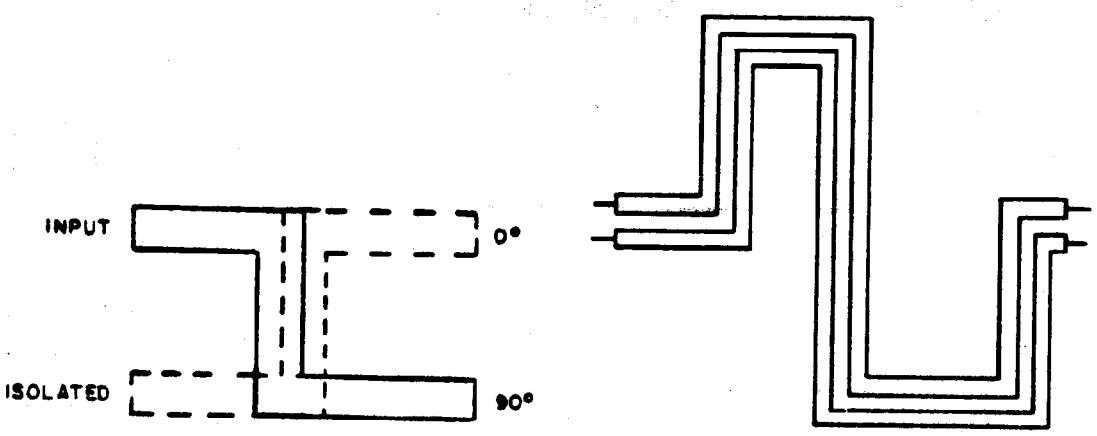


FIGURE 81. Stripline hybrids bent back to save space.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

At lower frequencies, quadrature hybrids are usually made of a mesh network of discrete inductors and capacitors intended to give equal-amplitude outputs having a near-constant phase difference of 90 degrees. Generally, the more complex the network, the nearer the ideal conditions can be approached.

Both low and high frequency "n-way" divider/combiners may be made; the simplest appearing approach being the Wilkinson configuration (Figure 81). In this configuration, sections of quarter-wavelength transmission line with an impedance equal to Zo (n)1/2 are used to provide the impedance match. A somewhat similar configuration is used at lower frequencies, with the transmission lines replaced with transformer-appearing structures.

In each configuration, parts are commonly soldered into position, and when that is not possible, attachment to the case with epoxy or other adhesive is favored by the manufacturer. Use of adhesives requires careful consideration of the materials being bonded together, as well as the temperature, shock, and vibration environments to be encountered, and the electrical effect of the adhesive.

Most often, the divider/combiner is within a shielded enclosure. Sometimes this is only the pair of ground-planes of a stripline, but it is more often continuous. It should be remembered that a good electrical shield is not necessarily a hermetic seal. It should also be remembered that commonly used conductive epoxy may not be a very good rf shield.

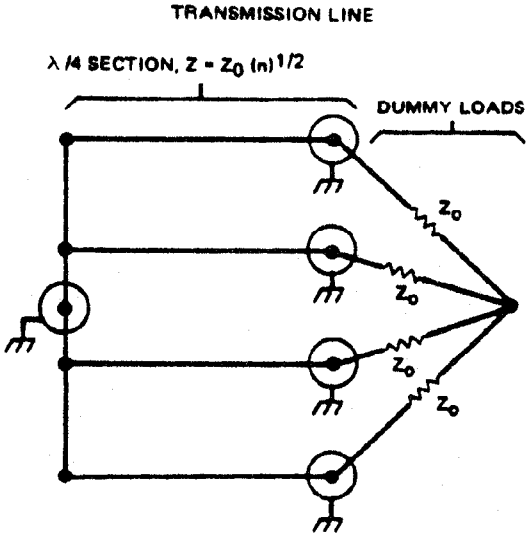


FIGURE 82. Wilkinson n-way hybrid all ports Z.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

Waveguide quadrature hybrids are constructed by paralleling two waveguides with a common wall, and having one or more openings in the common wall. If the narrow wall is common, the device is a "side-wall" hybrid; if the broad wall is common, the device is a "top-wall" hybrid. The size and positioning of the slot or slots between the waveguides affect the electrical characteristics.

6.7.4 Military designation. The military general specification for power combiners and divider/combiners is MIL-P-23971.

6.7.5 Electrical characteristics. The prime purpose of the hybrids and other power dividers/combiners under discussion is to provide equal known outputs with known phase relations. A secondary purpose is to provide isolation between several sources or loads is desired. Thirdly, though still important, is the desire to fit the device into a system having a certain characteristic impedance.

The impedance relations of Figure 74 define the target value of the resistances in the resistive power divider. The desire to match the characteristic impedance of the system results in the additional power loss shown in Table IX over that shown in Table X.

TABLE IX. Resistive dividers/combiners

Number of Ports	Output Level, dB
3	- 6.02
4	- 9.54
5	-11.14
6	-14.0
7	-15.56
8	-16.90

Note that the isolation from one port to another is no more or less than the tabulated value of output level. For identical construction in each of the legs of the star, all outputs would have the same amplitude and phase when feeding identical loads. Out of necessity, there will be small differences in the internal resistances and dimensions, and the loads will not be a perfect match. Therefore, it is necessary to specify how close to a perfect match of the nominal system characteristic impedance the test loads are. A maximum load VSWR of 1.01 is suggested.

The degree of amplitude and phase balance should be specified. The amount of loss in the combiner or divider in excess of the stated values of Table X should be specified. A fair approximation can be made by averaging the output levels and taking the difference between the average and the ideal.

The same characteristics are important with the hybrids, except that they can be expected to be more frequency-dependent. Additionally, the hybrid has the isolation feature, which although high is not infinite. Combined with poor loads, this feature may cause reduced isolation proportional to the mismatch. It is important to note in the specification of hybrids that both the internal impedance of the generator and the termination of the isolated port affect isolation. Usually isolation is specified as "the insertion loss between a pair of ports with the other pair of ports terminated in the nominal impedance."

The simplest quadrature hybrid is very a narrow band for phase and amplitude. The common 1/4-wave coupled hybrid will have two crossover points and a fair match, otherwise, over a 2:1 frequency range. The 3/4-wave coupled hybrid will have three cross-over points, and a fair match over, perhaps, a 4.5:1 frequency range.

The zero/180 degree hybrids are somewhat more wideband in the low-frequency types, and component-limited at the higher frequencies. It is important to note in the specification that "zero-degree" ports may, in fact, be 360-degree ports, and the desired effect may not be obtainable on a nonrecurring wave.

Likewise, some forms of hybrid are "slow-wave" and others "fast-wave." Mixing the two forms in a single path may cause a nonlinear phase shift with frequency dispersion.

The linking in specification of the hybrid technology of directional couplers to that of the power dividers and combiners has caused some confusion in nomenclature. It is perhaps natural with a 1/4-wave quadrature power divider to think of the energy in the second line as "coupled" and the energy reduction in the first line as "coupling loss"; yet, in most cases, the truly significant factors are amplitude equality and phase deviation from ideal. The amplitude with ideal division (Table X) in "lossless" hybrids may be subtracted from the average of observed outputs of a number of hybrids to give close approximation of the net loss in the device.

TABLE X. Reactive (zero loss) dividers

Number of Inputs or Outputs	Output Level dB
2	-3.01
3	-4.77
4	-6.02
5	-6.99
6	-7.78
7	-8.45
8	-9.03

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

The internal power rating of the combiner/divider is an important characteristic. There are three effects that limit the amount of power that these devices can withstand.

Insertion loss. The total power dissipated in the power divider under matched conditions with balanced outputs is limited by the losses in ferrite cores and wire windings in rf devices. In microwave devices, the heating of copper traces and board material causes increased resistance and more heat is generated. Exceeding the peak power rating can cause failures due to voltage breakdown of the dielectric.

Amplitude unbalance. Input power dissipation in a power divider under matched conditions takes place in the internal loads because of amplitude unbalance. Ideally, no power would be dissipated across this load, but because of imperfect amplitude balance a small voltage differential would exist.

Mismatched loads. Reflections from mismatched loads at the outputs of power dividers can cause a considerably larger voltage to appear across the internal load.

Equal amplitude signals are applied to the port when the device is used as a combiner, and little or no power is dissipated in the internal load. Possibly, a condition may occur where one or more of the signal sources fail, causing (in the case of a 2-way combiner, 50 percent of the power supplied at the remaining port to be dissipated in the internal resistor. Thus, the power injected at each port should not exceed twice the rating of the internal load in order to avoid this condition.

6.7.6 Environmental considerations. Environmental conditions may affect the equal division of power with high isolation. Those parameters are dependent on electric and magnetic fields within the power divider/combiner. Any environment that will affect those fields is potentially degrading.

Environmental conditions may cause electrical or mechanical effects that are temporary and reversible when the environment approaches again the initial conditions; or the effect may be permanent. Examples of the temporary effect might be the temperature coefficient of a capacitor, inductor, or resistor, whereas more permanent effects result from the dismounting or displacement of internal parts resulting from thermal shock, vibration, and physical shock.

The effect of moisture and humidity is somewhat open to question. The common power divider/combiner is designed for a fairly low impedance level, and a modest amount of condensation might not be operationally visible. Certainly there should be little possibility for moisture entrapment; for even if a modest collection is not electrically visible, it may promote corrosion, mold, or fungus which, in turn, will disable the device.

