PRICE $2.50

MOTOROLA

high-speed... SWITCHING TRANSISTOR HANDBOOK

- TRANSISTOR THEORY
- WORST-CASE CIRCUIT DESIGN
- APPLICATIONS

SWITCHING TRANSISTOR HANDBOOK

first edition

Technical Editor

William D. Roehr
Applications Engineering

Associate Editor

Darrell Thorpe
Technical Information Center

Contributors

Applications Engineering Dept.
William D. Roehr
James F. Kane

Quality Assurance Dept.
Joseph Flood

University of Arizona
Douglas J. Hamilton

Technical Assistance

T. Byrd, F. Loewy, J. Schroeder
Switching Circuits Group

Editorial Assistance

A. Ralph Greenburg, Leo L. Lehner, Lothar Stern,
Reuben Wechsler, Charles S. Meyer

Copyright 1963
by
MOTOROLA, INC

Circuit diagrams are included as a means of illustrating typical semiconductor applications, and the complete information necessary for constructional purposes is not necessarily given. The information in this handbook has been carefully checked, and is believed to be entirely reliable, but no responsibility is assumed for inaccuracies. Furthermore, such information does not convey to the purchaser of the semiconductor devices described any license under the patent rights of Motorola, Inc. or others.

PREFACE

The intent of this book is to supply answers to questions most often asked by the circuit design engineer about transistor switching characteristics. Engineering students and senior technicians will also find this book valuable because emphasis is placed upon filling the void between transistor theory and practical applications.

In Chapter 1, the fundamental differences between various transistor types are explained and the reasons for compromises between certain characteristics are discussed. To provide an introductory background, an elementary transistor model and the significance of resistivity profiles and geometry are introduced. Old-timers in the transistor field as well as the novice should find this treatment interesting.

Chapter 2 is written primarily for those unfamiliar with switching circuits and is designed to familiarize them with the differences between various switching modes and coupling techniques.

The dc on and off characteristics of a transistor are fully discussed in Chapters 3 and 4. The primary objective here is to equip the circuit designer with sufficient background information about transistor principles to enable him to obtain limit values of important transistor characteristics from a standard data sheet.

The transient characteristics of transistors are developed in Chapter 5 using a charge control viewpoint which proves to be a powerful tool for simplifying the presentation. Discrepancies between measured switching times and those computed from standard formulas are illustrated and discussed in this chapter.

Chapter 6 is designed to make reliability — which is often a vague term — meaningful to the circuit designer. Methods of enhancing overall system reliability by proper circuit design are discussed.

A unique feature of this book is the worst-case design procedures and illustrated examples in Chapter 7. The synthesis techniques developed not only provide a "cook book" method for building standard "workhorse" circuits, but also illustrate the type of thinking which is required to make a paper design actually meet performance specifications.

Chapters 8 and 9 illustrate many of the features and performance characteristics which can be obtained by operation in the current mode or avalanche mode.

If this book enables you to do your job better, or improves your understanding of transistors, it has fulfilled its purpose.

CONTENTS

page

LIST OF SYMBOLS vii

CHAPTER I *EFFECTS OF TRANSISTOR CONSTRUCTION ON ELECTRICAL CHARACTERISTICS* 1

- **1-1** THE IDEAL SWITCH 1
- **1-2** BASIC TRANSISTOR LIMITATIONS 2
- **1-3** BASIC TRANSISTOR PHYSICS AND CHARACTERISTICS 4
- **1-4** COMPARISON OF TRANSISTOR TYPES 14

CHAPTER II *SWITCHING MODES* 29

- **2-1** SATURATED MODE OPERATION 29
- **2-2** SATURATED MODE COUPLING CIRCUITS 33
- **2-3** CURRENT MODE OPERATION 37
- **2-4** COUPLING TECHNIQUES FOR CURRENT MODE CIRCUITS 39
- **2-5** AVALANCHE MODE 42

CHAPTER III *TRANSISTOR CHARACTERISTICS INFLUENCING OFF CONDITION DESIGN* 43

- **3-1** LEAKAGE CURRENTS AND THRESHOLD VOLTAGE 43
- **3-2** FACTORS CONTRIBUTING TO LEAKAGE CURRENTS IN A PN-JUNCTION 46
- **3-3** PREDICTING WORST-CASE I_{CBO} 51
- **3-4** PREDICTING WORST-CASE I_{EBO} 56
- **3-5** TRANSISTOR CUTOFF BEHAVIOR WITH BOTH JUNCTIONS CONNECTED 58
- **3-6** ESTIMATING THE BASE LEAKAGE CURRENT (I_{BL}) 62
- **3-7** RELATIONSHIP OF VOLTAGE BREAKDOWN AND SWITCHING LOAD LINES 63
- **3-8** AVALANCHE BREAKDOWN THEORY 65

CHAPTER IV *TRANSISTOR CHARACTERISTICS INFLUENCING ON CONDITION DESIGN* 69

- **4-1** THE COLLECTOR SATURATION REGION 70
- **4-2** BASE-EMITTER SATURATION CONSIDERATIONS 74
- **4-3** THE EMITTER-BASE JUNCTION 75
- **4-4** DERIVING LIMIT V_{BE} CURVES FROM TYPICAL DATA 78
- **4-5** THE COLLECTOR-BASE JUNCTION 82
- **4-6** THE COLLECTOR-EMITTER VOLTAGE 83
- **4-7** OBTAINING LIMIT COLLECTOR SATURATION CHARACTERISTICS 86

CHAPTER V *TRANSIENT CHARACTERISTICS OF TRANSISTORS* 91

- **5-1** DEFINITION OF TRANSITION TIMES 91
- **5-2** CHARGE CONTROL THEORY 93
- **5-3** TURN-ON DELAY TIME 94
- **5-4** RISE TIME 100
- **5-5** STORAGE TIME 115
- **5-6** FALL TIME 128
- **5-7** TOTAL CONTROL CHARGE (Q_T) 131

CHAPTER VI *RELIABILITY CONSIDERATIONS FOR THE CIRCUIT DESIGNER* 143

- **6-1** RELIABILITY OF TRANSISTORIZED EQUIPMENT 143

page

6-2 ACHIEVING TRANSISTOR RELIABILITY 145
6-3 CAUSES OF FAILURE 146
6-4 FAILURE RATES AS FUNCTION OF TIME 148
6-5 MEASUREMENT OF TRANSISTOR RELIABILITY 151
6-6 SPECIFYING RELIABILITY ASSURANCE 159
6-7 ACHIEVING RELIABLE SWITCHING SYSTEM PERFORMANCE 163
6-8 PRECAUTIONS FOR THE EQUIPMENT MANUFACTURER 167

CHAPTER VII *SATURATED MODE CIRCUITS* 171

7-1 INVERTER DESIGN 172
7-1-1 THE LOAD CIRCUIT 173
7-1-2 COUPLING FOR MAXIMUM TRANSFER EFFICIENCY 182
7-1-3 THE INPUT NETWORK 185
7-1-4 MINIMIZING SWITCHING TIME OF RTL STAGES 186
7-1-5 TOLERANCES OF PASSIVE ELEMENTS 187
7-1-6 RCTL DESIGN EXAMPLE 189
7-2 FLIP-FLOP DESIGN 198
7-2-1 SYNTHESIS EQUATIONS 199
7-2-2 TRIGGERING 201
7-2-3 CROSS COUPLING 203
7-2-4 DESIGN VARIATIONS 204
7-2-5 FLIP-FLOP SYNTHESIS PROCEDURE 204
7-2-6 FLIP-FLOP DESIGN EXAMPLE 208
7-3 THE ASTABLE MULTIVIBRATOR 215
7-3-1 DC CIRCUIT ANALYSIS 217
7-3-2 TRANSIENT CIRCUIT ANALYSIS 218
7-3-3 DEVELOPMENT OF SYNTHESIS PROCEDURE 220
7-3-4 ASTABLE MULTIVIBRATOR DESIGN EXAMPLE 228
7-4 THE MONOSTABLE MULTIVIBRATOR 235
7-4-1 ANALYSIS 238
7-4-2 SYNTHESIS 240
7-4-3 USE OF A CLAMP CIRCUIT 241
7-4-4 MONOSTABLE MULTIVIBRATOR DESIGN PROCEDURE 245

CHAPTER VIII *CURRENT MODE CIRCUITS* 253

8-1 FUNDAMENTALS OF CURRENT MODE CIRCUITS 254
8-1-1 DC ANALYSIS OF THE CURRENT MODE INVERTER 254
8-1-2 ADVANTAGES OF CURRENT MODE CIRCUITS 256
8-1-3 TRANSIENT RESPONSE 257
8-1-4 DC COUPLING TECHNIQUES 261
8-2 CURRENT MODE INVERTER 265
8-2-1 SYNTHESIS 266
8-2-2 CURRENT MODE INVERTER DESIGN EXAMPLE 269
8-3 CURRENT MODE FLIP-FLOP 274
8-4 CURRENT MODE ASTABLE MULTIVIBRATOR 276
8-5 CURRENT MODE MONOSTABLE MULTIVIBRATOR 280
8-6 CURRENT MODE SCHMITT TRIGGER 283

page

CHAPTER IX *AVALANCHE MODE SWITCHING* 285
9-1 STATIC CHARACTERISTICS OF THE AVALANCHE REGION 286
9-2 TRANSIENT CHARACTERISTICS 288
9-3 OPERATION OF A SINGLE TRANSISTOR CIRCUIT 290
9-4 GENERAL CIRCUIT DESIGN CONSIDERATIONS 293
9-5 AVALANCHE MODE CIRCUITS 294
SPECIFICATIONS AND TRANSISTOR SELECTION INFORMATION 305
APPENDIX I 342

Abbreviations and Symbols

Symbol	*Description*
BV_{CBO}	Measured breakdown voltage, collector to base, emitter open
BV_{CEO}	Measured breakdown voltage, collector to emitter, base open
BV_{CER}	Measured breakdown voltage, collector to emitter, with specified resistance between base and emitter
BV_{CEX}	Measured breakdown voltage, collector to emitter, with specified circuit from base to emitter
C_A	Incremental avalanche capacitance
C_{De}	Transistor diffusion capacitance
C_f	Effective collector-base feedback capacitance
C_{ib}	Transistor input capacitance (common base)
C_{in}	Effective input capacity due to C_{ib} and C_{ob}
C_K	Coupling capacitor
C_{ob}	Transistor output capacitance (common base)
C_s	Stray capacitance
C_{sc}	Collector stray capacitance
C_{se}	Emitter stray capacitance
C_T	Transistor transition capacitance
C_{Tc}	Collector transition capacitance
C_{Te}	Emitter transition capacitance
dc	Direct current
e	Naperian base 2.718
e_i	Input voltage (variable)
e_o	Output voltage (variable)
E_1	dc level which turns on a current mode transistor circuit
E_0	dc level which turns off a current mode transistor circuit
erf	Error function
f_T	Current gain — Bandwidth Product
F	Fall time factor
h_{FE} or β	dc forward current transfer ratio (common emitter)
i	Current, time dependent
I_B	Base current (dc)
i_B	Base current (instantaneous)
I_{BL}	Base leakage current
I_{BR}	Reverse base current
I_{B1}	Turn-on base current
I_{B2}	Turn-off base current
I_{Bx}	Excess base current
i_{Bx}	Time variable excess base current
I_C	Collector current (dc)
i_C	Collector current (instantaneous)
I_{CB}	Bulk Collector cutoff current
I_{CBO}	Collector cutoff current, emitter open
I_{CD}	Collector reverse current due to diffusion
I_{CEX}	Collector current with specified circuit between base and emitter
I_{CF}	Collector injected current
I_{CG}	Collector charge generation current

Symbol	*Description*
I_{CL}	Collector leakage current
I_{CS}	Collector surface leakage current
I_D	Bulk diffusion current or reverse saturation current
I_D	Forward Diode Current
I_{DZ}	Zener diode current
I_E	Emitter current (dc)
i_E	Emitter current (instantaneous)
I_{EBO}	Emitter cutoff current, collector open
I_{ED}	Emitter reverse current due to diffusion
I_{EF}	Emitter injected current
I_{EG}	Emitter charge generation current
I_{EL}	Emitter leakage current
I_{ES}	Emitter surface leakage current
I_F	Forward current (dc)
I_G	Charge generation current
I_K	Input current
I_P	Peak current
I_R	Reverse current
I_S	Surface leakage current
I_O	Load current in "0" state
I_1	Load current in "1" state
k	Boltzmann's constant $= 8.63\ 10^{-5}$ ev/°K
m	Empirical determined avalanche constant
M	Avalanche multiplication factor
ma	milliamp (ac)
mA	milliamp (dc)

Symbol	*Description*
mv	millivolt (ac)
mV	Millivolt (dc)
mW	milliwatt (dc)
n	Exponent describing depletion layer behavior
n_P	Power supply tolerance
n_R	Resistor tolerance
N_R	Ratio of tolerances $= \frac{1 + n_R}{1 - n_R} =$ Resistance Tolerance Modifier
nS	Nanosecond (10^{-9} Seconds)
pC	Pico-coulombs
P_C	Transistor collector dissipation
P_D	Power dissipation
pF	Pico-farad
q	Electron charge $=$ 1.6×10^{-19} coulomb
Q	Charge
q_a	Active base charge, time variable
Q_A	Total active charge required by the transistor for I_C to reach 90% of its final value
Q_I	Charge required to supply I_C
Q_{OB}	Off bias charge stored in both junctions
Q_R	Recombination charge
Q_S	Total stored charge
Q_T	Total control charge
Q_V	C_{ob} charge
q_x	Excess base charge, time variable
Q_x	Excess stored charge
R	Rise time factor

Symbol	*Description*
r_A	Avalanche resistance
r_B	Bulk base resistance
r'_B	Internal base spreading resistance
R_B	External base resistor
r_C	Bulk collector resistance
R_C	External collector resistor
R'_C	Equivalent to R_C and R_L in parallel
r_e	AC emitter resistance
R_E	External emitter resistor
r_E	Transistor bulk emitter resistance
R_F	Effective bulk collector to emitter saturation resistance
R_K	Coupling resistor, or drive resistor
R_L	Load resistor
r_s	Surface resistance, base-emitter junction
R_S	Source resistance
S	Store
S_B	Base store
S_C	Collector store
SV_{CE}	Collector saturation voltage ($V_{CE(sat.)}$)
t	Time
T	Absolute temperature
T_A	Ambient temperature
t_d	Delay time
T_D	Pulse duration
t_f	Fall time
T_J	Junction temperature
t_{off}	Turn-off time
t_{on}	Turn-on time

Symbol	*Description*
t_r	Rise time
T_R	Recovery time
t_s	Storage time
t_t	Total switching time
V_O	Low output voltage level (near ground)
V_1	High output voltage level
V_B	Avalanche breakdown voltage
V_{BE}	Base-emitter voltage (dc)
v_C	Capacitance voltage, time variable
V_{CB}	Voltage, collector to base (dc)
V_{CC}	Collector supply voltage
V_{CE}	Collector to emitter voltage (dc)
$V_{CE(sat)}$	Collector to emitter saturation voltage (SV_{CE})
V_D	Forward diode drop
V_{EB}	Voltage, emitter to base (dc)
V_{EE}	Emitter supply voltage
v_F	Forward voltage
V_F	Final voltage
v_i	Voltage input pulse, time dependent
V_I	Initial voltage
V_K	Clamp diode voltage supply
$V\alpha_M$	Collector-emitter breakdown voltage where $\alpha M = 1$
V_{OB}	Reverse base bias voltage
V_P	Projected offset voltage
V_{PT}	Punch-through voltage
V_R	Reverse voltage
V_{TF}	Forward voltage at threshold of conduction

V_{BB}	Base supply voltage	V_{TR}	Reverse threshold voltage
v_{BE}	Time dependent total value of base-emitter voltage	x_m	Depletion layer thickness

GREEK ALPHABET

Symbol	*Description*
α or α_N	$\frac{I_C}{I_E}$ Common base forward transfer ratio
α_I	$\frac{I_E}{I_C}$ Common base inverse transfer ratio
β	$\frac{I_C}{I_B}$ = Common emitter current gain in transition region
β_C	Ratio of I_C to I_{B2} (cutoff gain)
β_F	$\frac{I_C}{I_B}$ = Current gain in saturation region
β_{FS}	A specified value of β_F which will provide a sufficient penetration of the saturation region to allow V_{CE} to be linearly related to I_C.
β_o	$\frac{I_C}{I_B}$ = Current gain at edge of saturation
γ	Ratio of current through collector resistor for a zero to that for a one.
Δ	(Delta) indicates a small change in the variable with which it is associated
θ_J	Thermal resistance
θ_{VB}	Temperature coefficient of base-emitter saturation voltage
θ_{VC}	Temperature coefficient of collector-emitter saturation voltage
ρ	Material resistivity
τ	Lifetime or time constant
τ_a	Active base charge lifetime

τ_A	Active region time constant, describing time response to 90% point
τ_{BS}	Effective time constant of excess carriers
τ_{CR}	Effective collector recovery time constant
τ_L	Load time constant
τ_x	Excess base charge lifetime
ϕ_C	Theoretical collector-base junction voltage
ϕ_{CE}	Theoretical voltage across collector and emitter terminals
ϕ_E	Theoretical emitter-base junction voltage
ϕ_P	Collector-emitter offset voltage
ω_α	Alpha cutoff frequency (Rad/Sec.)
ω_I	Inverse alpha cutoff frequency (Rad/Sec.)
ω_T	Current gain-bandwidth product ($\omega_T = 2\pi f_T$)
∞	Infinity

Special Symbols

$\frac{kT}{q}$	$= 26$ mV @ T $= 27°C$ (increases 86 $\mu V/°C$)
—	Above a symbol indicates a maximum value
—	Below a symbol indicates a minimum value
~	Above a symbol indicates a typical value

MOTOROLA INC., Semiconductor Products Division

5005 EAST McDOWELL ROAD · PHOENIX 8, ARIZONA

CHAPTER 1

Effects of Transistor Construction on Electrical Characteristics

Since the introduction of the earliest transistor in 1948 there have been numerous breakthroughs in the basic knowledge of the physics, chemistry, and mechanics of transistor design. The earliest transistors, barely able to amplify and acting as very crude switches, have evolved into a line of devices whose inherent capabilities, in some instances, have outstripped even the circuits technology of the equipment in which they are employed. And in the process, the transistor has undergone such drastic changes that today's units bear little internal resemblance to their early predecessors.

From a circuit designer's viewpoint, the transistor can usually be treated as a proverbial "black box" whose input and output characteristics, as specified on data sheets, permit them to be designed into specific circuits. The electrical characteristics alone, however, do not tell the whole story. An understanding of the physical properties of such devices can be extremely valuable in the selection of transistors best suited for specific applications. A general acquaintance with the various kinds of transistors and their fabrication techniques could reveal inherent differences in ruggedness and fundamental differences in electrical characteristics which would make a particular transistor type superior to all others for specific applications.

1-1 — The Ideal Switch

A theoretically perfect switch is a device which has no power loss when used for the purpose of interrupting current through a load and one which can change its state, say from the on to the off condition, in zero time.

This rather simple definition has some far-reaching implications. It requires, for example, that in the closed position the switch must have no voltage drop across its terminals. In the open position, the switch must cut off all current flow through the load. And, finally, the repetitive cycling speed must be infinite, corresponding to zero time delay between one switch setting and the other.

Although a switch meeting the above criteria would be ideal as far as the load is concerned, actuation of the switch will require the expenditure of energy. The human is naturally more pleased when the energy expended is small, but some small finite amount should be required. Otherwise every stray signal would cause the switch to react producing a chaotic situation. Thus, a high but finite power gain is desirable

$$\left(\text{power gain} = \frac{\text{power switched}}{\text{actuating power}}\right).$$

The perfect switch, of course, remains still to be invented. Mechanical switches and solenoid relays, while closely approaching the first two criteria, cannot hope to meet the switching-speed requirements of today's applications. Mercury relays, while capable of extremely fast switching from one condition to another, have a relatively slow repetition rate. Vacuum tubes, while they approach the ideal switch in the off position and can operate at relatively high repetition rates, are severely limited by a high voltage drop in the on condition. This loss, coupled with the large power loss in the heater, limits the usefulness of these devices in modern applications. Transistors, though they fall far short of perfection, represent the best available components for switching applications, especially in high-speed equipment.

1-2 — Basic Transistor Limitations

The switching capabilities of a transistor stem from the capability of the device to rapidly change the dc resistance between its output terminals, the collector and emitter, from a very high to a very low value, in response to a small current injected into the control (base) terminal.

A simplified schematic of a basic transistor switching circuit is shown in Figure 1-1. A voltage pulse (v_i) applied to the input terminal causes a current flow (i_B) through the base resistance (R_K) which in turn reduces the collector-emitter resistance to a very low value and permits a large current flow (i_C) from the battery through the load. If the input voltage is removed, the reverse bias voltage (V_{OB}) causes the collector-emitter resistance to increase greatly, and significantly reduces the load current.

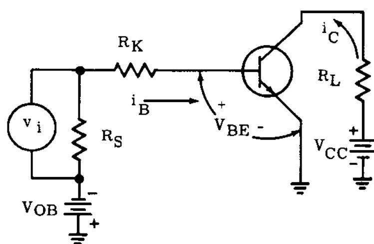


Figure 1-1 — Basic Transistor Switching Circuit

Transistor limitations, with respect to a perfect switch, can easily be determined from the graph in Figure 1-2a, which shows the collector current and voltage relationships for various values of base current, and from the waveform shown in Figure 1-2b. When a load line is drawn on the graph, it is obvious that under conditions where the base current is zero, a small amount of collector current still flows in the circuit. This residual current, called leakage current, prevents full cut-off of load current and is one characteristic which keeps the transistor from simulating the off condition of a perfect switch.

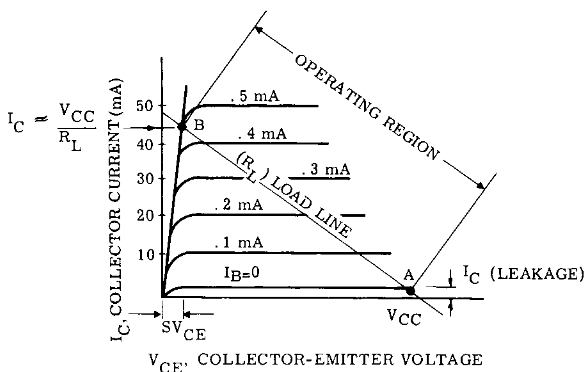


Figure 1-2a — Transistor dc Output Characteristic

Although leakage current can never be entirely eliminated, application of reverse bias can reduce its value to approximately that of I_{CBO} (see Chapter 3) which is generally negligibly small in comparison with the load current.

To simulate a perfect switch in the on position, the voltage drop across the transistor collector-emitter junction should be zero when a base driving signal is applied. As shown in Figure 1-2a, this condition can never be achieved entirely although it can be approached within a few millivolts. The residual on voltage drop across the collector-emitter terminals is called collector saturation voltage, ($V_{CE(sat)}$ or SV_{CE}). (See Chapter 4).

There are two situations in which an electronic device may have a zero power input requirement yet still require a finite control signal. If the input impedance is infinite, a voltage is required for control. If the input impedance is zero, a current is required for control. In each case, zero input power is required. The vacuum tube having a very high input impedance and the transistor having a very low input impedance approach the requirement for zero input power.

Because of its low input impedance, the transistor is actuated by a current. Its current gain is a significant characteristic and is defined as

$$\beta = I_C / I_B. \qquad (1\text{-}1)$$

From the family of curves in Figure 1-2a, note that for any given base current curve the collector current is larger by a factor of 100, $\therefore \beta = 100$.

Up to this point, only the steady state or dc characteristics that prevent a transistor from duplicating a perfect switch have been discussed. Of equal importance, are the dynamic or transient characteristics which affect switching speed. The transient limitations are illustrated in Figure 1-2b.

As illustrated by the waveforms, the collector current (the output waveform) does not respond immediately to changes in the input signal. At time (t_0), the input signal rises instantaneously to its maximum value. At this instant, the transistor is in the off condition because of the reverse bias (V_{OB}). The collector

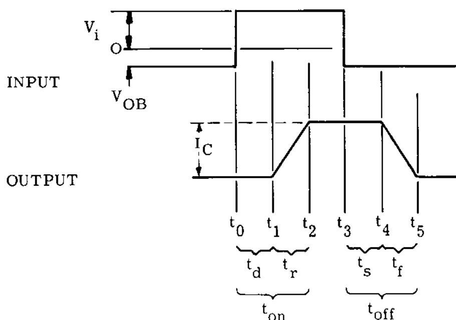


Figure 1-2b — Transistor Output Waveform

current does not begin to increase until time (t_1). The interval between t_0 and t_1 is called the *delay time* (t_d) and is defined as the time required to bring the transistor from the initial off condition to the edge of conduction. At time t_1, the operating point of the transistor is at the beginning of the active region and the collector current starts to increase toward its saturation value. However, it does not reach its maximum value until time t_2. The interval between t_1 and t_2 is defined as the *rise time* (t_r) of the collector current. The sum ($t_d + t_r$) is called turn on time, t_{on}.

The transistor will remain in the on state as long as the input signal is maintained. At time t_3 the input signal drops instantly, but it is observed that the collector current does not respond until time t_4. The time interval between t_3 and t_4 is referred to as the *storage time* (t_s). Finally, at time t_4, the transistor comes out of saturation and the collector current falls to its off value at time t_5. The interval of time between t_4 and t_5, is defined as the *fall time* (t_f). The sum ($t_s + t_f$) is called turn-off time, t_{off}.

The delays in the response of the collector current to changes in the input signal are attributed to various inherent transistor capacitances. These capacitances are discussed thoroughly later in this chapter.

In summary, the factors that prevent a transistor from duplicating an ideal switch are:

1. A residual leakage current when the transistor is off.
2. A residual collector-emitter saturation voltage when the transistor is on.
3. Time delays involved in the response of collector current to changes in the input signal.

1-3 — Basic Transistor Physics and Characteristics

In order to evaluate the effects of transistor physical characteristics upon the electrical characteristics of the devices, a basic understanding of transistor physics is required.

An ideal PN junction diode can be defined as a diode which has no reactive components and follows the voltage-current relationship, predicted by the simple first order theory as developed by Shockley[1] This relationship is expressed by the following equations:

$$I_F = I_R \left(e^{\frac{qV}{kT}} - 1\right) \quad (1\text{-}2a)$$

or

$$V = \frac{kT}{q} \ln \left(1 + \frac{I_F}{I_R}\right) \quad (1\text{-}2b)$$

where I_F = forward junction current

I_R = reverse junction current

$\frac{kT}{q}$ = a common semiconductor constant which equals 26 mV at 27 °C

V = voltage across the junction

A plot of these ideal diode characteristics is shown in Figure 1-3. The graph shows that any reverse voltage (V_R) (in excess of a few tenths of a volt) produces a small reverse current which remains constant. When a forward voltage (V_F) is applied, the forward current (I_F) increases exponentially.

In order to create a diode having 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 high resistivity corresponding to a low doping level. If the P-region is more heavily doped with impurities than the N-region, it will have the greater number of current carriers, and becomes the emitter. In this instance, when a forward voltage is applied to the junction, the forward current consists mainly of holes that are injected from the P-region (where they are plentiful) into the N-region. In addition, there exists a small current flow of electrons from the lightly doped N-region into the P-layer.

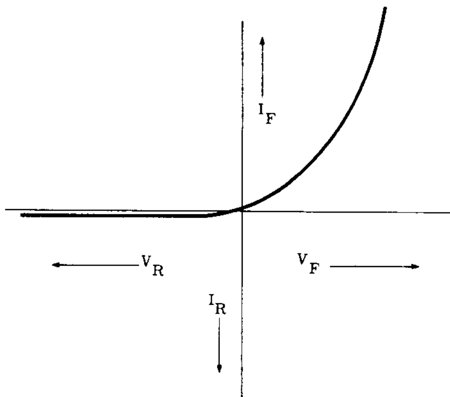


Figure 1-3 — Ideal Diode Characteristics

When the junction is reverse biased, a depletion layer is produced principally in the *high resistivity* side of the junction. That is, in order to preserve charge neutrality under the reverse-bias condition, the excess charges (holes in the P-region, electrons in the N-region) move away from the junction, leaving exposed ions in a region near the junction. Thus, this region is depleted of mobile carriers.

The reverse current (I_R) consists mainly of those minority carriers (holes) in the N-region which are close enough to the junction to be swept across by the electric field.

A transistor can be considered as two diodes of the same type connected back-to-back, with a narrow common region (the base) separating the two "emitters." When both junctions are reverse biased, the transistor behaves much like two reverse biased diodes and the sum of the reverse currents flows out of the base terminal. When both junctions are forward biased, the sum of the forward currents flows out of the base. When one junction is reverse biased and the other forward biased, normal transistor action occurs and the device is said to be operating in the active region.

In a PNP transistor, where the principal carriers are holes, a forward biased emitter injects holes into the base where some will recombine with the electrons. However, most of the injected holes reach the collector depletion layer where they are swept to the collector by the negative collector potential. The ratio of the emitter current (I_E) to collector current (I_C) is called α_N, the common base current transfer ratio in the normal connection.

To reduce recombination of holes with electrons in the base, it should be evident that the base must be narrow. Also, since the base emits electrons to the emitter, it is necessary to have the resistivity of the base high (low doping) so that the source of electrons is small. These two electron currents, i.e. electron current due to recombination and due to base injection, flow through the base region and the ratio of collector current (I_C) to base current (I_B) is the common emitter current gain, (β). By definition, the ideal transistor then fulfills these criteria: (1) terminal voltages and currents are given by equation 1-2 for both junctions forward biased or both reverse biased, (2) when one junction is forward biased and the other reversed biased, equation 1-2 applies to only the forward biased junction. The current out of the reversed biased junction is governed by the current transfer ratio.

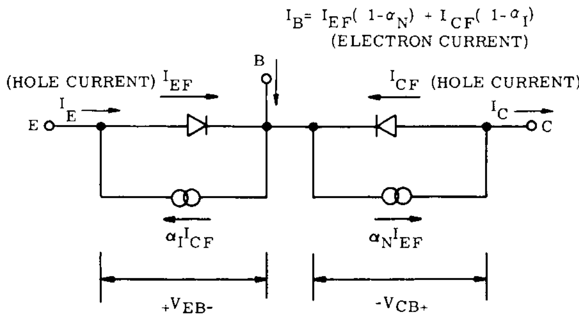


Figure 1-4 — Two-Diode Ideal Transistor Model

From the above discussion, it is evident that a transistor can be represented with the model shown in Figure 1-4 where the current generators represent the fraction of injected current that passes through the base. By applying the reasoning previously given, one could deduce that the equations governing the currents are:*

$$I_C = \alpha_N I_{EF} - I_{CF} \tag{1-3}$$

$$I_E = I_{EF} - \alpha_I I_{CF} \tag{1-4}$$

Where I_E = the emitter terminal current

I_{EF} = the current injected from the emitter

α_I = Inverted or reverse α (fraction of I_{CF} reaching emitter)

I_{CF} = the current injected from the collector

I_C = the collector terminal current

α_N = normal or forward α (fraction of I_{EF} reaching collector).

The terminal voltages may found by:

$$V_{EB} = \frac{kT}{q} \ln \left(1 + \frac{I_{EF}}{I_R}\right) \tag{1-5}$$

$$V_{CB} = \frac{kT}{q} \ln \left(1 + \frac{I_{CF}}{I_R}\right). \tag{1-6}$$

In the normal connection, the base-emitter junction is forward biased and equation 1-5 describes V_{EB}; when the collector is reverse biased, equation 1-3 describes the collector current. For the inverted connection (the function of the emitter and the collector are interchanged) the collector-base is forward biased and equation 1-6 describes V_{CB}. When the emitter is reverse biased, equation 1-4 describes the emitter current. If both junctions become forward biased, the transistor is in saturation and all four equations apply.

As an example of saturation behavior, assume that a transistor is made from two identical diodes, then $\alpha_N = \alpha_I$. Assume that $\alpha_I = \alpha_N = .95$, $I_R = 1\mu A$, $I_C = 10$ mA, $I_E = 11$ mA. Using the developed relations it is found that $I_{EF} = 15.4$ mA and $I_{CF} = 4.6$ mA, $V_{EB} = .25$ volt and $V_{CB} = .22$ volt. The voltage from collector to emitter (SV_{CE}) is $V_{EB} - V_{CB} = .03$ volt. The base current is $I_E - I_C = 1$ mA. Thus, it is seen that when a transistor is used in the common-emitter configuration, the output or saturation voltage is less than the input voltage V_{EB}. This saturation voltage is so small that it is convenient, and not too inaccurate, to think of a saturated transistor as a short circuit. The voltage V_{EB} is also low, varying from 0.2 to 0.5 volt for germanium and 0.5 to 0.8 volt for silicon for the usual range of collector currents. (I_R is several orders of magnitude lower for silicon devices which accounts for the higher values of V_{EB}).

*Exact relations as given by Ebers & Moll will be discussed fully in later chapters. For purposes of this chapter this intuitive approach will suffice.

Returning again to the diode, as reverse voltage is raised, current increases without limit as a voltage called the *avalanche breakdown voltage* (V_B) is approached. The higher the resistivity of the high-resistivity side, the higher the avalanche breakdown voltage. Avalanche breakdown is attributed to the fact that the high field across a reverse biased junction accelerates any moving particle, which, if moving fast enough, may have sufficient energy to free additional particles by collision with atoms. This results in a multiplication of carriers which proceeds at an extremely high rate at the breakdown voltage.

Another effect of reverse bias results from oppositely charged particles being close to each other at the junction. This results in a capacitive effect (transition capacitance) similar to that of a parallel-plate capacitor. Higher resistivity materials result in a wider depletion layer at a given voltage and thus have a lower capacitance per unit area.

Therefore, the undesirable effects of avalanche breakdown and transition capacitance are reduced as resistivity is increased. However, these characteristics are improved at the expense of the forward characteristics.

The passage of current through a diode will produce additional voltage drops across the bulk resistivity. This resistance is given by the familiar:

$$R = \frac{\rho l}{A} \tag{1-7}$$

where ρ = resistivity of material

l = length

A = area

It is apparent that the voltage drop across a high resistivity region can add a voltage which is considerably larger than that predicted by the ideal diode formula. (Equation 1-2)

Another important mechanism is charge storage. When a forward current is flowing, a carrier gradient is produced in the high resistivity side of a junction resulting in a storage of charge. If the source of forward bias is suddenly changed to a reverse bias, this stored charge maintains the current flow until the charge is depleted. Thus, the phenomenon of storage or recovery time is another departure from an ideal diode. The storage time becomes less of a problem as forward current is reduced, because the gradient producing these excess carriers is less resulting in less stored charge. Also, an increase of reverse bias will hasten the depletion of this stored charge. Making diodes from materials having a low lifetime (time that an isolated charge can exist before recombination) results in less stored charge.

These various effects (resistance, capacitance, voltage breakdown and storage) can be added to the ideal diode as external elements to get the equivalent circuit shown in Figure 1-5.

The diode, which follows equation 1-2, is given the avalanche or zener symbol as a reminder that it has a voltage limit V_B. The capacitance represents the transition capacitance (C_T). Remember that C_T decreases and V_B increases with increases in resistivity. The series resistance (r), however, is proportional

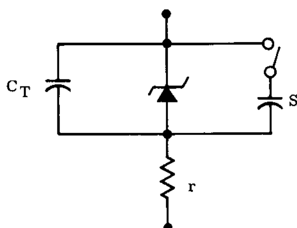


Figure 1-5 — Equivalent Circuit of a Diode

to resistivity. The store (S) represents an infinite capacitor, in a sense, because during storage time current flows but the diode voltage remains essentially constant. The switch indicates that the store is present only after current flows and is disconnected at the end of the storage time interval. It is as if instructions were given to the carriers — first one in — turn on the switch — last one out — turn it off! In general, increases in resistivity cause increases in the effective storage capacitance (S) since carrier lifetime increases with resistivity. This is because low resistivity material has more impurities and hence, more recombination centers. A wide high-resistivity region would also increase S since the volume available for storage is larger.

Returning to the two-diode transistor analogy, (Figure 1-4) external elements of a single diode also apply to the transistor model, but additional elements are added due to transistor action, as shown in Figure 1-6. This model will be used to illustrate the effects of all these non-ideal elements upon the on, off, and transient states of a transistor switch. The transistor symbol now represents two ideal diodes connected back-to-back by means of a common narrow base region. The resulting device exhibits constant current gain (β). The undesirable elements are shown external to the ideal transistor.

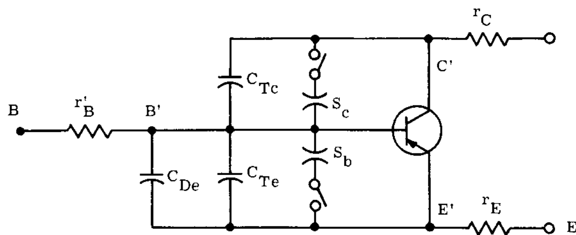


Figure 1-6 — Equivalent Circuit of a Transistor

One of these additional elements is the diffusion capacitance C_{De}. This capacitance accounts for the time delay experienced by carriers as they travel through the base enroute from emitter to collector. The value of C_{De} is reduced as the base width is decreased.

Another element peculiar to the transistor is the "base spreading-resistance" r'_B.* This additional resistance becomes important because it produces a voltage drop as the base current flows out from the transistor through the narrow base region. This transverse voltage drop is responsible for a number of transistor problems.

*The term "base-spreading" resistance was "coined" to describe the base resistance of alloy transistors. The base of alloy transistors "spreads out" from the center of the junction to the periphery of the transistor.

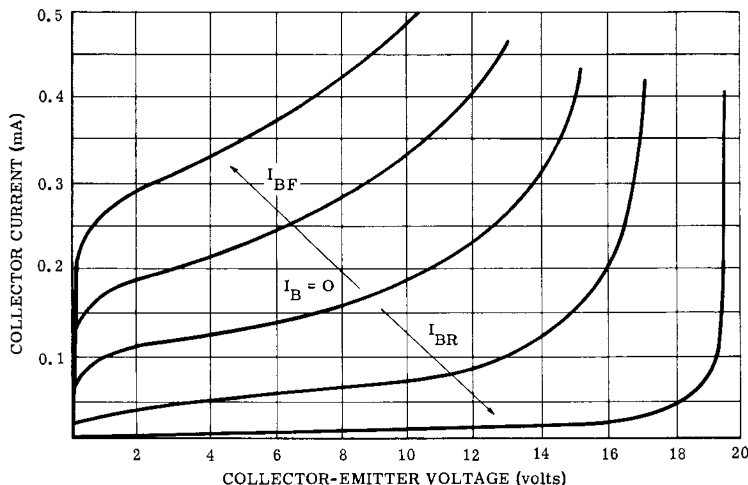


Figure 1-7 — Transistor Output Characteristics Near Cutoff

Figure 1-7 shows the output characteristics of a transistor in the low current region near the off point which is referred to as the cutoff region. Notice that collector current can never be entirely eliminated, but that application of a reverse bias can reduce its value to approximately that of the collector-base junction leakage current* which is generally negligible in comparison with the load current.

Note also that the application of reverse base current reduces the conductance as well as the collector leakage current. Thus, in the off condition, the resistance between the collector and emitter terminals is not infinite but is very large. The maximum amount of collector voltage, that can be applied to a transistor in the off state, is limited by the avalanche breakdown voltage of the collector-base junction as indicated by a rapid increase in collector current as this voltage is approached.

In the on condition the various voltage drops, caused by current flow through the internal resistances of the transistor, prevent the device from duplicating a perfect switch. As shown in Figure 1-8, which illustrates the dc current flow of an alloy transistor, the emitter current flows through the emitter, base, and collector regions and any resistance in these regions will produce voltage drops.

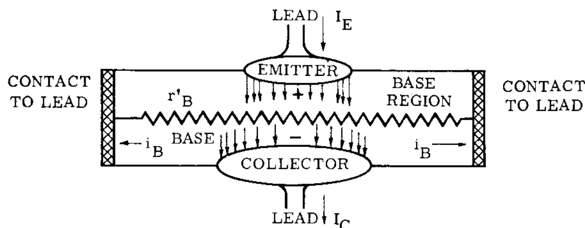


Figure 1-8 — dc Current Flow in an Alloy Transistor

*I_{CBO} — see Chapter 3 for particulars.

The base current is flowing transverse to the emitter current and produces a voltage drop on its way to the base contact. This drop, accounted for by r'_B, not only adds to the base terminal voltage, but the polarity of this voltage is such that it adds to the forward bias at the edge of the emitter, which tends to concentrate the emitter current near the emitter edge. As a result the emitter perimeter becomes more important than the area as a factor in determining current capability, and gain, at high currents. Mathematically, it is difficult to express r'_B, but it should be apparent from Figure 1-8 that a wide, low-resistivity base region is needed to make r'_B low. However, the effect of r'_B upon switching speeds is significant only when a transistor is used near its switching speed limit. It is not usually detrimental in the majority of switching circuits, and low values of r'_B are usually sacrificed to improve other characteristics such as C_{De} and emitter breakdown voltage.

There is a voltage drop in the base caused by the current flow which drops from I_E at the emitter junction to αI_E at the collector junction as a result of recombination and base injection. It is convenient to lump the voltage drop from collector to emitter into just two components $I_E r_E$ and $I_C r_C$. However, since I_C and I_E are approximately equal, the bulk voltage drop from collector to emitter can be simply expressed as $I_C R_F$, where R_F is the sum of the resistances from collector to emitter and is defined as the effective collector to emitter saturation resistance.

The expanded graph in Figure 1-9 shows the transistor characteristics in the saturation region representing the on condition of the switch. At any *on* point, a small amount of voltage is still present across the collector-emitter terminals. This is a result of voltage drops across the effective bulk resistance (R_F) plus a small collector-emitter voltage which is always present while a transistor is in operation. Notice that all the characteristic curves do not converge at zero but rather at some small offset voltage (ϕ_P). Also, note that under conditions of heavy base drive, ($I_B > > I_C/\beta$) a nearly linear relationship between changes in I_C and V_{CE} is evident. This slope is $1/R_F$. A line has been drawn on Figure 1-9 connecting the points where each base current line intersects a collector current line that is a given multiple of the base current. The resulting line also has a slope equal to $1/R_F$ but intersects the abscissa at an effective offset voltage (V_P), which is much larger than the offset voltage (ϕ_P) when I_C is zero.

The voltage (V_P) is the voltage difference between the two forward biased diodes of the ideal transistor as discussed in conjunction with the model of Figure 1-4. That is, V_P is the theoretical collector-emitter saturation voltage of the ideal transistor which is a function of the ratio of I_C to I_B. This voltage* is small and does not differ greatly for different transistor types under identical drive conditions. However, R_F, which is due to the bulk resistances, differs greatly from type to type and will be considered as the various processes are discussed. Thus, the transistor does not exhibit a total short in the on condition but has a small finite resistance and an offset voltage.

*The Ebers-Moll on voltage

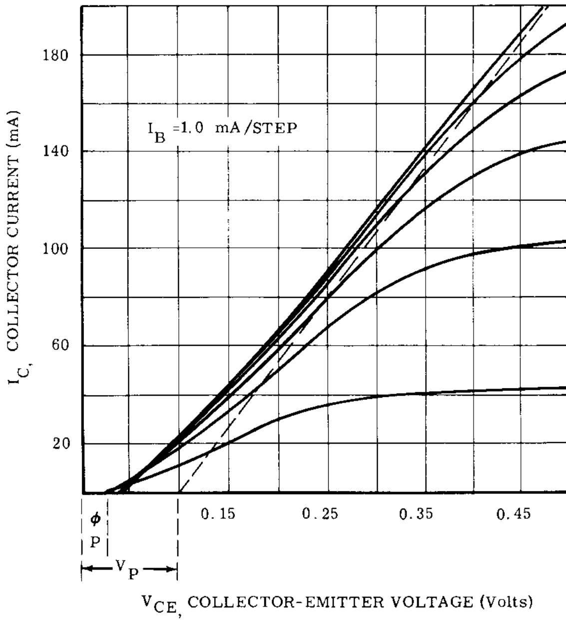


Figure 1-9 — Transistor Output Characteristics Near Saturation

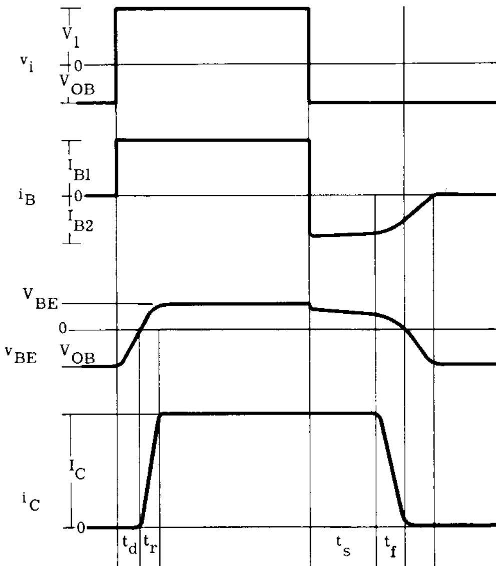


Figure 1-10 — Switching Wave Forms

The transient response, illustrated in Figure 1-10, can now be examined in greater detail. When the transistor is in the off state, it appears to the input circuit only as a capacitance which consists of C_{Tc} and C_{Te} in parallel (see Figure 1-6). These capacitors are charged by the reverse bias (V_{OB}) and, before the transistor can be turned on, this charge must be removed. The time required for the input signal to change the base voltage from the reverse bias condition to a voltage at the threshold of conduction constitutes delay time (t_d).

Delay time (t_d) increases with the magnitude of V_{OB} and the values of C_{Tc} or C_{Te}. It is inversely proportional to the magnitude of the turn-on current (I_{B1}).

As emitter current begins to flow, two additional effects occur. The diffusion capacitance (C_{De}), which is proportional to emitter current, increases and the collector transition capacitance (C_{Tc}) feeds back a current from collector to the base. This feedback current is proportional to the voltage gain from B′ to C′ and may be quite large.

All three capacitances (C_{De}, C_{Te} and C_{Tc}) influence rise time (t_r). The rise time is the time required for the collector current to reach its limiting value, approximately V_{CC}/R_L.

To further explain storage time, recall that in the diode, the stores accumulated a charge whenever the diode was forward biased. In the case of the transistor, however, the stores do not accumulate a charge unless the transistor is driven into saturation because the reverse bias across the collector junction provides a field that sweeps out the carriers if the forward bias on the base-emitter junction is removed. When the transistor is driven into saturation, the stores accumulate a charge because no reverse bias is present. The deeper the transistor is driven into saturation, the more charges are accumulated in the stores. The turn-off delay (storage time) is the time required to deplete the stores. Depletion is hastened by increasing the reverse drive voltage, thus allowing a larger reverse-current flow. Note, in Figure 1-10, that the transistor currents and voltages are nearly constant during the storage time interval.

Depletion of the stores initiates the fall time portion of the switching cycle. Fall time can be reduced by the injection of a reverse current into the transistor base, which speeds up the process of discharging the capacitances. During the fall-time interval the base voltage and current remain nearly constant until the collector current is completely shut off. The transition capacitances, then, are charged to V_{OB}.

From the discussion so far, it becomes clear that in any transistor design, a number of compromises are necessary. A change in a particular physical property can improve one electrical characteristic but may have a detrimental effect on another. A brief summary of some of these interrelations is given in Table 1-1, and these will be discussed in greater detail as each transistor fabrication process is considered.

TABLE 1-1

Change	Increase	Decrease	Switching Speed Effect
Decrease Emitter Resistivity	β	r_E (decrease SV_{CE})	——
Increase Base Resistivity	β V_B (emitter and collector junction) r'_B S_B r_E and r_C	V_{PT} C_{Te} C_{Tc}	Increase storage time, small increase in other intervals due to r'_B
Increase Collector Resistivity	V_B (collector) S_C r_C (increase SV_{CE})	C_{Tc}	Increase storage time, small decrease in other intervals
Decrease Base Width	β r'_B	C_{De} S_b V_{PT} r_E and r_C	Decrease rise time Decrease fall time Decrease storage time

1-4 — Comparison of Transistor Types

The physical geometry of a device and the type and quantity of impurities in the semiconductor crystal are the basic considerations that influence the electrical characteristics of any transistor. Therefore, to compare transistors, it is convenient to discuss the transistor in terms of its geometry and the resistivity of its emitter, base, and collector regions.

Many characteristics of the device are directly related to its geometry. The base width, for example, affects gain (β), the diffusion capacitance (C_{De}) and base-spreading resistance (r'_B). The area of the junction affects the transition capacitances (C_{Te} and C_{Tc}) and emitter and collector resistance (r_E and r_C).

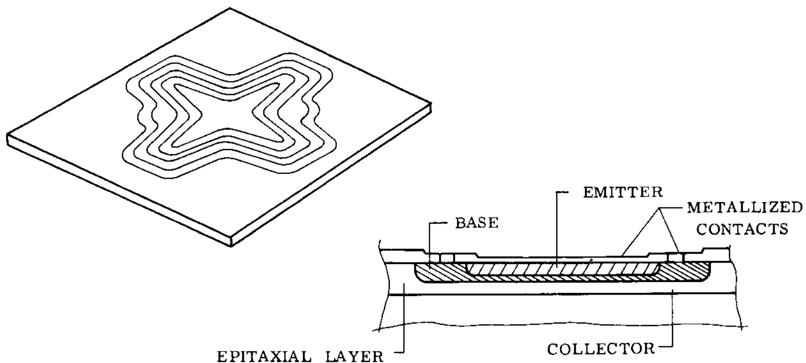


Figure 1-11 — Star Transistor Geometry

An example of how geometry can be utilized to attain improved characteristics is the "Star" transistor illustrated in Figure 1-11. Since high current capability requires a large emitter perimeter which in turn means higher capacitances and slower rise time, a compromise is required when determining the emitter area in devices where the emitter shape is fixed. However, in the mesa, planar, and annular transistors, the emitter can be shaped to provide necessary periphery for high gain at high currents without too great a sacrifice in rise time. The star geometry has a much smaller ratio of area to perimeter than that of a circle, for example.

In addition to the geometry effects, the electrical characteristics are also closely related to the resistivity or conductivity of the emitter, base and collector regions. This relationship can be conveniently illustrated using a resistivity profile diagram, which is a sketch of the resistivity of the transistor from the emitter, through the base region, and into the collector. The resistivity profile for an alloy transistor is shown in Figure 1-12.

The resistivity of the emitter and collector is indicated above the reference line. The farther away from the reference line the lower the resistivity. The base region is of oppositely doped material, therefore, it is shown below the reference line. Likewise, the farther below the reference line, the lower the base resistivity.

Figure 1-12 also shows that the base region of an alloy transistor has the same resistivity at both the emitter and collector junctions. Thus, the impurity distribution is uniform throughout the base region. A transistor, with this type of base region, is commonly referred to as a uniform base, or step junction transistor.

However, transistors can be made with a non-uniform distribution of impurities in the base, so that the base resistivity can be considerably different at the two junctions. The distribution of impurities in the base can be graded, and such transistors are known as graded base devices. Since they are made by a diffusion process, they are commonly referred to as diffused base transistors.

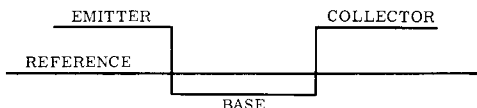


Figure 1-12 — Resistivity Profile for an Alloy Transistor

Resistivity levels and the impurity distribution affect breakdown voltage, gain, response times, capacitance per unit area, and resistance per unit area. Figure 1-13 shows the effect of pivoting the base resistivity upon the electrical characteristics of a transistor. The changes indicated are in relationship to the alloy transistor.

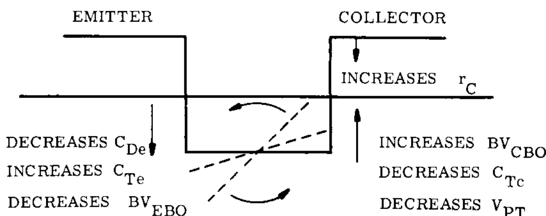


Figure 1-13 — Effect of Pivoting Base Resistivity Upon Transistor Characteristics

The resistivity profiles of several modern transistors are shown in Figure 1-14. This chart of resistivity profiles will be used in the following comparison of transistor types. In order to compare the effects of resistivity profiles on transistor characteristics, the following three principles of transistor physics should be kept in mind:

1. The amount of carriers injected from the base to the emitter is proportional to the ratio of resistivity of the emitter to that of the base. To reduce this undesired injection, which lowers current gain (β), emitters are always heavily doped and the contribution of the emitter resistivity to r_E is low.
2. Avalanche voltage breakdown increases and the transition capacitance per unit area decreases as resistivity increases and these characteristics are affected by the "slope" of the profile at the junction.
3. The depletion layer (x_m) extends principally into the high resistivity side of the junction. At a given voltage, the depletion layer is wider for high resistivity material.

TRANSISTOR PROCESS	RESISTIVITY PROFILE	BASE IMPURITY DISTRIBUTION	CROSS SECTION
A ALLOY	E C B	UNIFORM	ALLOY MAT E C E C B B
B DIFFUSED BASE-ALLOY	E C B	GRADED	DRIFT MADT E C E C B B
C DIFFUSED-BASE MESA TYPES	E C B	GRADED	B E C (MESA CONSTRUCTION)
D DIFFUSED-BASE EPITAXIAL	E C B	GRADED	SUBSTRATE (LOW RESISTIVITY) EPITAXIAL LAYER (HIGH RESISTIVITY) E B C (PLANAR CONSTRUCTION)

Figure 1-14 — Resistivity Profile Chart

ALLOY TRANSISTORS: The fabrication of alloy transistors, like that of all other transistors, begins with the drawing of a single crystal from a semiconductor melt. The entire crystal is lightly doped during the growing process resulting in a high-resistivity N or P-type material. The crystal is then sectioned or "diced" into squares with dimensions about ⅛-inch square and approximately 0.01 inch thick. Each of these squares or dice represents the base region of a transistor.

To form PNP germanium transistors, N-type germanium dice are heated and small round dots of indium, containing aluminum or gallium, are melted against both sides of the base substrates. These alloy into the germanium to form a saturated liquid solution on both sides of the base. Then, as the assembly cools, the germanium refreezes onto the surface of the base crystal. The recrystallized layers, however, are heavily doped (low resistivity) with P-type material, thus forming a PN junction on each side of the base.

Some advantages of the alloy process are high gain, low saturation resistance, and high avalanche voltage breakdown. Gain, and the breakdown voltage of both junctions is high because of the lightly doped base region; saturation resistance is usually low, because the heavily doped collector region has low resistance per unit area.

A major disadvantage of the alloy transistor is its susceptibility to "punch-through". Punch-through results from the spreading of the depletion layer into the high resistivity base region as the reverse voltage across the collector-base junction is increased. In effect, this causes base width to decrease as voltage increases. At the punch-through voltage (V_{PT}) the effective base width is reduced to zero as the collector depletion layer reaches the emitter region, and the base is effectively short circuited. Once base resistivity has been optimized the voltage at which punch-through occurs can be increased, but only by increasing the physical base width with its attendant detrimental increase in C_{De} (diffusion capacitance) and S_B (base store).

A second disadvantage of the simple alloy process is its lack of precise control. Because alloying depth is somewhat irregular, transistors of the same type often vary considerably in the values of C_{De}, S_B and V_{PT}. Furthermore, the irregular alloying depth prevents the design of transistors with the extremely narrow base widths needed for high speed switching. In general, the switching rate of alloy units is limited to about 500 kc.

Significant improvements in base-width control of simple alloy devices can be achieved by electrochemically etching a deep pit for the collector and a shallow pit for the emitter into the base region prior to alloying. With this process, the base thickness can be controlled to within one ten-thousandth of an inch through the use of light transmission during the etching operation. Very thin layers of indium, or similar materials, are plated to the bottom of these pits and alloyed into the base with a very shallow penetration.

Transistors made with this process (referred to as the micro-alloy transistor), have an extremely thin base layer with a corresponding improvement of switching time. Although there is a slight increase in base resistance, which is not particularly harmful, collector and emitter resistances are relatively unchanged. The major difficulty with such units is caused by the very thin base region which reduces the voltage rating. However, by limiting collector voltage to 6 volts, switching rates to about 2 mcs can be achieved. These devices are rather fragile because the collector and emitter contacts are supported only on a very thin membrane of semiconductor material.

DIFFUSED-BASE ALLOY TRANSISTORS: A further improvement of the alloy transistor resulted from a graded distribution of impurities in the transistor base region. This process differs from the alloy transistor process only in that the original crystal pulled from the melt consists of a very high purity material. The impurity atoms are introduced subsequently by a diffusion process which permits a graded impurity structure. A graded base results in a resistivity which is low at the junction of the emitter and base regions. The resistivity gradually increases as the collector terminal is approached, as shown in Figure 1-14. In transistors of this type, the depletion layer penetration into the base region increases rapidly with applied voltage due to the very high resistivity near the collector junction. However, as the depletion layer advances, its rate of penetration with increasing voltage is reduced by the lower resistivity material near the emitter. This effect yields devices with much higher voltage ratings than those of other types of alloy transistors. The "built-in" electric field associated with the gradation also reduces the transit time of carriers through the base region, resulting in lower values of C_{De} than available with uniform base regions.

The penalties paid for these improvements are a reduction of β and emitter avalanche breakdown voltage, and an increase of C_{Te} due to the relatively low overall resistivity of the area on the emitter side of the base region. These effects, however, are normally not significantly detrimental to switching applications.

Application of electro-chemical etching and plating techniques to the diffused-base alloy process have yielded excellent switching transistors i.e. the micro-alloy diffused-base transistor. Such devices, however, are still relatively fragile, and for extremely high speed designs, a compromise between voltage rating and rise time is required.

A common disadvantage of all alloy types of transistors, is the lack of control over the shape of the emitter area. With these processes, the emitter is circular, resulting in a decreasing ratio of perimeter to area, as the radius is increased. Since high current transistors require a large emitter perimeter, they necessarily have a large area which yields large capacitance values and reduces switching speeds. Accordingly, switching speed is reduced as current-handling capability is increased.

MESA TRANSISTORS: A significant breakthrough in transistor design came in 1956, with the introduction of the MESA transistor. Although it took several years to develop the mesa to its present capability, these units are characterized by higher switching speeds, greater power handling capability, and far more rugged construction than their alloyed counterparts.

In the basic mesa transistor fabrication process, the original crystal is drawn from a lightly doped P-type or N-type melt. The crystal is then sliced into very thin wafers, each of which serves as the collector region for hundreds of individual transistors.

After each wafer is carefully polished and cleaned, an opposite-polarity impurity is diffused into the collector crystal to form the base region. The entire wafer, then, consists of a PN junction composed of a relatively thick collector region — approximately .005″ — and a very thin diffused base layer, about .0001″. The emitter junction is formed on the base layer by the vacuum evaporation of an impurity, of the type used for the collector, through a metal mask con-

taining a series of openings such as rectangular slots. The masks are prepared by a photo-etch process which permits precise dimensioning and spacing of the slots, so that as many as 400 separate emitter stripes can be deposited simultaneously on a single wafer to form as many separate transistors. The distinctive physical fact about the mesa transistor is that the junction areas can be precisely controlled.

The emitter stripe evaporation is commonly done by the cross evaporation method so that an adjacent base-contact stripe may be formed during the same operation. By cross-evaporating the emitter and later the base contact stripes through the same slot, the extremely close spacing required for good high-frequency response can be achieved. By subjecting the resulting wafer to a short high-temperature cycle, all stripes alloy, to form junctions that have very shallow penetration into the diffused base layer.

The last step in the wafer process involves the selective etching of the mesa, which reduces the active area of the collector junction to a small rectangle which encloses the pair of stripes. The wafer is then scribed and broken into individual dice, each forming a complete transistor, ready for mounting on the individual headers.

The excellent precision with which these processes can be carried out, plus the fact that each process is performed simultaneously on several hundred transistors, results in a relatively high yield and a high degree of uniformity in transistor characteristics.

The resistivity profile results from the diffusion of impurities of one type into a high resistivity region of opposite type to form the graded base region. The emitter is then shallowly alloyed or diffused. Since the collector is of higher resistivity than the base, this profile has the unique advantage over the alloy processes, in that the depletion layer extends into the collector. Thus, the punch-through vs. base-width compromise largely disappears. Furthermore, the diffusion process lends itself to precise control. Thus, very small base widths are possible with the result that mesa transistors have lower C_{De} and lower base storage, S_B, than their alloyed counterparts. Also, the emitter geometry is not restricted to a dot shape as in the alloy types. Therefore, a more optimum geometry can be used in the Mesa transistor to fabricate higher current devices, and maintain good frequency response.

The standard mesa transistor has some limitations. Because of the high resistivity of the lightly doped collector region needed to achieve the transistor's high reverse voltage breakdown characteristic, the collector series resistance is relatively high. This results in a fairly high saturation voltage which in turn places a limitation on the amount of current that can be drawn from the device without exceeding the power rating. A more serious limitation is the storage time resulting from the presence of this bulky collector region. The only solution is to reduce the resistivity which reduces the voltage rating, or employ (in the case of silicon) some method of special doping to reduce the lifetime, which also has the undesirable effect of reducing β.

The standard mesa transistors have switching speeds approximately equal to those of the micro-alloy diffused units. However, because the bulk of the semiconductor material is in the collector region rather than in the base, and because most of the heat which a transistor must dissipate is generated in the collector junction, the bulk collector of the mesa permits a low thermal impedance between the junction and the heat sink which gives mesa transistors far greater power-handling capabilities. The mesa has longer storage times but compensates for

this somewhat by having faster rise times. Moreover, the mesa geometry provides satisfactory performance at higher currents, and can withstand more severe environmental conditions.

The disadvantages of the Mesa types can be overcome by reducing the width of the collector region to a value just wide enough to handle the required depletion layer. In this manner, the problems associated with a high r_C and S_C would disappear. To keep the physical size within the limitations of physical handling would require the profile of Figure 1-14D. This profile — the most ideal of all the possibilities — has been made a practical reality with the introduction of the epitaxial process.

DIFFERENCES BETWEEN A MESA AND AN ALLOY TRANSISTOR: The two fundamental processes for making junction transistors are exhibited by the mesa and the alloy techniques. As shown in Figures 1-15 and 1-16, one of the basic differences is that both junctions are put into the semiconductor material from the same side in the mesa process, while in the alloy transistor the junctions are formed from opposite sides. In the mesa processes, the bulk of the semiconductor die is the collector instead of the base region as it is in the alloy process. Since most of the heat which a transistor must dissipate is generated in the collector-base junction, the single-sided geometry and bulk collector of the mesa permit a low thermal impedance between junction and heat sink.

Whereas the common alloy transistor contains junctions formed by recrystallization of molten material , the all-diffused mesa transistor contains junctions formed by the high temperature solid state diffusion of the doping impurities. The diffusion process yields flat sheet junctions which can cover an entire wafer or be separated into distinct precisely shaped regions, as desired.

THE EPITAXIAL MESA TRANSISTOR: The epitaxial process provides a means of growing very thin, high-purity, single-crystal layers of semiconductor material, on a very heavily doped crystal wafer of the same material. The epitaxial layer is a true continuation of the single crystal structure of the wafer.

The construction of transistors using the epitaxial process is very similar to that of the basic mesa except that an epitaxial layer of material is grown onto the basic wafer to form the effective collector region. This produces a number of important advantages.

1. The quality of the active semiconductor material is greatly improved.
2. The interrelation and compromises between physical parameters and electrical characteristics are minimized.
3. Compromises between interrelating electrical characteristics are minimized.

Quality in semiconductor material is dependent on (1) low surface dislocation density (relatively few flaws in the orderly pattern of the semiconductor lattice), (2) good control over the average number of doping atoms intentionally placed in the lattice, and (3) the uniformity of distribution of the doping atoms (i.e.,uni-

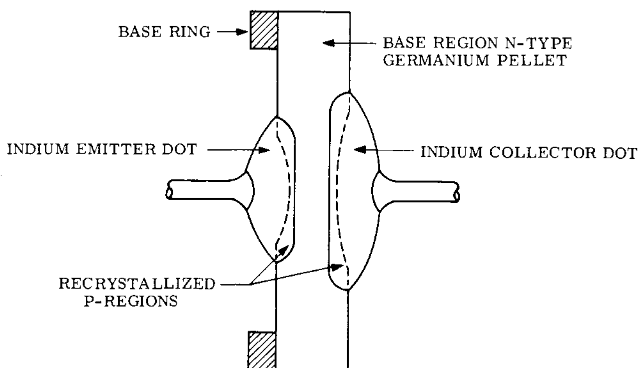


Figure 1-15 — Alloy Transistor Cross-section

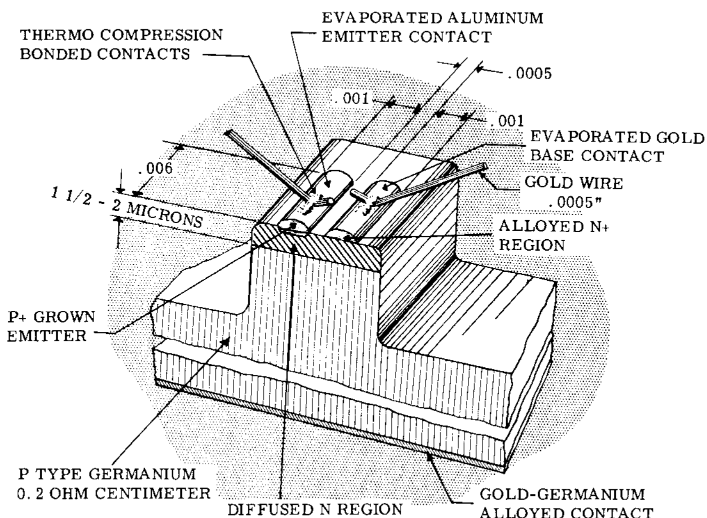


Figure 1-16 — Mesa Transistor Cross-section

formity of resistivity) across the face of the semiconductor wafer. The quality of semiconductor material (as just defined) obtainable from the epitaxial process is equal to or better than the best material that can be produced by any other means.

The amount of material actually needed for transistor action is far less than the smallest amount of material that can be conveniently handled or mounted to a header. The epitaxial process permits the use of the finest quality material to form the active region of the transistor, while supporting this active region on a much less-expensive substrate, which adds the physical strength and bulk to the transistor but contributes nothing to the performance.

A significant electrical compromise is the compromise between low switching losses in the on position (in saturation) and a reasonably high breakdown voltage across the switch in the off condition. The voltage drop across a transistor switch in the on condition (SV_{CE}) is only partly due to the transistor junction. Much of it is simple IR drop across the ohmic resistance of the collector, emitter, and base regions as the current flows from the high conductivity emitter lead through the active region of the transistor, through the semiconductor die, and finally out the high-conductivity collector case terminal. Since the emitter, base, and the active region of the collector are all very thin in the mesa transistor, the bulk of this IR drop is produced by the non-active portion of the collector region. Normally this non-active portion is some fifty times the thickness of the total active region of the transistor because it is necessary as a mechanical support for the device. The epitaxial process allows this mechanical supporting substrate to be of one impurity content while the active region of the transistor is formed in the higher purity (and hence higher resistivity) grown surface layer. The compromise mentioned above is resolved since the high resistivity active region, which is needed for high reverse breakdown voltage, is so thin that it adds very little to the on saturation voltage (SV_{CE}), while a low resistivity collector substrate provides a low resistance path from near the collector junction to the case.

An additional problem, which is greatly reduced by the epitaxial process, is that of excess charge stored in the collector region while the transistor is in the heavily conducting state. The use of a low resistivity (and thus high impurity density and low lifetime) material for the collector substrate, results in the restriction of the collector volume available for troublesome storage of excess charge, to the thin epitaxial layer (with its low impurity density). The low resistivity material of the bulky substrate, in an epitaxial mesa, will not long support the excess charges in their free state and, therefore, switch turn-off is much faster than in basic mesa transistors, where the relatively thick active collector region is made of high-resistivity material.

PLANAR TRANSISTORS: In the planar transistor, the area and shape of both the collector-base junction and the base-emitter junction is controlled during the diffusion process by selective masking. Here the "mesa" or active region of the transistor is turned downward into the wafer surface and the junctions are bent upward to terminate under a protective layer of oxide. All the fundamental physical characteristics that apply to a mesa transistor are present in the planar version; all junctions are formed from the same side; the semiconductor body is the collector; the transistor junctions are diffused, rather than alloyed; and an entire sheet of as many as 1,000 transistors is fabricated at one time. The electrical performance and the transistor design compromises are essentially the same as in the mesa, and in the epitaxial mesa when the planar uses epitaxial construction. The chief constructional difference between the two types is in the method of

controlling the size and shape of the active base region. In the conventional mesa, the control is accomplished by the chemical removal of the unwanted base material. In the planar transistor, the base is diffused through an oxide mask which limits the area available to the diffusant. Both construction technologies permit extremely close control of layer thicknesses, impurity profiles, and active cross-sectional areas.

ANNULAR TRANSISTORS: The Annular* process, a recently developed method of manufacturing high-frequency silicon transistors, circumvents some of the problems encountered in the planar process.

Present silicon transistors made by the planar process are limited in breakdown voltage by a phenomenon called "channeling". This is particularly evident in PNP devices, where the use of high-resistivity material for the collector region causes a reversal of material polarity in a narrow strip of collector material near the surface of the device. This polarity reversal produces an inversion layer at the surface which actually converts the surface portion of the *collector* region into an extension of the *base* region. The base-collector junction, therefore, rather than terminating at the top surface of the device, where it is covered by a protective oxide, is shifted to the edge of the transistor where it is exposed to the environment. This exposed surface causes an increase in leakage current far in excess of values that are tolerable for high-quality transistors.

The "channel," being uncontrolled, is quite erratic, both as to depth and apparent resistivity. This gives rise to severe instability problems even in applications where the high leakage current might be tolerable. As a result of these problems, silicon PNP planar devices are normally made with low resistivity collector regions, which retard channeling, but yield relatively low breakdown voltages.

The annular process circumvents the channeling problems prevalent in planar devices. Basically, it consists of a transistor structure having a deliberately induced channel with controlled characteristics which are relatively immune from the erratic variations of an uncontrolled channel. This induced channel is then terminated close to the base region of the transistor by means of a diffused annular ring, so that the channel cannot extend to the edge of the transistor. The collector-base junction, therefore, is terminated at the surface where it is protected from the environment by the oxide coating. This solves the problem of high leakage current while, at the same time, permitting the use of optimized resistivity material to obtain the desired transistor characteristics.

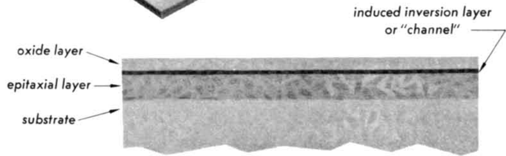
*Patents Pending.

1-5 — Fabrication Process for Annular Transistors

EPITAXIAL GROWTH AND OXIDE FORMATION: The various process steps are illustrated in Figure 1-17a thru i.

The epitaxial growth of the active region of the slice is done under ultra-clean conditions in an induction furnace. Silicon tetrachloride is carried in the vapor phase over the hot and polished surfaces of the substrate wafers. Under proper conditions of temperature, vapor pressure of the tetrachloride in its hydrogen carrier gas, and flow rate, a continuation of the existing substrate lattice will begin to grow, atom by atom, on the surface of the wafers. The dopant concentration of the epitaxial layer is controlled by the vapor pressure of dopant atoms in the gas flowing over the wafers. The desired thickness of the epi-layer is about 12 microns. When this thickness is reached, the composition of the gas flowing over the wafers is changed and an oxide is formed on all exposed semiconductor surfaces. An important feature of the process is that the oxide is formed directly on the ultra-clean epitaxial surface, without a chance for contamination.

Figure 1-17a



BASE MASK PREPARATION: The oxide-covered wafer is then coated with photo-resist* and exposed with the pattern of the desired base regions as shown in Figure 1-17b. After development of the photo-resist, the wafer is immersed in an etch that removes the oxide layer from those portions of the wafer that are to undergo base diffusion . The remainder of the oxide is protected by the layer of photo-resist and remains intact. After etching the base diffusion mask into the oxide-coated wafer, all traces of photo-resist are removed and the wafer is ready for base diffusion.

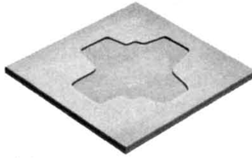


Figure 1-17b

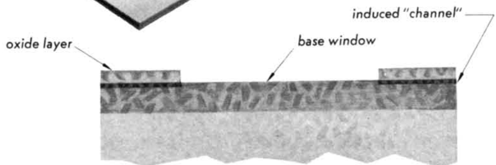
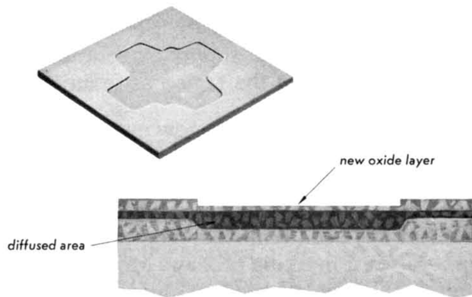


*Photo resist refers to a photographic process in which the material for treatment is coated with a photographic emulsion which is resistant to acid for short periods of time. By exposing and fixing selected patterns, portions of the material can be selectively etched.

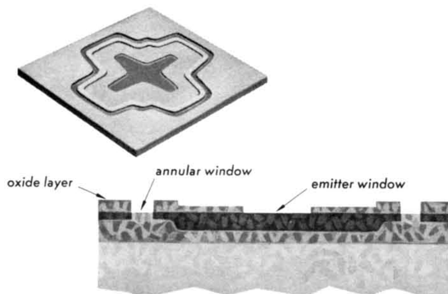
BASE DIFFUSION AND OXIDE FORMATION: Base diffusion takes place in a sealed furnace with carefully controlled amounts of the diffusant present. The oxide prevents diffusion into the wafer except where an etched window lets it through. An important characteristic of the process is that the diffusion proceeds in very uniform fashion from the window into the material. The mechanism is such that the diffused region spreads out under the oxide beyond the area of the etched window. The cross section in Figure 1-17c shows the diffused region to be larger in area than the window and the surrounding P-N junction is completely covered by the original oxide layer. Since the original oxide layer was applied as part of the epitaxial growth process, there is almost no possibility of a foreign impurity or ionized particle existing on the oxide-semiconductor interface at the point where the collector-base junction turns up to the "surface." At the conclusion of the base diffusion, oxygen is introduced into the furnace and a new layer of oxide is grown. The wafer is now again completely protected from outside contamination. However, during the formation of the base junction a "channel" is usually formed between the oxide layer and the collector region (especially if it is of high resistivity, as in the case of a high voltage transistor). The "channel" must be terminated by an Annular band during emitter junction formation.

Figure 1-17c



EMITTER AND ANNULAR BAND MASK PREPARATION: In exactly the same manner as the base mask was prepared, photo-resist is applied, exposed, and developed, to permit an acid solution to etch smaller windows in the new oxide as shown in Figure 1-17d. After the new windows are cut through to the semiconductor, using an etch that attacks only the oxide and not the semiconductor, all traces of the photo-resist are removed as before and the wafer is ready for diffusion.

Figure 1-17d



EMITTER AND ANNULAR BAND DIFFUSION: Diffusion is carried out in exactly the same manner as the base diffusion except that the diffusant is of the opposite impurity type. The cross section in Figure 1-17e shows that the emitter and annular band diffusion spreads out under the oxide as did the base diffusion. A new (third) layer of oxide is then grown.

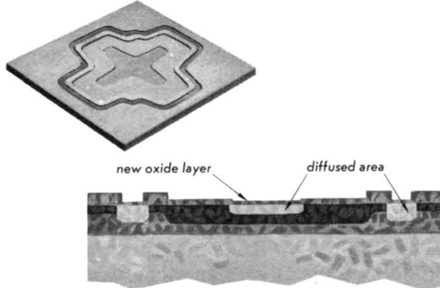


Figure 1-17e

OHMIC CONTACT MASK PREPARATION: Since the wafer of transistors is completely covered with oxide, the photo-resist masking process is again employed to cut windows for the application of ohmic contacts. The areas so exposed must not bridge or even approach too closely either of the lines where a junction comes to the surface, but since they will supply current to their related region, the contact area must cover as much of the region as practical.

METALLIZING CONTACTS: Metal is now evaporated in a thin layer over the entire active side of the wafer. In the contact areas the metal falls on semiconductor; over the rest of the wafer it falls on oxide. Since the oxide has as great an affinity for the metal as the semiconductor, the contacts cannot be alloyed until the unwanted metal has been removed. The excess metal is removed by a mild etch using the photo-resist masking process. Metal remains only on those areas of raw semiconductor which are to become ohmic contact areas. The metal is cleaned at the precise places where the thin connecting wires will be attached. These steps are detailed in Figure 1-17f.

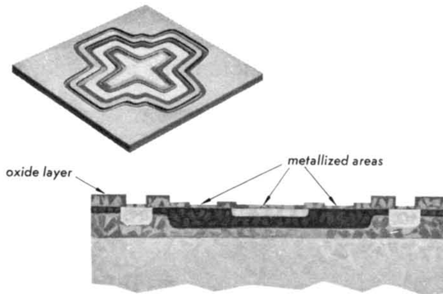


Figure 1-17f

SCRIBING AND DICING: All operations up to this point have been on an entire wafer at a time. Since a wafer contains up to 1000 transistors, it is eventually necessary to cut or break it up into individual transistor dice. This is done by a scribing operation in which a diamond point is pulled across the face of the wafer between rows and columns of transistors. After a chess-board pattern has been scribed onto the surface, the individual dice are separated by breaking the wafer along the scribe lines, in much the same way as window glass is cut to size.

ATTACHMENT OF DIE TO HEADER: In transistor types where it is desired to have the collector of the transistor electrically connected to the transistor can, the attachment of the die to the header is accomplished by a simple ohmic alloy process. Before dicing, the under side of the wafer is coated with a thin layer of gold. This coated side of the individual die is now alloyed to the gold plated header. In addition to making a good electrical connection, this method of mounting provides a low thermal path from the collector to the header. Figure 1-17g shows the die mounted to the header.

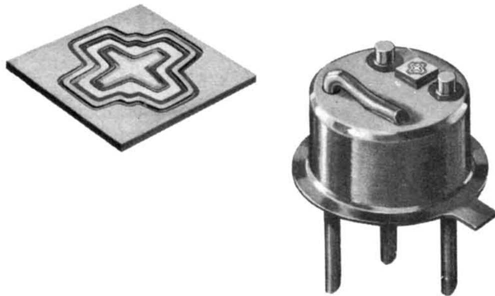


Figure 1-17g

ATTACHMENT OF LEADS TO DIE: Once the transistor die is fastened to the header, it is possible to move and orient it with precision. Jigs holding many transistors are fed into special thermo-compression bonding machines in which, by the combination of pressure and temperature, a 0.001 inch diameter connecting wire is secured between the contact area and header post. The detail in Figure 1-17h indicates the completed step.

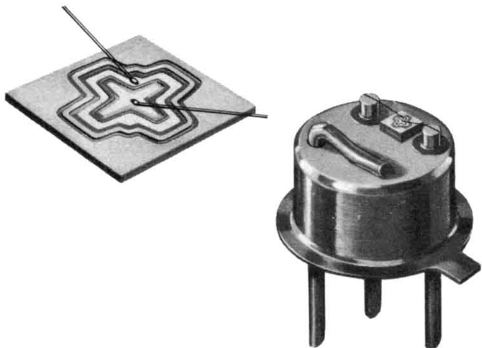
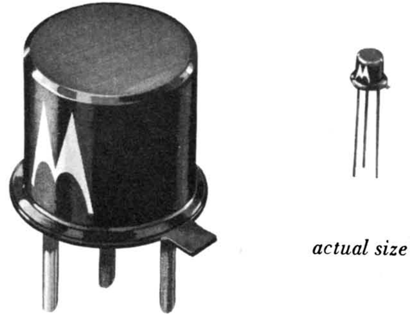


Figure 1-17h

HERMETIC ENCAPSULATION: At the present state of the art, almost all transistors are protected by means of the familiar hermetically sealed metal enclosure. Before being sealed, the unit on its header is baked and flushed with inert gas to remove residual moisture or other contamination. After the flushing and baking process, metal caps are placed over the transistor header and welded on. An enlarged view and the actual size of a type TO-18 transistor case is shown in Figure 1-17i.

Figure 1-17i



actual size

REFERENCES

1. Shockley, W.: *"The Theory of P-N Junctions in Semiconductors and P-N Junction Transistors" Bell System Technical Journal*, vol. 28, pp 435-489, July 1949.

CHAPTER 2

Switching Modes

Transistor switching circuits generally fall into three basic categories, depending on their operating mode. These are broadly classified as *saturated mode, current mode* and *avalanche mode,* and are determined by the portion of the transistor output characteristic curve utilized.

The operating regions for various transistor switching modes are shown in Figure 2-1. Here, it is evident that, for all modes, the switch-off condition is characterized by an excursion of the load line into the cut-off region of the transistor. The operating mode, therefore, is determined principally by the dc circuit state in the switch on condition, and by the location of the operating points.

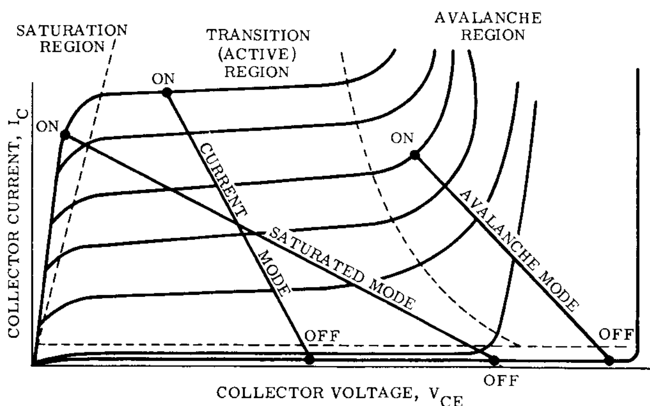


Figure 2-1 — Operation Regions for Switching Modes

2-1 — Saturated Mode Operation

Saturated mode operation most nearly duplicates the function of a mechanical switch. In the off condition, current through the switch is extremely small. In the on condition, the transistor, is driven into the saturation region which is distinguished by the fact that both the collector and the emitter are injecting carriers into the base, so that the transistor exhibits a virtual short circuit between its emitter and collector terminals.

Saturated mode operation has enjoyed widespread popularity. It is capable of producing a high voltage output, causes low transistor power dissipation in both the off and the on condition, and requires relatively few parts and simple circuitry. However, because the transistor is driven into saturation, it is troubled by storage delay time which limits switching speeds.

A common saturated mode switching circuit is illustrated in Figure 2-2. When the level of the input signal is at V_0, a voltage at or near ground potential, the base-emitter junction of the transistor is reverse biased by the action of V_{BB}, R_B and R_K. The transistor, therefore, is held in the cutoff condition, as indicated by Point B in the graph of Figure 2-3, and very little current flows through load resistor (R_L).

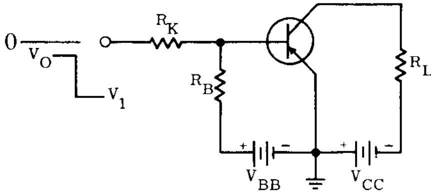


Figure 2-2 — Saturated Mode Switching Circuit

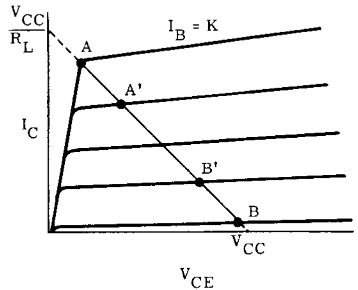


Figure 2-3 — The Saturated Mode Load Line

When the input is at V_1, a voltage that generally ranges in magnitude from 3 to 12 volts, the input current through R_K overcomes the bias current from V_{BB} and produces sufficient forward base current (I_B) to drive the transistor into saturation, placing operation at Point A on Figure 2-3. This constitutes the on condition and permits a load current flow approaching the limiting value of V_{CC}/R_L.

The basic circuit of Figure 2-2 is subject to a number of variations. In most practical applications, the load must have one terminal grounded and cannot be inserted between the collector and the power source. Under these conditions, the circuit in Figure 2-4 is normally employed.

In this circuit, the switching transistor is placed in parallel with the load so that the load is virtually short circuited when the transistor is turned on. With the transistor turned off, its output represents a virtual open circuit across the load, and load current approaches the limiting value of $V_{CC}/R_C + R_L$.

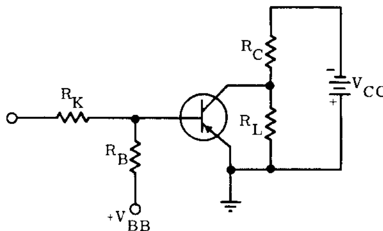


Figure 2-4 — Saturated Mode Switching Circuit

With such circuits, it is evident that the voltage across R_L is dependent upon the ratio of R_C to R_L. In applications such as computer logic systems, R_L is often a variable, depending upon the number of subsequent circuits that are turned on during a given logic operation. It is necessary to provide a nearly constant output voltage to assure enough, yet avoid excessive, drive into subsequent stages, regardless of their number. A nearly constant output level is achieved by the addition of a clamping diode circuit, as shown in Figure 2-5. When the transistor is in the on state and represents a virtual short circuit across the load, the diode is reverse biased and has no effect on circuit performance. At this point, the output voltage (V_O) is equal to the saturation voltage (SV_{CE}) of the transistor which is normally a fraction of a volt. When the transistor is cut off, representing an open circuit, the voltage from V_{CC} divides between R_C and R_L in accordance with the resistance ratio between these units. If the load voltage attempts to exceed the value of V_K, the diode becomes conductive and clamps the output voltage (V_1) to a value of $V_K + V_D$ where V_D represents the diode drop.

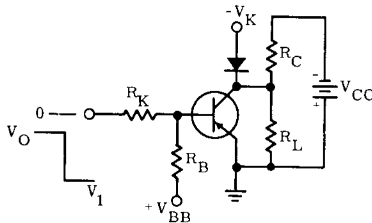


Figure 2-5 — Saturated Mode Circuit with Clamp Diode

In practice, it is usually possible to design the circuit so that this condition exists for all anticipated values of R_L, so that the output voltage will vary from virtually zero (with the transistor in saturation) to some fixed value, $V_K + V_D$ (with the transistor cut off).

The dc design of saturated circuits is fairly straight forward. The principal problem revolves about the selection of R_K and R_B so that the proper off and on conditions for the transistor are met. This problem is resolved by solving a pair of simultaneous equations once the transistor parameter limits are accurately known. It is necessary to accurately know limits for the saturation voltage SV_{CE} and the input voltage V_{BE} as affected by collector current, base current and temperature. The limit of base leakage current in the off condition must also be known at the maximum temperature. A design procedure, together with an example worked out in detail for this type of circuit is given in Chapter 7, Section 1.

A serious drawback of saturated mode operation is storage delay time. That is, there is a finite, often undesirably long, delay between change of input voltage from V_1 to V_0 and response of the transistor collector current. This delay results from overdrive, or excess base current, which is used to drive the transistor into the saturation region. The excess base current results in an accumulation of stored charge in the base and/or collector, which must be removed before collector current can change.

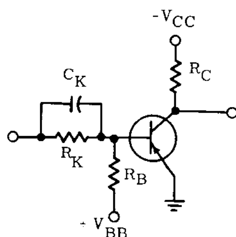


Figure 2-6 — Use of a Capacitor to Nullify Stored Charge

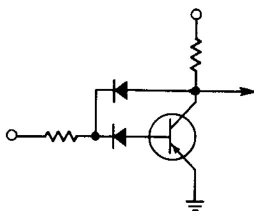


Figure 2-7 — The Baker Clamp Circuit

Various methods have been devised to overcome the storage time problem. The simplest method is to use a capacitor connected in parallel with the drive resistor, R_K, as shown in Figure 2-6. Since the voltage drop across R_K is fairly high, the capacitor provides a place for charge to be stored when the transistor is on. As the input signal changes, the charge on the capacitor is forced into the base of the transistor. This charge can effectively cancel the transistor stored charge, resulting in a reduction of storage time. This method is very effective if the output impedance of the preceding stage is low so that the peak reverse current into the transistor is high.

Another method of reducing turn-off delay time is simply to keep the transistor out of saturation. The use of diodes in a feedback arrangement devised by Baker[1], as in Figure 2-7, is quite effective. The voltage drops across the diodes are such that the collector junction can never become forward biased and any base driving current in excess of that required to bring the transistor to the edge of saturation is simply channeled around the base through the upper diode.

This method greatly complicates the dc design of the circuit because device specifications must be such as to keep the transistor out of saturation under worst-case conditions. Storage time problems are now transferred to the diode which must have a much faster recovery time than the transistor if this method is to be of much value.

Another approach, developed by Pressman[2], uses resistors in a linear feedback network, as shown in Figure 2-8. Since this circuit requires high transistor gain, it is common to use two transistors in a compound or Darlington connection. The circuit is particularly effective in computer logic circuits using diodes, where the base circuit is open when the transistor switch is on.

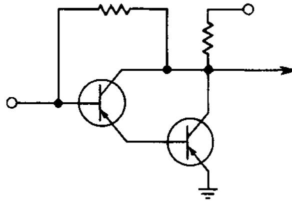


Figure 2-8 — Resistor Coupled Anti-Saturation Circuit

Both of these feedback methods sacrifice dc design simplicity and low on voltage for higher speed. They prevent the operating point from moving into the saturation region by setting the on point at or near Point A′ of Figure 2-3. Another method of preventing saturation is to change the mode of operation to that of the current mode, which will be discussed shortly.

2-2 — Saturated Mode Coupling Circuits

Most switching circuits are direct coupled. This not only increases circuit efficiency, but is generally necessary because circuits must remain in one state or the other for periods of time which are too long to make ac coupling practical.

The dc coupling technique plays a vital role in determining the switching characteristics of a circuit. The differences in coupling techniques are particularly significant in logic circuits, as used in digital computers. These circuits are characterized by a "fan-in", where one circuit is driven by several inputs and a "fan-out" where the circuit must deliver outputs to a number of subsequent stages. The function of logic circuits or gates is to be either off or on, corresponding to a "1" (high output) or a "0" (low output) depending upon the state of the various inputs.

The circuit behavior is described by Boolean algebra, the mathematical expressions for the language of logic. These expressions are shown on Figures 2-9 through 2-11. The significance of these equations will be explained in each case. In all cases, the function of logic circuits, or gates, is to provide an output only for a predetermined combination of input signals.

RESISTANCE COUPLING: Circuits similar to those of Figures 2-4 and 2-5 are examples of resistance coupling and, when used in logic circuits, are called resistor-transistor logic circuits, or RTL.

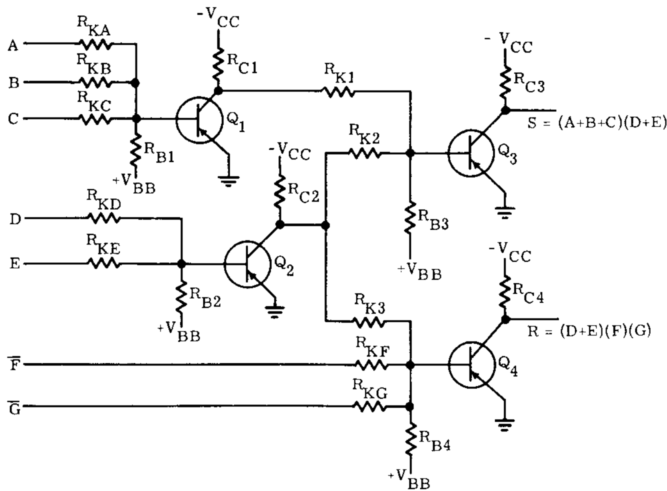


Figure 2-9 — Resistor — Transistor Logic (RTL)

The advantages and limitations of this coupling method can be evaluated from the schematic of a typical RTL gate shown in Figure 2-9. In this circuit, transistor Q_1 represents a logic gate that is turned on (output is "0") when any or all of the signals at Points A, B, and C go negative, which represents a "1" input. Q_2 is turned on when either signal D or E, or both, go negative. Both Q_1 and Q_2 are called "Nor" circuits. A "Nor" circuit produces a "0" output when any or all inputs are "1". Q_3 goes off, thereby producing a "1" output only when Q_1 and Q_2 are energized by a "1" at A *or* B *or* C *and* D *or* E. Q_4 is turned off when D or E signals are "1" provided that $\overline{F}$ *and* $\overline{G}$ are "0". Q_3 and Q_4 are also called "Nor" circuits because they produce a "0" output only when any or all their inputs are "1".