6.7 MICROWAVE DEVICES, RADIO FREQUENCY POWER DIVIDERS AND COMBINERS

Most dividers/combiners seem to be used at power levels of one watt or less. Under such circumstances altitude has little effect, except to cause stress on seals and sealed cases, and to cause unsealed enclosures to breathe. At higher power levels there may be a tendency for a voltage discharge of a Paschen or multipactor type, unless the device is sealed and/or pressurized. There is also a possibility that dummy resistor ratings, always less in lower-pressure air, may be exceeded.

One additional comment about shock is desirable: The electrical characteristics and environmental integrity of a device should always be suspect after a shock from dropping. A drop of a few inches onto a workbench may cause shock of a few thousand to several thousand G acceleration. A power divider/combiner should be handled with care.

In view of reliability, the divider/combiner should be considered to be an electronic assembly. The waveguide types are an exception, and should be treated as a moderately critical mechanical assembly.

Each electrical joint in the circuit and the shield should be inspected before it is hidden by subsequent assembly steps. This does not mean an inspection after each application of the soldering iron, but, rather, while the joint is still visible.

Prime failure modes noted to date have been the breaking loose or displacement of parts and wiring. This has most commonly been the result of improper soldering, improper wire treatment (anchoring or nicking), or improper choice or use of adhesives for anchoring. Design and assembly care and inspection are the only cures, but a regimen of thermal shock, vibration and mechanical shock tests also make a good screen. It is good to make "delta" electrical measurements (before and after, then calculate differences), but the limit to the delta is the repeatability possible in the measurement. For any requirement for small delta, there should be exact repetition of test equipment and personnel.

There is some advantage in the screening of parts that are to be used in a device--particularly discrete resistors and capacitors. At frequencies where inductors and transformers are more than a very few turns, these parts may also benefit. Please refer to the appropriate section of this manual for more details.

6.7.7 Reliability considerations. Specification and examination of seal integrity is not easy, and the result of poor seals is open to some discussion. A poor seal is usually a sign of poor design, poor workmanship, or abuse. Seal test, to the degree specified by the procurement documentation, is recommended on a 100-percent basis until it is determined that risk is sufficiently low to justify revision of the requirement.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

6.8 Electrical mixers.

6.8.1 Introduction. "Mixing" is a term commonly used in the subject of radio frequency to describe the practice of multiplying two electrical voltage or current waves or electromagnetic fields, to produce an output that is proportional to at least a part of the resulting product. This discussion is limited to devices that perform such a function.

Multiplication involving radio or higher frequencies almost invariably makes use of analogs which are functions of electrical circuits that provide continuous product output for inputs of two or more electrical functions. Probably the most common analog so used is the approximation to the logarithm. Two or more electrical functions are changed to their logarithms and added to derive the antilogarithm. This, if accurately done, gives the product of the original electrical functions. Some elements, such as particular classes of semiconductor diodes, have closely logarithmic current/voltage relationships over a ratio of perhaps 1,000,000:1, Figure 83. Often, the approximation shown will give sufficiently good results if the circuit includes wave filtering to remove products that are unwanted but produced by the actual relationship.

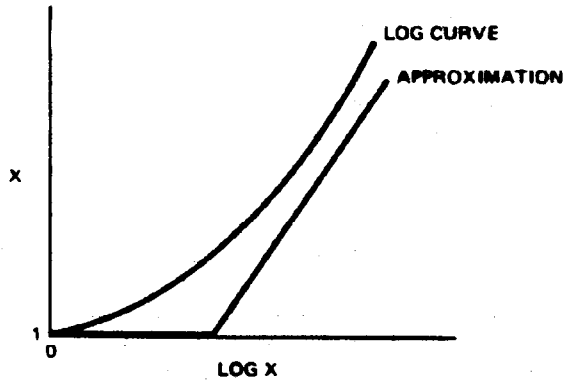


FIGURE 83. Current voltage curve for a mixer diode.

Often, unwanted products (including harmonics) can be minimized by connecting two such analog devices in a push-pull arrangement, commonly reducing even harmonics and other undesired products by 20 dB or more (balanced mixers).

The analog device is by no means limited to two-terminal components. Magnetic amplifiers, dual-gate FETs (transistors), and various connections and biasings of tubes and transistors will produce suitable characteristics. For instance, the Class C amplifier makes a good mixer. Mixers are available in monolithic and hybrid microcircuit forms; however, this discussion is limited to diode and transistor mixers.

Though there are many applications for mixers, some of the characteristics often sought are that the generators of the functions not affect each other by connection to the mixer, that the mixer present a known and constant load to those generators, that the mixer produce a useful amount of output, and that the output be the product of the inputs with a minimum of other outputs produced. These are called isolation, VSWR, conversion loss, and harmonics.

Approaches other than the logarithmic or logarithmic approximations (including the switching approximation) are possible. One that has received little recent emphasis is the negative-resistance mixer. The first examples that come to mind are the regenerative detector and the tunnel diode mixer; Figure 84 shows a way in which the multiplication function of a mixer can be theoretically obtained. As one of the goals of using any device in equipment is to obtain stable results, and since negative resistance in a linear circuit is not easy to maintain, it is doubtful that much use will be made of this multiplying procedure.

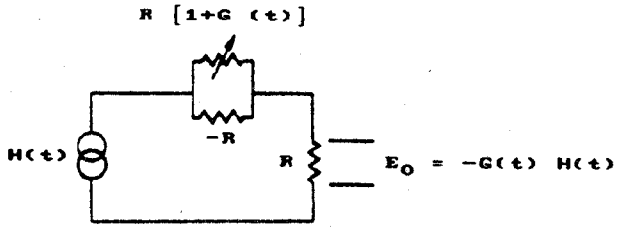


FIGURE 84. Tunnel diode mixer.

Time-varying electrical functions occupy bandwidth in the frequency domain. Though there is no bandwidth limitation in the concept of multiplication of electrical functions, there are limitations in the bandwidth of practical parts and circuits. Frequencies higher than the highest frequency present in the input to the mixer are present in its output, and frequencies approaching zero, or direct current are not uncommon.

Noise may be a problem at the desired frequencies and also at the harmonics. Electrical noise is usually described in terms of "noise figure," which is the ratio of actual extraneous noise power compared with the thermal agitation noise generated in an ideal resistor having the same impedance as the mixer input. Semiconductors and some other parts display "flicker" noise, characterized by an increase in per-Hertz bandwidth power inversely proportional to frequency. Another term for flicker noise is "1/f" noise. Above frequencies where 1/f noise is significant, the noise figure will rise slowly with the frequency, if the design is appropriately scaled. Finally, a frequency will be reached where parasitic inductances, capacitances, and resistances affect the circuit design, and the noise figure will begin a steeper climb. The noise figure is the limiting measure to the weakest electrical function input that will give useful output.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

Commonly, there is an advantage to setting the power level of one of the inputs much higher than that of the other inputs. This practice establishes the operating level of the mixer, hopefully, in some region that gives the desired outputs with a minimum of unwanted outputs. The ratio is usually 10 dB or more.

There is a tendency for some of the input energy to feed through to the output port; in many circuits the input currents will flow through the output circuit. In these cases, the user may have to terminate the circuit so that the input currents will develop no (or very little) power in the output load. These and other considerations are discussed later under Electrical characteristics, paragraph 6.3.5.

Establishing of the operating level of the mixer by the higher-power input will do much to reduce output distortions (intermodulation, compression, and "intercept point" - to be discussed later) and establish the other input and output impedances of the mixer.

6.8.2 Usual applications. The mixer function of multiplication has application from virtually zero frequency through millimeter waves. Some nonelectrical mixers are even used into the light-frequency regions in laser work.

Considering the simplest multiplication function where C = k f(A) X f(B), consider the case where f(A) can have any value from zero to one. When f(A) = 1, C = k f(B), so f(B) may be said to have been passed without change, except for the amplitude factor k. As f(A) approaches zero, the value of C will also approach zero, the output C will be zero, and the mixer will have acted as an electrically variable attenuator capable of giving infinite attenuation. (Mixers giving this characteristic are called "balanced," in that at least one input may be used to balance itself out in the event of a zero value for the other input.)

If, in the above example, f(A) were allowed to progress from zero to -1, the output would reverse the polarity of f(B) and increase to the same absolute magnitude as for f(A) = 1. Schematically, this could appear as in Figure 85 This is the foundation for three widely used applications.

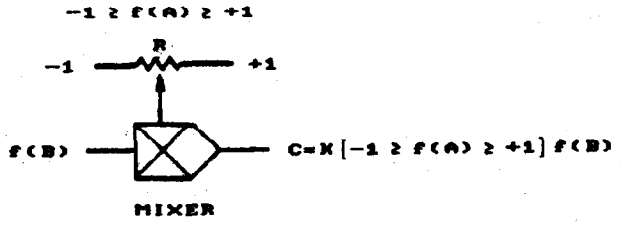


FIGURE 85. Simple mixer.

On-off switching of f(B) can be accomplished by switching f(A) between +1 and zero.

Phase-reversal of f(B) can be done by switching f(A) between +1 and -1.

Amplitude modulation can be accomplished by considering the use of electrical sine waves for f(A) and f(B). Let f(A) = cos A, f(B) = cos B, then, consider the original multiplication function,

$C = \frac{k}{2} \cos (A+B) + \frac{k}{2} \cos (A-B)$, which is the expression for double-sideband amplitude modulation without carrier wave (Figure 86).

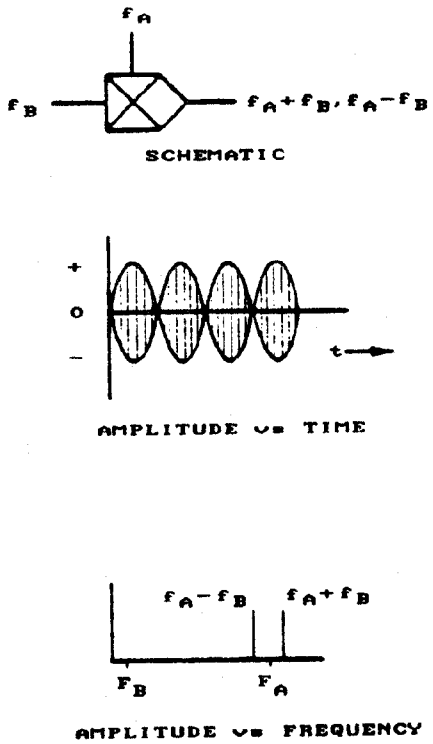


FIGURE 86. Double-sideband amplitude modulation.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

Amplitude modulation with carrier (conventional broadcast AM) can be obtained by letting

$f(A) = 1 + \cos A$, resulting in $C = k \cos B + \frac{k}{2} \cos (A+B) + \frac{k}{2} \cos (A-B)$

The (A+B) and (A-B) terms are the "sum-and-difference" or "upper and lower" frequency sidebands, where the angles A and B are

$A = 2\pi fAt$ and $B = 2\pi fBt$, the frequency f is in Hertz and time t in seconds

Note that in "amplitude modulation with carrier" the total bandwidth involved is from zero frequency (or dc, the "1" term) to f_A+f_B. A properly chosen and applied mixer must consider each of these frequencies (Figure 87).

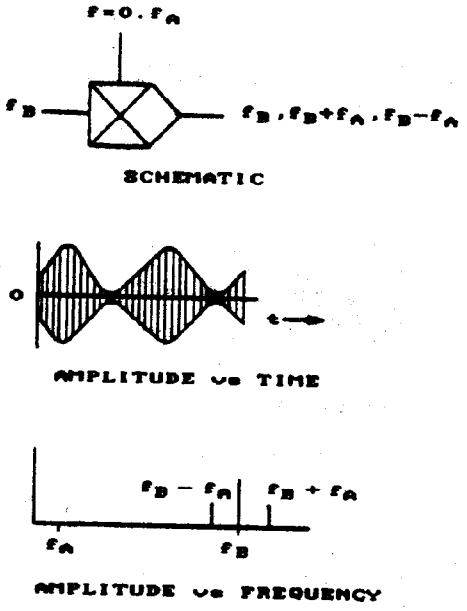


FIGURE 87. Modulator with carrier.

Phase detector use of mixers is possible by having one of the frequencies in f(A) the same as one frequency in f(B). A direct current component of the output results which is proportional to the cosine of the phase angle between the two inputs and the product of the magnitudes of the single common frequency (Figure 88).

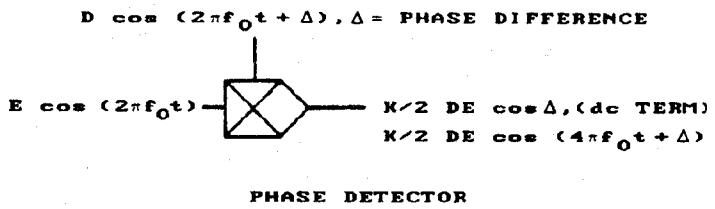
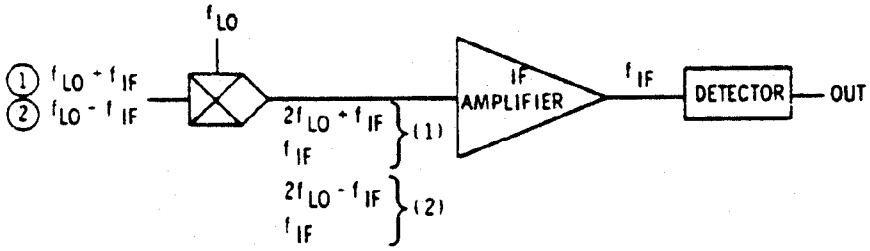
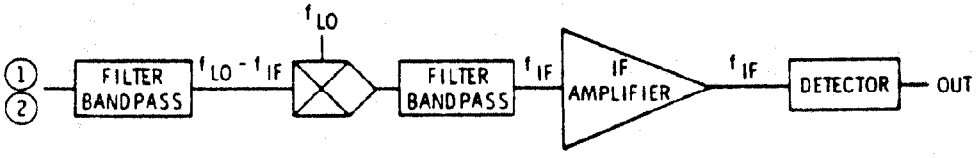


FIGURE 88. Phase detector.

Phase detector output from the mixer that is not a function of phase alone (except the zero output, or odd multiple of 90 phase difference) and is not proportional to phase angle is awkward. Filters and amplitude limiters on each input help select the frequency of interest. Clippers to make approximate square waves of the input waves will, with a wide-band mixer, give output proportional to phase angle with fairly good accuracy. The mixers must be wide-band, as making square waves of the basic frequency will require distortionless processing up to 10 or more times the basic frequency in the mixer.



A. Simple superheterodyne



B. Superheterodyne with filter selectivity

FIGURE 89. Superheterodyne with filter selectivity.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

Frequency mixing for use in superheterodyne receivers is probably the most common use of mixers (Figure 89). Here the received signal is mixed with a locally generated source of radio frequency power of much greater level, and the difference frequency (known as the intermediate frequency (IF)) is selected as the output. If the local oscillator is made variable, the receiver will be sensitive to various frequencies (for a constant intermediate frequency) by tuning the local oscillator. The IF may be amplified and narrowed by filtering to any desired degree, giving a receiving system of constantly high gain. Unfortunately, for each local oscillator (LO) frequency, there are two received frequencies giving the same IF. The unwanted frequency is called the "image frequency." Image responses may be reduced either by filtering the antenna line or by single sideband cancellation techniques.

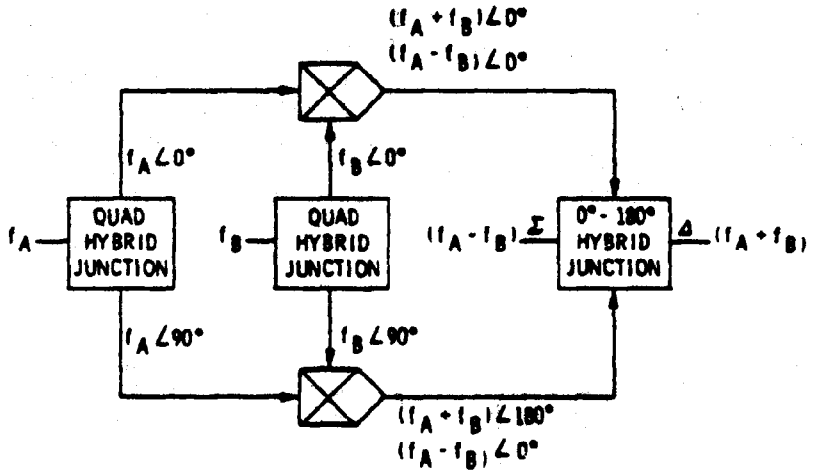


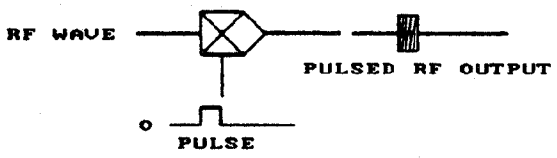
FIGURE 90. Single sideband generator.

Single sideband signals may be generated by the phasing system as shown in Figure 90. Here, if the signal amplitudes are kept equal in the corresponding parts of the two paths, the lower sideband will emerge from the summation port, and the upper sideband will emerge from the difference port. Depending on the balance and phase accuracy of the hybrids, the suppression of the unwanted sideband can be 20 dB or more. Such a system may exceed 40 dB suppression of the unwanted sideband if operated on fixed frequencies, but this degree of suppression becomes more difficult as bandwidth and operating frequency increase.

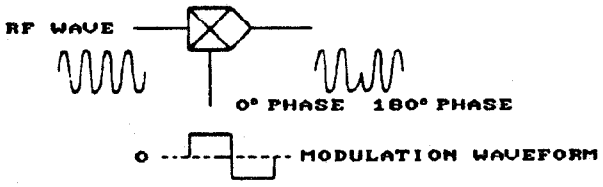
The image frequency of superheterodyne reception may likewise be suppressed.

Coherent detection of signals having symmetrical sidebands or a carrier is likewise attainable by feedback control of the frequency and phase of f_B (a voltage-controlled oscillator) by locking f_B to either the optimum phase for

detection of the sidebands or to the carrier. Noncoherent interference above or below f_B may be suppressed without loss of desired sideband content. (In voice communications circuits, the two outputs have been presented binaurally to allow additional mental discrimination by the apparent directional difference between the desired signal and interference.)



A. Pulse modulation or gating



B. NRZ (non return-to-zero) modulation

FIGURE 91. Modulators.

Pulse modulation or gating (Figure 91) may be easily accomplished by feeding the signal to be modulated or gated into one port and feeding the gating signal into the other. As the gating signal ratio of "on" to "off" time may range from extremely large to very small, the port to which the gating pulse is applied should be extremely wideband. Earlier when the switching possibilities of the mixers were discussed, this factor was not mentioned.