Cascaded "Nor" circuits produce the "And" function. The output of Q_3 can be stated: S is a "1" when at least A *or* B *or* C *and* either D *or* E is a "1". The output of Q_4 can be stated: R is a "1" when at least D *or* E *and* F *and* G is a "1". The symbols $\overline{F}$ and $\overline{G}$ indicates a "complement" of F and G; that is, when F is a one, $\overline{F}$ is a zero, etc.

The chief advantages of RTL design are that the circuit is quite simple and uses a minimum number of transistors and other parts. The disadvantage becomes evident when considering the design of the stages individually. For example, if Q_1 must be designed so that it will be driven into saturation by only a single input "1", then the application of three simultaneous "1" levels would move the operating point far into the saturation region and would result in a long storage time should the three inputs change to "0" simultaneously. An attempt to compensate for this increase in storage delay by increasing reverse-base-current (turn-off current) results in driving the transistor deeper into the cut-off region (during the off state) and causes an increase of turn-on delay time. Increased reverse base current from the V_{BB} — R_B source also results in reduced circuit gain since this increase in current through R_B must be overcome by an increase in current

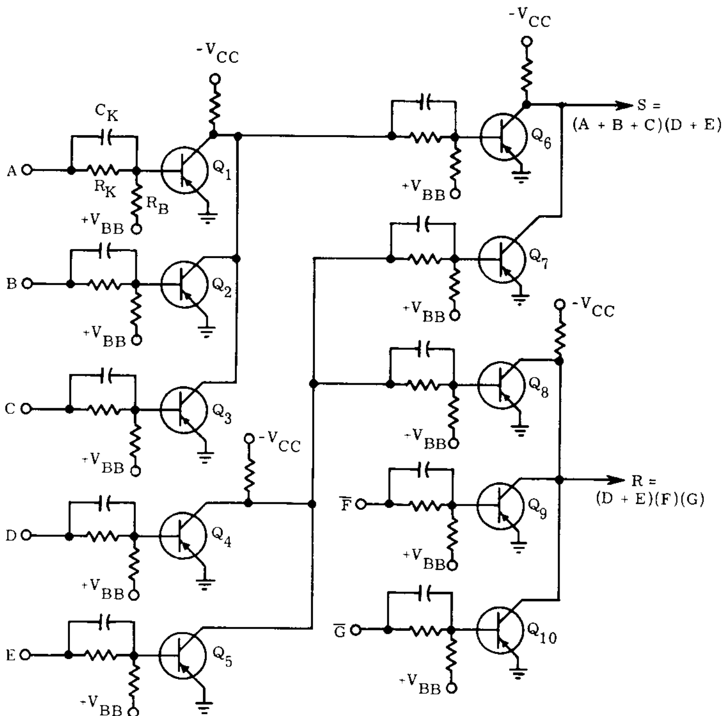


Figure 2-10 — Resistor — Capacitor Transistor Logic (RCTL)

through R_K in order to drive the transistor into saturation. While it is possible to design an optimum input network for any given number of fan-ins, resistance coupling still results in relatively slow switching speeds.

RESISTANCE-CAPACITANCE COUPLING: It was mentioned previously that a capacitor could be connected across the coupling resistors to enhance switching speeds by providing high peak currents when the input signal changes state. It is normally not practical to do this by simply adding capacitors to the RTL configuration of Figure 2-9, because of the extreme sensitivity to noise which this produces. If, for example, all the coupling resistors (the R_K's) were bypassed with capacitors, noise pulses at any of the inputs would be coupled directly to the transistor bases through the coupling capacitors and could easily produce spurious triggering. This problem is normally so severe that the speed of RTL circuits is never enhanced in this way.

As a rule, the use of capacitors leads to the configuration in Figure 2-10, normally referred to as resistor-capacitor-transistor logic (RCTL). This circuit results in the same type of logic as the previous one, but the number of transistors has been greatly increased. With each input circuit being isolated from the others through a transistor, noise problems are not severe.

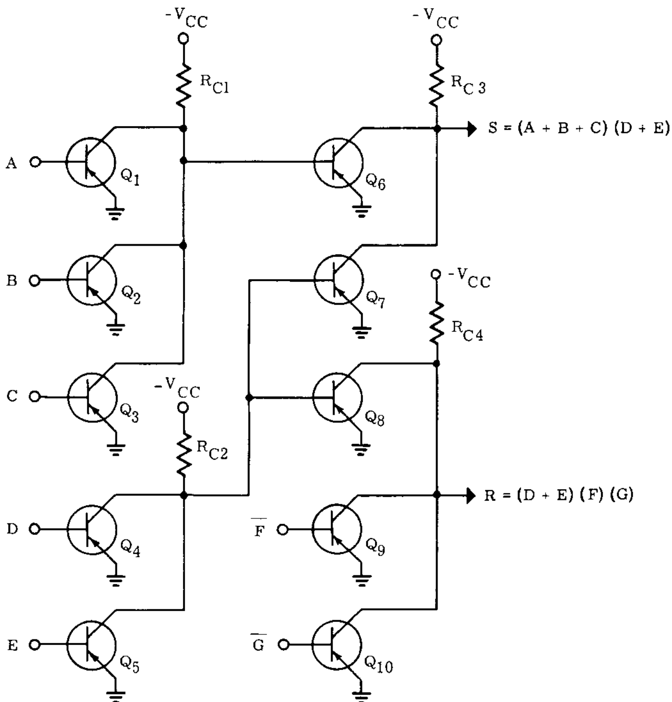


Figure 2-11 — Direct Coupled Transistor Logic (DCTL)

RCTL usually permits repetition rates of ten times that of RTL. The upper limit upon repetition frequency is determined by the size of the capacitor which in turn is determined by the stored charge of the transistor. It is necessary for the capacitor to reach equilibrium before any change of state occurs. When a transistor has turned off, some charge is usually left on the capacitor. This charge must decay through R_K and C_K, and circuit recovery time can be long.

DIRECT TRANSISTOR COUPLING: A simplification of the RCTL circuit, called direct-coupled-transistor logic (DCTL) is shown in Figure 2-11. This configuration offers about the same order of switching speed as RCTL and uses the same number of transistors, but requires fewer resistors and power supply sources. On the other hand, DCTL is characterized by an extremely small output signal, a high susceptibility to ground noise and a critical dependence upon a tight distribution of several transistor characteristics. These limitations become evident in an analysis of the circuit.

With zeros applied at Points A, B, and C, transistors Q_1, Q_2 and Q_3 are cut off and the collector voltage normally would attain a value of V_{CC}. However, due to the load circuit represented by the base-emitter diode of Q_6, the collector voltage is clamped to the base voltage of Q_6. Transistor Q_6 is turned on hard by the heavy drive current and is driven deep into saturation which results in a very low saturation voltage and fast rise time but also, unfortunately, in a high stored charge.

Although the stored charge is discharged rapidly through the low output resistance of Q_1, Q_2 or Q_3, when one of these is turned on, the storage time delay for DCTL circuits is longer than for comparable RCTL configurations. On the other hand, the elimination of R_K and C_K in DCTL circuits eliminates recovery time considerations so that the maximum switching speeds attainable with the two circuits are quite similar.

Fan-out of DCTL configurations is severely limited due to a problem called "current hogging". In the schematic it will be noted that the load for transistor Q_4 (or Q_5) is the parallel inputs of transistors Q_7 and Q_8. If one of these load transistors has a lower turn-on voltage than the other, the output voltage of the driver transistor may be clamped to a value which is insufficient to turn on the second load transistor. The problem becomes more serious as more load transistors are added and imposes stringent tolerances on the value of V_{BE}.

The noise susceptibility of DCTL circuits results from the small signal levels and the absence of turn-off bias. With all transistors operating at the edge of conduction in the off condition, only a small amount of ground noise injection at the emitter can cause a spurious signal output which could trigger subsequent stages.

SUMMARY: Depending on the specific application, each coupling method has some advantages and limitations. RTL, for example, requires the fewest active elements and permits high fan-out; but it is relatively slow. DCTL offers high speed, but requires many active elements and is critical of transistor parameters. RCTL, which is perhaps the most popular coupling method, offers high speed and high fan-out, but requires a large number of active elements and is somewhat more complicated than the other forms. It is generally used in conjunction with diodes, to perform logic functions, but the basic RCTL configuration forms the heart of more complex switching circuits such as multivibrators and trigger circuits.

Many combinations of diodes and transistors to perform logic have been developed. A treatment of many of these is given in reference 3.

2-3 — Current Mode Operation

As mentioned earlier, high switching speed can be obtained as a result of eliminating storage delay time by limiting the on excursion of the transistor load line to Point A′ in Figure 2-3. This can be accomplished most effectively by operating in the current mode.

The basic current mode circuit is illustrated in Figure 2-12. Here a voltage source, V_{EE}, produces a current flow through diode D_E and resistor R_E. The voltage drop across the forward biased diode, a few tenths of a volt, appears between the emitter of transistor Q_1 and ground; hence a slightly positive signal voltage on the base of Q_1 is required to keep the transistor in the cutoff state. A negative voltage on the base of Q_1 turns on the transistor which turns off the diode.

Analysis of the circuit reveals that, if V_{EE} is much larger than the signal voltage, the current produced by V_{EE} is approximately V_{EE}/R_E, and is relatively constant regardless of the transistor state. With Q_1 cut off, this current flows through diode D_E; with Q_1 turned on, it flows into the emitter of the transistor.

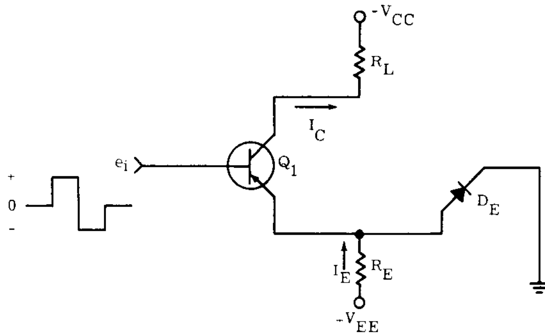


Figure 2-12 — Basic Current Mode Circuit

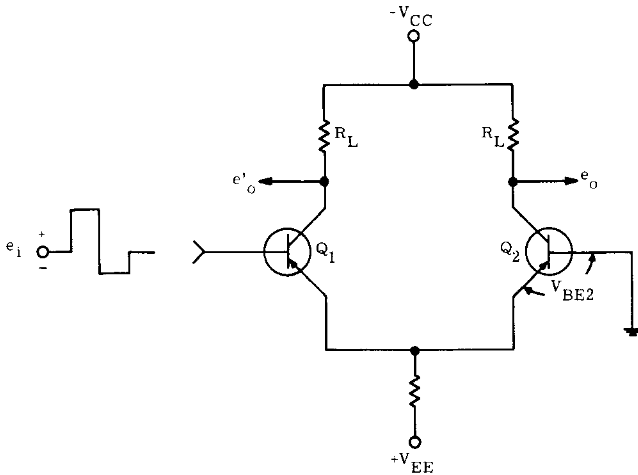


Figure 2-13 — Current Mode Circuit Having Complementary Outputs

As a rough approximation, if the values of V_{CC} and R_L are selected so that V_{CC}/R_L (the limit of collector current in saturated mode circuits) is greater than V_{EE}/R_E (the current that flows into the emitter), it should be evident that the collector current cannot enter the saturation region, and is limited to a point represented by A′ in Figure 2-3. Hence, current mode operation prohibits transistor saturation, which eliminates storage time as a speed limiting factor.

A more versatile current mode circuit, one that provides two out of phase output signals (complementary signals) simultaneously, can be obtained by replacing the diode of Figure 2-12 with a transistor, as shown in Figure 2-13. The operation of the circuit is exactly as when the diode is used.

If transistor Q_2 is conducting, the emitter bus will assume a potential of $+V_{BE2}$ due to the voltage drop across the base-emitter junction of Q_2. For germanium mesa transistors, this potential is approximately 0.5 volt . With the emitters positive, the input signal (e_i) to Q_1 must be positive by at least ($V_{BE} - V_T$)* in order to keep this transistor in a cutoff condition.

As e_i goes negative, current from the emitter-current source (V_{EE}) divides between Q_1 and Q_2 until e_i reaches a value of $-(V_{BE} - V_T)$. At this point, the emitter bus will be at a potential of V_T, causing the grounded-base transistor (Q_2) to cut off and permitting all the current from V_{EE} to flow through the base-emitter junction of Q_1.

The output voltage, e_o, obtained from the collector of Q_2 is in phase with the input signal, while e'_0, the output of Q_1, is 180° out of phase. Thus, complementary output signals are available at the same time.

In addition to high speed, current mode operation produces a number of significant benefits. Among these are: excellent dc stability, high noise immunity, and non-critical transistor parameters. For logic applications, this type of circuit has the disadvantage of requiring a relatively large number of transistors in comparison with saturated mode operation. This problem, today, is largely resolved by the availability of low-cost transistors and by the development of integrated circuits wherein multiple transistors can be manufactured as inexpensively as individual units.

2-4 — Coupling Techniques for Current Mode Circuits

In saturated mode circuits, the output voltage varies from virtually zero (when the transistor is in saturation) to V_{CC} or some clamp level (when the transistor is cut off). This output signal is of the proper polarity and has the required level to drive the following stages without requiring any special coupling considerations. In current mode operation, however, the output signal consists of voltage levels which vary about a reference level different from the input reference level. Direct coupling, therefore, cannot be employed without special coupling techniques to translate the output signal to the proper input level.

One common coupling method employs alternate PNP and NPN blocks in a configuration as shown in Figure 2-14. In this circuit, the base of each NPN transistor stage is returned to a -3 volt reference level rather than ground so that the input signal must vary above and below this -3-volt reference level. When transistor Q_2 is on, a current $\frac{V_{EE} - V_{BE}}{R_E}$, approximately 20 mA, flows through the load resistor R_L. Simultaneously, an opposite-polarity current $\frac{V_{CC} - V_L}{R_C + R_L}$, of 10 mA, flows through this resistor. As a result of the net current, a potential of 1 volt appears across R_L. This voltage drop adds $+1$-volt to the -3 volt reference level of Q_3, turning it on.

When Q_2 is cut off, only the current $\frac{V_{CC} - V_L}{R_C + R_L}$ flows through R_L. The resultant drop across R_L adds -1-volt to the -3 volt reference level, causing Q_4 to conduct and Q_3 to cut off.

*V_T is defined as the base-emitter voltage at the threshold of conduction where I_C is negligibly small.

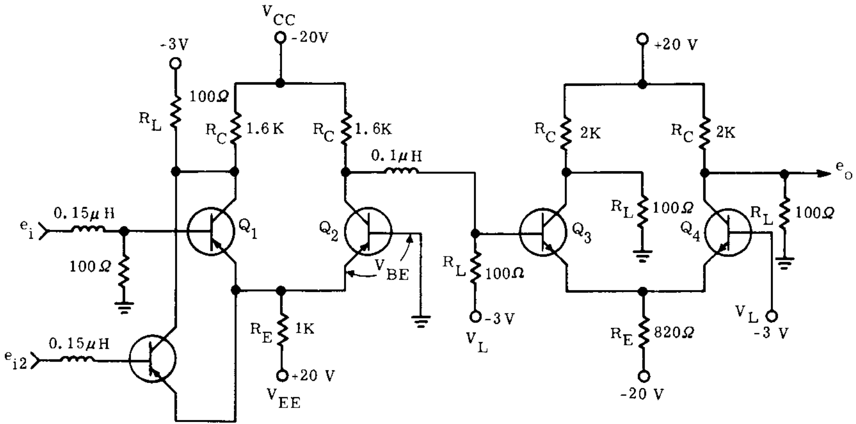


Figure 2-14 — Cascade Complementary Gate

It is important to note that the voltage at the collector of Q_2 never drops below —2 volts with respect to ground. Since the base of Q_2 is grounded, the collector-base junction of this transistor is always reverse biased and is kept out of the saturation region.

Another coupling method commonly employed for level translation in current mode circuits is shown in Figure 2-15. This method utilizes zener diodes as coupling elements and permits the use of transistors of the same polarity in cascaded current mode stages.

In this circuit, the base of each transistor is referenced to ground so that the input signal is required to vary above and below the zero reference level. If the load voltage is to vary, for example, from + 1 volt to — 1 volt, then with the 3.3 volt zener diodes shown, the collector voltage level must vary from —2.3 to —4.3 volts. The collector level with these potentials is always comfortably above the edge of saturation.

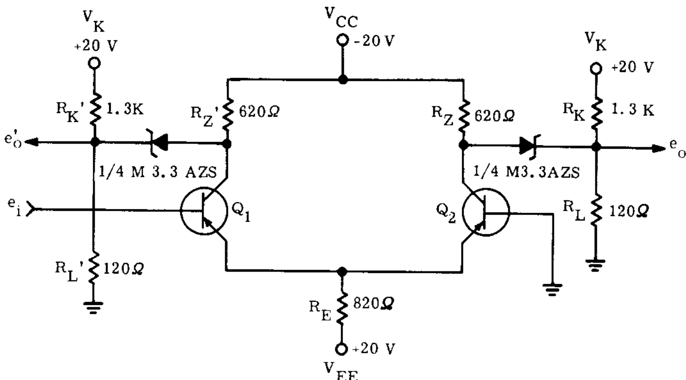


Figure 2-15 — Zener Diode Coupled Inverter

Operation of the circuit is as follows: When transistor Q_2 is off, two opposing currents flow through the load resistor R_L. The voltage source (V_{CC}) produces a load current of approximately 25 mA, while the voltage source V_K produces a load current of approximately 16 mA. The net current through R_L causes a voltage drop of about —1 volt to appear across this resistor.

When Q_2 is turned on, a transistor current of approximately 24 mA is injected into the node at the collector terminal in opposition to the current caused by V_{CC}. This reduces the zener diode current to about 4 mA, which is still sufficient to keep the diode in conduction. The load current now consists of a V_{CC} current of about 4 mA and an opposing V_K current of about 15 mA. The net current through R_L, therefore, still produces a 1-volt drop, but this time the V_K current predominates so that the load voltage is positive with respect to ground. Hence, the conditions for level translation are fulfilled.

In all current mode circuits, transistor saturation is avoided by limiting the emitter current to a value that is less than the normal collector current limiting condition represented by V_{CC}/R_L. When designed for a specific load, excessive variations in load conditions could result in improper performance. If the load is too light, it would be possible to enter the saturation region. If it is too heavy, the available output voltage might be insufficient to assure proper turn-on of the succeeding stage. Thus, the load must remain within fixed limits. The resulting problems differ for the two types of coupling methods described and an analysis is provided in Chapter 8.

A current mode logic circuit using complementary transistors is shown in Figure 2-16. Note that the logical outputs are the same as for the previous circuits. While the current mode configuration uses more transistors than the saturated mode RCTL circuit, its speed is approximately five times greater.

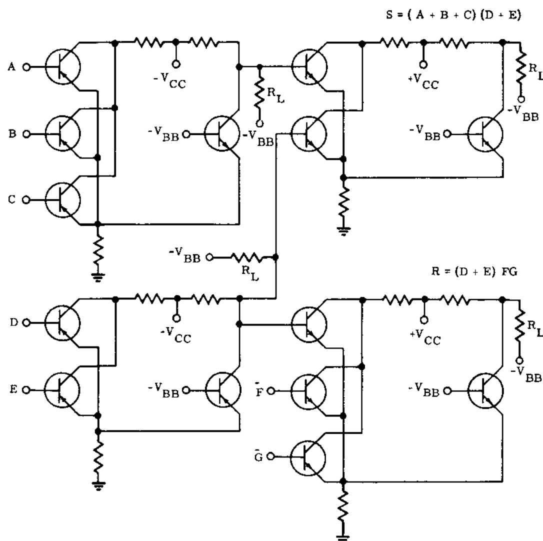


Figure 2-16 — Current Mode Logic Circuit Using Complementary Transistors

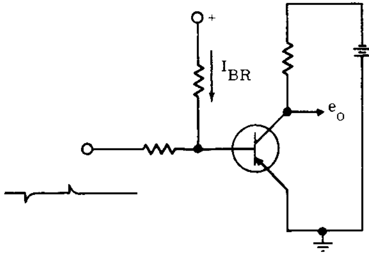


Figure 2-17 — Basic Avalanche Mode Circuit

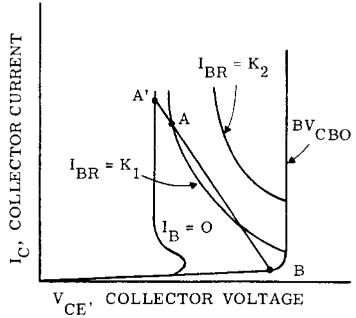


Figure 2-18 — Load Line for an Avalanche Mode Switch

2-5 — Avalanche Mode

Operation in the avalanche mode utilizes the negative-resistance characteristics of transistors, which result from operation in the common-emitter breakdown region. Figure 2-17 shows an elementary circuit capable of operating a PNP transistor in the avalanche mode. Its operation may be understood with the aid of Figure 2-18.

Assume, initially, that $I_{BR} = K_1$ is a small reverse current which holds the transistor operating point at Point B. If a negative trigger voltage is applied, so that the base current is reduced to zero, the operating point shifts to Point A′ on the $I_B = 0$ curve. The extreme speed with which it does this is the chief attraction of avalanche mode operation. When the trigger pulse disappears, the operating point shifts slightly to Point A and remains at that level. The switch, therefore, has two stable states. To return the switch to Point B it is necessary to apply sufficient reverse current to allow only a single stable condition. A small positive trigger accomplishes this, as indicated by $I_{BR} = K_2$.

Because of instability problems associated with the negative resistance region, avalanche mode circuits are normally ac coupled and do not find general use in logic systems. Obviously, a thorough characterization of a transistor's avalanche region is necessary to properly utilize operation in the avalanche mode. Primary applications and a more detailed analysis of this operating mode are given in Chapter 9.

REFERENCES

1. Baker, R. H., *"Maximum Efficiency Switching Circuits"*, MIT Lincoln Lab Report TR-110 (1956).
2. Ditkossy, H. and Pressman, A. B., *"A High Speed Anti-Saturation Inverter Logic Circuit"*. IRE Wescon Convention, 1961.
3. Pressman, A. B., *"Design of Transistorized Circuits for Digital Computers"*, John F. Rider Publishers, New York, N.Y. 1959.

CHAPTER 3

Transistor Characteristics Influencing Off Condition Design

As pointed out in Chapter 2, the off condition for the three modes of operation is the same, although the load line for each mode is different. The off region is defined as the state where the current through the transistor is at its minimum value and the voltage across the transistor is at its maximum value.

When a transistor switch is in the off state, there are three principal transistor characteristics that affect circuit design. They are: (1) leakage currents, (2) turn-on threshold voltage, and (3) avalanche breakdown (latch-up).

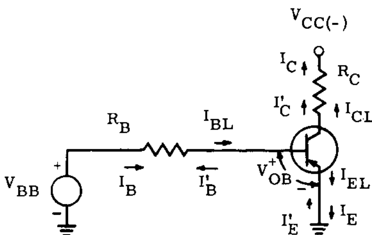
Most switch designs are based upon worst-case techniques which means that limit values are required for the dc characteristics of the transistor. However, most data sheets do not provide the necessary limit design characteristics for direct design application. For this reason, Motorola has devised a new data sheet — called a Designer's Data Sheet — to provide limit design data which permits complete switching circuits to be designed from the given data.

The concept of the Motorola Designer's Data sheet is relatively new. Therefore, the early part of this chapter discusses the off condition design requirements, and the appropriate off condition characteristics as obtained from the 2N964A Designer Data Sheet. The latter portion is devoted to explaining how to obtain limit design data from the conventional data sheet.

3-1 — Leakage Currents and Threshold Voltage

When a transistor switch is in the off condition, there are always some residual transistor leakage currents flowing in various legs of the circuit. Some of these currents are minute compared to other circuit currents and can be considered negligible. However, in certain portions of the circuit the leakage currents are not negligible in comparison with the other currents and therefore they must be considered as part of the design.

The currents which flow in a transistor are indicated in Figure 3-1. For a transistor to be truly off, only the residual cutoff or leakage currents should flow; i.e., $I'_C = I'_E = I'_B = 0$.



I_{CL} = The collector cutoff current with the collector junction reversed biased. I_{CL} is generated by the same components that compose I_{CBO}. However, I_{CBO} is measured with the emitter open and thus does not express the leakage current which is normally slightly less than I_{CBO} when the emitter is connected in a practical circuit.

I_C' = Portion of injected emitter current reaching the collector $(\alpha I_E')$

I_C = Total collector current $(I_{CL} + I'_C)$

I_{EL} = The emitter cutoff current with the emitter junction reversed biased. I_{EL} is generated by the same components that compose I_{EBO}. However, I_{EBO} is measured with the collector open and thus does not express the leakage current which is normally less than I_{EBO} when the collector is connected in a practical circuit.

I'_E = Total emitter current $(I_{EL} - I_E')$

I_{BL} = Base leakage current $(I_{CL} + I_{EL})$

I_B' = "On" base current (I_C'/β)

I_B = Total base current $(I_{BL} - I_B')$

Figure 3-1 — Current Flow Near the Cutoff Region

The typical behavior of the collector currents and base currents for a germanium transistor, in the vicinity of the cutoff region, is shown on Figure 3-2. It is evident, from the figure, that when $V_{BE} = 0$ the transistor is conducting slightly and is not truly off. Therefore if a transistor switch is to be truly off, a reverse bias voltage (V_{OB}) must be applied. A reverse bias equal to V_{TR} can be regarded as the minimum reverse bias voltage that must be maintained to keep the transistor cutoff. With reverse bias voltages greater than V_{TR}, the base current, I_B, equals I_{BL} and the collector current, I_C, equals I_{CL}. The effect of the currents I_{CL} and I_{EL} is usually negligible in the output circuit. However, I_{BL} normally flows through a relatively high resistance (R_B on Figure 3-1). The voltage drop across R_B developed by I_{BL} causes a small forward base-emitter bias which turns the transistor on slightly. Therefore, I_{BL} and V_{TR} under worst case conditions must be considered when designing the dc off condition for a transistor switch.

For circuit analysis, it is convenient to define a voltage, V_{TF}, at the threshold of conduction in the forward direction. V_{TF} is not a well defined point but can arbitrarily be taken at the point where I_C has increased about an order of magnitude from its cutoff value. For the transistor used to prepare Figure 3-2, V_{TF} is approximately 0.1 volt.

I_{BL} is related to the collector base diode leakage current (I_{CBO}) and the emitter-base diode leakage current (I_{EBO}). The relationship is a little involved, and will be developed later.

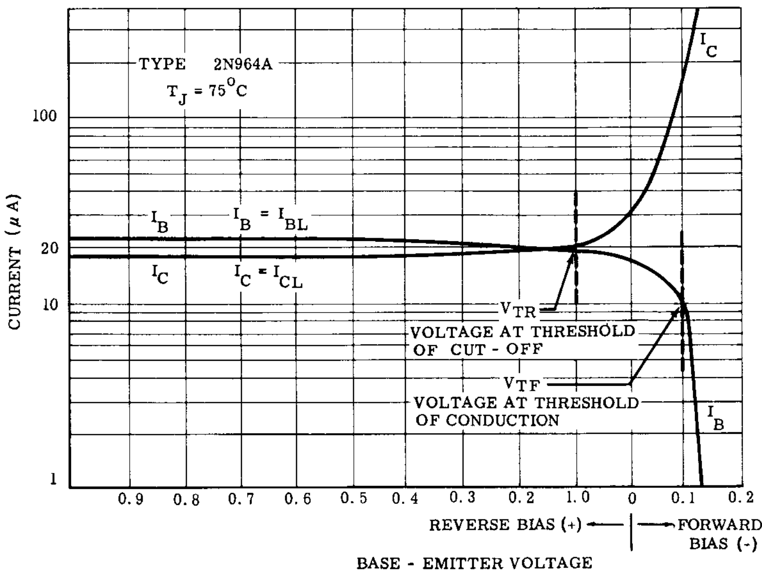


Figure 3-2 — Effect of Base Voltage Upon Current Flow

Complete worst-case design data for I_{BL} and V_{TR} are conveniently shown in the graph of Figure 3-3 for a 2N964A germanium transistor. As can be seen from the graph, both V_{TR} and I_{BL} are temperature sensitive, and I_{BL} increases with increasing reverse bias. Any reverse bias voltage (V_{OB}) selected to the right of the V_{TR} axis will maintain the transistor switch in the cutoff region. I_{BL} is specified at values of reverse bias greater than the minimum required, to handle those occasions where a reserve reverse bias is desired in order to provide noise immunity, or where large reverse biases are encountered as in multivibrator circuits.

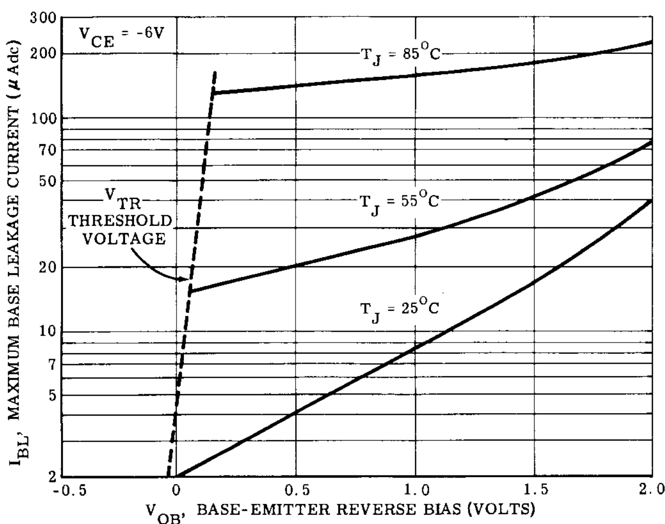


Figure 3-3 — Base Leakage Current for 2N964A

I_{BL} can be obtained at temperatures other than those specified by drawing additional curves on Figure 3-3. When $V_{OB} = V_{TR}$, I_{BL} increases exponentially with temperature; since the ordinate is a log scale for current, the line V_{TR} also represents a linear temperature scale for the beginning of the I_{BL} curves. Therefore, it can be easily marked to indicate 5 or 10 degree increments. Then, using the given curves as a guide, additional curves can be drawn at desired junction temperatures.

Since most data sheets do not specify the base leakage current, it must be estimated from I_{CBO} and I_{EBO} specifications. However, to estimate the worst-case I_{BL} from I_{CBO} and I_{EBO} specifications, the worst-case values of these specified leakage currents must be determined at the desired operating conditions.

3-2 — Factors Contributing to Leakage Currents in a PN Junction*

Whenever a semiconductor PN junction is reverse biased, a reverse or leakage current (I_R) flows across the junction. The reverse current is the sum of three currents: I_D, due to the diffusion, I_G, due to charge generation and I_S, due to surface leakage. Therefore:

$$I_R = I_D + I_G + I_S. \tag{3-1}$$

The diffusion current, (I_D) is caused by minority carriers diffusing across the junction. The total quantity is dependent upon temperature and is relatively independent of the applied voltage. The terminal current, due to I_D, can be found from the "ideal diode" equation

$$I = I_D \left(e^{\frac{qV}{kT}} - 1\right) \tag{3-2}$$

where: I = the diode current
I_D = the reverse diffusion current
q = electronic charge
k = Boltzmann's constant
T = Absolute temperature
V = Applied voltage.

(q, k, T): $\frac{kT}{q} = 26$ mV @ 27°C (300°K)

A plot of equation 3-2, with I_D taken as unity, is shown in Figure 3-4. The reverse voltage range above 0.1 volt is called the voltage saturation range since the current essentially becomes constant with voltage. For this reason, I_D is sometimes referred to as the reverse saturation current.

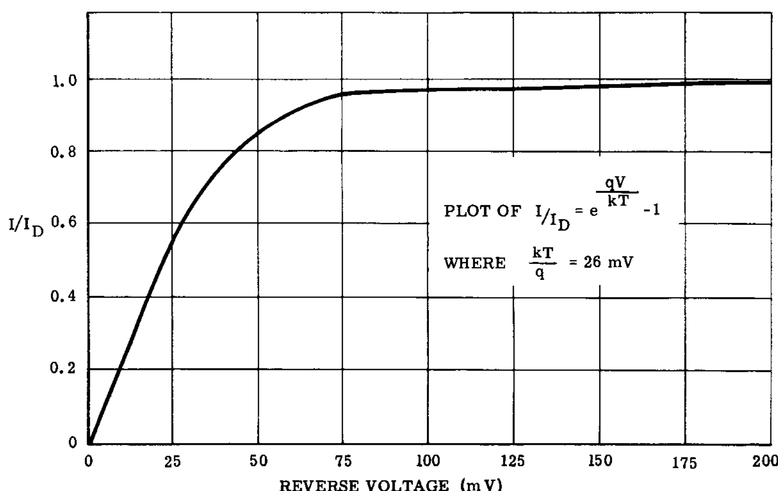


Figure 3-4 — Behavior of the Ideal PN Junction in the Reverse Direction

* See Reference Number 1.

The reverse current due to I_D is only constant with voltage in junctions in which the high resistivity side is very wide. This situation is normally not often encountered even in diodes, as wide high resistivity regions are a source of voltage drops. However, as pointed out in Chapter 1, standard mesa types do have a wide high resistivity collector region, and therefore, have a reverse leakage current due to I_D which is constant with voltage. The width of the high resistivity layer is narrow in most devices, and an analysis of the PN junction shows that the reverse current is proportional to the electrical junction width (W).

$$W = W_O - x_m$$

where W = the effective junction width
W_O = the physical junction width
x_m = the depletion-layer thickness.

The depletion layer thickness (x_m) varies as the square, cube or in general the nth root of applied voltage depending upon the type of junction, i.e., $x_m \propto \sqrt[n]{V}$. Therefore, to account for these effects, write

$$I_D(W) = I_D(W_o)\frac{W_o}{W}.$$

Substituting the previous relations:

$$I = \frac{I_D(W_o)}{1 - \frac{K}{W_o}\sqrt[n]{V}}\left(e^{\frac{qV}{kT}} - 1\right). \qquad (3\text{-}3)$$

where K signifies a transistor constant based upon depletion layer spreading.

This correction factor causes a finite slope to appear on the reverse characteristic as shown in Figure 3-5. This plot was determined with $n = 2$, which applies to a step (alloy) junction. The variation of I_D with voltage is less severe for the graded junction where $n = 3$.

Note as the voltage becomes large, the current increases rapidly. This occurs because the junction width is becoming extremely small. The point where the junction width becomes zero is called the punch-through voltage (V_{PT}).

The diffusion current is the dominant leakage current in germanium devices, particularly at high temperatures. However, in silicon devices the charge generation current, due to impurity ions in the depletion layer, is the dominant temperature sensitive current. It is proportional to the width of the depletion layer and is given by

$$I_G = K_V K_I \sqrt[n]{V} \qquad (3\text{-}4)$$

where: I_G = charge generation current
K_V = an empirical factor which approaches unity for voltages greater than 0.1 volt
K_I = a proportionality constant determined primarily by geometry, resistivity, and the impurities in the depletion layer
V = applied voltage
n = exponent describing depletion layer behavior.

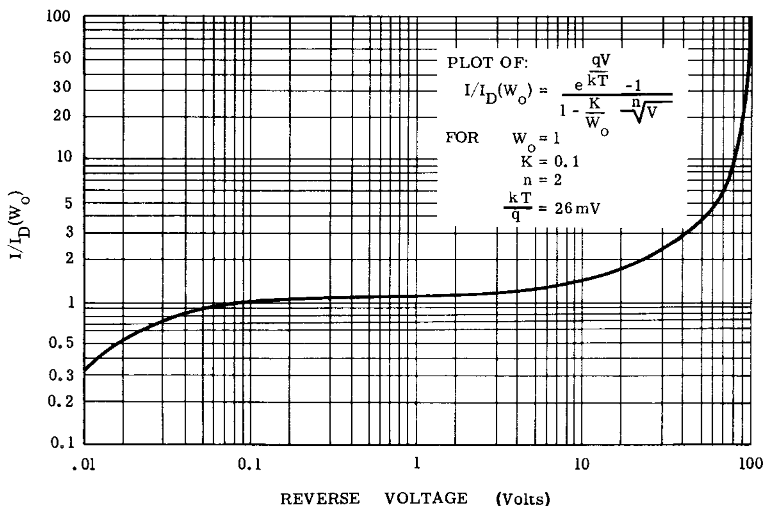


Figure 3-5 — The Effect of Voltage on the Reverse Current Due to Diffusion

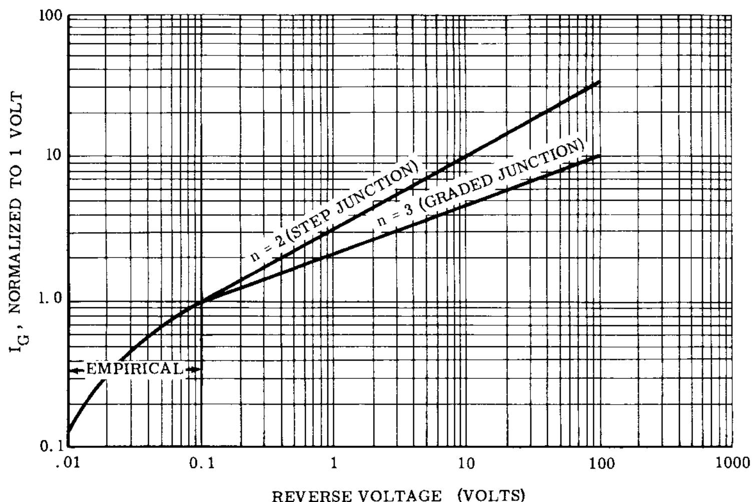


Figure 3-6 — Effect of Reverse Voltage Upon the Reverse Current Due to Charge Generation

Figure 3-6 shows the behavior of the charge generation current with voltage which is quite different from that of the diffusion current shown in Figure 3-5.

These currents (I_D and I_G) are often referred to as the bulk leakage currents as they originate in the body of the semiconductor material. Both I_D and I_G increase rapidly with temperature as shown in Figure 3-7. A rough "rule of thumb" is to describe the bulk current increase as doubling every 10°C.

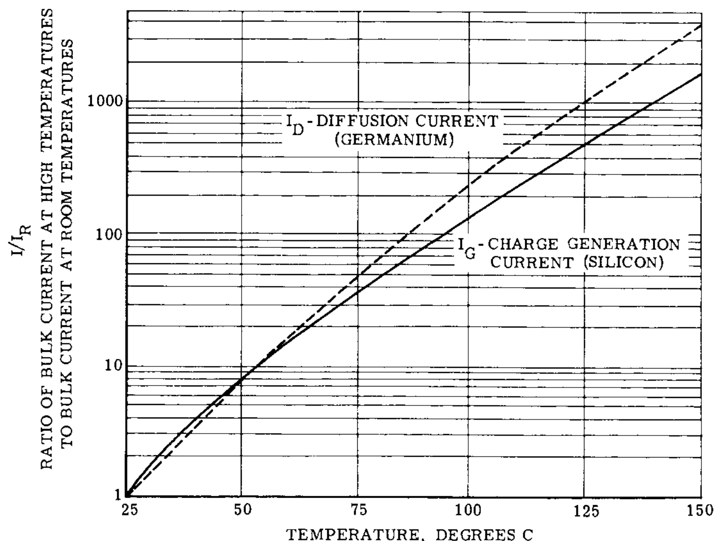


Figure 3-7 — Theoretical Variation of Bulk Reverse Current with Temperature

The remaining contributing factor to reverse current is surface leakage (I_S) which appears as an additive component to I_D and I_G. The surface leakage current may be considered as resulting from a resistance path across the junction. Since the factors contributing to I_S are resistive in nature, the value of I_S is voltage dependent; at high voltages it can add considerably to the total reverse current. Surface leakage may be considered to increase with temperature at about one-half the rate of the bulk leakage current. A graph of I_S vs. temperature based upon this premise is shown in Figure 3-8. This graph is only an estimate of I_S and should be used with discretion because in some transistors, surface leakage may change radically at certain temperatures or voltages.

Generally, at room temperature, the bulk current may be swamped by the surface leakage current which is comparatively independent of temperature. But, at higher temperatures, the diffusion and charge generation currents become dominant because they increase rapidly as temperature increases.

Increases in reverse current due to avalanche multiplication should also be considered. Avalanche effects are considered more completely in the latch-up section. For now, it is sufficient to establish that as the reverse voltage across a junction increases, the carriers are accelerated by the increased field to the point where some of them have sufficient energy to break valence bonds, thereby generating additional electron hole pairs. Given sufficient voltage, the current becomes very high and the junction is said to have "broken down".

As the breakdown voltage is approached, the leakage current begins to increase rapidly. This increase is designated by the multiplication factor (M). The multiplication factor is a property of the bulk current only (I_D or I_G) and does not affect the surface leakage current. Whenever estimating a reverse current

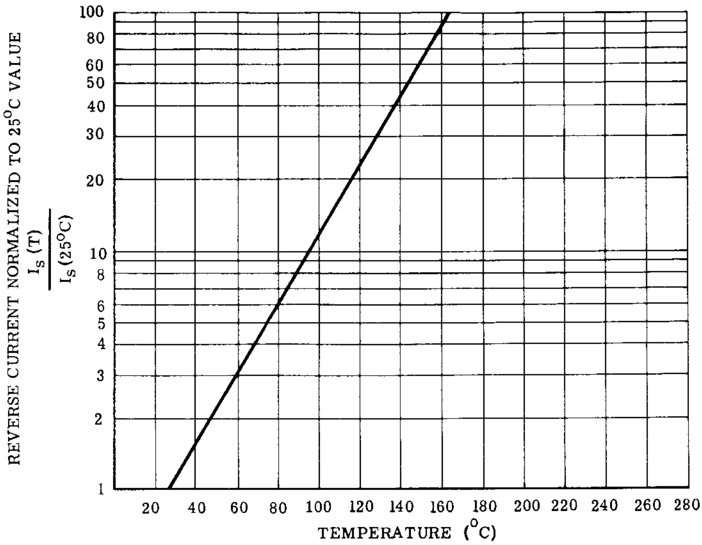


Figure 3-8 — Behavior of Surface Leakage Current with Temperature

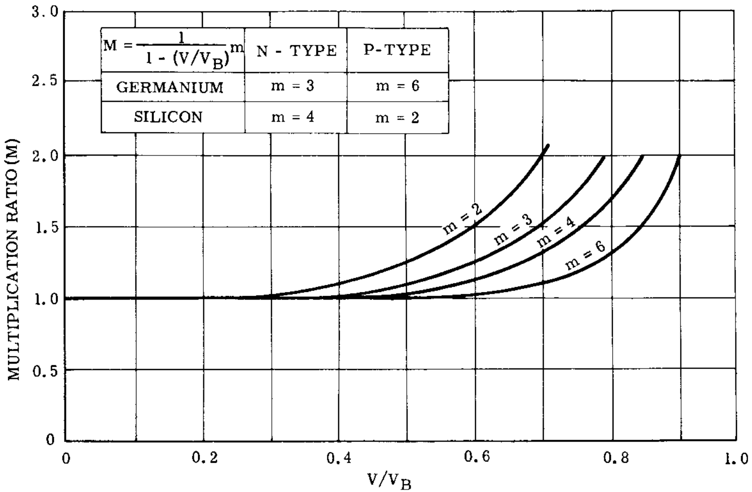


Figure 3-9 — Multiplication Ratio

at a higher voltage from that at which it was specified on the data sheet, this increase in leakage current due to multiplication must be considered. The empirical relationship between V_B and M as developed by Miller[2] is given and values are plotted in Figure 3-9. To use the chart, it is necessary to know whether the avalanche effect is occurring in the N or P region. That is, whether a transistor structure is basically of the alloy type, where breakdown occurs in the base, or of the mesa type, where breakdown occurs in the collector, must be known, as discussed in Chapter 1. When calculating the effect of M, the actual avalanche breakdown voltage (V_B) must also be known. The breakdown voltage (V_B) does vary with temperature. In devices having a breakdown greater than 6 volts, V_B increases with temperature so that for a given voltage, M would effectively decrease with temperature. Therefore, the added complication of V_B varying with temperature need not be considered when estimating I_{CBO} at high temperatures from a low temperature value.

SUMMARY OF FACTORS CONTRIBUTING TO LEAKAGE CURRENTS: Three components of current need to be considered in describing reverse current behavior of a PN junction, (1) the bulk current due to diffusion (I_D) which is dominant in germanium devices and insignificant in silicon devices, (2) the bulk current due to charge generation (I_G) which is noticeable in silicon devices but insignificant in germanium devices, and (3) the surface leakage (I_S) which affects both silicon and germanium devices, but is a more noticeable factor in silicon.

The bulk currents are primarily temperature sensitive, but are affected by voltage, which changes the depletion layer and causes avalanche multiplication. The surface current is primarily voltage sensitive, but generally shows some slight increase with temperature.

3-3 — Predicting Worst-Case I_{CBO}

Utilizing the background of factors contributing to leakage currents established in the preceding paragraphs, several examples* of how to obtain a worst-case value for I_{CBO} from the information provided on the data sheet will be given.

For the first example, consider the worst possible condition where I_{CBO} is specified at only room temperature, as is the case with the 2N962, a PNP germanium epitaxial mesa transistor.

EXAMPLE 1-A:

Specified: $\overline{I}_{CBO}$ @ 6 V and 25°C = 3 μA
$\underline{B}V_{CBO}$ @ 100 μA and 25°C = 12 V.
Obtain: $\overline{I}_{CBO}$ @ 9 V and 65°C.

From the limited information provided, to be absolutely safe, it should be assumed that the I_{CBO} specified in this case is bulk current. Since this is a germanium device the bulk current is a diffusion current.

*In these examples, a desired value or condition will be designated by an (*), while a specified value or condition will be designated by an ('). A bar over a term indicates a maximum value, a bar under a term indicates a minimum value, and a tilde indicates a typical value.

1. From Figure 3-7 read the temperature ratio

$$\frac{I^*_D}{I'_D} \text{ (at T} = 65°\text{C)} = 22.$$

2. From Figure 3-9, obtain the multiplication ratio at the specified voltage and at the desired voltage. The 2N962 is a germanium PNP mesa type. Therefore, breakdown occurs in the P-type collector and the $m = 6$ curve should be used. It is not possible to factor in the depletion layer effects.

   Assume $\underline{B}V_{CBO} = \underline{V}_B$

$$(V'/\underline{V}_B = 6/12 = 0.5 \text{ and } V^*/\underline{V}_B = 9/12 = 0.75$$

$$M' = 1.0 \text{ @ } 6 \text{ V}$$

$$M^* = 1.2 \text{ @ } 9 \text{ V}.$$

3. $\bar{I}^*_{CBO} = \bar{I}'_{CBO} \left(\frac{I^*_D}{I'_D}\right)\left(\frac{M^*}{M'}\right)$

   Therefore:

$$\bar{I}_{CBO} = (3)\ (22) \left(\frac{1.2}{1.0}\right) = 79\ \mu\text{A}.$$

EXAMPLE 1-B: Consider this same transistor with more complete I_{CBO} specifications. Assume that $\bar{I}_{CBO}$ had also been specified at 1 volt as 1 μA. At this low voltage, $M = 1$ and surface effects can be assumed negligible. With this amount of information, the leakage current can be separated into the bulk component I_D, and the surface component I_S. This should provide a more realistic value for I_{CBO} since the surface component is not affected as greatly by temperature as is the bulk component.

1. The bulk component (I_D) at 6 volts would be

   $\bar{I}_D = \bar{I}_{CBO}$ (at low voltage) times M (M at 6 V was determined in the previous example)

   $\bar{I}_D = (1)\ (1.0) = 1.0\ \mu\text{A}.$

2. Therefore, the remainder of the 3 μA would be surface current.

   $\bar{I}'_S = 3 - 1.0 = 2.0\ \mu\text{A}.$

3. The bulk saturation current I_D at the desired condition (9 V and 65°C) can be calculated from

$$\bar{I}^*_D = \bar{I}'_D \left(\frac{I^*_D}{I'_D}\right)\left(\frac{M^*}{M'}\right)$$

$$\bar{I}^*_D = (1.0)\ (22) \left(\frac{1.2}{1.0}\right) = 26.4\ \mu\text{A}.$$

4. The surface leakage current (I_S), at the desired voltage, can be calculated by assuming surface leakage is caused by a linear resistance.

$$\bar{I}^*_S = \bar{I}'_S \left(\frac{V^*}{V'}\right)$$

$$\bar{I}^*_S = (2.0) \left(\frac{9}{6}\right) = 3.0\ \mu\text{A}.$$

5. Modifying this value to include temperature effects:
   From Figure 3-8

$$I^*_S/I'_S = 3.5 \text{ at } 65^\circ C.$$

$$\text{then: } \bar{I}^*_S = (3.0)\ (3.5) = 10.5$$

6. I_{CBO} is equal to the sum of the leakage currents. Therefore,

   $\bar{I}^*_{CBO} = \bar{I}^*_D + \bar{I}^*_S$

   $\bar{I}_{CBO} = 26.4 + 10.5 = 36.9\ \mu A.$

The above approach assumes that the surface component does not increase excessively with voltage or temperature and that BV_{CBO} is the true avalanche breakdown voltage. If desired, the calculations can be made at several conditions, and a maximum curve of I_{CBO} plotted.

EXAMPLE 2: A 2N834 NPN silicon transistor is used in this example.

$$\text{Specified: } \bar{I}_{CBO} = 0.5\ \mu A\ @\ 20V \text{ and } 25^\circ C$$

$$\bar{I}_{CBO} = 30\ \mu A\ @\ 20V \text{ and } 150^\circ C$$

$$\underline{B}V_{CBO} = 40V\ @\ 10\ \mu A \text{ and } 25^\circ C.$$

Obtain: $\bar{I}_{CBO}$ at 100°C and 30 V

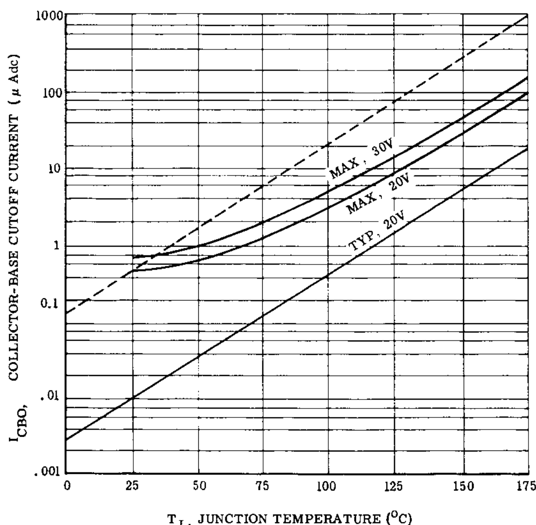


Figure 3-10 — I_{CBO} Behavior for a 2N834 Transistor

This device has enough information provided to accurately predict the behavior of I_{CBO}. Using the I_{CBO} graph from the data sheet (Figure 3-10), the two given maximum I_{CBO} points are plotted and a curve is drawn connecting the two points. This provides a maximum curve at 20 volts. To construct a 30 volt curve, the avalanche,depletion layer and surface effects must be included. (Note: if the high temperature point had not been specified, it would be necessary to assume

that the room temperature value of 0.5 μA is all charge generation currents which would result in a curve parallel to the typical curve. This is shown by the dashed curve which would have resulted in an I_{CBO} of ten times the actual specified value at 150° C).

1. Obtain the avalanche multiplication ratio from Figure 3-9. Since this is an NPN silicon mesa structure, breakdown occurs in the N-type collector and the m = 4 curve is used.

$$\frac{V^*}{\underline{V}_B} = \frac{30}{40} = 0.75 \quad \text{and} \quad \frac{V'}{\underline{V}_B} = \frac{20}{40} = .5$$

$$\therefore M^* = 1.5 \qquad \therefore M' = 1.05$$

The avalanche multiplication ratio is

$$\frac{M^*}{M'} = \frac{1.5}{1.05} = 1.43$$

2. As previously discussed, the depletion layer effect must be considered for a silicon device. The depletion layer factor is (using square root behavior to be safe)

$$\sqrt{\frac{V^*}{V'}} = \sqrt{\frac{30}{20}} = 1.22.$$

3. A leakage current specified at high temperature is primarily composed of the bulk current. Therefore, to predict I_{CBO} at 30 V and 150°C, it can be assumed that the leakage current specified at 150°C is entirely bulk current. To determine the high temperature point on the maximum I_{CBO} curve at 30 volts, multiply the given I_{CBO} conditions by the correction factors calculated in steps 1 and 2.

$$\bar{I}^*_{CBO} = \bar{I}'_{CBO}\left(\frac{M^*}{M'}\right)\left(\sqrt{\frac{V^*}{V'}}\right)$$

$$\bar{I}_{CBO} = (30)\,(1.43)\,(1.22) = 52\ \mu\text{A at } 150°\text{C}.$$

4. The I_{CBO} specification at room temperature is almost entirely surface leakage which may be taken directly proportional to voltage. To calculate the surface component at room temperature, observe from Figure 3-7 that I_G (150°C) = 1500 I_G (25°C).

Thus, $I_G\ (25°\text{C}) = \dfrac{30\ \mu\text{A}}{1500} = 20\ \text{nA}.$

Even when multiplied by the factors due to avalanche multiplication and the depletion layer, I_G is certainly small compared to the specified maximum of 500 nA. Therefore I_{CBO} equals the surface component at 25° C which can be calculated as

$$\bar{I}^*_S\ (25°\text{C}) = \bar{I}'_S\ (25°\text{C})\ \frac{V^*}{V'} = 500\ \frac{30}{20} = 750\ \text{nA}$$

Using the points calculated in steps 3 & 4, construct the curve of I_{CBO} at 30 volts as shown in Figure 3-10.

The required $\bar{I}_{CBO}$ conditions for this example are 100°C and 30-volts. From the 30-volt maximum curve just drawn $\bar{I}_{CBO}$ is equal to 5 μA.

EXAMPLE 3: a 2N2222 NPN "Star" silicon transistor is used.

Specified: $\bar{I}_{CBO} = .01\ \mu A$ @ 50 V and 25°C
$\bar{I}_{CBO} = 10\ \mu A$ @ 50 V and 150°C
$\underline{B}V_{CBO} = 60$ V @ 10μA and 25°C
A typical curve of I_{CBO} @ 50-volts is given.

Obtain: $\bar{I}_{CBO}$ @ 10 V and 100°C.

Since an I_{CBO} graph is provided (Figure 3-11), these two given maximum I_{CBO} points are used to draw a maximum curve. Compared to the typical curve provided, the high temperature maximum point appears to be unreasonably high. This could be due to the manufacturer allowing an extra guardband margin, or it could be due to abnormal surface leakage. However, the straight line curve through these two points does provide safe working values. Since the required I_{CBO} is at a lower voltage than that specified, it is conservative to neglect the effects of avalanche multiplication and surface leakage. From the constructed curve, I_{CBO} at 100°C and 50 V = 800 nanoamps. Since the desired I_{CBO} is at a voltage lower than the specified I_{CBO}, a cube root function for depletion layer correction will give a conservative value. Therefore, $\bar{I}_{CBO}$ at 100°C and 10 V (the desired condition) is given by

$$\bar{I}^*_{CBO} = \sqrt[3]{\frac{V^*}{V'}}\ \bar{I}'_{CBO} = \sqrt[3]{\frac{10}{50}}\ (800) = 470 \text{ nanoamps.}$$

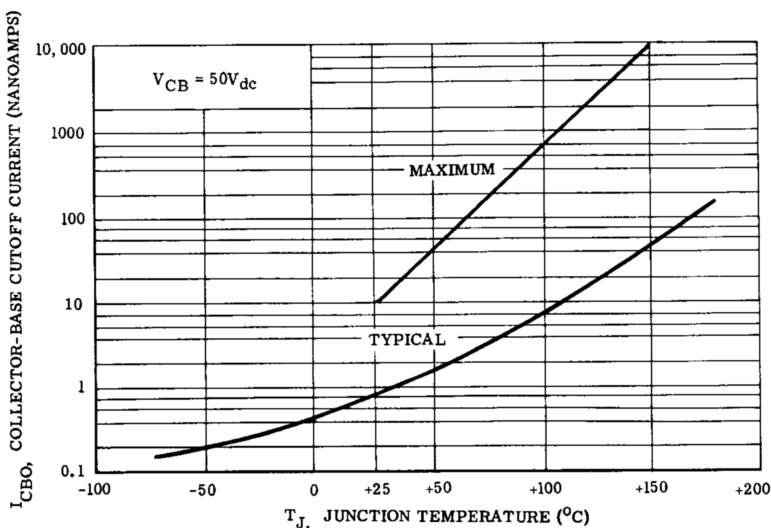


Figure 3-11 — I_{CBO} Behavior for a 2N2222 Transistor

The three examples have illustrated how to estimate a maximum working value for I_{CBO}, and should cover a broad range of conditions for which a maximum value of I_{CBO} must be obtained.

3-4 — Predicting Worst-Case I_{EBO}

In general, the same principles that were used to estimate the worst-case I_{CBO} also apply to estimating I_{EBO}. However, since I_{EBO} is seldom specified on the data sheet, it is considerably more difficult to determine I_{EBO}. Transistor theory will be of some help in this instance. To estimate the leakage current due to the diffusion component (germanium), use the following relationship which was developed by Ebers & Moll[3].

$$\alpha_I I_{CD} = \alpha_N I_{ED} \qquad (3\text{-}4)$$

where I_{CD} = collector diffusion current

I_{ED} = emitter diffusion current.

The forward current gain (α_N) can be considered as unity and the following values for α_I are useful as a "rule of thumb":

0.8 — uniform base alloy
0.5 — diffused base alloy
0.3 — diffused base mesa.

As previously discussed, the charge generation current (I_G) in silicon devices is proportional to the width of the depletion layer. In an alloy type, the base sustains the depletion layer caused by reverse bias on either junction.

$$\therefore I_{EG} = I_{CG} \text{ when } V_{CB} = V_{EB}$$

where I_{EG} = emitter charge generation current

I_{CG} = collector charge generation current.

In mesa types, however, the depletion layers extend into the collector for reverse collector voltage and into the base for reverse emitter voltages. Since the collector resistivity is high compared to that of the base, the depletion layer is much wider in the collector than in the base. A useful "rule of thumb" for silicon mesa devices is:

$$I_{EG} \approx 0.1\, I_{CG} \text{ when } V_{CB} = V_{EB}.$$

Since, at high reverse voltages, the depletion layer and avalanche effects influence the bulk current, these equations should only be used at low voltages. The effect of depletion layer can be included for high emitter voltages by using the square root behavior, regardless of the type of device.

In order to include the effects of avalanche multiplication, the avalanche breakdown voltage must be known. For uniform base alloy types, the emitter and collector breakdown voltages are, of course, identical. Usually, in high frequency mesa types, the breakdown voltage of the emitter is between 6 and 10 volts.

Surface leakage is often a problem with diffused base transistors having an alloyed emitter. To estimate the I_{EBO} surface leakage current, determine a very pessimistic value for the base-emitter junction surface resistance from

$$r_s = \frac{BV_{EBO}}{I_E}$$

where I_E indicates the current at which BV_{EBO} was measured.

With a value for r_s, the surface leakage currents at various base-emitter voltages can be estimated from $I_S = V_{OB}/r_S$. This surface leakage current can then be modified for temperature effects as described for I_{CBO}.

EXAMPLE 4: Determine I_{EBO} for a 2N962 germanium transistor at 0.5V and 65°C. This is the transistor that was used in example 1-B for calculating worst-case I_{CBO} at 9.0 V and 65°C where it was assumed that I_{CBO} was specified at 1 volt.

The data sheet specifies: $\underline{B}V_{EBO} = 1.25$ V @ 100 μA.

1. Obtain I_{CD} at a low voltage where avalanche and depletion layer effects are negligible.

   This device had an I_{CBO} specified as 1 μA at 1 volt and 25°C. An I_{CBO} specified at this low voltage can be considered as bulk leakage current. Therefore, $I'_{CD} \cong 1\ \mu A$

2. Obtain I_{CD} @ 65°C.

   The temperature correction ratio was determined from Figure 3-7 to be 22 at 65°C.

   Therefore, $\bar{I}^*_{CD}$ at 65°C $= 1\ (22) = 22\ \mu A$.

3. Determine I_{ED} using the appropriate relationship

   $$I_{ED} = I_{CD} \frac{\alpha_I}{\alpha_N}$$

   $\alpha_I \approx 0.3$ for mesa transistors, α_N may be taken as unity

   $\therefore I^*_{ED} = (22)\ (0.3) = 6.6\ \mu A$

4. Determine the emitter-base junction surface resistance.

   $$r_s = \frac{BV_{EBO}}{I_E} = \frac{1.25}{100\ \mu A} = 12.5\ K$$

5. From Figure 3-8 determine the temperature correction factor for the surface leakage current.

   $$\frac{I^*_S}{I'_S} = 3.5 \text{ at } 65°C$$

6. Determine $\bar{I}^*_{ES}$.

   $$\bar{I}^*_{ES} = \frac{V_{OB}}{r_s} \times \frac{I^*_S}{I'_S} = \frac{0.5}{12500}(3.5) = 140\ \mu A$$

7. Determine $\bar{I}_{EBO}$.

   $\bar{I}^*_{EBO} = \bar{I}^*_{ED} + \bar{I}^*_{ES}$

   $\bar{I}_{EBO} = 6.6 + 140 = 146.6\ \mu A.$

The unrefined method used to estimate r_s results in large surface leakage current, but unless more data is available, this method must be used in order to be safe.

3-5 — Transistor Cut-Off Behavior With Both Junctions Connected

When both junctions are reverse biased, it would normally be expected that the reverse currents of both junctions would simply add in the base terminal. This is not completely true, because the current due to diffusion is affected by transistor action. Therefore, before calculating the base leakage current (I_{BL}), an analysis of the ideal transistor with both junctions reverse biased is discussed in this section. It will be helpful to refer to Figure 3-1 as it indicates the sign convention used.

The leakage currents flowing in the collector and emitter circuit legs have been thoroughly analyzed by Ebers and Moll* for an *ideal transistor* which follows the diffusion equations and which has no surface leakage. They found that:

$$\alpha_I I_{CD} = \alpha_N I_{ED}. \qquad (3\text{-}5)$$

and with both junctions reverse biased:

$$I_C = \frac{I_{CD}(1 - \alpha_I)}{1 - \alpha_I \alpha_N} \qquad (3\text{-}6a)$$

$$I_E = \frac{I_{ED}(1 - \alpha_N)}{1 - \alpha_I \alpha_N}. \qquad (3\text{-}6b)$$

Where:
- I_C = Collector current
- I_E = Emitter current
- I_{CD} = Bulk saturation leakage current, due to diffusion, of the collector-base junction with the emitter open
- I_{ED} = Bulk saturation leakage current, due to diffusion of the emitter-base junction with the collector open
- α_I = Inverted current gain (I_E/I_C) (collector used as emitter)
- α_N = Normal current gain (I_C/I_E).

By combining equations 3-6a and 3-6b the base current (I_B) can be expressed by:

$$I_B = \frac{I_{CD}(1 - \alpha_I) + I_{ED}(1 - \alpha_N)}{1 - \alpha_N \alpha_I} \qquad (3\text{-}7)$$

The behavior of I_C and I_B with changes in α_N and α_I is complicated to visualize from the equations and, therefore, is shown in Figure 3-12. Equations 3-5, 3-6a and 3-7 have been combined and arranged so that I_C and I_B are expressed in relationship to a normalized I_{CD}.

From Figure 3-12, it is evident that for any reasonable β, I_B is minutely greater than I_{CD}, regardless of α_I. However, I_C is always less than I_{CD} and may be significantly less when β is low and α_I high.

From the preceding discussion, the following can be stated:

$$I_{CD} + I_{ED} > I_B \geqq I_{CD} > I_C > I_E. \qquad (3\text{-}8)$$

* See Reference Number 3 and Appendix I.

The above statement is based upon an analysis of only the diffusion currents of devices and is true only when sufficient reverse bias is present on both junctions. It also assumes that $\alpha_N > \alpha_I$, a condition always fulfilled.

The effect of changes in reverse bias voltage upon leakage currents of a germanium transistor is illustrated in Figure 3-2. Several important points on the figure can be found from Ebers and Moll's work. It can be shown that when $V_{BE} = 0$:

$$I_C = \frac{I_{CD}}{1 - \alpha_N \alpha_I} \tag{3-9}$$

$$I_E = - \frac{\alpha_I I_{CD}}{1 - \alpha_N \alpha_I}. \tag{3-10}$$

The equation for I_E has a negative sign indicating that injection is taking place. That is, the net emitter current is opposite in sign to the leakage current. An example is informative. Take: $\alpha_N = 0.8$, $\alpha_I = 0.4$, $I_{CD} = 1\ \mu A$ and $I_{ED} = 0.5\ \mu A$. I_C calculates to be 1.47 μA and $I_E = -.59\ \mu A$. The base current is 0.88 μA. With $V_{BE} = 0$, I_C is greater than I_{CD} and I_B is less than I_{CD}. Note this condition on Figure 3-2. I_C would become higher as α_I and α_N approach unity.

At some small reverse voltage (ϕ_{TR}) between 0 and V_{TR}, $I_B = I_{CD}$. This voltage can be found by setting $I_B = I_{CD}$ and solving for V_{BE} in Moll's general equations. This theoretical voltage is often referred to as the reverse threshold voltage and is given by

$$\phi_{TR} = \frac{kT}{q} \ln \frac{1}{(1 - \alpha_N)} \tag{3-11a}$$

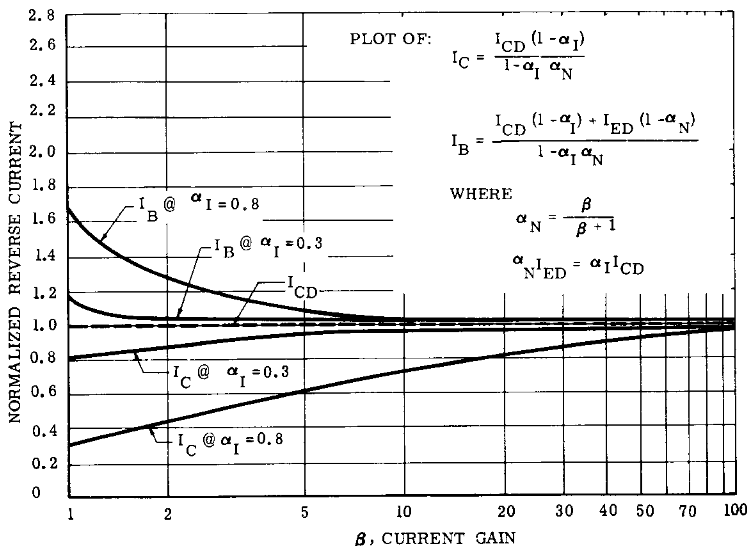


Figure 3-12 — Effect of Gain on Reverse Current

or in terms of β

$$\phi_{TR} = \frac{kT}{q} \ln (1 + \beta). \tag{3-11b}$$

When $V_{BE} = \phi_{TR}$, $I_C = I_B = I_{CD}$ and $I_E = 0$. The voltage ϕ_{TR} calculates to be in the range of 0.1 to 0.2 volt for practical values of β. According to equation 3-11, ϕ_{TR} increases 86 μV/°C which is a negligible amount. V_{TR}, the voltage which reduces I_C to its minimum value, is a slightly larger reverse voltage than ϕ_{TR} at high temperatures. At low temperatures the diffusion current is masked by the effects of the surface and charge generation currents, which effectively moves V_{TR} in the forward direction.

One further condition is of interest, the condition when $I_B = 0$, which occurs when the base circuit is open. Under these conditions

$$V_{BE} = \frac{kT}{q} \ln \left[1 + \frac{\alpha_N (1 - \alpha_I)}{\alpha_I (1 - \alpha_N)} \right] \tag{3-12}$$

and

$$I_C = \frac{I_{CD}}{1 - \alpha_N} \tag{3-13a}$$

or in terms of β

$$I_C = I_{CD} (\beta + 1). \tag{3-13b}$$

This condition could hardly be considered off and I_C will increase very rapidly with temperature and voltage. This point is near, and sometimes it may be convenient to consider it as, the threshold of forward conduction, (V_{TF}). A plot of equation 3-12 is shown in Figure 3-13.

Since transistor action does not influence the charge generation process, the emitter and collector charge generation currents should simply add, at the base terminal. Measurements have indicated, at temperatures where I_S is insignificant, that:

$$I_C = I_{CG} = I_{CBO} \text{ and } I_B = I_{EG} + I_{CG}.$$

Figure 3-14a shows that, for a silicon transistor, the collector current does not begin to increase significantly until the forward bias exceeds 0.1 volt. This occurs because the injection current, which is determined by I_D, is extremely small at this small forward potential and is masked by I_G. Therefore, for silicon devices, V_{TR} can be considered as 0. However, at very high temperatures, the diffusion current can be significant because the diffusion current increases at a faster rate, with temperature, than does the charge generation current. (See Figure 3-7.) Therefore, at temperatures near 200°C, a minimum reverse bias of 0.1 to 0.2 volt may be required with silicon transistors to keep leakage current at a minimum as Figure 3-14b shows.

Surface leakage may be regarded as a resistor connected externally to the transistor. As such, it simply adds a component of current and is considered in much the same way as the charge generation current. Both the charge generation and the surface currents mask the diffusion current and give an apparent shift of the threshold voltage in the direction of forward bias which becomes particularly evident at low temperatures.

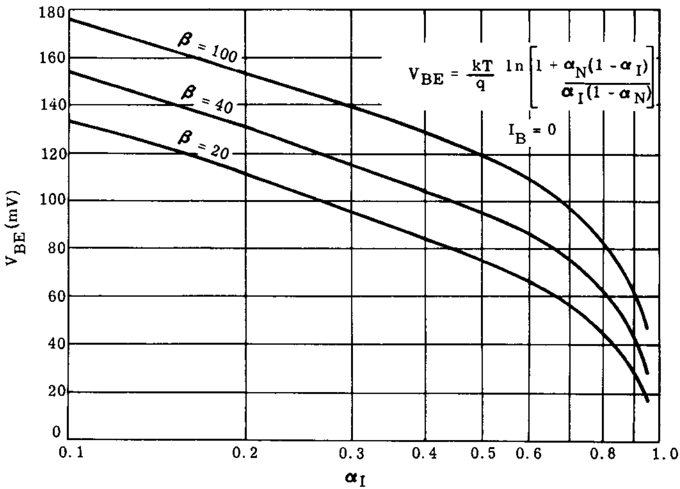


Figure 3-13 — Base-Emitter Voltage with Base Open

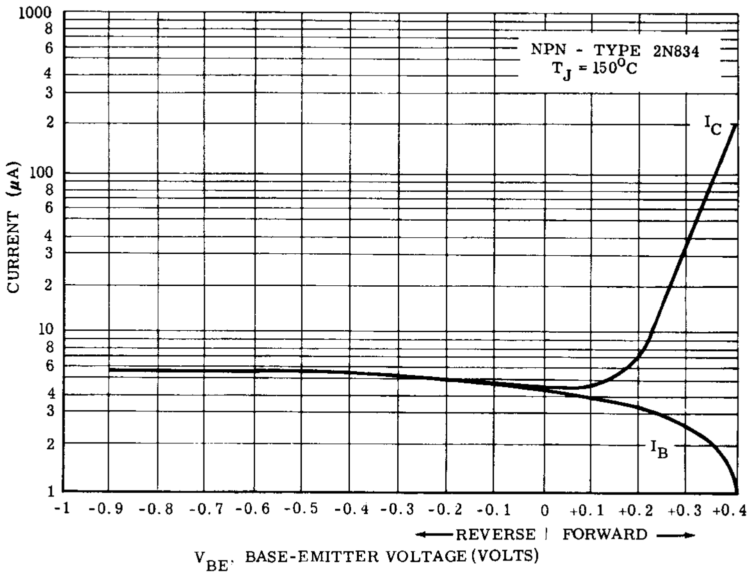


Figure 3-14a — Effect of Base Voltage Upon Transistor Current in the Cut-Off Region of a Silicon Transistor

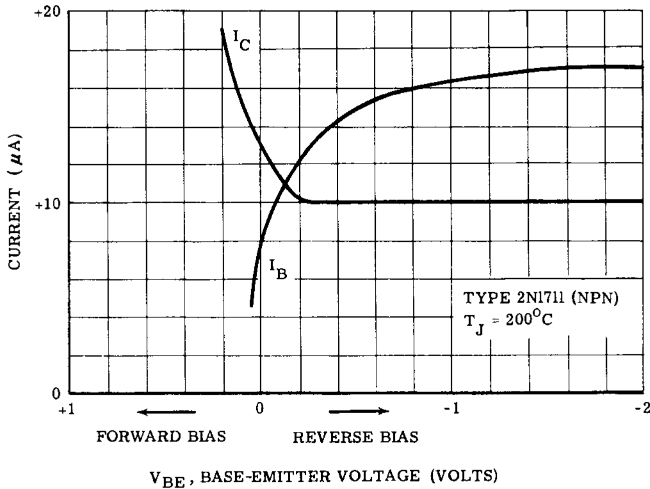


Figure 3-14b — Effect of Base Voltage Upon Transistor Current in the Cut-Off Region of a Silicon Transistor

3-6 — Estimating the Base Leakage Current (I_{BL})

From the foregoing discussion, it may be stated that a transistor switch should normally be operated with a few tenths of a volt reverse bias, from base to emitter.

With both junctions reverse biased, it has been shown that $I_B \approx I_{CD}$ when diffusion currents only are considered. The effects of charge generation and surface leakage are additive so that the base current is approximately equal to:

$$I_{BL} \cong I_{CD} + I_{CG} + I_{CS} + I_{EG} + I_{ES} \qquad (3\text{-}14a)$$

or

$$I_{BL} \cong I_{CBO} + I_{ES} + I_{EG}. \qquad (3\text{-}14b)$$

EXAMPLE 5:

Determine I_{BL} @ 65°C, $V_{CE} = 9V$, $V_{OB} = 0.5V$; for the 2N962 transistor.

$I_{CBO} = 36.9\ \mu A$ @ 65°C and $V_{CE} = 9V$ (From Example 1-B)

$I_{ES} = 140\ \mu A$ @ 65°C and $V_{OB} = 0.5V$ (surface component of I_{EBO} from Example 4)

$I_{EG} \cong 0$ (for germanium devices)

$\therefore I_{BL} = 36.9 + 140 = 176.9\ \mu A.$

3-7 — Relationship of Voltage Breakdown and Switching Load Lines

The conventional methods of specifying transistor voltage breakdown do not provide the designer very much data to actually determine that a switching load line will remain free of the avalanche voltage breakdown region. Although the BV_{CBO} specification does provide a maximum collector voltage limit, there is no assurance that the voltage-current excursions of a switch being turned off to BV_{CBO} will not enter the avalanche region. Figure 3-15 is a comprehensive transistor characteristic graph showing the avalanche region characteristics in addition to the normal operating characteristics. The breakdown characteristics switch back from BV_{CBO} towards BV_{CEO}.

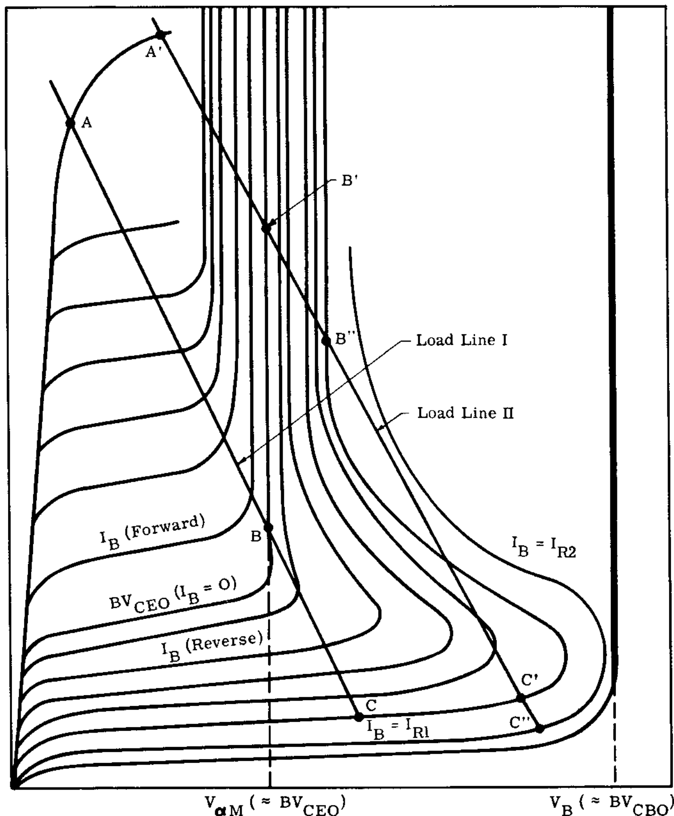


Figure 3-15 — Comprehensive Transistor Characteristics

If the load line of a switching circuit intercepts any of these curves in the breakdown region, it is possible that a stable operating point will result in the breakdown region. Examine the turnoff of a transistor switch with load line I and an on point at A. As drive is reduced, the operating point moves down the

load line, each intersecting point corresponding to an intersection of the load line with a collector current line determined by the base current at that instant. When $I_B = 0$, operation is at Point B, in the avalanche region. However, as reverse base current is increased to the final value I_{R1}, operation moves to Point C in the cutoff region. Examine load line II with an on condition point of A′. As current is reduced to zero, operation moves to B′ in the avalanche region. However, in this case the application of reverse bias which has the same final value of I_{R1} only moves operation to B″ in the avalanche region instead of to the desired off point, C′. The transistor is then in a condition which is called latch-up as it is locked at a stable on point in the avalanche region. The collector voltage has not reached the desired off value and the collector current is much greater than I_{CL}. Latch-up not only causes circuit malfunction, but could result in damage to the transistor, if the product of voltage and current at the point of latch-up is high enough to exceed the power ratings of the transistor. To allow the transistor to turn off, the reverse base current is increased to I_{R2}. There is now only one intersection of the load line with the base current curve and it is at point C″.

Since breakdown voltage specifications alone are not sufficient to forecast "latch-up" conditions, charts similar to the one shown in Figure 3-16 have been devised to provide a method of checking switching circuit load lines. This chart has three discrete areas indicated (1) a safe or latch-free load line area, (2) a conditionally safe area of operation, and (3) a forbidden or latch-up area.

The part of several representative load lines during turn off are shown on Figure 3-16 which applies to a 2N964A transistor. Load line "A" is a resistive load line and it lies entirely within the latch-free load line area. Load line "B" is also a resistive load, but it transverses the conditionally safe area, and could cause trouble if the fall time of the output pulse exceeds 15 nanoseconds. Load line "C" is a capacitive load line (collector current leads the collector voltage).

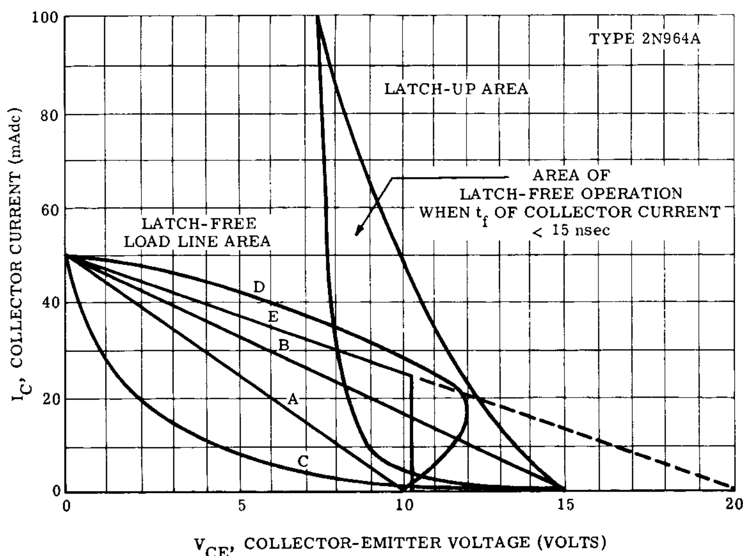


Figure 3-16 — Area of Permissible Load Loci Chart with Representative Load Lines

Capacitive load lines, generally, are latch-free since they have a shape such that they slip under the conditional area. Increasing turn off current (I_{B2}) increases the capacitive effect.

Load line "D" is an inductive load line because voltage leads the current. It also exhibits an inductive "kick". To be trouble-free, the fall time must be less than 15 nanoseconds. However, if the fall time is slower than 15 nanoseconds, latch-up conditions would depend upon the resistive component of the load line. For example, if the resistive component is load line "A", there will be a temporary latch-up condition until the energy in the inductance is dissipated. This temporary latch-up would result in an abnormally long fall time and considerable peak power dissipation. However, the transistor would eventually turn off.

The remaining load line "E" is the load line of an output circuit which uses a clamp diode to establish the off level. Since this load line lies within the conditionally safe area, the fall time must be less than 15 nanoseconds.

If an Area of Permissible Load Loci Chart is not supplied, one can be constructed by conducting tests on some low voltage transistor samples. Generally, it is necessary for the circuit designer to check individual circuits, using low limit BV_{CEO} samples to determine if latch-up can occur. To provide a more complete picture of the relationship of transistor voltage breakdown ratings and latch-up, the following section is a brief review of voltage breakdown in transistors.

3-8 — Avalanche Breakdown Theory

Avalanche breakdown occurs when the reverse bias applied to a semiconductor junction produces an electric field in excess of approximately 10^5 volts per centimeter. Under this condition, carriers are accelerated sufficiently to excite additional carriers by impact ionization with the atoms in the crystal lattice. Since this occurs at high fields and therefore high carrier velocities, recombination can be neglected, and the effect is regenerative.

This process can be described by the multiplication factor previously discussed (M) which Miller[2] has shown to be approximately

$$M = \frac{1}{1 - \left(\frac{V}{V_B}\right)^m} \qquad (3\text{-}14)$$

Where: $V =$ the applied voltage

$V_B =$ the avalanche breakdown voltage (See Note 1)

$m =$ empirical determined constant

A graph of this relationship is shown in Figure 3-9, where M is plotted as a function of the voltage ratio V/V_B. As $V \rightarrow V_B$, M increases without limit.

Note 1

The actual collector-base breakdown voltage is defined as V_B and the collector-emitter breakdown as $V_{\alpha M}$. These terms refer to the true breakdown voltages, i.e., to a voltage which will cause an infinite, or nearly so, increase in current if exceeded. It is common to measure V_B in a circuit with the emitter open and with a constant current forced through the junction. This is called a BV_{CBO} test. Note, however, from Figure 3-15, that the current used for the test must be large enough to put the operating point over the knee, or the true breakdown voltage, V_B will not be measured. The same general comment is true regarding BV_{CEO} and $V_{\alpha M}$. Both BV_{CBO} and BV_{CEO} represent a locus of points, while V_B and $V_{\alpha M}$ represent a definite breakdown voltage.

Tests designated BV_{CEX} and BV_{CER} are also used. The symbol x indicates some reverse bias and resistance are used from base to emitter (which must be specified) while R indicates a resistance alone is used.

When a transistor is operating in the common emitter connection, the effects of M become serious at relatively low voltages. This is because M effectively multiplies the current gain, α. Under conditions of zero or forward base current, the collector current becomes

$$I_C = \frac{M}{1 - \alpha M} (\alpha I_B + I_{CL}). \qquad (3\text{-}15)$$

When I_B is zero, the collector current is larger than the reverse current of the diode alone, by a factor somewhat larger than that of the common emitter current gain. As αM approaches unity I_C increases without limit. Using equation 3-14 and solving for the collector voltage ($V_{\alpha M}$) where $\alpha M = 1$, it is found that:

$$V_{\alpha M} = V_B (1 - \alpha)^{1/m}. \qquad (3\text{-}16a)$$

In terms of common emitter gain:

$$V_{\alpha M} = V_B/(\beta + 1)^{1/m}. \qquad (3\text{-}16b)$$

The significance of this equation is shown in Figure 3-17. In some cases, $V_{\alpha M}$ is a small fraction of the collector diode breakdown V_B.

When αM is greater than unity, base current must reverse in order to hold I_C, as given by equation 3-15, to a constant value. Thus, a family of collector characteristics exists for reverse values of base current, as previously shown on Figure 3-15. Therefore, voltages in excess of $V_{\alpha M}$ definitely place operation in the avalanche region of the transistor's characteristics. However, operation in the avalanche region will not cause latch-up, unless a stable operating point is found on the avalanche characteristics.

Switching speed, as well as the amount of turn-off bias and the shape of the load line, has been observed to affect latch-up. The picture thus becomes complex

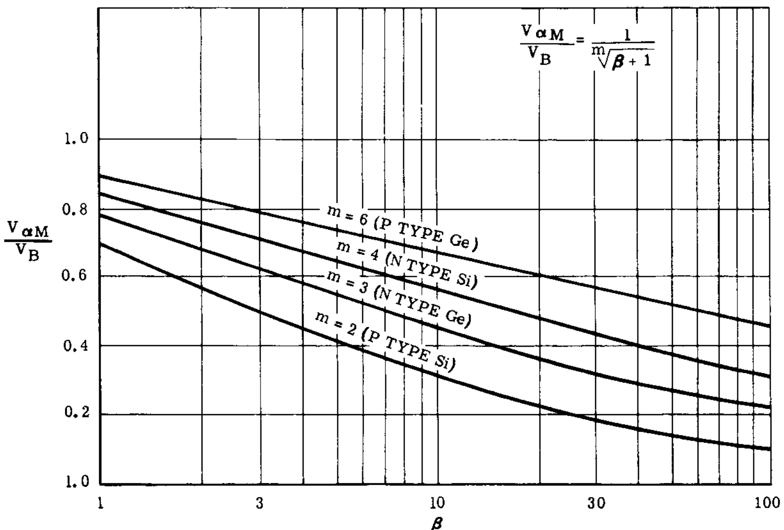


Figure 3-17 — Relationship of V_α to V_B

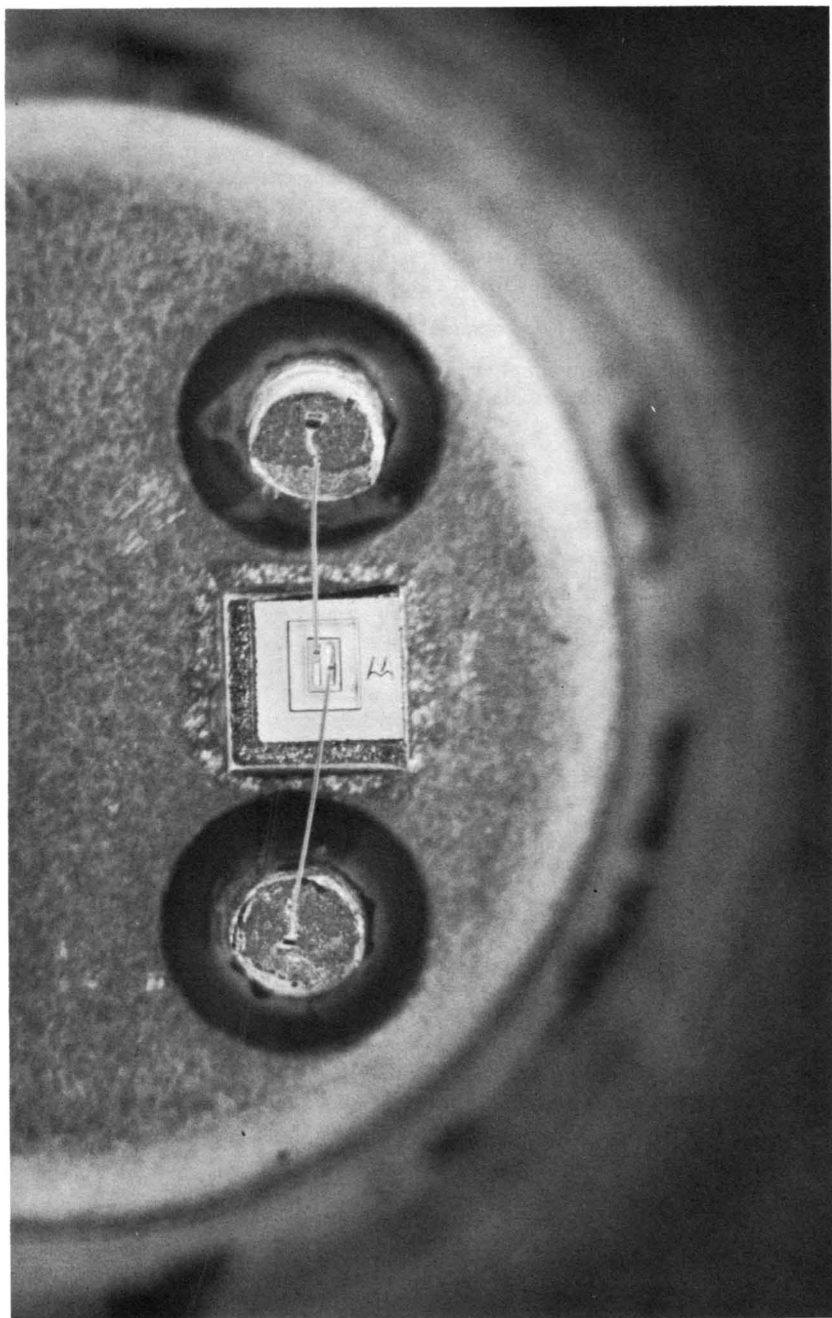
but a few ground rules can be given which will be of assistance particularly if a latch-up chart is available.

High current, slow speed switching over resistive or inductive load lines should not allow v_{CE} to exceed $V_{\alpha M}$. Fast low current switching, particularly in circuits having capacitive load lines, can often safely allow voltages up to V_B to be used. It would be well to check any circuit where v_{CE} exceeds $V_{\alpha M}$ for the possibility of latch-up using transistor samples with $V_{\alpha M}$ at the low limit. The circuit should be arranged for worst case circuit conditions which are:

1) Highest temperature
2) Minimum turn-off drive
3) Maximum collector supply voltage
4) Loads which produce a load line that is most inductive in shape.

REFERENCES

1. Phillips, A. B., *"Transistor Engineering"*, Chapter 6, McGraw Hill Book Company, Inc., New York, New York, 1962.
2. Miller, S. L., *"Avalanche Breakdown in Germanium"*, Physical Review, Vol. 99, pp. 1234-1241, August, 1955.
3. Ebers, J. J. and Moll, J. L., *"Large-signal Behavior of Junction Transistors"*, Proceedings of the IRE, Vol. 42, pp. 1761-1772, December, 1954.



MICROPHOTOGRAPH OF MESA TRANSISTOR JUNCTION AREA

CHAPTER 4

Transistor Characteristics Influencing "on" Condition Design

In Chapter 2, the operating conditions, in terms of collector voltage and collector current, for the three modes of transistor switching operation were outlined. A brief review of these operating conditions (Figure 2-1) reveals that the primary difference between the three modes of operation is in the on condition. In saturated-mode circuits, the transistor is driven into the saturation region. In current-mode circuits (and other non-saturated circuits) the on operating point is generally close to, but not within, the saturation region. Therefore for saturated-mode and current-mode circuit design, the transistor's characteristics in or near the saturation region must be thoroughly defined.

The on condition for avalanche mode circuits is considerably different from the previous two modes of operation and therefore must be treated as a special case. The avalanche-mode on condition is discussed completely in Chapter 9.

In this chapter, the circuit design requirements as related to the saturation region are discussed. The first section of this chapter covers the on condition characteristics of the transistor and introduces several new graphs which explicitly define the transistor characteristics in saturation. Then, the theoretical principles upon which these graphs are based are discussed. Finally, since the concept of these graphs is new, a procedure is given for developing these graphs from data available on most data sheets.

The on region of a saturated switch is characterized by the operating conditions in which the voltage across the transistor is at a minimum and the collector-current is at a maximum. The voltage and current conditions associated with a saturated transistor are illustrated in Figure 4-1 a . (Any saturated mode switching circuit can be reduced to this simple equivalent circuit). As shown, the transistor is turned on by a current I_B supplied to its base causing a voltage SV_{CE} (collector saturation voltage) to appear across the output terminals of the transistor. The voltage SV_{CE} is important because it determines power dissipation and sets a level which is coupled to succeeding stages and must be considered in the design of the following stage's input network. This voltage is determined by the transistor characteristics in the saturation region which are primarily functions of the collector-current (I_C) and the turn-on current (I_B).

Since switching circuits operating in the saturated mode are concerned mainly with the transferring of information contained within distinct voltage levels, it is usually desired to maintain SV_{CE} below a specified limit at a particular collector current. To maintain this specified limit, a minimum drive current (I_B) must be supplied under the anticipated worst-case operating conditions.

Another voltage (V_{BE}) is developed across the base-emitter diode of an on transistor switch. This voltage is in series with V_{IN} and opposes I_B; ie: ($I_B = \frac{V_{IN} - V_{BE}}{R_K}$). Therefore, the worst-case condition of V_{BE} must be considered when determining the minimum base drive current (I_B).