Nonreturn-to-zero (NRZ) modulation is often more useful pulse modulation than the on-off form of amplitude modulation. In Figure 91b, this principle is shown in a phase-reversal scheme frequently used in data communications. The modulation waveform bandwidth again must be considered. Some of the value of the approach may be seen in that the NRZ system virtually eliminates problems that might be caused by a moderate leakage of the rf wave through the mixer. The full-value output cancels or masks all but massive leakage.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

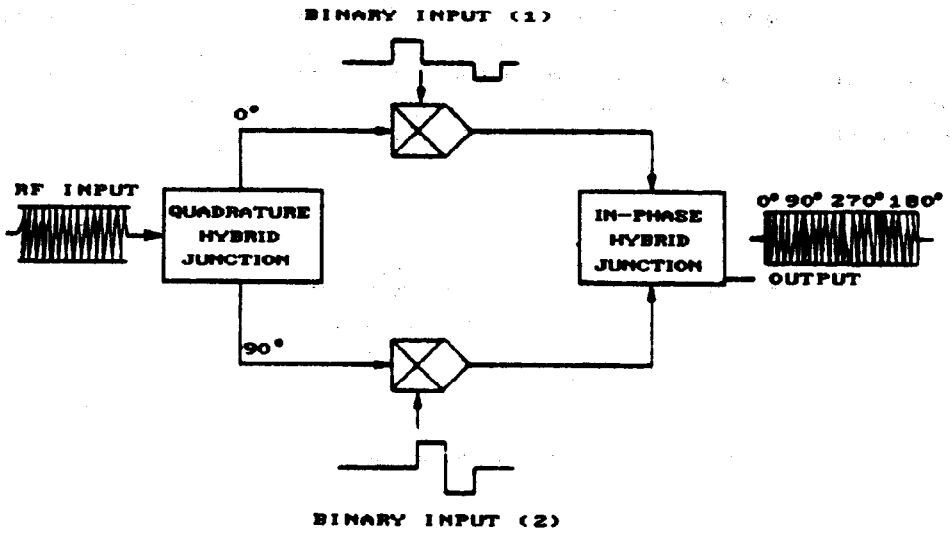


FIGURE 92. Quadrature four-phase modulation.

One attractive form of modulation is quadrature or four-phase modulation (Figure 92), where any quadrant can be electronically selected. The use of hybrid junctions not only supplies the 90 degree phase increments and easy addition, but also helps isolate the mixers from each other electrically.

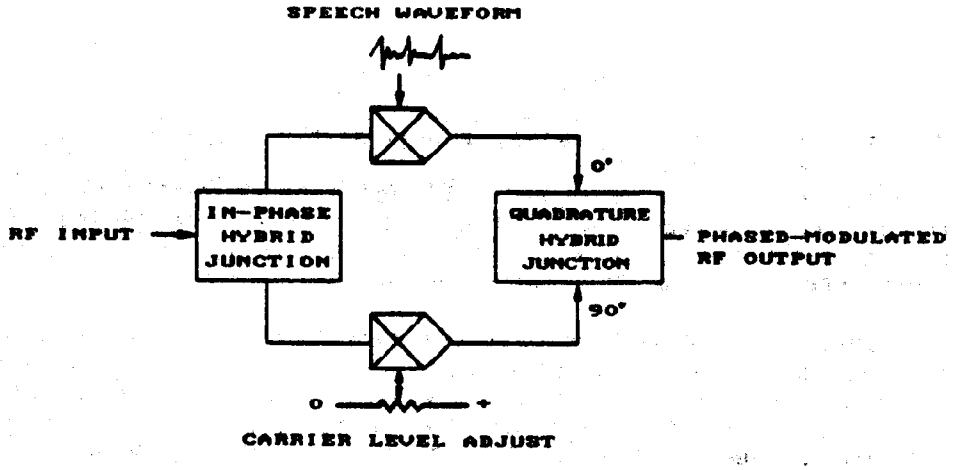


FIGURE 93. Phase modulation (voice communications).

Voice communications phase modulation illustrates (Figure 93) how the combination of mixers and hybrids can provide simultaneous but different processing in additive paths. The upper path provides voice-frequency sidebands, whereas the lower path gives electrically variable attentuation and shifts the carrier wave 90 degrees to form a conventional phase-modulated communications signal.

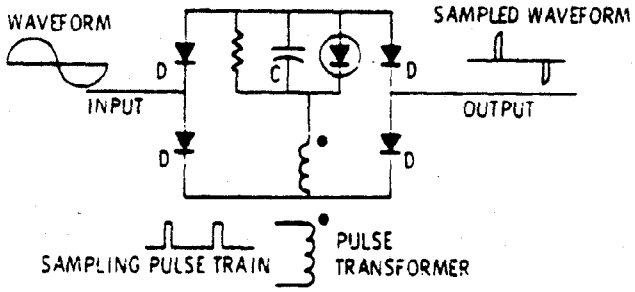


FIGURE 94. Sampling bridge.

A sampling-bridge (Figure 94) is often used to monitor very rapid periodic waveforms. This circuit (sometimes called a "boxcar" circuit) acts as a very high-speed switch, connecting input and output during a positive-going sampling pulse, and disconnecting during the remainder of the train until the next sampling pulse arrives. Followed by an integrator, the circuit is useful in automatic frequency control application. Similar circuits are often used in "sampling oscilloscopes." A more conventional mixer may also be used for sampling, but the configuration shown has advantages when the diodes "D" are well chosen for type and balance and the pulse transformer provides good isolation.

6.8.3 Physical construction. There is no single mixer configuration; this discussion is limited to diode and transistor mixers. Although the simplest mixer is no more than a single diode, this device normally contains additional elements to achieve the operations previously discussed.

At lower frequencies, the nonlinear element in which the product is generated is usually assisted with resistors, capacitors, inductors, and transformers to provide efficient coupling and filtering of the various inputs and outputs, to control the amount of power and biasing that the nonlinear element needs, and to confine the action of the mixer to the desired volume of space.

Such additional elements may not be so obvious at microwave frequencies, for pegs or ridges (by being significant parts of a wavelength) may act as capacitors, inductors, and transformers. The closed nature of the coaxial transmission line and waveguide makes the (equivalent circuit) complexity of microwave mixers even less obvious.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

The outward physical appearance of mixers will vary with frequency and power level. An RF mixer up to 2 GHz may be packaged in various microstrip or stripline flatpacks, TO cans or "bathtub" DIPs, with round pins. These parts may be further packaged in a housing with rf connectors, or connectors may be bolted right onto the flatpack. Microwave mixers are usually a stripline or microstrip assembly in a package with rf connectors, and of course they may be implemented in waveguide. To achieve high isolation, careful circuit design or packages with rf connectors should be used.

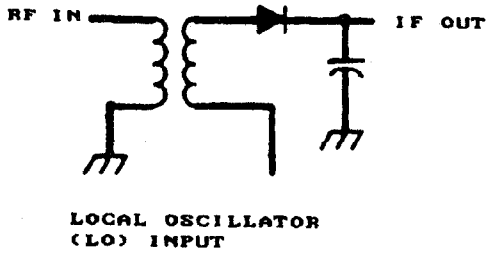
The package prevents the mixer from radiating unwanted signals and protects the mixer circuit from electromagnetic interference. Open components such as chip capacitors, thin film resistors, toroids and semiconductors are protected by the seal of the package. Flatpacks and TO cans are usually seam welded or resistance welded. DIPs are either welded or solder sealed. Microwave mixers tend to be more sensitive to overload, so mixer diodes are often mounted so that when burned out by excess power, they may be replaced without disassembling the mixer. Some diodes must have bias power from a dc connector or pin to place the operation in the best part of the diode characteristics.

6.8.4 Military designations. The military general specification for mixers is MIL-M-28837.

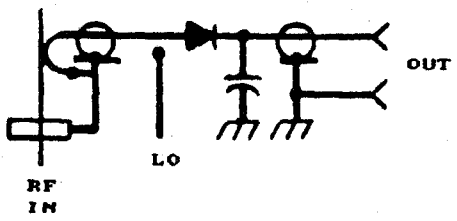
6.8.5 Electrical characteristics. The simplest form of two-input mixers uses a single diode. This is typical of use in superheterodyne receivers where varying input frequencies are translated into a fixed band of intermediate frequencies. Thus, in Figure 95A the two inputs are connected in series and the local oscillator (as previously mentioned) usually has a power level at least 10 times greater than the RF IN. A definite disadvantage exists; local oscillator current will flow back through the RF INPUT circuit. If filtering or other isolation against LO passage back to the RF source is not provided, LO energy may be radiated from the equipment antenna or other circuitry. This is presently illegal (in the general sense) in commercial and consumer equipment, and dangerously revealing in military equipment.

A practical form of shunt inputs is shown in Figure 95B. Here, the RF IN is coupled by a magnetic probe from a waveguide to the diode with a coaxial cable, whereas the LO is capacitively coupled to the coaxial line. The LO level signal determines the operating point of the diode mixer, and the difference in frequency is presented to the output. (The output circuit must provide a dc return path to ground.)

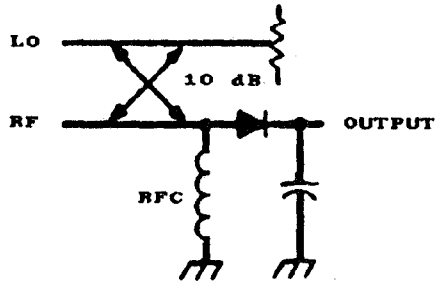
The series and shunt methods of coupling, as previously discussed, suffer from the possibility that the LO will feed back into the RF IN. Use of a directional coupler as in Figure 95C can reduce that effect. If the mixer assembly is perfectly matched to the directional coupler at the frequency of LO, all of the LO power will be absorbed in the mixer and none will reflect to the RF port of the directional coupler. (Most published examples show a 3 dB coupler, but if excess LO power is available, the 3 dB resulting loss in noise figure may be greatly reduced by using a 10 dB coupler.)