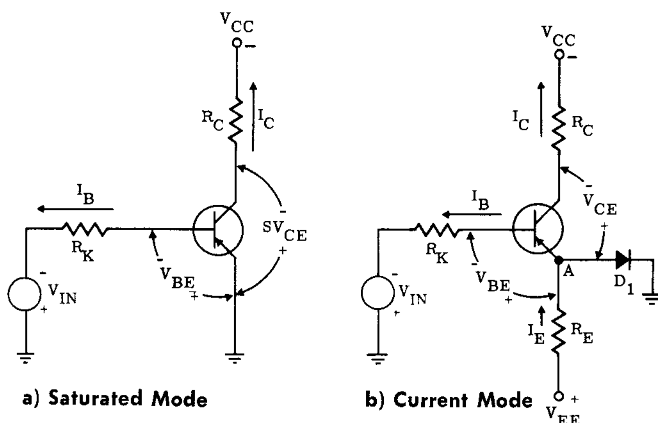


Figure 4-1 — Basic Switching Circuits

For current mode operation, the voltage drops and currents associated with the on condition are illustrated in Figure 4-1b. When the current mode switch is turned on, the diode (D_1) must be non-conducting. Therefore, in this circuit the voltage at Point "A" must be negative with respect to ground. This requirement is met when $V_{IN} \geqq V_{BE} + R_K I_B$. Thus, in current mode design, as in saturated mode design, the worst-case value of V_{BE} is an important characteristic.

In current mode circuits, the collector-emitter voltage for the on condition can be optimized for either maximum speed or for minimum power dissipation. If speed is the primary criterion, the collector-emitter voltage is high and, therefore, is not generally near the saturation region. To minimize power dissipation, however, the collector-emitter voltage must be low, which dictates operation near the saturation region, but operation should never enter the saturation region if the penalty of storage time is to be avoided. Thus, for current-mode circuit design, the edge of saturation must be thoroughly defined under worst-case conditions so that V_{CC} and R_C can be calculated to keep V_{CE} above the saturation level.

Therefore, to design the on state for either saturated-mode or current-mode switching circuits, the SV_{CE} and V_{BE} characteristics of the transistor must be known well enough so that worst-case conditions can be determined.

4-1 — The Collector Saturation Region

Design of the switch on condition for saturated mode operation requires detailed information regarding the behavior of two specific transistor characteristics — namely: (1) forward current transfer ratio or dc current gain (β) and (2) collector saturation voltage (SV_{CE}).

The dc current gain (β) is important because it predicts the input current (I_B) required to obtain a given value of collector current (I_C) as determined by the load. It is obviously desirable to have a high β in order to switch a large load current with a relatively small input current.

To characterize the variations of β and V_{CE} in the saturation region, characteristic curves as shown in Figure 4-2 are extremely helpful. To aid in interpreting them properly and in applying the information to actual design examples in subsequent chapters, the following definitions are used:

Current gain in the active region	$\beta = \frac{I_C}{I_B}$	V_{CE} is a specified voltage in the active or linear region
Current gain at the edge of saturation	$\beta_o = \frac{I_C}{I_B}$	V_{CE} is the voltage where saturation effects become noticeable
Current gain in saturation region or forced current gain	$\beta_F = \frac{I_C}{I_B}$	V_{CE} is a specified saturation voltage

In each definition I_C must be specified.

From Figure 4-2 it is evident that for any given value of collector current, β (in the transition region) is quite high and is relatively independent of collector voltage. At the edge of saturation current gain has decreased slightly from β to β_o. As a transistor is driven deeper into saturation, by an ever increasing drive current (I_B), β_F decreases with the increase of I_B, and SV_{CE} decreases. This decrease in SV_{CE} is limited only to a certain region of the curve; beyond this region SV_{CE} remains relatively constant regardless of the value of I_B. Thus, curves of this nature define the spread of I_B over which some SV_{CE} benefits can be obtained at the expense of β_F (circuit β), but beyond which any increase in I_B (overdrive) will contribute only to an increase in stored charge which reduces circuit speed. Final selection of $\underline{I_B}$* and $\overline{SV}_{CE}$ naturally involves additional considerations such as noise voltages generated by a changing SV_{CE} due to changes in the drive signal, transistor power dissipation, and input requirements of the following stage. A graph of this nature can help resolve these compromises.

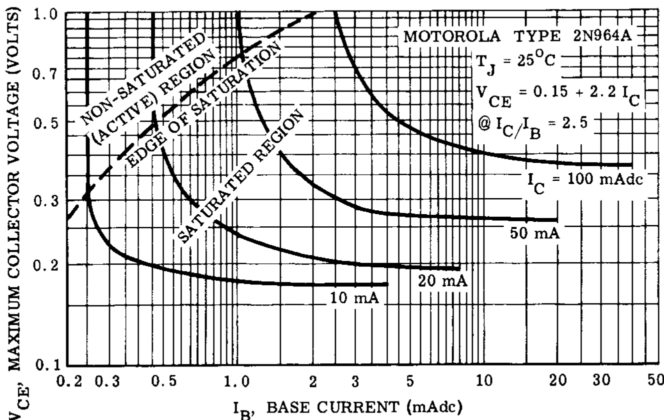


Figure 4-2 — Collector Output Characteristics

*Note: A $\overline{\text{bar}}$ over a term is used to indicate a maximum and a $\underline{\text{bar}}$ under a term is used to indicate a minimum.

The curves in Figure 4-2 show maximum limit values of V_{CE} for the Motorola 2N964A transistor for several values of collector current. However, these curves are plotted only for a specific temperature of 25°C. For worst-case design purposes, it is necessary to know V_{CE} — I_B variations at a specific collector current and at limit temperatures. Such additional curves can be constructed from information regarding two points of the desired curve.

To obtain the coordinates of these end points, data from two additional graphs is required. The active-region coordinates can be obtained from a graph of minimum β versus collector current at various temperatures, as shown in Figure 4-3 for the 2N964A transistor.

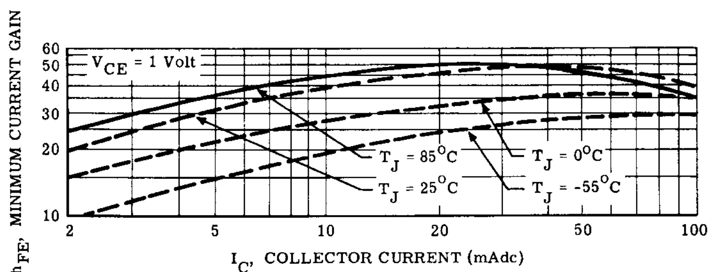


Figure 4-3 — Current Gain Characteristics for a 2N964A Transistor

The coordinates are

$$I_B = \frac{I_C}{\beta} \tag{4-4}$$

$$V_{CE} = V_{CE}\ (\text{specified})$$

V_{CE} is a collector voltage just above the edge of saturation at which β is characterized (as in Figure 4-3), and β is derived from the curve of the applicable temperature limit at the desired collector current.

The coordinates for the saturated-region point at 25° are*

$$I_B = \frac{I_C}{\beta_{FS}} \tag{4-5}$$

and

$$V_{CE} = V_P + R_F I_C \tag{4-6}$$

where

β_{FS} = a particular value of β_F which will provide a sufficient penetration of the saturation region to allow V_{CE} to be linearly related to I_C. Values of β_{FS} less than $\beta_o/2.5$ are normally satisfactory. Reasons for this will be discussed shortly.

V_P = an effective transistor offset voltage

R_F = the effective bulk resistance of the collector and emitter

I_C = the required collector current.

*See Fig. 1-9 and associated explanation.

For the 2N964A transistor a β_{FS} of 2.5 was employed and values for V_P and R_F are 0.15 and 2.2 respectively, as given in the inset of Figure 4-2. For transistors whose data sheets do not supply this type of information, the subsequent sections provide background information with which approximate values can be obtained from normally specified data-sheet information.

Equation 4-6 gives the coordinate for V_{CE} at a temperature of 25°C only. For other limit temperatures, the temperature coefficient of the saturation voltage must be added. Thus,

$$SV_{CE}(T_J) = SV_{CE}(25°) + \theta_{VC}(T_J - 25°) \tag{4-7}$$

where

$SV_{CE}(T_J)$ = collector-emitter saturation voltage at any given temperature
$SV_{CE}(25°)$ = collector-emitter saturation voltage at 25°C
θ_{VC} = temperature coefficient of saturation voltage.

The saturation-voltage temperature coefficient curves for the 2N964A transistor are shown in Figure 4-4.

Utilizing the points provided by the equations and graphs specified above, limit curves can be constructed for any temperature and collector current by connecting the end points with a curve of similar shape roughly paralleling the output characteristic curves as illustrated by Figure 4-2 for the 2N964A. This procedure is performed in detail in the inverter design example (Chapter 7).

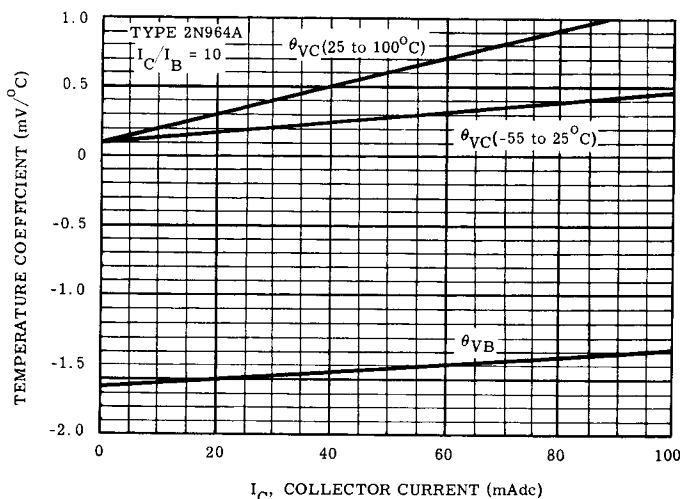


Figure 4-4 — Temperature Coefficients

4-2 — Base-Emitter Saturation Considerations

As was pointed out earlier, the design of the transistor input circuit requires a knowledge of the behavior of V_{BE}, since V_{BE} opposes the input signal voltage, thereby affecting the selection of R_K for a desired I_B. The voltage V_{BE} is primarily a function of collector current and base-emitter junction temperature. Limit values of V_{BE} for variations of collector current are plotted in Figure 4-5 and the base-emitter temperature coefficient is shown in Figure 4-4 for the 2N964A transistor.

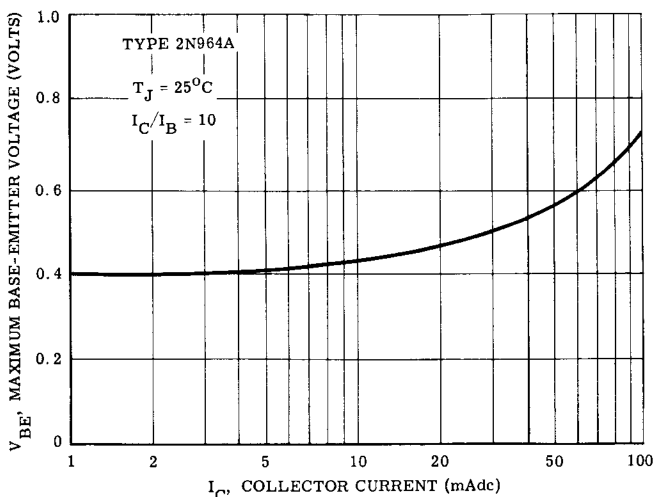


Figure 4-5 — Input Characteristics

It might appear reasonable that V_{BE} is also a function of I_B, so that the curve of Figure 4-5, which is based on an I_C/I_B ratio of 10, would not be valid for other points in the saturation region. However, the base resistance (r'_B) for a transistor when saturated is very low, due to the injection of carriers from the collector into the base. Thus, the input resistance is almost equal to the emitter resistance which is very small; therefore, variations of I_B have little effect on V_{BE} as will be shown.

Maximum V_{BE} occurs at minimum temperatures. Since Figure 4-5 shows maximum V_{BE} at a temperature of 25°C, a temperature coefficient curve is required to calculate $\overline{V}_{BE}$ at other temperatures. Therefore,

$$\overline{V}_{BE}(T_J) = \overline{V}_{BE}(25°) + \theta_{VB}(T_J - 25°) \qquad (4\text{-}8)$$

where $\overline{V}_{BE}(T_J)$ = base-emitter voltage at minimum junction temperature

$\overline{V}_{BE}(25°)$ = base-emitter voltage at 25°C

θ_{VB} = base-emitter temperature coefficient.

The foregoing method for determining SV_{CE} and V_{BE} involves the use of limit data and graphs which are not commonly given in all transistor data sheets. In the remainder of the chapter the validity of this method will be justified and a procedure for obtaining the required data, when it is not provided on the data sheet, is outlined.

4-3 — The Emitter-Base Junction

The theoretical concepts given in this chapter are based on a dc analysis by Ebers and Moll[1] modified to include the effects of bulk resistances. Their model is similar to the one which was developed in Chapter 1 using an elementary analysis. For convenience, Ebers and Moll's equations were modified so that magnitudes only are used. Therefore, the resulting equations are valid for both PNP and NPN transistors without having to use negative numbers. Voltages of interest are shown on Figure 4-6.

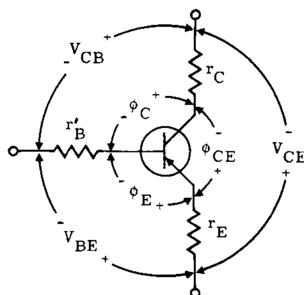


Figure 4-6 — Transistor Model Showing Voltage Conventions

The voltage across a transistor emitter-base junction is given by the expression

$$\phi_E = \frac{kT}{q} \ln \left[1 + \frac{I_E - \alpha_I I_C}{I_{ED}} \right] \tag{4-9}$$

where

ϕ_E = emitter-base junction voltage
I_E = emitter current
I_C = collector current
α_I = inverted current gain $\left[\frac{I_E}{I_C}\right]$ (Collector acting as emitter and vice versa)
$\frac{kT}{q}$ = a common transistor expression which has a numerical value of 26 mV at 27°C and increases 86 μV/°C { k = Boltzmann's constant; T = absolute temperature; q = electronic charge }
I_{ED} = diffusion saturation current.

Operation in the active region can be described by substituting for I_C its equivalent active region expression $\alpha_N I_E$, which yields

$$\phi_E = \frac{kT}{q} \ln \left[1 + \frac{I_E(1 - \alpha_I \alpha_N)}{I_{ED}} \right]. \tag{4-10}$$

This equation shows that the emitter-base junction voltage (ϕ_E) necessary to produce a given emitter current (I_E) is largely dependent upon the diffusion

* See Reference Number 1 and Appendix I.

saturation current (I_{ED}) and is modified slightly by the forward current gain (α_N) and the inverted current gain (α_I). The current I_{ED} is primarily dependent on the semiconductor material, the doping level (resistivity), the base width and the emitter area. For silicon, I_{ED} is not measurable (it is obscured by charge generation currents and surface effects) but it has been calculated to be several orders of magnitude smaller than for germanium.

As α_I and α_N approach unity the required base-emitter voltage for a given I_E is decreased. In most transistors α_N is very close to unity so that α_I is the principal factor of the $\alpha_I\alpha_N$ combination affecting ϕ_E.

The effects of the various factors in the equation are plotted in the graph of Figure 4-7. This graph, plus a knowledge of various transistor characteristics, permits a comparison between transistor types. Silicon devices, for example, with their much lower values of I_{ED} than germanium units, require higher voltages for comparable values of I_E. The difference, however, is not as great as might be supposed from the large difference in I_{ED}, because of the logarithmic nature of the curves. It is commonly said that silicon has a "band gap" voltage of 0.6 volt and germanium has a "band gap" voltage of 0.2 volt. This statement is a coarse approximation to the truth, as the two types will have the same voltage at the same ratio of I_E to I_{ED}. However, due to its lower I_{ED}, V_{BE} of silicon devices is generally 0.4 volt higher than germanium at identical current levels for types with similar geometry and resistivity.

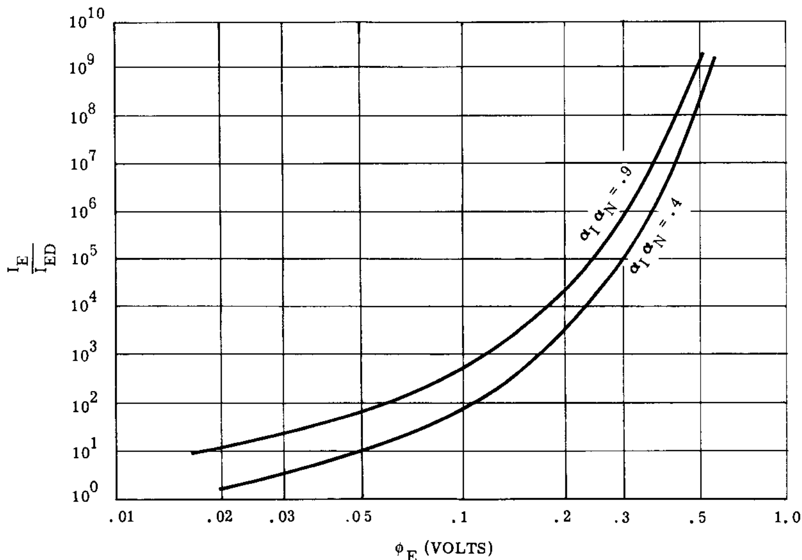


Figure 4-7 — Theoretical Transistor Input Characteristics

For transistors made of similar materials, I_{ED} varies directly with emitter area. Therefore, high-speed devices with relatively small areas require correspondingly higher voltages for a given I_E. Graded base units have lower α_I than

step junction devices because the built-in field appears as a retarding field in the inverse direction. Unsymmetrical transistor geometries have lower α_I than symmetrical units; thus mesa and planar transistors having a small area, a graded base, and an unsymmetrical geometry require higher base-emitter voltages for a given emitter current than do alloy types. The difference is small (approximately 100 mV), but it does enter into design considerations.

The actual base-emitter voltage (V_{BE}) at the base-emitter terminals, (see Figure 4-6), consists of ϕ_E, plus any voltage drops across the bulk series resistances r_B and r_E. A correction factor[2] λ, must be included in the expression for ϕ_E to account for deviations between theory and practice.

The correction factor (λ) varies between 1 and 2, approaching 1 in silicon devices at moderate current densities and approaching 2 at very low and very high current densities. For germanium, it is usually 1 at low current densities and nears 2 at high current densities. Thus, a complete expression for V_{BE} is

$$V_{BE} = \frac{\lambda kT}{q} \ln \left[1 + \frac{I_E - \alpha_I I_C}{I_{ED}} \right] + r_B I_B + r_E I_E \qquad (4\text{-}11)$$

In the active region, r_B equals r'_B but in the saturation region r_B is much less than r'_B, because the base resistivity is effectively lowered due to the injection of carriers from the collector into the base region.

Since r'_B drops as operation is moved from the active into the saturation region by increasing I_B, the voltage drop across r'_B may reduce if r'_B decreases more rapidly than I_B increases. This negative resistance can cause V_{BE} to decrease in the vicinity of the boundary between the active and saturation region as Figure 4-8 shows. After operation is well into the saturation region, r'_B becomes very small so that any increase in I_B does not result in an appreciable variation of V_{BE} until I_B becomes extremely large.

To obtain the correct design values for V_{BE} for current mode operation, which is the active region, a graph of V_{BE} vs I_C in the active region would be useful. Yet, since the negative resistance effect is normally slight, saturated values for V_{BE} at specific collector currents (Figure 4-5) can be employed without introducing appreciable error. (The steep rise of V_{BE} in the active region of Figure 4-8 is due to the increase in collector current as I_B increases.)

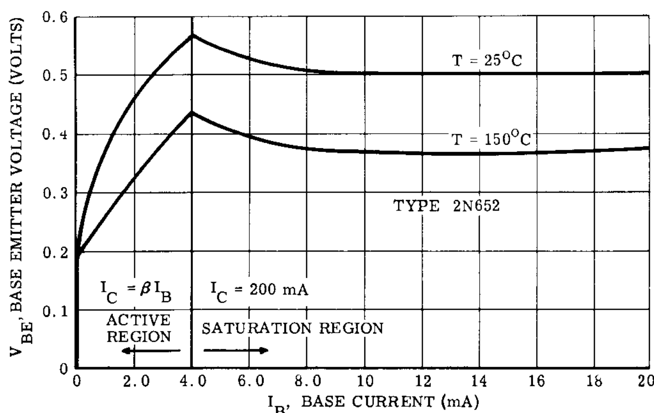


Figure 4-8 — Transistor Input Characteristics in the Region of Collector Saturation

TEMPERATURE EFFECTS UPON V_{BE}: Temperature also has a decided effect on the junction voltage required to produce a given emitter current. As described earlier, the expression $\frac{kT}{q}$ (equation 4-9) increases approximately 86μV/°C. This change, however, is insignificant compared with the variations of I_{ED}, which approximately doubles with every 10°C temperature increase as shown in Figure 3-7. The change in I_{ED} results in a temperature coefficient (θ_{VB}) of approximately —1.8 mV/°C (in the vicinity of room temperatures) for theoretical values of ϕ_E.

The bulk resistances r_B and r_E have temperature coefficients of their own which are of opposite polarity to the temperature coefficient of ϕ_E. The temperature coefficients of these bulk resistances result in a reduction of θ_{VB} as collector current increases, as indicated in Figure 4-4. Moreover, since r'_B becomes negligible in the saturation region, the effect of the temperature coefficient of r'_B can be disregarded. Thus, the overall temperature coefficient of θ_{VB} takes the form

$$\theta_{VB} = A - BI_E \qquad (4\text{-}12)$$

where

θ_{VB} = overall temperature coefficient
A = coefficient of ϕ_E (approximately 1.8 λ mV/°C)
B = coefficient of emitter resistance
I_E = emitter current.

It should be mentioned that θ_{VB} is not actually constant with temperature, because the rate of change of I_{ED} is not constant with temperature. Also, r_E may not change at a constant rate — depending upon its resistivity and the type of material. Normally, however, any error introduced by assuming a constant θ_{VB} (with temperature) is small.

4-4 — Deriving Limit V_{BE} Curves From Typical Data

Many transistor data sheets provide graphs of typical transistor characteristics in place of the limit curves provided for the 2N964A device. These typical curves can be converted into limit curves (required for worst-case design purposes) through the use of information normally given in the table of electrical characteristics plus some simple calculations. The accuracy of limit curves derived according to the procedure given depends primarily upon the extent of the given transistor specifications. The procedure outlined normally results in limits which are pessimistic.

A graph of typical V_{BE} variations with collector current, for various temperatures, as given on a 2N834 data sheet, is shown in Figure 4-9. The solid-line curves are the only ones appearing on the original data sheet graph. The dashed lines are limit curves which have been constructed for this device, using other available data and the procedure to be described.

The table of electrical specifications for the 2N834 lists typical V_{BE} = 0.74 volt (at 25°C) and maximum V_{BE} = 0.9 volt (at 25°C). These values yield a voltage difference of 0.16 volt and a ratio of maximum to typical of 1.22. The question is whether a constant difference should be added to all points of typical curves, or whether the typical values should be multiplied by the ratio, in order to obtain limit data. Investigation of the theoretical behavior of V_{BE} provides the answer.

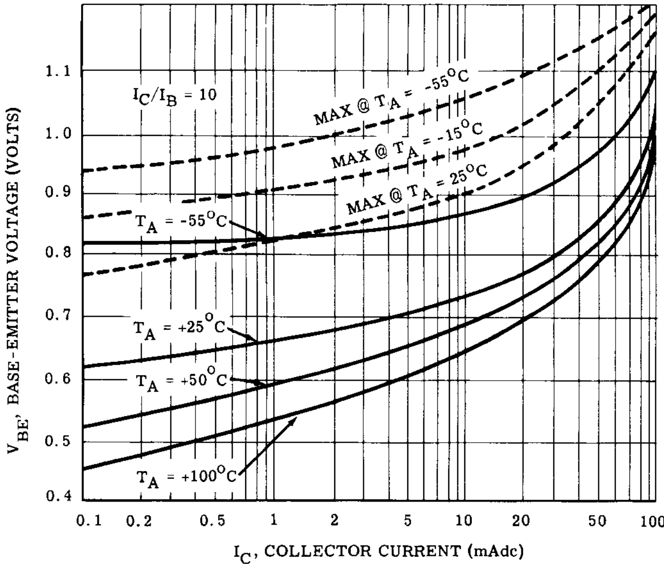


Figure 4-9 — Input Characteristics for a 2N834 Transistor

From equations 4-9 and 4-11 it is evident that

$$V_{BE} = \phi_E + I_B r_B + I_E r_E \qquad (4\text{-}13)$$

Since β_F is constant in the cases to be considered, equation 4-13 may be rewritten in terms of emitter current alone by substituting for I_B the equivalent expression $I_E/(\beta_F + 1)$, which yields:

$$V_{BE} = \phi_E + I_E \left(r_E + \frac{r_B}{\beta_F + 1}\right)$$

The term modifying I_E can be considered as a constant resistance r'_E. A typical value of V_{BE} can be written as

$$\tilde{V}^{*}_{BE} = \phi_E + I_E r'_E$$

There are two possible conditions of this equation that can express maximum V_{BE}.

$$\overline{V}_{BE1} = \underline{\phi}_E + \overline{r}'_E I_E$$
$$\overline{V}_{BE2} = \overline{\phi}_E + \underline{r}'_E I_E$$

The ϕ_E term changes very little with current compared to the $\underline{r}'_E I_E$ term. At emitter current levels below the test specification, $\overline{V}_{BE2}$ represents worst case since it would not drop as fast with decrease of emitter current as would $\overline{V}_{BE1}$ which was caused by a device having a significant r'_E.

Examination of the expression for ϕ_E (equation 4-9) reveals that the primary variable is I_{ED} since α_I and α_N change only slightly from one device to another. At any specific temperature, variations in I_{ED} between transistors are dependent only on variations of physical transistor differences, so that any specified difference

*NOTE: The tilde (~) above a symbol indicates a typical value.

between maximum and typical V_{BE} (as given in the data sheet characteristics table) can be expected to remain constant — from the standpoint of ϕ_E. Therefore at low collector currents where the drops due to $I_E r_E$ and $I_B r_B$ are negligible, the difference between $\overline{V}_{BE}$ and $\tilde{V}_{BE}$ will be constant. The lowest current where $\overline{V}_{BE}$ is specified may not be at a point where these bulk resistance drops are negligible. However, it is certainly safe to simply assume this is the case and add the difference between $\overline{V}_{BE}$ and $\tilde{V}_{BE}$ to $\tilde{V}_{BE}$ to obtain $\overline{V}_{BE}$ at current levels below this point. At emitter current levels above the test specification, the worst case would be represented by $\overline{V}_{BE1}$ since the increase in V_{BE} due to r'_E would be proportional to current whereas ϕ_E increases as the ln of current — a much lower rate. Thus, at currents above the highest current for which V_{BE} was specified, a reasonable method to obtain a maximum V_{BE} is to multiply all typical values by the ratio of $\overline{V}_{BE}$ to $\tilde{V}_{BE}$ at the test point.

This procedure has been applied to the 2N834 data and is shown as a dotted curve on Figure 4-9.

Once a limit curve for V_{BE} has been determined at 25°C, the temperature effect must be considered. In the case of the 2N834, the temperature behavior can easily be computed, since typical variations of V_{BE} with temperature are shown on the data sheet (See Figure 4-10). The slope of any one curve is constant with temperature which indicates that the temperature coefficient is constant with temperature at a given current. A temperature coefficient graph can be constructed by determining the mV/°C change in V_{BE} for each collector current. For example, at 50 mAdc collector current, V_{BE} changes 0.24 volt over the temperature range from −50°C to +150°C (total change 200°C). The change per degree is .24 /200= 1.2 mV/°C. The same procedure is applied to each collector current and the points are plotted, as in Figure 4-11, to obtain a temperature coefficient graph.

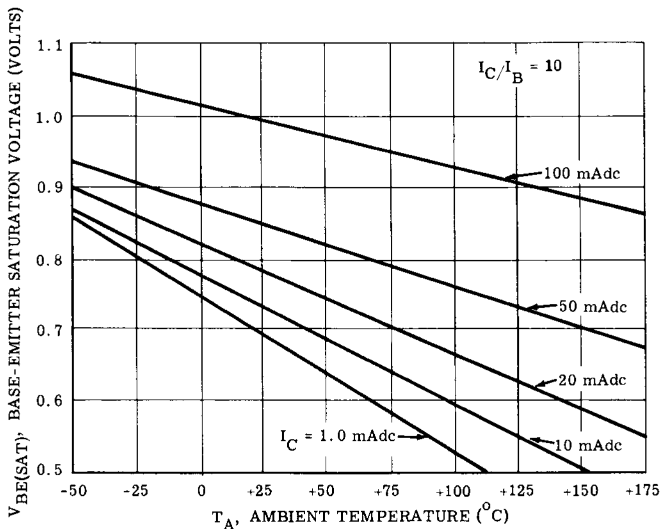


Figure 4-10 — V_{BE} Data for a 2N834 Transistor

This constructed temperature coefficient chart represents the behavior of typical transistors. At low emitter currents, all transistors of a given type have approximately the same temperature coefficient. This occurs because $V_{BE} \approx \phi_E$ whose temperature behavior is determined by I_{ED}, a characteristic which behaves in a predictable manner with temperature as shown in Figure 3-7. At high current levels, the effect of r'_E becomes noticeable. Normally, r'_E has a positive coefficient which in effect cancels a part of the negative coefficient of ϕ_E, in proportion to emitter current, which is indicated by the slope of the line in Figure 4-11.

The most often required limit of V_{BE} is $\overline{V}_{BE}$ at $\underline{T}_A$. At 25°C it is quite probable that all units having a maximum V_{BE} will have an r'_E higher than typical. Thus these units would have a lower temperature coefficient than the typical unit. Use of the typical data in this case results in a conservative value for $\overline{V}_{BE}$.

Using the temperature coefficient data and equation 4-8, a maximum curve can be constructed at any temperature. In Figure 4-9, this has been done at —15° and —55°C for the 2N834 transistor. The process simply consists of multiplying the temperature differences by the temperature coefficient (from Figure 4-11) at certain current values and adding the result to the 25°C maximum curve.

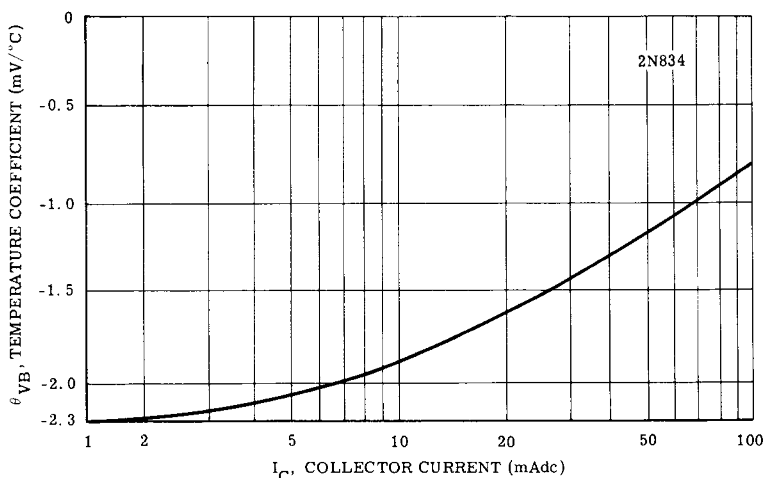


Figure 4-11 — Derived Temperature Coefficient

If typical data, such as Figure 4-9 and 4-10, are not given for a transistor type, it will have to be obtained by measurement. Theory does not permit computation with sufficient accuracy for design purposes since the details of a particular transistor design normally are not known.

4-5 — The Collector-Base Junction

The collector-base junction voltage follows an expression similar to equation 4-11.

$$V_{CB} = \frac{\lambda kT}{q} \ln \left[1 + \frac{\alpha_N I_E - I_C}{I_{CD}} \right] - I_C r_C + I_B r_B \qquad (4\text{-}14)$$

In the active region $I_C = \alpha_N I_E + I_{CD}$, and $V_{CB} \rightarrow -\infty$ according to equation 4-14. This means that the junction may assume any reverse voltage value as determined by the external circuit constants. In the saturation region, V_{CB} follows the curves of Figure 4-12. If I_C is held constant and I_E (and consequently I_B) is increased, then V_{CB} would increase without a limit, although its polarity has reversed from that in the active region. The line of $V_{CB} = 0$ theoretically represents the edge of saturation since the collector junction is forward biased when V_{CB} has the polarity shown on Figure 4-6. In practice, however, the curves do not bend (i.e., saturation does not occur) until injection from the collector becomes appreciable. This may require several tenths of a volt for a silicon transistor operated at high currents. Of course, if a high series collector resistance exists, it would be possible for a transistor to be in saturation even with V_{CB} in the reverse direction. This occurs in standard mesa types at high currents.

The previous comments for the emitter-base junction, regarding differences between transistor types and temperature behavior, would apply also to the collector-base junction. Since this information is required primarily for the design of common-base circuits, which are seldom used, it will not be further developed here.

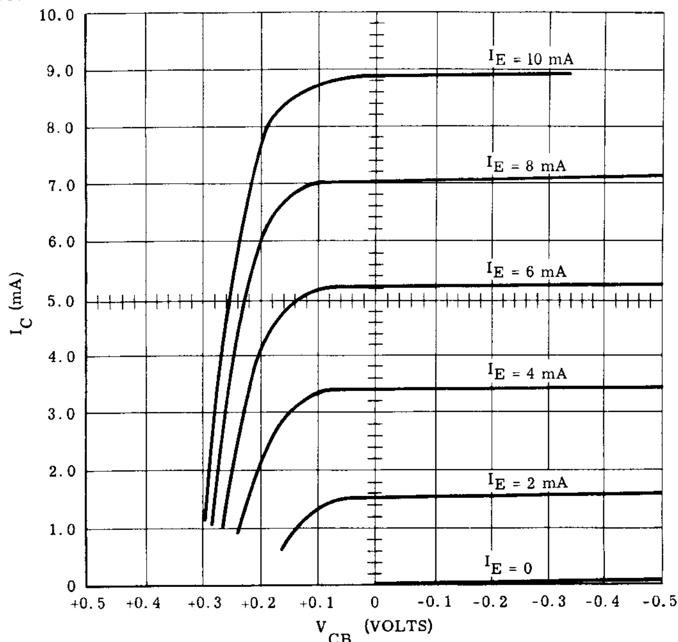


Figure 4-12 — Common-Base Output Characteristics

4-6 — The Collector-Emitter Voltage

The collector-emitter voltage is the difference between V_{BE} and V_{CB} and is given by

$$SV_{CE} = \frac{\lambda kT}{q} \ln \left(\frac{\alpha_N}{\alpha_I} \right) \left[\frac{I_B + I_C (1 - \alpha_I)}{\alpha_N I_B - I_C (1 - \alpha_N)} \right] + r_E I_E + r_C I_C \quad (4\text{-}15)$$

where

λ = the correction factor previously discussed with equation 4-11
k = Boltzmann's constant
T = absolute temperature
q = electronic charge
($\frac{kT}{q}$ = 26 mV at 27°C)
α_N = normal current gain
α_I = inverted current gain
I_C = collector current
I_B = base current
r_E = bulk emitter resistance
r_C = bulk collector resistance.

Equation 4-15 can be simplified as follows:

$$V_{CE} = \phi_{CE} + r_E I_E + r_C I_C \quad (4\text{-}16)$$

where ϕ_{CE} is the first term of equation 4-15 and represents the voltage across the collector and emitter junctions in the absence of bulk resistivities.

For the transistor saturation region, the behavior of ϕ_{CE} can be determined by substituting the following identities into equation 4-15:

$$\beta_o = \frac{\alpha_N}{1 - \alpha_N}$$

$$\beta_F = \frac{I_C}{I_B} = \text{circuit } \beta \text{ in the saturation region}$$

These substitutions, plus some algebraic manipulations yield

$$\phi_{CE} = \frac{\lambda kT}{q} \ln \left(\frac{1}{\alpha_I} \right) \left[\frac{1 + \beta_F (1 - \alpha_I)}{1 - \beta_F / \beta_o} \right]. \quad (4\text{-}17)$$

Note that the equation contains only gain terms, therefore ϕ_{CE} is independent of I_C. A plot of this equation, assuming λ of 1, is shown in Figure 4-13. Properly interpreted, this yields some interesting results when $\beta_F \ll \beta_o$

(1) At low values of β_F (the transistor driven deep into saturation), ϕ_{CE} is virtually independent of transistor gain, β_o.

(2) α_I is the most important factor affecting ϕ_{CE}.

The variations of α_I with current level for various types of transistors is illustrated in the graph of Figure 4-14. Transistors used to prepare this graph are type 2N651, a low-frequency alloy transistor, type 2N964A, an epitaxial mesa transistor, and type 2N501A, a micro-alloy diffused base transistor. The graph shows that, for the first two transistor types, α_I remains reasonably constant for wide variations of collector current, and that relatively large changes in α_I occur for the third transistor. These wide changes in α_I produce about a 100 mV change in ϕ_{CE} as indicated by the graph of Figure 4-13. It is also evident that ϕ_{CE} can vary considerably between transistors of different types as indicated by the rather large differences in α_I. The effect of α_I on SV_{CE} would be noticed only at low collector currents where the bulk drops are negligible.

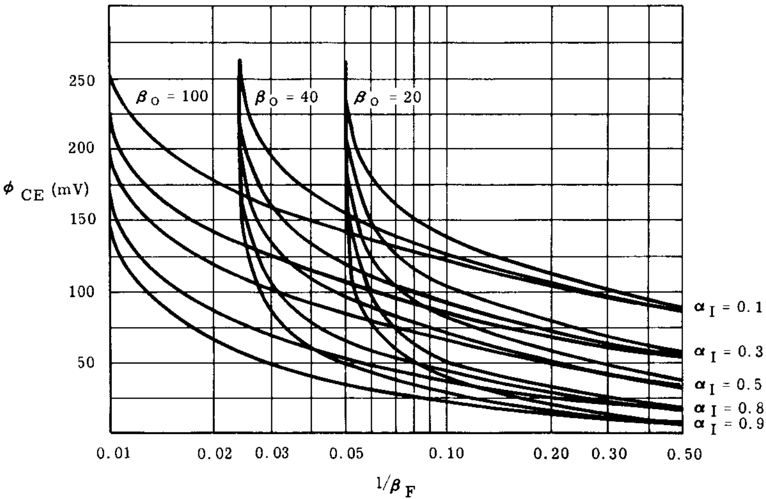


Figure 4-13 — Theoretical Knee Characteristics

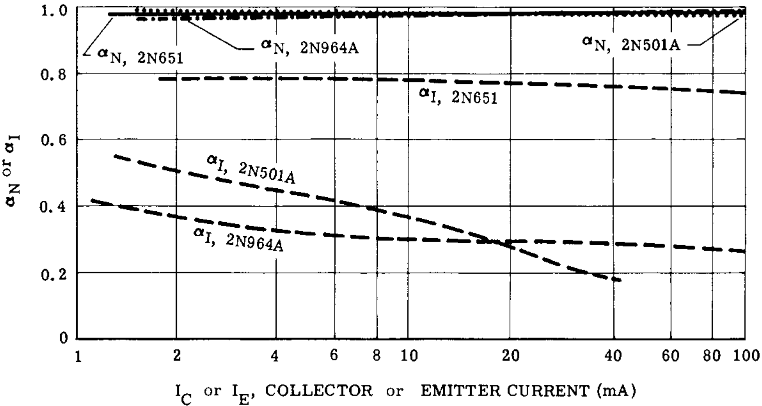


Figure 4-14 — Behavior of Forward and Inverse Current Transfer Ratio

Therefore, at moderate to high current levels, the expression for SV_{CE}, can be given as a function of collector current at a specific temperature when β_F is constant by the simple expression

$$SV_{CE} = V_P + R_F I_C \qquad \frac{I_C}{I_B} = \beta_{FS} < \frac{\beta_o}{2.5} \tag{4-18}$$

where V_P = projected offset voltage $\approx \phi_{CE}$ (considered constant)
R_F = effective saturation resistance $(r_E + r_C)$.

Most transistor measurements will fit equation 4-18, however, some transistor types do not. This occurs because R_F may vary with current as the distribution of current in a transistor is dependent upon current density. Also, R_F could be very low and α_I variations extreme which would result in non-conformance with equation 4-18.

TEMPERATURE EFFECTS UPON V_{CE}: As indicated in equation 4-15, absolute temperature, T, is a factor in determining SV_{CE}. Thus, when the effects of r_E and r_C are negligible, (low currents), SV_{CE} should be proportional to the absolute temperature, increasing at a rate of $86\mu V/°C$. This temperature coefficient is valid when operating deep in the saturation region, where the effects of temperature changes on parameters within the ln term of the equation are small. As operation moves toward the active region (β_F increases), the effects of β on ϕ_{CE} become increasingly pronounced, as indicated in Figure 4-13. Therefore, variations in β_o, due to temperature changes, begin to influence the temperature coefficient and, as the edge of saturation is approached, the temperature effects of SV_{CE}, for all practical purposes, are characterized by temperature variations of β alone.

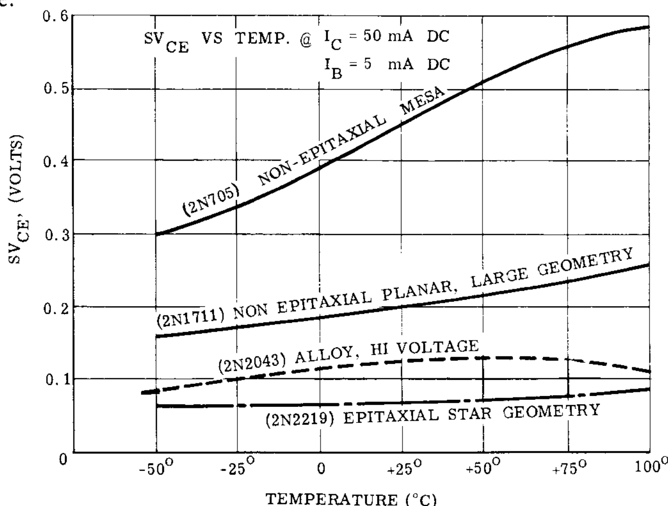


Figure 4-15 — Effect of Temperature Upon SV_{CE}

As collector current increases, the temperature coefficient of R_F becomes important. The temperature coefficient of $R_F(X)$ is expressed as a change in resistance with temperature. The overall transistor voltage temperature coefficient, therefore, takes the form

$$\theta_{VC} = W + XR_FI_C \qquad (4\text{-}19)$$

where W and X are the coefficients of ϕ_{CE}, and R_F respectively. Thus, θ_{VC}, like θ_{VB}, can be expressed as a linear function of I_C.

This temperature coefficient, however, may not be constant with temperature. Alloy transistors have a low collector resistivity but a high base resistivity. Mesa transistor types have a high collector resistivity and an intermediate base resistivity. Therefore the coefficient can vary with temperature in a number of ways, depending upon the resistivity in the various regions and must be determined by measurements. In the case of the 2N964A (see Figure 4-4) the temperature coefficient increases with temperature at temperatures above 25°C. Thus, it is necessary to have two curves of θ_{VC}. At temperatures between 25°C and 100°C, the given curve results in a slightly excessive θ_{VC}, but the error is not large and results in a conservative design. Figure 4-15 shows the temperature behavior of SV_{CE} for several different types of devices.

4-7 — Obtaining Limit Collector Saturation Characteristics

To obtain a complete picture of the saturation region for transistors not characterized by limit curves, as shown in Figure 4-2, it is necessary to know the typical behavior of V_{CE} in the knee region, as well as the limits of β and SV_{CE} with current and temperature.

A normalized plot of β for the 2N834 transistor is shown in Figure 4-16. If such a graph is not available, it will have to be obtained by measurement. Measurements have shown that the higher the β of a transistor, the more sensitive, on a percentage basis, it becomes with respect to current and temperature. Thus, if a data sheet specifies a minimum β somewhere near the peak of the h_{FE} — I_C curve, the typical normalized curve may be used to obtain minimum β at any other current level with the assurance that it will lead to conservative design.

The normalized β curves for the 2N834 show that β increases with temperature at low currents, but that it peaks at about room temperature at high currents. This is normal transistor behavior. With this graph, plus a $\underline{\beta}$ specification at one point, $\underline{\beta}$ can be estimated for any current and temperature by simply multiplying the specified $\underline{\beta}$ by the normalizing factor at the desired point. For example: the specified $\underline{\beta}$ is 25 at 10 mA and 25°C. It is desired to obtain $\underline{\beta}$ at 50 mA and —55°C. From Figure 4-16, it is seen that at the desired condition, normalized β, is 0.64 of the specified $\underline{\beta}$. Therefore, $\underline{\beta}$ at the desired condition is 0.64 x 25 = 16.

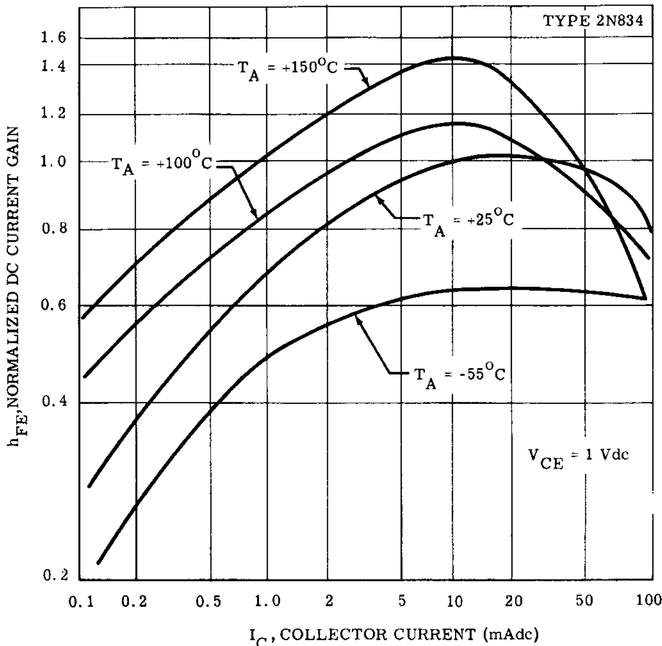


Figure 4-16 — Normalized Current Gain Characteristics

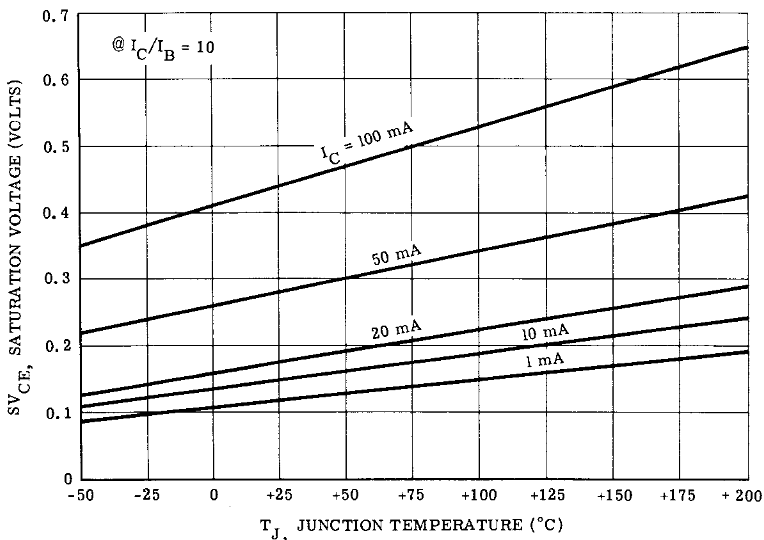


Figure 4-17 — Saturation Voltage Behavior for the 2N834

The typical behavior of SV_{CE} for a 2N834 transistor with current and temperature is shown in Figure 4-17. This data can be used to determine $\overline{SV}_{CE}$ at any current and temperature. Since the I_C/I_B ratio is fixed at 10, the typical β would have to be greater than 25, over the current and temperature range shown on the figure, for data to be described by equation 4-18. Typical β for the 2N834 is given as 40 at 10 mA and 25°C; inspection of Figure 4-16 shows that the $\underline{\beta}$ criteria is met at temperatures above 25°C and currents above 1 mA.

The data of Figure 4-17 can be converted to the typical curves shown in Figure 4-18, which plots SV_{CE} as a function of collector current at two temperatures, 25°C and 175°C. It is seen that the data results in straight lines conforming to equation 4-18.

At low current levels, $SV_{CE} \cong \phi_{CE}$. Therefore, a constant difference in millivolts exists between the two typical saturation voltages. At high levels however, $SV_{CE} \cong R_F I_C$. The resistance changes a given percentage with temperature. Therefore, the saturation voltage will change by a ratio. Thus, the curves of Figures 4-17 and 4-18 should fan upward, as shown.

The difference between $\overline{SV}_{CE}$ and $\widetilde{SV}_{CE}$, as I_C varies, must now be determined. Again, at low levels $SV_{CE} \cong \phi_{CE}$; therefore, the difference between a minimum and a typical device must be due to different values for α_I. Since the percentage change in α_I is not determined by its value, the difference in voltage between these transistors would remain constant over the temperature range. As current increases, the effect of R_F is evident with the result that the difference between $\overline{SV}_{CE}$ and $\widetilde{SV}_{CE}$ is greater at high currents. Furthermore, $\overline{SV}_{CE}$ increases faster than $\widetilde{SV}_{CE}$ with increases in temperature. This reasoning is verified experimentally. When comparing a transistor with a high SV_{CE} to one with a low SV_{CE}, it is found that at low current levels, the transistors remain a constant number of millivolts apart over the temperature range, while at high currents they bear a nearly constant ratio to each other.

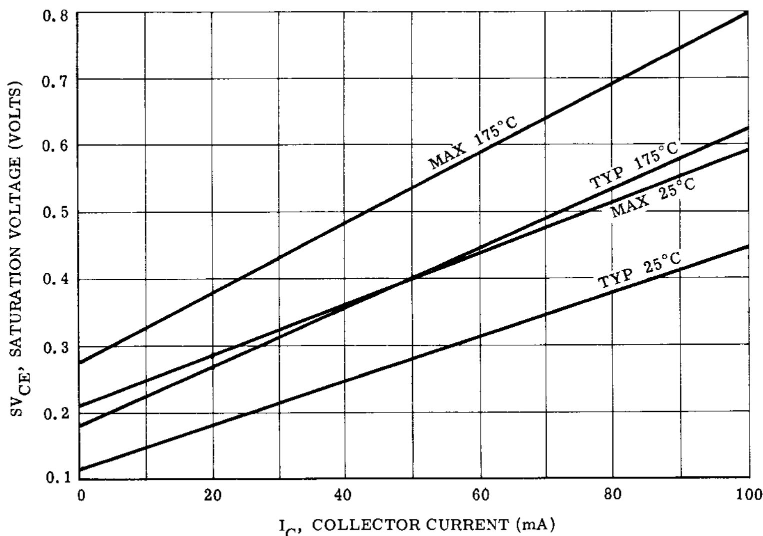


Figure 4-18 — Constructed Saturation Curves for a 2N834 Transistor

Since $\overline{S}V_{CE}$ is of concern for design purposes, maximum curves must be constructed. The specified values for $\overline{S}V_{CE}$ given on the data sheet are: 0.25 volt at 10 mA and 25°C, and 0.4 volt at 50 mA and 25°C. A straight line can be drawn through these two points on Figure 4-17 to construct the $\overline{S}V_{CE}$ curve at 25°C. To construct the $\overline{S}V_{CE}$ curve at 175°C, the difference between the typical 25°C curve and the maximum 25°C curve must be added to the typical 175°C curve.

Notice that the maximum curves have a slope about 20% greater than the typical curves. This information can be used, as a rough guide, to construct a maximum curve if $\overline{S}V_{CE}$ is specified at only one point on the data sheet; i.e., the slope of the maximum line should be approximately 20% steeper than the slope of the typical line.

From the maximum 25°C line, a defining equation for $\overline{S}V_{CE}$ can be written by noting the Y intercept which gives V_P, and the slope of the line which gives R_F. For the example under consideration, the Y intercept is at 0.21 volt and the slope of the line changes from 0.21 to 0.59 — a total change of 0.38 volt over a range of 100 mA. Then, $R_F = 0.38$ V. $\div$ 100 mA $= 3.8\Omega$. Therefore,

$$\overline{S}V_{CE} = 0.21 + 3.8\, I_C \text{ at } 25°C\ (I_C/I_B = 10).$$

In a similar manner, an equation for the maximum 175°C line can be written yielding

$$\overline{S}V_{CE} = 0.271 + 5.3\, I_C \text{ at } 175°C.$$

To obtain the temperature coefficient, the difference between the two equations must be divided by the difference in temperature. This produces the temperature coefficient equation

$$\theta_{VC}(\text{mV/°C}) = 0.4 + 10\, I_C$$

where I_C in these equations is in units of amperes.

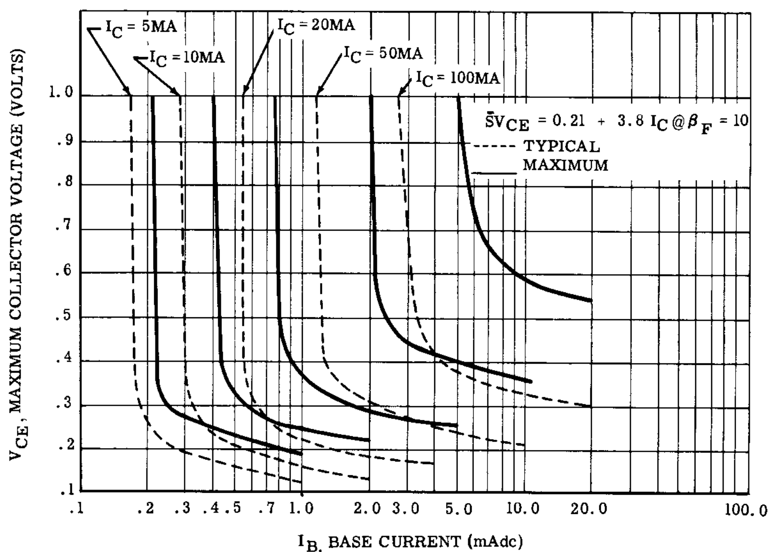


Figure 4-19 — Knee Characteristics for a 2N834 Transistor

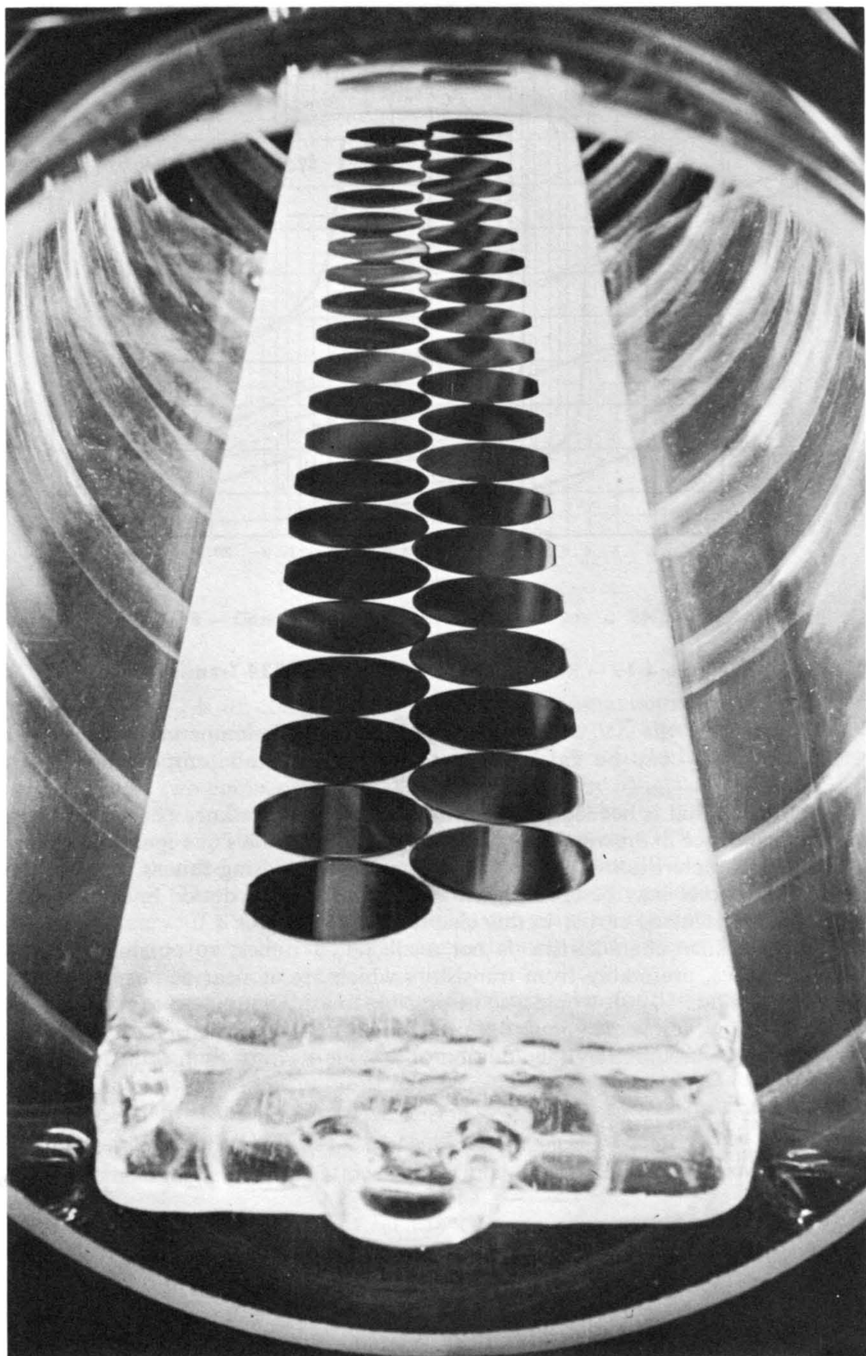
Thus, with the $\bar{S}V_{CE}$ equation at 25°C and the temperature coefficient equation, $\bar{S}V_{CE}$ can be calculated at any current and temperature when $I_C/I_B = 10$.

Now, all that is needed is the normal behavior of the knee characteristic to provide guidance in drawing the limit curves. The solid lines of Figure 4-19 show the knee characteristics of a low-β 2N834 transistor. Using this as a guide, the maximum curves may be constructed as indicated by the dotted lines using the proceedure discussed earlier in this chapter and inChapter 7.

If this knee characteristic is not available, it is best to obtain measured characteristics, preferably from transistors which are as near as possible to the specified β_0 and $\bar{S}V_{CE}$. It would also be possible to use Figure 4-13 to predict knee behavior. However, some knowledge of α_I is needed to pick the right curve which would then be modified by the addition of a value $R_F I_C$ to each ordinate.

REFERENCES

1. Ebers, J. J., and I. L. Moll: *"Large-signal Behavior of Junction Transistors"*, Proceedings of the IRE, vol. 42, pp 1761-1772, December, 1954.
2. Pritchard, R. L.: *"Advances in the Understanding of the P-N Junction Triode"* Proc. IRE, vol. 46, pp 1130-1140.



EPITAXIAL GROWTH ON SILICON WAFERS

CHAPTER 5

Transient Characteristics of Transistors

The characteristics of transistors in the on and off states were described in detail in Chapters 3 and 4. Although the circuit design of these static states is based upon the transistor's dc characteristics, the effect of the dc conditions upon the transient response must be considered. In some circuits, such as those which use resistance coupling, the dc design may be based entirely upon the transient performance. In many circuits, the dc conditions are dictated largely by load requirements, and the speed is enhanced by using special techniques, such as a speed-up capacitor. The problems and approaches are innumerable, but rare is the circuit where the transient response does not have to be investigated. Many of these inter-relations were discussed in Chapter 2.

In Chapter 1, a model of a transistor was developed in which external elements were added to a "perfect" transistor. These external elements — capacitances and stores — were used to explain the transient behavior in a qualitative manner and permit a simple discussion of device physics. That is about all the model is good for. The problem is that the elements are so non-linear that calculations of large signal or switching performance by classical methods become very complex.

In this chapter, a simple approach is developed from the charge control theory. The usefulness of this method stems from the fact that the variation of non-linear elements with time does not have to be considered. It is only necessary to be concerned with the movement of a discrete amount of charge.

5-1 — Definition of Transition Times

Before discussing the charge control concept of determining transition response, the idealized transient behavior of a transistor switch will be reviewed. In Figure 5-1, the transistor is being switched from off to on and back to off by the application of a step pulse to the transistor base. It is obvious that the output pulse is far from being an exact duplicate of the input pulse. Reasons for this waveshape distortion and some important definitions are given in the following paragraphs.

At time t_0 the pulse generator delivers a step of base current (I_{B1}) to the transistor. At this instant, the transistor is in the off condition because of the emitter reverse-bias voltage (V_{OB}). The collector current that is flowing is extremely low (approximately I_{CBO}); the voltage on the collector-base junction is equal to the sum of the off level collector voltage (V_1) and the reverse-bias voltage (V_{OB}). At time t_0 the base current rises immediately to I_{B1} but it may be observed that the collector current does not begin to increase until t'_1. The time between t_0 and t'_1 is called the turn on delay time (t_d), and is defined as the time required to bring the transistor from the initial off condition to the edge of conduction, i.e.,

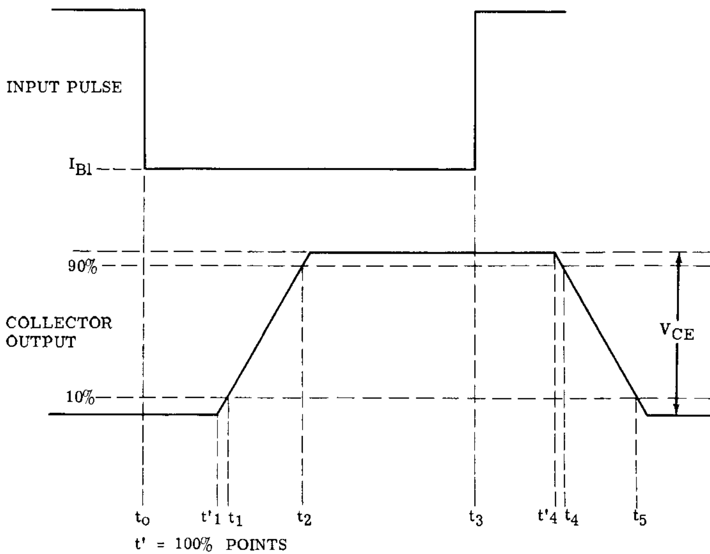


Figure 5-1 — Output Waveform of a Transistor Switch

the beginning of the active region. This may be defined as that instant of time t'_1 at which the applied base-to-emitter voltage is equal to V_{TF}, a small forward voltage which barely causes significant conduction. Physically, the finite time required for t_d comes about because of the reverse bias on both the emitter and collector junction transition capacitances. The delay time is simply the time required to charge these capacitances to the new voltage level. It should be apparent that if $V_{OB} = V_{TF}$, then $t_d = 0$*. Since in practice waveforms do not show sharp corners, for measurement ease, t_d is usually measured from the 10% point of the input pulse amplitude to the 10% point of output pulse amplitude (t_1).

At time t'_1 the operating point of the transistor is at the beginning of the active region; the emitter starts to become forward biased and begins to inject current into the base. The collector current begins to increase toward its saturation value, corresponding approximately to V_{CC}/R_C. A finite time elapses as the collector current increases from 10% to 90% of the final value which occurs at time t_2; the time interval $t_2 - t_1$ is defined as the rise time (t_r) of the collector current pulse. Rise time is caused by a finite transit delay between the base and the collector currents, due to the emitter diffusion capacitance and both the emitter and collector transition capacitances.

The transistor will remain in the on state as long as the input base current I_{B1} is maintained. At time t_3 the base input pulse "steps-off" immediately; however, it is observed that the collector pulse does not respond until time t'_4. The

*A delay time due to carrier transit time exists but it can be neglected for all practical purposes.

time interval between t_3 and t'_4 is referred to as the storage time (t_s). The storage time is a measure of the time required for the minority carriers in the base and collector to recombine back to the level corresponding to the boundary between the active and saturation regions. These excess carriers arise because the collector junction becomes forward-biased when the base current I_{B1} is greater than the I_B necessary to produce I_C; i.e.; $I_{B1} > I_C/\beta_0$.* Thus, storage time is related to a carrier recombination process, and is a measure of the minority carrier lifetime in the base and collector regions. Again, for measurement ease the point t_4, where the collector current has dropped 10% is usually used.

The transistor comes out of saturation at time t'_4 and the operating point traverses the load line again through the active region to the off state. This is the turn-off portion of the collector waveform, and the time interval between t_4 and t_5 is defined as the *fall time* (t_f). At t_5 the collector current has reduced to 10% of its on value. The description of the switching process for fall time is similar to that for rise time, except that the active region is traversed in the reverse direction.

5-2 — Charge Control Theory

The basic equation of the charge control concept is the charge continuity equation. [1,2]

$$i = \frac{dq}{dt} + \frac{q}{\tau}dt. \tag{5-1a}$$

where i = current
q = charge
t = time
τ = lifetime. (Time an isolated charge can exist before recombining)

Integrated over a given time interval, the equation becomes

$$\int_0^t idt = \int_{Q_1}^{Q_2} dq + \int_0^t \frac{q}{\tau}dt. \tag{5-1b}$$

This equation states that the total charge supplied during the interval t is equal to the charge necessary to change the current in the volume to a new level, plus the amount necessary to replenish that lost by recombination. The usefulness of the charge control approach lies in the fact that in solving equation 5-1 it is only necessary to be concerned with the absolute changes during the time interval determined by the limits of integration. The manner in which the charge varies with time is of little concern.

Applying the continuity equation to the transistor it is found that the input current must supply six components:

*β_0 was defined in chapter 4 as the current gain at the edge of saturation.

1. The current due to the rate of change of the active base charge q_a.
2. The current due to recombination of the active base charge having a lifetime τ_a.
3. The current due to the rate of change of the excess base charge q_x.
4. The current due to recombination of the excess base charge having a lifetime τ_x.
5. The current required to charge the sum of the emitter transition and the stray capacitance (C_{ib}).
6. The current required to charge the sum of the collector transition and the stray capacitance (C_{ob}).

Expressed mathematically

$$i_B = \frac{q_a}{\tau_a} + \frac{dq_a}{dt} + \frac{q_x}{\tau_x} + \frac{dq_x}{dt} + C_{ib}\frac{dv_{BE}}{dt} + C_{ob}\frac{dv_{CB}}{dt} \tag{5-2a}$$

Upon integrating and using limits, equation 5-2a becomes 5-2b

$$\int_0^t i_B dt = \int_{Q_{a1}}^{Q_{a2}} dq_a + \int_0^t \frac{q_a}{\tau_a} dt + \int_{Q_{x1}}^{Q_{x2}} dq_x + \int_0^t \frac{q_x}{\tau_x} dt + \int_{V_{BE1}}^{V_{BE2}} C_{ib}\, dv_{BE} + \int_{V_{CB1}}^{V_{CB2}} C_{ob}\, dv_{CB}$$

Although these equations appear quite formidable, in practice several terms are zero when deriving a solution for a particular transient.

5-3 — Turn-On Delay Time

Turn-on delay results when a switching transistor is being turned on from the off condition with both the emitter-base and collector-base junctions reverse biased. Under these off conditions, the internal transistor emitter and collector junction depletion layer capacitances C_{Te} and C_{Tc} plus any stray capacitance (C_s) become charged. When the transistor is turned on, the initial current that flows supplies charges to these capacitances and thus no collector current is produced until the stored charges are removed and the emitter junction becomes slightly forward biased. The time required to supply this charge is the delay time and for the case of a step of base current can easily be written as

$$t'_d = \frac{Q_{OB}}{I_{B1}} \tag{5-3}$$

where

t'_d = the actual turn-on delay (delay time as usually measured contains 10% of the rise time)

Q_{OB} = the off bias charge stored in both junctions

I_{B1} = the magnitude of the turn-on current step.

Referring to the continuity equation, it should be clear that q_a and q_x are zero and that recombination does not exist with both junctions reversed biased.

From the continuity equation, Q_{OB} is simply the charge required to change the voltage on the emitter and collector transition capacitances (C_{Te} and C_{Tc}) and any base to ground, or base to collector or supply stray capacitance. It is given by

$$Q_{OB} = \int_{V_{OB}}^{V_{TF}} C_{ib}\, dv_{BE} + \int_{V_1 + V_{OB}}^{V_1 - V_{TF}} C_{ob}\, dv_{CB} \qquad (5\text{-}4)$$

where

V_{OB} = the base off-bias (a reverse voltage).

V_{TF} = the threshold of conduction voltage (a small forward voltage).

V_1 = the off voltage level at the collector.

Equations similar to this are covered in the literature[3].

There are several methods of finding Q_{OB}. One method is to measure t_d as a function of V_{OB}, V_1, and I_{B1} and calculate Q_{OB} from equation 5-3. The results can be plotted and would be found to be independent of I_{B1}. This is attractive and a convenient method to use. However, with high speed transistors it is extremely difficult to obtain an input signal with a fast enough rise time so that it will not influence the measurement.

Another approach is to use equation 5-4 and find Q_{OB} analytically. To do this, the behavior of C_{ib} and C_{ob} vs. voltage must be known. Both of these capacitances are the sum of a transition and a stray capacitance; i.e., $C_{ib} = C_{Te} + C_{se}$, and $C_{ob} = C_{Tc} + C_{sc}$. Usually C_{Te} follows a square root behavior while C_{Tc} varies approximately as the cube root for graded junctions, the square root for step junctions, and is nearly constant for narrow epitaxial collector-base junctions. That is:

$$C_T \propto 1/(V)^n.$$

This relationship states that as $V \rightarrow 0$, $C_T \rightarrow \infty$. This inconsistency with fact is usually resolved by saying that the "actual" junction voltage is added to the "contact potential", therefore V never goes to zero. However, the value to use for the contact potential is questionable making the results of the analytical approach considerably in error.

A straightforward approach, which does not have the limitations of the previous methods, is to obtain a plot of the input and output capacitances vs. voltage and perform a graphical integration. This can usually be done by using information normally given on the data sheet. The C_{ib} and C_{ob} data can be used because the stray capacitance of the transistor case is usually included in this information. Examples will be given to illustrate some of these ideas. In the following examples, a bar over a term ($\overline{C}$) is used to indicate a maximum value, a bar under a term ($\underline{C}$) indicates a minimum value and a tilde over a term ($\widetilde{C}$) indicates a typical value.

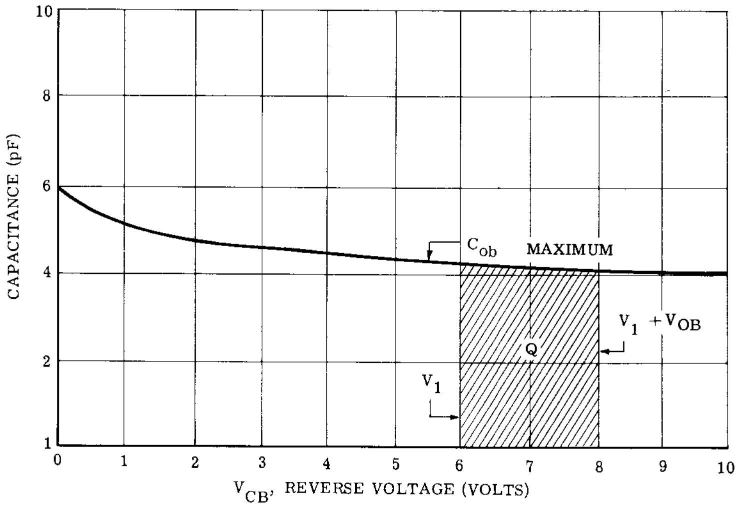


Figure 5-2 — C_{ob} Behavior for a 2N964A Transistor

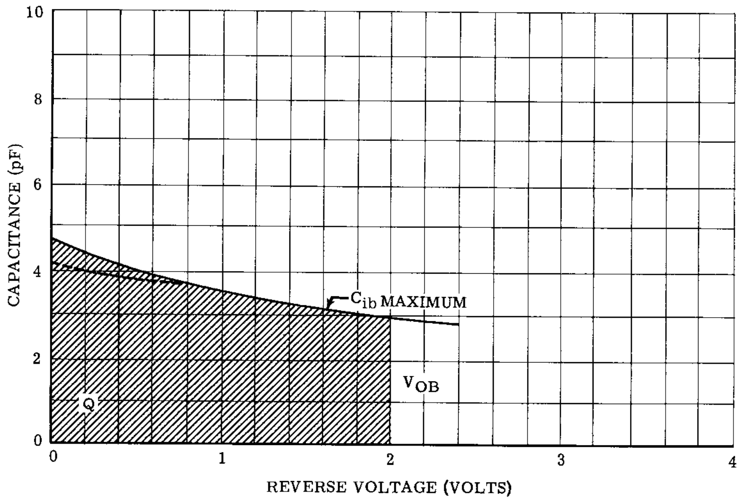


Figure 5-3 — C_{ib} Behavior for a 2N964A

Example 1: From C_{ob} and C_{ib} data for the 2N964A (Figures 5-2 and 5-3) determine Q_{OB} when the reverse bias (V_{OB}) is 2 volts and the collector voltage (V_1) is 6 volts.

Solution:

Since the 2N964A is a germanium transistor, assume that $V_{TF} = 0$. Examination of the area indicated in Figure 5-2 shows that the average capacitance is 4.2 pF, and the voltage differential is 2 volts.

$$\therefore Q_{Cob} = \triangle(CV) = (4.2)\,(2) = 8.4 \text{ pC.}$$

From Figure 5-3 notice that the effective capacitance for Q_{Cib} can be considered as a rectangle of 3 pF x 2V, plus the area of two small triangles.

$$\therefore Q_{Cib} = (CV) = (3x2) + \left(\frac{1.1}{2}\right)(2) + (1/2)\,(.5) = 7.35 \text{ pC}$$

and $Q_{OB} = 8.4 + 7.35 = 15.75$ pC.

Example 2: From information provided on the 2N834 data sheet, develop graphs similar to Figures 5-2 and 5-3 so that Q_{OB} can be determined.
The data sheet provides: $C_{ob} = 4$ pF, at $V_{CB} = 10V$, and Figure 5-4.
Solution:

The limited specifications make this problem more difficult than the previous example, however, a satisfactory estimate can be obtained.

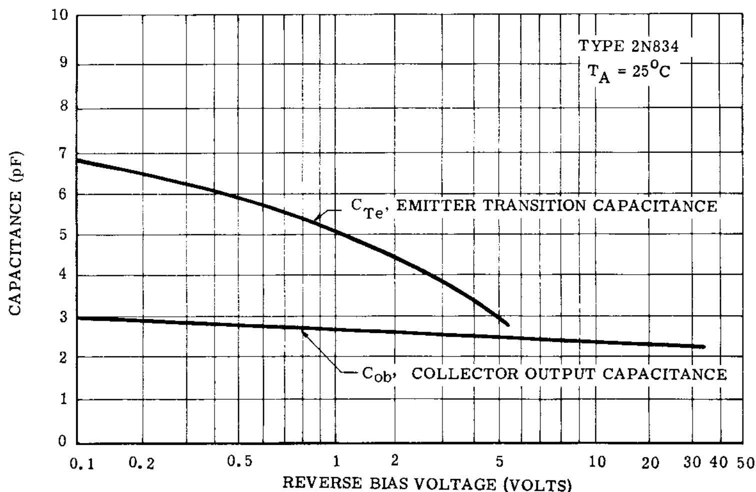


Figure 5-4 — Collector and Emitter Capacitance Characteristics

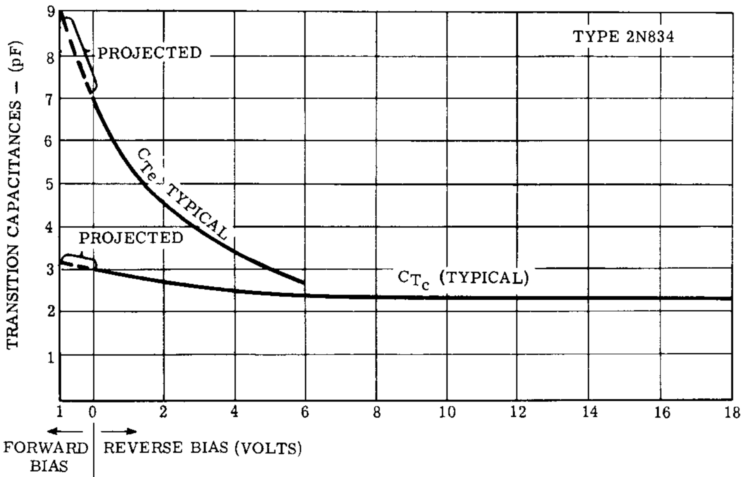


Figure 5-5 — Linear Plot of C_{Te} & C_{Tc}

1. Plot a linear graph of typical C_{Te} and C_{Tc} vs. voltage as shown in Figure 5-5. This can be done for C_{Tc} by subtracting the value of the stray header and can capacitances (see Table 5-1) from several points on the C_{ob} curve and plotting the new values. Since C_{Te} curve is given it can be transferred directly to the linear graph.

2. To obtain maximum values, apply the multiplying factor based upon the ratio of maximum to typical transition capacitance. The multiplying factor is given by:

$$M_C = \frac{\bar{C}_{ob} - C_{sc}}{\tilde{C}_{ob} - C_{sc}} = \frac{4 - .7}{2.4 - .7} = 1.94$$

All points on the C_{Tc} curve should be multiplied by M_C. This practice makes the valid assumption that there is no relationship between the transition capacitance value and the variation of transition capacitance with voltage. No maximum C_{ib} or C_{Te} data is given. However, experience shows that it is usually about twice the typical also.

If desired, maximum transition capacitance curves can be plotted, or maximum Q_{OB} can be computed from the typical curves of Figure 5-5 as

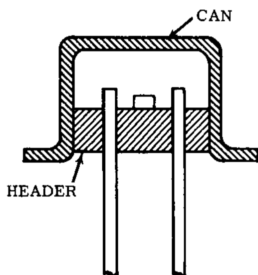
$$Q_{OB} = 2\,(Q_{CTe} + Q_{CTc}) + (C_{sc} + C_{se})\,(V_{OB} + V_{TF}).$$

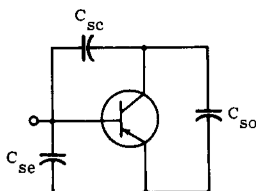
Alternately the stray capacitances could be added in, a maximum C_{ob} and C_{ib} curve drawn, and Q_{OB} computed as in the previous example.

TABLE 5-1

STRAY HEADER AND CAN CAPACITANCE FOR MOTOROLA SWITCHING TRANSISTORS

Case	C_{sc}	C_{so}	C_{se}
TO-5	0.6	0.6	0.0
TO-18	0.7	0.7	0.0





In switching transistors, it is common for the collector to be electrically connected to the case. Thus, the main source of capacitance is between the header and the leads to the base and emitter contacts. The lead-to-lead capacitance between the base and emitter is negligible.

The proper way to measure C_{ob} is to ground the collector to avoid stray pickup. Then,

$$C_{ob} = C_{Tc} + C_{sc}$$

When measuring C_{ib}, C_{sc} can be included in the measurement by grounding all other leads, or it can be nulled out using a guard voltage.

For delay time calculations, obviously only C_{sc} is of significance. However, the transistor socket, base bias resistor, and wiring all contribute to a stray capacitance from base to ground.

In silicon transistors, V_{TF} is on the order of 0.5 volt and therefore, it must be considered. The part of Q_{OB} due to V_{TF} is significant in relation to the part of Q_{OB} due to V_{OB}. This occurs because C_{Te} is rather high in the forward bias region as the projection of the data sheet information in Figure 5-5 shows. Experimentally it has been found that using a value of V_{TF} which is 0.1 to 0.2 volt less than SV_{BE} works well. The higher the collector current being switched, the greater the difference becomes between SV_{BE} and V_{TF}.

Finally, the validity of equation 5-3, which shows the relationship between t'_d and Q_{OB} for a step of constant current I_{B1}, is questionable in a practical circuit. For equation 5-3 to become accurate, the input pulse rise time must be less than a tenth of the delay time and its amplitude must be greater than ten times $(V_{OB} + V_{TF})$. These conditions are seldom encountered in practice. Therefore, equation 5-3 should be regarded as an optimistic approximation to the real case. Nevertheless, Q_{OB} constitutes the charge which must be removed during the delay time period, and forms a useful basis of comparison between transistor types. The simplest way to find delay time is to analyze the driving circuit to determine the time necessary to deliver the charge Q_{OB} to the input capacitance of the transistor.

5-4 — Rise Time

In order to turn on a transistor, sufficient charge must be supplied to do the following:

1. Charge the input transition capacitance C_{ib} to V_{BE}, and establish the proper carrier gradient in the base region to permit the desired collector current to flow. The charge required to do this function is called Q_I.
2. Change the voltage on the collector capacitance (C_{ob}) from the off level at the beginning of injection ($V_1 - V_{TF}$) to a new level ($V_O - V_{BE}$). The charge required for this function is called Q_V.
3. Allow for recombination. This charge will be called Q_R.

Referring to the continuity equation, note that for the rise time interval the terms involving the excess charge q_x are zero and the equation becomes

$$\int_0^t i_B dt = \int_0^{Q_a} dq_a + \int_0^t \frac{q_a}{\tau_a} dt + \int_{V_{BE1}}^{V_{BE2}} C_{ib}\, dv_{BE} + \int_{V_{CB1}}^{V_{CB2}} C_{ob}\, dv_{CB}. \tag{5-5a}$$

Using the correspondence between terms previously discussed and the continuity equation, and lumping the dq_a and C_{ib} terms into Q_I, the equation for a step of input current I_{B1} is

$$I_{B1} t_r = Q_I + Q_V + Q_R. \tag{5-5b}$$

In Chapter 1, it was explained that the base current is a result of injection into the emitter and recombination. In transistors operated at not more than three times the collector current at the point of maximum gain, injection from the base is negligible and recombination is the dominant factor in determining the base current. The recombination current is simply i_C/β and the recombination charge during the rise time interval is

$$Q_R = \int_0^{t_r} i_C/\beta \, dt = \int_0^{t_r} \frac{q_a}{\tau_a} dt. \tag{5-6}$$

To evaluate equation 5-6 the variation of i_C with time must be known. It actually increases approximately exponentially, however, if sufficient overdrive is applied it will be nearly linear. This is evident by examining Figure 5-6. Notice the exponential rise time when I_{B1} is just large enough to allow a current $I_C \approx V_1/R'_C$ to flow. (V_1 is the off level and R'_C is the Thevenin equivalent of the load circuit). The rise time (t_{r1}) from zero to the 90% point will be 2.3 transistor time constants.

Now consider the same transistor being turned on by a current equal to twice the previous value. The output current tries to rise to a value twice as great as the previous current but is limited by V_1/R'_C to the same value as before. By overdriving, the slowly rising portion of the exponential is clipped thereby substantially shortening rise time and making the response of i_C as a function of

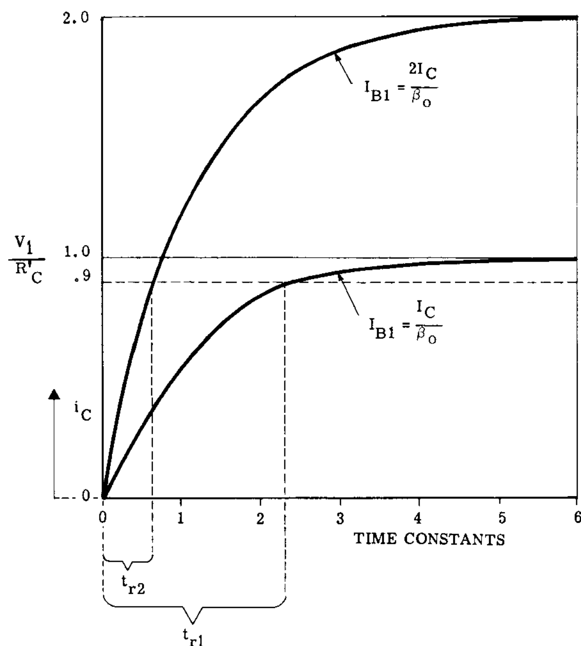


Figure 5-6 — Effect of Overdrive upon Rise Time

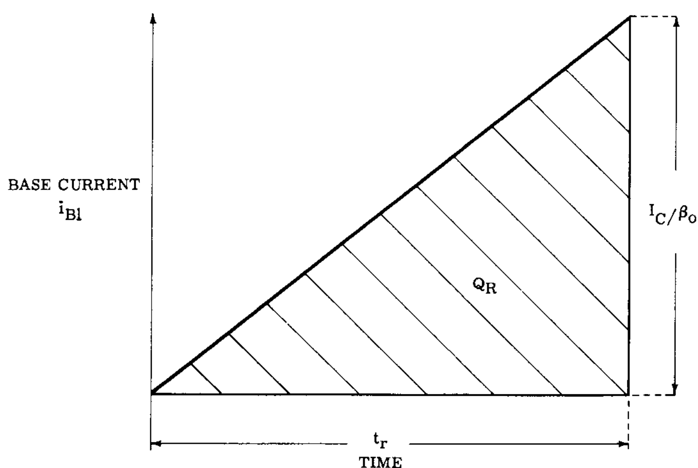


Figure 5-7 — Graphical Method for Determining Recombination Charge

time appear nearly linear. Thus, rise time could be made as short as desired* by the use of heavy overdrive currents, the reduction in time being approximately proportional to the drive current once an overdrive of 2:1 is exceeded.

Assume that some overdrive is used and find Q_R by assuming i_C increases linearly with time. This is accomplished by a graphical integration using Figure 5-7. Q_R is simply the area under the curve which is given as

$$Q_R = \frac{t_r I_C}{2\beta_o}. \tag{5-7}$$

Substituting this result in equation 5-5b and solving for t_r

$$t_r = \frac{Q_I + Q_V}{I_{B1} - I_C/2\beta_o}. \tag{5-8}$$

Since Q_I and Q_V are not normally specified on data sheets and are difficult to measure directly, it is instructive to put equation 5-8 into a more familiar form. Essentially Q_I represents the charge required to supply C_{Te} and C_{De}. Phillips[4] shows that

$$Q_I = \frac{I_C}{\omega_T}. \tag{5-9}$$

Where ω_T, the gain-bandwidth product, represents the cutoff frequency which C_{De} and C_{ib} form with the small signal emitter resistance (r_e).

i.e., $\omega_T = \dfrac{1}{r_e(C_{De} + C_{ib})}$, using a small signal equivalent representation.

The charge Q_V is simply $\triangle[(C_{Tc} + C_{sc})V_{CB}]$. Since C_{Tc} varies inversely with voltage, Q_V can be represented as $(K_Q C_{Tc} + C_{sc})\triangle V_{CB}$; where C_{Tc} is a constant if the value of K_Q is chosen so that the change in charge is the same as $\triangle(C_{Tc}V_{CB})$ for a given voltage swing. In other words:

$$C_{Tc}(\text{eff}) = \frac{\triangle Q}{\triangle V} = K_Q C_{Tc} \tag{5-10}$$

A general mathematical expression for this was developed by Narud and Aaron[5] who show that

$$K_Q = \frac{C_{Tc}(\text{eff})}{C_{Tc}} = \frac{1 - V_F/V_I}{(1-n)(1-V_F/V_I)} \tag{5-11}$$

where

C_{Tc} is measured at V_I, the initial high voltage.

V_F is the final low voltage.

n is the exponent governing the C_{Tc} vs. V relationship.

Table 5-2 gives K_Q values for a few important cases as calculated from equation 5-11.

*The lower limit upon rise time is approximately equal to the transit time of a carrier through the base.

Using the derived expressions an equivalent expression for Q_V may be written as

$$Q_V = (K_Q C_{Tc} + C_{sc}) \triangle V_{CB}. \quad (5\text{-}12)$$

Most authors have neglected the stray capacitance term, (C_{sc}). The bracketed term can be defined as an average collector-base feedback capacitance (C_f).

Since the rise time is usually expressed at the 90% point, by assuming $\triangle V_{CB} \approx \triangle V_{CE}$, $\triangle V_{CB}$ can be written as 0.9 $I_C R'_C$ where I_C is the final value of current which flows and R'_C is the Thevenin equivalent load resistance. The Q_I term should also use 0.9 I_C. Making these substitutions and also using equations 5-9 and 5-12 with equation 5-8 the rise time is expressed as:

$$t_r = \frac{0.9\, I_C\, (1/\omega_T + C_f R'_C)}{I_{B1} - I_C/2\beta_o}. \quad (5\text{-}13)$$

TABLE 5-2 — VALUES OF K_Q

JUNCTION TYPE	STEP ($n = 1/2$)	GRADED ($n = 1/3$)	EPITAXIAL ($n = 1/10$)
$K_{Q1}\begin{pmatrix} 0 - 90\% \\ V_F = 0.1\, V_I \end{pmatrix}$	1.52	1.31	1.08
$K_{Q2}\begin{pmatrix} 0 - 100\% \\ V_F = 0 \end{pmatrix}$	2.0	1.50	1.11

Note that if I_{B1} is large compared to the recombination term the expression becomes:

$$t_r = 0.9\, \frac{I_C}{I_{B1}}\left(\frac{1}{\omega_T} + C_f\, R'_C\right).$$

It is convenient to regard $0.9\left(\frac{1}{\omega_T} + C_f\, R'_C\right)$ as an **active region time constant** τ_A describing the time response to the 90% point. τ_A should not be confused with the active region *lifetime* τ_a. Divide both sides of the fraction in equation 5-13 by I_{B1}, substitute β_F for I_C/I_{B1}, and substitute τ_A for its equivalent expression to obtain a simple equation for rise time.

$$t_r = \frac{\beta_F\, \tau_A}{1 - \beta_F/2\beta_o}. \quad (5\text{-}14)$$

The τ_A term is a fundamental transistor property while the term $\frac{1}{1 - \beta_F/2\beta_o}$ is a current drive term or a universal rise time factor.

This expression assumes that $\beta_o = \beta$ and is constant over the load line traversed, and also that emitter efficiency is close to unity. These assumptions seldom cause significant error in practice.

It is interesting to compare these results to those obtained by various authors in the field using a formal method. The approach just outlined is essentially that used by Hwang, Cleverly, and Monsour.[6]

In the first and still the classic paper in the field, John Moll[7] described the transient behavior of junction transistors for the special case where a transistor is assumed to be linear, the effect of collector capacitance, C_{ob}, is negligible and a step of base current is applied to the input. Using his results but a different notation:

$$t_r = \frac{1}{(1-\alpha_N)\,\omega_\alpha} \ln \frac{I_{B1}}{I_{B1} - 0.9 I_C \left(\frac{1-\alpha_N}{\alpha_N}\right)}. \tag{5-15}$$

where
t_r = the rise time (0 to 90%)
α_N = common base forward current gain
I_{B1} = steady state base on current (magnitude of input current step)
I_C = steady state collector on current
ω_α = alpha cutoff frequency (rad/sec).

Easley[8] showed that the effect of C_{Tc} and the load resistor, R_L, could be included in most cases by the addition of a simple correction factor. Thus, he developed:

$$t_r = \frac{1 + \omega_\alpha R_L C_{Tc}}{(1-\alpha_N)\,\omega_\alpha} \ln \frac{I_{B1}}{I_{B1} - 0.9\, I_C \left(\frac{1-\alpha_N}{\alpha_N}\right)}. \tag{5-16}$$

When the $R_L\,C_{Tc}$ product is small compared to $\frac{1}{\omega_\alpha}$ equation 5-16 degenerates to equation 5-15.

In equation 5-16, C_{Tc} must be assigned the effective or average value C since it is a function of voltage as Bashkow[9] has pointed out. Narud[5] has generalized Bashkow's results as was indicated earlier.

Other authors have shown that ω_α should be modified by the factor 1.22. Using the charge control approach developed by Beaufoy and Sparks[1], Phillips[4] has illustrated that ω_T the gain-bandwidth product is a more accurate term to use than ω_α.

After reviewing the various literature, it has been concluded that the best expression for rise time is:

$$t_r = (\beta_o + 1)\left(\frac{1}{\omega_T} + R_L\,C_f\right) \ln \frac{I_{B1}}{I_{B1} - 0.9\, I_C/\beta_o} \tag{5-17a}$$

where:
$$\beta_o = \frac{\alpha_N}{1-\alpha_N}.$$

*Unity emitter efficiency means that injection from the base into the emitter is zero.