A. Series input series diode



B. Shunt inputs series diodes microwave



C. Shunt inputs series diodes directional coupler isolation

FIGURE 95. Single diode mixers.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

Several simple transistor mixers are shown in Figure 96. Bipolar transistors are shown in A and C, whereas field effect transistors are used in B and D. Shunt feed of RF INPUT and LO INPUT is used in A and B, with consequent possible interaction of the tuned RF INPUT circuit pulling the LO frequency. A degree of isolation is provided by the LO injection methods of C and D, but the results may not be satisfactory in critical applications.

Even less isolation has been used in some inexpensive transistor radios where the mixer transistor was also acting as its own local oscillator. In addition to the pulling effect, there is always the possibility of those mixers, and the circuits shown in Figure 96, permitting an undesirable amount of radiation at the LO frequency.

Transistor mixers are adjusted to the most efficient mixing point by a combination of dc bias (sometimes self-generated) and LO drive level. (As with other types of mixers, the LO INPUT level should be one or more orders of magnitude greater than the largest expected amount of RF INPUT.)

The field-effect transistor (FET) has become very attractive because of low-noise capabilities; however, in mixer service, the FET is not particularly distinguished by a low noise figure.

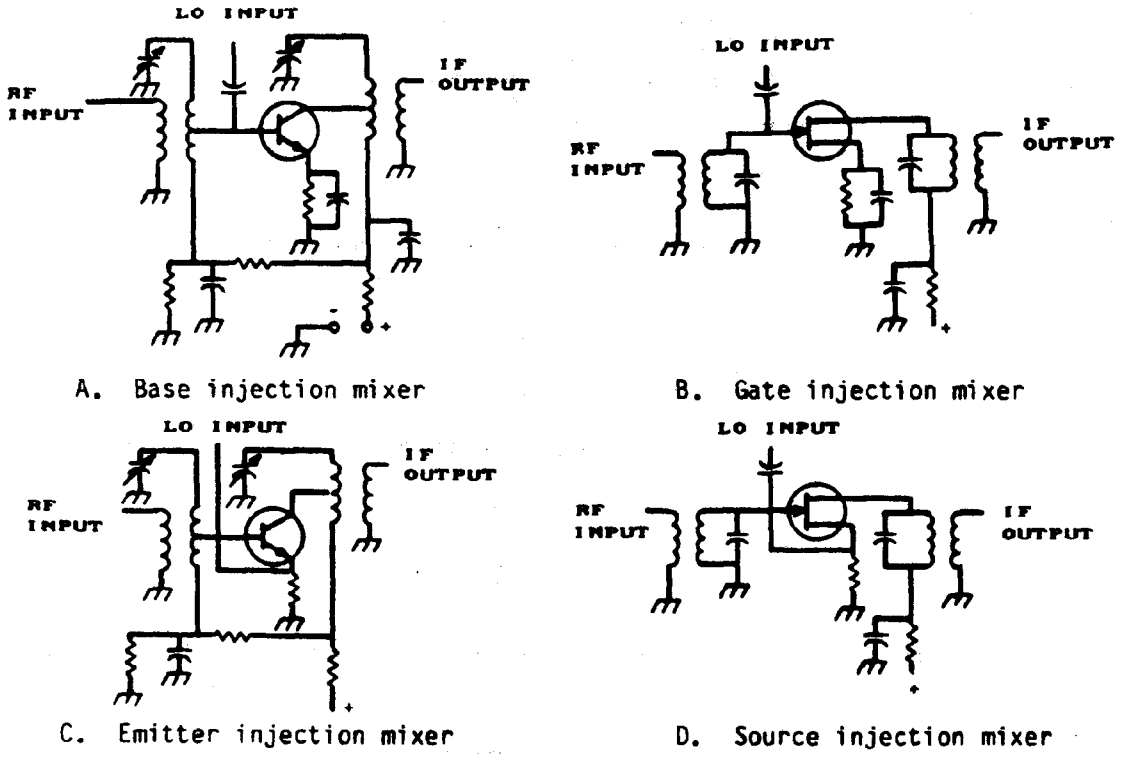


FIGURE 96. Transistor mixers.

The noise figure is the ratio of noise contributed by a device or system to the noise produced by thermal agitation in the resistive parts of the characteristic impedance of that device or system. The noise figure is usually referenced to the input circuit to make easier the understanding of the unwanted competition that the incoming signal will face.

A more serious cause of noise in mixers intended for use at microwave frequencies is noise generated by the local oscillator. Balanced modulators or mixers can be used to reduce the effect of local oscillator noise by 15 dB or more by use of two mixers in which near-equal but opposing noise components are created by interaction of the local oscillator and incoming signal. The mixer outputs, when added, perform the cancellation.

Balanced mixers have an additional advantage. Either of the inputs may be electrically isolated from the output or the other input. The various advantages are not available in all combinations in all types of balanced mixers, but the need for one or more is the usual reason for choosing the balanced configuration.

A single-balanced mixer is shown in Figure 97A. The RF IN does not appear in the output, although the LO IN does. The LO IN must provide a dc path so the diodes may conduct, and it is well if the LO IN presents a very low impedance at the desired output frequencies. Similarly, for the maximum LO power to be presented to the mixer diodes, the output circuit should present a very low impedance to the LO frequencies.

A double-balanced mixer (ring modulator) is shown in Figure 97B. Here each of the three ports is isolated from the others. A number of the unwanted outputs are cancelled in this arrangement if each port is properly terminated. For test purposes it is best to terminate each port in its characteristic impedance at all frequencies, as otherwise there may be unpredictable reflections from the termination back to the diodes. With proper terminations and high values of LO IN, low values of two-tone intermodulation (-60 to -100 dB or more) and standing-wave-ratio of the RF IN and IF OUT ports are possible. The SWR of the LO IN port will probably be much higher (perhaps 3:1) if the mixer is used over a wide frequency range.

Note that there is no inherent frequency dependence in either the single- or double-balanced diode mixers so far discussed, except that transformer-coupled ports will not pass zero frequency.

Balanced mixers do not have to be inherently wideband. The single-balanced mixer often benefits from restricted-bandwidth ports. The 10-dB conversion loss listed by one author for the wideband mixer of Figure 97A drops to the order of 2-3 dB with frequency-selective port terminations.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

A. Single-balanced

B. Double-balanced

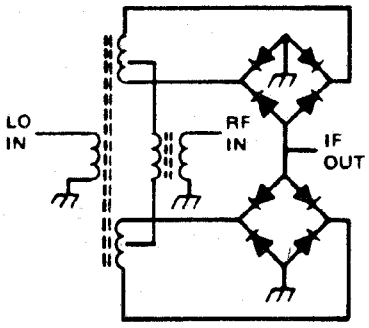
FIGURE 97. Balanced mixers.

Somewhat more complex double-balanced mixers are shown in Figure 98. A double-ring mixer is shown in Figure 98A, capable of operation from 10 MHz through 3 GHz. Though the IF OUT is dc-coupled, the response of this port falls off below 10 MHz because of dc shunting effects. This bridge is characterized by fairly high compression and desensitization level and third-order intercept point and reasonably low third-order intermodulation. These various attributes are desirable in the direction indicated, and in most types of mixers are--within limits--dependent on a high level of local oscillator power. The mixer shown has good impedance match over the frequency range on the RF IN and IF OUT ports, but will increase to perhaps 3:1 VSWR on the LO IN port.

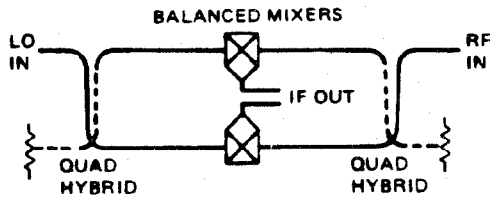
The quadrature input, double-balanced mixer shown in Figure 98B is specifically designed to present a low VSWR over a full octave (2:1) bandwidth. The LO IN and RF IN ports will have a 1.3:1, or better VSWR because of the action of the quadrature hybrids, which diverts reflected power from the mixers to dummy loads R1 and R2, respectively. The quadrature hybrids also improve the LO IN to RF IN isolation by perhaps 20 dB above that provided by the individual balanced mixers. The IF OUT ports may be connected to provide an unbalanced output, or the diodes in one mixer reversed to provide two-conductor balanced transmission line output.

The double-balanced FET mixer of Figure 98C is relatively narrow band because of the impedance-matching (Z-match) networks transforming the low-impedance output power-combiner to a higher impedance for the drains of the FETs. The combination of four FETs may be procured either as a single sealed or a single unsealed assembly.

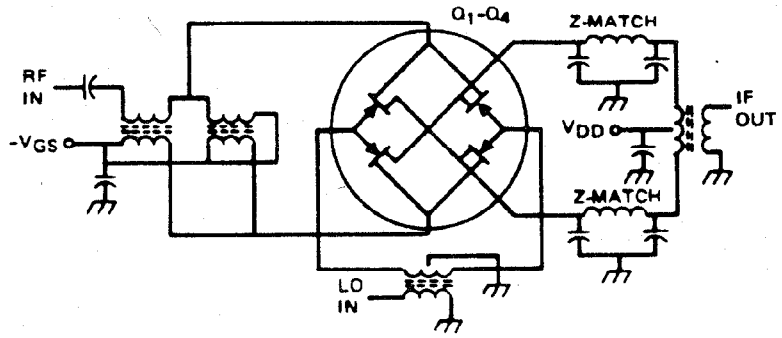
The double-balanced FET mixers are not inherently narrow band. One vendor advertises such a mixer with bandwidth limits of 200 KHz and 100 MHz. This unit, with 2 W of local oscillator power, is claimed to have a dynamic range of 155 dB, with -30 dBm third- and fifth-order intermodulation levels 140 dB below the desired products.