Since $\beta_F = I_C/I_{B1}$, and normally it can be assumed that $\beta_o \gg 1$, then equation 5-17a may be written as:

$$t_r = \beta_o \left(\frac{1}{\omega_T} + R_L C_f \right) \ln \frac{1}{1 - 0.9\, \beta_F/\beta_o}.$$

By making the substitution for τ_A and manipulating terms:

$$t_r = (\tau_A \beta_F) \frac{\beta_o}{0.9\, \beta_F} \ln \left(\frac{1}{1 - 0.9\, \beta_F/\beta_o} \right). \qquad (5\text{-}17b)$$

This equation is of the same general form as equation 5-14 except that the drive term or rise time factor is more complicated. That is, both equations can be written:

$$t_r = (\tau_A \beta_F)\, R \qquad (5\text{-}18)$$

where $R = \dfrac{1}{1 - \beta_F/2\beta_o}$ (assuming a linear rise time)

or $R = \dfrac{\beta_o}{0.9\, \beta_F} \ln \left(\dfrac{1}{1 - 0.9\, \beta_F/\beta_o} \right)$ (assuming an exponential rise time which is clipped).

Since the R term contains only the ratio of forced gain β_F to current gain β_o, it can be plotted as a function of this ratio as shown in Figure 5-8 . Notice that the two curves are nearly identical until β_o/β_F is less than 1.5. This amount of overdrive is almost always used, in which case the simple approximation is accurate enough for engineering calculations. Both R factors contain the assumption that β is constant over the load line. This is, of course, not true but the error introduced by a varying β is normally not significant.

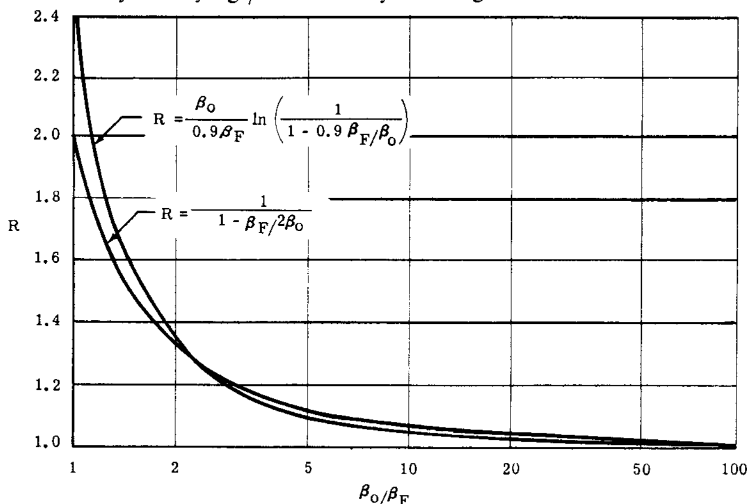


Figure 5-8 — Comparison of the Two Rise Time Factors

Equation 5-14, is readily solved for the value of β_F required to obtain a desired rise time, by simply manipulating terms.

$$\beta_F = \frac{t_r}{\tau_A + t_r/2\beta_o}. \tag{5-19}$$

SUMMARY OF RISE TIME CHARACTERISTICS: It has been shown that fundamentally the input circuit must deliver enough charge to:

1. Cause collector current to assume a new value.
2. Change the collector voltage.
3. Allow for recombination.

The charges involved in each of these functions were called Q_I, Q_V and Q_R respectively. Expressions were developed relating Q_I to ω_T (the gain-bandwidth product), and Q_V to C_{ob} (the collector output capacitance). The values of these charges depend upon the operating points traversed during the rise time interval (the load line). The time involved does not influence their value. However Q_R is time dependent and also very dependent upon the value of β which is influenced by the load line.

These charges were then related to a fundamental time constant τ_A and the effect of recombination (β) was incorporated in a drive factor R. By comparing the R factor derived rigorously, to that obtained from an intuitive approach which assumed that I_C was a linearly increasing function of time, it was shown that all approaches in the literature are essentially the same.

At this point, it might be well to question the usefulness of the R factors, as they were developed assuming that the input signal was a step of constant current which is never met in practice. However, if the amplitude of the input has reached its final value during the turn-on delay interval, then, during rise time, the transistor does see a constant level which is as though a voltage step were initially applied. The input current can be considered constant as long as $V_{BE} - V_{TF}$ is 1/10 or less of the applied signal and the impedance in the base is constant. In practice, this means that the input signal must be greater than 3 volts and the circuits must be resistance coupled, unless the coupling capacitor is so large that it can be considered a short-circuit during the turn-on period. Fortunately, a number of circuits do fulfill these requirements.

USE OF THE CHARGE AND TIME CONSTANT CHARACTERISTICS: In order to compute rise time, τ_A and β as a function of voltage and current must be known. If ω_T and C_{Tc} or C_{ob} information is given, or the rise time in a test circuit specified, then these data may be used to find τ_A. From the previously developed equations, it should be clear that the following highly usable relationship holds.

$$Q_A = \tau_A I_C = Q_I + Q_V = \frac{0.9\, I_C}{\omega_T} + \triangle V_{CB} C_f \tag{5-20a}$$

Q_A is the total active charge required by the transistor for I_C to reach 90% of its final value.

ω_T and C_f must be average or effective values over the load line. The following examples illustrate how rise time is determined using these principles.

EXAMPLE 1: Determine the rise time of a 2N964A transistor at an on collector current of 20 mA and an off collector voltage of 10 volts, when a step function of input current of 1.0 mA is applied. The junction temperature is 0°C. Solution:

From the curve of τ_A vs. I_C (Figure 5-9), at 20 mA, $\tilde{\tau}_A = 0.5$ nS and $\overline{\tau}_A = 1.2$ nS. However, this is given at a collector off level (V_1) of 5.5 V.

For this device, ω_T is relatively independent of V_{CB} but τ_A, as given by the curve, must be corrected for the extra charge due to the increase of $\triangle V_{CB}$. A graph of C_{ob} vs. V_{CB} is shown in Figure 5-10 which is taken directly from the data

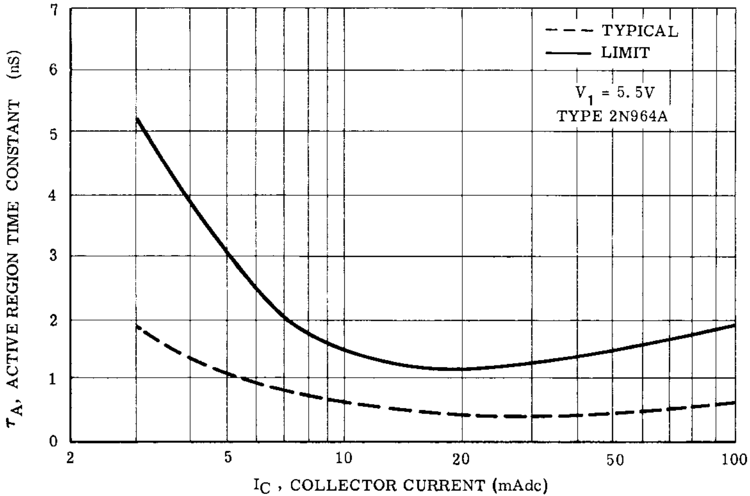


Figure 5-9 — Active Region Time Constant

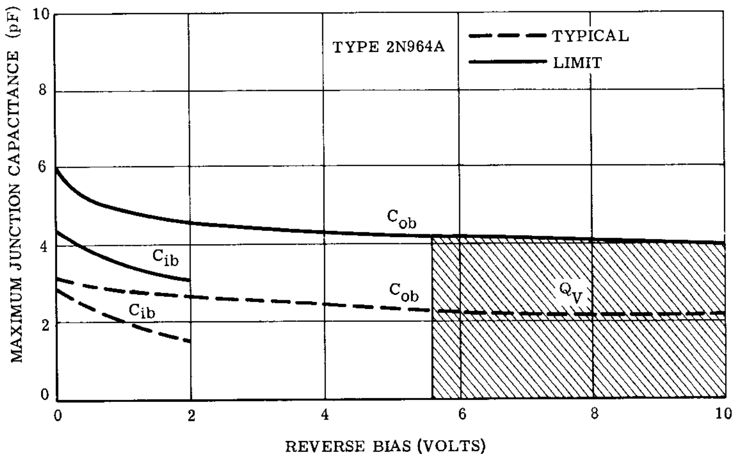


Figure 5-10 — Junction Capacitance Variations

sheet. From this graph, the amount of Q_V (shaded area) to be added is

$$\triangle\overline{Q}_V = (4.1)\,(4.5) = 18.4 \text{ pC}$$
$$\triangle\tilde{Q}_V = (2.3)\,(4.5) = 10.4 \text{ pC.}$$

At 5.5 volts and 20 mA

$$\overline{Q}_A = \overline{\tau}_A\, I_C = (1.2)\,(20) = 24 \text{ pC}$$
$$\tilde{Q}_A = \tau_A\, I_C = (0.5)\,(20) = 10 \text{ pC.}$$

At 10 volts and 20 mA

$$\overline{Q}_A = 18.4 + 24 = 42.4 \text{ pC}$$
$$\tilde{Q}_A = 10.4 + 10 = 20.4 \text{ pC.}$$

τ_A at 10 volts and 20 mA is

$$\overline{\tau}_A = \overline{Q}_A/I_C = \frac{42.4}{20} = 2.12 \text{ nS}$$

$$\tilde{\tau}_A = \tilde{Q}_A/I_C = \frac{20.4}{20} = 1.02 \text{ nS.}$$

The rise times are easily estimated by finding $\underline{\beta}_o$ and $\tilde{\beta}_o$ at 0°C. From the current gain curve, for this transistor, shown in Figure 5-11, $\underline{\beta}_o = 32$ at 20 mA and 0°C and the typical value is twice the minimum value as indicated in the tabular data at the rear of this handbook. The equation developed for rise time is

$$t_r = \tau_A \beta_F\, R.$$

From Figure 5-8, $\overline{R} = 1.5$ corresponding to a $\dfrac{\underline{\beta}_o}{\beta_F} = \dfrac{32}{20} = 1.6$

and $\tilde{R} = 1.19$ corresponding to a $\dfrac{\tilde{\beta}_o}{\beta_F} = \dfrac{64}{20} = 3.2.$

Therefore:

$$\overline{t}_r = \overline{\tau}_A \beta_F\, \overline{R} = (2.12)\,(10)\,(1.6) = 34 \text{ nS}$$
$$\tilde{t}_r = \tilde{\tau}_A \beta_F\, \tilde{R} = (1.02)\,(10)\,(1.19) = 12.1 \text{ nS.}$$

CALCULATING τ_A FROM CONVENTIONAL DATA: If the active region time constant τ_A is not specified, a value for τ_A can be approximated by using

$$\tau_A = \frac{.9\, I_C/\omega_T + \triangle\, V_{CB}(K_Q C_{Tc} + C_{sc})}{I_C}. \qquad \text{(5-20b)}$$

The gain-bandwidth product ($\omega_T = 2\pi f_T$) may be obtained from curves similar to the ones shown in Figure 5-12.

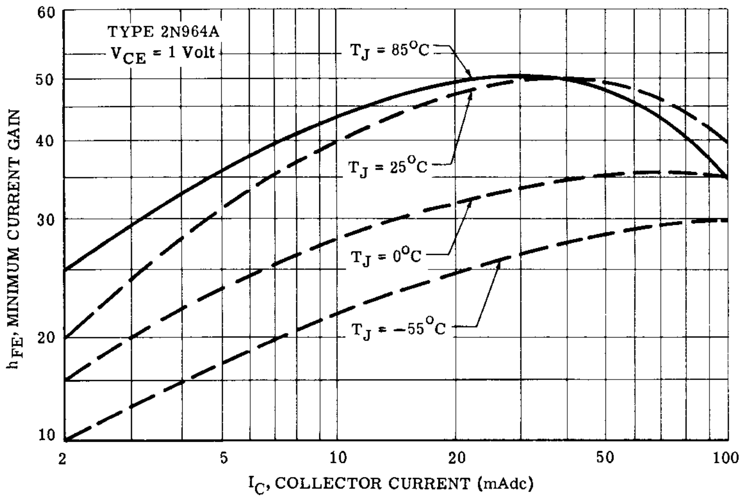


Figure 5-11 — Current Gain Characteristics

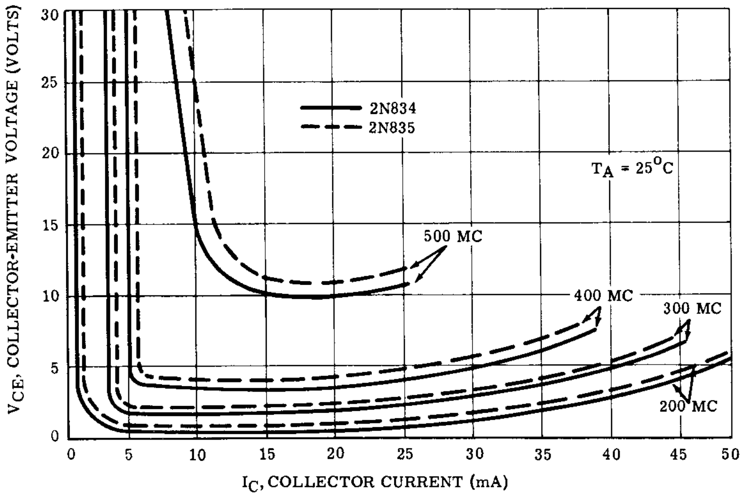


Figure 5-12 — Typical f_T Characteristics

The specified f_T curves on data sheets are usually typical curves and therefore, the value obtained must be normalized to obtain a minimum. Since behavior of ω_T with operating bias is not dependent upon its absolute value, a simple ratio can be used. Calculate ω_T from

$$\underline{\omega}_T \text{ (ave)} = [2\pi f_T \text{ (ave. from curves)}] \left[\frac{\underline{f}_T}{\tilde{f}_T} \text{(at point where } f_T \text{ is specified)}\right].$$

The collector capacitance can be obtained from data sheet charts similar to the one shown in Figure 5-13. Since these are usually typical curves, the data must be normalized to obtain a worst-case value of C_{Tc}. Since the variation of C_{Tc} with voltage is not dependent upon its absolute value, the worst case value can be calculated from

$$\overline{C}_{Tc} = [\tilde{C}_{Tc} \text{ (at desired condition)}] \left[\frac{\overline{C}_{Tc}}{\tilde{C}_{Tc}} \text{(at specific test condition)}\right].$$

The voltage $\triangle V_{CB}$ which represents the collector-base voltage at the 90% point of I_C requires some discussion. Figure 5-14 illustrates the voltages on the transistor at the beginning and end of the rise time interval. From the figure, it can be determined that

$$\triangle V_{CB} = V_{CB1} - V_{CB2} = V_1 - V_{TF} - SV_{CE} + V_{BE}. \qquad \text{(5-21a)}$$

The voltage V_{TF} is normally only 1 or 2 tenths of a volt less than V_{BE}, and SV_{CE} normally is in the range of 1 or 2 tenths of a volt; $\therefore V_{BE} \approx V_{TF} + SV_{CE}$. Thus, all of the terms cancel except V_1. Therefore, $\triangle V_{CB} \approx V_1$ for the full swing of I_C. At the 90% point used in computing rise time

$$\triangle V_{CB} \approx 0.9\ V_1 \qquad \text{(5-21b)}$$

with the high voltage point at $V_1 - V_{TF}$.

EXAMPLE 2: Determine τ_A for a 2N834 with a load line of $V_1 = 20$ volts and $I_C = 25$ mA using Figures 5-12 and 5-13.

SOLUTION:

1. Calculate Q_I to the 90% point.
   a) Determine the ratio $\underline{f}_T/\tilde{f}_T$.
   From the electrical characteristics chart, of the 2N834 @ 20 V & 10 mA, $\underline{f}_T = 350$ mcs, $\tilde{f}_T = 500$ mcs.
   $$\therefore \frac{\underline{f}_T}{\tilde{f}_T} = \frac{350}{500} = .7.$$
   b) Estimate an average $\tilde{f}_T$ over the load line from Figure 5-12. A reasonable estimate is $\tilde{f}_T = 420$ mcs.
   Then:
   $$\underline{\omega}_T = 2\pi\ (420)\ (0.7) = 1840 \text{ mcs}$$
   and $\overline{Q}_I = 0.9\ I_C/\underline{\omega}_T = 0.9\ (25)/1840 = 12.2$ pC.

2. Calculate Q_V.

   From the electrical characteristics chart
   $\overline{C}_{ob} = 4$ pF at 10V
   $\tilde{C}_{ob} = 2.4$ pF at 10V.

Subtracting the stray capacitance of 0.7 pF (See Table 5-1)

$$\overline{C}_{Tc} = 3.3 \quad \& \quad \frac{\overline{C}_{Tc}}{\tilde{C}_{Tc}} = 1.94$$
$$\tilde{C}_{Tc} = 1.7$$

Construct a typical C_{Tc} curve by subtracting the stray capacitance of 0.7pF. Multiply points on this curve by 1.94 to get a curve of C_{Tc}. Then add-in

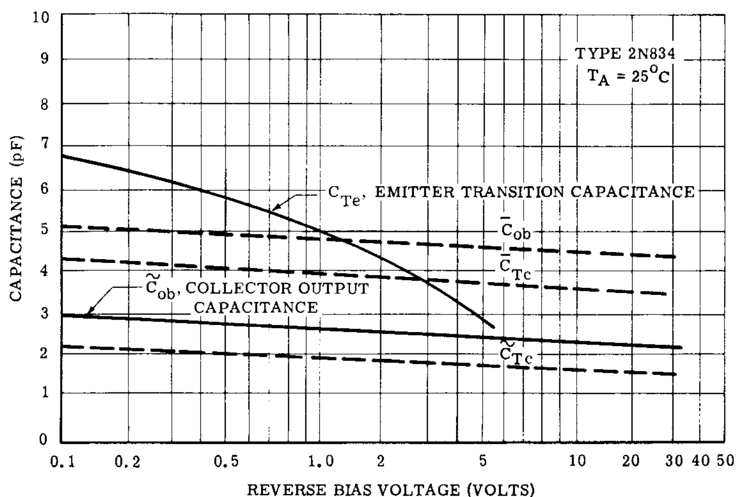


Figure 5-13 — Typical and Constructed Maximum Capacitance Variations

Figure 5-14 — Method of Determining ΔV_{CB}

the stray capacitance to obtain a curve of $\overline{C}_{ob}$. These are shown as broken line curves on Figure 5-13.

Since $V_1 = 20$ V, $\triangle V_{CB} = .9\ V_1 = 18$ V. Take $V_{TF} = 0.5$ V and find the area under the C_{ob} curve from 19.5 to 1.5 volts. From the constructed curve, $\overline{Q}_V = (4.6)(18) = 82.8$ pC.

3. Calculate $\overline{\tau}_A$.

$$\overline{\tau}_A = \frac{\overline{Q}_I + \overline{Q}_V}{I_C} = \frac{12.2 + 82.8}{25} = 3.8 \text{ nS.}$$

It should be evident that τ_A will vary somewhat with the load line. Figures 5-15a through 5-15d illustrate measured τ_A behavior for a number of transistor types. Notice that the two curves of τ_A vs. I_C, with different voltage, cross. Several effects are responsible for this behavior. As current is raised f_T decreases and becomes more affected by the collector voltage. This effect is shown on Figure 5-16, which indicates small signal f_T behavior. Since the average f_T over the load line is important in switching applications, the increase in τ_A due to a drop in f_T becomes apparent at currents much larger than the peak of small signal f_T shown in Figure 5-16. At high currents Q_I is much larger than Q_V, hence increases in V_1 manifest themselves in reducing Q_I while the increase in Q_V is negligible by comparison. At low currents the primary factor is C_{ob}; as shown in the previous example, the effect of ω_T is small. Thus τ_A becomes proportional to voltage.

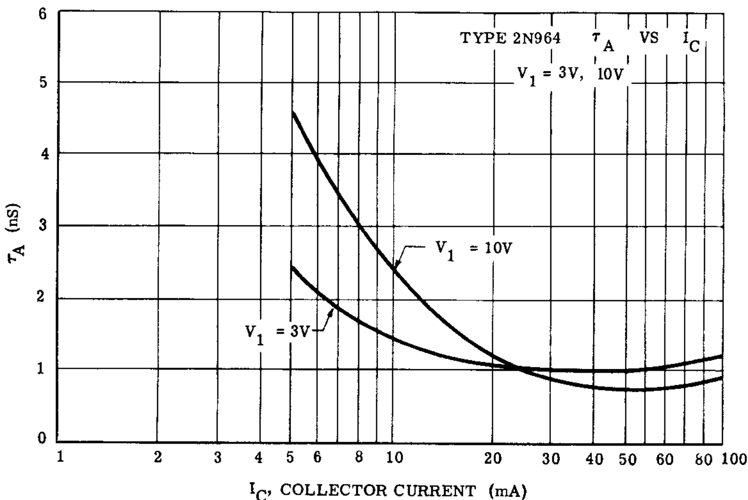


Figure 5-15a — Active Region Time Constant Characteristics for a 2N964 Transistor

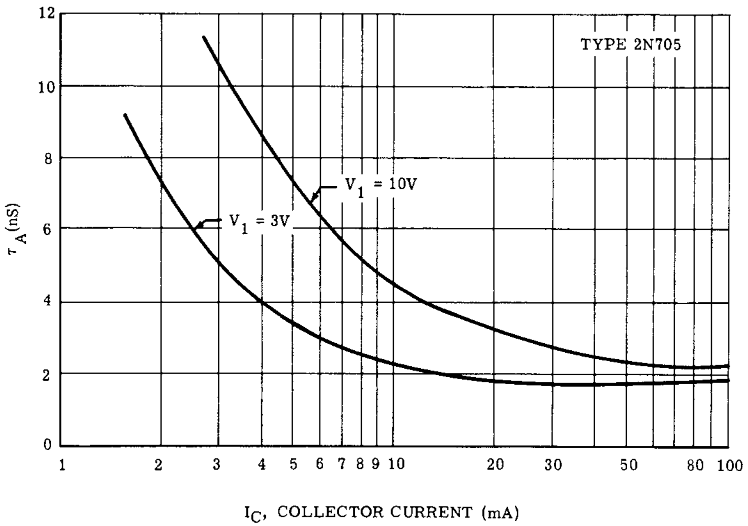


Figure 5-15b — Active Region Time Constant Characteristics for a 2N705 Transistor

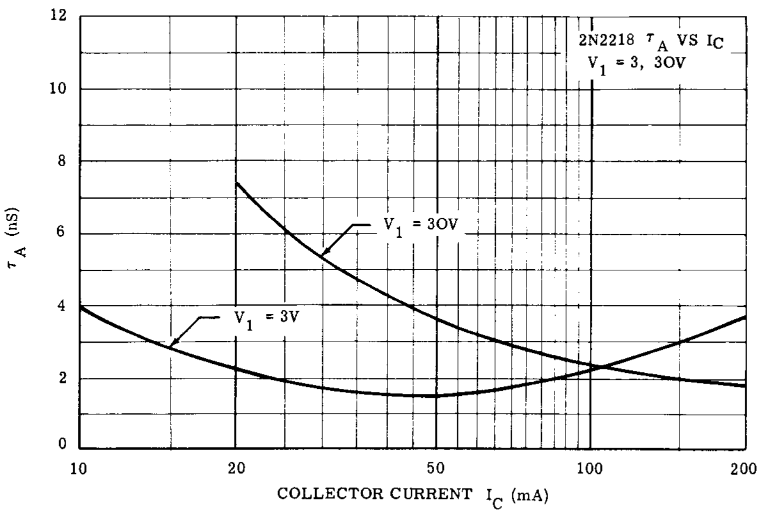


Figure 5-15c — Active Region Time Constant Characteristics for a 2N2218 Transistor

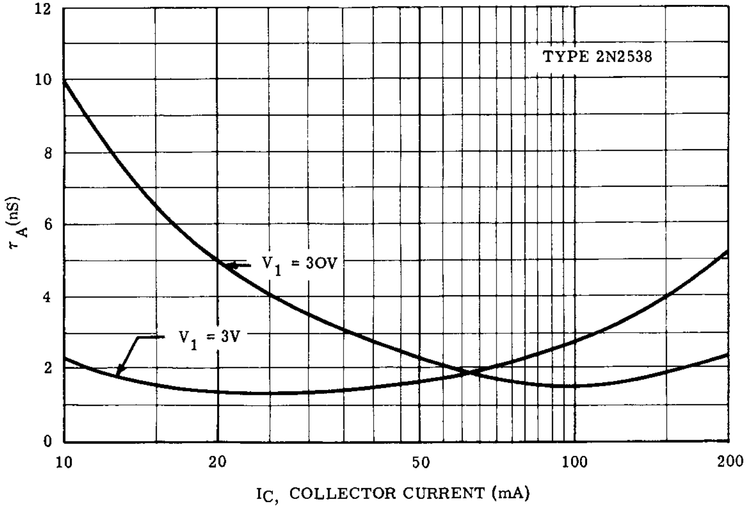


Figure 5-15d — Active Region Time Constant Characteristics for a 2N2538 Transistor

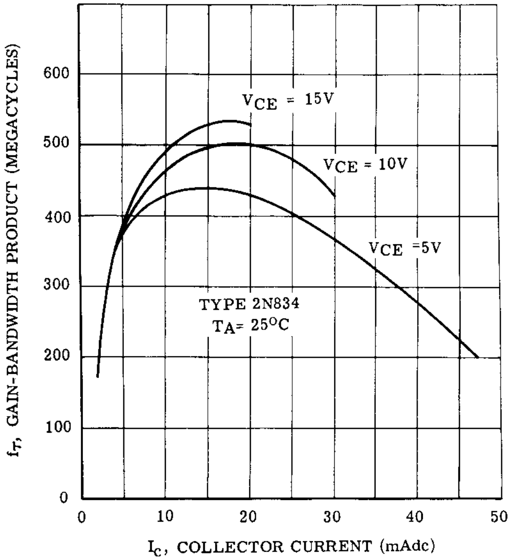


Figure 5-16 — Typical f_T Characteristics

5-5 — Storage Time

Storage time results from a transistor being driven into saturation by a turn on signal greater than that required to produce the collector current limited by $\frac{V_{CC}}{R_C}$. That is, $I_{B1} > I_C/\beta_o$. If the transistor is driven into saturation, the collector-base junction becomes forward biased and the collector begins to act like an emitter and injects carriers into the base. As a result of both junctions being forward biased, excess carriers, which form a charge, accumulate in the base, and also in the collector if its resistivity is appreciable. This excess charge must be removed for the transistor to turn off.

An expression for storage time can be obtained from the basic charge continuity equation. Since the voltages across the transistor junctions are constant during the storage time interval, the effects of C_{Te} and C_{Tc} need not be considered. The collector current does not change so charge associated with it does not affect storage time. The continuity equation during storage time becomes

$$\int_0^{t_s} i_B dt = \int_{Q_x}^{0} dq_x + \int_0^{t_s} \frac{q_x dt}{\tau_x} + \int_0^{t_s} \frac{q_a dt}{\tau_a}. \tag{5-22}$$

Equation 5-22 states that the input charge must equal the internal charges which are the following:

1. The excess charge caused by the presence of minority carriers in the base region. The amount of this charge, when on, will be called Q_x and is proportional to the excess base current I_{Bx} by the constant of proportionality, τ_x.
2. The recombination charge which is in two parts. The first, Q_{Rx}, is caused by recombination of the excess carriers having a time constant or lifetime τ_x. The second part is caused by recombination of the "active" carriers; that is, the carriers needed to maintain the collector current. This charge will be called Q_{Ra} and its lifetime τ_a.

   For a step of reverse base current I_{B2}

$$-I_{B2}\, t_s = -Q_x + Q_{Rx} + Q_{Ra}. \tag{5-23}$$

Let us now attempt to grasp an intuitive picture of what happens when a transistor is in saturation, and its behavior during the storage time interval. Consider the PNP transistor and its charge diagram shown in Figure 5-17b. When in saturation, an electron current (I_C/β_o) is drawn into the base as a result of recombination of the active charge (Q_a) having a lifetime (τ_a). Another electron current (I_{Bx}) is required to supply recombination of the excess charge (Q_x) having a lifetime (τ_x). The net emitter and collector currents are hole currents; if the direction of the hole current is taken as a reference (from + to —) then a current of $I_{Bx} + I_C/\beta_o = I_{B1}$ would flow out of the base terminal. However, it is important to remember that in reality, these are electron currents flowing into the base as shown in Figure 5-17a.

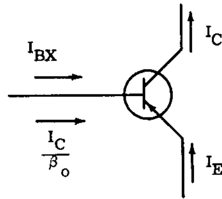


Figure 5-17a — Current Flow When a Transistor is in Saturation

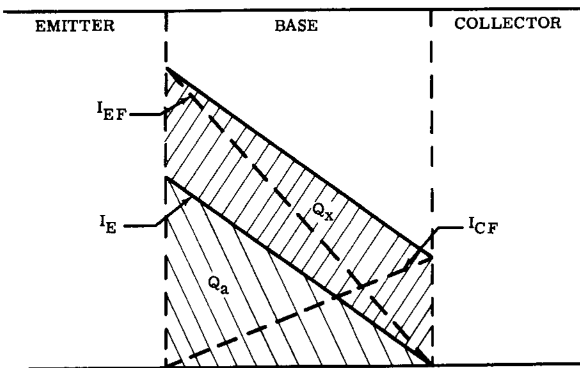


Figure 5-17b — Charge Distribution when a Transistor is in Saturation

If the base lead is opened, no base current can flow which means that turn-off is accomplished by internal recombination. In order to maintain I_C, a constant recombination current of $I_C/\beta_o = \frac{Q_a}{\tau_a}$ must flow. It flows internally by recombining with the excess carriers represented by Q_x. Since I_C/β_o is constant, during storage time

$$Q_{Ra} = \frac{t_s I_C}{\beta_o}. \tag{5-24a}$$

Furthermore, recombination within Q_x occurs causing an internal current to flow which is proportional to the amount of charge remaining, that is, $i_{Bx} = \frac{q_x}{\tau_x}$. In the absence of any other currents to decrease q_x (ie.,$I_C/\beta_o \approx$ o), the current i_{Bx} would be expected to decrease exponentially with time as shown in Figure 5-18. At $t = 0$, $i_{Bx} = I_{Bx}$ and decreases with a time constant of τ_x.

However, the presence of I_C/β_o hastens the recombination process and not only considerably shortens storage time, but makes the decrease of i_{Bx} with time become more linear.

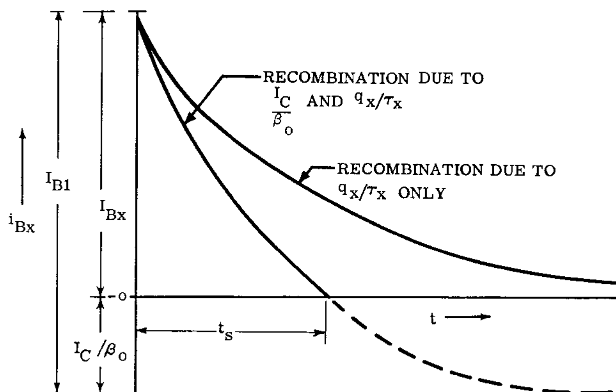


Figure 5-18 — Behavior of i_{Bx} with Time

Now suppose that instead of opening the base at time t = 0, the input voltage is changed instantaneously to a reverse potential. This potential would be positive for the PNP transistor shown in Figure 5-17a and cannot supply the electron currents required by recombination, instead it offers a field which repels the excess holes from the base to the emitter causing a current I_{B2} to flow. This causes q_x to decrease even faster with time than if the base were open-circuited, with a corresponding decrease in storage time. It also causes q_x to decrease more nearly linearly with time.

Assume that conditions are such that the decrease of q_x and i_{Bx} with time is linear. Referring to the charge diagram of Figure 5-19, it is seen that the excess recombination charge is

$$Q_{Rx} = \frac{I_{Bx}\, t_s}{2}. \tag{5-24b}$$

It was previously shown that $Q_{Ra} = t_s\, I_C/\beta_o$. Therefore, these relationships can be substituted into equation 5-23, and

$$-I_{B2}\, t_s = -Q_x + \frac{I_{Bx}\, t_s}{2} + \frac{I_C\, t_s}{\beta_o}. \tag{5-25}$$

Substituting $I_{Bx}\,\tau_x$ for Q_x and $I_{B1} - I_C/\beta_o$ for I_{Bx} and solving for storage time

$$t_s = \tau_x \frac{(I_{B1} - I_C/\beta_o)}{I_{B2} + \dfrac{I_{B1} + I_C/\beta_o}{2}}. \tag{5-26}$$

An interesting case occurs when $I_C/\beta_o \ll I_{B1}$, a worst-case condition in switching circuits. It is

$$t_s = \tau_x \frac{1}{\dfrac{I_{B2}}{I_{B1}} + \dfrac{1}{2}}. \tag{5-27}$$

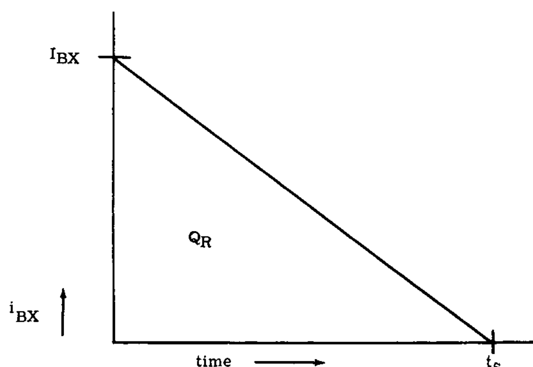


Figure 5-19 — Graphical Method of Determining Q when I_{BX} Decreases Linearly With Time

Storage time can be obtained in the general case by simply integrating the continuity equation. This leads to the expression often given in the literature [1, 4, 6, 7, 10]

$$t_s = \tau_x \ln \frac{I_{B1} + I_{B2}}{I_C/\beta_o + I_{B2}}. \tag{5-28}$$

Equations 5-26 and 5-28 give approximately the same result. This is shown by the various curves given in Figure 5-20. Note that agreement is very good when $I_{B1}/I_{B2} < 4$.

The object of deriving equation 5-26 was to show that by grasping an insight into the physical process, it is not too difficult to write equations which are fairly accurate when the charge control approach is used. When slide rules or tables are not handy, equation 5-26 can be used to determine a reasonably accurate storage time.

Considerable attention has been given to theory, in order to relate τ_x to device physics, and to measurement methods of obtaining τ_x.

In Moll's[7] analysis, a model of the transistor as two diodes back-to-back was used and it was found

$$\tau_{BS} = \tau_x = \frac{\omega_I + \omega_\alpha}{\omega_I\, \omega_\alpha (1 - \alpha_I \alpha_N)}. \tag{5-29}$$

Beaufoy and Sparkes[1] using a slightly different approach find

$$\tau_{BS} = \tau_x = \frac{1.22\,(\omega_I + \alpha_N \omega_\alpha)}{\omega_\alpha\, \omega_I\, (1 - \alpha_I \alpha_N)} \tag{5-30}$$

where τ_{BS} = Effective lifetime of excess carriers recombining in the base.
ω_I = inverse alpha cutoff frequency (transistor operated in inverted connection)
ω_α = forward alpha cutoff frequency
α_N = forward current gain
α_I = inverse current gain.

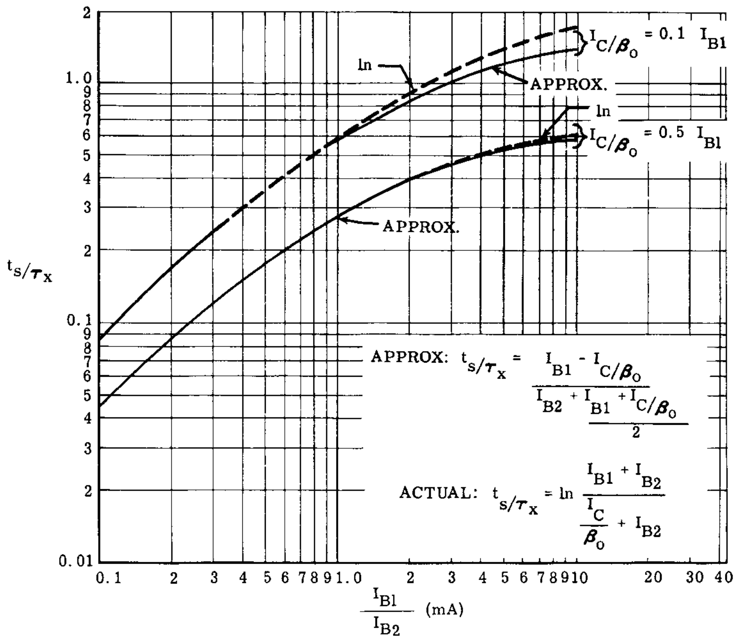


Figure 5-20 — Comparison of Storage Time Equations (Base Storage)

Equations 5-29 and 5-30 are not particularly helpful to the circuit designer as they involve measurement of four quantities. It is obviously more logical to make a direct measurement of storage time and calculate τ_x or use some techniques for measuring Q_x directly. Several approaches have been tried; these are summarized very well by Nanavati[10] whose conclusions seem supported by other workers in the field. In general storage time cannot be predicted very accurately at conditions far removed from a measurement point because of several effects which must be ignored in order to obtain a tractable solution for τ_x.

Some of these effects will briefly be considered.

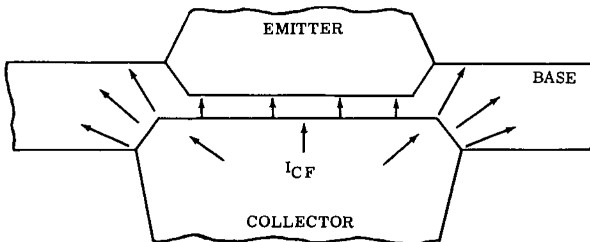


Figure 5-21 — Regions in an Alloy Transistor Base Which Serve as a Place for Stored Charge

If a device were perfectly symmetrical, then $\omega_I = \omega_\alpha$ and $\alpha_I = \alpha_N$. If α_N is close to unity, then Moll's equation yields $\tau_x = \beta/\omega_\alpha$ which is the active region lifetime τ_a. However, when fitting measured values to equation 5-28, it is found that $\tau_x \ll \tau_a$. Figure 5-21 shows a cross section of a typical alloy transistor. Notice that a large fraction of the charge injected from the collector could be stored in regions of the base distant from the emitter where surface lifetime could have an appreciable effect. The amount of charge in the area influenced by the surface would be dependent upon the bias conditions. The effect of the surface lifetime would affect the parameters of the Moll equation, but the method of measuring these to get a true value for τ_x is uncertain.

In mesa type* transistors, storage of charge in the collector can considerably modify these results. Phillips[4] has shown that the amount of collector charge can be lumped into that stored in the base. This approach essentially assumes that the carriers stored in the collector move back into the base before recombining. The additional charge modifies the value used for τ_x. Although this approach works well if the amount stored in the collector is small, in general other effects considerably complicate the picture.

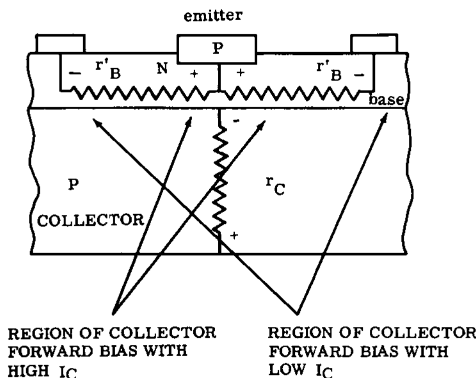


Figure 5-22 — Model of a Mesa Type Transistor

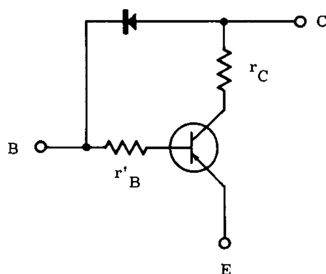


Figure 5-23 — Equivalent Circuit of a Mesa Type Transistor

Consider the model shown in Figure 5-22. For this discussion assume that I_{B1} is fixed. If I_C is low, the voltage drop across the collector series resistance r_C is negligible. The drop across r'_B causes the base terminal to be more negative than the area of the base under the emitter. Thus, as the transistor is driven into saturation and the collector becomes forward biased, the injected holes are attracted to the more negative region around the base terminal. The active base region, under the emitter, does not serve as a place for stored carriers. They are stored in the collector and in the base region near the base contact. More carriers are stored in the collector than in the base because it is made of higher resistivity material than the base, and its volume is many times larger. This same situation could also exist in an alloy transistor if r'_B is high as is normal in very high speed devices, but in this case storage would be confined to the high resistivity base region.

*The term "mesa types" is used to designate all transistors made with mesa, planar, and annular processes.

The equivalent circuit of Figure 5-23 describes this situation. It is similar to the Baker clamp circuit discussed in Chapter 2. As the magnitude of the voltage at the collector attempts to drop below the magnitude of the input voltage, the "diode" starts conducting and holds the "transistor" out of saturation. Storage time is now the time for the stored charge to exit from the "diode."

The behavior of the stored charge during the recovery or storage time of diodes is exceedingly complex. Lax and Neustadter[12] analyzed a particular class of diodes under conditions where a step of reverse current is applied and found

$$\text{erf}^* \sqrt{t_s/\tau_x} = \frac{I_F}{I_F + I_R}. \quad (5\text{-}31)$$

Using the equivalent circuit of Figure 5-23, I_F and I_R can be put in terms of I_{B1}, I_{B2} and I_C/β_0 an idea first published by Grinich and Noyce[13]. Thus, in the transistor, storage time is given as

$$\text{erf} \sqrt{t_s/\tau_x} = \frac{I_{B1} - I_C/\beta_0}{I_{B1} + I_{B2}}. \quad (5\text{-}32)$$

Experimentally it has been found that some transistors, under certain bias conditions, closely follow the relation of equation 5-32. However, as current is increased many mesa types begin to exhibit quite different behavior.

Referring to Figure 5-22, as collector current is increased, the drop across r_C increases altering the internal biasing to cause more of the collector injection to occur in the region directly under the emitter. More of the active part of the transistor is now in saturation causing storage time to be partly determined by the rate of diffusion of carriers, stored in the collector, back into the base. Some carriers would also be stored in the base, and recombination there would influence storage time. It should be clear that τ_x is not a constant and really is composed of the lifetime of various regions in a transistor. For simplicity the following definitions will be used:

$\tau_x = \tau_{BS}$, if measured transistor behavior can be fitted to equation 5-28 which assumes that all storage occurs in the base region between the emitter and collector.

$\tau_x = \tau_{CR}$, if measured transistor behavior can be fitted to equation 5-32 which assumes that all storage occurs in the collector or base region distant from the emitter, such that the equivalent circuit of Figure 5-23 applies.

Figure 5-24 shows data measured for a variety of transistors. This data is not intended to be representative of the given type but rather to show how an individual transistor behaves as current is varied. In each case, a value of τ_{BS} and τ_{CR} was computed at one measured point and the appropriate relationship plotted. Note that τ_{CR} is a larger number than τ_{BS} and that both of these "constants" vary significantly with bias. Neither relationship describes storage time behavior exactly.

*erf indicates the error function or probability integral. It is defined as: $\text{Erf}\,(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dt$
Tables are available for numerical values of this function.

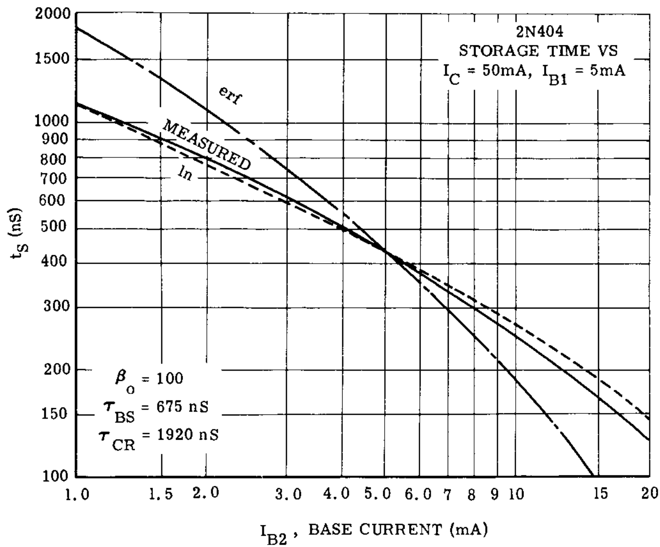


Figure 5-24a — 2N404 Storage Time Characteristics (Collector Current Constant)

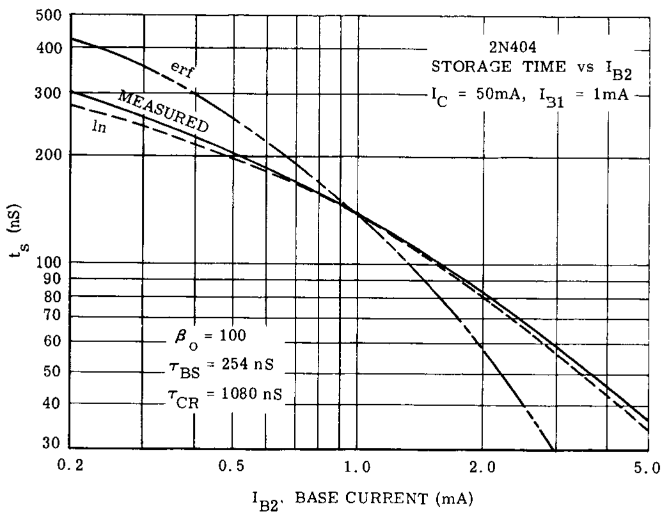


Figure 5-24b — 2N404 Storage Time Characteristics (Collector Current Constant)

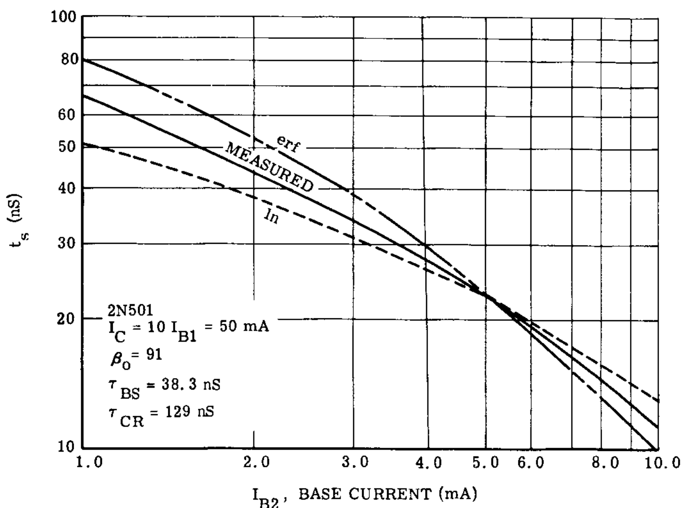


Figure 5-24c — 2N 501 Storage Time Characteristics (Collector Current Constant)

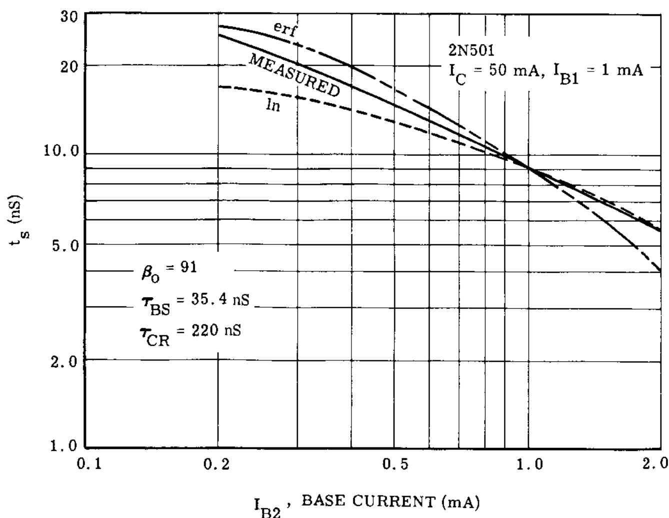


Figure 5-24d — 2N 501 Storage Time Characteristics (Collector Current Constant)

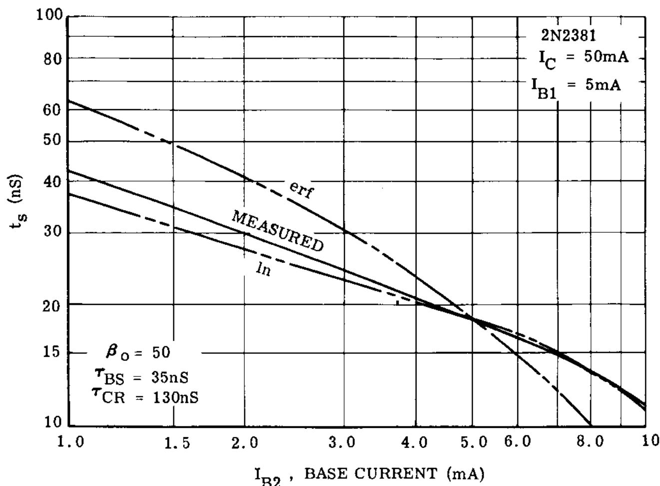


Figure 5-24e — 2N2381 Storage Time Characteristics (Collector Current Constant)

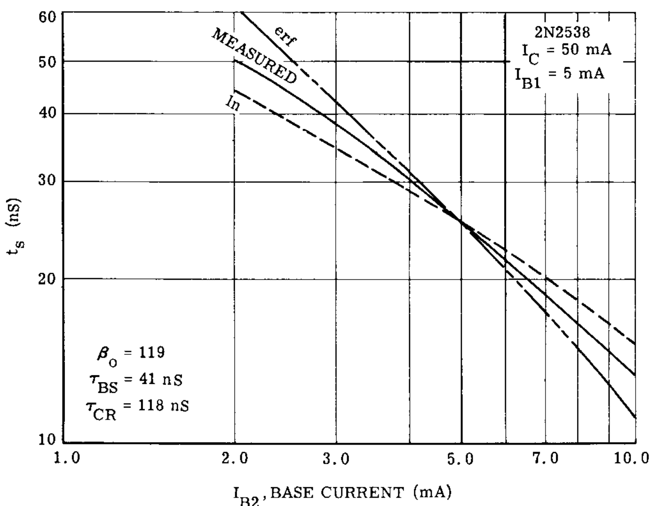


Figure 5-24f — 2N2538 Storage Time Characteristics (Collector Current Constant)

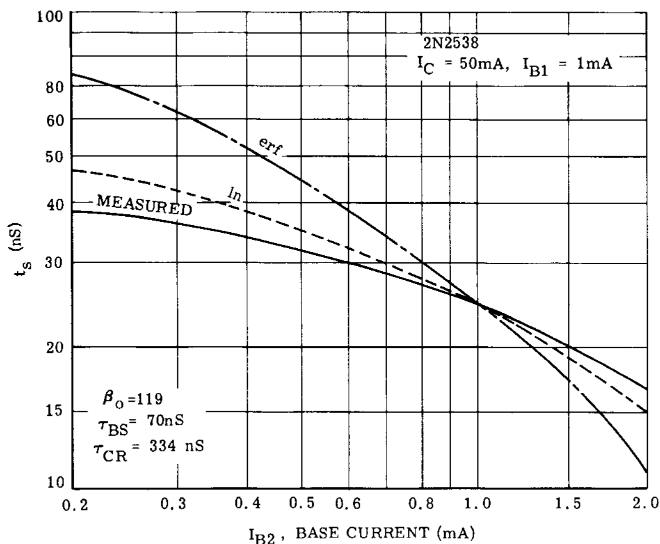


Figure 5-24g — 2N2538 Storage Time Characteristics (Collector Current Constant)

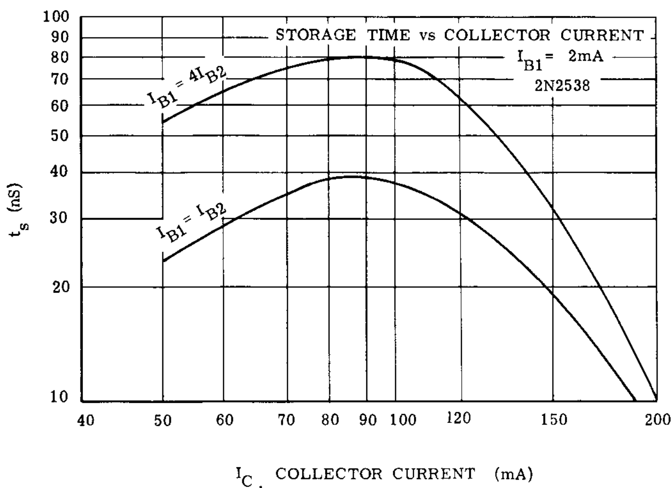


Figure 5-24h — 2N2538 Storage Time Characteristics (Base Current Constant)

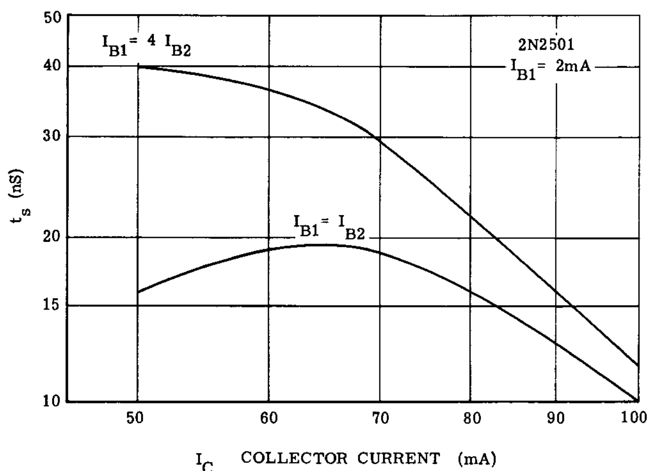


Figure 5-24i — 2N2501 Storage Time Characteristics (Base Current Constant)

TABLE 5-3 — IDENTIFICATION OF TRANSISTOR TYPES

Type Number	Relative Emitter Area	Material	Basic Process
2N404	large	germanium	alloy, uniform base
2N501	small	germanium	micro alloy, graded base
2N705	small	germanium	standard mesa type
2N706	small	silicon	standard mesa type
2N834	small	silicon	epitaxial mesa type
2N914	small	silicon	epitaxial mesa type
2N964	small	germanium	epitaxial mesa type
2N968	small	germanium	standard mesa type
2N2218	large	silicon	epitaxial mesa type
2N2381	large	germanium	standard mesa type
2N2538	large	silicon	epitaxial mesa type
2N2501	small	silicon	epitaxial mesa type

Note: "Mesa type" refers to transistors made by the mesa, planar or annular processes. All mesa type transistors have a graded base.

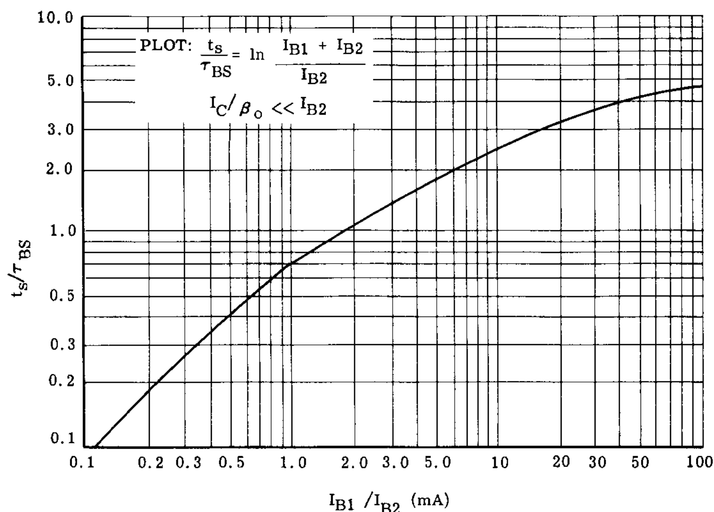


Figure 5-25 — Normalized Base Storage Time vs. Circuit Drive Ratio

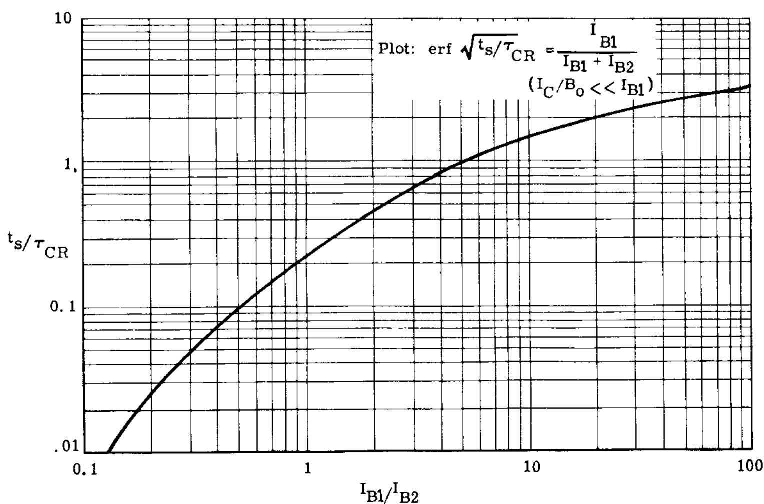


Figure 5-26 — Normalized Collector Storage Time vs. Circuit Drive Ratio

For transistor characterization purposes it is normally desirable to select I_C so that I_C/β_o is small compared to either I_{B1} or I_{B2}, whichever is appropriate, in order to minimize the effect of β_o on t_s. Then the measured t_s will bear a truer relationship to τ_x and for a given circuit drive ratio, I_{B1}/I_{B2}, a constant proportionality will exist between τ_x and t_s. Figure 5-25 illustrates this for the case of base stored charge while Figure 5-26 shows collector stored charge. These curves will be useful to predict storage time as I_{B2} is varied from the value where storage time is

measured. As I_{B2} is reduced from the test value, the curve will yield a pessimistic value for t_s because I_C/β_o is neglected. For the same reason as I_{B2} is increased the predicted value will be optimistic.

The time constants τ_{CR} and τ_{BS} are quite different for identical storage times. That is, for a particular ratio of I_{B1}/I_{B2}, $\tau_{CR} > \tau_{BS}$ by an amount dependent upon the I_{B1}/I_{B2} ratio. This is shown in Figure 5-27 which shows the ratio of τ_{CR}/τ_{BS} as a function of I_{B1}/I_{B2} and is obtained by dividing ordinates from Figure 5-25 by ordinates from Figure 5-26 for a given I_{B1}/I_{B2}. Note that $\tau_{CR} > 1.5\,\tau_{BS}$ for all drive ratios.

For simplicity, it will often be convenient to assume that storage time behavior is described adequately by the ln relationship and regard τ_{BS} as fairly constant. Even though errors will be inevitable, several important concepts can still be developed.

As with delay and rise time, the various storage time equations derived were based upon the assumption that the input signal was a step function of current which remains constant during the storage time interval. In practice this means that the rise time of the input pulse must be short compared to the storage time interval. Since v_{BE} is virtually constant during storage time, any resistance driven circuit normally fulfills the constant current assumption. Analysis reveals that the worst effect of a finite transition time, t_1, between I_{B1} and I_{B2} is to lengthen the storage time by t_1 but in most cases, the increase in t_s is much smaller than t_1.

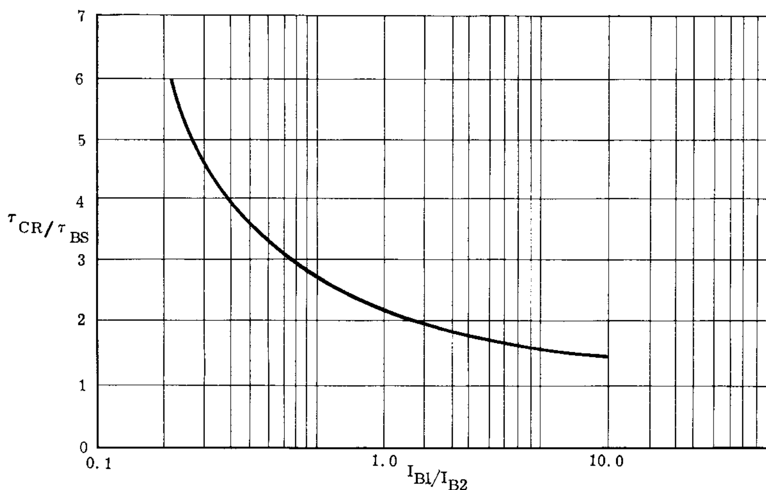


Figure 5-27 — Ratio of Collector Recovery Time Constant to Base Storage Time Constant

5-6 — Fall Time

Since the fall time occurs in the active region, τ_A, Q_I, and Q_V apply. However, recombination aids turn-off. In the same manner as the rise time expression, (equation 5-13), was developed, it is found that

$$t_f = \frac{0.9\, I_C \left(\frac{1}{\omega_T} + R'_C\, C_f\right)}{I_{B2} + I_C/2\beta_o}. \tag{5-33}$$

By defining and substituting a cut-off gain $\beta_C = I_C/I_{B2}$ and also substituting τ_A for its equivalent quantity

$$t_f = \frac{\tau_A\, \beta_C}{1 + \beta_C/2\beta_o}. \tag{5-34}$$

In a manner analogous to the rise time factor, the denominator may be regarded as a fall time factor, F.

After reviewing the literature* it was concluded that the best expression for fall time is

$$t_f = \left(\frac{1}{\omega_T} + R'_C\, C_f\right)(\beta_o + 1)\ln\left(\frac{\beta I_{B2} + I_C}{\beta I_{B2} + 0.1\, I_C}\right). \tag{5-35}$$

By substituting β_C and τ_A for their equivalent quantities this equation may be written (if $\beta_o \gg 1$) as

$$t_f = \tau_A\, \beta_C \left[\frac{\beta_o}{0.9\,\beta_C}\, \ln\left(\frac{1 + \beta_o/\beta_C}{0.1 + \beta_o/\beta_C}\right)\right] \tag{5-36}$$

where the bracketed term is the F factor.

Both of these factors are plotted in Figures 5-28a and b (a and b differ only in the scale used for the ordinate) where it is seen that the two factors do not differ substantially whenever $\beta_o/\beta_C > 0.1$. The worst case occurs when $\beta_o \gg \beta_C$ making $F \approx 1$.

Whenever Q_x is all stored such that it can exit during storage time these fall time expressions hold. However, storage in regions removed from the emitter considerably changes turn-off behavior because not all the excess charge leaves during storage time. The decay of the remaining charge would be determined by the lifetime of the minority carriers in the region, yielding behavior similar to that of a diode. The mathematics describing the turn-off of a diode are involved; the situation becomes exceedingly cumbersome when the fall time of the transistor is also occurring at the same time. No formulas have been developed which may be used to predict fall time in mesa transistors in situations where the charge stored in the collector is a considerable fraction of the active charge $\tau_A I_C$. Thus, fall times for mesa transistors may contain a few surprises for designers accustomed to working with standard alloy types. Some data and explanations for the measured behavior will be given in the next section when total control charge is discussed.

The fall time expressions derived assumed that I_{B2} was in the form of a constant current step. This situation is normally found in the case of resistance driven circuits because the change from I_{BI} to I_{B2} occurred during the storage time; thus, the input voltage is constant during the fall time. As long as $V_{in} \gg (V_{BE} - V_{TF})$ then I_{B2} will be essentially constant during the fall time interval.

*See rise time references

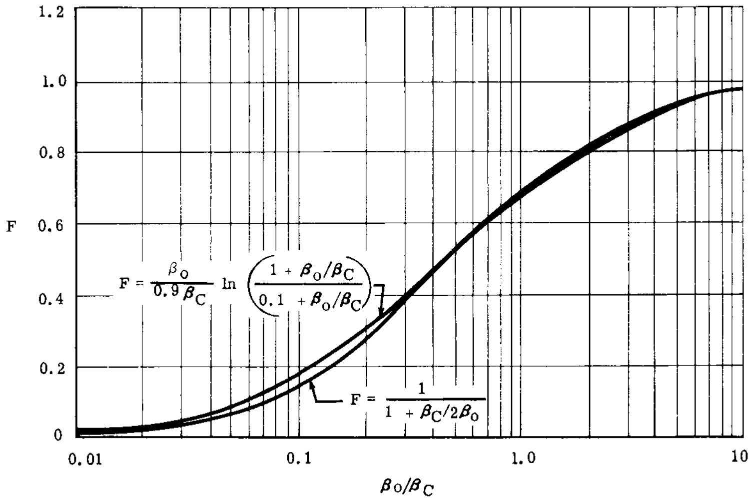


Figure 5-28a —Fall Time Factor vs. Overdrive Factor

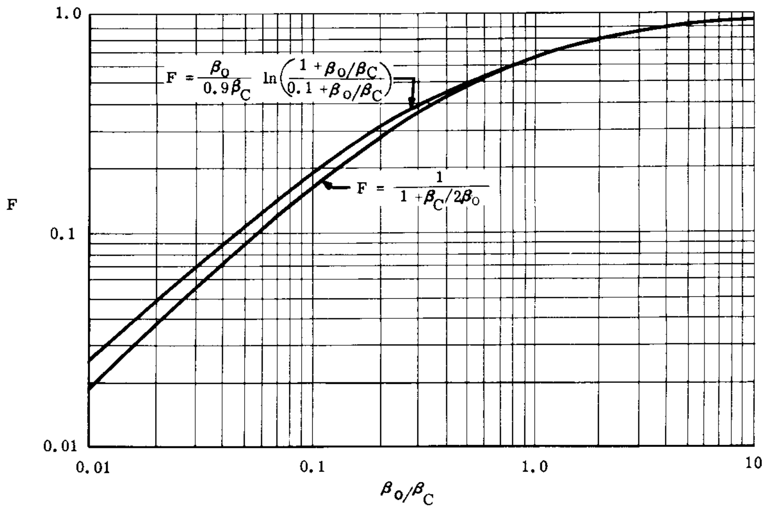


Figure 5-28b —Fall Time Factor vs. Overdrive Factor

5-7 — Total Control Charge (Q_T)

Previously, it was shown that to turn on a transistor from the threshold of conduction to 90% of the final value of collector current involved movement of a charge $\tau_A I_C$. This same amount of charge is also involved during the fall time. When driven into saturation, excess charge Q_x results. Thus, to turn off the transistor the total stored charge Q_s must be removed, which is given as

$$Q_s = Q_x + \tau_A I_C. \tag{5-37}$$

A popular method of measuring Q_s is shown in Figure 5-29. The capacitor (C) is adjusted to the minimum value which will produce a turn off waveform similar to the one shown in Figure 5-30 where $C = C_{opt}$. The optimum capacitance is obtained when the "bumps" just disappear. The charge on the capacitor which will be called Q_T is simply

$$Q_T = C_{opt} \triangle V_{in}. \tag{5-38}$$

It is necessary to determine if Q_T actually represents the total stored charge Q_s of the transistor.

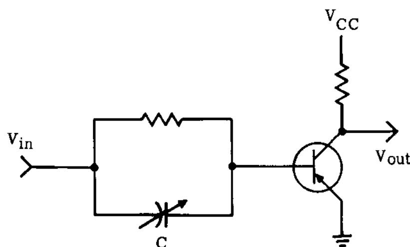


Figure 5-29 — Q_T Test Circuit

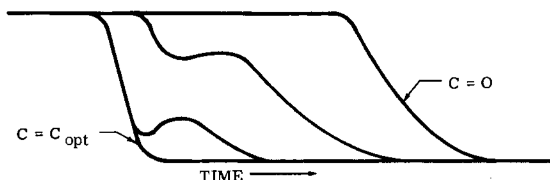


Figure 5-30 — Turn-Off Waveform (PNP Transistor)

ALLOY TRANSISTOR BEHAVIOR: In the alloy transistor it was shown that, during storage time, Q_x is reduced due to recombination of excess carriers with a lifetime τ_x, recombination of excess carriers with a lifetime τ_a and movement of charge due to the reverse current in the base-emitter junction. Following depletion of the excess charge, the active charge $\tau_A I_C$ is reduced by recombination of the active carriers and movement of charge due to I_{B2}. Thus, to measure the stored charge, conditions must be such that the loss of charge due to recombination is negligible. This means that turn-off must be very rapid which is synonymous with making I_{B2} very high.

The test circuit can fulfill this requirement if the charge (Q_T) stored on C is larger than Q_s and the resistance of the source (R_S) is low. The base current is given by

$$i_{B2} = \frac{v_C - V_{BE}}{R_S + r'_B}. \tag{5-39}$$

The voltage (v_C) on the capacitor will decrease exponentially with time as Q_x is reduced. There must be sufficient voltage across the capacitor to insure a high i_{B2} during fall time. If not, recombination will be the chief mechanism determining turn-off and it will be slow.

Suppose that the voltage on C fell to zero at some point during the fall time. Then, i_{B2} would be small and the fall time would exhibit a long "tail". In most cases, however, this condition also results in the bumps as shown on Figure 5-30. This is undoubtedly caused by excess carriers stored in places remote from the active base-emitter region, which are removed at a slower rate than the normal excess carriers. Since turn-off of collector current has already commenced, these carriers constitute an active charge when they enter the active base-emitter region and result in an increase of collector current.

It is easy to see that Q_T can provide a good measure of Q_s in alloy devices. However, this does not necessarily apply to mesa type devices. In order to understand turn-off of mesa type devices, the turn-off mechanism in diodes must be studied in detail.

DIODE BEHAVIOR: When driven from a current source the turn-off transient response of a diode may be divided into two phases, the constant current phase and the recovery phase. Immediately following reversal of the input signal, the diode continues to exhibit a low impedance, approximately equal to that which it had in the forward direction. During this period the current is determined by the external circuit. This period may be so short in some diodes that it is not measurable. To fulfill the constant current requirement the source voltage need only exceed a few volts and the source resistance need only exceed a few tens of ohms.

The constant current phase continues until all the excess carriers are removed from the region immediately adjacent to the junction, at which time the impedance of the junction starts to rise rapidly to a very high value, and the recovery phase begins. During the recovery phase, the impedance of the diode is so high that in a practical case, the current cannot be controlled by the external circuit; that is, the diode current is essentially independent of diode voltage. The voltage across the diode is determined entirely by the voltage drop across the source impedance, and approaches the source supply as the current decreases. The waveforms depicting diode turn-off behavior are shown in Figure 5-31.

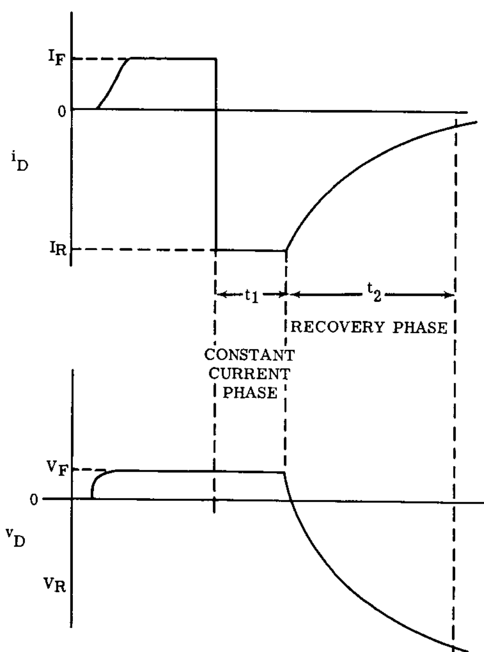


Figure 5-31 — Diode Turnoff Wave Forms

A qualitative picture of the turn-off process may be gained by examining the sketches of charge distribution shown in Figure 5-32. Figure 5-32 shows charge distribution, i.e., number of stored carriers as a function of the distance from the junction. The stored charge is virtually all confined to the high resistivity side of the junction.

The ordinate P indicates the number of carriers and the abscissa is distance x normalized to the diffusion length L. Diffusion length is the average distance a carrier can travel before recombining. The curve $t/\tau = 0$, shows the charge distribution in a PN junction which has been conducting a forward current (I_F) for a time which is much longer than the lifetime (τ) so that a state of charge equilibrium is attained. Since conduction is by diffusion, charges move from regions of high charge density to those of lower density. Current is in the direction of, and is proportional to, the gradient of the charge density. Therefore, at the junction, the slopes of the lines are proportional to the diode current.

In Figure 5-32a, the diode is open-circuited at $t = 0$, therefore the charge gradient at the junction ($x/L = 0$) drops immediately to zero and the charge decays solely by internal recombination as $e^{-t/\tau}$. In Figure 5-32b the initial forward biased conditions are the same as in Figure 5-32a but the diode is switched such that $I_R = I_F$. Charge distributions during the constant current phase are shown as solid lines, those during the recovery phase as dotted lines. Upon reversal of input voltage polarity in Figure 5-32b, the slope at the junction ($x/L = 0$) during the constant current phase is the negative of the slope of the line at $t = 0$.

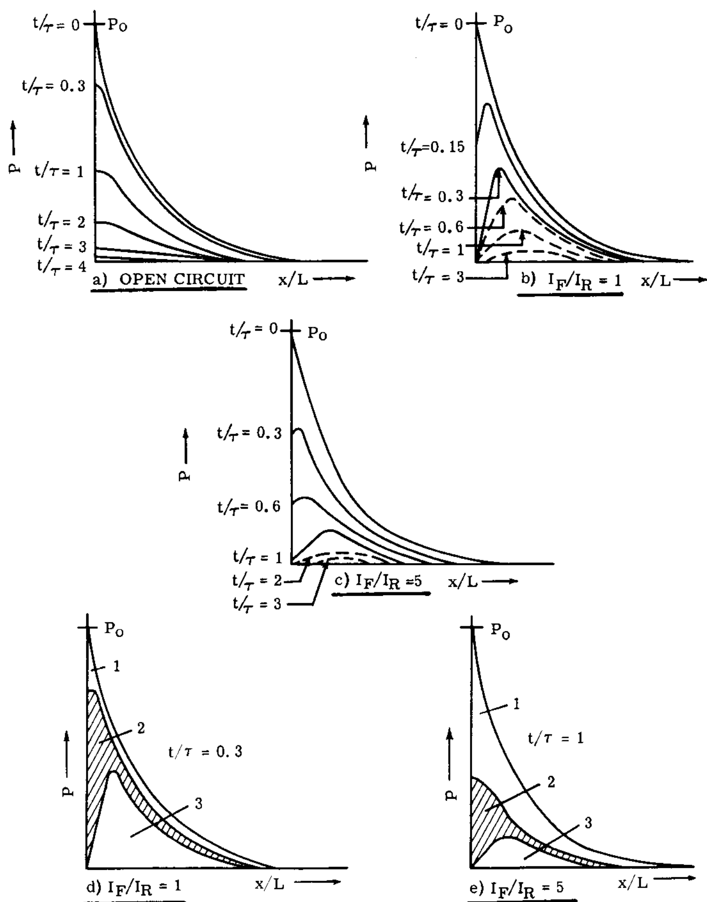


Figure 5-32 a,b,c — Charge Distribution in a Diode
d,e — Charge Removed During Constant Current Phase

The slope, and therefore I_R, remains constant until the charge concentration is zero at the junction. Then the slope decreases during the recovery phase as the diode current reduces to zero. In Figure 5-32c the diode is switched such that $I_R = I_F/5$. Note that the slope at the junction is less than that shown in Figure 5-32b.

Figures 5-32d and 5-32e compare the charge distributions at the end of the constant current phase — which represents storage time — for a given current ratio with what the distribution would have been had charge been lost by recombination alone. The difference between the areas of these two distributions must be proportional to the charge $I_R t_s$ which was removed due to external current flow. For example, from Figure 5-32b, it is seen that $t/\tau = 0.3$ at the end of the

constant current phase. The boundary between areas 2 and 3 of Figure 5-32d is the charge distribution line at $t/\tau = 0.3$ taken from Figure 5-32b. The boundary between areas 1 and 2 is the charge distribution line at $t/\tau = 0.3$ as indicated on Figure 5-32a. Area 1 therefore, represents the charge lost by recombination, area 2 the charge which appeared in the external circuit, and area 3 the charge left in the diode, at the end of the constant current phase under conditions of $I_R = I_F$. In a similar manner, Figure 5-32e was constructed.

In Figure 5-32d, since t_s is less, much less recombination has taken place than in Figure 5-32e. Similarly, less time is available for charge to diffuse toward the junction so that I_R consists mainly of the removal of charge which was relatively close to the junction at $t = 0$. As reverse current is increased further, t_s becomes smaller, the slope of the line where $P = 0$ becomes steeper and rises higher. As $I_R \rightarrow \infty$, $t_s \rightarrow 0$ and the charge available for conduction reduces to a vanishingly thin layer next to the junction. Thus, the charge removed ($t_s I_R$) approaches zero.

The total charge in the diode is given by the area under the $t = 0$ curve. This curve represents the carrier concentration P and is given by

$$P = P_0 e^{-x/L}$$

where

P_0 is the concentration at the junction under forward bias conditions
x is the distance in the junction
L is the diffusion length for carriers, which is the average distance a carrier will travel before recombining.

Integrating this curve from zero to infinity and combining the result with other properties of the junction, it is found that

$$Q = \tau I_F. \tag{5-40}$$

As illustrated, during storage time this charge never appears completely in the external circuit and external circuit charge approaches zero as I_R approaches infinity. Therefore during a Q_T measurement, all of the charge that appears in the external circuit must exit during the *recovery phase* of the diode.

This amount of charge can be calculated and is found to be

$$Q_\infty = K_D \tau I_F \tag{5-41}$$

where

Q_∞ signifies the charge which appears in the external circuit as $I_R \rightarrow \infty$ and is the maximum charge which can ever appear in the external circuit. K_D is always equal to or less than 0.5.

The value of K_D depends upon the ratio W/L where W is the physical width of the junction. The charge which does not appear in the external circuit is lost by internal recombination. *The total charge given by equation 5-40, never, under any circumstances, appears in the external circuit.*

MESA TRANSISTOR BEHAVIOR: When the results of the discussion on diodes are extended to the mesa transistor during the Q_T test, the following facts become apparent.

(1) The charge which exits during the storage time interval, consists of excess carriers from the base. Storage time is very short since the charge stored in the base is very small and I_{B2} is high.

(2) During the fall time, which represents the recovery phase of the "diode", the diode stored charge exits. This charge is always less than the total charge stored in the diode. Also, $\tau_A I_C$ exits during fall time.

(3) The Q_T measurement, though it does not yield a value equal to the total stored charge Q_s, does give a value of charge which is the maximum that can ever occur in the external circuit. As such, it is a valuable piece of design information.

In circuits that are resistance driven, the fall time of mesa transistors will increase over that predicted by the classic fall time formulas as the turn-off drive is increased, because more of the stored charge in the diode must exit during the recovery phase. Increasing the turn-on drive will also increase the fall time because this increases the diode stored charge Q.

The theoretical fall time behavior can be written from equation 5-33 as

$$t_f = \frac{\tau_A}{I_{B2} + I_C/2\beta_o}. \tag{5-42}$$

If t_f is plotted vs. $1/(I_{B2} + I_C/2\beta_o)$ a straight line results. Therefore, if measured data is plotted in this manner, any deviations between classical theory and the actual case become obvious. The data in Figure 5-33 shows that equation 5-42 is an approximation to the true behavior, and the expected deviations are evident.

At the present time, it is not possible to accurately predict the turn-off behavior of mesa type transistors. However, considerable information can be gained from the Q_T specification.

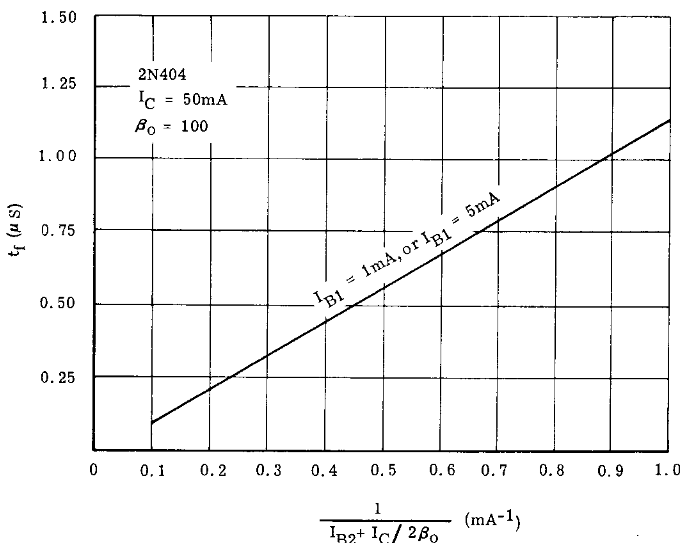


Figure 5-33a — 2N404 Fall Time Characteristics

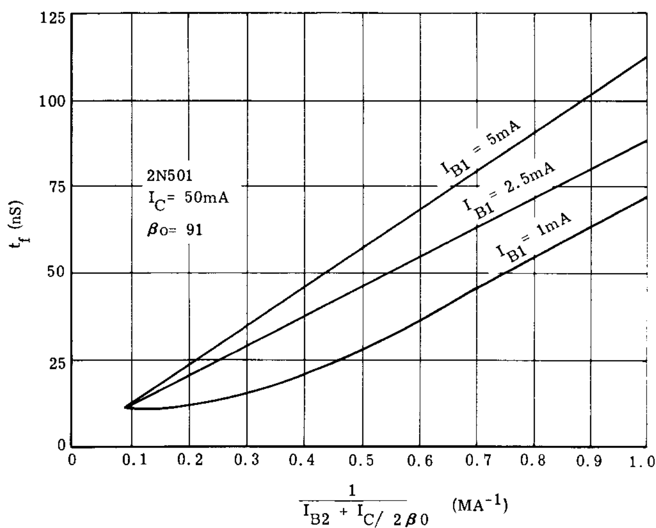


Figure 5-33b — 2N501 Fall Time Characteristics

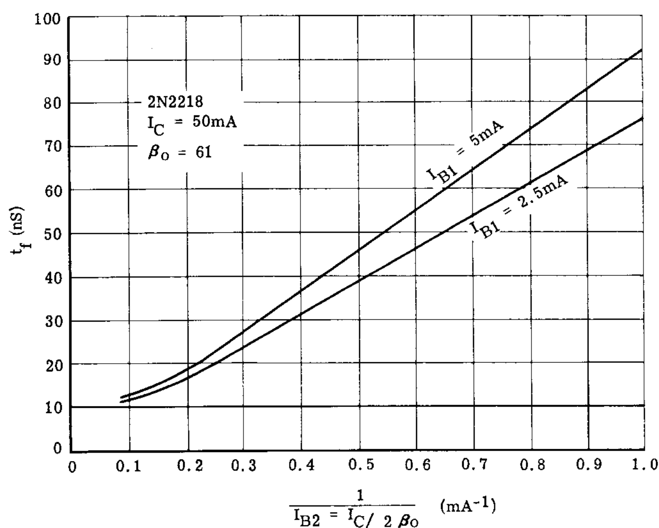


Figure 5-33c — 2N2218 Fall Time Characteristics

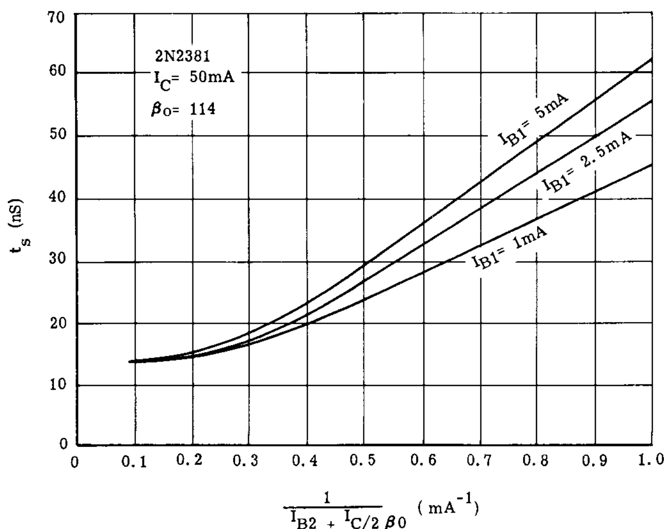


Figure 5-33d — 2N2381 Fall Time Characteristics

Figure 5-34 shows Q_T data for several representative transistors where β_F is a constant. This data can be used to find Q_T at other values of β_F. Rewriting equation 5-37 and substituting $K_T(I_{B1} - I_C/\beta_o)\tau_x$ for Q_x (where K_T indicates the fraction of Q_x which appears in the base circuit) then

$$Q_T = K_T I_{B1} \tau_x + I_C(\tau_A - K_T \tau_x/\beta_o). \tag{5-43}$$

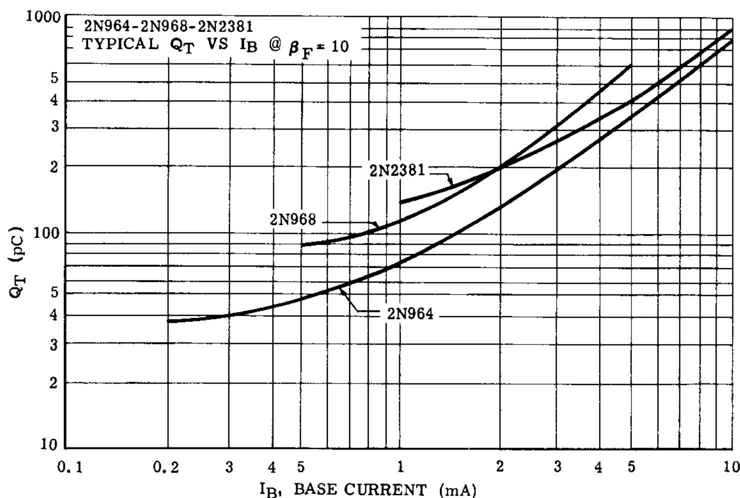


Figure 5-34a — Total Control Charge Data

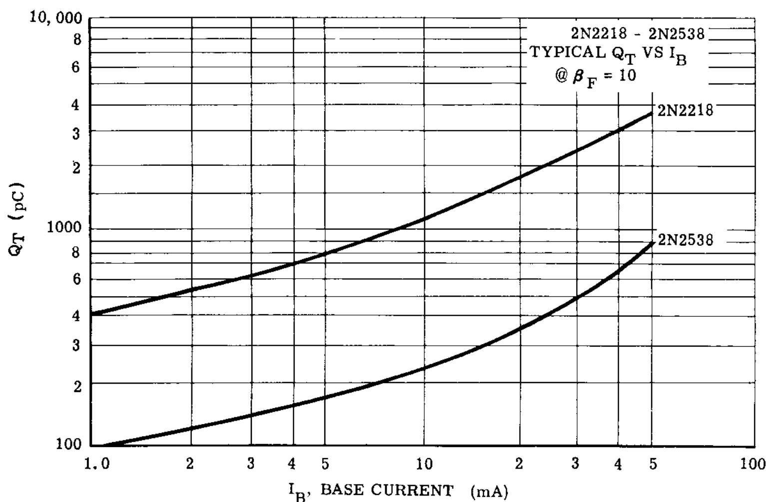


Figure 5-34b — Total Control Charge Data

The time constant (τ_x) varies with I_{Bx} and τ_A varies with I_C and V_{CC}. Therefore, if Q_T is to be known exactly at other operating points, all variations of these characteristics must be known because Q_T could increase or decrease as I_C is increased, depending upon the relative values of K_T, τ_x, τ_A and β_o, as the curves of Figure 5-35 illustrate.

To find a worst-case possibility, assume $K_T\tau_x/\beta_o \ll \tau_A$. Let the data from the Q_T curve be denoted by a subscript 1 and data at some new collector current by a subscript 2. Then

$$Q_{T2} = Q_{T1} + \tau_{A2}I_{C2} - \tau_{A1}I_{C1}. \tag{5-44}$$

Thus τ_A and Q_T curves can be used to obtain Q_T at any condition, which is very useful information for the design of RCTL circuits. Use of equation 5-44 with Q_T and τ_A data is shown in detail in the RCTL example inverter of Chapter 7.

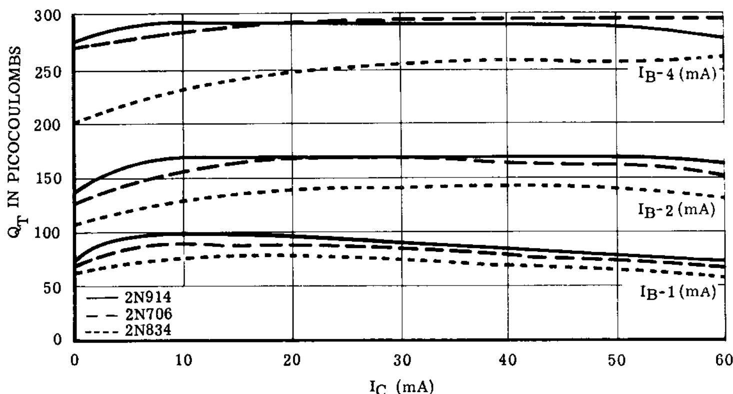


Figure 5-35a — Typical Q_T vs I_C Data

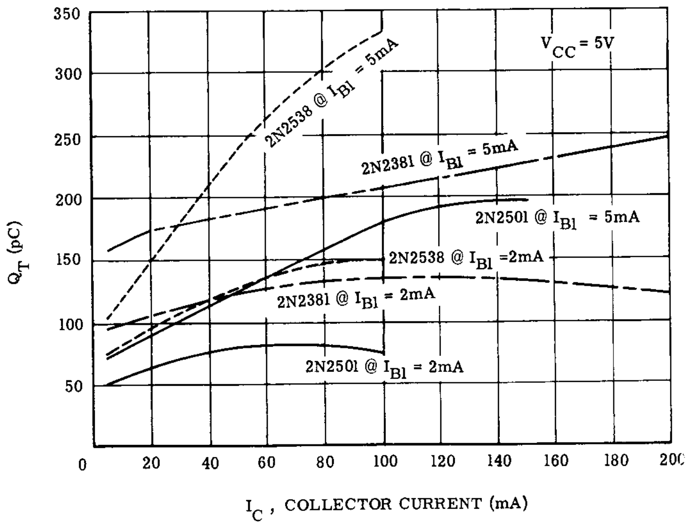


Figure 5-35b — Typical Q_T vs I_C Data

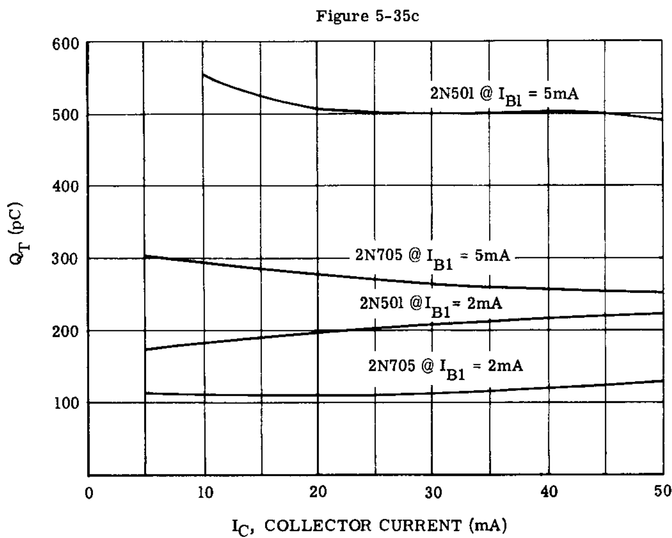


Figure 5-35c — Typical Q_T vs I_C Data

REFERENCES:

1. Beaufoy, R. and J. J. Sparkes: The Junction Transistor as a Charge Controlled Device, ATE Journal, Vol. B, pp 310-327, October, 1957.
2. Baker, A. N.: Charge Analysis of Transistor Operation, Proc. IRE, Vol. 49, pp 636-637, March, 1961.
3. Wilfinger, R. J.: Predicting Transistor Turn-on Delay Time in the Common Emitter Configuration, pp 34-41, Solid State Design, April, 1962.
4. Phillips, A. B.: Transistor Engineering, pp 337-340, McGraw-Hill Book Co., Inc., New York, N.Y., 1962.
5. Narud, J. A. and M. R. Aaron, "Analysis and Design of a Transistor Blocking Oscillator Including Inherent Non-linearities." BSTJ, Vol. 38, No. 3, May, 1959, pp 785-852 (p. 792).
6. Hwang, Y. C., D. S. Cleverly and D. J. Monsour: Transistor Switching Speed from Base Storage Charges and Their Lifetimes, Electronic Design, Part I, pp 52-55, March 1961, Part II, pp 50-53, April, 1963.
7. Moll, J. L., Large Signal Transient Response of Junction Transistors, Proc. IRE, Vol. 42, pp 1773-1784, December, 1954.
8. Easley, J. W.: The Effect of Collector Capacity on the Transient Response of Junction Transistors, IRE Transactions on Electron Devices, Vol. ED-4, pp 6-14, January, 1957.
9. Bashkow, T. R., "Effect of Nonlinear Collector Capacitance on Collector Current Rise Time." IRE Transactions on Electron Devices, Vol. ED-3, October, 1956, p. 167.
10. Nanavati, R. P.: Prediction of Storage Time in Junction Transistors, IRE Transactions on Electron Devices, Vol. ED- pp 9-15, January, 1960.
11. Stewart, J. M.: Charge Control Method Analysis Switching Times of Junction Transistors, Canadian Electronics Engineering, pp 45-50, February, 1962.
12. Lax, B. and S. F. Neustadter, "Transient Response of a p-n Junction," Journal of Applied Physics, Vol. 25, No. 9, September, 1954, pp 1148-1154.
13. Grinich & Noyce, "Switching Time Calculations for Diffused Base Transistors," IRE Wescon Convention Record, Part 3, Electron Devices, pp 141-147, August, 1958.

AUTOMATED TESTER

CHAPTER 6

Reliability Considerations for the Circuit Designer

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 defined percentage of "down time" due to equipment malfunction. To do this, the designer must have a knowledge of all the facets of reliability which contribute to system reliability, some of which are:

1) The relationship of transistor and component reliability to system reliability.
2) The causes of component failure.
3) How reliability is measured.
4) The various methods of specifying reliability assurance.
5) The factors involved in selecting components.
6) The effect of circuit design upon overall system reliability.

These facets of reliability, as pertaining to the transistor, will be discussed in this chapter. In a general way, much of the discussion can be applied to other components as well.

6-1 — Reliability of Transistorized Equipment

The ultimate measure of reliability is the degree to which an equipment performs the function for which it was designed.

A general method of expressing equipment reliability is Mean Time Between Failures (MTBF) which is equal to operating time divided by the number of failures. A number of other measures of equipment reliability are used, but MTBF is probably the most commonly accepted and can be related to the others mathematically. The reciprocal of MTBF is the failure rate. The MTBF of an equipment is, of course, dependent upon the number of components used in the equipment as well as the reliability of the individual components under the stresses encountered. In comparing reliability of different equipments, a useful method of normalizing is to consider failures per component hour. This is the number of failures divided by the product of the number of components and the hours of equipment operation.

The MTBF for an equipment is not a constant over its entire life. The MTBF will probably be relatively short during the "debugging" phase early in an equipment's life, until the early life failures due to both component manufacturing and equipment assembly faults are eliminated, and late in life when the failure rate increases due to component wear out failures. During the mid-portion of life, when the failures are random in nature, MTBF should be a maximum.

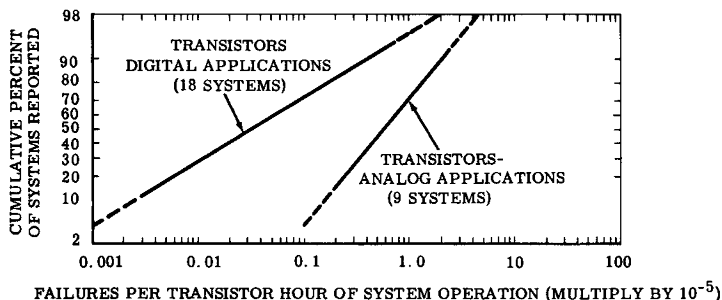


Figure 6-1. Failure Rate of Transistors in Digital and Analog Computer Application

Transistorized computers have a good reliability history.[1] Figure 6-1 shows reported field failure rates of transistors by class and circuit mode of operation. The median failure rate per transistor in digital applications is of the order of 0.03×10^{-5} failures per transistor hour for the eighteen systems reported, while the failure rate in analog applications in the nine systems reported was approximately twenty times greater.

This compilation for equipment reliability was made in 1961 and represented experience with equipments in service before that time. There is reason to believe that the reliability factor is much greater for today's digital equipment which use transistors with higher reliability than those available several years ago.

The reliability goal for the Airborne Computer of the Minuteman missile system is 7.76×10^{-5} failures per system hour for a computer with approximately 20,000 components.[2] Such high equipment reliability is possible only with proper circuit design and extremely reliable components.

Transistor reliability is usually expressed as failure rate in percent per 1000 hours. Since failure rate is the reciprocal of MTBF it is also possible to express transistor failure rates in terms of MTBF, but such an expression can be extremely misleading. For example, standard transistors with failure rates on the order of 1.0% to 0.1% for 1000 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 or over 114 years.

This is a rather meaningless figure for a number of reasons, some of which will be discussed in the following sections. All transistor reliability will, therefore, be expressed in terms of failure rate throughout this chapter.

Although the transistor's potential high reliability has been recognized since its invention, it has only been during the last few years that its actual reliability capabilities have been demonstrated. Probably the most comprehensive program to procure highly reliable transistors has been in conjunction with the guidance and control system for the Minuteman missile. As an example of the reliability required for Minuteman, the goal for germanium diffused base switching transistors was a maximum failure rate of .0007%/1000 hours with a confidence level of 60%. Motorola germanium mesa transistors have met and exceeded this goal.

6-2 — Achieving Transistor Reliability

Three major factors contribute to transistor reliability. There are:

1. Good basic device design.
2. Good manufacturing processes.
3. Quality and Reliability Control.

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

A transistor can be no more reliable than the basic reliability of its design. Therefore, the design must be capable of withstanding all the stresses which will be encountered in use. A number of transistor design features are important to reliability. Among the more important of these are surface stability, mechanical strength, and uniformly low thermal resistances from junction to environment. In addition, a transistor design which requires a critical manufacturing process is not a reliable design and must be avoided. Presuming that critical steps in a manufacturing process have been minimized, some of the attributes for the manufacture of highly reliable transistors are:

1) An effective system of vendor parts approval and incoming material control.
2) Thorough and complete training of manufacturing operators with maximum emphasis on quality of workmanship rather than speed of assembly.
3) A written process description which is adequate, exact, and up to date.
4) Adequate control of manufacturing environment.
5) Minimization of assembly steps.
6) A conscientious program of preventive machine and tool maintenance.
7) Management emphasis upon the importance of reliability in every transistor manufacturing operation.

Reliability emphasis during every phase of the manufacturing operation is the key which assures that every device will approach the reliability level inherent in the basic transistor design.

The final element essential to the fabrication of reliable transistors is a program of Quality and Reliability Control. Although the measurement of the quality of finished transistors is a vital element in assuring transistor reliability, it is only one segment of an effective quality control program.

The evaluation of reliability must accompany each step of the transistor design cycle, if the design is to be optimized for reliability. A number of evaluation techniques, some of which will be discussed in this chapter, have been developed to evaluate transistor reliability under various accelerated stress conditions. The application of these tests during the transistor development cycle results in data which can be factored into the device design to optimize reliability.

Quality Control must have a role in every step of the transistor manufacturing cycle if maximum reliability is to be realized. Constant verification of quality at each step in transistor manufacture from inspection of parts and subassemblies through device assembly will assure good process optimization and control. Good in-process quality control is essential to improve yield which will

generally result in high transistor reliability, because the same production uniformity which gives high yields can also minimize deviations from the reliability inherent in the basic design.

6-3 — Causes of Failure

A knowledge of the causes of semiconductor device failure is essential to an understanding of transistor reliability. Since a complete analysis of semiconductor device failure mechanisms is beyond the scope of this chapter, only the most general aspects of those failure mechanisms which are pertinent to switching transistors of the diffused base mesa types will be considered.

Transistor failure mechanisms can be broadly classified as follows:

1) Surface Defects
2) Mechanical Defects
3) Bulk Defects

SURFACE DEFECTS: The most prevalent cause of poor transistor reliability is failure due to the condition of the semiconductor surface. A surface condition leading to poor reliability may be caused either directly by imperfections within the encapsulated transistor itself, or by failure of the transistor housing which causes the semiconductor surface to be subject to the external environment, or a combination of both these factors. During transistor 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 transistors in an inert atmosphere (such as nitrogen) to reduce possibility of chemical reaction with the semiconductor surface, 2) the use of getters which absorb moisture (such as molecular sieve) to maintain low partial vapor pressure within the transistor housing, 3) the use of silicone varnish to seal the semiconductor surface from environment, and 4, 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. The recent use of the Annular* Process to prevent "channeling" also limits surface as well as bulk changes.

Stresses, which cause a change in the state of the semiconductor surface during transistor life, are a potential source of poor transistor reliability. Among the factors which can introduce mechanisms to change the state of the transistor surface are:

1) Entrapment of moisture or other contaminants within the transistor during encapsulation.
2) Escape of gases from the can or internal parts of the transistor during transistor life.
3) Loss of the hermetic seal due to encapsulation or glass seal leaks which were present at the time the transistor was manufactured or which occurred during subsequent transistor life.

*Patents Pending

Among the mechanisms of failure are the creation of conductive shunt paths or high series resistance paths. These can be above or below the oxide surface of passivated silicon transistors.

The surface passivation of silicon transistors by the growth of silicon dioxide films, which are chemically bonded to the surface, affords a greater degree of surface protection than has previously been available. However, unless properly designed and manufactured, even this class of transistors 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 necessary, even for passivated transistors, when maximum reliability is desired.

Collector base reverse current (I_{CBO}) is the most sensitive measure of transistor surface defects. Measurement of low level current gain (h_{FE} or h_{fe}) also often reveals useful information about the surface.

MECHANICAL DEFECTS: The mechanical defects which can occur in diffused base switching transistors are relatively straightforward to analyze. Among these are:

1) Poor bonding of die to header.
2) Poor electrode contact.
3) Defects associated with the fine wire which connects the header post and the electrode contact.
4) Lack of hermetic seal.

Those failure mechanisms associated with thermal compression bonding have received considerable attention with the advent of silicon diffused base mesa, planar, and annular transistors with 300°C storage temperature ratings. Thermal compression bonding is the technique by which a fine wire — about 0.001 inch in diameter, usually of aluminum or gold, is attached to the evaporated metal emitter and base contacts by pressure at a temperature below the melting point of the metals. Since this is a critical fabrication process, reliable thermal compression bonds require intensive process control, visual inspection, and testing of bond strength, to eliminate potentially unreliable transistors.

A problem can occur when gold is used for the fine electrode wire in silicon transistors. Embrittlement of the gold wire occurs at the point of attachment to the aluminum contact, because of formation of a gold-aluminum compound. This failure mechanism is stressed particularly at temperatures of 300°C and above. The wire embrittlement occurs at a very much slower rate under normal operating temperatures. This cause of poor reliability is eliminated by the use of the same material for both the contacts and the electrode wires.

Another potential source of poor reliability associated with this fine wire is its inability to carry high currents. Care must be taken to limit current, in this class of transistor, to levels which can be reliably conducted by the fine electrode wire. For example, for the 1.0 mil wires commonly used, exceeding 800 mA dc will often cause failure.

Poor contact of the die to the header may increase the thermal resistance of the transistor resulting in high junction temperatures during high power operation. Poor electrode contacts may cause hot spots but this is of secondary importance for relatively low level switching applications.

The effects of poor package sealing on surface stability have been reviewed in the previous section.

BULK DEFECTS: Bulk defects in switching transistors 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 transistor fabrication process, although they are usually detected at the final electrical test.

Included in this classification of defects are crystal imperfections which can cause non-uniform diffusion, resulting in high current concentrations and hot spots, and undesired impurities which can result in uneven voltage gradients. 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 the well designed and fabricated transistor, but could be a contributing cause to transistor "wear-out."

FAILURE ANALYSIS: Complete transistor failure analysis is quite complex and requires extensive facilities and a thorough knowledge of semiconductor theory and transistor fabrication methods. However, preliminary analysis at the equipment manufacturer's plant can prove very helpful in improving the reliability of transistorized equipment.

When a transistor failure is detected at any stage of equipment manufacture, from incoming parts inspection to final equipment test, a complete record should be compiled describing the indication of failure, the stage of manufacture at which the failure occurred, the circuit in which the transistor failed, the stress applied and any other information which will help complete the history of the failure. When the transistor is returned to the analysis laboratory, it should be visually examined for any possible indication of mishandling. Then it should be measured for electrical characteristics to determine if it truly is a failure.

If electrical test indicates the transistor is inoperative, the failure is probably mechanical in nature and the device should be X-rayed in an attempt to see the cause before the transistor is cut open. Opening a transistor case should be the last operation in failure analysis because no matter how much care is exercised, additional damage may be done which may mask the original cause of failure. Once the transistor is opened, the cause of mechanical failure will usually become apparent under microscopic examination.

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

If the transistor shows excessive leakage, the case should be thoroughly washed to remove any low resistance path that may have formed externally.

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

With the possible addition of a leak detection test, this is probably as far as failure analysis can be practically carried, outside of a semiconductor laboratory.

6-4 — Failure Rate as a Function of Time

The idealized curve of component failure rate versus time is shown in Figure 6-2. Several features of this familiar "bathtub" curve are important in any consideration of transistor reliability. The first portion of this curve indi-

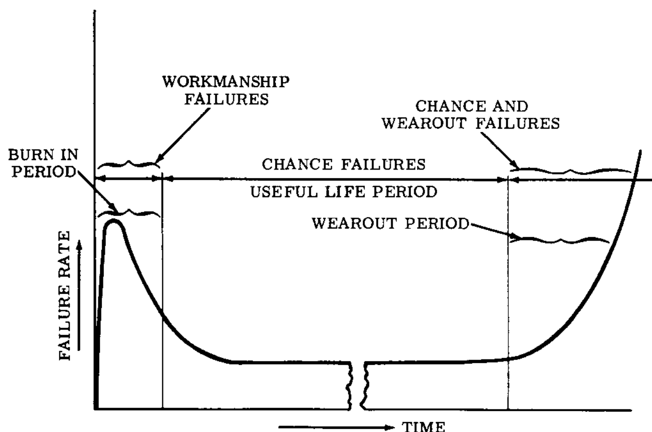


Figure 6-2 Failure Rate as a Function of Time

cates a sharply increasing and then a steadily decreasing failure rate during the "burn-in portion" of transistor life. The increasing failure rate for the very early life portion of Figure 6-2 may not always be seen. The portion of this curve which shows a decreasing failure rate for transistors has been repeatedly demonstrated.[3,4,5] These early life failures are generally classified as a result of poor workmanship.

The failure rate during the very early life depends upon a number of factors. Among these are the actual zero time in the life of the transistor, the definition of failure and, of course, the inherent reliability of the transistor. Actually, the life of a transistor begins when the encapsulating process is completed. Subsequent to encapsulation, a period of stressing, usually at elevated temperature, is required to stabilize the transistor'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 very early portion of the failure rate versus life curve.

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 transistor type 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 transistors will have a higher early life failure rate than they would have if tested to a specification with life test limits relaxed from initial limits. If the transistor parameters continue to drift with time, even the more relaxed life test limits will be exceeded and the total number of failures will be the same, regardless of the specified limits. However, if the transistors stabilize after a short period, as is often the case, then the total number of failures will often be less to the relaxed life test points than to the tighter limits.

The idealized failure rate versus time curve 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 6-2 shows an increasing failure rate indicated as "wear-out". This portion of the failure rate versus time curve is extremely difficult to define and will vary widely depending on transistor 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 or germanium die and the case header due to repeated cycling of junction temperature while the transistor case is at more or less a fixed temperature, or by glass hermetic seal failures due to environmental cycling, or by fatigue of the fine electrode wires due to mechanical stress, or by bulk defects. Little data is available from either transistor life tests or equipment field tests to permit an accurate picture of this portion of the failure rate versus time curve. Contrary to the early life failures which may be characterized as workmanship faults, the failures which occur in the wear-out period are a result of the basic design limitations of a transistor.

The fact that failure rate is not constant with time throughout transistor life, dictates that any statement of failure rate must refer to the time period considered. In this chapter all failure rates are based upon the first 1000 hours of transistor life test unless otherwise stated. This changing failure rate during the transistor life is a reason for not using MTBF as measure of reliability on an individual transistor basis.

The Minuteman transistor specifications have a requirement that samples must periodically be put on a three year operating life test under equipment use conditions. It is of interest to note that by the end of 1962 Motorola germanium mesa transistors subjected to these tests logged over 9.43 million operating hours with the early samples having over 23,000 operating hours with no indication of an increasing failure rate due to wear-out. In fact, only a total of four failures occurred and all of these were relatively early in life.

Since many early life transistor failures are the result of manufacturing flaws, it is quite possible to develop screening procedures to improve transistor reliability. Actually most reputable transistor manufacturers employ screening procedures as a regular part of the transistor fabrication process. The effectiveness of any screening procedure must be carefully verified for the particular transistor type under consideration.

All transistors receive a stress to stabilize characteristics. The stress applied should be one that has proven effective for the particular transistor type being manufactured. In many cases, high temperature storage for a period of time has been determined to be adequate. For other transistor types, operation with power applied, or high temperature storage with voltage applied has been demonstrated to be most effective.

The use of 100% leak detection is quite universal in the transistor industry. Subjecting all silicon transistors, using thermal compression bonds, to a centrifuge test has become standard with many manufacturers.

All transistors 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 transistor life. For example, extreme mechanical stresses may not only destroy weak units, but may weaken good units.

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

1. Relatively short term life tests, (e.g. 1000 hours) are sufficient to assure transistor reliability for long time use.
2. Transistor quality can be enhanced by the use of screening procedures to eliminate workmanship failures.

6-5 — Measurement of Transistor Reliability

The ultimate measure of transistor reliability is performance in intended applications. However, since long term tests at use conditions are not feasible from either time or cost considerations, more practical test procedures have been developed to assure transistor reliability.

The intent of these tests is to compress time so that a measure of a transistor's long term reliability may be obtained in a relatively short time, (e.g. 1000 hours) and to accelerate stresses so that a relatively small number of transistors may be tested at high stress levels to assure very low failure rates under normal use conditions. A test procedure used to develop the acceleration factors between high stress tests and stresses encountered under use conditions is known as matrix testing.

MATRIX TESTING: Matrix testing was used to establish acceleration factors for the MINUTEMAN Reliability Improvement Program. A matrix program includes the testing of a number of devices under a range of test conditions designed to stress the potential device failure mechanisms. Aside from mechanical and environmental stresses, the conditions a semiconductor device will encounter during use which could cause failure are voltage, current, ambient temperature, and junction temperature. These conditions are not independent. In fact, they are closely inter-related.

To determine the extent of the effect of these stresses on semiconductor device reliability, an experiment is designed to test devices under various combinations of these stresses. Statistical analysis of these test results at specific stress points permits the prediction of failure rates at other stress conditions, and provides a relative measure of the effects of various stress conditions, i.e. develops acceleration factors.

A matrix test which was developed by Motorola as a part of its MINUTEMAN Transistor Reliability Improvement Program is shown in Figure 6-3. A total of 9675 devices were tested in this Matrix. Five power levels from 0 to 133% of rated dissipation were tested under 8 ambient temperature conditions. Two voltage levels were used at each dissipation level. It should be noted that test conditions below, at, and above device ratings are included. The tests were conducted for 4000 hours.

Percent of rated 25°C Power Dissipation and Test Conditions									
	0%	33%		66%		100%		133%	
AMBIENT TEMPERATURE °C	0 Volts	5Vdc 10mAdc	15Vdc 3.3mAdc	5Vdc 20mAdc	15Vdc 6.7mAdc	5Vdc 30mAdc	15Vdc 10mAdc	5Vdc 40mAdc	15Vdc 13.3mAdc
25		1500	1500	500	500	200	200	100	100
50	1500	500	500	200	200	100	100	75	75
75	500	200	200	100	100	75	75	50	50
100	200								
125	100								
150	75								
175	50								
200	50								

Figures in chart are number of transistors tested at each stress level.

Figure 6-3 Minuteman Matrix Test

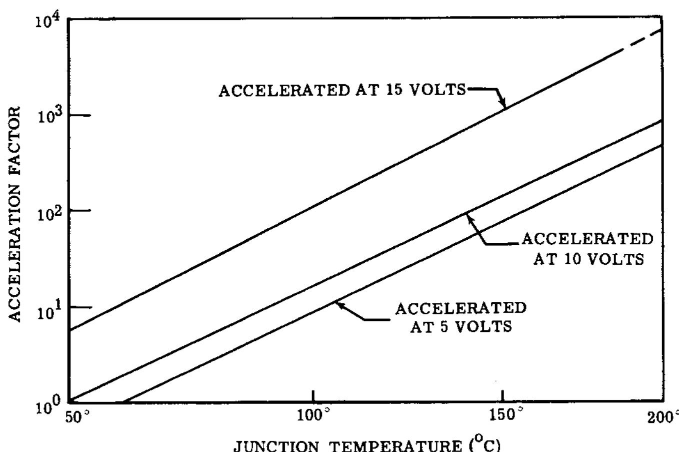


Figure 6-4 Minuteman Matrix II Acceleration Factor Referenced to 50°C, and 10 Volt Nominal

From the relationship of failure rate vs. temperature obtained from the matrix test, acceleration factors versus junction temperature as shown in Figure 6-4 were determined. It will be noted that these plots are straight lines. These acceleration factors were referenced to a junction temperature of 50°C and $V_{CB} = 10V$, specified MINUTEMAN usage conditions. From these curves it is possible to determine an acceleration factor of a wide range of test conditions. A junction temperature of 100°C is used for the acceptance tests and the failure rate verification tests under the MINUTEMAN program. The acceleration factor between 100°C and 50°C, at 10 volts, is 28. By utilizing this acceleration factor, a failure rate of 0.0007%/1000 hours at 50°C usage condition can be verified by demonstrating a failure rate of 28 x 0.0007%/1000 hours or 0.0196%/1000 hours at 100°C.

Such extensive matrix testing programs as were conducted for Minuteman cannot often be economically justified. However, even in a much simpler form such approaches can give significant results in determining a relationship between failure rates under high stress conditions and those at use conditions.

A vital precaution which must be observed in matrix testing or any other accelerated test plan is to assure that no new failure mechanisms are introduced by the accelerated stress which will not be encountered in normal transistor circuit use. If the high stress tests introduce new failure mechanisms, then they lose validity in predicting long term life.

STEP-STRESS TESTING: The step-stress test method, pioneered by Bell Telephone Laboratories, has the advantage over matrix testing in that it is a relatively short time test.

Step-stress testing consists of subjecting the devices being evaluated to successively increasing levels of stress until a majority of the devices have failed. Step-stressing can be done for mechanical stresses such as constant acceleration, electrical stresses such as surge current or power dissipation, and ambient stress such as temperature.

Figure 6-5 illustrates the results of step-stress analysis performed on a Motorola developmental germanium epitaxial switch transistor. The power was increased in 50 milliwatt increments and applied for 5 hours at each step. The failures at each step were noted. The power was converted to junction temperature based upon the rated thermal resistance of 0.5°C/mW and plotted with the stress temperature in basic units, e.g. reciprocal of absolute temperature as the ordinate and a Gaussian probability failure scale as the abscissa.

The linear relationship will be noted when the points are joined. The fact that this plot is a straight line indicates that only one failure mechanism is actuated by

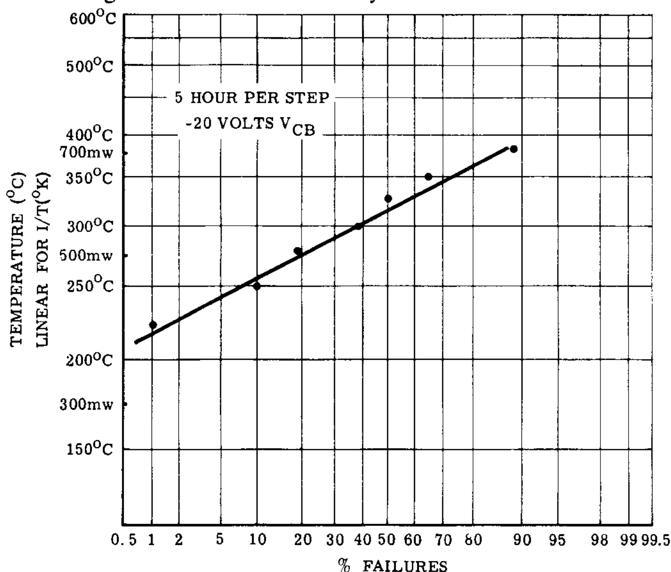


Figure 6-5 Power Step-Stress Test for Developmental Germanium Switching Transistor

the stress applied. If a second failure mechanism was introduced at the higher stresses the slope of the line would be discontinuous.

A number of other applications have been proposed for step-stress analysis[6], however, its application for relatively fast comparative analysis is most useful. Step-stress analysis can be used to determine,

1. The comparative effect of transistor manufacturing process changes on reliability.
2. The variation in reliability of transistor lots manufactured at different times.
3. Comparative analysis of similar transistor types supplied by different manufacturers.

ACCEPTANCE TESTING: To assure the reliability of transistors, an acceptance test program is necessary. Although the inherent reliability of a transistor type may have been demonstrated, it is necessary to regularly test devices to assure a uniformly high quality product being produced. A more or less standard array of acceptance tests have been adopted in the military specifications for transistors. Usually each lot of transistors for delivery against military specifications is sampled to assure compliance with specified electrical characteristics and capability to withstand environmental, mechanical, and life test stresses. These tests are adequately described in MIL-S-19500, the general military specification for semiconductor devices, and MIL-STD-750 which specifies test methods for semiconductor devices.

These tests are generally conducted at the maximum transistor ratings. If reliable equipment performance is to be assured, the stresses which the transistor encounters during life will be less than the maximum ratings. Thus, normal acceptance tests are accelerated.

It has been proposed[6] that transistor maximum ratings and acceptance tests be based upon stresses which produce a failure rate of 1 per cent. This is probably a realistic figure for the maximum reliability which can be economically assured by acceptance testing on a regular basis.

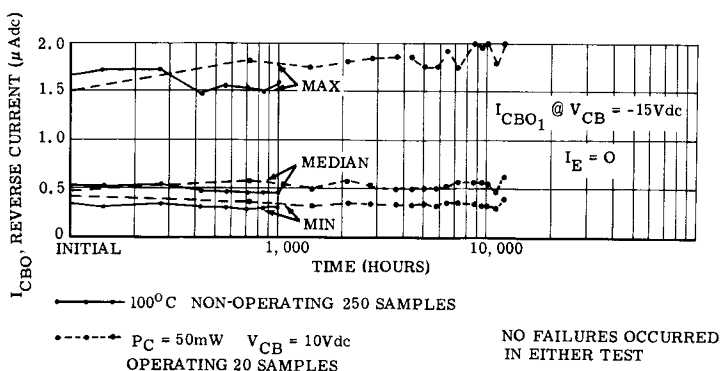
The adequacy of accelerated acceptance testing in assuring highly reliable performance for transistorized equipment has been shown by low equipment failure rates which have been experienced. A specific example of the adequacy of accelerated testing may be found in Motorola's program of supplying germanium mesa transistors for Minuteman. A basic quality assurance test to verify the quality of transistors delivered for Minuteman is the Degradation B life test which must be regularly performed on samples randomly selected from each day's production. For the germanium mesa transistor, this accelerated test consists of a 100°C non-operating life test for 1000 hours. Figure 6-6 and Table 6-1 illustrate the performance of samples from Lot 21 which was manufactured during the summer of 1961. Fifty samples were randomly selected from each day's production to make up a sample of 250 transistors for the 1000 hour 100°C life test. In addition to this accelerated life test, four samples per day were selected from Lot 21 to make up a weekly sample of 20 for the three year life test under typical use conditions of 50 milliwatts (junction temperature of 50°C) at $V_{CB}=10$ Vdc. This life test will be continued for 36 months.

Figure 6-6 shows the results of the I_{CBO} and h_{FE} measurements made during the 1000 hour life test at 100°C and the same parameters for the 50 milliwatt life test for the 12,240 hours completed by the end of 1962. The results of measure-

ments for the other parameters are shown in Table 6-1. Initial and 1000 hour measurements are listed for the 100°C non-operating life test and initial and 12,240 hour measurements for the three year operating life test. No failures occurred in either life test.

No significant differences in the results of the two life tests are apparent. It must be borne in mind that for an accelerated test to be significant, the stress applied must be of the same type but greater in magnitude than those which will be encountered during life.

Data for other Motorola products is shown in Tables 6-2 and 6-3. The excellent stability of the measured characteristics shows that little degradation occurs with time. Therefore, it would be unwise to apply a safety factor to these characteristics if the safety factor results in an increased number of components in the system.



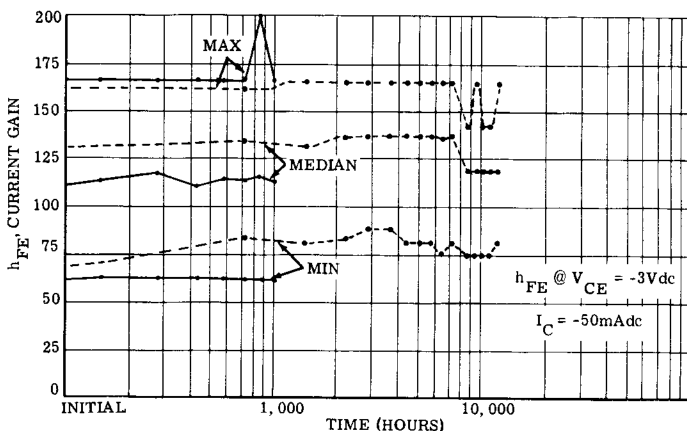


Figure 6-6a and b Results of I_{CBO} and h_{FE} Measurements During a 12,240 Hour Life Test

TABLE 6-1 — PARAMETER DISTRIBUTIONS DURING LIFE TEST OF MOTOROLA GERMANIUM MESA TRANSISTORS TYPE 101M (Lot 21)

Parameter	Conditions	High Temperature Non-Operating Life $T_A = 100°C$ Sample Size = 250										Operating Life Test $P_C = 50mWdc$ $V_{CB} = 10Vdc$ Sample Size = 20										Units
		Initial					1000 Hours					Initial					12,240 Hours					
		Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	
I_{CBO1}	$V_{CB} = 15Vdc$, $I_E = 0$	.38	.41	.52	.76	1.7	.34	.37	.48	.68	1.5	.35	.39	.49	.88	1.5	.45	.45	.60	1.3	2.1	μAdc
I_{CBO2}	$V_{CB} = 1Vdc$, $I_E = 0$	.13	.14	.18	.22	.26	.11	.12	.15	.19	.24	.12	.12	.17	.22	.24	.14	.14	.20	.26	.27	μAdc
I_{EBO}	$V_{EB} = 5Vdc$, $I_E = 0$	3	8	29	58	83	3	9	32	62	86	10	11	28	65	73	3	3	9	50	60	μAdc
h_{FE}	$V_{CE} = 3Vdc$, $I_C = 50mAdc$	63	70	111	160	166	63	71	114	161	166	71	74	132	161	166	83	89	119	161	166	—
SV_{CE}	$I_C = 50mAdc$, $I_B = 1.5mAdc$	.41	.44	.49	.58	.71	.41	.43	.48	.57	.71	.43	.43	.46	.52	.53	.43	.43	.46	.52	.53	Vdc
V_{BE1}	$I_C = 50mAdc$, $I_B = 4mAdc$	.51	.53	.57	.65	.71	.53	.53	.58	.65	.79	.53	.53	.56	.68	.71	.55	.55	.58	.68	.73	Vdc
V_{BE2}	$I_C = 5mAdc$, $I_B = 1.5mAdc$	.35	.35	.37	.38	.40	.35	.36	.37	.39	.41	.35	.35	.36	.38	.38	.36	.36	.37	.39	.39	Vdc

TABLE 6-2 — PARAMETER DISTRIBUTIONS DURING LIFE TEST OF MOTOROLA GERMANIUM EPITAXIAL MESA TRANSISTORS

GL-5 Line Source for Types 2N828, 2N960-2N967 (Lot 210)

Parameter	Conditions	High Temperature Non-Operating Life $T_A = 100°C$ Sample Size = 40										Operating Life Test $P_C = 150mWdc$ $V_{CB} = 10Vdc$ Sample Size = 100										Units
		Initial					1000 Hours					Initial					1000 Hours					
		Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	
BV_{CBO}	$I_C = 100\mu Adc$	14	16	26	36	43	14	16	26	36	43	16	17	25	34	43	16	17	25	33	43	Vdc
I_{CBO}	$V_{CB} = 6Vdc$ $I_E = 0$	.08	.08	.13	.98	2.7	.06	.06	.08	1.6	3.0	.08	.08	.12	.90	2.9	.08	.08	.12	.98	1.60	μAdc
BV_{EBO}	$I_E = 100\mu Adc$	2.3	2.5	4.7	6.1	6.5	2.2	2.5	4.9	6.4	6.6	2.5	4.0	5.0	6.1	6.5	2.8	4.4	5.3	6.3	7.0	Vdc
h_{FE}	$V_{CE} = .3Vdc$ $I_C = 10mAdc$	25	28	58	83	83	25	27	67	83	83	27	28	53	83	83	36	38	74	83	83	—
SV_{CE1}	$I_C = 10mAdc$ $I_B = 1mAdc$	.08	.08	.09	.13	.14	.08	.08	.10	.13	.14	.07	.08	.10	.13	.14	.07	.08	.10	.13	.14	Vdc
SV_{CE2}	$I_C = 50mAdc$ $I_B = 5mAdc$	.11	.12	.16	.25	.31	.11	.12	.17	.25	.32	.11	.12	.15	.25	.36	.11	.12	.15	.24	.36	Vdc
V_{BE1}	$I_C = 10mAdc$ $I_B = 1mAdc$	.36	.36	.38	.41	.42	.36	.37	.38	.41	.43	.35	.36	.38	.41	.42	.36	.37	.38	.40	.42	Vdc
V_{BE2}	$I_{C1} = 50mAdc$ $I_B = 5mAdc$	.46	.46	.49	.59	.65	.45	.46	.49	.59	.65	.44	.46	.48	.58	.64	.44	.46	.48	.58	.64	Vdc

TABLE 6-3 — PARAMETER DISTRIBUTIONS DURING LIFE TEST OF MOTOROLA SILICON EPITAXIAL TRANSISTORS
SL-1 Line, Source for Types 2N834, 2N835, 2N2501,
(Lots 214, 215, 216)

Parameter	Conditions	Operating Life Test $P_C = 300mW$ $V_{CB} = 15Vdc$ Sample Size = 298										Units
		Initial					1000 Hours					
		Min	5%	50%	95%	Max	Min	5%	50%	95%	Max	
I_{CBO}	$V_{CB} = 15Vdc$, $I_E = 0$	2.6	3.2	5.0	17	71	3.1	3.7	5.5	14	94	nAdc
BV_{CBO}	$I_C = 10\mu Adc$	40	48	72	98	142	41	49	74	100	142	Vdc
BV_{EBO}	$I_E = 10\mu Adc$	6.6	7.1	7.6	8.1	8.7	6.7	7.1	7.8	8.2	8.6	Vdc
h_{FE}	$V_{CE} = 1Vdc$, $I_C = 10mAdc$	51	61	95	143	192	38	60	90	133	167	—
SV_{CE1}	$I_C = 10mAdc$, $I_B = 1mAdc$	.12	.13	.15	.18	.20	.13	.13	.15	.18	.21	Vdc
SV_{CE2}	$I_C = 50mAdc$, $I_B = 5mAdc$	.15	.16	.20	.26	.30	.16	.16	.20	.27	.33	Vdc
SV_{BE}	$I_C = 10mAdc$, $I_B = 1mAdc$	.72	.72	.74	.77	.83	.71	.72	.74	.78	.83	Vdc

6-6 — Specifying Reliability Assurance

An understanding of the techniques used to specify reliability assurance is essential. The factors which influence the degree of reliability assurance obtained by testing a sample of transistors are:

1) The stress applied.
2) The sample plan used.
3) The criteria of failure.
4) The number of failures permitted.

All of these factors must be specified if adequate reliability verification is to be assured.

STRESS: The stresses applied should be chosen to accelerate the transistor failure mechanisms which can cause failures during equipment life. Acceptance testing is almost universally conducted under maximum rated conditions. Since the stresses the transistor encounters during life in well designed equipment are less than the maximum rated, the acceptance life test is an accelerated test.

SAMPLE PLAN: In any plan by which the quality of a large population of devices is assured by testing a sample of that population, there is an element of risk that the measured quality of the sample will not give an accurate picture of the quality of the total population. The smaller the absolute size of the sample, and the smaller the sample is in relation to the total population, the greater the risk that the measured quality of the sample is not the true quality of the total population. The sample test results may give an accurate, a pessimistic, or optimistic picture of the true quality of the total population. The sample plan must be selected to give as accurate a picture of the total population as cost and time limitations permit.

The accuracy with which the sample test results measure the quality of the total population is known as the confidence level.

If it is desired to use the results of a sample test to state a reliability level for an entire lot, then the maximum failure rate assured becomes lower, as the confidence level with which it is assured becomes higher. Thus, any statement of failure rate must include information as to whether it is a measured failure rate or whether it is a maximum failure rate. If it is a maximum failure rate, then the associated confidence level must also be stated.

Two basic methods of sampling quality assurance are in use in the transistor industry today. These are the AQL and the LTPD plans.

The AQL (Acceptable Quality Level) procedure has been in use for a number of years. Under it, an inspection level and an AQL are specified. For each lot size the inspection level specified determines the number of samples required. The number of samples to be tested increases as the lot size increases but the ratio of sample size to lot size decreases for larger lots. MIL-STD-105 "Sampling Procedures and Tables for Inspection by Attributes" specifies the sample size for any inspection level and lot size and stipulates the number of failures permitted for any AQL. The AQL value is roughly the maximum average per cent defective permitted if 19 out of 20 lot submissions are to be accepted.

The AQL system is known as a "producers risk" plan because the producers risk is specified while the risk the consumer is taking is not specified. The

manufacturer has an approximately 5% chance of having a lot rejected if the percent of defective devices is less than the specified AQL. The lots accepted, however, could have a considerably higher per cent defectives than the AQL indicates. This method of quality assurance is especially unsatisfactory if the sample size is small.

The LTPD (Lot Tolerance Per Cent Defective) method of quality assurance has gained increased acceptance in recent years. Under this procedure, assurance is given that only infrequently, (generally 10% of the time) lots with a poorer quality than that specified will be passed. Since under this plan, the consumer is protected against receiving poor quality 90% of the time in comparison to the AQL system which protects the manufacturer from rejecting good quality product 95% of the time, the LTPD system is a "consumer risk" plan because the risk of the consumer is specified.

The relationship between AQL and LTPD is illustrated by Figure 6-7. This curve is the operating characteristic for an AQL of 4.0% at sample sizes 15 and 150. An operating characteristic may be said to be a measure of ability of an acceptance plan to distinguish between acceptable and reject lots. The ideal operating characteristic is a vertical line intersecting the abscissa at the desired quality level. This ideal operating characteristic can only be achieved by 100 percent inspection. At less than 100% inspection, the operating characteristic is a measure of the degree with which the results of the sample test assure the quality of the total lot. For smaller sample sizes, the effectiveness of the AQL procedure in assuring quality becomes quite poor. In Figure 6-7 for a 4.0% AQL, a sample size of 150 will permit 1 lot out of 10 with an 11 percent defective to pass, while a sample size of 15 will permit a lot with 25% defectives to pass 10% of the time.

Under the LTPD procedure, the sampling plan is such that the lower end of the operating characteristic is controlled so that no more than 1 in 10 lots can pass if the specified LTPD is exceeded.

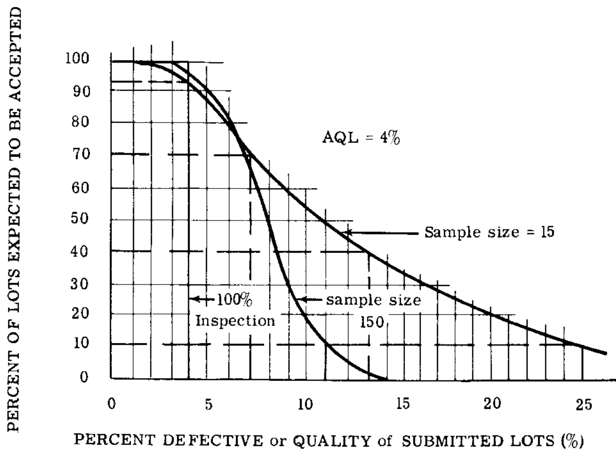


Figure 6-7 Operating Characteristics for a 4% AQL

Under the LTPD plan, sample size is independent of lot size. MIL-S-19500 lists sample sizes and number of rejects permitted for various LTPD's. The per-cent defective, permitted in a sample size required to assure a specified LTPD, increases as the sample size increases. Under a pure LTPD plan, the per cent defective permitted in the sample approaches the actual specified LTPD as the sample size approaches 100% inspection.

The LTPD's included in typical military specifications are generally larger than the AQL's formerly specified. Thus, for large sample sizes greater quality assurance protection might be obtained from a typical AQL military specification than from an LTPD specification. Since the purpose of the military specification groups in adopting the LTPD system was not to reduce the known quality of transistors accepted, but to reduce the risk of material of poor quality being unknowingly passed due to sampling risk, a modified LTPD plan is in general use today. Under this modified LTPD plan a maximum acceptance number or minimum rejection number is specified. (Minimum rejection number equals maximum acceptance number plus one.) Under this procedure the quality may be verified by selecting a sample not larger than a specified size. The lot may be accepted by testing smaller sample sizes, but the permitted per cent defective of the sample is less.

This modified LTPD plan gives added consumer protection against the possibility of receiving poor quality because of the risks involved with small lot and small sample sizes which might occur under the AQL system. Yet this plan limits the minimum quality which can be shipped to that which would be assured by an equivalent AQL for large lot sizes.

The LTPD levels which are typical for current military specifications are an LTPD of 5 with a minimum rejection number of 5 for major electrical characteristics, an LTPD of 10, minimum rejection number of 5 for mechanical and environmental tests, and an LTPD per 1000 hours* of 5 or 10 for life test. No minimum rejection number is usually given for life test because test cost will limit the size of the samples which can be life tested for 1000 hours. The cost of testing is a definite limitation upon the level of reliability assurance which may be verified by acceptance testing. This is especially true when acceptance tests are performed by sampling lots of transistors which have been accumulated in response to specific customer orders. The degree of reliability assurance which can be economically provided for specific orders by this method is probably limited to an LTPD of 10%.

At Motorola, the concept of reliability assurance is carried one step further by the "Line Acceptance Program". Under this procedure transistor reliability is assured by regularly sampling production lines at final electrical test, and subjecting the samples to the full military accelerated test sequence. This program not only assures the quality of transistors shipped, but it provides immediate feedback to the production lines of any change in transistor quality so that remedial action can be taken immediately.

Even under the "Line Acceptance" program, it becomes impractical to assure the operating life reliability of a given lot of transistors much below the $\lambda = 5$ level. For example, to demonstrate a λ of 1 (with 90% confidence that the lots will have no worse a failure rate than 1% per 1000 hours) a sample

*LTPD per thousand hours is designated λ.

size of 231 may be tested with no rejects permitted, or 1 reject would be allowed in a sample size of 390. The sample size would have to be 1,421 transistors if 10 rejects were to be permitted. To demonstrate very low failure rates for lot acceptance, testing costs necessitate the use of sequential test procedures in which the sample test results of a given lot are combined with test results of preceding lots to determine the actual reliability level.

It must be remembered that these acceptance tests are accelerated tests. The failure rates which are assured by these tests are higher than those which will be experienced in equipment life. Acceleration factors may be used to relate the acceptance test failure rate to expected failure rate in equipment.

END POINTS: The criteria of failure (end points) is the third factor in specifying reliability assurance. Three methods of specifying failure are in common usage today. These are:

1) The end point limits are the same as the initial electrical limits, e.g. $h_{FE} = 20$ min., 40 max., initial and end of life.

2) The end point limits are relaxed from the initial limits, e.g. $h_{FE} = 20$ min., 40 max., initial and $h_{FE} = 15$ min., 50 max., end of life.

3) The maximum shift of parameter characteristics are specified on an individual basis, e.g. the change of h_{FE} for any transistor must be less than $\pm$ 20% during the life test.

The second method has been the most widely used. The first method is often used for "high reliability" specifications, but its value is somewhat questionable. This method of specifying end of life limits does not take into account any possible inaccuracy in repeating parameter measurements during a several week life test. If minor shifts in transistor characteristics are causing problems in meeting the end point limits, the manufacturer can institute a parameter screen to select transistors to tighter than the specified initial limits. Thus, in effect, no greater parameter stability than given by method 2 is assured. Specifications which have identical initial and end of life limits may dictate that the quality assurance provisions, (i.e. AQL or LTPD), be quite loose in order to avoid lot rejection due to relatively minor parameter shifts or inability to precisely repeat measurements.

The third method, which is to specify permitted parameter shift on an individual transistor basis, has considerable merit for assuring the delivery of stable transistors. This method is the most expensive of the three to implement, because it requires that data on each characteristic be recorded and that calculations be performed on the shift of each characteristic for each transistor to determine if the lot meets the specified quality assurance provisions. A precaution which must be observed when this method is used, is to be sure that measurement accuracy is much greater than the parameter shift permitted. For example, many silicon passivated transistors have a maximum I_{CBO} limit of 10 nanoamps (10×10^{-9} Amperes) with median of the distribution being a fraction of a nanoamp. The measurement of even a 50% change in reverse current, which was a fraction of a nanoamp initially, after 1000 hours of life is most impractical. A specification using parameter shift as a criteria should specify a per cent shift or an absolute value, whichever is greater. For example, for a silicon annular transistor with an initial limit for I_{CBO} of 10 nanoamps, the end of life limit in relation to the initial values could be specified as +50% or +2 nanoamps, whichever is greater.

In the specification of end points, a careful compromise must be reached between making the end points tight enough so that poor reliability will be detected and yet not so tight that any minor shift in characteristics will reject the lot. Probably the best compromise is the use of double end points. This would consist of a relatively tight limit, perhaps a maximum shift of parameters, to a relatively loose LTPD, and looser limits to a tighter LTPD. For lot acceptance both criteria must be met. It should be borne in mind that with this method the sample size which must be tested is dependent upon the tighter LTPD(or AQL) so this will govern the test cost.

In addition to the elements of stress applied, sampling used, and criteria of failure; the number of rejects permitted remains to be considered in order to adequately specify acceptance procedures for reliable transistors. The number of rejects permitted, of course, depends upon the quality assurance required. The use of a reliability assurance plan which permits no rejects should be avoided because the possibility of a random failure exists even in the most reliable product and the most carefully conducted tests.

6-7 — Achieving Reliable Switching System Performance

SELECTING TRANSISTORS: The foundation upon which any reliable equipment design must be based is reliable components. Without reliable components even the most careful design cannot result in maximum equipment reliability. Of course, the prime consideration in the choice of a transistor type is its capability to perform the electronic function required. Generally, at the circuit design stage, any one of a number of transistor types could be selected to give satisfactory performance. However, the reliability of these transistor types may not be equal. A number of factors must be considered in the choice of a transistor type when reliability is of prime importance. These are:

1) Has the reliability of the device under consideration been proven? New transistor 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 borne in mind that it takes time to adequately prove a transistor's reliability and that generally the reliability of newer transistor types has not been verified to the extent of older types.
2) Has the transistor been in production long enough for any problems which may adversely affect reliability to have been eliminated? Early in the production phase of a transistor type, major emphasis is often given to process improvement to optimize electrical characteristics. As the production process settles down and yields improve, reliability will generally also improve.
3) Is the transistor type under consideration a major portion of the manufacturers yield? A characteristic of the semiconductor industry has been that a number of transistor 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 transistor 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) Is the transistor type one which will receive wide usage? Unless a transistor will have high volume application, it may not be in continuous production or its production rate may remain low. Under the circumstances of intermittent production or low production rates, it is extremely difficult to optimize the manufacturing process for reliability.

5) Does the transistor manufacturer have a good reliability image? The semiconductor technology is not a simple one. Unless a manufacturer has a broad background of semiconductor technology and experience, and a history of reliable products, the resources necessary to optimize transistor reliability may not be available. A manufacturer possessing these resources has the facility to accurately measure reliability and the skills to take the corrective action necessary to eliminate the factors causing poor reliability. A history of products with good reliability is the best assurance of the reliability of new products.

CIRCUIT DESIGN CONSIDERATIONS: The selection of the most reliable transistor to perform the required function is basic, but is only the first step in assuring the reliable operation of digital circuitry. Several circuit design considerations to assure reliable transistor performance follow:

1. When possible, circuit performance should be based upon the most stable transistor parameters. This is much more feasible in digital than in analog circuitry, and is probably the major contributing factor to the higher reliability of digital equipment.
2. 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.
3. Circuit design which is dependent upon transistor characteristics that are uncontrolled can lead to poor reliability and should be avoided. If circuit performance is dependent upon transistor characteristics which are not specified, and thus not controlled, there is no assurance that subsequent production will have the same characteristics.
4. 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 optimize reliability.
5. The environment which the transistor, circuit and equipment encounters during assembly, testing, and use, must be controlled to assure maximum reliability.

Each of these considerations for reliability will be considered in detail.

TRANSISTOR STABILITY: The basic properties of a transistor used in switching applications are its high impedance in the off state and its amplifying characteristics in the on state. The two characteristics which are usually measured to verify transistor quality are reverse current (usually I_{CBO}) and current gain (h_{FE}). Measurement of these characteristics verify the off and on reliability of the transistor.

These parameters, particularly reverse current, are largely dependent upon the condition of the semiconductor surface. Thus I_{CBO} and h_{FE} are also the most sensitive parameters for detecting poor transistor reliability and the most important parameters for digital circuit performance.

Figure 6-6 and Table 6-1 showed the excellent stability of the important characteristics for digital applications of Motorola germanium mesa transistors. Tables 6-2 and 6-3 give similar data for Motorola germanium epitaxial and silicon passivated switching transistors, for 1000 hours of operating life at maximum ratings. It will be noted that all the characteristics show excellent stability, but that the breakdown voltage, emitter-base and collector-emitter saturation voltages, which are primarily dependent upon bulk characteristics, show practically no change during life. Experience has demonstrated that the switching speed will show practically no change during life in a well designed and fabricated transistor, except for that caused by current gain instability, which would be detected by the h_{FE} measurement.

USING PROPER DESIGN LIMITS: The preceding section on parameter stability indicates that gain and reverse current are characteristics which are least stable. The other characteristics, which are determined by bulk properties, are quite stable with life but do vary with temperature. Fortunately, the variations of these bulk sensitive parameters with temperature are quite predictable so that given limit data, safety margins are not required.

Current gain and reverse current are also temperature sensitive, although the degree of their changes with temperature are generally more controlled by the bulk characteristics of the transistor than the surface characteristics and hence are predictable to a fair degree. Lacking adequate life test data, it is common practice to design digital circuitry to perform if the current gain decreases to 50% of its minimum specified value and reverse current increases to 5 times its maximum specified value.

Effort should be made to secure aging information from component manufacturers. Use of narrow limits can lead to a high probability of individual circuit failure, while use of wide limits will usually result in a greater number of components to perform a given electronic function which increases the probability of random failures. Either extreme can cause poor system reliability. For example, if it is known that h_{FE} decreases with operating time, an insufficient derating of h_{FE} will result in poor reliability. However, if h_{FE} is known to be stable or to increase with life, as often happens, derating h_{FE} will result in a greater number of transistors in the system which also can result in poor reliability.

UNCONTROLLED CHARACTERISTICS: A major source of poor reliability in transistorized digital circuitry has been the use of uncontrolled or unspecified transistor characteristics. In some cases, this resulted from inadequate specification on the part of the transistor manufacturer. Todays transistor specifications are much more complete in providing controls on the parameters necessary for digital switching applications. On occasion, circuits have been designed with transistor types which were never intended for the mode of operation used. Examples of this are the use of standard switching transistors in avalanche mode circuits or in chopper circuits. The fact that a given sample of transistors happen to work in such circuits is no assurance that their operation will be reliable, or that the next shipment will give satisfactory performance.

The use of transistors in modes of operation and at operating points significantly different from those at which parameters are specified, and therefore controlled, must be avoided if maximum reliability is desired.

DERATING: The relationship of derating and reliability has been introduced in previous sections. The transistor manufacturer must conduct his reliability tests under accelerated conditions at or above maximum device ratings because of time and cost limitations. Furthermore it is necessary to obtain data quickly which can be fed-back into the manufacturing line to enable corrective action to be taken if necessary.

The amount of transistor derating which should be employed for any digital application depends upon a number of factors:

1) The system reliability requirements.
2) System design constraints such as size, weight, power supply capacity, etc.
3) The crossover point between the transistor reliability gained by derating, and the loss of reliability by added circuit complexity.
4) The point of diminishing returns where added derating will not increase transistor reliability.
5) The cost of components having specifications better than that dictated solely by electrical requirements.

Of course these questions can only be answered for a particular equipment design and for a particular transistor type. However, some general rules can be stated as guides.

1) Junction temperature is probably the most significant factor affecting transistor reliability. Limiting the maximum junction temperature rise to approximately 50% of maximum ratings is probably the most effective method of improving reliability. A fact which must be remembered in any consideration of temperature derating is the method of verifying transistor dissipation ratings. Most transistors are life-tested under rated dissipation at room temperature, and by non-operating life test at or above rated junction temperature. This method of life testing is valid to guarantee the derating curve only if the room temperature operating test brings the junction to maximum operating temperature. If rated junction temperature is not reached during the operating life test, a higher failure rate may be encountered than anticipated if the transistor is operated at an ambient temperature higher than 25°C.
2) Most transistor surface defects are to some extent voltage sensitive. The amount of voltage derating necessary for maximum reliability depends upon the transistor type being considered, because some are more voltage sensitive than others. For maximum reliability it is a good rule to derate collector-base voltage so the transistor is never subjected to any voltage in excess of its collector-emitter voltage rating. The amount of voltage derating used depends to a large extent upon the surge voltage conditions which may be encountered in the application, and the amount of current limiting included in the circuit.
3) For general digital applications, current derating is not a major reliability consideration. However, surge current limitation is very important for modern switching transistors with their relatively small diameter connecting wires.
4) The maximum feasible derating of mechanical stresses is desirable for maximum transistor reliability.

6-8 — Precautions for the Equipment Manufacturer

To insure maximum transistor reliability from incoming inspection through outgoing equipment final test, a number of precautions should be observed. Among these are:

HANDLING PRECAUTIONS:

1. Transistors should be handled in a manner which avoids the possibility of sudden shocks being applied, such as those encountered in dropping from a work bench to a hard floor. Damage done to the transistor by such shocks may not be detected by subsequent testing, yet may cause poor equipment reliability.
2. Any lead trimming or other handling operations such as the attachment of plastic lead spacers should be done with care to avoid damaging the leads or the glass header seals. Hand trimming of leads with pliers should be avoided unless care is taken to avoid pulling the leads.
3. Care must be taken during all soldering operations. Hand soldering should be avoided if possible. If hand soldering is done, a heat sink such as a pair of pliers should be clamped on the lead between the point of application of the soldering iron or gun, and the transistor. Dip soldering should be limited to the minimum time and temperature required to make reliable connections. It is unsafe to exceed the general specification to which transistors are tested for solderability. This is 10^{+2}_{-0} seconds at a temperature of 230° ± 5°C at a point 1/16 ± 1/32 inch from the transistor body.

   Precautions should be taken to prevent solder or flux bridging and causing a conductive path across the bottom of the transistor header.
4. Ultra-sonic cleaning of printed circuit boards should be carefully controlled. The energy level used should be the minimum possible. The presence of standing waves in the bath should be avoided, perhaps by the use of a source with slightly varying frequency. The board should be held as firmly as possible to minimize ultrasonic vibration. The particular method of ultra-sonic cleaning to be used should be thoroughly evaluated to assure that it does not damage transistors.
5. The discharge of static electrical charge through a transistor should be avoided. The charge accumulated by an assembler walking across a floor or even turning on a chair can be sufficient to cause failure if discharged through a transistor.

TESTING PRECAUTIONS:

1) Voltage and current surges must be avoided at any equipment test station. All the transistor leads should be grounded directly at the socket during any test equipment switching or card punching operation. The transmission of surge voltages through common power lines to test equipment has caused transistor failures.
2) The high gain-bandwidth product of presently available switching transistors has led to transistor testing problems as oscillations can occur in test circuits. For example, if a transistor is connected for oper-

ation in the active region using long lead lengths there is a strong possibility of oscillation at some frequency near the self-resonant frequency of the circuit. The emitter and collector wires cause capacity coupling between the emitter and collector, producing positive feedback which may be enough to cause oscillation. This problem arises when measuring the temperature variation of transistor parameters, where the test equipment is placed outside of the test chamber and long leads are run to the transistor. It is also encountered in life test facilities where many transistors are operated using common collector, emitter, or base lead wires.

Any oscillation causes erroneous measurements and may cause transistor burnout. The usual procedure to prevent oscillations is to place isolation resistors in the emitter and collector leads and have the base connected to ground. These isolation resistors should be placed immediately at the transistor socket. The resistors should be noninductive; carbon or deposited film resistors are preferable. The test voltages across the transistor should be read by means of additional isolation resistors connected between the transistor terminals and the voltmeter. To avoid measurement error high impedance measuring equipment must be used.

3) Because of the sensitivity of current gain and saturation voltages to junction temperature, accurate measurements can only be made by short duration pulses that cause no appreciable heating. Pulse duration must be much less than the thermal time constant of the transistor (about 10 milliseconds for low level mesa transistors). Tests done with dc equipment or with curve tracers, which normally sweep at a 60 cps rate, are not suitable for tests where accurate and reproducible measurements are required. For example, the initial and end of life measurements of current gain have to be accurate to within a few per cent to yield meaningful life test data.

4) Tests for I_{CBO}, I_{EBO}, I_{CEX} etc. should use leads which are carefully dressed and shielded so that there is no stray pick-up which could produce serious errors in readings. Electronic micro-microammeters, never dc meter movements, should be used, because the internal inductance of dc meters can generate voltage spikes which can damage the transistor junction.

   For all leakage tests, a suitable resistor should be placed in series with the transistor under test and the supply to limit the current in case the transistor has high leakage or is shorted. If this is not done, complete destruction of the transistor could occur which will prevent further analysis of the failure.

5) When measuring the gain-bandwidth product f_T, the input current source should have an impedance that is high compared to the input impedance of the transistor, in order to approximate a constant current source. At low currents the transistor input impedance may be rather high, with the result that the source does not approximate a constant current source with conventional f_T test fixtures as shown in Figure 6-8. In this case, a tuned circuit should be used to obtain the high impedance

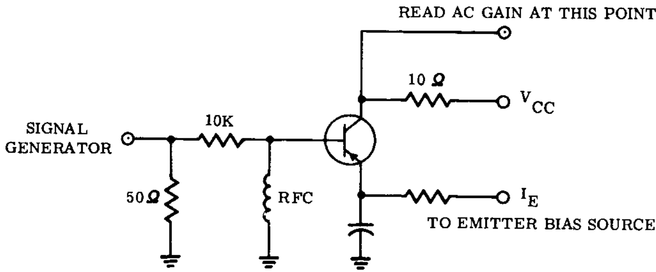


Figure 6-8 — Conventional f_T Test Circuit

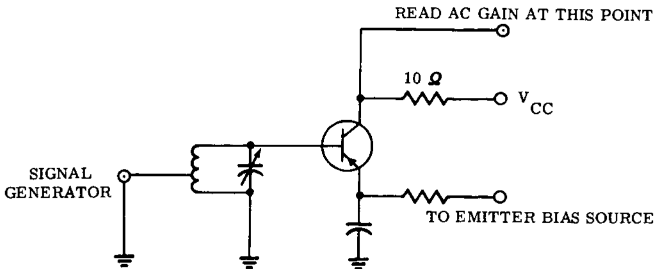


Figure 6-9 — Tuned Input f_T Test Circuit

required for constant current drive to the base, as shown in Figure 6-9. The collector sampling resistor, in either case, should be adjusted so that it is non-reactive at the test frequency and should also be of such a low magnitude that its resistance times the C_{ob} of the transistor forms a time constant that is much less than $1/\omega_T$. Good engineering practice requires that the measurement frequency be in the region where the transistor's current gain is between 2 and 5. f_T is then the product of the measurement frequency and the measured gain.

6) Because of the relatively high speeds of present day switching transistors, carefully constructed testing equipment is required to accurately measure the transient response. The generator driving the input must be terminated carefully to avoid overshoot or ringing on the input pulse.

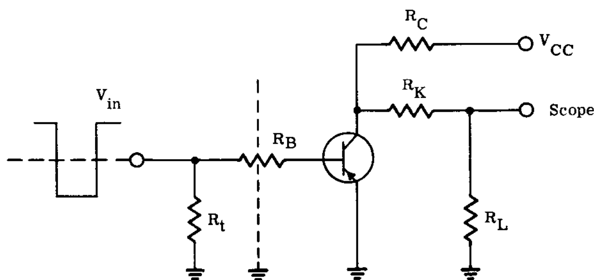


Figure 6-10 Transient Response Test Circuit

In the test circuit, shown in Figure 6-10, R_B is inserted through a hole in a ground plane. This is to prevent the speed-up effect of the end-to-end resistor capacity from causing erroneous readings.

The oscilloscope probe is not placed directly on the collector which would cause capacitance loading but is used in conjunction with a divider. The time constant the divider forms with the probe must be much less than any expected transient time of the transistor.

Precautions regarding the transient response and amplitude of the input pulse were described in Chapter 5.

REFERENCES

1) Von Alven, William H. Introduction *Semiconductor Reliability* Volume II. Engineering Publishers, Elizabeth, New Jersey.

2) West, W. J. "Achievement of High Reliability Levels for Minutemen Solid-State Devices" *Semiconductor Reliability* Volume II. Engineering Publishers, Elizabeth, New Jersey.

3) Peck, D. S. "Uses of Semiconductor Life Distributions" *Semiconductor Reliability* Volume II. Engineering Publishers, Elizabeth, New Jersey.

4) Procassini, A. A. and Romano, A. "Use of the Weibul Distribution Function in the Analysis of Multivariate Life Test Results" Motorola Inc. Semiconductor Products Division, Phoenix, Arizona.

5) Zierdt, C. H. Jr. "Comparison of Operating Life Tests and Storage Tests" *Semiconductor Reliability* Volume I, Engineering Publishers, Elizabeth, New Jersey.

6) Embree, M. L. "Semiconductor Device Ratings — Their Determination and Use in Reliable System Design" *Semiconductor Reliability* Volume II. Engineering Publishers, Elizabeth, New Jersey.

CHAPTER 7

Saturated Mode Circuits

The advantage of saturated mode circuitry can be readily summarized as follows:

1) Low transistor power dissipation
2) Clamped output levels
3) dc conditions that are nearly independent of transistor characteristics

Its chief disadvantage is that saturation does result in storage time which places an upper limit upon switching speed.

This chapter consists essentially of design examples, applying the principles described in detail in Chapters 3, 4, 5, and 6 to worst-case saturated-mode circuit design. The examples are intended to illustrate design principles rather than standard practices. For this reason computer logic techniques[1] are not covered, on the assumption that the logic-circuit designer can reduce the input and output circuit configurations to their simplified equivalent circuits.

This chapter covers design procedures for:

1) Inverters
2) Flip-flops
3) Astable multivibrators
4) Monostable multivibrators

which can be considered the basic family of circuits for any switching system.

Worst case design procedures are used in each example. When all limits used are absolute worst case, an extremely conservative circuit design results which ordinarily increases the total number of parts and the power dissipation of the system. In large systems to achieve high reliability it may be necessary to modify the procedure in accordance to some statistical method, such as Taylor Worst Case[2] or Quantized Probability[3]. The design procedure in this chapter is not modified by this change in the limits. In the examples, it is assumed that worst-case transistor limits are available. These can be obtained directly from data sheets such as the one supplied for the 2N964A transistor, or they can be calculated from typical data sheet information in accordance with procedures described in Chapters 3, 4, and 5.

A $\overline{\text{bar}}$ over a term indicates a maximum value and a $\underline{\text{bar}}$ under a term indicates minimum. These usually are not absolute maximum or minimum limits but rather limits under a particular set of worst-case conditions.

Section 7-1 — Inverter Design

The function of the inverter is to invert the polarity of an input signal producing an output within specified maximum and minimum voltage levels. It also provides gain. That is, the current which it can deliver to a load is A times the current required by its input. It can be designed to provide pulse restoration, or "squaring".

A basic saturated-mode inverter circuit is shown in Figure 7-1-1. When the input signal is at V_1 (a negative voltage for the PNP transistor shown), sufficient current flows through R_K to overcome the current from the $V_{BB} — R_B$ source and supply enough base current to drive the transistor into the saturation region. The output level is $V_O = SV_{CE}$ and maximum current is delivered to load #2. When the input voltage is at V_O, a level near ground, current from the $V_{BB} — R_B$ source flows through R_K and causes a reverse bias to be developed at the base which cuts off the transistor. The current from the $V_{CC} — R_C$ source now flows into the diode producing an output level $V_1 = V_K + V_D$, making current available to load #1. The input current may be many times less than that required by the loads so that the inverter provides current gain.

If the source impedance of the preceding stage is low, the speed-up capacitor C_K, may be very effective in reducing circuit response time. C_K can be chosen from Q_T data.

The analysis of the inverter is straightforward and only the results will be shown. The most difficult problem is to select variables which may appear to be an arbitrary choice. The design procedure will also be different depending upon the criterion which is used to produce an optimum design. The criterion depends upon which is most important; gain, switching speed, or power dissipation and whether the inverter must drive some defined load.

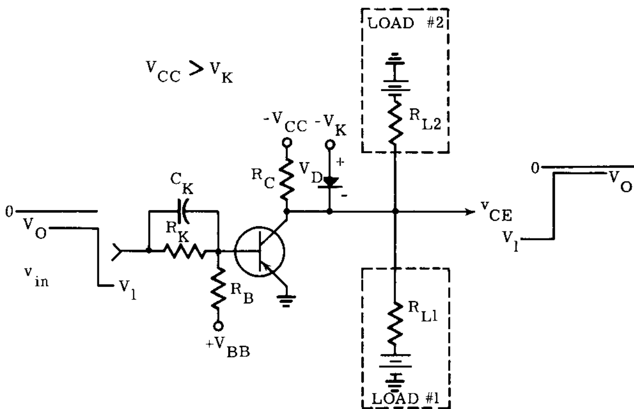


Figure 7-1-1 — Basic RCTL Inverter Circuit

The switching speed as well as the dc gain is almost always maximum if the transistor is operated at its point of maximum gain. The reasoning is quite simply stated:
Since very high gain — as well as minimum gain — transistors can be expected in a typical lot of transistors, the worst-case excess stored charge will be

$Q_X \approx \tau_{BS} I_{B1}$ Where I_{B1} is the base on current as determined by the circuit and τ_{BS} is the storage time constant of the transistor

This charge and therefore storage time will always reduce if the circuit can be altered to reduce I_{B1}. The minimum I_{B1} is, of course, slightly in excess of $\bar{I}_C/\beta_o$ to insure saturation. However, if I_C is reduced in order to lower I_{B1}, τ_A — and therefore rise time — increases for a given circuit gain. Therefore high gain is important.

When starting a new design, a transistor should be picked which has a suitable transient response and high gain, and it should be operated near the peak of its β vs I_E curve. Choosing I_E largely fixes the amount of load current which can be developed. Often, though, the transistor must drive some load which has already been developed. In this case the collector current is determined by the load and a transistor is selected to have high gain at that current. In either case, an understanding of the output circuit is necessary.

7-1-1 — The Load Circuit

A general load circuit is shown in Figure 7-1-2; that is, any conceivable output circuit is of this form. For example, R_{L1} might represent a diode "and" gate load and R_{L2} a diode "or" gate load. In general, these loads would vary over a wide range depending upon the state of other circuits in the system. With variable loads, it is almost always necessary to employ clamps to hold the output levels within reasonable limits. The transistor, when on, clamps the "0" level, while the diode clamps the "1" level when the transistor is off. Under the special case of a nearly constant resistance load, such as occurs with RTL logic, the diode may be omitted.

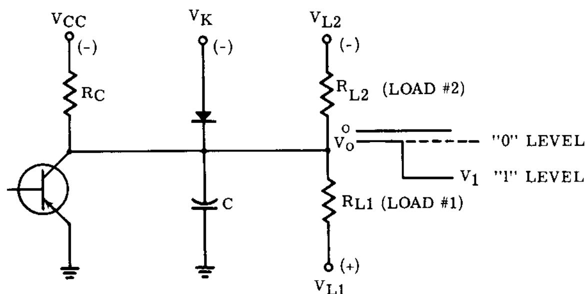


Figure 7-1-2 — The General Load Circuit

VOLTAGE LEVELS: The first problem to be resolved is to choose the nominal levels and tolerance for the output levels, V_0 and V_1. The charge moved when switching from V_0 to V_1 is $\int i\,dt = C\,(V_1 - V_0)$. If switching speed is to be increased, either the signal swing must be reduced or current levels increased to reduce the effect of capacitance. For the common values of 10 to 50 mA of collector current employed with transistors of the 50 to 200 mW class, signal swings of 3 to 12 volts have been used. It should be obvious that as the level is reduced the tolerance must be tightened to preserve a reasonable amplitude difference between $\underline{V}_1$ and $\overline{V}_0$. This difference is a very significant figure; all stages must be designed to be in the proper state with either level applied, and must change state within a given amount of time, when the input changes from one level to the other.

Suitable system levels, for low level high speed transistors, which will be used in some of the design examples in this chapter are listed in Table 7-1-1.

TABLE 7-1-1

Useful Voltages for System Levels		
	Min.	Max.
V_0	0	1.0
V_1	5	7.2

In this table, V_1 has a larger tolerance than V_0 since it is derived from the clamp diode and its power supply which will have much more variation due to temperature and loading conditions than V_0 which is derived from a saturated transistor.

The next step in the design would be to establish limits to be used in the design of the input and output circuits of stages. Since the levels usually deteriorate in transmission through a system, the tolerance at the input could be expected to be greater than that at the output. The specifications for an inverter might look as shown in Table 7-1-2.

TABLE 7-1-2

Available Input Levels			Required Output Levels	
	Min.	Max.	Min.	Max.
V_0	0.0	1.0	0.0	0.5
V_1	5.0	6.9	6.0	7.2

The choice of levels must be made with reasonable power supply voltages and tolerances, the dc saturation and gain characteristics of a transistor having the desired transient response, and the forward drop and speed of a suitable diode in mind. The breakdown rating of each semiconductor element should also be considered. These characteristics are easily found having designer's data available as discussed previously in Chapters 3, 4 and 5.

DETERMING V_{CC} AND R_C: Once suitable levels and active devices have been selected, resistor R_C and the power supply voltage V_{CC} must be determined. This problem can be solved by analyzing the output circuit as shown in Figure 7-1-3. Here the loads have been replaced simply by the currents that they are required

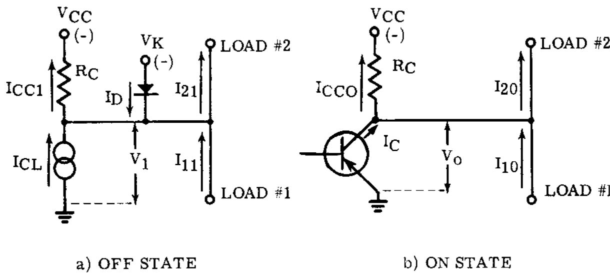


Figure 7-1-3 — Equivalent Load Circuit During the Two States

to draw when either a "1" or "0" is present at the output. That is, I_{11} refers to the current drawn by load 1 when the output is at V_1, etc.

Consider the action of the circuit. When the transistor is off, (Figure 7-1-3a) the output is at V_1; and the maximum current required by load 1 is being delivered. Some minimum current may be flowing from load 2 which will assist in supplying I_{11}; the remainder of load current as well as a small amount to keep the diode in conduction and to compensate for I_{CL} must be supplied by R_C from V_{CC}. Thus

$$\overline{R}_C = \frac{\underline{V}_{CC} - V'_1}{\underline{I}_D + \overline{I}_{11} - \underline{I}_{21} + \overline{I}_{CL}} \tag{7-1-1}$$

where the denominator is the minimum current ($\underline{I}_{CC1}$) through R_C needed to supply the load when V_{CC} is minimum. $\overline{I}_{11} - \underline{I}_{21}$ may be thought of as a net maximum load current $\overline{I}_1$ delivered in the "1" state. The normal case here occurs when $\overline{I}_{11} > \underline{I}_{21}$. If this were not so, load 2 could supply all the current required by load 1 and R_C would be unnecessary, but this condition is not often encountered. The voltage V'_1 is a particular value of V_1. The exact value of V'_1 depends upon several factors which will be discussed shortly.

When the transistor is on, (Figure 7-1-3b) the output is V_0. In this case attention is focused upon supplying the maximum current required by load 2. Some minimum current will be flowing from load 1 to assist; the rest of the load current plus that demanded by R_C will have to be supplied by the transistor. In this case, $\overline{I}_C$ is given by

$$\overline{I}_C = \frac{\overline{V}_{CC} - \underline{V}_0}{\underline{R}_C} + \overline{I}_{20} - \underline{I}_{10}. \tag{7-1-2}$$

Here the difference between $\overline{I}_{20}$ and $\underline{I}_{10}$ may be thought of as a net maximum load current $\overline{I}_0$ delivered in the "0" state.

CLAMP DIODE: The primary purpose of the clamp diode is to prevent $\overline{V}_1$ from becoming very large when the load current is minimum. As the load current decreases, the drop across R_C also decreases and V_1 increases unless the diode clamps. Use of the clamp diode also allows V_{CC} to be several times V_1 which is desirable in order to have the collector current low for a given load current. The diode greatly minimizes trouble due to the excessive drive which would occur in other transistor stages being driven from the inverter output. The value to use for V'_1 in equation 7-1-1 is $\underline{V}_1$ if the diode is not used or if it is permissible to have the diode cut off when maximum load current is being delivered. However, a higher value for V'_1 is required if the diode is to remain always in conduction even when the load current is maximum. This condition is desirable because it keeps the collector circuit impedance low (in the order of 26 ohms/mA of I_D), thereby minimizing noise problems.

Imagine a theveninized equivalent of the general output configuration as seen by the diode. It will be simply an effective voltage V'_1 in series with an effective resistance R'. If the diode is to conduct a given amount, it should be clear that the voltage V'_1 must equal $\overline{V}_D + \overline{V}_K$. However, $\underline{V}_1 = \underline{V}_K + \underline{V}_D$. This means that the actual $\underline{V}_1$ cannot be used in equation 7-1-1 to calculate R_C if it is desired to keep the diode always in conduction. Rather $V'_1 = \overline{V}_K + \overline{V}_D$ must be used where $\overline{V}_D$ is determined at minimum diode current which is selected to be several mA to keep the diode impedance low.

To select diodes and determine their stored charge, maximum diode current must be known. It flows when the load current is minimum and may be expressed as:

$$\overline{I}_D = \overline{I}_{21} - \underline{I}_{11} + \frac{\overline{V}_{CC} - \underline{V}_K - \underline{V}_D}{\underline{R}_C} \qquad (7\text{-}1\text{-}3)$$

The last term in equation 7-1-3 is the current $\overline{I}_{CC1}$ which flows when R_C is minimum and V_{CC} is maximum.

RATIO OF COLLECTOR CURRENT TO LOAD CURRENT: In the following discussion, it will be shown that it is preferable to have V_{CC} and R_C large from a circuit efficiency viewpoint.

In order to obtain a useful solution it is necessary to define:

$$\overline{R}_C = (1 + n_R)\, R_C, \qquad \underline{R}_C = (1 - n_R)\, R_C$$
$$\overline{V}_{CC} = (1 + n_P)\, V_{CC}, \qquad \underline{V}_{CC} = (1 - n_P)\, V_{CC}$$

Where n_R and n_P are the resistor and power supply tolerances respectively.

Since the resistor tolerance appears in a number of equations, it is convenient to define a resistance tolerance modifier as:

$$N_R = \frac{1 + n_R}{1 - n_R} = \frac{\overline{R}}{\underline{R}}\,.$$

In this expression n_R is expressed as a decimal. With n_R expressed as a percent, N_R is plotted in Figure 7-1-4. Since this ratio is encountered often, Figure 7-1-4 is a useful design tool.

Using the previous definitions for $\overline{I}_0$ and $\underline{I}_{CC1}$, write equation 7-1-1 as

$$(1 + n_R)\, R_C = \frac{(1 - n_P)\, V_{CC} - V'_1}{\underline{I}_{CC1}}$$

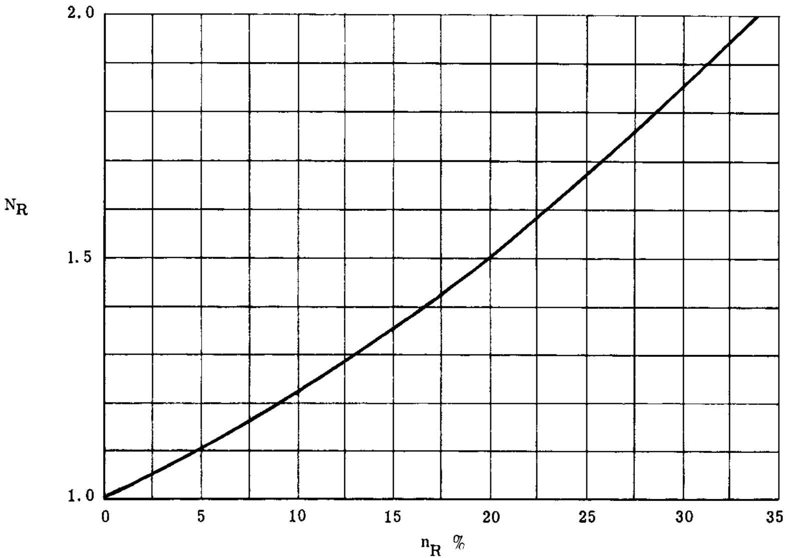


Figure 7-1-4 — Tolerance Multiplier

and write equation 7-1-2 as ($V_0 = 0$)

$$\bar{I}_C = \frac{(1 + n_P)\, V_{CC}}{(1 - n_R)\; R_C} + \bar{I}_0$$

Combining these equations to solve for a term γ which is defined as the ratio of the current through R_C for a "0" ($\bar{I}_{CCO} = \bar{I}_C - \bar{I}_0$) to the current for a "1" ($\underline{I}_{CC1}$) it is found that

$$\gamma = \frac{\bar{I}_C - \bar{I}_0}{\underline{I}_{CC1}} = N_R \frac{1 + n_P}{1 - n_P - V'_1/V_{CC}} \tag{7-1-4}$$

Note that the resistor tolerance N_R appears as a constant multiplier of considerable importance. Equation 7-1-4 is plotted in Figure 7-1-5 with power supply tolerance as a parameter. In many circuits I_0 is zero and I_{CC1} essentially equals I_1 which means that γ is the ratio of collector current to load #1 current. This graph clearly shows the benefits of making V_{CC} — R_C approach a current source. Note the severely high ratio of current through R_C and the collector, to the current through the load, which results as V_{CC} approaches V_1.

MINIMIZING POWER DISSIPATION: The power dissipated in R_C imposes a practical upper limit upon V_{CC} and R_C. Notice frcm Figure 7-1-5 that as V'_1/V_{CC} becomes small, γ approaches unity. This indicates that the $V_{CC}R_C$ source is approaching a current source because the current through R_C for a "0" output level approaches that for a "1" output level. The drop across R_C, therefore, approaches V_{CC} making the power dissipated in R_C proportional to V_{CC}. At the other limit, as $V'_1V/_{CC}$ approaches unity, γ becomes very large, making the current through R_C become very large at the "0" output level. At this limit the

dissipation in R_C becomes proportional to γ. Somewhere between these extremes there is a point where the dissipation in R_C is minimum. In a given circuit, maximum dissipation occurs under conditions of "0" output and can be written with the aid of equation 7-1-4

$$P_D = \overline{I}_{CC0}\overline{V}_{CC} = \gamma \underline{I}_{CC1}\overline{V}_{CC} = N_R \frac{(1 + n_P)\underline{I}_{CC1}V_{CC}}{1 - n_P - V'_1/V_{CC}}.$$

This equation can be solved for the value of V_{CC}/V'_1, which will minimize P_D, by taking the derivative, holding $\underline{I}_{CC1}$ constant, and setting the derivative equal to zero. This yields

$$\frac{V_{CC}}{V'_1} = \frac{2}{1 - n_P} \qquad (7\text{-}1\text{-}5)$$

This result simply says that for minimum dissipation in R_C, the minimum collector supply voltage should be twice the minimum output voltage. However, the current and accordingly the dissipation in the transistor approaches twice what it would be with $V_{CC} >> 2V'_1$ as seen from Figure 7-1-5. Usually it is preferable to keep collector current as low as possible rather than to minimize circuit dissipation, which makes use of a V_{CC} from 3 to 6 times V'_1 a good compromise. Figure 7-1-5 is valuable as an aid in the determination of suitable output networks.

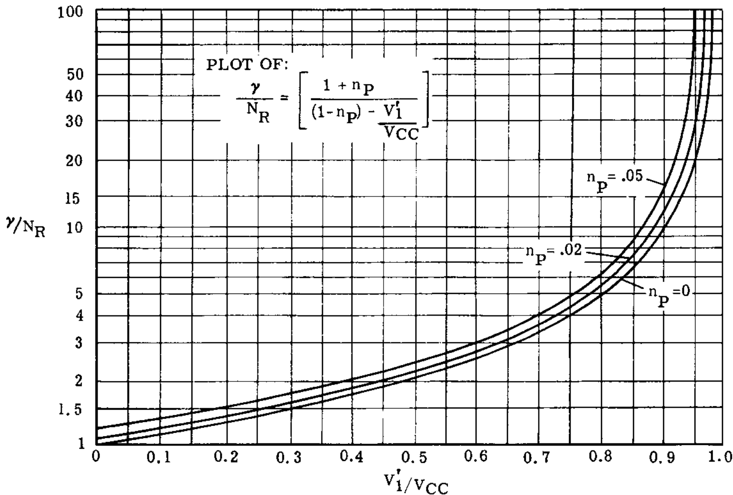


Figure 7-1-5 — Ratio of Collector to Load Current vs. Ratio of Output to Supply Voltage

EFFECT OF LOAD CAPACITANCE: In many cases, the load circuit has a capacitor to ground. To improve switching speed the collector resistor can be chosen on the basis of transient response rather than dc conditions. The capacitance could be stray or line capacitance to ground, or a speed-up capacitor used in an inverter or flip-flop, or a timing capacitor used in a multivibrator. (Speed-up or timing

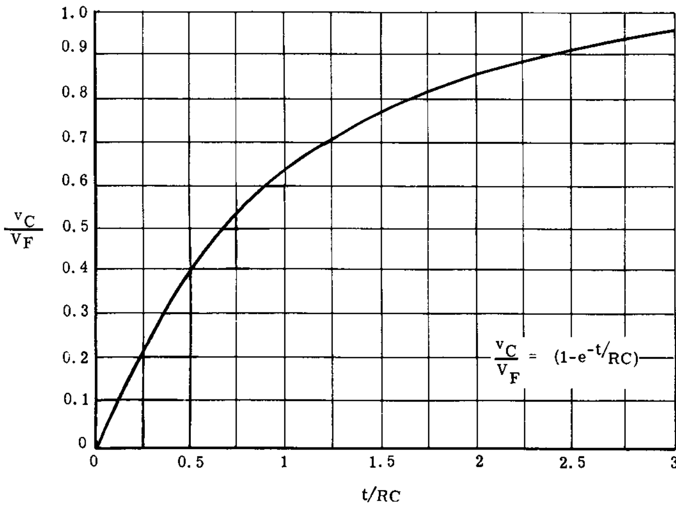


Figure 7-1-6 — Plot of Rise time Function

capacitors have one end connected to the base of a transistor which is a virtual ground when the transistor is on.)

When a transistor is being turned on, it can easily drive a capacitive load, since the capacitor appears to the input signal reduced in magnitude by the current gain of the device. However, when the transistor is being turned off, once its current drops to zero the transistor can no longer drive the load capacitor. The capacitor decay then is governed only by the R-C time constant of the load.

The general R-C circuit behavior is given by

$$v_C = V_F(1 - e^{-t/RC}) \qquad (7\text{-}1\text{-}6a)$$

or

$$t = RC \ln \frac{V_F}{V_F - v_C} \qquad (7\text{-}1\text{-}6b)$$

where v_C is the capacitor voltage as a function of time t
V_F is the final voltage which the capacitor can achieve; i.e.: the voltage which it "sees".

These expressions assume that initially the capacitor voltage is zero.

Notice from Figure 7-1-6, which shows general R-C circuit behavior, that the time taken to reach 90% of the final voltage is 2.3 time constants. This time can be considerably shortened if a clipping circuit is used to remove the slowly changing portion of the exponential. The clamped output circuit can accomplish this clipping if R_C is chosen so that the level V_1 would assume in the absence of the diode is greater than the actual V_1 desired. The diode is often called a "catch" diode, when used in this manner.

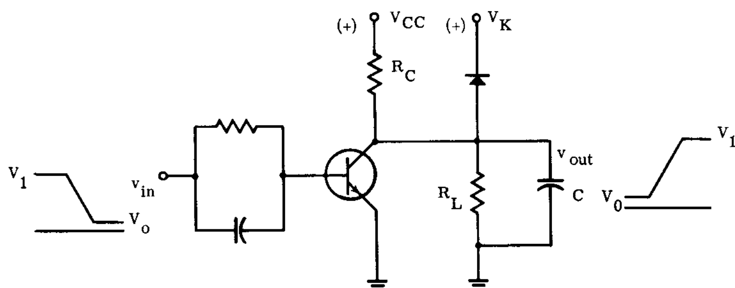


Figure 7-1-7a — Circuit Used in Transient Analysis

The circuit of Figure 7-1-7a will be analyzed to obtain design guides in order to choose values for the components in the output network.

In the absence of the diode, the 90% rise time will be:

$$t_r = 2.3\, R'_C C \tag{7-1-7}$$

where:

$$R'_C = \frac{R_C R_L}{R_C + R_L}.$$

The output voltage V_1 will equal the final voltage V_F.

$$V_F = \frac{R_L V_{CC}}{R_L + R_C} \tag{7-1-8}$$

A comparison between a clipped and an unclipped case is informative. If it is desired to keep V_1 at the same point in each case, V_K must be chosen so that $V_K = V_1 - V_D$. Assume that V_K has been determined according to the previous criterion, then, V_{CC} is doubled, and all other values held fixed. In this case, the voltage which the capacitor is trying to charge toward would be $2V_1$. However, the diode "catches" the voltage at V_1. From Figure 7-1-6, observe that the time to bring v_C to .5 V_F is .69 RC. The on current through the switch has doubled since V_{CC} has doubled. However, the rise time has decreased by a factor of 3.34, which represents a considerable improvement. Rise time could be further improved by increasing V_{CC} even more, but the improvement would be roughly proportional to the increase in current through the switch since the slope of the curve is becoming more nearly constant.

As intuitively expected from the exponential behavior, raising V_F a slight amount above V_1 results in a considerable improvement in rise time. Further increases of V_F result in progressively smaller decreases in time on a percentage basis. Figure 7-1-7b shows the nature of the compromise where the amount of improvement in rise time (to the 90% point), due to clamping, divided by the increase of switch current, is plotted against the ratio of final voltage to clamp voltage. It is seen that there is little percentage increase in the clipping improvement factor (F_K) for $V_F/V_1 > 3$.

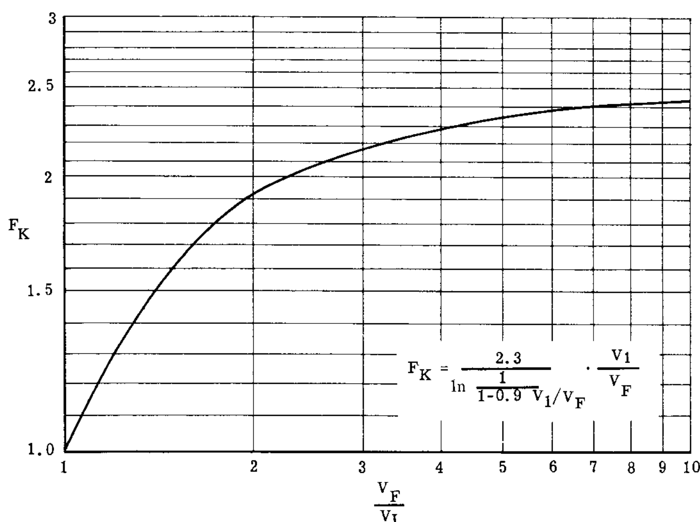


Figure 7-1-7b — Improvement Factor Resulting from Clamping

Figure 7-1-7b was plotted by taking the ratio of rise time to the 90% point for the unclipped waveform ($V_1 = 0.9\ V_F$) to that obtained from the general equation under conditions of V_1/V_F variable and dividing that result by the increase in the on switch current which results. That is,

$t_{r1} = 2.3\ R'_C\,C$ when $V_1 = 0.9\ V_F$, the unclipped case.

$$t_{r2} = R'_C\ C\ \ln \frac{1}{1 - 0.9\ V_1/V_F},\ \text{for variable } V_1/V_F. \qquad (7\text{-}1\text{-}9)$$

Under all conditions:

$$I_C = V_F/R'_C = V_{CC}/R_C \qquad (7\text{-}1\text{-}10)$$

When clipping is not employed, $V_F = V_1$.

Therefore, $I_{C1} = \frac{V_1}{R'_C}$ for the unclipped case and $I_{C2} = \frac{V_F}{R'_C}$, for variable V_1/V_F.

Combining these relations

$$F_K = \frac{t_{r1}/t_{r2}}{I_{C1}/I_{C2}} = \frac{\dfrac{2.3\ R'_C C}{R'_C C \ln\left(\dfrac{1}{1 - 0.9V_1/V_F}\right)}}{\dfrac{V_1}{R'_C} \Big/ \dfrac{V_F}{R'_C}}$$

which simplifies to:

$$F_K = \frac{2.3\left(\frac{V_F}{V_1}\right)}{\ln\left(\frac{1}{1 - 0.9V_1/V_F}\right)} \qquad (7\text{-}1\text{-}11)$$

These results can be extended into a useful design guide for cases when it is necessary to change the level on a capacitive load within a fixed time.

Substituting equation 7-1-8 into equation 7-1-9 and rearranging yields

$$t_{r2} = R'_C C \ln \frac{\frac{V_{CC}}{V_1} R_L}{R_L\left(\frac{V_{CC}}{V_1} - 0.9\right) - 0.9\, R_C} \qquad (7\text{-}1\text{-}12)$$

The load current is

$$I_1 = \frac{V_1}{R_L}.$$

Combining this expression with equations 7-1-7 and 7-1-12, the result is

$$t_{r2} = \frac{V_{CC}}{V_1}\,\frac{R_L C}{(V_{CC}/V_1 + I_C/I_1)} \ln \frac{V_{CC}/V_1}{(V_{CC}/V_1 - 0.9) - \frac{0.9\, V_{CC} I_1}{V_1\, I_C}} \qquad (7\text{-}1\text{-}13)$$

It is convenient to define a load time constant τ_L as

$$\tau_L = R_L C$$

from which the ratio of t_r to τ_L can be written

$$t_r/\tau_L = \frac{V_{CC}/V_1}{V_{CC}/V_1 + I_C/I_1} \ln \frac{V_{CC}/V_1}{(V_{CC}/V_1 - 0.9)\frac{V_{CC}/V_1}{I_C/I_1}}. \qquad (7\text{-}1\text{-}14)$$

This result is plotted as Figure 7-1-8. From this graph, the required collector current to charge the load capacitance in a given time can be easily found. The place where the curves end, as t_r/τ_L increases, is the point where the clipping circuit has no effect, i.e.: $V_1 = V_F$. These curves also reflect the result of Figure 7-1-7b; less collector current is required to charge the capacitor in a given time interval as V_{CC} is allowed to become higher.