A. Double-ring



B. Quadrature input



C. Double-balanced FET

FIGURE 98. Double-balanced mixers.

The simple mixer may be adequate for many applications. Relatively little power is required if frequency-selective networks are used at the input and output ports. Even if such networks are used, there is a strong tendency for the inputs to affect other inputs, either by energy feeding out the other ports or by "pulling" (shifting oscillator frequency) the circuits connected to other ports. Most distortion-free operation is usually obtained with one input (usually "local oscillator") having 10 or more times the power of the other input. Intermodulation and crosstalk tend to be high.

The single-balanced mixer requires perhaps twice the "local oscillator" power of the simple mixer, but offers the advantage that most local oscillator noise may be suppressed by utilizing a 90- or 180-degree hybrid divider at the input.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

The double-balanced mixer may require up to four times the local oscillator power of the simple mixer but offers the advantages of good isolation between all ports, much cancellation of noise from the local oscillator, and great reduction of intermodulation distortion. The port impedances of double-balanced mixers may be established within fairly close limits. The quadrature input double balanced mixer offers the above advantages regardless of the LO power and termination impedances.

None of the four mixer types is optimum for all parameters. The relative advantages and disadvantages of each type for a leading microwave mixer manufacturer is summarized in Table XI.

In each type of mixer, not only intermodulation but also compression and desensitization are to be expected. Each of these effects is dependent on the level of the local oscillator power available, the level or levels of one or more other inputs, and the terminations of each port. Compression deals with the effects of just one other input. Desensitization deals with one other desired input and one other undesired input. Intermodulation deals with two other desired inputs. ("Desired" here means that the inputs are in the passband of desired signals.)

Compression

Compression is usually defined as the minimum input level at which the consequent output will become 1 dB less than straight-line prediction of a proportional output-power/input-power ratio.

Desensitization level

Desensitization level is the minimum level at which the output level, corresponding to a weak desired input, is reduced by 1 dB, compared with when the undesired signal is absent.

Intermodulation

The two-tone method is the usual form of intermodulation test, wherein two equal-power desired frequencies are the other input, and particularly the third and fifth order intermodulation products are examined. Other products may on occasion be important, but the third and fifth order are most often annoying. The third order products have the frequencies $2f_1 \pm f_2$ and $2f_2 \pm f_1$. The fifth order products may include harmonics, the sum of whose orders add up to five. In some mixers it is possible to detect intermodulation products above the 250th.

TABLE XI. Performance comparison for four basic mixer types*

Mixer Type 1/	No. of Diodes 2/	VSWR 3/	Conversion Loss 4/	LO/RF Isolation 5/	Bias	Spurious Rejection |m-n| = 1 6/	Harmonic Suppression 7/	Intercept Point (dBm) 8/	Dynamic Range	IF Bandwidth	Image Focusing Port 11/
90° hybrid	2	Good	Lowest	Poor	No	Poor	Poor	+15	High	Wide	LO
single balanced	2	Good	Lowest	Poor	Yes	Fair	Fair	+30 9/	Highest	Wide	LO
180° hybrid	2	Poor	Low	Good	Yes	Fair	Good	+15	High	Wide	RF
single balanced	2	Poor	Low	Good	Yes	Fair	Good	+30 9/	Highest	Wide	RF
Quad input double balanced	4	Good	Low	Very good	No	Good	Fair	+18	High	Wide	Internal load
Double balanced	4	Poor	Low	Very good	No	Good	Very Good	+18	High	Extremely wide 10/	RF
	4	Poor	Low	Excellent	No	Good	Very Good	+18	High	Extremely wide	RF

See following page for notes.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

TABLE XI. (Continued)

NOTES:

1/ Data is from typical L-band or S-band units.

2/ 2-diode types typically require +7 dBm LO power for best performance; 4-diode types require +10 dBm LO power.

3/ VSWR: Poor: 2.5:1 typical
Good: 1.3:1 typical

4/ Conversion loss: Lowest: 5-7 dB typical
Low 7-9 dB typical

5/ LO/RF isolation: Poor: 10 dB typical
Good: 20 dB typical
Very Good: 25-30 dB typical
Best: 35-40 dB typical

6/ Spurious rejection (m x n): where $|m - n| = 1$; i.e., 1 x 2, 2 x 1, 2 x 3, 3 x 2, etc.
Poor: partial rejection of most $|m - n| = 1$ spurs
Fair: partial rejection of most $|m - n| = 1$ spurs (bias adjust will suppress some spurs even further)
Good: potentially rejects all $|m - n| = 1$ spurs.

7/ Harmonic suppression: Poor: partial rejection of LO/RF even harmonics
Fair: partial rejection of LO/RF even harmonics (bias adjust will suppress some harmonics even further)
Good: can reject all LO even harmonics
Very Good: can reject all LO and RF even harmonics

8/ Intercept point: Typical third-order intercept point is 6 to 9 dB above the LO power.

9/ This intercept point can only be achieved by using the optimum load line biasing technique and increasing the LO power to approximately +23 dBm.

10/ The IF OUTPUT bandwidth for this series of balanced mixers overlaps the RF INPUT range of the units. Bandwidths from dc to 4 GHz are available.

11/ Image focusing: defines where image signal energy is directed to the LO port, RF port, or the internal termination. Where this image signal is focused can affect mixer performance in phase and amplitude tracking systems.

VSWR

The VSWR specifies the degree of impedance match between the mixer LO and RF ports and the 50-Ω system in which the mixer is used.

One of the most desirable characteristics of the 90-degree hybrid mixer is its excellent VSWR over the full performance range of the hybrid. When a signal is fed into either input, any reflections from the similar mixer diodes will combine at the other input, providing a low VSWR at either port. The input reflection coefficient for a 90-degree hybrid mixer having diode reflection coefficient, ρ_1, ρ_2 is given by:

$$\rho(\text{mixer}, 90°) = \frac{\rho_1 - \rho_2}{2}$$

If $\rho_1 = \rho_2$ (if the diodes are balanced), whatever their value, the input remains matched.

If the mixer uses a 180-degree hybrid, any reflections from the similar mixer diodes will be focused back to the input port. Unless the diodes look like a very good match, the input VSWR will be poor. The input reflection coefficient for a 180-degree hybrid mixer having diode reflection coefficients ρ_1, ρ_2 is given by:

$$\rho(\text{mixer}, 180°) = \frac{\rho_1 - \rho_2}{2}$$

Diode impedance is a function of LO power; therefore, VSWR will be affected by LO drive level for this mixer. If the diode impedances are equal, the input VSWR equals the return loss for one diode.

A quadrature input double-balanced mixer has the same good VSWR characteristics as the 90-degree hybrid mixer, because it uses 90-degree hybrids as the coupling mechanism to the diodes.

A double-balanced mixer exhibits poor VSWR characteristics. It uses 180-degree hybrids as the coupling networks to its diodes and, therefore, depends on a good diode match to provide good input VSWR. In addition, the LO VSWR is frequently degraded even further by the necessity for extracting the IF signal from the LO coupling network. VSWRs of 3:1 or greater are not uncommon for double-balanced mixers.

LO/RF leakage

The LO/RF isolation is a measure of the LO leakage at the RF port of the mixer.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

The 90-degree hybrid mixer has poor isolation characteristics because LO/RF isolation is dependent upon diode match. Any LO energy not absorbed by the diodes is directed to the RF port. The LO/RF isolation for a 90-degree hybrid mixer having diode reflection coefficients ρ_1, ρ_2 is given by:

$$\text{Isolation (dB)} = 20 \log_{10} \left(\frac{2}{\rho_1 + \rho_2} \right)$$

If the diode impedances are equal, the isolation equals the return loss for one diode. Because diode impedance is a function of LO power, the isolation will be sensitive to LO drive.

A 180-degree hybrid mixer has good isolation characteristics; dependent only on the diode balance, regardless of their impedance values. The LO/RF isolation for a 180-degree hybrid mixer having diode reflection coefficients ρ_1 and ρ_2 is given by:

$$\text{Isolation (dB)} = 20 \log_{10} \left(\frac{2}{\rho_1 - \rho_2} \right)$$

LO power will not affect isolation if the diodes continue to track each other.

The quadrature input double-balanced mixer will have good LO/RF isolation, regardless of LO power, because the phasing of the LO and RF networks directs any unused LO energy to an internal termination; it will never show up at the RF port as LO/RF leakage. The good LO/RF isolation is independent of LO power level. The quadrature input double-balanced mixer is the only one of the four basic mixers that will provide, simultaneously, good VSWR and good isolation regardless of LO drive level.

The double-balanced mixer has the same good LO/RF isolation characteristics as the 180-degree hybrid-balanced mixer because it uses 180-degree hybrids as the coupling networks to its diodes. As long as the diodes track each other, the LO/RF isolation will be good and will remain independent of LO power.

For many receiving applications, the amount of LO power leaking out the RF port is important because it will ultimately be reradiated out of the antennas. A biased mixer can be very useful in such applications because the LO power and resultant reradiation can be significantly lowered while still maintaining good conversion loss. In fact, using biased 90-degree balanced mixers may actually result in lower radiated energy from the antenna and better overall performance than when using many double-balanced mixers. Double-balanced mixers, at first, would appear to offer better results because of their higher LO/RF isolation specifications, but the following comparison, shown in Table XII, will serve to illustrate why this may not be the case.