7-1-2 — Coupling for Maximum Transfer Efficiency

Cascaded inverters are required when it is necessary to obtain more gain than a single stage inverter can supply. Also, a second stage is often necessary to isolate a variable load from a multivibrator. In these situations, the interstage voltage (V_1) is not required to be at any particular level; therefore an opportunity is provided to minimize the collector current of the first stage by making R_C as

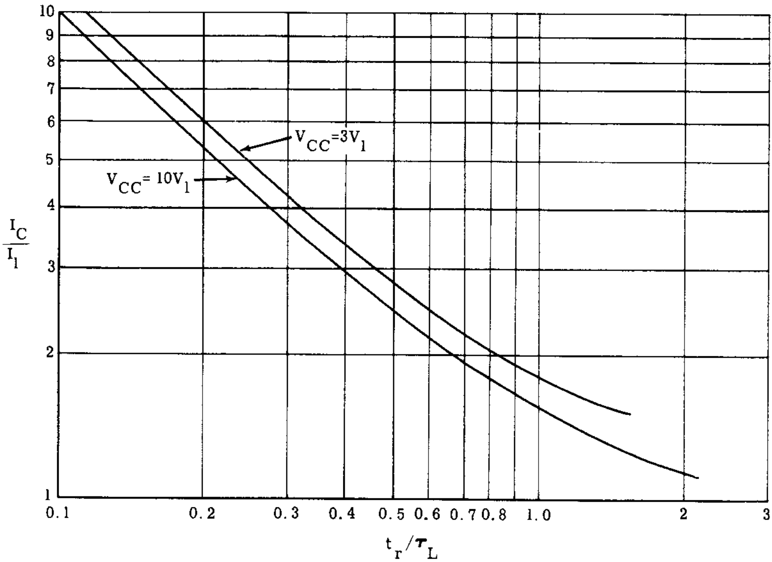


Figure 7-1-8 — Ratio of Collector Current to Load Current

large as possible. The only requirement is that conditions be such that sufficient base current in the on state, and off bias,in the off state,is applied to Q_2 as shown in Figure 7-1-9.

The governing equations can easily be written from Figure 7-1-9

$$\frac{\underline{V}_{CC} - \overline{V}_{BE}}{\overline{R}_K + \overline{R}_C} - \frac{\overline{V}_{BB} + \overline{V}_{BE}}{\underline{R}_B} = \underline{I}_B \qquad (7\text{-}1\text{-}15)$$

$$-\frac{\underline{V}_{OB} + \overline{S}V_{CE}}{\underline{R}_K} + \frac{\underline{V}_{BB} - \underline{V}_{OB}}{\overline{R}_B} = \overline{I}_{BL} \qquad (7\text{-}1\text{-}16)$$

Equations 7-1-15 and 7-1-16 need to be combined into a single equation so that a differentiation can be performed, the result set equal to zero, and optimum conditions found for R_B, R_K, and R_C. Solving for R_B in equation 7-1-16, substituting it into equation 7-1-15, and rearranging, it is found:

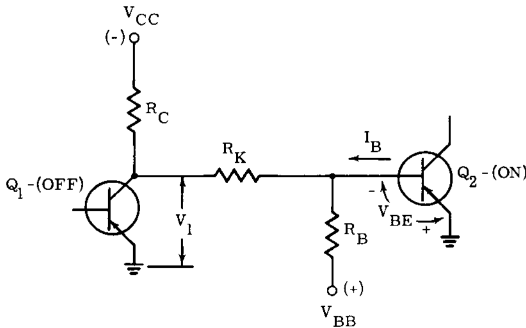
$$\frac{1}{1 + \overline{R}_C/\overline{R}_K} = \frac{1}{(\underline{V}_{CC} - \overline{V}_{BE})}\left[\underline{I}_B + \overline{I}_{BL}(1 + n_B)\left(\frac{\overline{V}_{BB} + \overline{V}_{BE}}{\underline{V}_{BB} - \underline{V}_{OB}}\right)\right]\overline{R}_K$$

$$+ \frac{(\overline{V}_{BB} + \overline{V}_{BE})\ (\underline{V}_{OB} + \overline{S}V_{CE})}{(\underline{V}_{BB} - \underline{V}_{OB})\ (\underline{V}_{CC} - \overline{V}_{BE})}(1 + n_B)N_K\,.$$

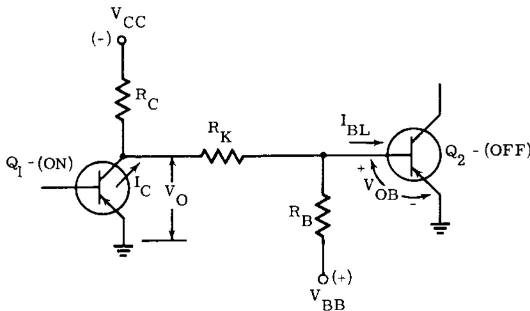
which is of the form

$$\frac{1}{1 + \overline{R}_C/\overline{R}_K} = A\,\overline{R}_K + B.$$

where A is the coefficient of R_K and B is the constant. A and B contain only terms which are known. Solving for $\overline{R}_C$



a) CONDITIONS WITH Q_1 OFF and Q_2 ON



b) CONDITIONS WITH Q_1 ON and Q_2 OFF

Figure 7-1-9 — Interstage Network

$$\bar{R}_C = \frac{\bar{R}^2_K(1 + B) - A\,\bar{R}^2_K}{A\,\bar{R}_K + B}. \tag{7-1-17}$$

Taking the derivative yields

$$\frac{d\bar{R}_C}{d\bar{R}_K} = \frac{(A\,\bar{R}_K + B)(1 + B - 2A\,\bar{R}_K) - \bar{R}_K(1 + B - A\,\bar{R}^2_K)\,A}{(A\,\bar{R}_K + B)^2}.$$

Setting this expression equal to zero and solving

$$\bar{R}_K = -B/A \pm \sqrt{\left(\frac{B}{A}\right)^2 + \frac{B}{A^2}(1 + B)}\,.$$

Since R_K must be positive, only the positive radical is significant and upon simplification the result is

$$\bar{R}_K = \frac{B}{A}\left(\sqrt{2 + 1/B} - 1\right). \tag{7-1-18}$$

The value computed for R_K could be inserted in equations 7-1-15 and 7-1-16 and they could be solved simultaneously for R_C and R_B. Or since A and B must be computed in order to solve equation 7-1-17, R_C may be solved by putting the expression for $\bar{R}_K$ into equation 7-1-17 which yields

$$\bar{R}_C = \frac{\frac{(1+B)}{A}(\sqrt{2+1/B}-1) - \frac{B}{A}(\sqrt{2+1/B}-1)^2}{\sqrt{2+1/B}} \qquad (7\text{-}1\text{-}19)$$

Further considerations involved in this coupling circuit apply as discussed in the following section on the input network.

7-1-3 — The Input Network

Design of the input network when V_1 and V_0 have discrete values allows less choice than the output network and is less complex. It can be designed solely on the basis of dc conditions, assuming that if necessary, a capacitor can be used to enhance speed. If resistance driving is employed, the input network can be designed on the basis of transient response which will be discussed later.

The input network can be found by solving the following simultaneous equations which are easily written from Figure 7-1-9.

$$\frac{\underline{V}_{OB} + \bar{V}_o}{R_K(1-n)} - \frac{\underline{V}_{BB} - \underline{V}_{OB}}{R_B(1+n)} + \bar{I}_{BL} = 0 \qquad (7\text{-}1\text{-}20)$$

$$\frac{\underline{V}_1 - \bar{V}_{BE}}{R_K(1+n)} - \frac{\bar{V}_{BB} + \bar{V}_{BE}}{R_B(1-n)} - \underline{I}_B = 0 \qquad (7\text{-}1\text{-}21)$$

Depending upon the transistor characteristic, either the high or the low temperature extreme could be worst-case. The off bias voltage (V_{OB}) must be greater than V_{TR}, which is defined as the reverse base-emitter voltage required to hold I_{CL} to a value near its minimum. $\bar{V}_{TR}$ & $\bar{I}_{BL}$ always occur at high temperature limit and must be determined from transistor data. When determining high temperature data only, the junction temperature to be used should include the increase over ambient caused by power dissipation in the on condition. $\bar{V}_{BE}$ always occurs at the low temperature limit. However, $\underline{I}_B$ and $\bar{S}V_{CE}$ occur at high temperatures, then $\bar{V}_{BE}$ used in the equation should also be a high temperature limit.

The bias source V_{BB} may be open to choice. Reasoning similar to that applied to the output network will show that V_{BB} — R_B should approach a current source; that is V_{BB} should be large compared to the voltage shift at the base.

Regardless of the values used for R_B and V_{BB} a specifiied current through R_B is required to establish V_{OB}. For an example, assume that V_{BB} is just a few times larger than V_{OB}. When the input level goes to V_1, the voltage at the base must reverse polarity and increase to V_{BE}. This voltage shift will increase the current through R_B, resulting in an unnecessary reduction in drive current as all the current through R_B must be overcome by the input current to provide base current. A useful "rule of thumb" is to make V_{BB} at least as large as V_1. Power dissipation in R_B is low and therefore is usually not a factor.

The coupling capacitor C_K can be selected using Q_T data as outlined in Chapter 5.

The principles so far discussed are illustrated in the inverter design example. This example is done in detail to provide a guide for using the designer's data sheet to obtain worst-case values and also, to illustrate the principles of inverter design.

7-1-4 — Minimizing Switching Time of RTL Stages:

By using the relationships derived in Chapter 5, the proper relationship between I_{B1} and I_{B2} can be found to minimize the switching time for a given stage gain, A. In RTL circuits where speed is important, overdrive is heavy and recombination can be neglected. The appropriate relations to use are then:

$$t_r = \tau_A\, I_C/I_{B1} \qquad t_s = \tau_{BS}\frac{I_{B1}}{I_{B2} + I_{B1}/2}$$

$$t_f = \tau_A\, I_C/I_{B2} \qquad t_d = Q_{OB}/I_{B1}$$

As shown in Figure 7-1-10, in RTL circuits, the input current I_K must overcome I_{B2} from the bias source and provide I_{B1} to the transistor. Defining the stage gain as

$$A = \frac{I_C}{I_K} = \frac{I_C}{I_{B1} + I_{B2}} \qquad (7\text{-}1\text{-}22)$$

I_C can be written as

$$I_C = A\, I_{B1} + A\, I_{B2}.$$

The off bias voltage is given approximately as $I_{B2}\, R_K$. Thus

$$Q_{OB} = I_{B2} R_K\, C_{in}.$$

where C_{in} is the effective input capacitance due to C_{ib} and C_{ob}. The product $C_{in}\, R_K$ can be regarded as an input time constant τ_{in}.

Substituting these relationships in the RTL transient equations, the following equations result

$$t_r = \tau_A \frac{A\, I_{B1} + A\, I_{B2}}{I_{B1}} \qquad t_s = \tau_{BS}\frac{I_{B1}}{I_{B2} + I_{B1}/2}$$

$$t_f = \tau_A \frac{A\, I_{B1} + A\, I_{B2}}{I_{B2}} \qquad t_d = \frac{\tau_{in}\, I_{B2}}{I_{B1}}$$

In order to find the proper relationship between I_{B1} and I_{B2} define the drive ratio D as

$$D = I_{B1}/I_{B2} \qquad (7\text{-}1\text{-}23)$$

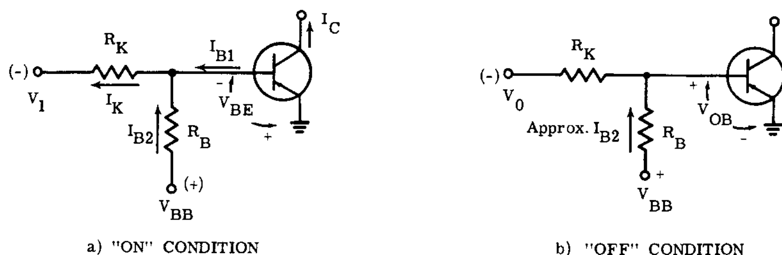


Figure 7-1-10 — Conditions for an RTL Circuit

By using the drive ratio D and summing all the transistor times the total switching time (t_t) can be found.

$$t_t = t_r + t_f + t_s + t_d$$

$$= \tau_A(A + A/D + AD + A) + \tau_{BS}\frac{D}{1 + D/2} + \frac{\tau_{in}}{D}$$

In order to get a tractable solution the D/2 term in the t_s expression is neglected. This over-emphasizes the importance of I_{B2} in minimizing storage time and leads to a result which yields I_{B2} higher than desirable. However, the solution is still useful as a "rule of thumb".

Taking the derivative yields

$$\frac{dt_t}{dD} = \tau_A\left(-\frac{A}{D^2} + A\right) + \tau_{BS} - \frac{\tau_{in}}{D^2}$$

Setting the derivative equal to zero and solving

$$D_{opt} = \sqrt{\frac{A\,\tau_A + \tau_{in}}{A\,\tau_A + \tau_{BS}}} \qquad (7\text{-}1\text{-}24)$$

The result shows that when $\tau_{in} > \tau_{BS}$, I_{B1} should be larger than I_{B2}; for $\tau_{in} < \tau_{BS}$ I_{B1} should be less than I_{B2}. If τ_{in} and τ_{BS} are negligible $I_{B1} = I_{B2}$. In a general way, the equation is in conformity with what intuitive reasoning would conclude. However, the result should not be considered as an exact answer since the storage time equation was oversimplified and the effects of recombination neglected.

Once A is chosen and D_{opt} is determined, the switching times are determined for a given transistor type. The forced gain (β_F) can be found by combining equations 7-1-22 and 7-1-23.

$$\beta_F = \frac{I_C}{I_{B1}} = A\,(1 + 1/D) \qquad (7\text{-}1\text{-}25)$$

Using equations 7-1-23 and 7-1-25 to solve for I_{B1} and I_{B2}, the input network values can be determined from the loop equations which can be written from Figure 7-1-10.

$$I_{B2} = \frac{V_{BB} - V_{BE}}{R_B} + \frac{V_{BE} - V_O}{R_K} \qquad (7\text{-}1\text{-}26)$$

$$I_{B1} = \frac{V_1 - V_{BE}}{R_K} - \frac{V_{BB} - V_{BE}}{R_B} \qquad (7\text{-}1\text{-}27)$$

Once nominal values are found, worst case conditions can be calculated and maximum switching times computed.

7-1-5 — Tolerances of Passive Elements

In the design procedures to be given in this chapter, often two or three simultaneous equations must be solved. Rarely will the value obtained for a component fall very close to a standard value. It is important, therefore that the tolerances used in the equations not only include the effects of temperature, aging, and initial deviation from nominal, but also include a factor representing the worst deviation from a nominal standard part value that a component might calculate to be.

The standard values for high quality precision resistors are in approximately 1% increments. Thus the maximum error between a calculated value and a standard part would be ½% which is normally of little concern. It is standard practice to allow worst case limits for these resistors to be ± 5% to account for this initial tolerance problem as well as for aging and variations due to temperature. By obtaining information for a specific type, this overall tolerance might be found to be considerably less.

For carbon resistors, the standard EIA values are in 10% increments. Thus, even if 5% tolerance resistors are used, a maximum difference of 5% could exist between a calculated value and a standard part value. These resistors are also less stable with temperature and age. To account for all these deviations it is standard practice to use a tolerance of ±20% for 5% components. When using these resistors it is very important to obtain tolerance data as the use of wide tolerance resistors imposes a severe penalty upon circuit performance. This problem is illustrated by the example flip-flop design.

Capacitors cause less trouble because in circuits the tolerance does not multiply as resistor tolerance does. Normally only minimum values are required at some worst case temperature. Silver mica capacitors fill the needs of most trigger and coupling capacitors while inexpensive ceramic types serve adequately as bypass capacitors.

TABLE 7-1-3 INVERTER DESIGN DATA

	AVAILABLE INPUT LEVELS (VOLTS)		REQUIRED OUTPUT LEVELS (VOLTS)	
	Min.	Max.	Min.	Max
V_0	0	—1.0	0.0	—0.5
V_1	—5.0	—6.9	—6.0	—7.2

LOAD CURRENT REQUIREMENTS (mA)

	I_1		I_2	
	Min.	Max.	Min.	Max.
At Output Level V_0	1	20	4	16
At Output Level V_1	2	30	2	8

AVAILABLE POWER SUPPLY VOLTAGES

Minimum	—17.1	+17.1	—11.4	+11.4	—5.4	+5.4
Nominal	—18	+18	—12	+12	—5.5	+5.5
Maximum	—18.9	+18.9	—12.6	+12.6	—5.6	+5.6

AMBIENT TEMPERATURE OF INVERTER ENVIRONMENT

—55 to +65°C

COMPONENT TOLERANCES

Resistor tolerances ±5% end of life

(±1% initial) ∴ $N_R = 1.105$.

7-1-6 — RCTL Design Example

Procedure	Example
Step 1 List specifications.	1. Specifications are shown in Table 7-1-3.
Step 2 Obtain a fairly accurate estimate of the maximum collector current. Use $\underline{I}_{CC1} = \overline{I}_{11} - \underline{I}_{21} + \underline{I}_D$ (neglect I_{CL}) $I_0 = \overline{I}_{20} - \underline{I}_{10}$ Use Fig. 7-1-5 to choose V_{CC} — Find γ Then $\overline{I}_C = (\gamma \underline{I}_{CC1} N_R + \overline{I}_O)$.	2. $\underline{I}_{CC1} = 30 - 2 + 2 = 30$ mA 2 mA was assumed for I_D. The wide variations of I_1 and the tight tolerance of V_1 dictate the use of a diode. $\overline{I}_0 = 16 - 1 = 15$ mA. Use the -18 V supply to keep $\overline{I}_C$ as low as possible. Calculate $\underline{V}_1/V_{CC} = 6/18 = .333$ and read $\gamma/N_R = 1.7$ for $n_P = .05$. Then $\overline{I}_C = (1.7)(30)(1.1) + 15 = 71.1$ mA
Step 3 Obtain an estimate of the maximum diode current. $\overline{I}_D = \overline{I}_{21} - \underline{I}_{11} + N_R \underline{I}_{CC1}$.	3. $\overline{I}_D = 8 - 2 + (1.1)(30) = 39$ mA.
Step 4 Select suitable semiconductors.	4. Assuming this is a fairly high speed application, a 2N964A transistor and a 1N3605 diode are used. They can easily handle the currents required.
Step 5 Select a suitable minimum diode current and determine $\overline{V'}_1$. If the diode is to be kept in conduction at all times when the output is at the "1" level use $V'_1 = \overline{V}_D + \overline{V}_K$ If diode is allowed to cut off use $V'_1 = \underline{V}_1$	5. An I_D of 2 mA will provide a low impedance and is a negligible value in comparison to the load currents. At 2 mA, $\overline{V}_D$ is 0.86 volt as read from Figure 7-1-11. $V'_1 = 0.86 + 5.6 = 6.46$ V

Step 6

Calculate $\overline{R}_C$ from equation 7-1-1. Find R_C and $\underline{R}_C$.

$$\overline{R}_C = \frac{\underline{V}_{CC} - V'_1}{\underline{I}_{CC1}}$$

In cases where I_{CL} is not negligible it would have to be found at $\overline{T}_J$ which in turn is found from $\overline{T}_J = \overline{T}_A + \theta_J{}^*\overline{I}_C\overline{V}_O$

$$R_C = \frac{\overline{R}_C}{1 + n_R} \qquad \underline{R}_C = \frac{\overline{R}_C}{N_R}$$

6. $\overline{R}_C = \dfrac{17.1 - 6.46}{30} = 0.356\text{K}\Omega$

$$R_C = \frac{356}{1.05} = 339\Omega \qquad \underline{R}_C = \frac{356}{1.1} = 324\Omega$$

Step 7

Find $\overline{I}_C$. Figure 7-1-5 could be used, but as a check on previous calculations use equation 7-1-2

$$\overline{I}_C = \frac{\overline{V}_{CC} - \underline{V}_0}{\underline{R}_C} + \overline{I}_{20} - \underline{I}_{10}$$

7. $\overline{I}_C = \dfrac{18.9}{324} + 16 - 1 = 73.5$ mA

This value is higher than the estimate because V_1 was raised in order to keep the diode in conduction at all times. This resulted in an additional 2.4 mA of I_C or a 3% increase.

Step 8

Find $\overline{V}_1$ from,

$$\overline{V}_1 = \overline{V}_K + \overline{V}_D$$

Where $\overline{V}_D$ is determined by $\overline{I}_D$.

$$\overline{I}_D = \overline{I}_{21} - \underline{I}_{11} + \frac{\overline{V}_{CC} - \overline{V}_K - \overline{V}_D}{\underline{R}_C}$$

For this equation, V_D is found at the estimated current calculated in Step 3. Then, with a value for $\overline{I}_D$, an accurate value can be determined for $\overline{V}_D$ which can be used to solve for $\overline{V}_1$.

8. $\overline{V}_D$ is 1.09 V at 39 mA as found from figure 7-1-11.

$$\overline{I}_D = 8 - 2 + \frac{18.9 - 5.6 - 1.09}{0.324} = 44 \text{ mA}$$

at 44 mA, $\overline{V}_D = 1.10$ V

$\overline{V}_1 = 5.6 + 1.10 = 6.7$ V

*θ_J —Thermal resistance.

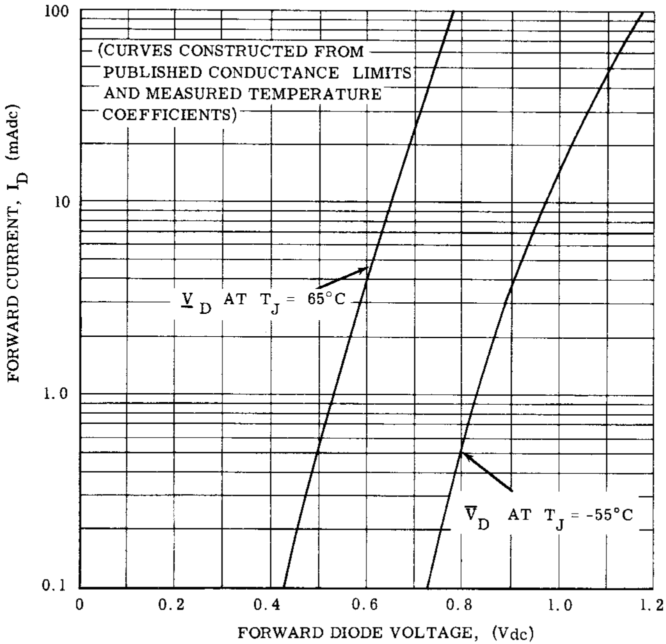


Figure 7-1-11 — Estimated Forward Conduction Characteristics for a 1N3605 Diode

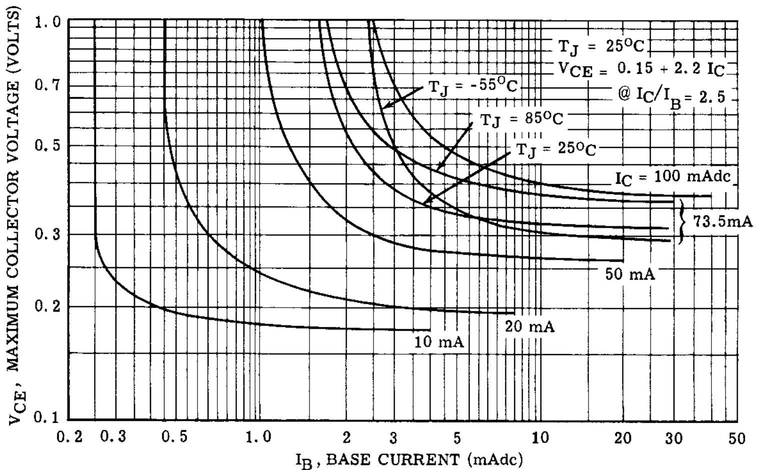


Figure 7-1-12 — DC Output Characteristics Showing Construction for RCTL Inverter Example

Step	Example
Step 9 Determine the suitability of the previously calculated R_C and V_1 by plotting the load line of R_C on a chart of Area of Permissable Load Loci. The load line must remain within the "latch-free" operating area.	9. The load line resulting from the calculated output circuit is within the permissable operating area for the 2N964A.
Step 10 Determine the minimum value of base current ($\underline{I}_B$) that will drive the transistor to the maximum allowable saturation voltage ($\bar{S}V_{CE}$) (system specification given as $\overline{V}_0$ under "worst-case" conditions from Table 7-1-3). This step involves manipulating the transistor data. Worst case output characteristics showing the saturation and knee region must be available. The curves in the given data are used only for reference purposes. To obtain the required $\underline{I}_B$, three additional curves (for 25°C, and the two specified temperature limits) must be constructed for the correct value of I_C as follows: *Curve 1* (for 25°C). Plot a point with the coordinates $V_{CE} = V_P + R_F I_C$ (volts) at $\beta_F = \beta_{FS}$ and $I_B = \frac{I_C}{\beta_{FS}}$ (mA). Plot a second point at $V_{CE} = V_{CE}$ specified for h_{FE} and $I_B = \frac{I_C}{h_{FE}}$ Connect the two points using the contours of the given curves as a guide.	10. To illustrate the use of the designers data for the 2N964A transistor, the steps necessary to obtain $\underline{I}_B$ will be done in detail. The output characteristic for the 2N964A is reproduced as Figure 7-1-12. The three curves at 73.5 mA are constructed using the procedure to be described, h_{FE} (β) data is taken from Figure 4-3 and θ_{VC} data is from Figure 4-4. For the 2N964A, these equations and the current for the example yields: *Curve 1* $V_{CE} = .15 + 2.2\ (.0735) = .3312V$ @ $\beta_{FS} = 2.5$ $I_B = \frac{73.5}{2.5} = 29.4$ mA. V_{CE} at which h_{FE} is specified is 1 volt $h_{FE} = 46$ @ 73.5 mA and 25°C $I_B = \frac{73.5}{46} = 1.6$ mA.

Curve 2 (for maximum junction temperature). Draw a curve at $\overline{T}_J = \overline{T}_A + \theta_J(\overline{I}_C\overline{S}V_{CE})$.	*Curve 2* θ_J is obtained from the data sheet as 0.5°C/mW. $T_J = 65° + .5\,(73.5)\,(0.5) = 83.8°C$ (Use 85°C for simplicity)
θ_J is the thermal resistance of the device as specified on the data sheet, and $\overline{S}V_{CE}$, the maximum allowable saturation voltage, may be taken equal to the circuit output voltage V_O originally specified.	$\theta_{VC} = 0.85$ mV/°C at 73.5 mA.
Plot point 1 using the coordinates $\overline{V}_{CE}(\text{at } \overline{T}_J) = \overline{V}_{CE}(\text{at } 25°) + \theta_{VC}\,(\overline{T}_J - 25°C)$	$\overline{V}_{CE} = 312 + 0.85(85\text{-}25) = 362$ mV
and $I_B = \dfrac{I_C}{\beta_{FS}}$.	$I_B = \dfrac{73.5}{2.5} = 29.4$ mA
Plot point 2 using the coordinates $V_{CE} = V_{CE}$ at which h_{FE} is specified	$V_{CE} = 1$ Volt
and $I_B = \dfrac{I_C}{h_{FE}\,(\text{at } \overline{T}_J)}$.	$h_{FE} = 43$ @ 73.5 mA and 85°C $I_B = \dfrac{73.5}{43} = 1.71$ mA @ $T_J = 85°C$
Connect the two points as for Curve 1.	
Curve 3 (for minimum junction temperature). Plot Point 1 using the coordinates	*Curve 3*
$\overline{V}_{CE}\,(\text{at } \underline{T}_A) = \overline{V}_{CE}\,(\text{at } 25°C) + \theta_{VC}\,(\underline{T}_A - 25°C)$	$\overline{V}_{CE} = 312 + 0.27(\text{-}55\text{-}25) = 290$ mV
and	

$I_B = \frac{I_C}{\beta_{FS}}$. Plot point 2 using the coordinates $V_{CE} = V_{CE}$ at which h_{FE} is specified and $I_B = \frac{I_C}{h_{FE}}$ (at($\underline{T}_A$). Connect the two points, as before. To determine the "worst case" condition locate the constructed curve that requires the larger I_B to maintain the desired SV_{CE}. Read I_B. This is the value of I_B used for calculating R_B and R_K in step 12.	$I_B = \frac{73.5}{2.5} = 29.4$ mA $V_{CE} = 1$ volt $h_{FE} = 30$ @ 73.5 mA and —55°C. $I_B = \frac{73.5}{30} = 2.45$ mA @ $\underline{T}_A = -55$°C A minimum value of 3 mA for I_B is required as seen on Figure 7-1-12.
Step 11 Determine the maximum base leakage current ($\bar{I}_{BL}$) and the minimum base emitter reverse bias voltage ($\underline{V}_{OB}$) from the intersection of the maximum junction temperature curve and the threshold voltage (V_{TR}) line.	11. From the transistor data (Figure 3-3) $\bar{I}_{BL} = 130\ \mu A$ and $\underline{V}_{OB} = 0.2V$.
Step 12 Calculate the maximum base-emitter voltage for the on condition at the minimum ambient temperature using: $\bar{V}_{BE}(\text{at } \underline{T}_A) = \bar{V}_{BE}(\text{at } 25°C) + \theta_{VB}(\underline{T}_A - 25°C)$ where V_{BE} (at 25°C) and θ_{VB} are obtained from transistor data.	12. From Figure 4-5, $\bar{V}_{BE} = 0.62$ From Figure 4-4, $\theta_{VB} = -1.5$ mV/°C $\bar{V}_{BE}$ (at $\underline{T}_A$) = 620 — 1.5 (—55 — 25) = 740 mV.

Step	Solution
Step 13 Using values determined in steps 9, 10, and 11 and originally specified system conditions, calculate resistor R_B and R_K by solving equations 7-1-20 and 7-1-21: $\frac{(\underline{V}_{OB} + \overline{V}_{O(In)})}{R_K(1-n)} - \frac{(\underline{V}_{BB} - \underline{V}_{OB})}{R_B(1+n)} + \overline{I}_{BL} = 0$ $\frac{(\underline{V}_1 - \overline{V}_{BE})}{R_K(1+n)} - \frac{(\overline{V}_{BB} + \overline{V}_{BE})}{R_B(1-n)} - \underline{I}_B = 0.$	13. $\frac{(0.2 + 1.0)}{R_K(0.95)} - \frac{(17.1 - 0.2)}{R_B(1.05)} + 0.130 \times 10^{-3} = 0$ $\frac{18.9 + 0.74}{R_B(0.95)} - \frac{(5 - 0.74)}{R_K(1.05)} - 3.0 \times 10^{-3} = 0$ $R_B = 9.08\ K \qquad R_K = 771\Omega$ The nearest standard values available in 1% tolerance resistors are 9.09KΩ and 768 ohms. For convenience in solving subsequent equations, limit values for R_B and R_K are determined below where end of life tolerances have been used. $\overline{R}_B = 9.09\,(1.05) = 9.54\ K$ $\underline{R}_B = 9.09\,(.95) = 8.64\ K$ $\overline{R}_K = 768\,(1.05) = 806\ \Omega$ $\underline{R}_K = 768\,(.95) = 730\ \Omega$
Step 14 It is convenient to have a checking equation to confirm the values obtained from R_B and R_K. From equation 7-1-20 $\underline{V}_{OB} = \frac{\underline{V}_{BB}\,\underline{R}_K - \overline{V}_O\,\overline{R}_B - \overline{I}_{BL}\,\underline{R}_K\,\overline{R}_B}{\overline{R}_K + \underline{R}_B}.$	14. $\underline{V}_{OB} = \frac{(17.1)\,(.730) - (1.0)\,(9.54) - (.13)\,(.73)\,(9.54)}{.73 + 9.54}$ $\underline{V}_{OB} = .196V$ $\underline{V}_{OB}$ was chosen to be 0.2V, therefore, solution checks.

Step 15 Calculate $\overline{V}_{OB}$ to be sure it does not exceed the emitter-base breakdown rating, by using the above equation with the opposite worst case conditions imposed. (I_{BL} can be assumed zero). $$\overline{V}_{OB} = \frac{\overline{V}_{BB}\,\overline{R}_K - \underline{V}_O\,\underline{R}_B}{\overline{R}_K + \underline{R}_B}.$$ If $\overline{V}_{OB}$ exceeds the rating of the transistor, it will be necessary to tighten the tolerances of the system or to select a different transistor.	15. $$\overline{V}_{OB} = \frac{(18.9)\,(.806) - 0}{.806 + 8.64} = 1.61\ \text{V}$$ This value is within the rating of the 2N964A.
Step 16 At $\overline{T}_J$, calculate $\underline{V}_{BE}$, $\overline{I}_B$, and then calculate $\overline{Q}_T$. These are the worst case conditions needed to determine C_K. (a) Calculate $\underline{V}_{BE}$ from equation 4-6 $$\underline{V}_{BE}\,(\text{at } \overline{T}_J) = \underline{V}_{BE}\,(\text{at } 25°\text{C}) + \theta_{VB}\,(\overline{T}_J - 25°\text{C})$$ where $\underline{V}_{BE}$ (at 25°C) is the minimum specified value obtained from the data sheet at $\overline{I}_C$. (b) Calculate $\overline{I}_B$ by altering equation 7-1-21. $$\overline{I}_B = \frac{\underline{V}_{1\,(in)} - \underline{V}_{BE}\,(\text{at } \overline{T}_J)}{(1-n)\,R_K} - \frac{V_{BB} + \underline{V}_{BE}\,(\text{at } \overline{T}_J)}{(1+n)\,R_B}$$ Worst case occurs at $\underline{V}_{1(in)}$ because the charge stored on the capacitor decreases faster with a drop in V_1 than does the charge stored in the transistor.	16. (a) $$\underline{V}_{BE}\,(85°\text{C}) = 0.4 - (.0015)\,(85\text{-}25) = .31\ \text{V}$$ (b) $$\overline{I}_B = \frac{5 - .31}{.730} - \frac{17.1 + .31}{9.54} = 4.6\ \text{mA}$$

(c) Calculate $\overline{Q}_T$ from equation 5-44. $\overline{Q}_{T2} = \overline{Q}_{T1} + \overline{\tau}_{A2}\, \overline{I}_{C2} - \overline{\tau}_{A1}\, I_{C1}$ where Q_{T1} is found at $\overline{I}_B$ τ_{A1} is found at I_{C1} where I_{C1} is the value used for the Q_T curve τ_{A2} is found at $\overline{I}_C = I_{C2}$.	(c) Using the following conditions: $I_B = 4.6$ mA, $T_J = 85°C$ From the data sheet, $\overline{Q}_{T1} = 390$ pC $\tau_{A1} = 1.4$ nS ($I_C = 46$ mA) $\tau_{A2} = 1.6$ nS ($I_C = 73.5$ mA) $\therefore \overline{Q}_{T2} = 390 + (1.6)(73.5) - (1.4)(46) = 444$ pC
Step 17 Calculate the value of the speed-up capacitor (C_K) $\underline{C}_K = \dfrac{\overline{Q}_{T2}}{(\underline{V}_1 - \overline{V}_O)_{(\ln)}}$.	17. $\underline{C}_K = \dfrac{444}{4} = 111$ pF A 120 pF silver mica capacitor will meet this requirement.

SECTION 2 — FLIP-FLOP DESIGN

The basic flip-flop, or bistable multivibrator circuit is shown in Figure 7-2-1. It has two stables states — Q_1 on and Q_2 off, or Q_1 off and Q_2 on. Application of a trigger causes the flip-flop to change state. For every two triggers the flip-flop will be in the same state; thus it counts by two's and is often called a binary circuit.

The flip-flop is used in electronic counters as the basic counter unit and in timing circuits as a frequency divider. The widest usage is found in digital computers where it functions as a storage unit in performing logic. In this application a system of gates is used at the trigger so that the flip-flop changes state for only certain combinations of inputs. In computers, it is generally used with clamp diodes or output amplifiers, but the circuit of Figure 7-2-1 may be considered the heart of any flip-flop.

To produce a reliable and useful flip-flop, the circuit must fulfill the following conditions under some worst-case combination of component and power supply tolerances:

1. The flip-flop must remain in a stable state unless triggered. To insure this condition, the coupling resistor (R_K) and the bias resistor (R_B) must be so chosen that when one side is on (collector voltage at SV_{CE}) the other side is biased off and simultaneously the output voltage at the off side allows sufficient base current to drive the on side into the saturation region.
2. The output voltage level must remain within fixed limits while allowing for a maximum load current to flow.
3. The flip-flop must change state within a prescribed time after application of a trigger pulse.

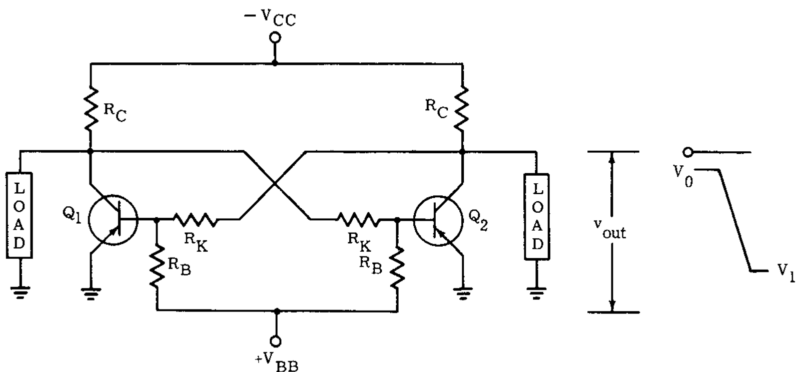


Figure 7-2-1 — Basic Flip-Flop Circuit

7-2-1 — Synthesis Equations

Since the flip-flop can be considered as two cross-coupled inverters, the resulting design equations are very similar. The chief difference is that the current through R_C when the output is at V_1, must not only supply the load current I_1 but also the feedback current I_K, which holds the opposite side on.

For simplicity, let $V_{BE} = 0$, since this is the worst case, and by inspection of Figure 7-2-2a write

$$\overline{I}_1 = \frac{\underline{V}_{CC} - \underline{V}_1}{\overline{R}_C} - \overline{I}_{CL} - \frac{\underline{V}_1}{\underline{R}_K}. \tag{7-2-1}$$

As with the inverter, the input network is described by the equations

$$\frac{\underline{V}_{OB} + \overline{SV}_{CE}}{\underline{R}_K} - \frac{\underline{V}_{BB} - \underline{V}_{OB}}{\overline{R}_B} + \overline{I}_{BL} = 0 \tag{7-2-2}$$

and

$$\frac{\underline{V}_1 - \overline{V}_{BE}}{\overline{R}_K} - \frac{\overline{V}_{BB} + \overline{V}_{BE}}{\underline{R}_B} - \underline{I}_B = 0 \tag{7-2-3}$$

which may be written by inspection of Figure 7-2-2 a & b. These three equations can be solved simultaneously by using matrix techniques for the unknown resistors. A problem arises because the collector current is unknown; and it determines I_B and, to a large extent, values for the other transistor characteristics in the equations. Therefore, a preliminary method of estimating I_C is necessary before a circuit can be designed.

Of course, the three equations could be solved by guessing values for the transistor characteristics. Then I_C could be solved for, which would permit accurate transistor values to be found. Then, using these transistor values the matrix could be resolved. This would work well since the saturation voltages normally do not affect circuit currents significantly. However, the currents in the equations, particularly I_B, normally have a large effect upon the required resistor values. This problem can be overcome by using the ratio of collector to base current (β_F) in equation 7-2-3. That is

$$\underline{I}_B = \frac{\overline{I}_C}{\underline{\beta}_F} = \frac{\overline{V}_{CC}}{\underline{R}_C\underline{\beta}_F} + \frac{\overline{I}_o}{\underline{\beta}_F} \text{ (neglecting } SV_{CE}) \tag{7-2-4}$$

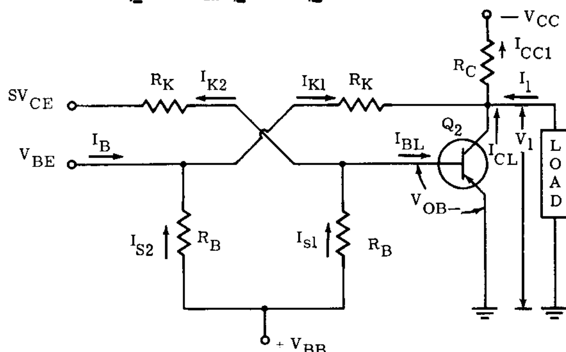


Figure 7-2-2a — Conditions with Q_2 Off

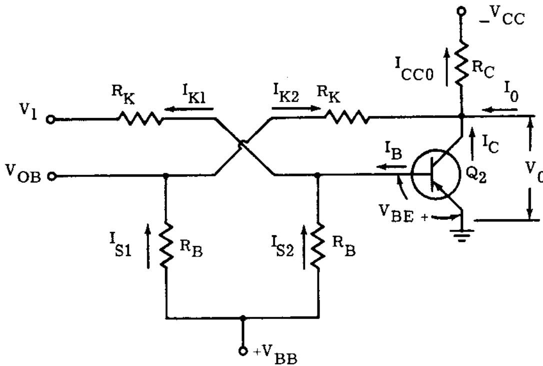


Figure 7-2-2b — Conditions with Q_2 On

where I_o = maximum load current with the transistor on.

As long as β_F is held constant, errors in the estimated collector current will not significantly affect the values for SV_{CE} and V_{BE}.

The values of $\overline{I}_{BL}$ and $\underline{V}_{OB}$ used in equation 7-2-2 depend upon the junction temperature, which is affected by the power dissipation. Normally, in the saturated mode flip-flop, the dissipation is low and the junction temperature is only slightly above the ambient. Thus $\overline{I}_{BL}$ and $\underline{V}_{OB}$ also can be estimated even though T_J is not known exactly.

A simpler method of finding $\overline{I}_C$ consists of examining the output network and making an estimate of the coupling efficiency as follows.

For the inverter, the output circuit was analyzed and the following equation developed

$$\gamma = \frac{\overline{I}_C - \overline{I}_O}{\underline{I}_{CC1}} = N_{RC} \cdot \frac{1 + n_P}{1 - n_P - V'_1/V_{CC}} \qquad (7\text{-}2\text{-}5)$$

where

$\underline{I}_{CC1}$ = the minimum current through the collector resistor
$\overline{I}_C$ = the maximum collector current
$\overline{I}_0$ = the maximum load current with the transistor on
$\overline{I}_1$ = the maximum load current with the transistor off
N_{RC} = collector resistor max./min. ratio
n_P = power supply tolerance
V'_1 = minimum output voltage
V_{CC} = nominal collector supply voltage.

A graph of equation 7-2-5 is shown in Figure 7-1-5.

To start a design, I_C must be estimated. $I_{CC1} \approx I_1 + I_K$ but I_K is unknown. An estimate for I_K can be found as follows: Experience indicates that I_S is almost always less than I_B; being pessimistic assume they are equal. Thus,

$$\overline{I}_{K1} = \overline{I}_B + \overline{I}_S \approx 2\overline{I}_B = \frac{2\overline{I}_C}{\underline{\beta}_F}.$$

The current $\overline{I}_{K1}$ effectively adds to $\overline{I}_1$. Substituting $\underline{I}_{CC1} = \overline{I}_1 + \frac{2\overline{I}_C}{\overline{\beta}_F}$ into equation 7-2-5 find

$$\gamma = \frac{\overline{I}'_C + \overline{I}_o}{\overline{I}_1 + \frac{2\overline{I}_C}{\overline{\beta}_F}}$$

Where $\overline{I}'_C$ indicates an estimated $\overline{I}_C$ including the effect of $\overline{I}_{K1}$. Simplifying:

$$\overline{I}'_C = \frac{\gamma \overline{I}_1 + \overline{I}_o}{1 - \frac{2\gamma}{\overline{\beta}_F}}. \qquad (7\text{-}2\text{-}6)$$

This expression can be used with Figure 7-1-5 to estimate I_C and thereby obtain transistor values for use in the matrix.

After the matrix has been solved, other worst-case limits must be found to complete the design.

To check for latch-up problems $\overline{V}_1$ must be determined. The voltage (V_1) will be maximum when all currents are minimum and the supply voltage is maximum. ($I_{CL} \approx 0$ at low temperatures). Rewriting equation 7-2-1

$$\overline{V}_1 = \frac{\overline{V}_{CC}\,\overline{R}_K + \overline{V}_{BE}\,\overline{R}_L - \underline{I}_1\,\underline{R}_C\,\overline{R}_K}{\overline{R}_K + \underline{R}_C}. \qquad (7\text{-}2\text{-}7)$$

In order to find maximum stored charge and storage time, $\overline{I}_B$ must be determined. This can be done by rewriting equation 7-2-3

$$\overline{I}_B = \frac{\overline{V}_1 - \underline{V}_{BE}}{\underline{R}_K} - \frac{\underline{V}_{BB} - \underline{V}_{BE}}{\overline{R}_B} \qquad (7\text{-}2\text{-}8)$$

To check for the maximum reverse bias on the emitter junction, V_{OB} must be found. Equation 7-2-2 can be rewritten for this purpose. In most cases I_{BL} and SV_{CE} are negligible and the expression reduces to

$$\overline{V}_{OB} = \frac{\overline{V}_{BB}\,\overline{R}_K}{\overline{R}_K + \underline{R}_B} \qquad (7\text{-}2\text{-}9)$$

This completes the dc analysis and development of the necessary synthesis equations.

7-2-2 — Triggering

A good deal of the problems associated with flip-flop operation can be traced to difficulties with triggering. The trigger signal should inject just enough charge to change state of one of the transistors. Also, the trigger circuit should be self gated; that is, as the flip-flop changes state, it should remove the trigger signal so that there would be no tendency for the trigger to cause the flip-flop to revert to its original state.

One of the most satisfactory methods of triggering is shown in Figure 7-2-3. The circuit operates as follows: When the leading edge of the trigger pulse arrives, the diode D_C connected to the on side provides a low impedance path to charge C_T to the pulse amplitude, (—5V in this circuit) less the diode drop. When the trigger returns to ground, the anode of D_T is raised to + 5V; since its cathode is near ground potential it conducts heavily discharging C_T

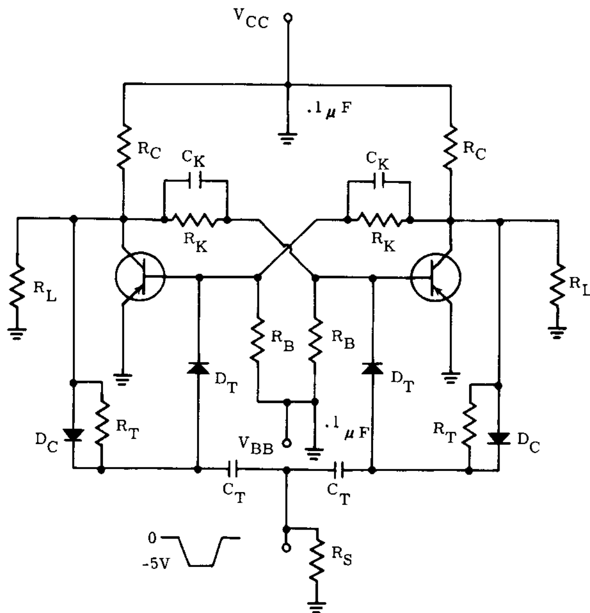


Figure 7-2-3 — Flip-Flop Showing Trigger Circuit

into the base of the on transistor. If the charge on C_T is greater than the total control charge Q_T of the transistor and R_S is small, then turn-off will be very rapid. Note that the flip-flop is actuated on the trailing edge of the input pulse.

There is a compromise involved in selecting the value of R_T. Notice that when a side is being turned off, R_T appears in series with D_T from base to collector causing an undesirable shunt. Thus, as the voltage on the collector starts to rise, a negative feedback current flows through R_T to the base and slows the fall time. This problem is alleviated if R_T is high.

Note, however, that D_T will continue to conduct until the voltage across it has reached zero. Recovery time problems arise because normally not all the charge on C_T is used in turning off the transistor. Suppose, for example, that 3 volts were left on C_T. The transistor is now cut off, so the circuit impedance is fairly high resulting in a long discharge time for C_T. However, by making R_T small, the remaining 3 volts of trigger charge can be completely removed and then C_T can be charged to the collector voltage, V_1 at a faster rate. It is not necessary that the voltage at the junction of C_T and D_T be at V_1 when the next trigger pulse appears, but it should be at a voltage which cuts off the diode D_T. When the leading edge of the trigger pulse arrives, it will bring this junction point voltage up to V_1, thereby back biasing D_T. If sufficient reverse bias is not present on D_T it will conduct on the trailing edge and inject a positive voltage into the off side which opposes the regenerative action. Therefore, R_T should be low to reduce this problem.

Usually, making R_T about 2 to 4 times the value of R_K is a suitable compromise, however, it is better to experimentally optimize its value.

Note that C_T must charge through the collector of an on transistor. The required current can be appreciable and may cause a transistor to come out of

saturation prior to the trailing edge of the trigger. Therefore, this current should be computed and added to the value of $\bar{I}_C$ used in the design.

The value of the trigger capacitor can be found from the general equation:

$$\underline{C}_T = \frac{\bar{Q}_T}{\Delta \underline{V}_T}. \tag{7-2-10}$$

$\Delta \underline{V}_T$ is given by:

$$\Delta \underline{V}_T = (\underline{V}_T - \bar{S}V_{CE} - \bar{V}_{DC} - \bar{V}_{DT})$$

where: V_T is the trigger amplitude
SV_{CE} is collector saturation voltage
V_{DC} drop of diode D_C
V_{DT} drop of diode D_T.

Since Q_T and SV_{CE} increase with temperature and the diode drops decrease, values at minimum and maximum temperature must be found in order to determine which is the worst case value.

7-2-3 — Cross Coupling

Selecting the cross-coupling capacitor also involves a compromise. Since the turn-off of collector current is rapid due to the injection of Q_T from a low impedance source, the decay time of the collector voltage will be due to the time constant C_K forms with R_K, R_C, and R_L in parallel, and could be long if C_K is large. However, the turn-on delay and rise time of the other side will be improved if C_K is large. Usually, a value of C_K selected so that its change in charge is about 2 to 5 times $\tau_A I_C + Q_{OB}$ is a satisfactory estimate, and the circuit may be experimentally optimized.

Table 7-2-1 shows typical measured switching times, for the design example, without C_K and with $C_K = 20$ pF (the optimum value). The table also illustrates the effect of making R_T too small.

TABLE 7-2-1

Typical Flip Flop Performance (5% tolerance circuit)					
	t_d (ns)	t_r (ns)	t_s (ns)	t_f (ns)	t_{total}
$C_K = 0$, $R_T = 20K$	49	36	6	33	124
$C_K = 20$ pF, $R_T = 20K$	14	15	6	62	97
$C_K = 0$, $R_T = 5K$	47	44	6	78	175
$C_K = 20$ pF, $R_T = 5K$	14	22	6	150	192

7-2-4 — Design Variations

As Table 7-2-1 clearly shows, the longest switching interval is fall time if C_K is used and delay and rise time if it is omitted. There are variations of the design procedure which can be used to shorten these times. However, they all come at the expense of more trigger energy. In digital computers this is not much of a problem since more clock power can usually be easily provided. In a counter system, however, this could mean employing trigger amplifiers.

Fall time can be improved by using clamp diodes, and selecting R_L and R_C so that the value, which V_1 would assume in the absence of the diodes, is greater than 1.5 times the clamp level. The "catching" action of the clamps will linearize the fall time and clip off the slowly changing portion of the exponential. This procedure was fully discussed in the inverter section. The results of using diodes are shown in Table 7-2-2, where the clamp level was adjusted so that $V_1 = 5V$. As R_L is increased, V_1 tries to find a higher level and the catching action of the diodes is improved. All other component values are unchanged.

Delay and rise time can also be improved by selecting β_F to be much lower than that based upon the dc conditions, in which case the cross-coupling capacitors can be omitted. Delay time may be improved by using diodes from the bases to ground to prevent V_{OB} from exceeding few tenths of a volt. Fast germanium diodes should be used for this application. The diodes also hasten recovery time since the impedance at the base is kept low at all times, thereby speeding recovery of C_K and C_T.

TABLE 7-2-2

Improvements Due to Use of Clamp Diodes

R_L	t_d(ns)	t_r (ns)	t_s(ns)	t_f(ns)	t_{total}
1.68K *	14	15	6	62	97
5.0K	14	8	6	37	65
∞	14	5	6	17	42

*No Clamp Diodes

7-2-5 — Flip-Flop Synthesis Procedure

There are basically two different approaches to synthesize a flip-flop.

1. The flip-flop can be designed by simply using two inverters connected back-to-back. It is presumed that the inverter was designed and optimized according to some criteria; that is the transistor is operating near its maximum gain point or point of maximum speed, etc. Then, the flip-flop would also be of similar optimum design. This is often done in digital systems and the procedure used for inverter synthesis would be used.
2. The flip-flop may be required to drive a given dc load. Unless there is some restriction on output impedance, then the flip-flop can be designed solely on the basis of supplying a given load current at a given output voltage. Using this criteria, I_C may be kept to a minimum. Making I_C as low as possible will lower the required energy necessary to trigger the flip-flop. This is the design procedure used in the example.

As discussed with the inverter, in any dc system, it is necessary to establish limits for the two signal levels. The level near ground is called the zero ("0") level, designated V_0, and the high level is called the one ("1") level, designated V_1. The design procedure described insures that the flip-flop will deliver a given current to the load, while maintaining the "1" level above a required minimum, and that the "0" level will be below a specified maximum. The design is based upon just meeting these requirements under worst-case conditions, thereby avoiding the inefficiency, expense, and poor reliability of overdesign.

PRELIMINARY DESIGN CONSIDERATIONS: The output levels must be chosen with the transistors to be used in mind or vice versa. The transistor must be able to provide the saturation voltage, SV_{CE}, required for V_0 with a reasonable gain and, of course, the breakdown voltage rating must exceed the maximum limit for V_1. In the discussion to follow, a bar over a term indicates a maximum and a bar under a term indicates a minimum condition. The symbol n is used to indicate tolerance, assumed $\pm$ by the same amount.

The advantages of keeping the ratio of V'_1/V_{CC} low were discussed in the section on the inverter. However, in the flip-flop it should also be clear that the V'_1/V_{CC} ratio imposes a lower limit upon the transistor current gain. Consider a case where no dc load is present, i.e., $I_1 = I_K$. Further assume that the efficiency of the R_B — R_K network is 50%, i.e., the current I_S resulting from establishing the off condition is equal to I_B, the on base current. Then the ratio of $\frac{\overline{I}_C}{\overline{I}_1} = \frac{\beta}{2}$.* Thus, by referring to Figure 7-1-5, it is seen that if the V'_1/V_{CC} ratio is near unity, very high gain transistors become necessary. In the usual case, where a dc load current I_1 is required in addition to the feedback current I_K, the current gain requirement increases greatly in comparison to the simple case just discussed.

The advantages of a low V'_1/V_{CC} ratio are not without price because this makes the V_{CC} — R_C source approach a current source. Therefore, variations in the load R_L will result in changes in the output voltage. A high output voltage results in excessive overdrive into the on side of the flip-flop and any transistor stages driven from the output which causes longer storage time. However, the solution to this problem is not to make R_C low; it is much better to achieve output voltage stabilization by using clamp diodes at the output as shown in Figure 7-2-4 if the change in output voltage becomes troublesome. The design procedure with clamp diodes is essentially the same as without them, since the procedure is based upon providing a minimum I_1 to a load when V_1 is minimum. With clamp diodes, the current I_1 effectively must be increased by a small amount in order to have current available to keep the clamps always in conduction. Further considerations when using clamps were discussed with the inverter.

The clamp power supply voltage is determined by

$$V_K = (V'_1 + \overline{V}_D) / (1 - n_P) \qquad (7\text{-}2\text{-}11)$$

and the maximum output voltage is given by:

$$\overline{V}_1 = (1 + n_P)\, V_K - \underline{V}_D \qquad (7\text{-}2\text{-}12)$$

$* \; I_1 = I_K = 2I_B = \frac{2I_C}{\beta}$

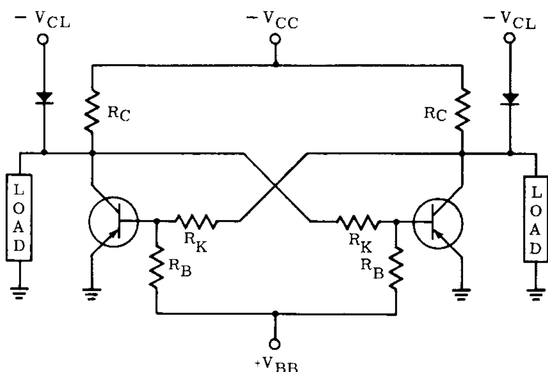


Figure 7-2-4 — Basic Flip-Flop with Clamps

The diode voltage drops ($\underline{V}_D$ and $\overline{V}_D$) are obtained from the diode characteristic curves at conditions of $\underline{I}_D$ at the high temperature limit, and $\overline{I}_D$ at the low temperature limit respectively.

If clamps are not used, it is not possible to give an accurate value for $\overline{V}_1$ for a variable load until the design is complete. If the minimum load current were zero, $\overline{V}_1$ would approach V_{CC}.

DESIGN PROCEDURE: The first step in design is to select a suitable $\underline{V}'_1/V_{CC}$ ratio based upon $\overline{I}_1$ and the supply tolerances. Then, neglecting I_K, $\overline{I}_C$ can be estimated. This estimated value is called $\overline{I}'_C$; with it known, a transistor with suitable characteristics can be selected. Then, by obtaining a value for $\underline{\beta}_F$, the forced gain or circuit gain, a closer estimate of $\overline{I}_C$ is found by estimating the feedback current I_K. With this information, the significant transistor data can be found, to use in a matrix solution, to obtain values for R_C, R_K and R_B. In this solution I_C is treated as an unknown and β_F is used, rather than a value for I_B and I_C. This approach works well because transistor gain is a slowly varying function of emitter current as are SV_{CE} and V_{BE} when β_F is constant. Once the resistor values are known, exact values of I_C and I_B can be calculated and limits of SV_{CE} and V_{BE} found. Usually, the exact collector current is close to the estimated current and the values of the transistor characteristics used in the matrix solution are satisfactory. If a closer solution is desired, the process can be repeated using new values for the transistor characteristics.

Finally, other worst-case limits can be calculated such as maximum collector voltage, $\overline{V}_1$, maximum drive current, $\overline{I}_B$, and maximum off bias, $\overline{V}_{OB}$. With these quantities known, operation can be checked for latch-up and the stored charges calculated so that a transient solution can be obtained.

The following step-by-step procedure has been prepared for ease in designing flip-flop circuits. Two solutions using different tolerance resistors are developed in detail to illustrate the tremendous importance of this factor.

TABLE 7-2-3

Required Output Levels (Volts)				
	Min.		**Max.**	
V_0	0		−0.3	
V_1	−5.0			
Load Current Requirements (mA)				
	Min.		**Max.**	
At Output Level V_0	0		0	
At Output Level V_1	0		3	
Available Power Supply Voltages				
Minimum	+ 5.94	−5.94	+ 11.88	−11.88
Nominal	+ 6	−6	+ 12	−12
Maximum	+ 6.06	−6.06	+ 12.12	−12.12
Ambient Temperature Range				
−55 TO + 85°C				
Resistor Tolerances*				
Two Solutions	1)	± 5%	(± 1% Initial), ∴ $N_R = 1.105$	
	2)	± 20%	(± 5% Initial), ∴ $N_R = 1.5$	

*See Inverter Section for a discussion of tolerances

7-2-6 — Design Example

General Procedure	Example
Step 1 List specifications.	1. Specifications are shown in Table 7-2-3.
Step 2 Select V'_1/V_{CC} and determine I'_C by assuming a value for β_F $\bar{I}'_C = \dfrac{\gamma \bar{I}_1 + \bar{I}_O}{1 - \dfrac{2\gamma}{\bar{\beta}_F}}$.	2. Use a 12 V supply to keep I_C low, $V'_1/V_{CC} = \dfrac{5}{12} = .415.$ from Figure 7-1-5 $\dfrac{\gamma}{N_R} = 1.80.$ With $n_R = .05$, $\gamma = 1.92$ — With $n_R = .20$, $\gamma = 2.63$ Assume a β_F of 20 $I'_{C1} = \dfrac{(1.92)(3)}{1 - \dfrac{2(1.92)}{20}} = 7.1$ mA — $I'_{C2} = \dfrac{(2.63)(3)}{1 - \dfrac{2(2.63)}{20}} = 10.7$ mA (Since $\bar{V}_1$ is not specified a clamp diode is not needed)
Step 3 Select a suitable transistor & find $\bar{\beta}_F$ @ $\bar{I}'_C$ to produce $\bar{S}V_{CE} \leq \bar{V}_O$ with good gain.	3. A 2N964A type is suitable. At the worst-case temperature of $-55°C$, SV_{CE} will be below 0.3 volt @ $\bar{\beta}_{F1} = 17$ & $\beta_{F2} = 18$
Step 4 The $\bar{I}'_C$ just computed must be corrected using the new values of β_F Use $\bar{I}_C = \dfrac{\gamma \bar{I}_1 + \bar{I}_O}{1 - \dfrac{2\gamma}{\bar{\beta}_F}}$ (7-2-6)	4. $\bar{I}_{C1} = \dfrac{(1.92)(3) + 0}{1 - \dfrac{2(1.92)}{17}} = 7.45$ mA ($n_R = 5\%$) $\bar{I}_{C2} = \dfrac{(2.63)(3) + 0}{1 - \dfrac{2(2.63)}{18}} = 11.2$ mA ($n_R = 20\%$)

Step 5 Repeat Step 4 if β has changed significantly until $\bar{I}_C$ and β match transistor specifications. Whenever $\bar{I}_1 \gg \frac{\bar{I}_C}{\beta_F}$, one trial is usually enough.	5. The second estimate of I_C is close enough in both cases so that the β_F chosen in Step 3 can be used.
Step 6 Obtain the worst-case values for the transistor being used. The maximum temperature must include the effects of power dissipation. For this purpose use $\bar{T}_J = \bar{T}_A + \theta_{JA}\bar{I}_C\bar{V}_O$ where θ_{JA} is the thermal resistance. The lower limit must not include the temperature rise due to power dissipation. $\therefore \underline{T}_J = \underline{T}_A$	6. Limit curves are provided on the 2N964A data sheet. Using techniques previously described for the inverter, the limit values are easily found. Since the currents in these designs are 7.4 & 11.2 mA, for simplicity, construct min. & max. deviations of the $I_C = 10$ mA curve which is close enough to the design current and is plotted in the data sheet. $\bar{T}_J \approx 85 + .5\,(10)\,(.3) = 86.5°C$ For simplicity use 85°C. $\underline{T}_J = -55°C$ These limits are. $\bar{I}_{BL} = 140\mu A$ & $\underline{V}_{OB} = 0.2V = \bar{V}_{TR}$ $\bar{S}V_{CE} = .19$ @ $\beta_F = 18$ & $T_J = 85°C$ $\bar{S}V_{CE} = .23$ @ $\beta_F = 18$ & $T_J = -55°C$ $\bar{V}_{BE} = .586$ @ $\beta_F = 10$ & $T_J = -55°C$ (@ $\beta_F = 18$, $\bar{V}_{BE}$ will only be a few mV lower) The value of $\bar{S}V_{CE}$ at 85°C will be used in the design since it is coupled through R_K to the off side and it is at the high temperature limit where I_{BL} & V_{TR} are maximum.
Step 7 Now a matrix solution of the network equations developed previously as equations 7-2-1, 7-2-2, and 7-2-3 can be performed. For convenience in solving they are written as follows: $K_{11}\,G_C + K_{12}\,G_K + K_{13}\,G_B = A_1$ $K_{21}\,G_C + K_{22}\,G_K + K_{23}\,G_B = A_2$ $K_{31}\,G_C + K_{32}\,G_K + K_{33}\,G_P = A_3$ Where the terms are defined as $G_C = \frac{1}{R_C}$, $G_K = \frac{1}{R_K}$, $G_B = \frac{1}{R_B}$,	

7. Procedure	Solution I ($n_R = .05$)	Solution II ($n_R = .2$)
$K_{11} = \frac{V_{CC}(1\text{-}n_P) - \underline{V}_1}{1 + n_{RC}}$	$K_{11} = \frac{12(1-.01) - 5}{1 + .05} = 6.56$	$K_{11} = \frac{12(1-.01) - 5}{1 + .2} = 5.75$
$K_{12} = \frac{-\underline{V}_1}{1 - n_{RK}}$	$K_{12} = \frac{-5}{1 - .05} = -5.26$	$K_{12} = \frac{-5}{1 - .2} = -6.25$
$K_{13} = 0$	$K_{13} = 0$	$K_{13} = 0$
$K_{21} = 0$	$K_{21} = 0$	$K_{12} = 0$
$K_{22} = -\frac{(\bar{S}V_{CE} + \underline{V}_{OB})}{(1 - n_{RK})}$	$K_{22} = -\frac{(.19 + .2)}{1 - .05} = -.41$	$K_{22} = -\frac{(.19 + .2)}{1 - .2} = -.488$
$K_{23} = \frac{V_{BB}(1 - n_P) - \underline{V}_{OB}}{(1 + n_{RB})}$	$K_{23} = \frac{12(1-.01) - .2}{(1 + .05)} = 11.13$	$K_{23} = \frac{12(1-.01) - .2}{1 - .2} = 9.75$
$K_{31} = \frac{V_{CC}(1 + n_P)}{\beta_F(1 - n_{RC})}$	$K_{31} = \frac{12(1 + .01)}{17(1 - .05)} = .797$	$K_{31} = \frac{12(1 + .01)}{18(1 - .2)} = .841$
$K_{32} = -\frac{(\underline{V}_1 - \bar{V}_{BE})}{(1 + n_{RK})}$	$K_{32} = \frac{-(5 - .586)}{(1 + .05)} = -4.2$	$K_{32} = \frac{-(5 - .586)}{1 + .2} = -4.2$
$K_{33} = \frac{\bar{V}_{BE} + V_{BB}(1 + n_P)}{(1 - n_{RB})}$	$K_{33} = \frac{.586 + 12(1 + .01)}{1 - .05} = 13.4$	$K_{33} = \frac{.586 + 12(1 + .01)}{1 - .2} = 15.9$

$A_1 = \bar{I}_1 + \bar{I}_{CL}$	$A_1 = 3x10^{-3} + .14x10^{-3} = 3.14x10^{-3}$	$A_1 = 3x10^{-3} + .14x10^{-3} = 3.14x10^{-3}$
$A_2 = \bar{I}_{BL}$	$A_2 = .14 \text{ x } 10^{-3}$	$A_2 = .14 \text{ x } 10^{-3}$
$A_3 = 0$	$A_3 = 0$	$A_3 = 0$
	Substituting the above values and solving	
	$R_C = 1.6 \text{ K}\Omega$	$R_C = 1.09 \text{ K}\Omega$
	$R_K = 5.6 \text{ K}\Omega$	$R_K = 2.94 \text{ K}\Omega$
	$R_B = 52 \text{ K}\Omega$	$R_B = 31.6 \text{ K}\Omega$

Step 8

Check matrix solution by solving for

$$\underline{V}_{OB} = \frac{\underline{V}_{BB}\,\underline{R}_K - \bar{S}V_{CE}\,\bar{R}_B - \bar{I}_{BL}\,\bar{R}_B\,\underline{R}_K}{\underline{R}_K + \bar{R}_B}$$

8.

$n_R = 5\%$

$$\underline{V}_{OB} = \frac{12\,(.99)\,(5.31K) - (.19)\,(54.6K) - (.14x10^{-3})\,(54.6K)\,(5.31K)}{5.31K + 54.6K}$$

$\underline{V}_{OB} = 0.202$ Volt

$n_R = 20\%$

$$\underline{V}_{OB} = \frac{12\,(.99)\,(2.36K) - (.19)\,(37.8K) - (.14x10^{-3})\,(37.8K)\,(2.36K)}{2.36K + 37.8K}$$

$\underline{V}_{OB} = 0.207$ Volt

These values check sufficiently close to the desired V_{OB} so that the matrix computation is correct;

Step 9 Select nearest standard values Using given tolerances, solve for maximum and minimum resistance values.	9. Standard 1% resistors are: $R_C = 1.58\ K\Omega$, $R_K = 5.62\ K\Omega$, $R_B = 52.3\ K\Omega$ Standard 5% resistors are: $R_C = 1.1\ K\Omega$, $R_K = 3.0\ K\Omega$, $R_B = 33\ K\Omega$ Therefore: 1%: $\overline{R}_C = 1.58\,(1.04) = 1.64\ K\Omega$ $\underline{R}_C = 1.58\,(.96) = 1.52\ K\Omega$ $\overline{R}_K = 5.62\,(1.04) = 5.85\ K\Omega$ $\underline{R}_K = 5.62\,(.96) = 5.4\ K\Omega$ $\overline{R}_B = 52.3\,(1.04) = 54.4\ K\Omega$ $\underline{R}_B = 52.3\,(.96) = 50.2\ K\Omega$ 5%: $\overline{R}_C = 1.1\,(1.15) = 1.26\ K\Omega$ $\underline{R}_C = 1.1\,(.85) = 0.935\ K\Omega$ $\overline{R}_K = 3\,(1.15) = 3.45\ K\Omega$ $\underline{R}_K = 3\,(.85) = 2.55\ K\Omega$ $\overline{R}_B = 33\,(1.15) = 38\ K\Omega$ $\underline{R}_B = 33\,(.85) = 28\ K\Omega$
Step 10 Calculate actual $\overline{I}_C$ and compare to estimated $\overline{I}_C$ $\overline{I}_C = \dfrac{\overline{V}_{CC} - \overline{V}_O}{\underline{R}_C}$ If assumed values for SV_{CE}, V_{BE} and β_F are seriously in error, then new values should be found and the matrix solution (Step 7) repeated.	10. at $n_R = 5\%$ $\overline{I}_C$ (estimated) $= 7.4$ mA the actual value is $\overline{I}_C = \dfrac{12\,(1.01) - 0.19}{1.52\ K} = 7.88$ mA at $n_R = 20\%$ $\overline{I}_C$ (estimated) $= 11.2$ mA the actual value is $\overline{I}_C = \dfrac{12\,(1.01) - 0.19}{0.935\ K} = 12.7$ mA (A $\overline{V}_O$ of 0.19 volt was assumed when computing the cross-coupling network; therefore, this value must be used here)

Step	Solution
Step 11 $\overline{V}_1$ must be found to determine if latch-up problems will be encountered. Use $$\overline{V}_1 = \frac{\overline{V}_{CC}\,\overline{R}_K + \overline{V}_{BE}\,\underline{R}_C - \underline{I}_1\,\underline{R}_C\,\overline{R}_K}{\overline{R}_K + \underline{R}_C}$$	11. $n_R = 5\%$ $$\overline{V}_1 = \frac{12\,(1.01)\,(5.85\text{ K}) + (.5676)\,(1.52\text{ K})}{5.85\text{ K} + 1.52\text{ K}}$$ $\overline{V}_1 = 9.75$ Volts $n_R = 20\%$ $$\overline{V}_1 = \frac{12\,(1.01)\,(3.45\text{ K}) + (.606)\,(0.87\text{ K})}{3.45\text{ K} + 0.935}$$ $\overline{V}_1 = 9.67$ Volts
Step 12 Check the resulting load line to make sure it is free from latch-up.	12. Both of these voltage values and respective I_C were checked for possible latch-up. Both were found to be within the given latch-free operating area on the data sheet.
Step 13 To calculate the stored charge $\overline{I}_B$ must be known. Therefore, calculate: $$\overline{I}_B = \frac{\overline{V}_1 - \underline{V}_{BE}}{\underline{R}_K} - \frac{\underline{V}_{BB} - \underline{V}_{BE}}{\overline{R}_B}$$ $\underline{V}_{BE}$ must be found from the data sheet.	13. $\underline{V}_{BE}$ occurs at high temperatures, so its value must be found at 85°C. $\underline{V}_{BE}$ is specified at .3 volt at 25°C. Using the temperature coefficient curve of the 2N964A data sheet: $\underline{V}_{BE} = 0.197\text{V}.$ $n_R = 5\%$ $$\overline{I}_B = \frac{9.75 - .197}{5.4\text{ K}} - \frac{12\,(.99) - .197}{54.4\text{ K}} = 1.56\text{ mA}$$ $n_R = 20\%$ $$\overline{I}_B = \frac{9.67 - .197}{2.55\text{ K}} - \frac{12\,(.99) - .197}{38} = 3.41\text{ mA}$$

Step 14	14.
Check to see if maximum reverse bias is within limits. This value is also needed to determine turn-on delay time. $$\overline{V}_{OB} = \frac{\overline{V}_{BB}\,\overline{R}_K}{\overline{R}_K + \underline{R}_B}$$ If $\overline{V}_{OB}$ exceeds the limit for the chosen transistor, then either tolerance must be tightened or another transistor selected.	$$n_R = 5\%$$ $$\overline{V}_{OB} = \frac{12\,(1.01)\,(5.85\text{ K})}{5.85\text{ K} + 50.2\text{ K}} = 1.26\text{ Volt}$$ $$n_r = 20\%$$ $$\overline{V}_{OB} = \frac{12\,(1.01)\,(3.45\text{ K})}{3.45\text{ K} + 28\text{ K}} = 1.33\text{ Volt}$$ The values are within limits for the 2N964A transistor.

This completes the dc design of the flip-flop, and the values necessary for the transient design are also known. Note the price paid for using the 5% initial tolerance resistors in comparison to the 1% components.

1) I_C almost doubles. This results in:
   a) Greater possibility of latch-up difficulties.
   b) A greater charge requirement to produce a given rise time.
   c) Less efficient operation and more circuit power dissipation.
2) I_B increases by a factor of 2½. This causes more stored charge which must be removed during the storage time interval resulting in either a longer turn-off time or a reduction in maximum repetition rate.

These comments apply, of course, under the worst-case condition. The differences would be much less pronounced under nominal operation.

SECTION 7-3 – THE ASTABLE MULTIVIBRATOR

Although the simple astable multivibrator shown in Figure 7-3-1 has serious limitations as a pulse generator, it does find uses in applications where only one transition time need be fast and where the pulse duration can have wide tolerance. It is easily synchronized, however, which makes the total period very stable. Of most importance is the fact that a study of its operation and design illustrates the principles basic to any R-C timing Circuit.

Consider the circuit of Figure 7-3-1 with the action stopped at a point in the cycle when Q_2 is being held on by current through R_{B2} and Q_1 is being held off by a reverse bias voltage from C_1 as indicated by time A on the waveforms shown in Figure 7-3-2. Capacitors C_1 and C_2 will charge toward the voltage which they "see". For proper operation, C_2 must become fully charged before any change of state occurs. It charges through the parallel resistance of R_{C1} and R_{L1}, which will be referred to as R'_{C1}. A voltage of $V_{11} - V_{BE2}$ is developed across C_2. Capacitor C_1 is also charging but it sees a voltage of $V_{BB} - SV_{CE2}$. When C_1 develops a slightly positive voltage (V_{TF}) on the base of Q_1 with respect to ground, Q_1 begins to conduct which lowers the output level. This voltage drop is coupled through C_2 to Q_2 and starts to turn it off. As a result, the collector voltage of Q_2 rises; this change is coupled through C_1 to the base of Q_1 aiding turn on of Q_1. The process is regenerative and it proceeds rapidly until Q_1 is on and Q_2 off. From the waveforms, it is apparent that a pulse applied to a base in phase with the base timing voltage can cause early conduction and lock the multivibrator in synchronization with this external signal.

The regenerative transient is normally so rapid that the charges on the capacitors do not change during the transient. Thus, the voltage on C_1 is approximately zero, but C_2 carries a voltage of $V_{11} - V_{BE2}$ which biases Q_2 well into the cutoff region. Conditions are as when first examined, but the states of Q_1 and Q_2 have interchanged.

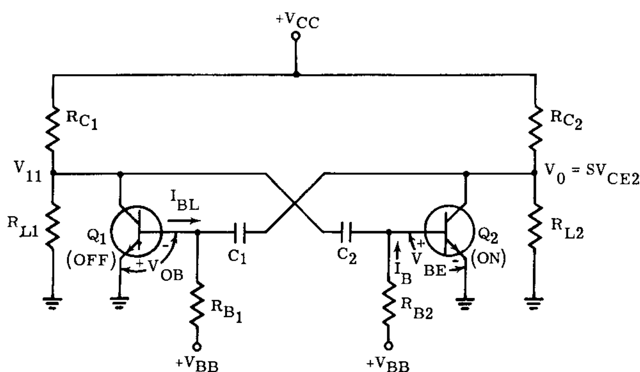


Figure 7-3-1 — Basic Astable Multivibrator

Inter-relations between component values become evident by closer examination of the waveforms shown in Figure 7-3-2. Notice that the recharge time of C_2 severely slows the rise time of the waveform at the collector of Q_1. The waveform can be made more rectangular by decreasing the $R'_{C1}C_2$ time constant. Since R'_{C1} is determined by the load, the only way to improve rise time is to decrease C. However C_2 and R_{B2} established the long timing waveform at the base of Q_2. Therefore if C is reduced, R_{B2} must increase. Since the load is fixed, increasing R_{B2} would require Q_2 to have a higher current gain because R_{B2} establishes the on base current to Q_2. Therefore, waveform improvement at one output comes only by increasing the current gain of the transistor which drives the other output. The more asymmetrical the timing requirement, the more difficult it is to obtain a fast transition time at the output which produces the short pulse.

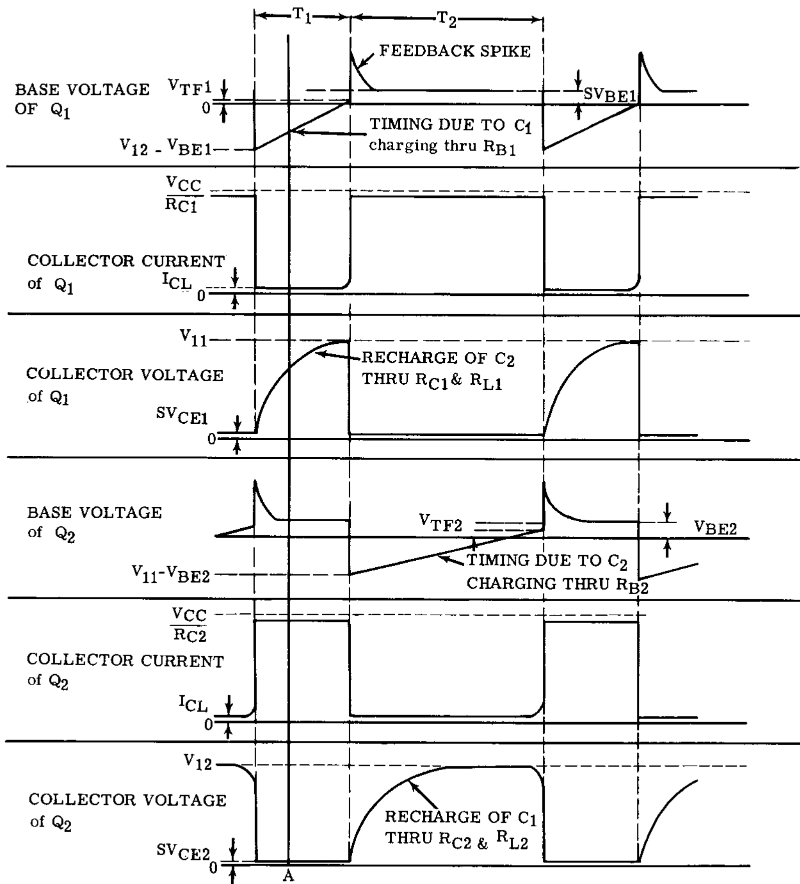


Figure 7-3-2 — Astable Multivibrator Wave Forms

The simple astable circuit can maintain fixed timing only when the load (R_L) is fixed because the output levels V_{11} and V_{12} appear as levels upon the capacitors during the initial part of the timing interval. If variable loads must be handled, then either output amplifiers or clamp diodes must be used if timing is to be fixed. However, in practice timing is never really fixed because of component variations with temperature and ageing.

7-3-1 — Circuit Analysis

The choice of values for R_C and V_{CC} is based upon the same considerations as with an inverter or flip-flop; therefore, Figure 7-1-5 can be used to select a suitable V'_1/V_{CC} ratio and R_C is determined by

$$R_C = \frac{V_{CC} - V'_1}{I_1}. \qquad (7\text{-}3\text{-}1)$$

When a transistor is on, it must supply the current required by R_C and the charging current to the timing capacitors. When the transistor just switches on, the timing capacitor has a voltage across it of $V_1 - V_{BE}$ which is of opposite polarity to V_{BB}. The charging current flows through the collector of the on transistor. Maximum collector current would occur if the saturation voltages were zero; therefore, for simplicity they will be neglected. Thus,

$$I_{C1} = \frac{V_{CC}}{R_{C1}} + \frac{V_{12} + V_{BB}}{R_{B2}} \qquad (7\text{-}3\text{-}2a)$$

and

$$I_{C2} = \frac{V_{CC}}{R_{C2}} + \frac{V_{11} + V_{BB}}{R_{B1}}. \qquad (7\text{-}3\text{-}2b)$$

When Q_1 is on, C_1 is recharging to V_{12} and this current also drives Q_1 deep into saturation. However, for proper operation, C_1 must be fully charged before a change of state occurs. Thus, the C_1 charging current cannot be depended upon to hold Q_1 on. Therefore, the design equations for the base currents are given simply by the dc conditions which are:

$$I_{B1} = \frac{V_{BB} - V_{BE1}}{R_{B1}} \qquad (7\text{-}3\text{-}3a)$$

and

$$I_{B2} = \frac{V_{BB} - V_{BE2}}{R_{B2}}. \qquad (7\text{-}3\text{-}3b)$$

The forced transistor gain has been defined as $\beta_F = I_C/I_B$ where $\beta_F < \beta_0$ (β_0 is the transistor gain at the edge of saturation). Intuitively it should be clear that β_F should be as high as possible to minimize the ratio of recharge time to delay time. The equations just developed govern the dc conditions.

The common case occurs when the circuit has identical output levels and uses identical transistors. Combining all the previous expressions in order to tie the dc conditions together:

$$I_C = \frac{V_{CC}}{R_C} \cdot \frac{1}{1 - \dfrac{1 + (V_1/V_{CC})}{\beta_F}} \qquad (7\text{-}3\text{-}4)$$

and

$$\overline{R}_B = \frac{\overline{\beta}_F}{\overline{I}_C}\left(\underline{V}_{CC} - \overline{V}_{BE}\right). \qquad (7\text{-}3\text{-}5)$$

In deriving equation 7-3-4, V_{BE} was neglected since it is a conservative simplification. Also V_{BB} was taken equal to V_{CC} and $\underline{V}_{CC}$ was used consistently since $\overline{V}_{CC}$ and $\underline{V}_{CC}$ cannot occur at the same time. The reason for doing this will be apparent when the timing equations are discussed.

7-3-2 — Transient Circuit Analysis

TIMING: Use will be made of the basic timing equation:

$$T = RC \ln \frac{V_F - V_I}{V_F - V_{C\,(T)}}, \qquad (7\text{-}3\text{-}6)$$

where

T is the interval of interest

RC is the circuit time constant

V_F is the voltage which the capacitor is charging toward

V_I is the initial capacitor voltage

$V_{C\,(T)}$ is the capacitor voltage which corresponds to time T.

Note that in the astable circuit the initial capacitor voltage is of opposite polarity to the final voltage and therefore V_I has a negative sign. The effect of the base leakage current I_{BL} must be included in the basic timing equation because it adds to the current through R_B, thus making the capacitor charge faster. In effect, I_{BL} makes the final voltage appear to be $V_{BB} + I_{BL}R_B$. Capacitor C_1 determines the off time of Q_1; this time will be called T_1. Applying the notation indicated on Figures 7-3-1 and 7-3-2 to the general timing equation, write

$$T_1 = R_{B1}C_1 \ln\left[\frac{(V_{BB} + I_{BL1}R_{B1} - SV_{CE2}) + (V_{12} - V_{BE1})}{(V_{BB} + I_{BL1}R_{B1} - SV_{CE2}) - (V_{TF1} - SV_{CE2})}\right]. \qquad (7\text{-}3\text{-}7)$$

This equation can be simplified and worst-case limits can be included. If the multivibrator is to be locked in synchronism with another signal, the worst-case is when the time is minimum since the trigger can be made to shorten a timing interval by causing a transistor to conduct early. (Using the trigger to lengthen a timing interval is not practical.) Therefore

$$\underline{T}_1 = \underline{R}_{B1}\underline{C}_1 \ln\left[\frac{\overline{V}_{BB} + \overline{I}_{BL1}\underline{R}_{B1} + \underline{V}_{12} - \overline{S}V_{CE2} - \overline{V}_{BE1}}{\overline{V}_{BB} + \overline{I}_{BL1}\underline{R}_{B1} - \underline{V}_{TF1}}\right]. \qquad (7\text{-}3\text{-}8)$$

The expression for T_2 is identical except for a change in subscripts. Actually, $\underline{V}_{BB}$ is the worst-case condition in the numerator and $\overline{V}_{BB}$ is the worst-case condition in the denominator of the equation. However, since $\overline{V}_{BB}$ and $\underline{V}_{BB}$ cannot occur simultaneously, $\overline{V}_{BB}$ represents the worst-case condition because the effect of V_{BB} is more pronounced in the denominator. Likewise, $\underline{R}_{B1}$ represents the worst-case condition for R_{B1} since its effect is greatest outside the

logarithm term. Note that $\overline{I}_{BL1}$, $\overline{V}_{BE1}$ and $\underline{V}_{TF1}$ cannot occur at the same temperature. Whether the high or low temperature limit represents worst-case depends upon the magnitudes of the quantities.

It is informative to put the timing equation in a slightly different form by dividing through by $\overline{V}_{BB} + \overline{I}_{BL1}\underline{R}_{B1}$:

$$\underline{T}_1 = \underline{R}_{B1}\underline{C}_1 \ln \frac{1 + \dfrac{\underline{V}_{12} - \overline{S}V_{CE2} - \overline{V}_{BE1}}{\overline{V}_{BB} + \overline{I}_{BL1}\underline{R}_{B1}}}{1 - \dfrac{\underline{V}_{TF1}}{\overline{V}_{BB} + \overline{I}_{BL1}\underline{R}_{B1}}}. \tag{7-3-9}$$

From this form, it is apparent that V_{BB} should be fairly large in order to cause timing to be independent of V_{TF}. When V_{BB} is large, V_{TF} is a voltage on the steeply rising portion of the exponential base voltage waveform. Note also that $\underline{V}_{12}$ should be large compared to changes in $\overline{S}V_{CE}$ and $\overline{V}_{BE}$ if their effect upon the timing is to be small. Also, $\overline{I}_{BL}\underline{R}_{B2}$ should be small compared to $\overline{V}_{BB}$ if its effect is to be small. However, the base current is approximately V_{BB}/R_B. In other words, the on base current should be many times $\overline{I}_{BL}$.

Further, note that if V_1 is derived from the V_{BB} supply through a resistor, then changes in V_{BB} will appear proportionally at V_1 if the load is also a resistor, making the timing independent of the supply voltage. In this case, $\underline{V}_{12}$ and $\overline{V}_{BB}$ cannot appear together and $\underline{V}_{12}$ should be multiplied by N_P to represent a realistic worst-case. In addition, when $V_{CC} = V_{BB}$, a compensating effect upon the circuit gain (β_F) occurs making β_F independent of the supply voltage since I_C and I_B would change proportionately with the supply. That is: $I_C \approx V_{CC}/R_C$, $I_B \approx V_{CC}/R_B$; therefore $\beta_F \approx R_B/R_C$ if the saturation voltages are neglected. The worst-case condition is given approximately by $\overline{\beta}_F = \overline{R}_B/\underline{R}_C$.

When V_1 is derived from a clamp diode, in order to make the time independent of V_1, the clamp supply V_K must be closely regulated since $V_1 = V_K + V_D$. If the diode is chosen to be of the same semiconductor material as the transistor, the voltages V_D and V_{BE} are approximately the same and normally they will have the same temperature coefficient which will increase the temperature stability of the timing.

In the case of identical output voltages, loads, and transistors, values for all quantities in the equations for T_1 and T_2 are identical except for the value of the capacitors. Under these special conditions,

$$\frac{T_1}{T_2} = \frac{C_1}{C_2}. \tag{7-3-10}$$

RECOVERY: During time T_1, when Q_1 is off, C_2 must recharge to V_{12} and similarly C_1 must recharge to V_{11} during T_2. This condition puts a restriction upon the $R'_C C$ time constant. In general

$$\underline{T}_1 \geqq \eta_1 \overline{R}'_{C1}\overline{C}_2 \tag{7-3-11a}$$

$$\underline{T}_2 \geqq \eta_2 \overline{R}'_{C2}\overline{C}_1 \tag{7-3-11b}$$

where η is a recovery factor.

η can approach 1 if a clamp circuit is used as discussed with the inverter. For a circuit not employing clamp diodes, η must be about 4 or greater if accurate

timing is to be performed ($\eta = 3.9$ would result in 98% of the final voltage being placed on C). η_1 and η_2 need not be the same — in fact they usually are not. As η is increased the output waveform will have sharper corners.

CHARGE CONSIDERATIONS: The minimum pulse width for a reliable saturated mode astable circuit can be found by establishing the criterion that the charge on the timing capacitor Q_C be at least ten times the charge stored in the transistor Q_T. The following relations apply:

$$\beta_F = I_C/I_B$$

$$Q_C = V_1 C \approx T\, I_B$$

$$\bar{Q}_T < Q_C/10$$

Combining these equations, and solving for $\bar{Q}_T$,

$$\bar{Q}_T \ll \frac{I_C T}{10\,\beta_F} \qquad (7\text{-}3\text{-}12a)$$

If storage is principally in the base as in an alloy device, the minimum pulse width can be related to τ_{BS} by assuming that Q_T is composed of only excess charge Q_x. The following relation can replace Q_T in the above expression:

$$Q_T \approx Q_x = I_B \tilde{\tau}_{BS}$$

and thus:

$$\tilde{\tau}_{BS} < \frac{T}{10} \qquad (7\text{-}3\text{-}12b)$$

The value of τ_{BS} usually decreases slightly as I_B is increased; therefore increasing the on drive current would reduce the minimum timing pulse available with a given transistor, but this slight improvement would require either reduced η or higher collector current.

7-3-3 — Development of Synthesis Procedure

DETERMINING β_F: Using relations for the case where no diodes are used and the load is resistive, some important concepts can be developed. For best timing stability choose $V_{CC} = V_{BB}$. Further, let $V_{12} = V_{11} = V_1$ and $R_{L1} = R_{L2} = R_L$, a restriction which makes $R_{C1} = R_{C2} = R_C$. Also assume that V_{CC} is large compared to the transistor voltages and neglect I_{BL}. The timing equations can be combined to enable the best choice of R_{B1}, R_{B2}, C_1, and C_2 to be made under this set of conditions and approximations which are typical of most astable multivibrator applications.

From equation 7-3-9,

$$T_1 = R_{B1} C_1 \ln\left(1 + \frac{R_L}{R_L + R_C}\right) \qquad (7\text{-}3\text{-}13a)$$

$$T_2 = R_{B2} C_2 \ln\left(1 + \frac{R_L}{R_L + R_C}\right) \qquad (7\text{-}3\text{-}13b)$$

Usually T_1 and T_2 must be kept within specified limits. A simple method for gaining timing stability is to adjust the capacitors in each circuit after con-

struction to meet nominal timing requirements. This greatly reduces the overall tolerance on T because trimming can compensate for initial component tolerances.

If the capacitors are trimmed to meet timing conditions, the resistors must be expected to be at worst-case conditions when choosing the timing capacitor and trimmer. When the resistor values of equation 7-3-13 are at their minimum, a capacitor is needed that can restore nominal timing. This is the largest capacitor value needed by the design, and determines one limit on the value of the capacitor. In general,

$$\overline{C}'' = \frac{\tilde{T}}{\underline{R}_B} \frac{1}{\ln\left[1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right]} \tag{7-3-14}$$

By similar reasoning, the other limit on the capacitor value is:

$$\underline{C}'' = \frac{\tilde{T}}{\overline{R}_B} \frac{1}{\ln\left[1 + \frac{\overline{R}_L}{\overline{R}_L + \underline{R}_C}\right]} \tag{7-3-15}$$

Recovery time needed to recharge C_1 must also be kept within limits. If C'' is at its maximum the design must still meet equations 7-3-11.

Even though $\overline{C}''$ is all that is ever needed, changes with temperature and ageing could increase its value with a corresponding increase in recovery time. To allow for full recovery, this tolerance should be included. The limits for the capacitors are

$$\overline{C}' = (1 + n_C)\overline{C}'' = (1 + n_C)\frac{\tilde{T}}{\underline{R}_B} \frac{1}{\ln\left[1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right]} \tag{7-3-16}$$

and

$$\underline{C}' = (1 - n_C)\,\underline{C}'' = (1 - n_C)\frac{\tilde{T}}{\overline{R}_B} \frac{1}{\ln\left[1 + \frac{\overline{R}_L}{\overline{R}_L + \underline{R}_C}\right]}. \tag{7-3-17}$$

Combine equation 7-3-11a and 7-3-11b with 7-3-16, and substitute the nominal values and tolerance for the minimum and maximum resistance and capacitance values. The capacitors can be eliminated from the equations to find

$$\underline{T}_1 \geqq \eta_1 N_R N_{C2} \tilde{T}_2 \frac{\overline{R}'_C}{\underline{R}_{B2}} \left[\frac{1}{\ln\left(1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right)}\right] \tag{7-3-18a}$$

$$\underline{T}_2 \geqq \eta_2 N_R N_{C1} \tilde{T}_1 \frac{\overline{R}'_C}{\underline{R}_{B1}} \left[\frac{1}{\ln\left(1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right)}\right]. \tag{7-3-18b}$$

Also note that $V_{CC}/R_B \approx I_B$ and $V_{CC}/R_C \approx I_C$; since $I_C/I_B = \beta_F \cong R_B/R_C$ we can solve for the β_F needed to give a specified duty cycle and recovery factor in terms of the tolerances. Since $\overline{\beta}_F$ which is $\overline{I}_C/\underline{I}_B$ is required, $\overline{R}_B$ and $\underline{R}_C$ must be used in the equations. This is easily done since $\overline{R}_B = (1 + n_{RB})\,R_B$ and $\underline{R}_C$ is $(1 - n_{RC})\,R_C$. Also, $\overline{R}'_C = \overline{R}_C\overline{R}_L/(\overline{R}_C + \overline{R}_L)$.

Making these substitutions and assuming the tolerance of all resistors is equal:

$$\overline{\beta}_{F1} \geqq \frac{\tilde{T}_1}{\underline{T}_2} N_C\, \eta_2\, N_R^2 \frac{\overline{R}_L}{(\overline{R}_L + \overline{R}_C) \ln\left(1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right)} \tag{7-3-19}$$

$$\overline{\beta}_{F2} \geqq \frac{\tilde{T}_2}{\underline{T}_1} N_C\, \eta_1\, N_R^2 \frac{\overline{R}_L}{(\overline{R}_L + \overline{R}_C) \ln\left(1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right)} \tag{7-3-20}$$

Note that the required gain is directly proportional to the ratio of the long pulse to the short pulse. The minimum transistor gain β_o must be larger than $\overline{\beta}_F$.

The significance of equations 7-3-19 and 7-3-20 lies in the fact that in order to sharpen the rise of the waveform at Q_2 (η_2 larger), the gain of Q_1 must increase. If $T_1 > T_2$ (resulting in unequal capacitor values), even higher gain is required to produce a fast rise time.

The resistor tolerances can add a considerable factor to the maximum β_F necessary. The terms involving R_C/R_L asymptotes to unity when $R_C \gg R_L$ and asymptotes to 1.44 when $R_L \gg R_C$. For a conservative $\overline{\beta}_{F1}$, equations 7-3-19 and 7-3-20 can be written as

$$\overline{\beta}_{F1} \geqq 1.44 \frac{\tilde{T}_1}{\underline{T}_2} \eta_2 N_C N_R^2 \tag{7-3-21}$$

$$\overline{\beta}_{F2} \geqq 1.44 \frac{\tilde{T}_2}{\underline{T}_1} \eta_1 N_C N_R^2 \tag{7-3-22}$$

In order to find β_F to meet nominal timing requirements, T_1 and T_2 in the previous equations must be expressed in terms of nominal values. The timing equation 7-3-13will be examined a little closer to determine what timing tolerance can be expected from the circuit. If trimmers are used, the capacitor tolerances affecting timing are due to temperature and ageing only. $\overline{T}$ and $\underline{T}$ are:

$$\overline{T} = \overline{R}_B \overline{C} \ln\left[1 + \frac{\overline{R}_L}{\overline{R}_L + \underline{R}_C}\right] \tag{7-3-23}$$

$$\underline{T} = \underline{R}_B \underline{C} \ln\left[1 + \frac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right]. \tag{7-3-24}$$

The tolerance of T can be found by dividing $\overline{T}$ and $\underline{T}$ and remembering that $\overline{R}_B/\underline{R}_B = N_{RB}$, and $\overline{C}/\underline{C} = N_C$. If $N_{RL} = N_{RC} = N_{RB} = N_R$, then,

$$\frac{\overline{T}}{\underline{T}} = N_R\, N_C \frac{\ln\left(1 + \frac{1}{1 + R_C/R_L\, N_R}\right)}{\ln\left(1 + \frac{1}{1 + N_R\, R_C/R_L}\right)}. \tag{7-3-25}$$

Examination of the ln term reveals an asymptote to N_R^2 when R_C is large and an asymptote to unity when R_L is large. Conservatively written, equation 7-3-25 simplifies to

$$\frac{\overline{T}}{\underline{T}} = N_C N_R^3 = N_X . \tag{7-3-26}$$

To obtain the nominal value write

$$\underline{T} = \widetilde{T}(1 - n_X).$$

Previously the term N was defined as $N = (n + 1)/(n - 1)$. Solving for n in terms of N it is found that

$$n = \frac{N - 1}{N + 1}.$$

Therefore

$$\underline{T} = \widetilde{T}\left(1 - \frac{N_C N_R^3 - 1}{N_C N_R^3 + 1}\right)$$

which simplifies to

$$\underline{T} = \widetilde{T}\left(\frac{2}{N_C N_R^3 + 1}\right). \tag{7-3-27}$$

Substituting 7-3-27 into 7-3-21 and 7-3-22

$$\beta_{F1} \geqq \frac{\widetilde{T}_1}{\widetilde{T}_2}\eta_2 N_T \tag{7-3-28a}$$

$$\beta_{F2} \geqq \frac{\widetilde{T}_2}{\widetilde{T}_1}\eta_1 N_T \tag{7-3-28b}$$

where $N_T = 0.72\,(N_R^3\, N_C + 1)\,(N_R^2\, N_C)$.

This result simply says that the maximum required circuit gain is proportional to the ratio of the long pulse to the short pulse, the recovery factor and a factor due to tolerance.

The tolerance factor N_T, is plotted on Figure 7-3-3. It varies from 1.44 for zero tolerance to 10.5 for the case where n_R is 20% and n_C is 10%. Tolerance is obviously an important factor. The tolerance factor given in equation 7-3-28 is conservative. If the factor includes the effect of ageing then assurance is given that the trimmer capacitor can always set the timing to the proper value. Normally, the tolerance N_C would only include temperature effects, on the assumption that timing would be reset to the proper value due to drift with age. However, if the timing were to be set only once and allowed to drift with age, then N_C should also include ageing effects.

In order to design circuits, it is necessary to have a relationship between t_r and T. From equation 7-3-11, assuming the timing is adjusted to nominal

$$T_1 = \underline{\eta}_1\, R'_{C1}\, C_2 \tag{7-3-29a}$$

$$T_2 = \underline{\eta}_2\, R'_{C2}\, C_1 \tag{7-3-29b}$$

where $\underline{\eta}_1$ and $\underline{\eta}_2$ represent the lowest values for η which could occur. The rise time is always 2.2 $R'_C C$. Therefore,

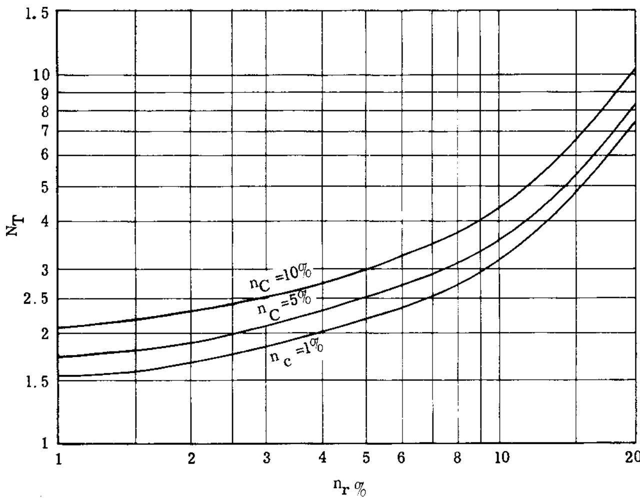


Figure 7-3-3 — Astable Multivibrator Tolerance Multiplier

$$\underline{t}_{r1}/\widetilde{T}_1 = \frac{2.2}{\underline{\eta}_1} \tag{7-3-30a}$$

$$\underline{t}_{r2}/\widetilde{T}_2 = \frac{2.2}{\underline{\eta}_2} \tag{7-3-30b}$$

Equations 7-3-28, 7-3-29 and 7-3-30 can be used to select a suitable transistor and make a preliminary design. Several examples are given to illustrate the use of the equations.

Example 1.

Given: $T_1 = T_2$. Required: β_{F1}, β_{F2}

Only output desired is a differentiated spike, therefore η may be as low as 4.

All resistor tolerances $= 5\%$. $\therefore N_T = 2.5$ from Figure 7-3-3

Capacitor tolerance $= 5\%$.

Solution:

Since $T_1 = T_2$, $\beta_{F1} = \beta_{F2} = \beta_F$, $\eta_1 = \eta_2 = \eta$.

Using equation 7-3-28

$$\beta_F = \frac{T_1}{T_2}\,\eta N_T = (1)(4)(2.5) = 10.$$

Any transistor could fulfill this requirement.

Example 2.

Given: $T_1 = 9\, T_2$. Required: Find β_o

Pulse to be used for gating, therefore, a flat top is desired and t_r may not exceed 20% of T_2.

Assume $n_C = 5\%$ and $n_R = 10\%$.

From Figure 7-3-3, $N_T = 3.5$.

Using equation 7-3-28a, $\bar{\beta}_{F1} \geq \frac{\tilde{T}_1}{\tilde{T}_2} \eta_2 N_T$.

From equation 7-30, $\underline{\eta_2} = \frac{2.2\, T_2}{t_{r2}} = (2.2)\,(5) = 11.$

Solution:

$\beta_o = \bar{\beta}_{F1} = (9)\,(11)\,(3.5) = 346$

No commercially available transistors have a β_o greater than 110, so either a Darlington connection of two transistors must be used or η must be compromised. Assume a transistor with a $\beta_o > 100$ is available and the tolerance can be tightened to $n_R = 5\%$ and $n_C = 1\%$ which makes $N_T = 2.2$. By solving for η from equation 7-3-28

$$\eta_2 = 5.05; \quad \eta_1 = 410.$$

Thus, the side producing the short pulse (side 2) has a vastly poorer waveshape as can be seen by using equation 7-3-30.

$$t_{r1}/T_1 = \frac{2.2}{\eta_1} = \frac{2.2}{410} = 0.00538$$

$$t_{r2}/T_2 = \frac{2.2}{\eta_2} = \frac{2.2}{5.05} = 0.435$$

The rise time of the long pulse is about 0.5% of its period but the rise time of the short pulse is almost half its period, under worst-case conditions.

SELECTING CAPACITORS: Once the circuit has been built, the fixed capacitor-trimmer combination is adjusted to set the timing to the desired nominal value. When all circuit values (except the capacitor) are such that minimum timing will occur, the fixed capacitor-trimmer combination must be increased to a value such that the timing will still be at the nominal value. Equation 7-3-14 indicates the maximum capacitor value necessary. Similarily equation 7-3-15 is an expression for the minimum capacitance. The fixed capacitor-trimmer combination must be adjustable over this range.

While a good deal of choice is available, there is an absolute maximum value allowed for the fixed capacitor. Assume that the minimum capacitance as determined by equation 7-3-15 is required in a particular circuit. If the fixed capacitor is chosen too large, the combination of trimmer and fixed capacitor may be too large even with the trimmer reduced to its minimum value. Thus, the maximum value for the fixed capacitor (including end of life tolerance and temperature) must be less than $\underline{C}''$ minus the minimum trimmer capacitance or:

$$\bar{C} \leq \underline{C}'' - \underline{C}_{TR} \qquad (7\text{-}3\text{-}31a)$$

At the other extreme, assume a maximum capacitance $\overline{C}''$ were needed for timing. The fixed capacitor must not be chosen too low such that even with the trimmer at its maximum the combination of trimmer and fixed capacitor could not reach $\overline{C}''$, or:

$$\overline{C}_{TR} + \underline{C}_1 \geq \overline{C}'' \tag{7-3-31b}$$

The range of trimmer capacitors varies with the value and the manufacturer. The minimum trimmer capacitance is usually not a mathematical function of the maximum value so that the trimmer capacitance may not be treated as a variable in equations 7-3-31a and 7-3-31b. It is necessary to use a cut and try approach to pick a suitable value for the fixed capacitor and trimmer.

First, let $\underline{C}_{TR} = 0$ in equation 7-3-31a. Then choose a standard capacitor less than $\underline{C}''$. For the standard capacitor chosen, determine its minimum value. Use this in equation 7-3-31b and solve for a $\overline{C}_{TR}$. Next choose a convenient trimmer with a maximum capacitance greater than the $\overline{C}_{TR}$ just determined. Now, the minimum trimmer capacitance can be found and used in equation 7-3-31a to determine a new value for $\overline{C}$. If the standard value previously selected is too large according to the new calculation, a lower capacitor value must be selected and the process repeated.

OTHER CONSIDERATIONS: For the astable multivibrator, γ (as defined in Section 7-1) can be written as

$$\gamma = \frac{\overline{I}_{CCO}}{\overline{I}_1} = \frac{\overline{V}_{CC}}{\underline{R}_C \overline{I}_1}$$

Therefore equation 7-3-4 can be written as

$$\overline{I}_C = \gamma \overline{I}_1 \frac{1}{1 - \dfrac{1 + \overline{V}_1/\underline{V}_{CC}}{\overline{\beta}_F}} \tag{7-3-32}$$

In order to select a transistor with adequate breakdown voltage and low enough Q_T, it is necessary to estimate values for $\overline{I}_B$ and $\overline{V}_1$. Both $\overline{I}_B$ and $\overline{V}_1$ can be determined by tolerance factors. Rewriting equation 7-3-3 and neglecting V_{BE}:

$$\overline{I}_B = N_P N_R \underline{I}_B. \tag{7-3-33}$$

When V_1 is derived from V_{CC} by means of a resistor divider, V_1 is given by:

$$V_1 = \frac{V_{CC} R_L}{R_L + R_C} \tag{7-3-34}$$

Maximum V_1 occurs when R_L and V_{CC} are maximum and R_C is a minimum. Minimum V_1 occurs when R_L and V_{CC} are minimum and R_C is a maximum. Including these factors in equation 7-3-34 write $\overline{V}_1/\underline{V}_1$ as:

$$\frac{\overline{V}_1}{\underline{V}_1} = N_P N_R \frac{\underline{R}_L + \overline{R}_C}{\overline{R}_L + \underline{R}_C} \tag{7-3-35}$$

The ratio of resistors was previously shown to approach N_R as R_C gets large. Thus, conservatively written:

$$\frac{\overline{V}_1}{\underline{V}_1} = N_P N_R^2 \tag{7-3-36}$$

The following astable multivibrator design example illustrates the principles discussed.

TABLE 7-3-1 — REQUIREMENTS OF ASTABLE MULTIVIBRATOR		
Output Voltage Levels		
	Max	Min
V_1	—	4.5V
V_O	0.5V	0
(Loads are connected to both outputs)		
Output Impedance		
$R_L = 100\ \Omega$		
$\overline{R}_L = 105\ \Omega$		
$\underline{R}_L = 95\ \Omega$		
Output Pulse Width		
$\tilde{T}_1 = 1.0\ \mu S$	After capacitor adjustment	
$\tilde{T}_2 = 4.0\ \mu S$	After capacitor adjustment	
$\underline{t}_r = 0.55\ \mu S$		
Temperature Range		
0 to + 60°C		
Available Power Supplies		
$\tilde{V}_{CC} = 24V$	$n_P \approx 1\%$ & $N_p \approx 1.02$	
$\overline{V}_{CC} = 24.2V$	$n_P \approx 1\%$ & $N_p \approx 1.02$	
$\underline{V}_{CC} = 23.8V$	$n_P \approx 1\%$ & $N_p \approx 1.02$	
Additional Information		
± 5% end of life tolerance resistors (1% initial tolerance) are used.		
± 5% initial tolerance capacitors are used (assume trimmers are used to adjust frequency) and tolerance due to temperature and ageing is ± 5%.		

7-3-4 — Astable Multivibrator Design Example

Step	Example
Step 1 List specifications.	1. Specifications are shown in Table 7-3-1.
Step 2 Determine $\overline{\beta}_{F1}$ from equation 7-3-28. $\overline{\beta}_{F1} \geqq \frac{\widetilde{T}_1}{\widetilde{T}_2} \eta_2 N_T$ where $\widetilde{T}_1$ is the long pulse, $\widetilde{T}_2$ the short pulse, N_T may be found from Figure 7-3-3, and η_2 may be found from $\underline{\eta_2} = 2.2\, \widetilde{T}_2/\overline{t_{r2}}$.	2. From Figure 7-3-3 for $n_C = .05$, $n_R = .05$, $N_T = 2.5$ $\underline{\eta_2} = (2.2)(1.0)/(0.55) = 4.0$ $\overline{\beta}_{F1} \geqq \frac{4}{1} (4.0)(2.5) = 40$
Step 3 Determine the characteristics to select a suitable transistor. To obtain I_C use equation 7-3-32. Obtain γ from Figure 7-1-5. Then $\overline{I}_C \approx \gamma I_1 N_R \left[\frac{1}{1 - \frac{1 + V_1/V_{CC}}{\beta_F}} \right]$ where $\overline{I}_1 = \underline{V}_1/\underline{R}_2$. Obtain $\underline{I}_B = (I_C/\overline{\beta}_F)$ use equation 7-3-33 and $\overline{I}_B \approx \underline{I}_B N_P N_R$ Find a transistor type which will handle the collector current, has $\underline{\beta}_o > \overline{\beta}_{F1}$ and check to see that	3. At $V_1/V_{CC} = 4.5/24$ and $n_P = .01$, read $\frac{\gamma}{N_R} = 1.25$ $\quad \overline{V}_1 = (1.02)(1.105)^2(4.5)$ $N_R = 1.105$ $\quad \overline{V}_1 = 5.6V$ $\overline{I}_1 = \frac{4.5}{(100)(0.95)} = 47.4\text{ mA}$ $\overline{I}_C = (1.25)(1.105)(47.4) \left[\frac{1}{1 - \frac{1 + 5.6/23.8}{40}} \right]$ $\overline{I}_C = 67.5\text{ mA}$ $\underline{I}_B = (67.5/40) = 1.68\text{ mA}$.

$\overline{I}_{BL} < \underline{I}_B/10$ $\underline{B}V_{EBO} > \overline{V}_1 \approx N_P N_R^2 \underline{V}_1$ $\overline{Q}_T < Q_C/10$ $\overline{Q}_C = I_C T_2/\beta_F$ where T_2 is the short pulse.	$\overline{I}_B = 1.68\,(1.02)\,(1.105) = 1.9$ mA A check of the 2N2501 data sheet reveals that it has sufficient current capability and that: $\underline{\beta}_o = 40$ at 0°C $\overline{I}_{BL} = 200$ nA at 60°C $< \frac{1.68 \text{ mA}}{10} = 168\ \mu A$ $\underline{B}V_{EBO} = 6V > 5.6V$ $\overline{Q}_C = I_C T_2/\beta_{F2} = (67.5)\,(1)/(40) = 1690$ pC $\overline{Q}_T$ is estimated as 90 at an I_B of 1.9 mA and 60°C which meets the design requirement. (The 2N2501 meets all the design requirements.)
Step 4 Calculate R_C from equation 7-3-1 $R_C = \frac{\underline{V}_{CC} - \underline{V}_{11}}{\overline{I}_1\,(1 + n_R)}$ and select a standard value equal to or less than the calculated value of R_C.	4. $R_C = \frac{23.8 - 4.5}{(47.4)\,(1.05)} = 0.388$ K 383 Ω is a standard 1% tolerance type; then $\underline{R}_C = (0.95)\,(383) = 364\ \Omega$ $\overline{R}_C = (1.05)\,(383) = 402\ \Omega$
Step 5 Calculate an exact value for I_C from equation 7-3-4 $\overline{I}_C = \frac{\underline{V}_{CC}}{\overline{R}_C} \frac{1}{1 - \frac{1 + \overline{V}_1/\underline{V}_{CC}}{\overline{\beta}_F}}$ where $\overline{V}_1$ is found at $\underline{V}_{CC}$.	5. $\overline{I}_C = \frac{23.8}{0.364} \frac{1}{1 - \frac{1 + (\frac{5.6}{1.02})/23.8}{40}} = 68$ mA

Step 6

Calculate R_B by modifying equation 7-3-5

$$R_B = \frac{\overline{\beta}_F}{\overline{I}_C} \cdot \frac{\underline{V}_{CC} - \overline{V}_{BE}}{1 + n_R}$$

$\overline{V}_{BE}$ must be found from the data sheet.

Select a close standard value and find

$$\underline{R}_B = (1 + n_R)\, R_B$$
$$\overline{R}_B = (1 - n_R)\, R_B$$

6.

V_{BE} has a temperature coefficient of 1.9 mV/°C and is 1.0 V maximum at $\underline{50}$ mA. From the typical curve, it can be assumed that $\overline{V}_{BE}$ is approximately 1.1 V at 67 mA and 25°C. The temperature of 0°C would add a voltage of $\underline{(25)}$ (1.9) = 48 mV.

Therefore, $\overline{V}_{BE} = 1.15$ V.

$$R_B = \frac{40}{68} \cdot \frac{23.8 - 1.15}{1.05} = 12.7 \text{ K}$$

12.7 K is a standard 1% value.

Then, using 5% end of life tolerances:

$$\underline{R}_B = (0.95)(12.7) = 12.1 \text{ K}$$
$$\overline{R}_B = (1.05)(12.7) = 13.3 \text{ K}$$

Step 7

Calculate the limits for the timing capacitors using equations (7-3-14) and (7-3-15).

$$\overline{C}'' = \frac{\tilde{T}}{\underline{R}_B} \frac{1}{\ln\left(1 + \dfrac{\underline{R}_L}{\underline{R}_L + \overline{R}_C}\right)}$$

$$\underline{C}'' = \frac{\tilde{T}}{\overline{R}_B} \frac{1}{\ln\left(1 + \dfrac{\overline{R}_L}{\overline{R}_L + \underline{R}_C}\right)}$$

7.

$$\overline{C}_1'' = \frac{4.0}{12.1} \frac{1}{\ln\left(1 + \dfrac{95}{95 + 364}\right)} = 1760 \text{ pF}$$

$$\overline{C}_2'' = \frac{1.0}{12.1} \frac{1}{\ln\left(1 + \dfrac{95}{95 + 364}\right)} = 440 \text{ pF}$$

$$\underline{C}_1'' = \frac{4.0}{13.3} \frac{1}{\ln\left(1 + \dfrac{105}{105 + 402}\right)} = 1600 \text{ pF}$$

$$\underline{C}_2'' = \frac{1.0}{13.3} \frac{1}{\ln\left(1 + \dfrac{105}{105 + 402}\right)} = 400 \text{ pF}$$

Step 8 Choose $\overline{C}$ and select trimmer. $C = \text{standard value} < \dfrac{\underline{C}''}{1 + n_C}$ $\overline{C}_{TR} \geqq \overline{C}'' - \underline{C}$ (7-3-30)	8. Assume 5% initial tolerance capacitors are to be selected. $C_1 < \dfrac{1600}{1.05} = 1520\ \text{pF} \quad C_2 < \dfrac{400}{1.05} = 380\ \text{pF}$ Select: $C_1 = 1500\ \text{pF} \qquad C_2 = 350\ \text{pF}$ $\overline{C}_{TR1} > 1760 - (1 - .05)(1500) = 335\ \text{pF}$ $\overline{C}_{TR2} > 440 - (1 - .05)(350) = 108\ \text{pF}$ Select a 380-45 pF trimmer for C_{TR1} and a 120-20 pF trimmer for C_{TR2}. Thus, $\underline{C}_{TR1} = 45\ \text{pF} \qquad \underline{C}_{TR2} = 20\ \text{pF}$
Step 9 Calculate $\overline{C}$ and select C. $\overline{C} = \underline{C}'' - \underline{C}_{TR}$ (7-3-31) $C = \dfrac{\overline{C}}{1 + n_C}$ Select a standard capacitance value equal to or less than the calculated value.	9. $\overline{C}_1 = 1600 - 45 = 1555\ \text{pF}$ $C_1 = 1475\ \text{pF}$ Standard 5% values are 1300 pF and 1500 pF. To meet tolerance with the selected trimmer a 1300 pF capacitor in parallel with a 150 pF capacitor can be used. Otherwise a new trimmer must be selected to be used with the 1300 pF capacitor. $\overline{C}_2 = 400 - 20 = 380\ \text{pF}$ $C_2 = 362\ \text{pF}$ A standard 5% value is 360 pF

CIRCUIT VARIATIONS

The recovery time can be considerably shortened by employing a clamp diode at the output and choosing the load resistor so that the voltage seen is at least 1.5 $\overline{V}_1$. In this way the capacitor will recharge in less than 4 time constants. However, with a clamp, the capacitor current remains high until the capacitor is charged and it must flow through the base which increases the stored charge of the transistor. It takes approximately a time of $3\tau_x$ for this stored charge to decay which causes a serious limitation at high speeds. This problem is discussed more fully in the monostable multivibrator section. The clamp also results in additional collector current and an additional delay when a transistor is turning on, as regeneration cannot occur until all the current has been switched from the diode to the transistor.

A serious disadvantage of the simple circuit is that the output voltage cannot exceed the emitter-base rating of the transistors. With diffused base transistors this rating is low and often additional modifications are needed to utilize this transistor type in a high speed astable multivibrator.

The circuit can be modified as shown in Figure 7-3-4 to permit V_1 to exceed the emitter-base rating. Disconnect-diodes D_{D1} and D_{D2} also cause a significant improvement in the output waveform by letting the capacitors recharge through R_K. The junction of R_K and C becomes disconnected from the load when a transistor is turned off. Resistors R_K and R_D form a voltage divider which prevents the full capacitor voltage from being applied to the transistor. Clamp diodes are not required, however, their use permits variable loads to be handled at the expense of the penalties previously discussed.

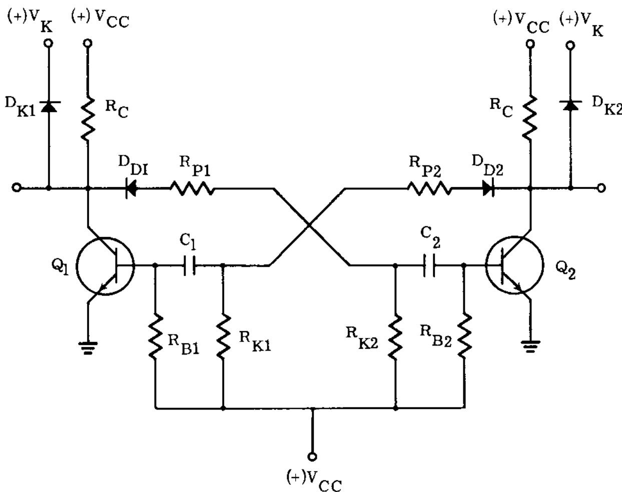


Figure 7-3-4 — "Fast Rise" Astable Circuit

The most satisfactory way of handling variable loads is to use output amplifiers. Used with the simple circuit of Figure 7-3-1 these amplifiers should be designed to switch fully on when the collector signal from the basic astable multivibrator has risen to less than one-half of its final value. In this way, considerable rise time improvement is obtained. A properly chosen speed-up capacitor can insure this condition, but it must be included in calculations for the recharge time of the timing capacitors.

Use of output amplifiers with the circuit of Figure 7-3-4 will provide even greater flexibility and better performance. The clamp diodes are unnecessary, of course, since the output amplifier is a fixed load. The disconnect-diodes will permit faster rise times and the speed-up capacitors of the output amplifiers will not slow the recharge time of the timing capacitors. A circuit utilizing these ideas is shown in Figure 7-3-5.

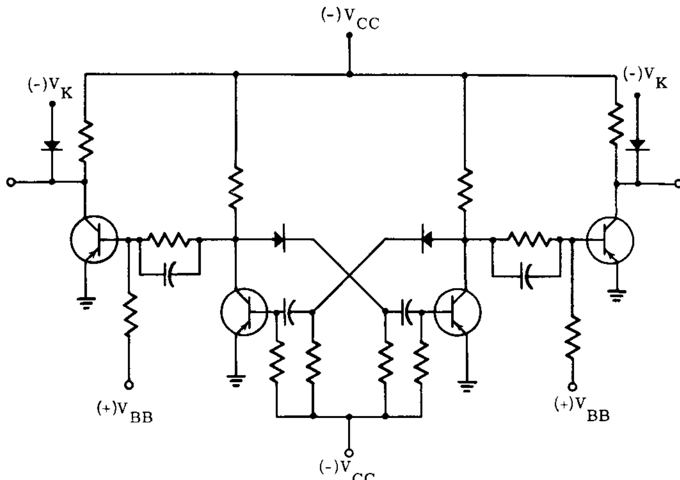


Figure 7-3-5 — Astable Circuit with Output Amplifiers

MOTOROLA STAR* TRANSISTOR
(INSIDE CONSTRUCTION)

* Trademark of Motorola Inc.

Section 7-4 — The Monostable Multivibrator

The monostable multivibrator, often called a one shot or delay flop, has one stable and one quasi-stable state. It must be triggered into the quasi-stable state, where after a given relaxation time it returns to the stable state. The pulse output, which is of a width equal to the relaxation time, is used in timing circuits when fairly long durations of moderate precision are required.

A standard monostable multivibrator is shown in Figure 7-4-1. Current through resistor R_D holds Q_2 normally on. C_D is charged to V_S through the output circuit of Q_1. When the circuit is triggered by turning Q_1 on, C_D discharges through Q_1 and Q_2 until Q_2 turns off. Q_2 is then reverse biased by a voltage of approximately V_S stored on C_D. Now, C_D starts to charge to V_{CC} through R_D. When the voltage at the base of Q_2 is such that it is slightly forward biased, collector current will start to flow. The resulting drop in v_{CE2} is coupled to Q_1 by C_K which causes Q_1 to turn off. When both transistors enter the active region, regeneration occurs which greatly speeds up the transition times.

Before the next trigger pulse arrives, the capacitor must become charged to V_S. V_S must be limited to a value less than the BV_{EBO} rating of Q_2. Use of the clamp circuit, shown dotted, can be used to speed recovery time as well as limiting V_S. Resistor R_S, also shown dotted, is used only when the clamp is not used and when V_{CC} is larger than BV_{EBO}. Since V_S must be limited to a value less than BV_{EBO} of Q_2, the output of Q_1 often cannot be used as an output because the level is too low.

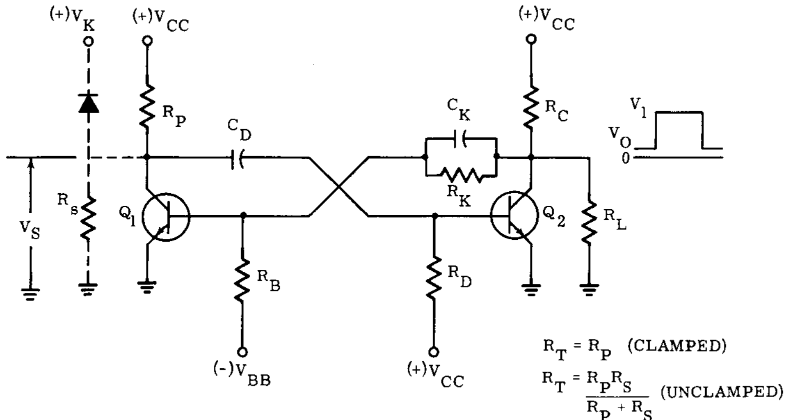


Figure 7-4-1 — Basic Monostable Circuit

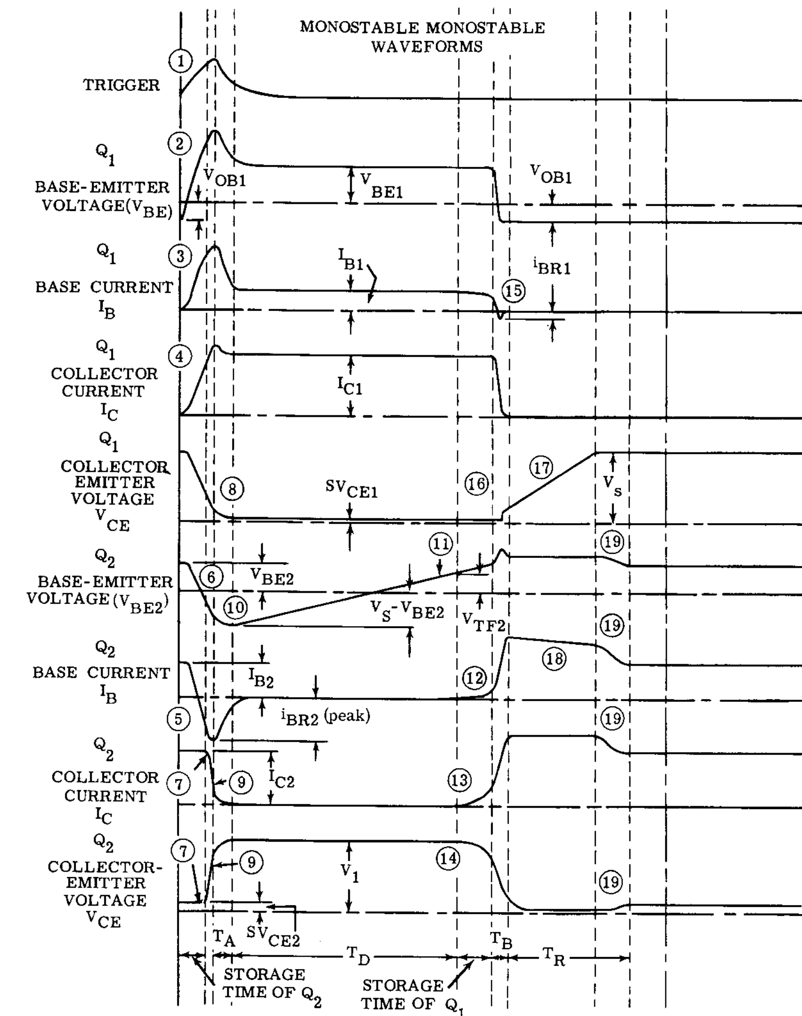


Figure 7-4-2 — Waveforms of the Monostable Multivibrator

Detailed waveforms are shown in Figure7-4-2 for a circuit using the clamp diode. Events are numbered in the order of occurrence.

Events 1 thru 8 occur during the storage time interval of Q_2.

1. The trigger is applied to the base of Q_1. This causes
2. the base voltage of Q_1 to change from its off bias value to a high forward value resulting in
3. considerable base current causing

4. the collector current to increase. It increases considerably above its quasi-stable state value due to the
5. reverse base current of Q_2 caused by the excess charge. The trigger signal must provide the base current to Q_1 to sustain its collector current. At the end of the storage time interval of Q_2 notice that:
6. the base of Q_2 is reverse biased because of the heavy reverse base current (i_{BR2}) which is flowing,
7. the collector voltage and current of Q_2 have not changed and
8. the collector voltage of Q_1 has not fully reached the SV_{CE} value. Now that Q_2 is in the active region, the high i_{BR2} causes its
9. collector current to quickly fall permitting the output voltage to rise and
10. the base voltage to assume the level of $V_S - V_{BE2}$. The rounding of the Q_2 waveforms results from the fall off of trigger signal and the presence of C_K which contributes little to regenerative action at this time. All waveforms are affected, this period is indicated on the waveforms as time interval T_A. At the end of interval T_A, Q_2 is completely cut off and the onset of the quasi-stable state occurs. Q_1 then assumes quasi-stable state conditions.
11. When the voltage at the base of Q_2 has reached V_{TF} at time T_D, it initiates the termination of the output pulse by causing
12. base current to flow in Q_2 which in turn causes
13. collector current to flow in Q_2 thereby,
14. dropping the collector voltage of Q_2. The turn-on of Q_2 results in
15. a lowering and then a reversal of the base current of Q_1. If the storage time of Q_1 is rather long, then the output will reach ground before regeneration has much effect and further action in the circuit will have negligible effect upon the output pulse. In the case shown, the output has dropped about 1/3 when regeneration occurs causing the transition times to speed up. During time T_B, regeneration brings the collector current of Q_2 to the steady state value and cuts off Q_1.

    As the collector current of Q_1 falls, it recharges the capacitor C_D very quickly by a small amount which accounts for
16. the small step on the collector waveform. When i_{C1} reaches zero, then,
17. the recharge of C_D is determined by R_T. As indicated by
18. the recharge current through the base of Q_2 in a direction to cause heavy forward bias resulting in the
19. steps on the waveforms as the capacitor becomes fully charged permitting the base current to decay.

Stable state conditions are now attained and the circuit is ready to receive another trigger. Improvements in switching times can be obtained if the storage time of Q_1 is very short so that the turn-off of Q_1 and the recharge current of C_D can be used to assist Q_2 in turning on. The circuit is not ready to receive a trigger until i_B has decayed from the value used to charge C_D. The time for a 90% decay is $2.3\tau_x$, where τ_x is the lifetime of the excess carriers. This additional time is not present in circuits which do not use a clamp circuit at the output of Q_1 as the base current would exponentally approach its final value.

7-4-1 — Analysis

The output circuit of Q_2 and the input circuit of Q_1 are exactly the same as in the flip-flop circuit. Using the terminology of Figure 7-4-3a and b the equations for the monostable multivibrator are:

$$\overline{I}_1 = \frac{\underline{V}_{CC} - \underline{V}_1}{\overline{R}_C} - \overline{I}_{CL2} - \frac{\underline{V}_1}{\underline{R}_K} \quad (7\text{-}4\text{-}1)$$

$$\frac{\underline{V}_{OB1} - \overline{SV}_{CE2}}{\underline{R}_K} - \frac{\underline{V}_{BB} - \underline{V}_{OB1}}{\overline{R}_B} + \overline{I}_{BL1} = 0 \quad (7\text{-}4\text{-}2)$$

$$\frac{\underline{V}_1 - \overline{V}_{BE1}}{\overline{R}_K} - \frac{\overline{V}_{BB} + \overline{V}_{BE1}}{\underline{R}_B} - \underline{I}_{B1} = 0 \quad (7\text{-}4\text{-}3)$$

where the subscripts 1 and 2 refer to characteristics of Q_1 and Q_2 respectively.

Using the terminology of Figure 7-4-3, the astable timing relations can be adapted to the monostable multivibrator. For minimum delay:

$$\underline{T}_D = \underline{R}_D\underline{C}_D \ln \frac{\overline{V}_{CC} + \overline{I}_{BL2}\underline{R}_D + \underline{V}_S - \overline{SV}_{CE1} - \overline{V}_{BE2}}{\overline{V}_{CC} - \underline{V}_{TF2} + \overline{I}_{BL2}\underline{R}_D} \quad (7\text{-}4\text{-}4a)$$

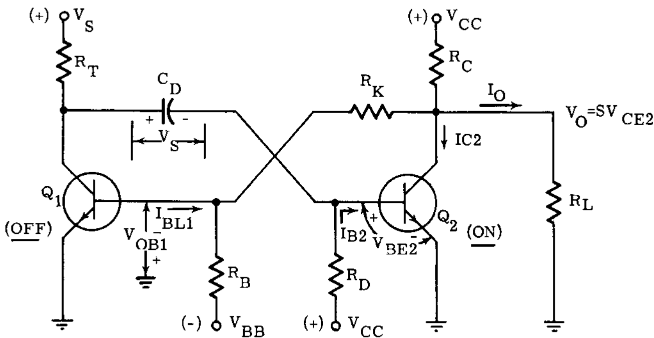


Figure 7-4-3a — Conditions During Stable State

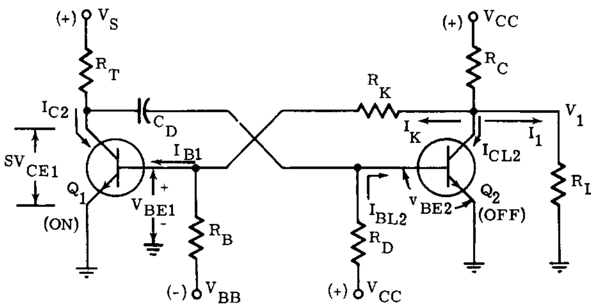


Figure 7-4-3b — Conditions During the Quasi-stable State

Note that $\overline{I}_{BL2}$ and $\overline{S}V_{CE1}$ and $\underline{V}_{TF2}$ occur at the high temperature limit while $\overline{V}_{BE2}$ occurs at the low temperature limit. The worst-case temperature, therefore, depends upon the values of these characteristics.

Since the ln term will appear in many calculations, its terms will be defined as $\underline{x}$. Thus,

$$\underline{T}_D = \underline{R}_D \underline{C}_D \ln \underline{x} \tag{7-4-4b}$$

Conditions must be at the stable state before another trigger signal is applied or the delay time will be shortened. Thus, the maximum recovery time $(\overline{T}_R)$ is determined by:

$$\overline{T}_R = \frac{1}{\overline{f}_I} - \overline{T}_D \tag{7-4-5}$$

where $\overline{f}_I$ = maximum repetition frequency of input trigger

$\overline{T}_D$ = maximum pulse duration

Actually, while Q_2 enters the active region when its base voltage reaches V_{TF}, recovery cannot commence until Q_1 has been turned off. This time lag is difficult to determine mathematically but experience shows that it can be approximated by the time it would take for the base voltage of Q_2 to go from V_{TF} to V_{BE} due to the timing network alone. Therefore, the equation for maximum delay time is obtained by inserting $\overline{V}_{BE2}$ for $\underline{V}_{TF2}$ in equation 7-4-4a and determining the proper worst-case limits for the other terms. This yields

$$\overline{T}_D = \overline{R}_D \overline{C}_D \ln \frac{\underline{V}_{CC} + \overline{V}_S - \underline{S}V_{CE1} - \overline{V}_{BE2}}{\underline{V}_{CC} - \overline{V}_{BE2}} \tag{7-4-6a}$$

This ln term will also appear in many calculations; its terms will be defined as $\overline{x}$. Therefore:

$$\overline{T}_D = \overline{R}_D \overline{C}_D \ln \overline{x}. \tag{7-4-6b}$$

The maximum recovery time equation is:

$$\overline{T}_R = \eta \overline{R}_T \overline{C}_D. \tag{7-4-7}$$

where η must be chosen to allow complete recovery. In the circuit without the clamp η must be at least 4.

When designing a circuit, R_T must be determined by the recovery time which in turn determines I_{C1} and hence I_{B1}. In essence, R_T is fixed by the dc conditions in the quasi-stable state. Thus, the actual value allowed for $\overline{C}_D$ must be determined by the trigger repetition frequency and $\overline{T}_D$ as seen by combining equations 7-4-5 and 7-4-7. Therefore,

$$\overline{C}_D = \frac{\dfrac{1}{\overline{f}_I} - \overline{T}_D}{\eta \overline{R}_T}$$

Substituting $\overline{T}_D$ as given by equation 7-4-6b and solving for $\overline{C}_D$, it is found

$$\overline{C}_D = \frac{1/\overline{f}_I}{\eta \overline{R}_T + \overline{R}_D \ln \overline{x}} \tag{7-4-8}$$

Besides the dc current flowing through Q_1, a transient current, as a result of C_D charging toward V_{CC}, also flows. Thus, the equation for I_{C1} is identical in form to that of the astable

$$\overline{I}_{C1} = \frac{\overline{V}_S}{\underline{R}_T} + \frac{\overline{V}_{CC} + \overline{V}_S}{\underline{R}_D} \tag{7-4-9a}$$

The proper value of $\underline{R}_T$ can be found by substituting $\underline{I}_{B1}\underline{\beta}_{F1}$ for $\overline{I}_{C1}$ and solving for $\underline{R}_T$

$$\underline{R}_T = \frac{\overline{V}_S}{\underline{\beta}_{F1}\underline{I}_{B1} - \dfrac{\overline{V}_{CC} + \overline{V}_S}{\underline{R}_D}}. \tag{7-4-9b}$$

The final equation required is that for R_D. This resistor, as discussed with the astable circuit, should be as high as possible and is given by

$$\overline{R}_D = \frac{\overline{\beta}_{F2}}{\overline{I}_{C2}}(\underline{V}_{CC} - \overline{V}_{BE2}). \tag{7-4-10a}$$

By substituting $\overline{I}_{C2} = \dfrac{\overline{V}_{CC}}{\underline{R}_{C2}}$, $\overline{R}_D$ assumes the form:

$$\overline{R}_D = \beta_{F2}\underline{R}_{C2}\left(\frac{\underline{V}_{CC} - \overline{V}_{BE2}}{\underline{V}_{CC}}\right). \tag{7-4-10b}$$

(Since $\overline{V}_{CC}$ and $\underline{V}_{CC}$ can not occur together, $\underline{V}_{CC}$ is used because it produces worst case conditions.)

7-4-2 — Synthesis Equations

In order to find transistor data it is necessary to first obtain estimates of I_{C1} and I_{C2} based upon the load current required and the timing equations. I_{C2} can be estimated from the load and cross-coupling conditions as discussed in the flip-flop section. I_{C1} can be put in terms of I_{C2} by a process involving the timing equations.

Combining equations 7-4-7 and 7-4-6b

$$\overline{R}_T = \frac{\overline{T}_R\overline{R}_D \ln \overline{x}}{\eta\overline{T}_D}$$

Using the tolerance factors N_{RS} and N_{RD} for the resistors and substituting the above expression into equation 7-4-9a write

$$\overline{I}_{C1} = \frac{\eta\overline{V}_S\overline{T}_D N_{RS}}{\overline{T}_R\overline{R}_D \ln \overline{x}} + \frac{\overline{V}_{CC} + \overline{V}_S}{\overline{R}_D} N_{RD}$$

Substituting equation 7-4-10b for $\overline{R}_D$ in the above expression, and simplifying, it is found that

$$\overline{I}_{C1} = \frac{\overline{I}_{C2}}{\overline{\beta}_{F2}}\left[\frac{\overline{V}_S}{\overline{V}_{CC}}\left(\frac{\eta\overline{T}_D N_{RS}}{\overline{T}_R \ln \overline{x}} + N_{RD}\right) + N_{RD}\right]. \tag{7-4-11}$$

Normally N_{RD} is small compared to the other terms. For a given delay to recovery ratio, it is seen that I_{C1} is inversely proportional to β_{F2}, which means that the current gain of Q_2 should be high to keep I_{C1} down. Also, I_{C1} is proportional to η, therefore, using a clamp circuit will also help reduce I_C as η could be less than 4. A high ratio of V_{CC} to V_S also assists in keeping I_{C1} low.

Equation 7-4-11 may be written as

$$\bar{I}_{C1} = \frac{\bar{I}_{C2}\,\lambda}{\bar{\beta}_{F2}} \qquad (7\text{-}4\text{-}12a)$$

where

$$\lambda = \frac{\bar{V}_S}{\bar{V}_{CC}}\left(\frac{\eta \bar{T}_D N_{RS}}{\bar{T}_R \ln \bar{x}} + N_{RD}\right) + N_{RD}. \qquad (7\text{-}4\text{-}12b)$$

The λ term contains only known quantities, or quantities which can easily be estimated.

As in the flip-flop analysis,

$$\bar{I}_{C2} = \gamma\,(\bar{I}_1 + \bar{I}_K) + \bar{I}_O \qquad (7\text{-}4\text{-}13)$$

and the current $\bar{I}_K$ can be estimated as $2\,\bar{I}_{C1}/\bar{\beta}_{F1}$.

Combining this estimate of $\bar{I}_K$ with equations 7-4-12a and 7-14-13

$$\bar{I}_{C2} = \gamma\,(\bar{I}_1 + \frac{2\,\bar{I}_{C2}\lambda}{\bar{\beta}_{F2}\,\bar{\beta}_{F1}}) + \bar{I}_O$$

which simplifies to

$$\bar{I}_{C2} = \frac{\gamma\,\bar{I}_1 + \bar{I}_O}{1 - \dfrac{2\gamma\lambda}{\bar{\beta}_{F1}\,\bar{\beta}_{F2}}}. \qquad (7\text{-}4\text{-}14)$$

Equations 7-4-12a and 7-4-14 can be used to obtain estimates for $\bar{I}_{C1}$ and $\bar{I}_{C2}$ so that transistor characteristics can be determined. These transistor characteristics are then used to calculate component values.

Also, since $\underline{I}_{B1} = \dfrac{\bar{I}_{C1}}{\bar{\beta}_{F1}}$, equation 7-4-12a can be used in equation 7-4-3 by substituting $\bar{I}_{C2} = \dfrac{\bar{V}_{CC}}{\underline{R}_{C2}} + \bar{I}_O$. This yields,

$$\frac{\underline{V}_1 - \bar{V}_{BE1}}{\bar{R}_K} - \frac{\bar{V}_{BB} + \bar{V}_{BE1}}{\underline{R}_B} - \frac{\lambda\,\bar{V}_{CC}}{\bar{\beta}_{F1}\underline{R}_{C2}\bar{\beta}_{F2}} = \frac{\lambda\,\bar{I}_O}{\bar{\beta}_{F1}\bar{\beta}_{F2}} \qquad (7\text{-}4\text{-}15)$$

7-4-3 — Use of a Clamp Circuit

There are some definite advantages in using a clamp circuit at the collector of Q_1 instead of a resistive network even though the output of Q_1 will not be used.

When using a divider or separate source, the maximum recovery time is given by equation 7-4-7 i.e., $\bar{T}_R = \eta\bar{R}_T\bar{C}_D$.

Using a clamp, the maximum recovery time can be written from the general timing equation, where V_F is $\underline{V}_{CC}$, V_T is $K\bar{V}_S$, and R_T is $\bar{R}_P$.

$$\overline{T}_R = \overline{R}_P \overline{C}_D \ln \frac{1}{1 - \dfrac{K\overline{V}_S}{\underline{V}_{CC}}}$$

Here, V_S indicates the clamping level and K indicates the fraction of V_S of interest. If I_C is to be equal in both the clamped and unclamped case, under worst condition for recovery, then, $\dfrac{\overline{V}_S}{\overline{R}_T} = \dfrac{\underline{V}_{CC}}{\overline{R}_P}$; and

$$\overline{T}_R = \frac{\underline{V}_{CC}}{\overline{V}_S} \overline{R}_T \overline{C}_D \ln \frac{1}{1 - \dfrac{K\,\overline{V}_S}{\underline{V}_{CC}}} \tag{7-4-16}$$

combining 7-14-16 with 7-4-7 find

$$\eta = \frac{\overline{T}_R}{\overline{R}_T\,\overline{C}_D} = \frac{\underline{V}_{CC}}{\overline{V}_S} \ln \frac{1}{1 - \dfrac{K\,\overline{V}_S}{\underline{V}_{CC}}} \tag{7-4-17}$$

Equation 7-4-17 is plotted as Figure 7-4-4. K was taken at 0.98, therefore, Figure 7-4-4 can be used to design for the total recovery time required in monostable circuits. The significance of clipping off the slowly rising part of the waveform is quite evident. Thus, use of the clamp circuit can speed the recovery by a factor of 3 to 4 over that of a resistive network with no increase in collector current, depending upon the $\underline{V}_{CC}/\overline{V}_S$ ratio used.

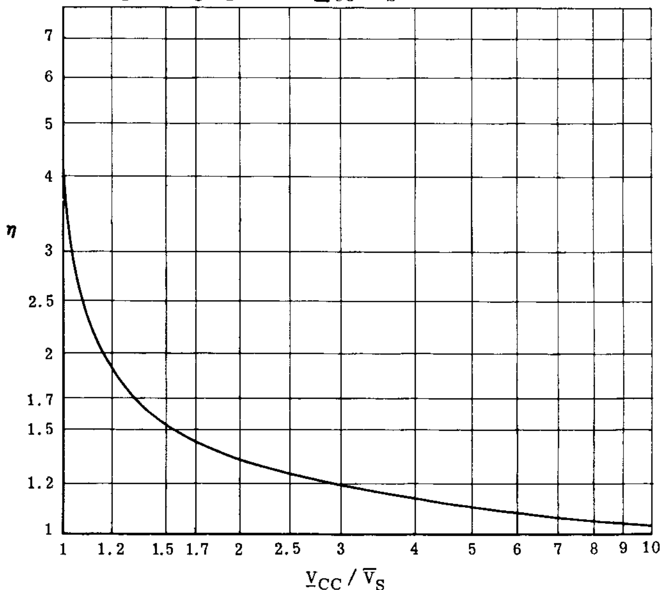


Figure 7-4-4 — Recharge Time Improvement Due to Clamp Circuit

This method also permits the diode to provide temperature compensation for the V_{BE} of the transistor. That is, using the clamp circuit the general timing equation becomes

$$T_D = R_D C_D \ln \frac{V_{CC} + I_{BL2} R_D + V_K + V_D - SV_{CE1} - V_{BE2}}{V_{CC} - V_{TF2} + I_{BL2} R_D}$$

The changes in V_D and V_{BE2} with temperature will be nearly equal leaving the difference between them nearly constant.

Whether the clamp circuit will provide improved performance is dependent upon the lifetime of the excess carriers in Q_2. As indicated in the discussion of the waveforms, the current used to charge C_D takes a time of approximately $2.3\,\tau_x$ to decay after C_D has been charged. The actual time for this current to decay depends somewhat upon C_D. The voltage on C_D changes as V_{BE} drops from this high forward potential to V_{BE} of the stable state. If this time forms a significant portion of the recovery period, it will be found that the unclamped circuit will have to be used in order to keep the recovery time within specifications.

Timing equations can now be written for both the case of the clamped and the unclamped circuit.

Case 1 — Clamped

$$\overline{T}_D = \overline{R}_D \overline{C}_D \ln \overline{x}_c \tag{7-4-18}$$

where

$$\overline{x}_c = \frac{\underline{V}_{CC} + \overline{V}_K + \overline{V}_D - \underline{S}V_{CE1} - \overline{V}_{BE2}}{\underline{V}_{CC} - \overline{V}_{BE2}}$$

$$\underline{T}_D = \underline{R}_D \underline{C}_D \ln \underline{x}_c \tag{7-4-19}$$

where

$$\underline{x}_c = \frac{\overline{V}_{CC} + \overline{I}_{BL2} \underline{R}_D + \underline{V}_K + \underline{V}_D - \overline{S}V_{CE1} - \overline{V}_{BE2}}{\overline{V}_{CC} - \underline{V}_{TF2} + \overline{I}_{BL2} \underline{R}_D}$$

Case 2 — Unclamped

$$\overline{T}_D = \overline{R}_D \overline{C}_D \ln \overline{x}_u \tag{7-4-20}$$

where

$$\overline{x}_u = \frac{\underline{V}_{CC} + \overline{V}_S - \underline{S}V_{CE1} - \overline{V}_{BE2}}{\underline{V}_{CC} - \overline{V}_{BE2}}$$

$$\underline{T}_D = \underline{R}_D \underline{C}_D \ln \underline{x}_u \tag{7-4-21}$$

where

$$\overline{x}_u = \frac{\overline{V}_{CC} + \overline{I}_{BL2} \underline{R}_D + \underline{V}_S - \overline{S}V_{CE1} - \overline{V}_{BE2}}{\overline{V}_{CC} - \underline{V}_{TF2} + \overline{I}_{BL2} \underline{R}_D}$$

The ratio of maximum to minimum delay time is

$$\frac{\overline{T}_D}{\underline{T}_D} = N_{RD} N_{CD} \frac{\ln \overline{x}}{\ln \underline{x}} \tag{7-4-22}$$

where the appropriate x term should be used. Notice that the temperature which gives a worst-case limit to the x terms is dependent upon the magnitudes

of the many transistor terms and may occur at either the high or low limit.

If V_S is derived by divider action from V_{CC}, V_S can be put in terms of R_P and R_S.

$$V_S = \frac{V_{CC}R_S}{R_S + R_P} \tag{7-4-23}$$

For this case, R_T is

$$R_T = \frac{R_S R_P}{R_S + R_P}. \tag{7-4-24}$$

An examination of the R_S, R_P divider network shows that, if the tolerance of R_S equals that of R_P, V_S has the tolerance

$$\frac{\overline{V}_S}{\underline{V}_S} = N_{CC}\, N_R^2 \tag{7-4-25}$$

TABLE 7-4-1

REQUIRED TIMING		
$\underline{T}_D = 450$ nS		$\overline{f}_I = 1$ mc
REQUIRED OUTPUT LEVELS (volts)		
	MIN	MAX
V_O	—	0.5
V_1	6	—
LOAD CURRENT REQUIREMENTS (mA)		
	MIN	MAX
I_O	0	0
I_1	—	9
AVAILABLE POWER SUPPLY VOLTAGE		
MINIMUM	+ 23.8	—5.93
NOMINAL	+ 24	—6
MAXIMUM	+ 24.2	—6.06
POWER SUPPLY TOLERANCE ±1% ($N_P = 1.02$)		
AMBIENT TEMPERATURE RANGE —55°C to +100°C		
RESISTOR TOLERANCE* ±5% (±1% initial) ($N_R = 1.105$)		
CAPACITOR TOLERANCE ± 10% (± 5% initial) ($N_C = 1.22$)		

*Refer to inverter section for a discussion of tolerances.

7-4-4 — Monostable Design Procedure

Step 1 State the requirements of the design	1. Specifications are shown in Table 7-4-1.
Step 2 Estimate $\overline{I}_{C2}$ from equation 7-4-13 $\overline{I}_{C2} = \gamma \overline{I}_1 + \overline{I}_O$ where γ is obtained from Figure 7-1-5.	2. From Figure 7-1-5 at V_1/V_{CC} ratio of 0.25 and $n_P = 0.1$ read $\frac{\gamma}{N_R} = 1.35$ $N_R = 1.1$ $I_{C2} = (1.35)(1.1)(9) = 13.3$ mA
Step 3 Find a transistor type suitable for the application.* Check for $\overline{I}_{BL} < \underline{I}_B/10 = \overline{I}_C/10\,\overline{\beta}_F$ $\overline{Q}_T < \underline{Q}_C/10.$ where $\underline{Q}_C = \underline{T}_D\,\underline{I}_B$ *Refer to discussion in the section on the astable multivibrator for background information.	3. A check of the 2N2501 data sheet shows that it has sufficient current capability. From the normalized β curve at −55°C, β is approximately constant between 5 and 20 mA. Thus, the specification at 10 mA and −55°C of $\beta = 20$, holds at 13.3 mA. This specification is at $V_{CE} = 1$ V, whereas V_o is to be about 0.5 V. Since the saturation knee is 0.3 V, a choice of $\overline{\beta}_{F2} = 20$ is reasonable. $\underline{I}_B = \frac{\overline{I}_C}{\overline{\beta}_F} = \frac{13.3}{20} = 0.665$mA $Q_C = \underline{I}_B T_D = (0.665)(450) = 300$ pC. From the data sheet, $\overline{I}_{BL} = 700$nA. @ 100°C $\overline{Q}_T \approx 40$pC @ $I_B = 0.665$ mA. Therefore the 2N2501 is suitable although Q_T may be marginal.

Step 4

Decide if a clamp circuit is to be used to recharge C_D. A clamp is usually desirable unless f_I is rather high.

From the data sheet, estimate the values for the terms in $\overline{x}$ and $\underline{x}$. Since R_D is not known, neglect the I_{BL} term and find $\overline{x}$ and $\underline{x}$.

For the unclamped case

$$\overline{x}_u = \frac{\underline{V}_{CC} + \overline{V}_S - \underline{SV}_{CE1} - \overline{V}_{BE2}}{\underline{V}_{CC} - \overline{V}_{BE2}} \qquad (7\text{-}4\text{-}20)$$

$$\underline{x}_u = \frac{\overline{V}_{CC} + \underline{V}_S - \overline{SV}_{CE1} - \overline{V}_{BE2}}{\overline{V}_{CC} - \underline{V}_{TF2}} \qquad (7\text{-}4\text{-}21)$$

where

$$\underline{V}_S = \frac{\overline{V}_S}{N_P N_R^2} \qquad (7\text{-}4\text{-}22)$$

and $\overline{V}_S \approx 0.9\,\underline{BV}_{EBO}$

4. Given $f_I = 1$mc, no clamp circuit will be used. For the 2N2501, at an $I_C = 13.3$mA, the following values for V_{BE}, SV_{CE}, and V_{TF} are roughly determined to be:

$\underline{V}_{BE} = 0.5$ V $\qquad \overline{V}_{BE} = 1.0$ V at $-55°$C temp.

$\underline{SV}_{CE} = 0.1$ V $\qquad \overline{SV}_{CE} = 0.3$ V

$\underline{V}_{TF} = 0$ V $\qquad \overline{V}_{TF} = 0.3$ V at $-55°$C temp.

$\underline{BV}_{EBO} = 6$V for the 2N2501. Select $\overline{V}_S = 5.4$V.

Since the timing is very sensitive to V_S, use 1% resistors for R_P and R_S with 5% end of life tolerance.

From Table 7-4-1, $N_R = 1.105$.

$$\text{Thus } \underline{V}_S = \frac{5.4}{(1.02)(1.02)^2} = 5.1\text{V}$$

$$\text{and } \overline{x}_u = \frac{23.8 + 5.4 - .1 - 1}{23.8 - 1} = 1.23$$

$$\underline{x}_u = \frac{24.2 + 5.1 - .3 - 1}{24.2 - 0.3} = 1.17$$

Step 5

Determine $\overline{T}_D$ and $\overline{T}_R$ from equations 7-4-23 and 7-4-5

$$\frac{\overline{T}_D}{\underline{T}_D} = N_{RD} N_C \frac{\ln \overline{x}_u}{\ln \underline{x}_u}$$

and

$$\overline{T}_R = \frac{1}{f_I} - \overline{T}_D$$

5. From Table 7-4-1, $N_C = 1.22$ and $f_I = 1$ mc:

$$\frac{\overline{T}_D}{\underline{T}_D} = (1.105)(1.22)\,\frac{\ln 1.23}{\ln 1.17} = 1.78$$

and

$$\overline{T}_D = (1.78)\,(450) = 800 \text{ nS}$$

$$\overline{T}_R = 1000 - 800 = 200 \text{ nS}$$

Step 6 Calculate λ from equation 7-4-12b $\lambda = \frac{\overline{V}_S}{\overline{V}_{CC}} \frac{\overline{T}_D}{\overline{T}_R}\left(\frac{\eta N_{RS}}{\ln(x)} + N_{RD}\right) + N_{RD}$	6. Since no clamp is used to develop V_S, $\eta = 4$ must be used. $\overline{x}_u = 1.23$ $\lambda = \frac{5.4}{24.2} \frac{800}{200}\left[\frac{(4)(1.105)}{\ln(1.23)} + 1.105\right] + 1.105$ $\lambda = 21.2$
Step 7 Estimate $\overline{I}_{C1}$ and $\overline{\beta}_{F1}$ using equation 7-4-12a and transistor data $\overline{I}_{C1} = \lambda \frac{\overline{I}_{C2}}{\overline{\beta}_{F2}}$	7. $\overline{I}_{C1} = \frac{(21.2)(13.3)}{20} = 14.1$ mA $\underline{\beta_o}$ at 14.1 mA and $-55°C = 20$ Let $\overline{\beta}_{F1} = 20$
Step 8 Recalculate $\overline{I}_{C2}$ from equation 7-4-14 and check $\overline{\beta}_{o2}$ $\overline{I}_{C2} = \frac{\overline{I}_1\gamma + \overline{I}_o}{1 - \frac{2\lambda\gamma}{\overline{\beta}_{F1}\overline{\beta}_{F2}}}$	8. $\overline{I}_{C2} = \frac{(10)(1.35)(1.1) + 0}{1 - \frac{(2)(21.2)(1.35)(1.1)}{(20)(20)}}$ $\overline{I}_{C2} = 17.4$ mA $\underline{\beta_o}$ at 17.4 mA and $-55°C$ is 20 Therefore $\overline{\beta}_{F2} = 20$ still holds.

Step	Solution
Step 9 Recalculate $\overline{I}_{C1}$ and check $\underline{\beta}_{o1}$ if a change in $\overline{\beta}_{F1}$ is necessary, $\overline{I}_{C2}$ must be calculated again. $\overline{I}_{C1} = \lambda \frac{\overline{I}_{C2}}{\overline{\beta}_{F2}}$	9. $\overline{I}_{C1} = \frac{(21.2)(17.4)}{20} = 18.4\ \text{mA}$ $\underline{\beta}_{o1}$ is not changed and $\overline{\beta}_{F1} = 20$ is acceptable. Since no change in $\overline{\beta}_{F1}$ or $\overline{\beta}_{F2}$ is necessary, the design can continue.
Step 10 Determine more accurate values for characteristics needed to solve the matrix. Determine: $\overline{SV}_{CE}$ at high temperature $\overline{V}_{BE}$ at low temperature $\overline{I}_{CL}$ at high temperature $\overline{I}_{BL}$ at high temperature	10. From the data sheet $\overline{SV}_{CE1} = \overline{SV}_{CE2}$ @ $100°C = 0.3V$ $\overline{V}_{BE1} = \overline{V}_{BE2}$ @ $-55°C = 0.95V$ $\overline{I}_{BL} \approx \overline{I}_{CL} \approx \overline{I}_{CBO}$ @ $100°C = 25\mu A$
Step 11 Calculate R_{C2}, R_K and R_B from the matrix equations: $\frac{\underline{V}_{CC} - \underline{V}_1}{(1+n_R)R_{C2}} - \frac{\underline{V}_1}{(1-n_R)R_K} = \overline{I}_1 + \overline{I}_{CL2}$ $\frac{\underline{V}_{BB} - \underline{V}_{OB1}}{(1+n_R)R_B} - \frac{\underline{V}_{OB1} + \overline{SV}_{CE2}}{(1-n_R)R_K} = \overline{I}_{BL1}$ $\frac{-\underline{V}_{BB} + \overline{V}_{BE1}}{(1+n_R)R_B} - \frac{\lambda \overline{V}_{CC}}{\overline{\beta}_{F1}\overline{\beta}_{F2}(1-n_R)R_{C2}} +$ $\frac{\underline{V}_1 - \overline{V}_{BE1}}{(1+n_R)R_K} = \frac{\lambda \overline{I}_O}{\overline{\beta}_{F1}\,\overline{\beta}_{F2}}$	11. To minimize the delay of time of Q_1, let $V_{OB} = 0$ $\frac{23.8 - 5.4}{(1.05)R_{C2}} - \frac{5.4}{(.95)R_K} = (10 + .025)$ $\frac{5.93}{(1.05)R_B} - \frac{0.3}{(.95)R_K} = .025$ $\frac{-(5.93 + .95)}{(1.05)R_B} - \frac{(21.2)(24.2)}{(20)(20)(.95)R_{C2}} +$ $\frac{5.4 - .95}{(1.05)R_K} = 0$

	From which $R_{C2} = 1.54K$ $R_K = 4.3K$ $\underline{R}_K = (1 - .05)(4.3) = 4.09K$ $R_B = 57.4K$ $\overline{R}_B = (1 + .05)(57.4) = 60.3K$
Step 12 Check matrix solution by calculating $\underline{V}_{OB}$ from equation 7-4-2. $\underline{V}_{OB} = \dfrac{\underline{V}_{BB}\underline{R}_K - \overline{SV}_{CE2}\overline{R}_B - \overline{I}_{BL}\underline{R}_K\overline{R}_B}{\underline{R}_K + \overline{R}_B}$	12. $\underline{V}_{OB} = \dfrac{(5.93)(4.09) - (0.3)(60.3) - (.025)(4.09)(60.3)}{(4.09 + 60.3)} = 0$ It was assumed $\underline{V}_{OB} = 0$; therefore the matrix checks.
Step 13 Select standard values for R_{C2}, R_K, R_B, and calculate worst case limit values.	13. Standard 1% resistors are: (Tolerance to be used is end of life limit ± 5%) $R_{C2} = 1.54K$ $\underline{R}_{C2} = 1.46K$, $\overline{R}_{C2} = 1.62K$ $R_K = 4.32K$ $\underline{R}_K = 4.1K$, $\overline{R}_K = 4.55K$ $R_B = 57.6K$ $\underline{R}_B = 54.8K$, $\overline{R}_B = 60.5K$

Step	Calculation
Step 14 Calculate R_D from equation 7-4-10b which describes the on condition of Q_2. $$R_D = \frac{(\underline{V}_{CC} - \overline{V}_{BE2})\overline{\beta}_{F2}\underline{R}_{C2}}{\underline{V}_{CC}(1 + n_{RD})\left(1 + \frac{\overline{I}_0 \underline{R}_{C2}}{\underline{V}_{CC}}\right)}$$	14. $$R_D = \frac{(23.8 - .95)}{(23.8)(1.05)}(20)(1.46)(.95)$$ $$R_D = 25.4\ K\Omega$$
Step 15 Select a standard resistor for R_D and calculate $\overline{R}_D$ and $\underline{R}_D$.	15. A standard 1% resistor for R_D is 25.5K $\therefore\ \underline{R}_D = 24.2K$ $\overline{R}_D = 26.8K$
Step 16 From equation 7-4-21, the timing equation, calculate $\underline{C}_D$: $$\underline{C}_D = \frac{T_D/\underline{R}_D}{\ln\left(\frac{\overline{V}_{CC} + \overline{I}_{BL}\underline{R}_D + \underline{V}_S - \overline{SV}_{CE1} - \overline{V}_{BE2}}{\overline{V}_{CC} - \underline{V}_{TF} + \overline{I}_{BL}\underline{R}_D}\right)}$$ Since $\overline{SV}_{CE}$ and $\overline{V}_{BE}$ do not occur at the same temperature $\underline{C}_D$ should be calculated at both temperature limits.	16. The worst case temperature is found to be 100°C. At this temperature $\overline{V}_{BE2} = 0.72V$ and $\underline{V}_{TF} = 0.3V$ Thus, $$\underline{C}_D = \frac{450/24.2}{\ln\left(\frac{24.2 + (.025)(24.2) + 5.1 - .3 - .72}{24.2 - .3 + (.025)(24.2)}\right)}$$ $\underline{C}_D = 113\ pF$
Step 17 Calculate $\underline{I}_{B1}$ from equation 7-4-3. $$\underline{I}_{B1} = \frac{\underline{V}_1 - \overline{V}_{BE1}}{\overline{R}_K} - \frac{\overline{V}_{BB} + \overline{V}_{BE1}}{\underline{R}_B}$$	17. $$\underline{I}_{B1} = \frac{6 - .95}{4.55} - \frac{6.06 + .95}{54.8} = 0.98\ mA$$

Step 18 Calculate $\underline{R}_T$ from equation 7-4-9b. $\underline{R}_T = \dfrac{\bar{V}_S}{\beta_{F1}\underline{I}_{B1} - \dfrac{\bar{V}_{CC} + \bar{V}_S}{\underline{R}_D}}$	18. $\underline{R}_T = \dfrac{5.4}{(20)\,(0.98) - \dfrac{24.2 + 5.4}{24.2}} = 0.294\ \text{K}$
Step 19 Calculate R_S and R_P from equations 7-4-24 and 7-4-25: $\bar{V}_S = \dfrac{\bar{V}_{CC}(1 + n_R)R_S}{(1 + n_R)R_S + (1 - n_R)R_P}$ $\underline{R}_T = \dfrac{(1 - n_R)R_P(1 - n_R)R_S}{(1 - n_R)R_P + (1 - n_R)R_S}$	19. $5.4 = \dfrac{(24.2)(1 + .05)R_S}{(1 + .05)R_S + (1 - .05)R_P}$ $0.294 = \dfrac{(1 - .05)R_P(1 - .05)R_S}{(1 - .05)R_P + (1 - .05)R_S}$ Simultaneous solution yields: $R_P = 1.53\ \text{K}$ $R_S = 0.397\ \text{K}$
Step 20 Select standard resistors for R_S and R_P and calculate $\bar{R}_T$ and $\underline{R}_T$. $\bar{R}_T = \dfrac{\bar{R}_P\bar{R}_S}{\bar{R}_P + \bar{R}_S}$ $\underline{R}_T = \dfrac{\underline{R}_P\underline{R}_S}{\underline{R}_P + \underline{R}_S}$	20. Standard 1% resistors for R_P and R_S are $R_S = 402\ \Omega$ $\underline{R}_S = 382\ \Omega$, $\bar{R}_S = 422\ \Omega$ $R_P = 1.54\ \text{K}$ $\underline{R}_P = 1.46\ \text{K}$, $\bar{R}_P = 1.62\ \text{K}$ $\therefore\ \underline{R}_T = \dfrac{(1.46)\,(0.382)}{1.46 + 0.382} = 0.303\ \text{K}$ $\bar{R}_T = \dfrac{(1.62)\,(0.422)}{1.62 + 0.422} = 0.335\ \text{K}$

Step	Calculation
Step 21 Calculate $\overline{C}_D$ from equation 7-4-8. $\overline{C}_D = \dfrac{1/\overline{f}_1}{\eta \overline{R}_T + \overline{R}_D \ln \overline{x}}$ where $\overline{x} = \dfrac{\underline{V}_{CC} + \overline{V}_S - \underline{SV}_{CE1} - \overline{V}_{BE2}}{\underline{V}_{CC} - \overline{V}_{BE2}}$	21. $\underline{SV}_{CE} = 0.1V$ $\overline{x} = \dfrac{23.8 + 5.4 - 0.1 - 0.95}{23.8 - 0.95} = 1.23$ $\overline{C}_D = \dfrac{1000}{(4)(0.335) + (26.8) \ln 1.23}$ $\overline{C}_D = 132 \text{ pF}$
Step 22 Select C_D nominal and its tolerance so that the limits of $\overline{C}_D$ and $\underline{C}_D$ are not exceeded.	22. $\underline{C}_D$ from step 16 was 113 pF. Choose C_D nominal = 130 pF 5% tolerance mica capacitor.
Step 23 Apply checking equations for $\overline{I}_{C1}$ and $\overline{I}_{C2}$ and compare to estimates made in steps 6 and 7. $\overline{I}_{C1} = \dfrac{\overline{V}_S}{\underline{R}_T} + \dfrac{\overline{V}_{CC} + \overline{V}_S}{\underline{R}_D}$ (7-4-9a) $\overline{I}_{C2} = \dfrac{\overline{V}_{CC}}{\underline{R}_{C2}} + \overline{I}_o$	23 $\overline{I}_{C1} = \dfrac{5.4}{0.303} + \dfrac{24.2 + 5.4}{24.2} = 19 \text{ mA}$ $\overline{I}_{C2} = \dfrac{24.2}{1.46} = 16.6 \text{ mA}$ Both estimates are close enough to the calculated values so a repeat of calculations is not necessary since β_o has not changed. The error in estimating collector current is almost entirely due to the estimate of $I_K = 2I_{B1}$.

CHAPTER 8

Current Mode Circuits

High-speed operation, excellent dc stability, good noise immunity, and circuit performance that is independent of many transistor characteristics are among the advantages offered by current-mode switches.

Current-mode operation is unsurpassed when used in any high speed circuit where regeneration is important, such as in monostable or astable multivibrators. Logic circuits present some difficulties because level transformation is required between stages. However, exceptional performance has been obtained.

Current Mode Switching has not enjoyed the wide popularity of saturated mode operation. The primary reason for this is that the major use of transistors is in the digital computer field and the speeds encountered could be handled easily by the conventional diode logic used with saturated mode circuits. However, with the emphasis of today's computer upon speed and miniaturization, coupled with the amenability of current mode circuits to adapt to integrated circuit techniques,[1] new interest is being generated in current mode logic circuits.

The idea of using current mode switching in computer logic was first proposed by Yourke[2]. At that time, the only transistors with a fairly high ω_T were of the "drift" type. These devices were designed for R.F. amplifier applications and had poor saturation and storage time characteristics. Furthermore, because they were designed to yield A.G.C. action by varying the collector voltage, ω_T decreased rapidly with decreasing voltage. This required the use of a high collector voltage, which resulted in high dissipation, in order to have a fast switching circuit.

A computer[3] was built using this approach, but it was not particularly successful. A primary reason for its abandonment was the fact that the transistor and circuit dissipation were too high. Today, the availability of mesa transistors which maintain a high ω_T at low collector voltages makes another look at current mode logic worthwhile.[4,5,6,7]

As with the saturated mode circuits, the basic amplifier or inverter will be discussed in detail, because it is the heart of any circuit. Then, the flip-flop and pulse generating circuits are discussed. A worst-case design procedure is given for an inverter.

Since Motorola 2N2256 through 2N2259 transistors are low cost units specified for current mode operation, they are used in the design examples. However, other type numbers can be used. Very fast switching can be obtained by using mesa R.F. transistors which have low r'_B and high ω_T at low collector voltages.

SECTION 8-1 — FUNDAMENTALS OF CURRENT MODE CIRCUITS

8-1-1 — DC Analysis of the Current Mode Inverter

A qualitative description of the operation of the current mode inverter was given in Chapter 2. Four significant worst-case conditions affecting circuit design are shown in Figure 8-1-1. In part a, the minimum input level to just turn off Q_1 and allow Q_2 to be on is shown. From the figure this condition is

$$\underline{E}_O = \overline{V}_{BE} - \underline{V}_T \tag{8-1-1a}$$

V_T is defined as a forward base-emitter voltage at the threshold of conduction where I_C is a negligible value. V_T is slightly less than the forward threshold voltage V_{TF} as defined and used in Chapters 5 and 7.

Observe also the case where Q_1 is on and Q_2 is off shown in part c. The condition for this case is

$$\underline{E}_1 = \overline{V}_{BE} - \underline{V}_T. \tag{8-1-1b}$$

The polarity of E_0 is opposite to that of E_1, as indicated by the sign convention on the figure. Establishment of both dc states depends upon a minimum signal level which varies (+) and (—) about the reference level (ground in the circuit shown) by an amount sufficient to overcome the base-emitter voltage of the conducting transistor.

The respective on emitter currents of Q_1 and Q_2 are:

$$I_{E1} = \frac{V_{EE} - V_{BE} + E_1}{R_E} \tag{8-1-2}$$

and

$$I_{E2} = \frac{V_{EE} - V_{BE}}{R_E} \tag{8-1-3}$$

The voltages used in equations 8-1-2 and 8-1-3 represent absolute magnitudes only and the expressions are valid for either PNP or NPN transistors.

To keep I_{E1} and I_{E2} essentially equal and reduce the effect of the level of E_1 upon the magnitude of I_{E1}, it can be seen that V_{EE} must be made substantially larger than changes in E_1. Thus, the common emitter supply approaches a constant current source.

It should be noted that V_T is normally a slight forward voltage in current mode circuits. Since the on collector current is large, a few hundred μA of leakage in the off condition is not detrimental to circuit performance. Measurements indicate that the term $(\overline{V}_{BE} - \underline{V}_T)$ is approximately 0.4 volt regardless of the transistor type used. The term is approximately constant with temperature also, since the temperature coefficient does not differ much between the normal collector on current and the current which flows with V_T applied.

It should also be evident that, relatively speaking, the driving source impedance should be low and the transistor current gain high. Otherwise, the base current which flows will reduce the voltage at the base below that of the source, thereby, altering the emitter current. The collector on current is αI_E, which reduces the output level from V_{CC}, when the transistor is off, to $V_{CC} - \alpha I_E R_L$ when a transistor is on. For any reasonable β, α is close to unity, therefore,current gain has only a minor effect upon circuit operation when the source impedance is low.

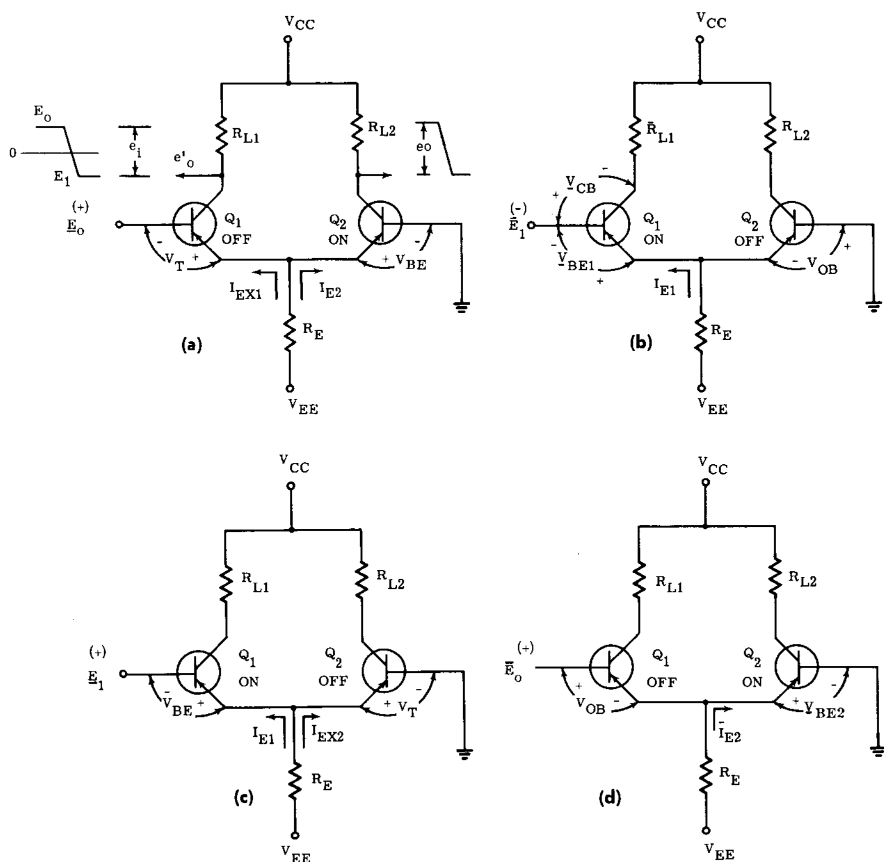


Figure 8-1-1 — Significant Worst Case Conditions
(a) Q_1 Just Cut-Off
(b) Maximum V_{OB} On Q_2
(c) Maximum V_{OB} On Q_1
(d) Q_2 Just Cut Off

To avoid saturation, it is necessary to observe the inequality

$$\underline{V}_{CB1} > \bar{E}_1 \qquad (8\text{-}1\text{-}4)$$

This worst-case condition is indicated on part b of Figure 8-1-1. Actually, since collector capacitance is inversely proportional to some power of the voltage, switching speed is enhanced if V_{CB} is greater than zero by a volt or two; curves of f_τ and C_{Tc} versus voltage will provide guidance in making a choice for $\underline{V}_{CB}$.

The output voltage e_0, obtained from the collector of Q_2, is in phase with the input signal, while e'_0, the output voltage of Q_1, is 180° out of phase or, in logic language, the complement of e_0. In cases where only e'_0 is required, Q_2 can be replaced by a diode. The ready availability of complementary signals, however, often simplifies logic problems.

When ac coupling cannot be employed, problems arise because the output is at a different dc reference level than the input making direct coupling of circuits in the form of Figure 8-1-1 impractical. By using current sources in the output, however, alternate PNP and NPN gates can easily be coupled. Another coupling method uses zener diodes to establish the proper dc levels which permits the use of transistors of the same polarity. These techniques will be discussed later.

8-1-2 — Advantages of Current Mode Circuits

HIGH DC STABILITY: It is easy to show[8] that, under the worst-case condition, i.e., where $h_{FE} \rightarrow \infty$, the current stability factor, for changes in I_{BL} of an on transistor is,

$$S_I = \frac{\partial I_E}{\partial I_{BL}} = \frac{R_B}{R_E}$$

where

R_E is the total dc resistance in the emitter circuit, and

R_B is the total dc resistance in the base circuit.

In the current mode switch, R_E is several times R_B with the result that emitter current change due to I_{BL} is negligible. Furthermore in low level transistors, I_{BL} is so low that dc stability need not be considered in the design.

HIGH SPEED OPERATION: The lack of storage time is the principal reason for improved high-speed operation. However, current mode operation also provides improvements in the other switching time intervals. Delay time is very short because the low impedance drive circuit quickly charges the input capacitance. Moreover, as the transistor turns on, this low impedance drive also aids rise time. Since the source impedance is low, an input signal just slightly greater than that given by equation 8-1-1 provides sufficient overdrive to achieve high speed operation.

Fall time is similarly improved, and the low voltage swing also enhances speed.

NON-CRITICAL TRANSISTOR REQUIREMENTS: Transistor requirements for current-mode switching are less demanding than those of a saturated switch.

The dc characteristics of transistors in current mode logic circuits are not critical. The excellent stability of the circuit virtually negates the normally important I_{BL} characteristic. For example, an I_{BL} value of 100 μamps across a 100-

ohm base resistor produces only 10 millivolts — a value which is negligible. Emitter breakdown normally presents no problem. Under the worst-case conditions,

$$\overline{V}_{OB} = \overline{E}_1 - \underline{V}_{BE1} \quad (8\text{-}1\text{-}5)$$

or

$$\overline{V}_{OB} = \overline{E}_O - \underline{V}_{BE2}$$

as can be seen from Figure 8-1-1b and d V_{OB} is ordinarily less than a volt.

Current gain assumes some importance in considering "fan-out" and this will be discussed in connection with gate circuits.

Since the collector-emitter saturation voltage, SV_{CE}, does not set a logic level in current mode switching, its value is not critical but the knee characteristic is of importance. In view of the high current levels usually employed, it is important to be able to work at low collector voltage to keep dissipation low. The values of ω_T and C_{ob} at low collector voltages may cause a reduction in speed. This factor coupled with the value of V_{CE} at the edge of saturation must be considered when setting the collector voltage in the on condition.

Of the transient characteristics, storage time is of no interest, however, f_T, C_{ob}, and C_{ib} carry their usual significance. Also, since the source impedance is low, the base spreading resistance, r'_B, becomes important.

Fortunately, the inherent characteristics of mesa transistors, such as high power ratings and a high gain-bandwidth product at low V_{CE} and high I_C, make these units ideally suited for current mode circuitry.

CONSTANT POWER SUPPLY LOADING: Regardless of which transistor is conducting, both the current from V_{EE}, flowing into an emitter, and the current αI_E, flowing from the collector to V_{CC}, are equal. During the switching interval, the current from V_{EE} divides between Q_1 and Q_2, but still results in the same amount of total current (αI_E) from V_{CC}.

As a result of this constant power supply loading, the problems of power supply design are reduced and generation of undesirable transients, which could cause spurious triggering in other parts of the system, are eliminated.

EXCELLENT NOISE IMMUNITY: The current mode system is immune to noise in many ways. Since the load impedance is low, coaxial cables can easily be used to couple from collector to load when long runs are needed. Since the output signal is in the form of a current from a high impedance source, resistance in the cable does not cause a loss of signal at the load. A low impedance coupling system, of course, is very resistant to stray signal pickup.

The circuits are similar to a differential amplifier. With the bases of both Q_1 and Q_2 referenced to the same potential, a change in this potential, or that of the common emitter supply, appears in the collector circuit, but is attenuated in value. (The common mode voltage gain is R_L/R_E which is much less than unity.) Inspection of the gate circuits to be described will reveal that noise on any power supply is attenuated when passing from supply to signal line.

8-1-3 — Transient Response

The transient response to a step function can be found in a manner analogous to that of a grounded emitter stage.

In Chapter 5, it was shown, that, for a saturated mode switch driven by a step of base current, the rise and fall times are given by

$$t_r = \frac{I_C \tau_A}{I_{B1} - I_C/2\beta_o} \quad \text{and} \quad t_f = \frac{I_C \tau_A}{I_{B2} + I_C/2\beta_o}$$

In the current mode switch, the rise time of one transistor occurs during the fall time of the other. Therefore, recombination effects, represented by the $I_C/2\beta_o$ terms tend to cancel and can be neglected in the analysis. Neglecting recombination is further justified because switching is normally fast.

Since the collector voltage does not enter the saturation region, τ_A is lower for a current mode switch than for a saturated mode switch operating at the same value of I_C. This arises because the large increase of C_{Tc} and the drop of ω_T as the saturation knee is approached does not occur in current mode switching.

RISE TIME EQUATIONS: The equations derived for saturated mode switching, which assumed that the transistor is driven from a constant current source, must be re-evaluated for current mode switching. Figure 8-1-2a shows conditions at the beginning of the rise time interval while Part b of the figure shows conditions at the end of the rise time interval. With the conditions shown, the input voltage is at a level where the off transistor is biased at the edge of forward injection. The voltages e_i and v_B are indicated on Figure 8-1-3 as a function of time. As Q_1 goes on and Q_2 goes off, v_B is shown changing linearly from V_{B1} to V_{B2}. In reality v_B changes nearly linearly, even though v_{BE} is exponentially related to i_E, because v_B is the difference between v_{BE1} and v_{BE2}.

The shaded area is related to the amount of charge delivered to the transistor during the rise time interval. Since $Q = idt = \frac{v}{R}dt$, the shaded area must equal the product of RQ_{in}. If the transistors used are identical i.e., $V_{BE2} - V_{TF1} = V_{BE2} - V_{TF2}$, then the small triangular areas above and below the center line are equal and the total input charge is given simply as

$$Q_{in} = t_r E_1/R.$$

At the end of the rise time interval, the input charge Q_{in} equals the transistor active charge Q_A. The term R is the total series resistance which includes that of the source as well as r'_B of both transistors, therefore

$$t_r = \frac{Q_A}{E_1/(R_S + 2r'_B)}. \tag{8-1-6}$$

Equation 8-1-6 is exactly the same as for a saturated mode circuit if V_{BE} is assumed zero.

In Chapter 5 the charge from zero to the 90% point was shown to be

$$Q_A = 0.9\left(\frac{1}{\omega_T} + R_L C_f\right) I_C$$

In current mode switching, the charge moved between the 10% and 90% points is of more interest, therefore, the factor, 0.9, should be replaced by 0.8. Since the voltage swings are small, C_f can be considered equal to C_{ob} and ω_T can be assumed independent of voltage. Therefore, the final form of the rise time equation is

$$t_r = \frac{0.8\,(1/\omega_T + R_L C_{ob})\, I_C}{E_1/(R_S + 2r'_B)} \tag{8-1-7}$$

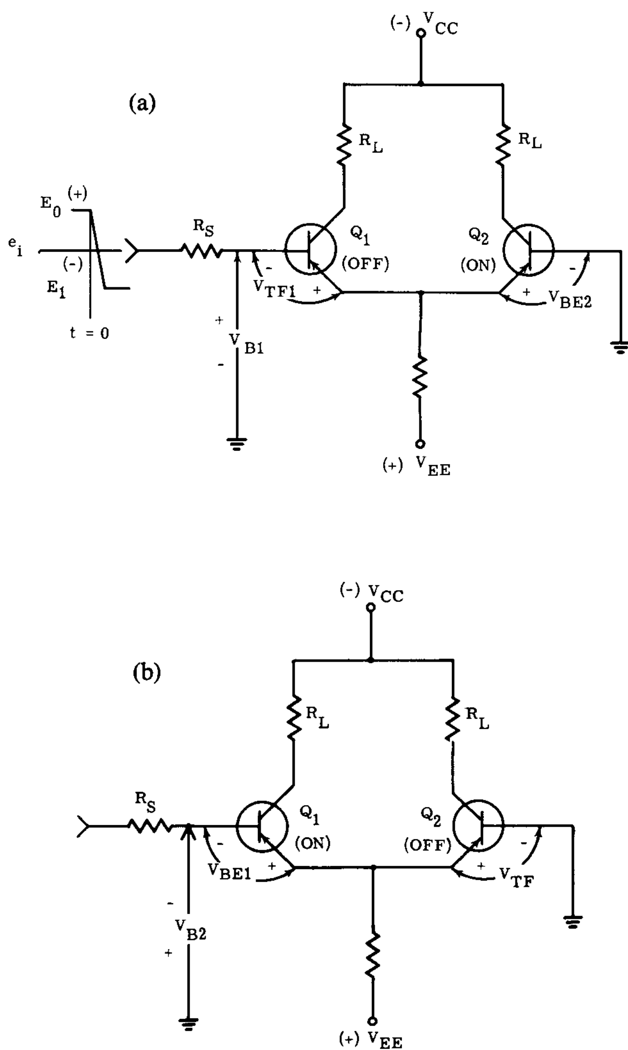


Figure 8-1-2 —
(a) Conditions at the threshold of conduction for Q_1
(b) Conditions at the threshold of conduction for Q_2

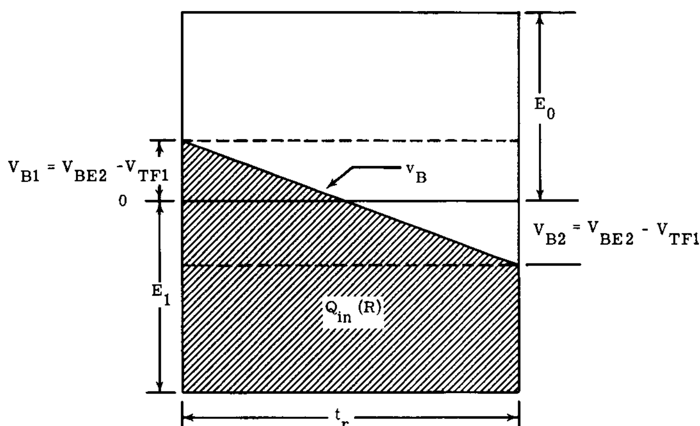


Figure 8-1-3 — Charge Diagram to Determine Input Charge

OPTIMUM SOURCE RESISTANCE: It is possible to minimize t_r by choosing an optimum value for R_S. R_S will normally equal R_L, but in cases where a larger voltage swing is desired at the output than at the input, R_L will bear some ratio to R_S. If the collector current of a preceding stage is equal to the collector current of the stage under consideration, then,

$$(E_1 + E_o)_{in} = R_S I_C$$
$$(E_1 + E_o)_{out} = R_L I_C.$$

The voltage gain therefore, is

$$A_V = \frac{R_L}{R_S} \tag{8-1-8}$$

Normally $E_o = E_1$. Therefore, the rise time may be written in terms of the input and output voltages as

$$t_r = 0.8\,(1/\omega_T + R_L C_{ob}) \left(\frac{R_S + 2r'_B}{R_L}\right)\left(\frac{2\,E_{1\,(out)}}{E_{1\,(in)}}\right) \tag{8-1-9}$$

Substituting 8-1-8 into 8-1-9, simplifying the result, taking the derivative and setting it equal to zero, the optimum value for R_S may be found

$$R_{S\,(opt)} = \sqrt{\frac{2r'_B}{A_V\ C_{ob}\ \omega_T}} \tag{8-1-10}$$

This result indicates that as the voltage gain (A_V) increases the magnitude of R_S should decrease for maximum speed. A physical explanation for this behavior is that as A_V increases, the input capacity is increased due to Miller Effect. However, the Miller Effect can be offset somewhat by lowering the source impedance.