TABLE XII. Application showing a comparison of biased versus unbiased mixer operation

Biased 90-degree Balanced Mixer		Typical Double-Balanced Mixer	
Conversion loss:	8.0 dB max (0 dBm LO) 8.5 dB max (-4 dBm LO)	Conversion loss:	9.0 dB max
LO power:	0 dBm at 1.8 mA bias -4 dBm at 1 mA bias	LO power:	+10 dBm (no bias option)
LO/RF isolation:	10 dB typical (6 dB min)	LO/RF isolation:	17 dB typical
LO leakage at RF port:		LO leakage at RF port:	-7 dBm (+10 dBm LO)
LO, RF VSWR:	1.65 typ	LO, RF VSWR	3.0 typ

The biased 90-degree balanced mixer radiates 7 dB less LO power at the antenna terminal than the double-balanced mixer. This improved leakage performance is even more impressive when the rest of the specifications are examined. When compared with the typical double-balanced mixer, the biased, 90-degree hybrid mixer has:

a. Lower LO/RF leakage levels

b. Equal or better conversion loss

c. 10 to 14 dB less LO drive required

d. Much lower VSWR.

Conversion loss and noise figure.

Conversion loss and noise figure are two of the most important and closely-related mixer parameters. The primary requirement for a mixer is to provide the maximum IF output power with the minimum RF INPUT power while generating the least amount of noise. Minimum losses in the RF, LO coupling networks, high quality Schottky-barrier (hot carrier) coupling networks to the diodes, broad-band diode matching and optimum local oscillator drive, all contribute to the good conversion loss and low noise figure.

Conversion loss of the mixer is a measure of the power at the IF frequency relative to the power at the rf input frequency. It is a function of the mixer alone.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

The mixer noise figure is commonly stated at a particular IF frequency when measured in a system containing an IF amplifier following the mixer. This IF amplifier contributes to the measured noise figure. The mixer diodes also can contribute a small amount of noise energy to the noise figure. The measured noise figure (NF_o) in decibels is related to the mixer conversion loss (L_c), the IF amplifier noise figure (F_{if}), and the mixer diodes noise-temperature ratio (T) by:

$$NF_o = L_c + 10 \log_{10} (F_{if} + T-1) \quad (1)$$

Where L_c is mixer conversion loss in dB
F_{if} is the noise factor (a numeric ratio) of the IF amplifier
T is the effective diode noise-temperature ratio.

T can be as low as 0.85 for modern high-quality hot-carrier diodes, but a more typical value is 1.0. The relationship between measured noise figure and conversion loss then simplifies to:

$$NF_o = L_c + NF_{if} \quad (2)$$

Where NF_{if} is IF preamplifier noise figure in dB.

When T = 1, the mixers' single-sideband noise figure is the same as its conversion loss. Equation (2) then shows that the overall noise figure is obtained by adding the noise figure of the IF preamplifier directly to the mixer conversion loss.

It is instructive to observe what happens to the overall noise figure when T = 0.85. By equation (1):

$$NF_o = L_c + 10 \log_{10} (F_{if} + 0.85 - 1)$$

When NF_{if} is equal to 1.5 dB:

$$NF_o = L_c + (1.413 - 0.15)$$
$$= L_c + 1.02 \text{ dB}$$

The combined preamplifier/diode noise contribution is only 1.02 dB, which is 0.48 dB less than that produced by the preamplifier alone.

Biased mixers

In many instances the optimum LO power is not available for best conversion loss. The conversion loss of an unbiased mixer degrades rapidly as the LO power decreases. This is due chiefly to diode mismatch at the rf and IF frequencies. The diode mismatch can be minimized and the conversion loss dramatically improved by biasing the diode with a dc current.

6.8.6 Environmental considerations. Every adverse environment will unfavorably affect some type of mixer. Generally balanced mixers will keep more desirable characteristics in the face of adverse environments than unbalanced mixers, but even their characteristics may be badly degraded.

Heat or cold causes at least temporary dimensional change, causing corresponding changes in inductance and capacitance. Values of permeability, dielectric constant, resistance, and the characteristics of semiconductors, likewise, change with temperature change. It is particularly important that balanced mixers be built so that the balance is maintained over the operating temperature range.

Repetitive temperature changes tend to stress mechanical structures both to reduce the strength of joints and break seals. In some cases, stresses will be relieved less disastrously (as in coil interturn stresses), and will result in a more stable, though somewhat changed, product.

With the exception of high-power modulated radio amplifiers, most mixers do not involve high enough voltages to be affected by high-altitude breakdowns. Barometric pressure changes, thus, have their major effect in aneroid stressing of sealed cases. Temperature cycling will also cycle pressure, except that it is the pressure within the case that is changed. Where cases have poor seals, the refrigeration effect of gas expansion will tend to condense water vapor within the case.

If not properly finished and sealed, the mixers can be corroded by salt spray and damaged by sand and dust.

Mixers (generally being aggregations of parts) will, even in the best cases, be adversely affected by the forces of shock, vibration, and acceleration in direct proportion to the degree of the force effect on the component parts. Where balanced circuits are involved, the effect may seem greatly magnified if one member of a balanced pair is affected while the other is not.

6.8.7 Reliability considerations.

6.8.7.1 Failure modes. Reliability of mixers as a topic literally covers the field of electronic reliability. It is an oversimplification to say that the reliability of a mixer is the product of the reliability of the component parts, unless the parts count is taken in the same way as the parts count of an equipment. In equipments, the reliability of each connection--as well as the reliability of the discrete parts--must be considered.

A major consideration is the method of construction. For instance: is the assembly on a printed circuit board? Is there conformal coating? Are the parts embedded in plastic encapsulation? Is there stress relief to prevent embedment crushing of the parts? These and similar questions must be answered if an accurate estimate of reliability and reliability factors is to be considered.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

A brief review of the basics of mixers reveals that the action of mixing takes place in one or more devices that do not have proportional current versus voltage relationships. A number of passive parts may be included in the mixer assembly to provide the proper magnitudes of voltages and currents, and proper phases to give the desired mixer action. As a general rule, the reliability of the mixing device is more in question than the reliability of passive (linear) parts.

The discussion of reliability may be divided into three parts. Most commonly reliability is discussed in terms of catastrophic failure, where a permanent and disabling fault occurs. There are also drift failures, where a desired function is permanently impaired because of a permanent shift of the characteristics outside of some critical parameter limit. Finally, there are reversible changes in a critical characteristic, usually in response to an environment such as temperature, often with full recovery of the desired function when the forcing stress is removed. All of these factors should be considered in a discussion of mixer reliability.

Because the field of mixers and the parts that go into making the assembly is broad, only some of the reliability factors are discussed in this section. The reader is urged to scan other sections of the manual covering parts used (diodes, inductors, and others) for discussion of some of the more subtle reliability considerations affecting those parts.

The input range within which diodes and transistors will operate properly is very limited. Certain minimum input levels (such as local oscillator input) are needed to bring the mixer device into the optimum operating range, but as little as two or three times the proper driving level may be sufficient to burn out the mixer device. Thus, an excessive drive level may cause catastrophic failure, and if the level is too low, it will cause a temporary degradation in operating characteristics until corrected.

Environments affect mixers, usually as a result of temperature changes, causing characteristics to shift, or by the adverse effect of moisture on structures or electrical resistances. Changes in temperature tend to cause stress and fatigue in mechanical joints, often resulting in the failure of hermetic seals. Moisture that reaches semiconductor junctions can cause direct or indirect contamination. These examples show the desirability of conducting seal leak tests after thermal cycling semiconductors.

Such screening is only a small part of the screening desirable with semiconductors. Screening provisions of MIL-S-19500 are very desirable for transistors and diodes. Although screening is intended to weed out individual devices of poor workmanship, it is also helpful in identifying poor designs or material. However, there is no substitute for good electrical and mechanical design as a foundation for reliability.

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

Often, particularly in the microwave frequency region, the discrete parts of the mixer assembly are packaged in containers that are bulky enough to cause excessive stray inductance and capacitance.

It is tempting, and in many cases essential, to use unpackaged discrete parts, but the part protection previously supplied by the package must be furnished in some way. This goal is commonly met by arranging mixer ports to accomodate transmission lines carrying signals into and out of a sealed enclosure containing the mixer assembly. A properly terminated transmission line does not either narrow the bandwidth of the port or add reactance. The usual low impedance eases the metal-ceramic seal problem of added shunt capacitance to conductors passing through an insulator mounted in the wall of a metal container. The sealing of a large container, though, can be much more difficult than the sealing of a diode or transistor package.

After the reliability factors of the parts making up the mixer assembly have been considered, the design and workmanship of the assembly itself deserve investigation. A sealed package must have a good mechanical design. The sealing operation itself must not damage the assembly and the finish of the package must meet paint or plating requirements.

Some aspects of poor design are inherent to the assembly; others are triggered by external factors. One example of the inherent poor design might be the use of a point contact diode when a Schottky barrier (perhaps with bias) will provide equivalent electrical performance with increased reliability. Specification MIL-S-45743 gives a great deal of guidance on aspects of conventional solder assembly which affect reliability.

Poor design triggered by external factors is common, and high temperature is, perhaps, the leading cause. An extremely common example, is the use of lead-tin solder in assemblies later processed at temperatures above 180 °C. Temperature shock and vibration often reveal that a bad choice of structure or materials was made. Some apparently good materials in an enclosure may outgas and contaminate the enclosure. Age affects materials, too, usually adversely.

When considering high-temperature environments, the combination of age and heat has several adverse effects. In semiconductors this combination is the accelerator of "purple plague," where a joint of aluminum and gold conductors swells into a brittle compound of gold and aluminum.

Wire insulation similarly has a temperature rating, with the temperature being the combination of the ambient plus the temperature rise of the device used. (In low-level mixers, the temperature rise may usually be ignored.) The temperature codes often used are as listed in the table herein:

6.8 MICROWAVE DEVICES, ELECTRICAL MIXERS

TABLE XIII. Wire insulation temperature rating

Temperature	Military	Commercial	Notes
85 °C	Q	--	Minimum limit for some types of organic insulations
90 °C	--	O	--
105 °C	R	A	--
130 °C	S	B	Usual limit for organic insulations
155 °C	V	--	Usual limit to avoid outgassing and weight loss
170 °C	T	--	Inorganic insulation (mica, asbestos, etc.) only
180 °C	--	H	

The military code U and commercial code C are occasionally used, but usually must be individually and additionally described.

6.8.7.2 Screening. Ideally, screened parts should be used in a mixer assembly to achieve the highest possible quality. However, that is not always possible. Therefore, the intent of the military specification is to perform screening of the entire mixer to achieve a high degree of quality and allow the manufacturer to use unscreened parts. Screening steps should include internal visual examination, high temperature storage for at least 24 hours at the maximum rated storage temperature, thermal shock for 10 cycles at minimum and maximum rated storage temperature, hermetic seal (if applicable) fine and gross leak, and high temperature power conditioning with power injected directly into the diodes at 60 Hz, if possible, or at the frequency of operation. Initial and final electrical tests with calculation of delta limits (percent change) should be specified. Additional tests might include radiographic inspection, and electrical tests at high and low operating temperatures.

6.9 Circulators and Isolators.

6.9.1 Introduction. A circulator is a three-port device used to direct and separate rf signals. An rf signal entering a port of a circulator interacts with a magnetized ferrite and is directed out the next port in the direction of rotation. A very small signal leaks through to the port in the opposite direction of rotation. This leakage signal, when compared with the output signal in the direction of rotation, is called the isolation. The loss of input signal at the output port due to internal mismatch and dissipation is called the insertion loss. If a short circuit is put on the first port in the direction of rotation, the signal travels to the next port in the same direction. The power will be reduced by the insertion loss of the two paths of travel.

When the third port of a circulator is terminated in a matched load, the device is an isolator. An isolator is a unidirectional device, somewhat like a diode, that presents a matched impedance to a source when it is connected to a mismatched load. The worst case of mismatch presented by the isolator may be measured by placing a variable phase short circuit on the output of the isolator and measuring the input VSWR as a function of the phase of the output short circuit.

6.9.2 Usual applications. Isolators are used predominately as matching devices and are almost mandatory in power and frequency generating circuits. Certain rf power amplifiers, varactor multiplier chains, and Gunn device circuits would not be possible without these devices.

Isolators and circulators are relatively narrow band devices, as can be seen from a plot of isolation versus frequency. Typically, temperature shifts this curve in frequency. Device characteristics must be controlled somewhat beyond the frequency band used, and the isolation should be tested at the operating temperature. Interaction of the device's magnetic field with nearby parts should not occur, nor should nearby parts interact magnetically with the isolator or circulator.

Circulators and isolators are generally low-power to medium-power devices. Air- and liquid-cooled devices are possible when necessary. Special consideration must be given to achieve adequate cooling of the ferrite and to assure that the power level will not cause ferrite limiting or arcing. A duplexer is a special case of a circulator using hybrids and ferrite in a four-port device to combine both transmit and receive signals to the antenna.

6.9.3 Physical construction. Coaxial circulators are usually a stripline configuration with coaxial connectors. The ground planes are the circulator body. A ferrite is placed in the center of the center conductor with magnets mounted into the ground planes. Connectors are usually flange-mounted with solder connection to the center conductor. Small pieces of dielectric material are sometimes epoxied on or near the center conductor for matching. Isolators use the same general construction with a built-in or connector-mounted resistive load.

6.9 MICROWAVE DEVICES, CIRCULATORS AND ISOLATORS

Waveguide circulators have the ferrite, usually on a stepped transition, mounted in the waveguide. A three-port waveguide Y-junction circulator is shown in Figure 99A. Magnets are mounted on the waveguide walls and sometimes use a "C" type magnet for magnetic field concentration. Isolators use insert or flange-mounted loads. A four-part waveguide, Faraday rotation circulator is shown in Figure 99B. Energy entering port 1 will leave at port 2, but energy entering port 2 will leave from port 3. Similarly, energy entering port 3 will leave at port 4, and energy entering port 4 will leave at port 1. Such a device is useful for coupling an antenna to a receiver-transmitter system.

Drop-in isolators are used in microstrip circuits to provide input and output matching of the circuit and internal components. They are purchased complete with magnets and a load in place. The flat input and output leads are matched and may be soldered directly to the microstrip trace. A stripline drop-in Y-junction circulator is shown in Figure 99C.

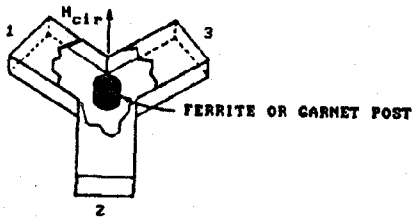
Air-gap in the ferrite-to-magnet interface is minimized for all types of construction; and ferrite placement, spacing, and conductor interfaces are critical for high performance units. Magnetic field strength, choice and placement of ferrite material, and matching of materials are the key to a good design.

6.9.4 Military designation. Military circulators are found in MIL-C-28790. Military isolators are found in MIL-C-28791.

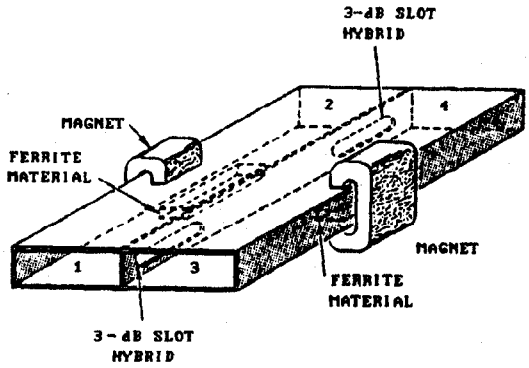
6.9.5 Electrical characteristics. Primary parameters of a circulator are VSWR, isolation, and insertion loss. Typical values for coax units are VSWR = 1.2:1 maximum, isolation = 20 dB (at band edge) minimum, and insertion loss = 0.4 dB maximum across the specified frequency range. Isolation should peak, and VSWR and insertion loss should be at a minimum at band center because the device is a tuned unit. Waveguide units are similar, except that VSWR will be less. Higher isolation can be achieved by cascading units, but at the expense of higher insertion loss. Broader bandwidth can be achieved by overlapping isolation curves of two units. Four-port (or more) units are available.

6.9.6 Environmental considerations. The environmental considerations most critical for circulators and isolators are mechanical- and temperature-related. Efficient rotation of the electromagnetic field will be degraded by anything which may cause the physical location of the ferrite or magnet to move. This includes mechanical shock, repeated thermal shocks or high accelerations.

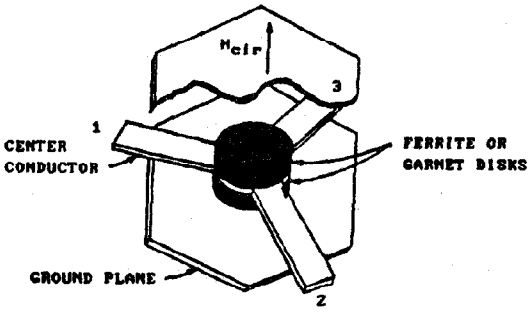
Solvents and extremely high or low temperatures may weaken adhesives used to mount ferrites and magnets. Braze joints, flanges, and bodies of waveguide circulators and isolators must withstand humidity and salt corrosion. High power circulators and isolators may have to be pressure tight and withstand high power at high altitudes. Temperature extremes affect VSWR, isolation, and insertion loss and derate the power handling capabilities of isolator loads.



A. Waveguide Y-junction circulator



B. Four-port Faraday circulator



C. Stripline Y-junction circulator

FIGURE 99. Construction of circulator.

6.9 MICROWAVE DEVICES, CIRCULATORS AND ISOLATORS

6.9.7 Reliability considerations.

6.9.7.1 Failure modes. Circulator and isolator failures usually exhibit themselves in reduced performance of isolation, VSWR, and insertion loss over the bandwidth of interest. The cause is usually mechanical in nature. Movement of the ferrite or magnet or tuning elements will cause rapid degradation of electrical parameters. Mismatch of an isolator load element will cause signal reflections to rotate to the output or input port. In high-peak power applications, arcing may occur due to spacing, burrs, or as a result of power and altitude conditions. Isolator load destruction may result from misapplication of power to the isolator, large reflections of power from the output, or inadequate heat dissipation due to lack of derating or poor heat sinking.

6.9.7.2 Screening. Initial electrical measurements of isolation, VSWR, and insertion loss followed by thermal shock (10 cycles) or temperature cycling (25 cycles) followed by final electrical measurements with calculation of delta parameters is an excellent screen for circulators and isolators. A high-power, high-temperature burn-in for isolators is effective for screening potential failures of load elements. If this is not possible, a high-temperature storage for 100 hours at the maximum rated temperature of the device should be done. Electrical testing of isolation, VSWR, and insertion loss at low and high temperatures may be required for extremely narrow band devices.

USA Information Systems, Inc.
(757)491-7525 / (800)872-8830
www.usainfo.com