A numerical example will put the results of the preceding discussion into

meaningful terms. Typical values for the significant characteristics of a 2N2259 transistor are:

$$r'_B = 75\ \Omega$$
$$C_{ob} = 4\ pF$$
$$f_T = 300\ mcs.$$

Inserting these values in equation 8-1-10, taking $R_S = R_L$, and solving for $R_{S(opt)}$, it is found that it equals 126 Ω.

Circuit voltages typical of current mode circuits are $E_{1\,(in)} = E_{1\,(out)} = 1$ volt. Taking $R_S = R_L = 126$ ohms, substituting the required values into equation 8-1-9 it is found that $t_r = 3.6$ nS. This value is in close agreement with experimentally measured data.

8-1-4 — DC Coupling Techniques

Two basic coupling methods are employed in current mode circuits. One uses alternate PNP and NPN blocks (See Figure 8-1-4) in which there are two different "1" and "0" levels; the other uses zener diodes (See Figure 8-1-5) and current sources in the collector circuit so that the output of a block is always at the proper level to be coupled to a similar block. The basic action of these coupling methods was discussed in Chapter 2.

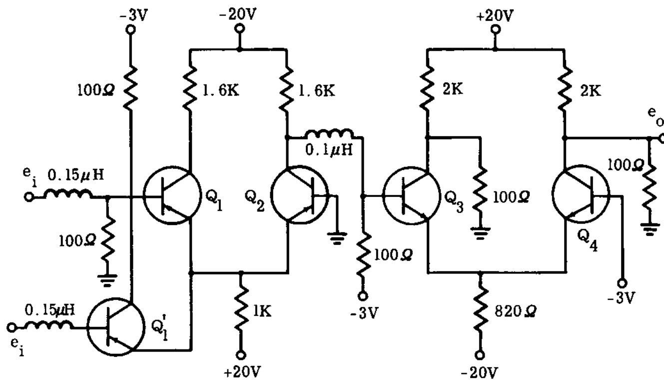


Figure 8-1-4 — Alternate NPN-PNP Coupling

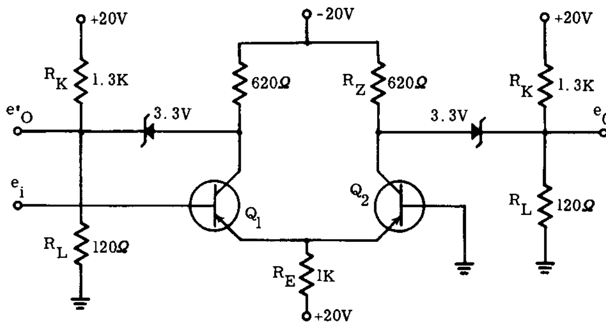


Figure 8-1-5 — Zener Diode Coupling

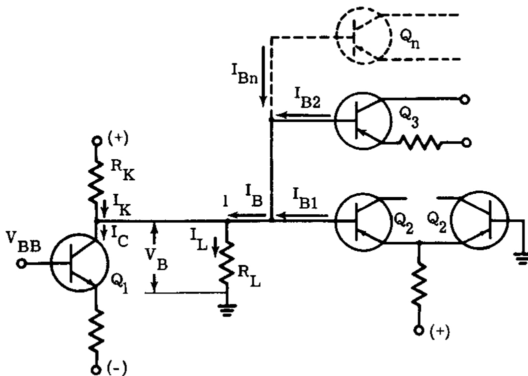


Figure 8-1-6 — Details of Complementary Transistor Coupling

COUPLING USING COMPLEMENTARY TYPES: When coupling from an NPN stage with its base referenced to V_{BB}, to a PNP stage having its base referenced to ground, as in Figure 8-1-6, the current through the load can be found by summing currents at node 1.

$$I_L = I_B + I_K - I_C \quad (8\text{-}1\text{-}11)$$

where I_B = the total base current from n number of stages. From this, the voltage at node 1 is V_{BO} when all the transistors are off.

$$V_{BO} = R_L I_L = R_L I_K \quad (8\text{-}1\text{-}12)$$

where V_{BO} is a positive voltage sufficient to satisfy the criterion given by equation 8-1-1.

When the transistors are on, the voltage at node 1 is

$$V_{B1} = I_L R_L = R_L (I_B + I_K - I_C) \quad (8\text{-}1\text{-}13)$$

and I_C must be larger than $I_K + I_B$ by a sufficient amount such that V_{B1} is negative and is larger than the value given by equation 8-1-1. Therefore, the voltage swing is

$$e_o = (I_C - I_B) R_L \quad (8\text{-}1\text{-}14)$$

The current I_K essentially represents waste current since it doesn't drive anything. Examining equation 8-1-11 it becomes evident that, in a sense, I_B is also waste current. Therefore, transistor current gain should be as high as possible.

Low gain transistors can be accommodated by increasing R_L to provide more voltage swing and reducing I_K to just satisfy equation 8-1-12. However, the voltage at the collector of Q_1 should never be allowed to drop below V_{BB} in order to prevent saturation. To increase "fan-out", an emitter-follower can be inserted in the line between nodes 1 and 2 as shown in Figure 8-1-7. It will be necessary to shift the level at node 1 to compensate for the drop through the base-emitter junction of Q_1. R_E insures that Q_1 is conducting at all times thereby providing the proper off-bias when the gate transistors are off. This system will provide a more universal coupling block.

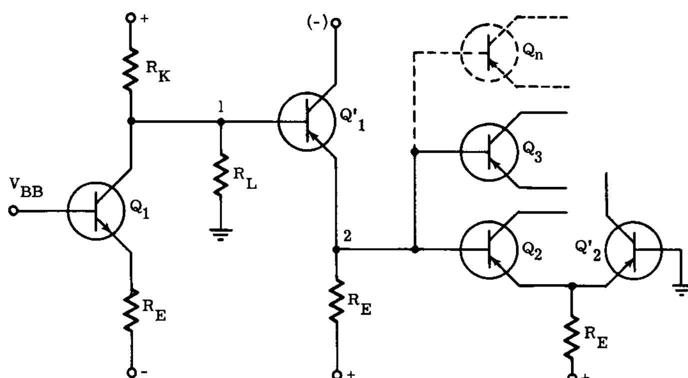


Figure 8-1-7 — Use of Emitter-Follower to Increase Fan-out

ZENER DIODE COUPLING: Using a similar procedure, the coupling system using zener diodes can be analyzed from the notations shown on Figure 8-1-8. It is found that

$$V_B = R_L\,(I_C + I_K + I_B - I_Z). \tag{8-1-15}$$

For the two distinct conditions

$$Q_2 \text{ off, } Q_1 \text{ on: } V_{BO} = R_L\,(I_C + I_K - I_Z) \tag{8-1-16}$$

$$Q_1 \text{ off, } Q_2 \text{ on: } V_{B1} = R_L\,(I_K + I_B - I_Z). \tag{8-1-17}$$

From Figure 8-1-8, note that, in order to keep the zener diode conducting, the condition

$$I_Z > I_C \tag{8-1-18}$$

must be fulfilled. A minimum of 2 to 5 mA is usually required for low voltage zener diodes. Therefore, to establish the off bias the current I_K is injected so that V_{BO} given by equation 8-1-16 is greater than $(\overline{V}_{BE} - \underline{V}_T)$ for the transistor. Examining equations 8-1-17, observe that I_Z provides the drive while I_K and I_B again represent waste current. Current I_C provides the voltage swing across R_L which is represented by

$$e_o = R_L(I_C - I_B) \tag{8-1-19}$$

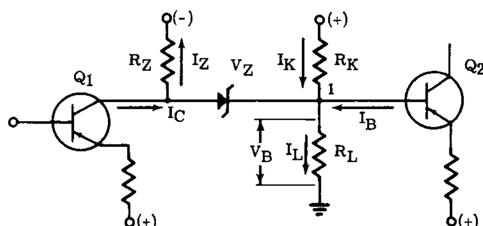


Figure 8-1-8 — Details of Zener Diode Coupling

To accommodate high "fan-out", an emitter follower amplifier, as previously discussed, may be used with the zener coupling method. This is the approach used by Buelow,[3] and basically the same approach used in MECL* integrated logic blocks.

IMPROVING CIRCUIT EFFICIENCY: With either coupling method, I_K represented waste current. By making R_L larger for the off bias condition, I_K can be reduced. Since the direction of I_L through R_L changes from one logic level to the other, it is a simple matter to make R_L vary with current as shown in Figure 8-1-9. By means of diodes, the proper load resistor is connected in the circuit as current is reversed. Thus, I_K can be greatly reduced. Generally, R'_L can be eliminated if the forward voltage drop of D_2 produces sufficient off bias. With I_K reduced, I_C can also be reduced, thus reducing power dissipation in the drive transistor and its emitter current source.

The load resistor R_L can also be eliminated if D_1 can develop the proper signal level. In this case, D_1 serves as a clamp or a current sink. "Fan-out" can easily be accomplished, because as additional transistor bases are added in parallel, their current requirement is borrowed from the diode. The waste of current through the load resistor has been eliminated and "fan-out" approaches the current gain of the transistor type used.

Against these improvements in circuit efficiency and "fan-out" capability must be weighed the added circuit complexity and recovery time problems in the diodes.

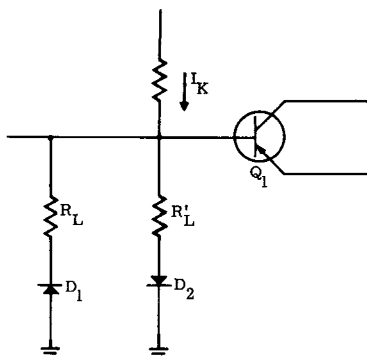


Figure 8-1-9 — Using Diodes to Increase Circuit Efficiency

REFERENCES:

1. J. A. Marvel, W. C. Seelbach and N. Miller, "Current Mode Logic with Integrated Circuitry", Electronic Industries, Nov.-Dec., 1962.
2. Hannon S. Yourke, "Millimicrosecond Transsistor Current Switching Circuits", IRE Transactions on Circuit Theory, September 1957, Vol. CT-4, No. 3, pp. 236-240.
3. Block, "Engineering Design of the Stretch Computer", Proceeding of the Eastern Joint Computer Conference, Dec. 1-3, 1959.
4. F. K. Buelow, "Improvements to Current Switching", 1960 Solid States Circuits Conference Digest of Technical Papers, pp. 30-31.
5. Carl M. Campbell, Jr., "New Configurations in Non-Saturating Complementary Current Switching Circuits", Semiconductor Products, July 1959, pp. 15-20.
6. William D. Roehr, "Techniques of Current Mode Logic Switching", Electronics Design, Sept. 13, 1962, pp. 54-59.
7. Dean C. Bailey, "Another Look at Current Mode Systems", Electronic Design, Nov. 22, 1962, pp. 48-55.
8. William D. Roehr, "Development of Gain-Stability Design Charts", Electronic Equipment Engineering, Oct. 1960.

*Registered trademark, Motorola, Inc.

SECTION 8-2 – CURRENT MODE INVERTER

Clamped outputs while adding circuit complexity, increase efficiency and stabilize the output levels and are, therefore, commonly used. A clamped inverter showing the dc on-off states is shown in Figure 8-2-1.

The use of clamped outputs complicates circuit design because a clamp supply voltage must be determined and current must be provided for the diode when the load current is maximum. Inverter design remains basically separate from the output clamp design once the clamp conditions are determined.

In the on condition, the three resistors (R_Z, R_E and R_K) must guarantee a minimum available load current and also must maintain the zener diode in conduction at all times. In the off condition, R_K and R_Z must be chosen so that a minimum available load current is assured.

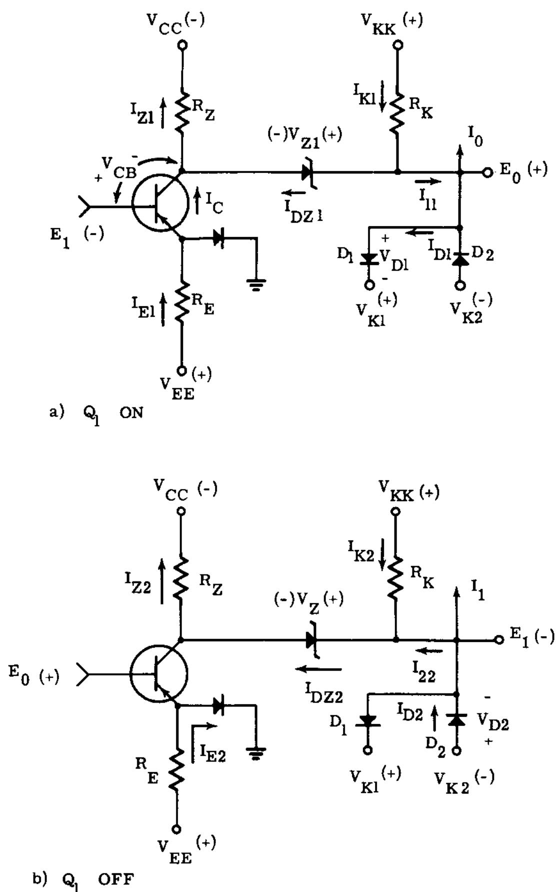


Figure 8-2-1 — DC States for Inverter Driving a Clamped Load

8-2-1 — Synthesis

The inverter must be designed to supply the required output currents and maintain the output levels within specified limits. If variable loads are encountered, clamp diodes must be used to limit shifts in the output levels.

From Figure 8-2-1, it can be seen that the minimum output levels occur when the diode drops are minimum. Thus, the minimum clamp level needed to meet circuit requirements is easily determined from:

$$\underline{V}_{K1} = \underline{E}_0 - \underline{V}_{D1} \quad (8\text{-}2\text{-}1)$$

$$\underline{V}_{K2} = \underline{E}_1 - \underline{V}_{D2} \quad (8\text{-}2\text{-}2)$$

A minimum diode current must be chosen to keep the diode in conduction, depending upon the diode characteristics. This current is then added to the required output current to set the requirement for the inverter current. Or:

$$\overline{I}_{11} = \overline{I}_0 + \underline{I}_{D1} \quad (8\text{-}2\text{-}3)$$

$$\overline{I}_{22} = \overline{I}_1 + \underline{I}_{D2} \quad (8\text{-}2\text{-}4)$$

The maximum $E_{0(out)}$ occurs when both V_{K1} and V_{D1} are maximum.

$$\overline{E}_{0(out)} = \overline{V}_{K1} + \overline{V}_{D1} \quad (8\text{-}2\text{-}5a)$$

$$\overline{E}_{1(out)} = \overline{V}_{K2} + \overline{V}_{D2} \quad (8\text{-}2\text{-}5b)$$

Where V_{D1} is determined at the maximum diode current of $(\overline{I}_0 - \underline{I}_0 + \underline{I}_{D1})$, and V_{D2} is determined at $(\overline{I}_1 - \underline{I}_1 + \underline{I}_{D2})$.

Considering the diodes as part of the load, the requirements for the inverter are:

$\overline{I}_{11}$ must be available when $e_0 = \underline{E}_0$

$\overline{I}_{22}$ must be available when $e_0 = \underline{E}_1$.

The zener provides the level shift from input to output. To avoid saturation $\underline{V}_{CB}$ should be set to insure a reverse biased collector junction. Quite simply then, the zener voltage is the difference between the input and output levels, plus a voltage to insure the proper $\underline{V}_{CB}$.

$$\underline{V}_Z = \overline{E}_{0(out)} + \overline{E}_{1(in)} + \underline{V}_{CB} \quad (8\text{-}2\text{-}6)$$

For the on condition, the worst case occurs when I_{K1} is supplying a minimum of current and I_{11} is required to be a maximum. If I_C drops too low, I_{DZ1} must increase which reduces the available I_{11}. To guarantee that $\overline{I}_{11}$ is available,

$$\overline{I}_{11} = \underline{I}_{K1} - \overline{I}_{Z1} + \underline{I}_C \quad (8\text{-}2\text{-}7)$$

or writing equation 8-2-7 in the notation of Figure 8-2-1:

$$\overline{I}_{11} = \frac{\underline{V}_{KK} - \overline{E}_{0\,(out)}}{\overline{R}_K} - \frac{\overline{V}_{CC} - \underline{V}_Z + \overline{E}_{0\,(out)}}{\underline{R}_Z} + \frac{\underline{V}_{EE} - \overline{V}_{BE} + \underline{E}_{1\,(in)}}{\overline{R}_E} \quad (8\text{-}2\text{-}8)$$

Also, when the transistor is on, a condition may occur where I_C increases to a maximum and I_{Z1} as determined by V_Z and E_0 is at a minimum. In this case, I_{DZ1} could be too low to keep the zener in conduction. Thus, to guarantee $\underline{I}_{DZ1}$.

$$\underline{I}_{DZ1} = \underline{I}_{Z1} - \overline{I}_C \quad (8\text{-}2\text{-}9)$$

Again, using the notation of Figure (8-2-1a):

$$\underline{I}_{DZ1} = \frac{\underline{V}_{CC} - \overline{V}_Z + \underline{E}_{0\,(out)}}{\overline{R}_Z} - \frac{\overline{V}_{EE} - \underline{V}_{BE} + \overline{E}_{1\,(in)}}{\underline{R}_E} \quad (8\text{-}2\text{-}10)$$

From the off conditions shown in Figure 8-2-1b, a minimum I_{22} must be made available. In this case, all the current goes through the zener and it is always in conduction. The worst-case occurs when I_{Z2} is a minimum and I_{K2} is supplying a maximum current.

$$\overline{I}_{22} = \underline{I}_{Z2} - \overline{I}_{K2} \tag{8-2-11}$$

Using the notation of Figure 8-2-1b:

$$\overline{I}_{22} = \frac{\underline{V}_{CC} - \overline{V}_{Z} - \overline{E}_{1\,(out)}}{\overline{R}_{Z}} - \frac{\overline{V}_{KK} + \overline{E}_{1\,(out)}}{\underline{R}_{K}} \tag{8-2-12}$$

Equations 8-2-8, 8-2-10 and 8-2-12 can be solved simultaneously for the three resistors, R_E, R_Z and R_K by matrix methods. The only transistor characteristic appearing is V_{BE}.

Summarizing the three equations and writing them in matrix form:

$$\begin{bmatrix} K_{11} & K_{12} & -K_{13} \\ K_{21} & -K_{22} & K_{23} \\ -K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{bmatrix} \frac{1}{R_Z} \\ \frac{1}{R_E} \\ \frac{1}{R_K} \end{bmatrix} = \begin{bmatrix} \overline{I}_{22} \\ \underline{I}_{DZ1} \\ \overline{I}_{11} \end{bmatrix} \tag{8-2-13}$$

where the following relationships apply:

$$K_{11} = \frac{1}{1 + n_{RZ}} (\underline{V}_{CC} - \overline{E}_1 - \overline{V}_Z)$$

$$K_{12} = 0$$

$$K_{13} = \frac{1}{1 - n_{RK}} (\overline{V}_{KK} + \underline{E}_1)$$

$$K_{21} = \frac{1}{1 + n_{RZ}} (\underline{V}_{CC} + \underline{E}_O - \overline{V}_Z)$$

$$K_{22} = \frac{1}{1 - n_{RE}} (\overline{V}_{EE} - \underline{V}_{BE} + \overline{E}_{1\,(in)})$$

$$K_{23} = 0$$

$$K_{31} = \frac{1}{1 - n_{RZ}} (\overline{V}_{CC} + \overline{E}_O - \underline{V}_Z)$$

$$K_{32} = \frac{1}{1 + n_{RE}} (\underline{V}_{EE} - \overline{V}_{BE} + \underline{E}_{1\,(in)})$$

$$K_{33} = \frac{1}{1 + n_{RK}} (\underline{V}_{KK} - \overline{E}_O)$$

E_1 and E_O are output levels unless otherwise noted.

To determine a value of $\overline{V}_{BE}$ an estimate of I_C is needed. An intuitive approach will be sufficiently accurate to obtain the estimate for I_C because the value used for $\overline{V}_{BE}$ does not significantly affect the values required for the resistors. Assume that $I_Z \approx I_{Z1} \approx I_{Z2}$ and $I_K \approx I_{K1} \approx I_{K2}$. When Q_1 is on, I_C opposes I_{Z1} and allows a residual zener diode current to flow (I_{DZ}). I_{K1} must supply the load current and the zener current. Therefore:

$$I_C = I_Z - I_{DZ}$$

$$I_K = \overline{I}_{11} + I_{DZ}$$

When Q_2 is off, I_{K2} and I_{22} are supplied by I_Z, that is

$$I_Z = I_K + \overline{I}_{22}$$

Combining these relations it is seen that

$$I_C = \overline{I}_{11} + \overline{I}_{22}$$

From experience with tolerances, it can be assumed that they will multiply the required value for I_C.

Three tolerance factors are involved:

1) The tolerance of the resistors (N_R)
2) The tolerance of the power supplies (N_P)
3) The tolerance of I_K and I_Z due to the shift of the output level from E_1 to E_O (N_V).

Using intuitive reasoning, an estimate of I_C should contain an N_R term for each of the three resistors, an N_P term for each of the three power supplies and an N_V term for both I_Z and I_K. Thus:

$$I_C \approx N_{VZ}\, N_{VK}\, N_P^3\, N_R^3\, (\overline{I}_{11} + \overline{I}_{22}) \tag{8-2-14}$$

Where

$$N_{VZ} = \frac{\overline{V}_{CC} - \underline{V}_Z + \overline{E}_O}{\underline{V}_{CC} - \overline{V}_Z - \overline{E}_1} \qquad N_{VK} = \frac{\overline{V}_{KK} + \overline{E}_O}{\underline{V}_{KK} - \overline{E}_1}$$

$$N_P = \frac{1 + n_P}{1 - n_P} \qquad N_R = \frac{1 + n_R}{1 - n_R}$$

(N_P and N_R were assumed identical for each resistor and power supply in equation 8-2-14)

Using equation 8-2-14 to find I_C, V_{BE} can be found from the data sheet and then equation 8-2-13 can be solved for the resistor values.

The principles involved are illustrated in the following design example. The problem is formulated in much the same manner as for the saturated mode inverter. The design proceeds step-by-step from output to input.

TABLE 8-2-1 — CURRENT MODE INVERTER DESIGN EXAMPLE

Available Input Levels (volts)			Required Output Levels (volts)		
	Min	Max		Min	Max
E_0 (+)	0.8	1.7	E_0 (+)	1.3	1.7
E_1 (−)	0.8	1.7	E_1 (−)	1.3	1.7
LOAD CURRENT REQUIREMENTS(mA)					
				Max	Min
	at Output Level E_0		I_0	5	0
	at Output Level E_1		I_1	8	0
AVAILABLE POWER SUPPLIES (volts)					
Min.		−23.8	+23.8		Clamp supplies to be determined
Nom.		−24.0	+24.0		
Max.		−24.2	+24.2		
AMBIENT TEMPERATURE OF INVERTER ENVIRONMENT					
0°C to 45°C					

8-2-2 — Design Example

Step	Solution
Step 1 State design requirements. Determine clamp supplies and select diodes to obtain minimum levels. $\underline{V}_{K1} = \underline{E}_O - \underline{V}_{D1}$ $\underline{V}_{K2} = \underline{E}_I - \underline{V}_{D2}$ $\underline{V}_D$ must be found at $\underline{I}_D$, a current to put operation over the "knee."	**1** The design requirements are shown in Table 8-2-1. $\underline{I}_D = 2$ mA is suitable $\underline{V}_D = 0.596$V for a 1N3605 diode at 2 mA and 45°C $\underline{V}_{K1} = 1.3 - 0.596 = 0.704$ V $\underline{V}_{K2} = 1.3 - 0.596 = 0.704$ V Thus, for 1% regulation the nominal voltage for V_{K1} and V_{K2} would be 0.712V.
Step 2 Determine the required current into the output network. $\bar{I}_{11} = \bar{I}_0 + \underline{I}_{D1}$ $\bar{I}_{22} = \bar{I}_1 + \underline{I}_{D2}$	**2** $\bar{I}_{11} = 5 + 2 = 7$ mA $\bar{I}_{22} = 8 + 2 = 10$ mA
Step 3 Determine maximum output levels $\bar{E}_O = \bar{V}_{K1} + \bar{V}_{D1}$ $\bar{E}_1 = \bar{V}_{K2} + \bar{V}_{D2}$ $\bar{V}_D$ must be found at $\bar{I}_D$. Use $\bar{I}_{D1} = \bar{I}_O - \underline{I}_O$ $\bar{I}_{D2} = \bar{I}_1 - \underline{I}_1$	**3** $\bar{I}_{D1} = 5 - 0 = 5$ mA $\bar{I}_{D2} = 8 - 0 = 8$ mA $\bar{E}_O = .72 + .83 = 1.55$V $\bar{E}_1 = .72 + .87 = 1.59$V from the diode characteristic, $\bar{V}_{D1} = 0.3$V $\bar{V}_{D2} = 0.87$V

Step	Example
Step 4 Select a zener diode to establish $\underline{V}_{CB}$ $\underline{V}_Z = \overline{E}_{O\ (out)} + \overline{E}_{1\ (in)} + \underline{V}_{CB}$ Select zener with $\underline{V}_Z$ greater than or equal to calculated value.	**4** $\underline{V}_{CB} = 1V$ will be used to avoid saturation and keep reasonable power dissipation and speed. $\underline{V}_Z = 1.55 + 1.7 + 1 = 4.25$ A 4.3V 5% tolerance zener will be used. $\therefore \overline{V}_Z = 4.52V$ and $\underline{V}_Z = 4.18V$
Step 5 Estimate $\overline{I}_C$ $\overline{I}_C = N_{VK}\, N_{VZ}\, N_P^3\, N_R^3\, (\overline{I}_{11} + \overline{I}_{22})$ $N_{VK} = \dfrac{\overline{V}_{KK} + \overline{E}_O}{\underline{V}_{KK} - \overline{E}_1}$ $N_{VZ} = \dfrac{\overline{V}_{CC} - \underline{V}_Z + \overline{E}_O}{\underline{V}_{CC} - \overline{V}_Z - \overline{E}_1}$ $N_P = \dfrac{1 + n_P}{1 - n_P}$ $N_R = \dfrac{1 + n_R}{1 - n_R}$	**5** Two examples, one using 5% tolerance resistors and one using 1% tolerance resistors are calculated. For convenience, the previously determined values are summarized below. $\overline{V}_{CC} = \overline{V}_{KK} = \overline{V}_{EE} = 24.2\ V$ $\underline{V}_{CC} = \underline{V}_{KK} = \underline{V}_{EE} = 23.8\ V$ $\overline{V}_Z = 4.52\ V$, $\underline{V}_Z = 4.18\ V$ $\overline{E}_1 = 1.59\ V$, $\underline{E}_1 = 1.3\ V$ $\overline{E}_O = 1.55\ V$, $\underline{E}_O = 1.3\ V$ $\overline{E}_{1\ (in)} = 1.7\ V$, $\underline{E}_{1\ (in)} = 0.8\ V$ $\underline{I}_{DZ1} = 4\ mA$ (chosen to keep zener in conduction) $\overline{I}_{22} = 10\ mA$, $\overline{I}_{11} = 7\ mA$ } as obtained from load requirements $N_{VK} = \dfrac{24.2 + 1.55}{23.8 - 1.59} = 1.16$ $N_{VZ} = \dfrac{24.2 - 4.18 + 1.55}{23.8 - 4.52 - 1.59} = 1.215$

$$N_P = \frac{1 + .02}{1 - .02} = 1.04$$

$$N_R\,(5\%) = \frac{1 + .05}{1 - .05} = 1.105$$

$$N_R\,(1\%) = \frac{1 + .01}{1 - .01} = 1.02$$

Solution 1 ($N_R = .05$)

$\overline{I}_C = (1.16)\,(1.215)\,(1.04)^3\,(1.105)^3\,(7 + 10) =$ 36.5 mA

Solution 2 ($N_R = .01$)

$\overline{I}_C = (1.16)\,(1.215)\,(1.04)^3\,(1.02)^3\,(7 + 10) =$ 28.6 mA

Step 6

Select a suitable transistor and find $\overline{V}_{BE}$ and $\underline{V}_{BE}$

6

A 2N2258 will be used. $\theta_{VB} = 1.5$ mV/°C

$\overline{V}_{BE} = 0.71 + (.0015)\,(25) = 0.75$V at 0°C and 36.5 mA

$\overline{V}_{BE} = 0.64 + (.0015)\,(25) = 0.68$V at 0°C and 28.6 mA

$\underline{V}_{BE} = 0.2$V for both 36.5 mA and 28.6 mA

Step 7

Procedure Calculating values for the K terms	Solution 1 ($n = .05$)	Solution 2 ($n = .01$)
$K_{11} = \frac{1}{1 + n_{RZ}} (\underline{V}_{CC} - \overline{E}_I - \overline{V}_Z)$	$K_{11} = \frac{1}{1 + .05} (23.8 - 1.59 - 4.52) = 16.85$	$K_{11} = \frac{1}{1 + .01} (23.8 - 1.59 - 4.25) = 17.5$
$K_{12} = 0$	$K_{12} = 0$	$K_{12} = 0$
$K_{13} = \frac{1}{1 - n_{RK}} (\overline{V}_{KK} + \underline{E}_1)$	$K_{13} = \frac{1}{1 - .05} (24.2 + 1.59) = 27.3$	$K_{13} = \frac{1}{(1 + .01)} (24.2 + 1.59) = 26.1$
$K_{21} = \frac{1}{1 + n_{RZ}} (\underline{V}_{CC} + \underline{E}_O - \overline{V}_Z)$	$K_{21} = \frac{1}{1 + .05} (23.8 + 1.3 - 4.52) = 19.6$	$K_{21} = \frac{1}{1 + .01} (23.8 + 1.3 - 4.52) = 20.35$
$K_{22} = \frac{1}{1 - n_{RE}} (\overline{V}_{EE} - \underline{V}_{BE} + \overline{E}_{1\,(in)})$	$K_{22} = \frac{1}{1 - .05} (24.2 - 0.2 + 1.7) = 27.1$	$K_{22} = \frac{1}{1 + .01} (24.2 - 0.1 + 1.7) = 26.1$
$K_{23} = 0$	$K_{23} = 0$	$K_{23} = 0$
$K_{31} = \frac{1}{1 - n_{RZ}} (\overline{V}_{CC} + \overline{E}_O - \underline{V}_Z)$	$K_{31} = \frac{1}{1 - .05} (24.2 + 1.55 - 4.18) = 22.7$	$K_{31} = \frac{1}{1 - .01} (24.2 + 1.55 - 4.18) = 22$
$K_{32} = \frac{1}{1 + n_{RE}} (\underline{V}_{EE} - \overline{V}_{BE} + \underline{E}_{1\,(in)})$	$K_{32} = \frac{1}{1 + .05} (23.8 - 0.75 + 0.8) = 22.8$	$K_{32} = \frac{1}{1 + .01} (23.8 - 0.68 + 0.8) = 23.6$
$K_{33} = \frac{1}{1 + n_{RK}} (\underline{V}_{KK} - \overline{E}_O)$	$K_{33} = \frac{1}{1 + .05} (23.8 - 1.55) = 21.1$	$K_{33} = \frac{1}{1 + .01} (23.8 - 1.55) = 21.95$

Step 8

Calculate R_Z, R_C, R_K

Procedure

$$\begin{bmatrix} K_{11} & K_{12} & -K_{13} \\ K_{21} & -K_{22} & K_{23} \\ -K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{bmatrix} \frac{1}{R_Z} \\ \frac{1}{R_E} \\ \frac{1}{R_K} \end{bmatrix} = \begin{bmatrix} I_{22} \\ I_{DZ} \\ I_{11} \end{bmatrix}$$

Solution 1 ($n = .05$)

$$\begin{bmatrix} 16.85 & 0 & -27.3 \\ 19.6 & -27.1 & 0 \\ -22.7 & 22.8 & 21.1 \end{bmatrix} \begin{bmatrix} \frac{1}{R_Z} \\ \frac{1}{R_E} \\ \frac{1}{R_K} \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 7 \end{bmatrix}$$

$R_Z = 386\Omega$

$R_E = 580\Omega$

$R_K = 810\Omega$

Solution 2 ($N_R = .01$)

$$\begin{bmatrix} 17.5 & 0 & -26.1 \\ 20.35 & -26.1 & 0 \\ -22 & 23.6 & 21.95 \end{bmatrix} \begin{bmatrix} \frac{1}{R_Z} \\ \frac{1}{R_E} \\ \frac{1}{R_K} \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 7 \end{bmatrix}$$

$R_Z = 600\Omega$

$R_E = 870\Omega$

$R_K = 1.36K\Omega$

Step 9

Calculate $\overline{I}_C$ and compare to the estimated I_C in Step 5.

$$\overline{I}_C = \frac{V_{EE} - V_{BE} + E_{1\,(in)}}{R_E}$$

Solution 1 ($n = .05$)

$$\overline{I}_C = \frac{24.2 + 1.7}{(580)(.95)} = 47 \text{ mA}$$

Solution 2 ($n = .01$)

$$\overline{I}_C = \frac{24.2 + 1.7}{520\,(.99)} = 30.1 \text{ mA}$$

Both answers are sufficiently close to the estimate of Step 5 and thus, a calculation based upon this value for $\overline{I}_C$ is unnecessary.

SECTION 8-3 – CURRENT MODE FLIP-FLOP

Very fast flip-flops can be built using current mode operation. One method consists of using two zener-coupled inverters connected back-to-back with a common emitter resistor as shown in Figure 8-3-1. The circuit shown has operated at speeds up to 100mc. This circuit is an R-S flip-flop* employing "pull down" transistors as trigger amplifiers. A relatively small voltage swing at either the set or the reset line causes a change of state in a few nanoseconds.

The zener diodes provide the level shift between input and output levels. Two output levels are available from this circuit depending upon which side of the zener is used as the output. Only the output level taken from the low impedance side of the zener diode will be regulated against zener voltage changes.

The triggering of this circuit proved to be somewhat tricky because the change of state is very rapid. It was necessary to use trigger peaking circuits as shown in the test circuit of Figure 8-3-2. The Schmitt trigger which supplies the trigger is discussed in Section 8-6 of this chapter. The 50 mc output waveform and the set input from the peaker are shown in Figure 8-3-3.

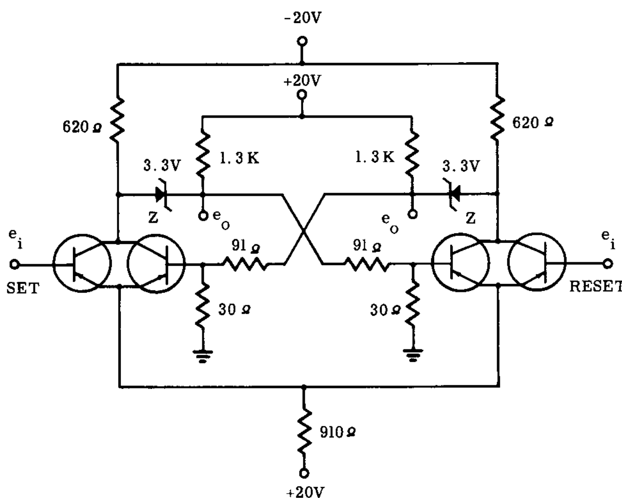


Figure 8-3-1 – 100 MC Current-Mode Flip-Flop

*R-S flip-flop refers to a flip-flop which is "reset," that is, the output is a "0," when a "1" is applied to the reset input. Similarly a "1" applied to the set input, sets the output to a "1."

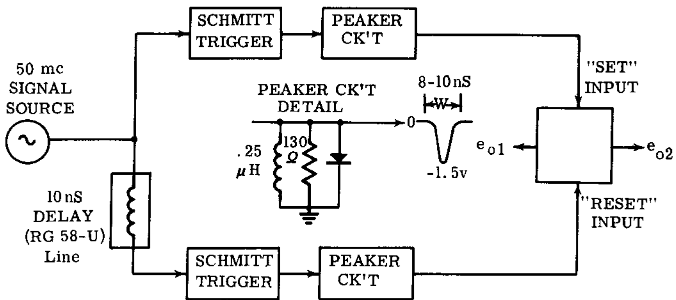


Figure 8-3-2 — Test System for 100 mc Current-Mode Flip-Flop

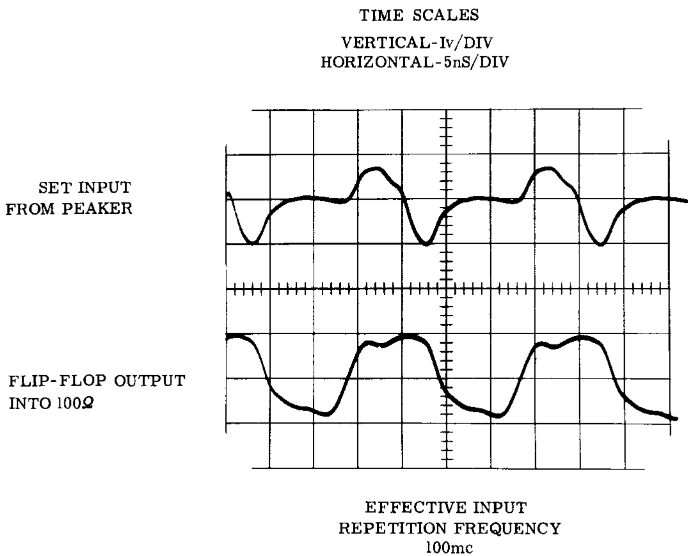


Figure 8-3-3 — Output and "Set" Waveforms for 100 mc Flip-Flop

8-4 — ASTABLE MULTIVIBRATOR

A fast rise astable multivibrator results from the current mode circuit shown in Figure 8-4-1. Placing the timing capacitor in the emitter greatly improves the output waveform compared to a conventional saturated mode circuit because the recharge voltage across the timing capacitor does not appear at the output.

The circuit is designed so that two quasi-stable states exist. Operation is similar to that of the current mode flip flop in that a constant current is switched between Q_1 and Q_2. V_{EE} and resistors R_{E1} and R_{E2} are chosen so that they represent current sources of I_1 and I_2 respectively. The voltage change across the capacitor is small. Examine Figure 8-4-2 at time T_0 when Q_1 is off and Q_2 is on. The emitter current of Q_2 is composed of the current (I_2) through R_{E2} plus the current (I_1) through R_{E1} which flows through the capacitor. As the voltage builds up across C, the emitter of Q_1 becomes more positive which decreases the off bias on Q_1. When the voltage across the base-emitter junction of Q_1 reaches V_{TF}, Q_1 begins to conduct. Conduction of Q_1 results in a change of its collector voltage which tends to turn off Q_2. Regeneration commences which rapidly turns off Q_2 and turns on Q_1.

The regeneration time is so short that the capacitor voltage does not change. Because of the shift in the dc level at the emitter of Q_2 as a result of the change of state, the voltage on C is of a polarity to place an off bias on Q_2. The current through R_{E1} now charges the capacitor and the next half cycle commences.

Note that the capacitor never recharges in the same sense as with the saturated mode circuit, but has a sawtooth waveform. The use of only one capacitor also eliminates the starting problem, common with the saturated mode circuits. It is impossible for both transistors to be in the same state.

DESIGN GUIDES: The load can form part of the collector resistance if the load is a constant resistance. A variable load should not be used because it affects the off bias which changes timing.

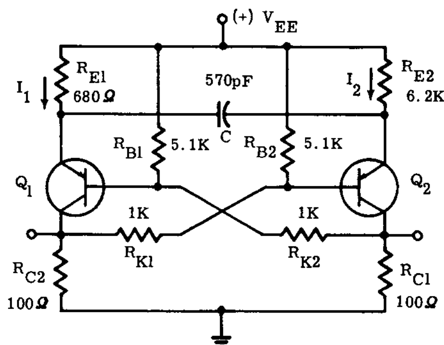


Figure 8-4-1 — Emitter Timed Astable Multivibrator

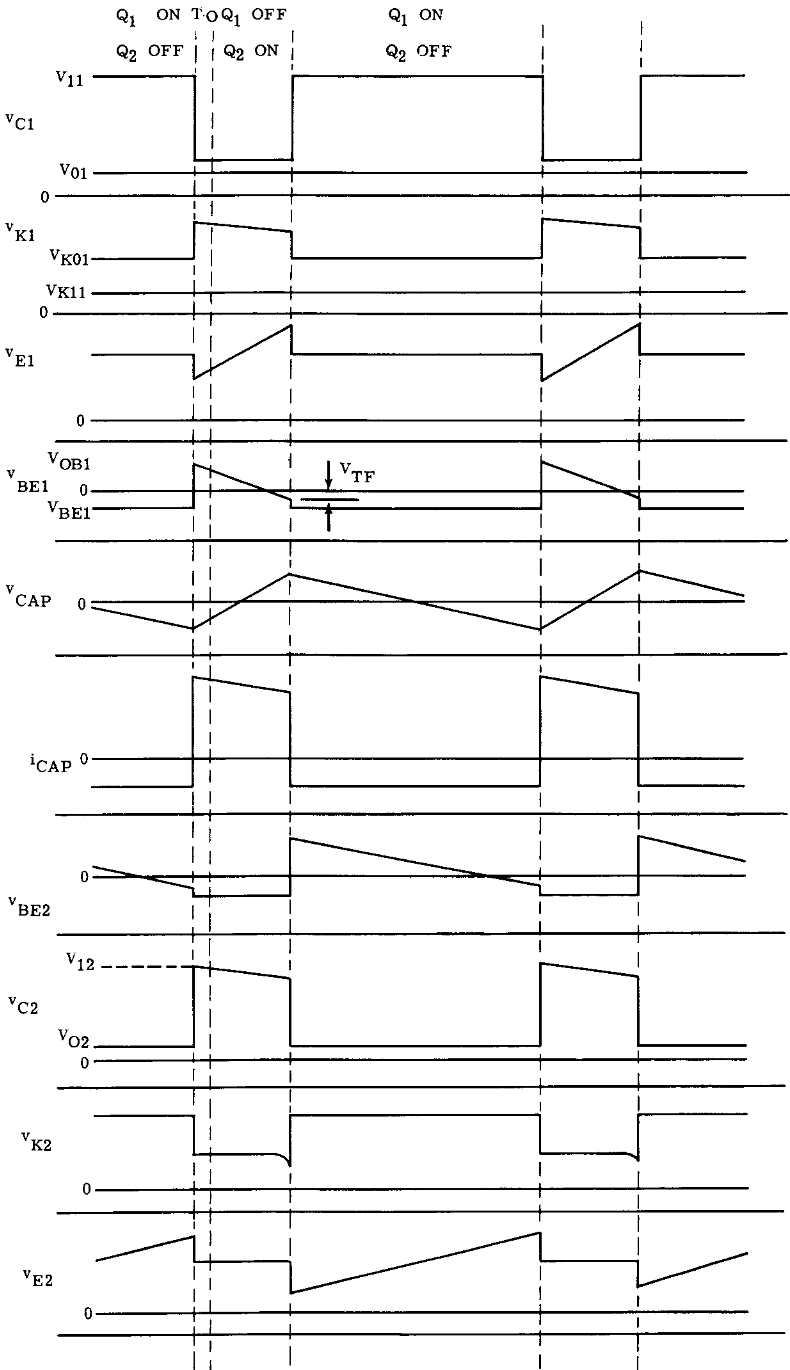


Figure 8-4-2 — Emitter Timed Astable Waveforms

The purpose of the cross coupling network is to provide the proper level at the base so that the transistors can never enter the saturation region and to provide a sufficient off bias for good timing stability. For good on current stability, the current through R_B and R_K should be large compared to the maximum base current required by a transistor.

The voltage swing (v_c) across the timing capacitor is approximately equal to V_{OB} which in turn is slightly less than the output voltage swing e_o. The values of R_K and R_B determine V_{OB}, which must be limited to a voltage less than BV_{EBO} of the transistor but should not be too low otherwise the timing will be sensitive to changes in V_{TF}.

The voltage at the emitter must change by V_{OB}, and some minimum V_{CB} is required to avoid saturation. To make I_1 and I_2 appear as current sources then the condition:

$$V_{EE} \gg V_{OB} + V_{CB} + V_1 \tag{8-4-1}$$

must be met.

All things considered, a reasonable value for V_{OB} is 2 volts. Having selected V_{OB} the ratio of R_K and R_B can be chosen, keeping in mind that the sum of R_B and R_K should be such that the current through them is large compared to the base current. With these constraints, suitable values for R_B and R_K may be found.

Assuming the currents I_1 and I_2 are nearly constant the timing is given by

$$T_1 \approx \frac{CV_{OB}}{I_2} \tag{8-4-2}$$

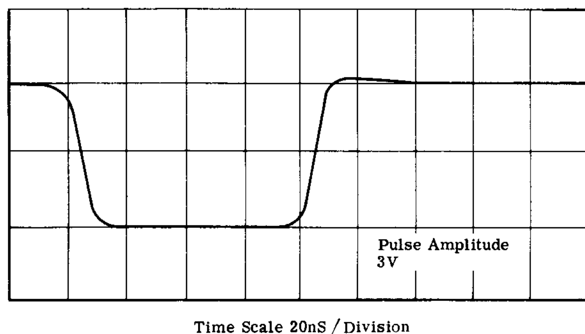
$$T_2 \approx \frac{CV_{OB}}{I_1} \tag{8-4-3}$$

The emitter current is given by

$$I_E = I_1 + I_2 \tag{8-4-4}$$

The sag in the output waveform (v_{C2}) as shown on Figure 8-4-2 occurs because the current I_1 does not remain absolutely constant as the capacitor charges. Since T_1 is shorter than T_2, I_1 is greater than I_2. Therefore, slight changes in I_1 are noticeable at the output of Q_2, as the sum of I_1 and I_2 flow through it when on. Conversely, slight changes of I_2 are not noticeable at the output of Q_1.

Figure 8-4-3 shows the output of Q_1 using the circuit of Figure 8-4-1. The 6nS transition times illustrated cannot be attained with conventional astable multivibrators.



Time Scale 20nS / Division

Figure 8-4-3 — Output Waveform of Astable Multivibrator

SECTION 8-5 — CURRENT MODE MONOSTABLE MULTIVIBRATOR

In the astable multivibrator, a capacitor connected between the emitters provides the timing mechanism. If one side of the capacitor is returned to ground instead of to an emitter, a current mode monostable multivibrator results. A circuit using this idea is shown in Figure 8-5-1.

In this circuit, diode D_1 is used to set a reference voltage for the emitter of Q_1. Since the emitter of Q_2 does not have a reference voltage, it is on in the stable state. The pertinent waveforms are shown in Figure 8-5-2. A trigger signal applied to the base of Q_2 turns it off, and its collector voltage rises. This change in voltage is coupled to the base of Q_1 causing regeneration which results in a rapid change of state for the transistors. When Q_1 is on, a reverse bias is placed on the base of Q_2. Timing starts as C_D charges through the emitter resistor of Q_2 toward 24 V and continues until C_D has charged to a voltage such that Q_2 becomes forward biased. As Q_2 conducts, its collector voltage changes and regeneration quickly turns Q_1 off bringing e_0 to ground. Then, C_D discharges through the base-emitter junction of Q_2 toward ground potential. Q_2 acts as an emitter follower in discharging C_D, however, the time constant involved in discharging C_D is large. The repetition rate is limited by the recovery time of the capacitor and in this case the recovery time is almost comparable to the timing interval. Maximum repetition rate for the circuit of 8-5-1 is 1 mc.

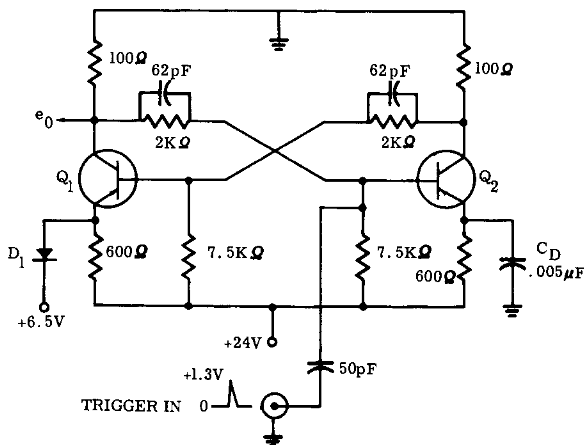


Figure 8-5-1 — Emitter-Timing, Current-Mode Monostable

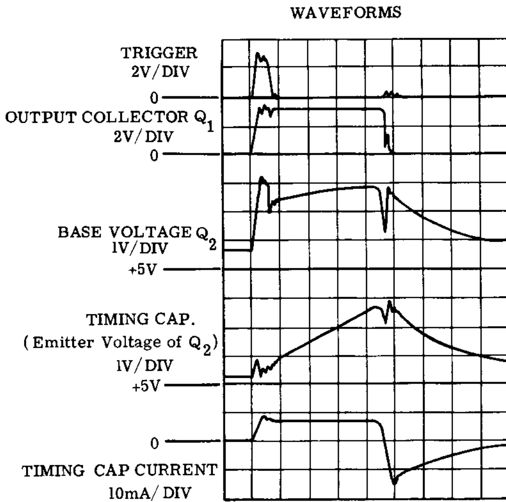


Figure 8-5-2 — Emitter Timing, Current Mode Monostable

Another type of monostable multivibrator can be created by using a common emitter resistor and supplying one side with base current to hold it on as shown in Figure 8-5-3. Current through R_B holds diode D_1 in conduction and base current to hold Q_2 on is borrowed from the diode. The zener diode and associated coupling network are chosen so that Q_1 is biased off.

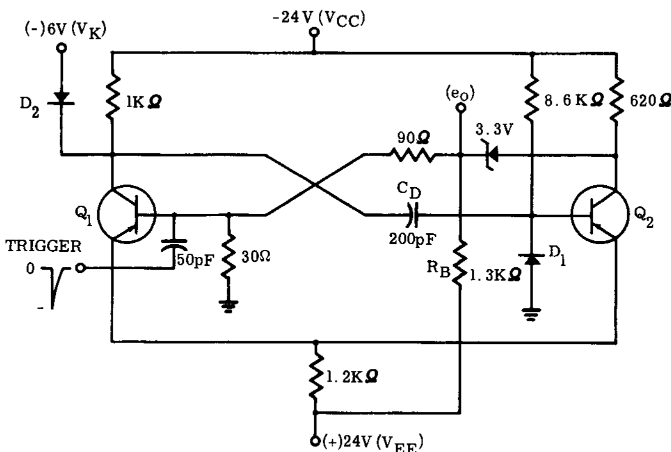


Figure 8-5-3 — Base Timing Current Mode Monostable

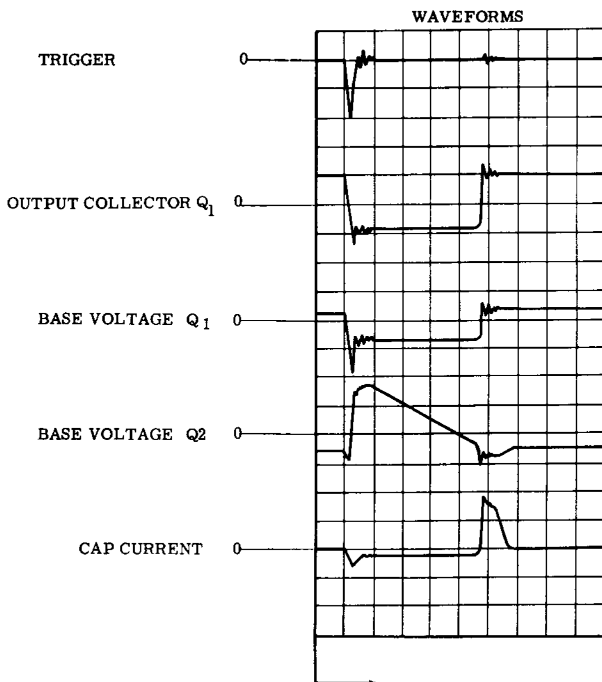


Figure 8-5-4 — Base Timing Current Mode Monostable

Operation is analogous to the saturated mode monostable circuit. As can be seen from the waveforms in Figure 8-5-4, application of a trigger turns Q_1 on causing its collector voltage to drop. This drop is coupled through the large timing capacitor, which appears as a short to the transient, to the base of Q_2 turning Q_2 off and regeneration causes a rapid change of state. Timing begins as C_D charges through R_B, and ends when Q_2 starts to conduct. A slight conduction of Q_2 causes regeneration to begin, thus returning conditions to those of the stable state. Diode D_2 and the 6V supply limit the voltage appearing across C_D and thus prevents excessive reverse emitter junction voltage from being applied to Q_2 when this voltage appears at its base during the timing cycle.

Recovery of this circuit is fast because the charge current is equal to the collector current of Q_2 and the voltage waveform during recovery is a clipped exponential due to the clipping action of diode D_2. Thus, charging is linear and only a small fraction of the RC charging time constant. The duty cycle of this circuit can be as high as 80%. In this regard, the base timing monostable is a better performing circuit than the emitter timing circuit previously described.

These designs have the advantage of being sensitive to small trigger signals but insensitive to level. Noise is not a problem, however, as trigger level may be set well above any noise level. Regeneration is extremely swift in these circuits resulting in very fast rise and fall times.

SECTION 8-6 — CURRENT MODE SCHMITT TRIGGER

As shown on Figure 8-6-1, a current mode Schmitt Trigger can be constructed from a simple inverter stage where the input reference voltage of one stage depends upon the collector voltage of the preceding stage. The second inverter is isolated from the input voltage and only responds to changes which are large enough to cause regeneration which results in a change of state.

Since the base voltage on Q_1 determines the states of Q_1 and Q_2, the bias potentiometer serves as a sensitivity-symmetry control. Variations in the nominal zener diode voltage will require trimming this adjustment. The zener coupling method used in the inverter of Section 8-1 could be employed to render the circuit insensitive to shifts in zener diode voltage. The zener diode is kept in conduction at all times by R_6. Resistor R_4 supplies a constant current into the 100 ohm output load so that the output levels are shifted to +1 volt and —1 volt from the levels of 0 and —2 volts, which would otherwise result. The shift in levels is often needed to properly trigger gates and flip-flops.

Using 2N2258 transistors, rise and fall times are 2 and 3 nS respectively, at a 50 mc input frequency. Figure 8-6-2 shows the output waveforms under different loading conditions. The rise and fall times with capacitive loading are governed primarily by the R-C time constant of the load; that is, $t_r = t_f = 2.3\ R_L\ (C_L + C_{ob})$, where C_L is the load capacitance.

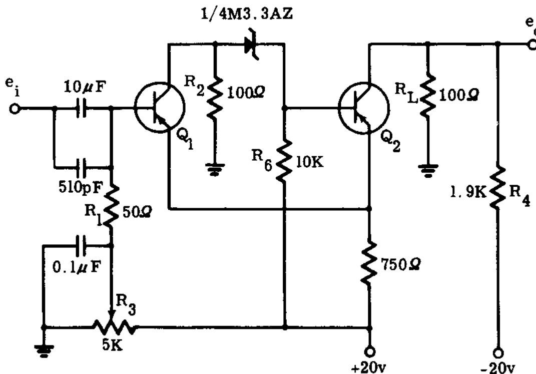
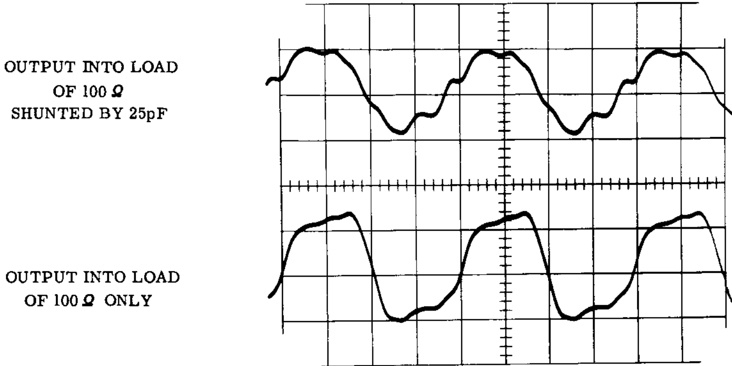


Figure 8-6-1 — 50 mc Current Mode Schmitt Trigger



REPETITION
FREQUENCY 50mc

Figure 8-6-2 — Output Waveforms of Current Mode Schmitt Trigger

CHAPTER 9

Avalanche Mode Switching

In conventional saturated mode or current mode transistor circuits, operation in the avalanche region has to be carefully controlled if latch-up is to be avoided. Under some conditions, operation in the avalanche region may even result in permanent damage to the device.

However, avalanche mode operation can provide extremely high switching speeds and is capable of producing a current output far in excess of that obtained from conventional circuits. This chapter discusses some of the considerations for avalanche mode operation, reviews behavior of transistors in the avalanche region, and presents some useful circuits.

Avalanche multiplication occurs in reverse-biased PN-junctions as a result of impact ionization produced by mobile charge carriers. The process — similar to Townsend breakdown in gases — may be visualized as follows:

An electron, thought of classically as a charged particle, moving through the crystal lattice in the depletion region of the junction, gains energy as it is accelerated by the electric field across the junction. Eventually it collides with one of the atoms of the lattice and, if it has gained sufficient energy prior to the collision, may disrupt the atomic bonds and release an electron by impact ionization. In so doing, it also releases a hole; there are now three charged particles, each of which may participate in one or more ionizations. Thus, an *avalanche* of charged particles is produced and a large current flows through the junction.

A multiplication factor (M) is defined as the number of hole-electron pairs produced per carrier entering the depletion layer. The analytical expression is extremely complex but Miller[1] has found that the behavior of M with reverse junction voltage can be approximated by

$$M = \frac{1}{1 - (V/V_B)^m} \tag{9-1}$$

where V is the reverse junction voltage, V_B is a critical voltage at which M becomes infinite and is called the *breakdown voltage,* and m is a constant whose value is determined by the type of semiconductor material. A plot of M vs. V/V_B is shown in Figure 9-1.

In the PN-junction, the ionizing carriers are thermally generated, however conditions necessary for avalanche multiplication can also be produced in the reverse-biased collector junction of a transistor where carriers are injected as well as thermally generated. Since the carriers injected into the collector depletion region can be controlled by the emitter and base currents, it might be anticipated that avalanche behavior in a transistor would be much more interesting than that of a single junction.

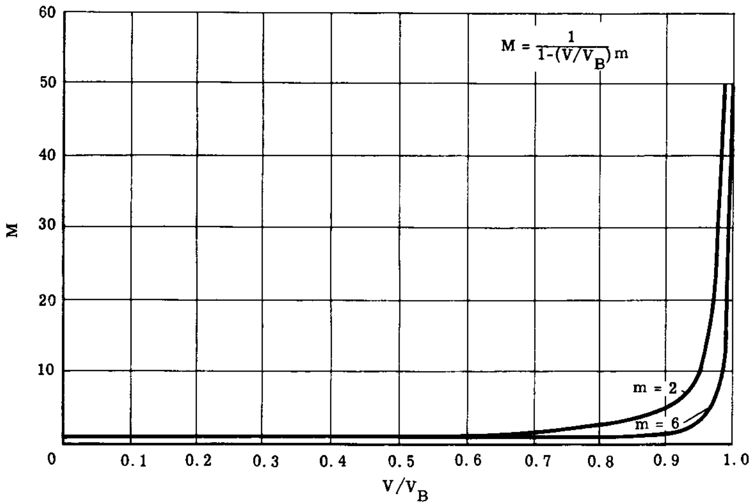


Figure 9-1 Avalanche Multiplication Factor (M)

9-1 — Static Characteristics of the Avalanche Region

The general behavior of a transistor when collector avalanche effects are considered can be found by simply incorporating the multiplication factor M into the general transistor equation. That is:

$$I_C = M\,(\alpha I_E + I_{CB})$$

where I_{CB} is the bulk reverse diffusion or charge generation current excluding multiplication effects. The surface current has been neglected.

Substituting $I_E = I_C + I_B$ and simplifying

$$I_C = \frac{M}{1 - \alpha M}\,(I_{CB} + \alpha\, I_B) \tag{9-2}$$

In Chapter 3, it was shown that entrance into the avalanche region occurs when αM is greater than unity in which case V_{CB} must exceed $V\alpha_M$ (the voltage where $\alpha M = 1$). If I_C is to be larger than I_{CB}, I_B must be negative in order to satisfy equation 9-2.

Substituting Miller's equation 9-1 into 9-2 and assuming $V_{CB} \approx V_{CE}$

$$I_C = \frac{I_{CB} - \alpha I_{BR}}{1 - \alpha - (V_{CE}/V_B)^m} \tag{9-3a}$$

Solving for the collector voltage V_{CE}

$$V_{CE} = V_B \left\{ 1 - \alpha + \alpha \left[\frac{I_{BR}}{I_C} - \frac{I_{CB}}{\alpha I_C} \right] \right\}^{1/m} \tag{9-3b}$$

where I_{BR} is understood to be the magnitude of the negative base current.

Since the base current is in the reverse direction and the emitter current is in the forward direction, the collector current can never be less than the base current, i.e.: $I_C = I_{BR} + I_E$.* When I_C becomes large compared to I_{BR} and I_{CB},

*That is, equation 9-3 is not valid when the emitter becomes reverse biased by any significant amount as the emitter reverse currents have been neglected in this analysis.

observe that V_{CE} becomes

$$V_{CE} = V_B [1 - \alpha]^{1/m}. \quad (9\text{-}4)$$

This relationship was derived in Chapter 3 (equation 3-16a) for the case when $\alpha M = 1$; the voltage at this point was defined as $V_{\alpha M}$. Therefore, regardless of conditions at the base, the collector voltage of a transistor, which is on in the avalanche mode, approaches $V_{\alpha M}$ when I_C is high. Thus, $V_{\alpha M}$ becomes a significant voltage for avalanche mode operation.

$V_{\alpha M}$ is not constant with current because both α and m are current sensitive. Usually, α rises to a peak at a moderate current level giving a dip in a plot of $V_{\alpha M}$ vs. current. Also, m increases (that is, multiplication decreases) as current level increases due to depletion layer widening effects. These effects explain the typical behavior of $V_{\alpha M}$ as shown in Figure 9-2.

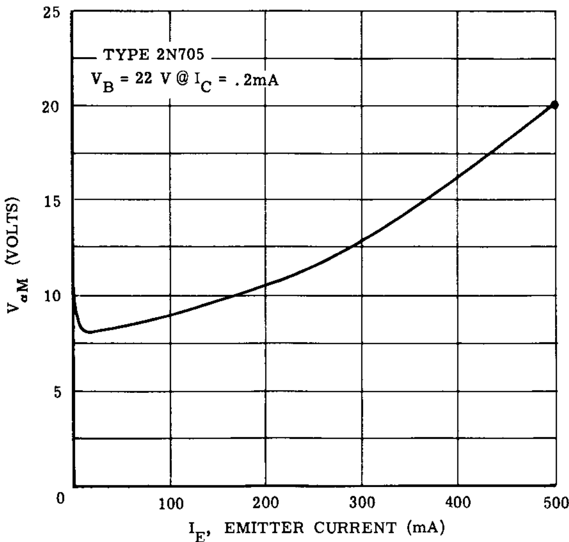


Figure 9-2 Typical $V_{\alpha M}$ Behavior with Emitter Current

From equation 9-3b, when $I_{BR} = I_C$ observe that V_{CE} approaches V_B as I_C is increased. The condition where $I_{BR} = I_C$ corresponds to a slight reverse bias voltage (ϕ_{TR}) on the base (previously given in equation 3-11) in order to keep I_E zero. For this special case, from equation 9-2, by setting $I_C = I_B$:

$$I_C = I_B = M\, I_{CB}.$$

This equation defines the conditions at the off point. The point where the collector current as given by equation 9-3a intersects the M I_{CB} curve indicates the point where the emitter begins to inject.

A particularly interesting condition occurs when $I_{BR} = I_{CB}/\alpha$. Then for $I_C > I_{CB}$, $V_{CB} = V_{\alpha M}$.

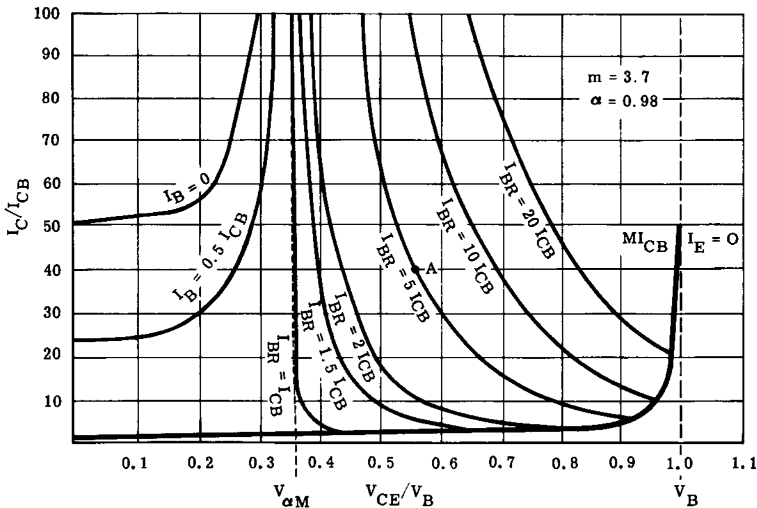


Figure 9-3 Theoretical Plot of Transistor Output Characteristics in the Avalanche Region

Behavior in the avalanche region, as represented by equations 9-3a and b, is plotted in Figure 9-3 for the case where m $=$ 3.7 (a typical measured value for the 2N705). Actual devices closely approximate this behavior. The chief departures arise because of the variations of m and α with current as previously discussed. The relative ratios of $V_{\alpha M}$ to V_B and the shape of the curves would vary slightly with changes in α and m due to the different device structures.

The heavy line corresponds to the case of zero emitter current where $I_C = I_{BR} = MI_{CB}$. Notice that the transistor exhibits a high negative resistance at the point of emitter injection. The incremental avalanche resistance ($r_A = dV_{CE}/dI_C$) is negative when $I_C > I_{BR}$ and decreases at a given I_C as I_{BR} increases. For large values of I_C, V_{CE} approaches $V_{\alpha M}$ and r_A approaches zero. For $I_{BR} = I_{CB}$, as the curve shows, V_{CE} is always slightly greater than $V_{\alpha M}$ for any $I_C > I_{CB}$.

9-2 — Transient Characteristics

An incremental model of a transistor operating in the avalanche mode can be developed in order to investigate its stability and operating points in circuits. The nature of this model can be qualitatively deduced by considering the effects of differential currents (i) and voltages (v). The incremental resistance r_A was defined as dv/di at a given operating point. In addition a capacitive effect is evident because during transient conditions the amount of stored charge is changing. This changing stored charge can be accounted for in a model by an incremental capacitance C_A which is defined as dq/dv.

Assume the operating point is located at Point A on Figure 9-3 and a negative increment of voltage $\triangle$v is applied to the terminals. This is accompanied by an increase of current $\triangle$i. Assume effects of C_{ob} are negligible, then to produce this increase of current $\triangle$i, the stored charge must increase an amount $\triangle$q.

Therefore,

$$C_A = \frac{\Delta q}{\Delta v} = -\left|\frac{\Delta q}{\Delta v}\right|, \qquad (9\text{-}5)$$

which is a negative value throughout the negative resistance region. If Point A is moved into a region of higher current, the same Δv is accompanied by a larger Δi. This means that C_A approaches minus infinity as I_C increases. Also, as Point A is moved to the very low current range the same reasoning reveals that C_A approaches zero.

In Chapter 5 it was shown that the change in charge required to change I_C by a small amount was given by:

$$\Delta q = \Delta i \left(\frac{1}{\omega_T} + \frac{\Delta v}{\Delta i} C_{ob}\right) \qquad (9\text{-}6)$$

Substituting 9-6 into 9-5 we have

$$C_A = \frac{\Delta i}{\Delta v} \frac{1}{\omega_T} + C_{ob}$$

but, $\Delta v / \Delta i$ is the negative resistance r_A, therefore,

$$C_A = -\left(\frac{1}{r_A \omega_T} - C_{ob}\right). \qquad (9\text{-}7)$$

Thus, when the collector capacity is included in the expression for C_A it is found that:

$$\text{at high } I_C,\ r_A \rightarrow 0 \quad \therefore C_A \rightarrow -\infty$$
$$\text{at low } I_C,\ r_A \rightarrow \infty \quad \therefore C_A \rightarrow C_{ob}.$$

Similar reasoning will show that both r_A and C_A are positive on the $I_E = 0$ line.

Since the differential change of terminal current accompanying a differential change of voltage is the sum of changes occurring in the resistance and capacitance, the equivalent circuit takes the form of r_A and C_A in parallel as shown in Figure 9-4.

In summary then, the salient features of operation in the avalanche region are:

1. $V_{\alpha M} < v_{CE} < V_B$
2. Negative nonlinear output resistance
3. Negative nonlinear output capacitance

(2 and 3: measured between collector and emitter terminals)

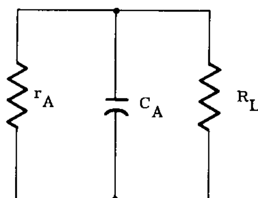


Figure 9-4 Equivalent Circuit of a Transistor in the Avalanche Region with Resistive Load

9-3 — Operation of a Single Transistor Circuit

Operation in the avalanche mode is regenerative, as inferred by the presence of the negative resistance. The regenerative behavior can be understood by means of a qualitative example. Consider a single transistor biased as shown in Figure 9-5a. Here the available reverse base current is constant at about 1 mA, and the collector supply voltage is approximately V_B. The emitter junction is slightly reverse biased, and the transistor operating point is at D as shown in Figure 9-5b. The multiplication factor M of the transistor is large since V_C is close to V_B.

Assume for simplicity that a small amount of current is injected into the emitter (by means not illustrated on the figure), causing the emitter junction to become forward biased. When the injected current reaches the collector, it is multiplied by M. However, the base current is held constant; therefore, the increase of collector current results in a similar increase of emitter current since the V_{CC} supply cannot store current. This regenerative process continues; as the current builds up, the collector voltage decreases which in turn decreases M: the build up proceeds more slowly until an equilibrium condition is reached, at the operating point A, where regeneration is no longer possible. The operating point will remain at A until some external trigger is applied to cause it to move back to D. Thus, the single transistor circuit of Figure 9-5a exhibits bistable operation, moving regeneratively from D to A and from A to D in response to an externally applied trigger.

Let us examine this process in more detail. In general there are three points of intersection of a load line with the transistor characteristic, (A, B and D on Figure 9-5b). It is helpful to investigate the potential stability of these points.

The stability of a static operating point can be found by representing the circuit by its incremental model and finding its natural frequencies. If any natural frequencies correspond to a growing transient, the operating point is unstable because any disturbing signal will grow with time. This will cause operation to move in the direction of the disturbing signal until a stable point is found. Conversely, a stable point is one whose natural frequencies correspond to decaying transients.

Using pole-zero-theory, the natural frequency of the transistor circuit is found by inspection of Figure 9-4 to be:

$$S = -\frac{g + G}{C_A} \qquad (9\text{-}8)$$

where $g = 1/r_A$ and $G = 1/R_L$

As long as S is negative, the pole is located in the left hand of the complex S plane and operation is stable. Clearly then on the $I_E = 0$ curve, where both g and C_A are positive, operation is always stable, an obvious conclusion. However, in the region where $I_C > M\ I_{CB}$ both g and C_A are negative. Therefore, in order to make S negative $|G| < |g|$. In terms of resistance

$R_L > r_A$ for a stable point

$R_L < r_A$ for an unstable point.

Obviously the point where $R_L = r_A$ (i.e.: $1/R_L$ is just tangent to the transistor characteristic curve) is the boundary between stable and unstable points. At Point A, $R_L > r_A$ and Point A is stable while at Point B, $R_L < r_A$ and Point B is unstable. The characteristics of the avalanche region are typical of a group

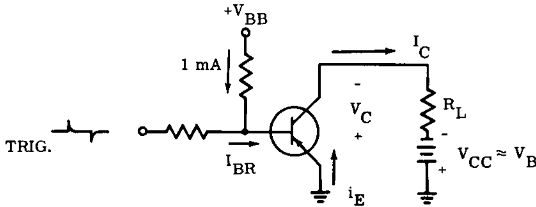


Figure 9-5a Basic Bistable Circuit Biased for Avalanche Operation

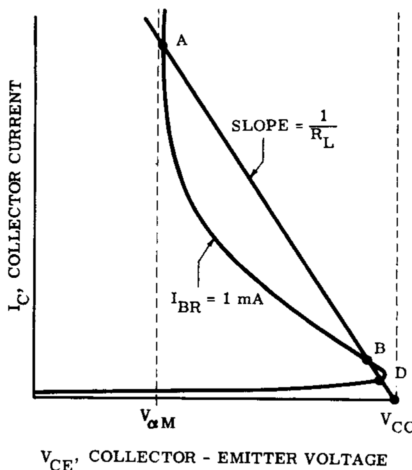


Figure 9-5b Load Line for the Basic Bistable Circuit

of negative resistance devices which are called open circuit stable devices. That is, they are stable at any point under conditions where $R_L \rightarrow \infty$.

Switching time is fairly independent of load resistance due to the peculiar inter-relationship[2] of the functional behavior of C_A and the v-i characteristic.

Consider how current must build up to move operation from Point D to Point A, which is the rise time. The trigger circuit causes injection of an amount (c) of carriers from the emitter into the base, which reach the collector depletion region after a transit time T*. Multiplication occurs and Mc carriers leave the collector. However, all carriers (since the base current is constant) that leave the collector must flow into the emitter. These Mc carriers again flow across the base, after another interval T they are multiplied and M^2c carriers leave the collector. Thus, a regenerative build-up occurs, its rate depending upon the transit time T and M.

*$T \approx 1/\omega_T$

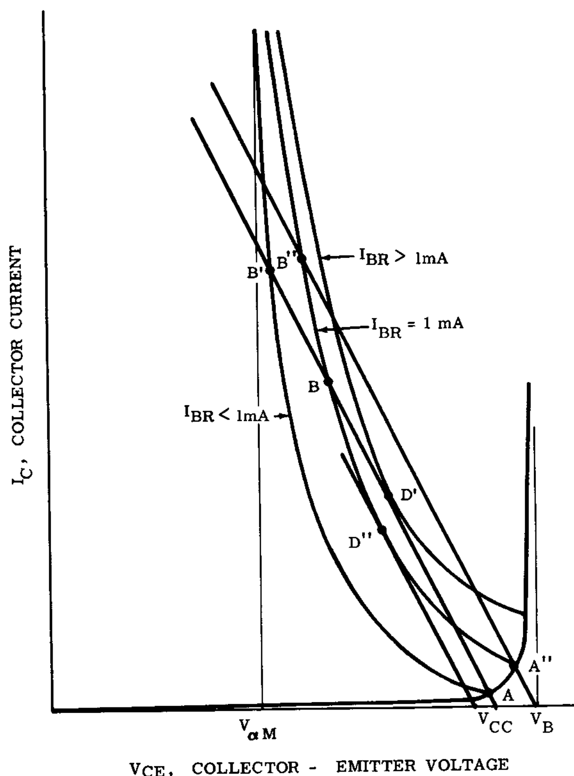


Figure 9-6 Voltage-Current Characteristic Curves and Load Lines in the Avalanche Region

As the switching interval is first initiated, because v_{CE} is high, M is large, and T is at its smallest value since the base width is at its narrowest.

Therefore, regeneration proceeds quite rapidly. However, as v_{CE} approaches $V\alpha_M$, T has decreased somewhat and M is only slightly greater than unity. This causes the rate of rise to become progressively less as $v_{CE} \rightarrow V\alpha_M$. The overall result is an approximately exponential response time.

A mathematical expression for rise time is extremely complex[2] and of doubtful value for the circuit designer. Measured values for a 2N705 transistor are typically 40nS which is approximately 30 T.

The decay of current which moves operation from A to D — the fall time interval — is the inverse of the current build up. Accordingly the switching response from A to D is characterized by a large delay and slow initial rate of rise, since M is small initially.

The action of the simple avalanche circuit, can now be reviewed, giving more attention to the details. Referring to Figure 9-6 assume operation is initially at Point A where $I_{BR} = 1\text{mA}$. In order to trigger the circuit a situation must be produced where $r_A > R_L$. This can be accomplished in two ways:

1. The base current could be lowered causing the operating point at Point

A to become unstable ($R_L < r_A$). Once operation is in the negative resistance region, base current could be restored to 1 mA, in which case the only stable operating point is at B. However, if the trigger is still present at the end of the switching time, then operation will be at B′.

2. Collector voltage could be increased causing the load line to shift to A″. At this point $R_L < r_A$ and the circuit will trigger. The stable point would again be at B with the trigger removed or at B″ with the trigger present.

In order to turn the transistor off to Point A, the condition $R_L < r_A$ must again be produced. This can be accomplished by the inverse of either of the methods used initially. That is I_{BR} could be increased, which would move the operating point to D′ or V_{CC} could be lowered which would move the operating point to D″. D′ and D″ are unstable points which would cause the circuit to trigger in the direction of the disturbance, i.e.,: back to Point A.

The operating path which the circuit follows must always lie on the load line. The difference in current between load line and v-i characteristic is the current which flows into the device capacitance. This, of course, assumes that the device can be represented by a single non-linear resistance in parallel with a non-linear capacitance.

9-4 — General Circuit Design Considerations

BIASING: The previous discussion has shown that the switching speed is determined in part by the value of M. It follows then that V_{CE} at the off point should be very close to V_B. That there is an optimum value for I_C to produce maximum M can be deduced by the following reasoning:

For $I_C = I_{CB}$, (i.e.: $I_D + I_G$), $M = 1$ and the collector voltage is small, as is the collector dissipation. As I_C increases, I_{CB} is approximately at its ambient temperature value; thus M increases. As I_C is further increased, making v_{CE} approach V_B, the dissipation rapidly increases, and the heating of the junction causes I_{CB} to increase. Eventually, the percentage change of I_C will be less than the percentage increase of I_{CB} which it causes; thus M decreases. An approximate analysis[3] indicates that the optimum bias current required to produce maximum M is

$$I_C\,(\text{opt.}) = \frac{kT_A^2}{E_g\,\theta V_B} \qquad (9\text{-}9)$$

and

$$M\,(\text{max.}) = \frac{I_C\,(\text{opt})}{e\,(I_{CBA})} \qquad (9\text{-}10)$$

where
- k is Boltzmann's constant (8.63×10^{-5} ev/°K)
- E_g is the gap energy for the semiconductor material (0.72 electron-volt for germanium and 1.1 electron-volt for silicon at 300°K — Room Temperature)
- θ is the thermal resistance in degrees C per watt
- T_A is the ambient temperature in degrees Kelvin
- I_{CBA} is the bulk current at the ambient temperature
- e is the Naperian base 2.718.

For a 2N705 transistor the optimum current calculates to be about 1 mA.

TRIGGERING: In the previous discussion it was shown that there are two basic methods of triggering; the base current can be lowered (i.e. biased more toward the forward direction) or the collector voltage can be increased. This section will show that the second method — collector triggering — is preferable from transient considerations[4].

If the base is triggered it is found that several difficulties arise. Long unpredictable delays ranging from 10 to 100 nanoseconds occur and the amplitude of the trigger pulse affects the output pulse waveshape. These problems are due mainly to two factors:

1. The most significant factor arises because the trigger lowers the collector-base voltage by an amount equal to the trigger amplitude, which results in a decrease of M. The decreased M considerably reduces circuit speed. (In this manner the trigger opposes the regenerative action.)

2. The emitter-base input capacitance (C_{ib}) must be charged through the base spreading resistance or lateral base resistance r'_B, which is fairly high because the high collector voltage causes the base width to be narrow. This effect is variable among transistors of the same type and is most severe in alloy type transistors. (See Chapter 1).

Conversely, collector triggering results in distinct advantages.

1. The trigger signal increases the collector to base voltage thus increasing multiplication.

2. The multiplied collector current flows into C_{ib} through a very small transverse base and collector resistance (approximately R_F — see Chapter 4) which causes rapid charging of this capacitance.

The resulting delay is found to be nearly independent of transistor type and to be of the order of a few nanoseconds.

SELECTING TRANSISTORS: From the discussions so far, it is clear that V_B and $V_{\alpha M}$ determine a great many of the characteristics of avalanche mode circuits. It follows that if avalanche mode circuits are to be built that do not need to be tailored to fit each transistor, a device type with a very close distribution of V_B and $V_{\alpha M}$ must be chosen. The transit time (T) must also be small for fast switching times. Most standard mesa switching transistors fulfill these requirements. The 2N705 is used in the circuits in this chapter.

9-5 — Avalanche Mode Circuits

Single transistor avalanche mode circuits are not particularly satisfactory. As has been pointed out for the resistive load circuit, rise time is long (30 T) and a long delay precedes switching from the on to off state. To minimize this delay a large turn-off pulse is required. An R-C circuit can be built to produce monostable operation, but the pulse shape is not rectangular and is difficult to control. The rise time, however, is much faster, on the order of 5T, because of the capacitive load[2].

Bistable operation has not proven practical. A single transistor circuit is slow and requires triggers which differ in polarity. A two transistor bistable circuit has been constructed which illustrates the remarkable speed attainable with avalanche mode circuits, but since it must be direct current coupled, the biasing problems introduced make the circuit unattractive.

Avalanche mode circuits find their greatest appeal in the generation of rectangular pulses. The single transistor circuit when used with a delay line provides a useful pulse shape because the delay line provides a substantial turn-off pulse. In general though, two transistors are needed; the circuit is arranged so that at least one device is operating with high M during the switching interval and so that the other device produces a large turn-off pulse. This technique greatly enhances circuit speed.

The circuits to be described[2,5,6] were developed in the Engineering Research Laboratories at the University of Arizona under a contract agreement with Motorola. Since no products are specified for worst-case avalanche mode operation, no attempt has been made to design the circuits on a worst-case basis.

The circuits should be of interest as they illustrate good design practice and have proven to be quite workable. The only transistor characteristics having major influence upon performance are V_B and V_{α_M}. In a Delay Line Generator, the pulse amplitude is determined by (V_B - V_{α_M}) and the duration is controlled by the delay line. In a high current monostable circuit, pulse width control is relatively independent of (V_B - V_{α_M}). However, the amplitude is determined by (V_B - V_{α_M}). This circuit is useful for the generation of large current pulses into a low impedance load. In a low level monostable circuit, very fast rise times are featured but the pulse duration, rather than the amplitude, is determined by (V_B - V_{α_M}).

DELAY-LINE PULSE GENERATOR: An effective circuit for generating either positive or negative pulses with controllable pulse amplitude and duration can be obtained by the use of a single transistor operating in the avalanche mode in conjunction with a delay line. The circuit is, in many respects, analogous to the familiar gas-tube pulse generator. The basic circuit of such a pulse generator is shown in Figure 9-7.

In this circuit, the output voltage is positive, however, negative voltage can also be obtained, if the load R_L is placed in series with the emitter-to-ground connection.

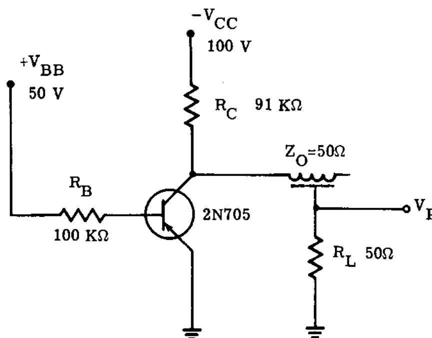


Figure 9-7 Delay Line Pulse Generator

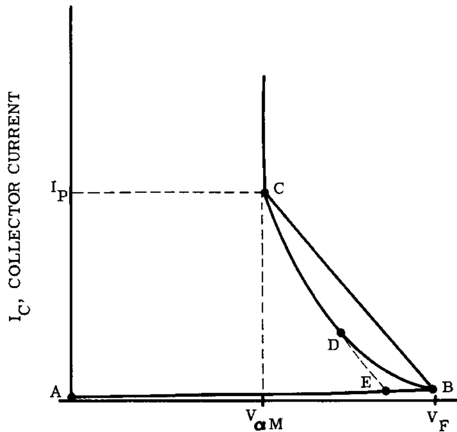


Figure 9-8 Static Characteristics of Delay Line Generator

Circuit Operation: Assume that initially the operating point is at A on Figure 9-8. The resistance R_C in conjunction with the bias supply V_{CC} charges the delay line, which behaves essentially as a capacitive load during the off-cycle, toward the supply voltage. When a critical voltage $V_F \approx V_B$ (as determined by I_B and I_C) is reached, the emitter begins to inject and the transistor exhibits a negative resistance which causes regeneration. During the regenerative transient the current through R_C can be neglected. The delay line appears to be a resistance equal to its characteristic impedance Z_o, thus the operating point moves to Point C. Figure 9-9 illustrates circuit condition at this point. The switching transient generates a positive transient on both the line and the load. When the transient on the line reaches the open end, it is reflected without a change of polarity. When it returns to the collector, it increases the voltage across that end of the line; since the line and load are in series with the fixed source of $V\alpha_M$, this results in a reduction of the voltage across R_L. In turn, the current is reduced to a value which can no longer sustain the collector voltage at $V\alpha_M$. Operation moves from Point C to Point D, an unstable point, which causes regeneration and moves the operating point to Point E. Reflections may now occur on the line but since the impedance of the transistor is high, they cannot appear at the output. The cycle is then repeated.

Circuit Analysis: The pulse duration is simply the two-way propagation time of the line or

$$T_D = 2\ T_P \tag{9-11}$$

The peak pulse current, (I_P) can be determined from

$$I_P = \frac{V_B - V\alpha_M}{R_L + Z_o} \tag{9-12}$$

and the output amplitude (V_P) from

$$V_P = (V_B - V\alpha_M)\ R_L/(Z_o + R_L) \tag{9-13}$$

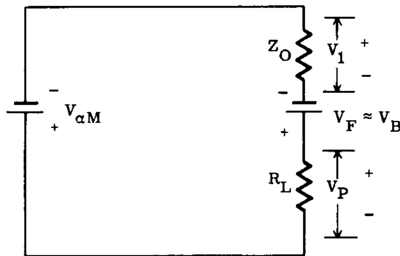


Figure 9-9 Equivalent Circuit for the Delay Line Generator in the On State

The maximum output voltage is limited to $(V_B - V\alpha_M)$ and occurs when $R_L \gg Z_o$. However, this condition will result in reflections on the load. This occurs because the amplitude of the pulse on the line (V_L) is given by:

$$V_L = (V_B - V\alpha_M)\, Z_o/(Z_o + R_L) \qquad (9\text{-}14)$$

The voltage V_L would be very small compared to V_P if $R_L \gg Z_o$. When this small voltage returns from its trip down the line, it is incapable of reducing the load current sufficiently to terminate the pulse. Thus, reflections will occur until the voltage at the output is reduced to approximately $V\alpha_M$. The load current will then be small enough to initiate regeneration to turn off the transistor. The criterion for no reflections to occur can be found by determining under what condition the voltage on the line is capable of reducing the load current to a value too low to maintain operation above point D. Obviously if $V_L \geqq V_P$, the load current will be reduced to zero and no reflections can occur. Therefore, the criterion for no reflections is approximately

$$R_L \leqq Z_o \qquad (9\text{-}15)$$

For fastest transition times, V_{CC} and R_C should be selected to give an optimum M in accordance with equation 9-10. V_{CC} and R_C affect the repetition rate; the current source in the base also gives some control. However, the repetition rate is primarily determined by the value of the line capacitance and therefore is mainly determined by the pulse width.

By choosing suitable transistors and line impedances, a variety of pulse amplitudes can be obtained. The load, R_L can be varied to obtain different voltage outputs, but if reflections are to be avoided, then R_L must be less than Z_o.

A HIGH-CURRENT PULSE GENERATOR: When high current pulses are required, it is necessary to allow the high output current to flow only during the pulse interval, to keep transistor dissipation low. It is desirable to have a rectangular pulse shape of controllable duration. For a monostable pulse generator, these specifications dictate a constant voltage source, rather than a constant current source, to supply the pulse current, a quasi-stable state during which the pulse amplitude remains constant, and a means of regeneratively terminating the pulse at some controllable time after its initiation. The two-transistor monostable circuit herein described is an attempt to satisfy these requirements. The general operation is similar to the delay-line pulse generator, differing in that a pulse from a second transistor, rather than a pulse reflected from an open ended delay line, is used to terminate the pulse.

Circuit Operation: Although the exact circuit configuration shown in Figure 9-10 is not practical, its basic operation is the same as the practical circuit introduced later and its simplicity warrants its use here. Transistor T_1, bias current I_1, and load resistor R_L comprise a circuit which is used to produce a rectangular high-current output pulse through R_L. Transistor T_2, bias current I_2, capacitors C and C_C and resistor R_L form a resistance-capacitance load circuit used to generate a pulse which is delivered to the base of T_1 to cause it to regeneratively turn off. The circuit operates in the monostable mode and is permanently stable while no pulses are being generated; this is termed the stable state. After being triggered on, the circuit cannot remain in the pulse-producing condition indefinitely but terminates its output and turn-off pulses. The circuit condition during which these pulses are being generated is thus termed the quasi-stable state. The waveforms generated by the circuit and the time periods corresponding to the stable and quasi-stable states are shown in Figure 9-11.

During the stable state, the base-collector junction of T_1 is biased in avalanche by I_1 and the base voltage is thus the avalanche voltage, V_{B1} of T_1, where the voltage I_1R_L is assumed to be negligible compared to V_{B1}. The value of the emitter voltage supply, V_E, is slightly less than V_{B1} thus providing the reverse emitter-base junction bias. A negative trigger, of magnitude larger than $V_{B1} - V_E$, applied to the base forward biases the emitter-base junction and causes the emitter to begin injecting holes into the base. These holes are multiplied by M at the collector junction and the regenerative current build-up occurs.

As in the elementary circuits previously discussed, the turn-on transient and conducting state operating point are described in terms of the collector characteristic curve. Regeneration moves the stable point to approximately $V\alpha_M$ generating an output voltage across R_L of approximately $V_{E1} - V\alpha_M$, thus the load current is:

$$I_P = \frac{V_{E1} - V\alpha_{M1}}{R_L} \tag{9-16}$$

where $V\alpha_{M1}$ is the $V\alpha_M$ of T_1.

During the stable state prior to turn-on, the collector of T_1 is above ground only by the amount of the negligible voltage I_1R_L. The emitter of T_2 is connected to the collector of T_1 through R and is thus also essentially at ground potential. The base of T_2 is at V_{B2} volts, and the base collector junction is biased in avalanche by current I_2. When T_1 is triggered on and the output pulse of $V_{E1} - V\alpha_{M1}$ volts appears across R_L, C begins charging exponentially toward $V_{E1} - V\alpha_{M1}$ with time constant RC. The bias voltage V_{B2} is less than $V_{E1} - V\alpha_{M1}$ and, therefore, the emitter-base junction of T_2 will become forward biased when C charges to V_{B2} volts. At this time, holes are injected from the emitter into the base of T_2. These holes are multiplied by M at the avalanche biased collector junction, thus initiating the regenerative current build-up.

The pulse generated by T_2 is essentially that of the elementary R-C load circuit previously mentioned where the resistive part of the load is R_L plus the bulk resistance of the transistor and the capacitive part is the series combination of C_C and C. This pulse is coupled through C_C to the base of T_1 and causes regenerative turn-off of T_1. As C discharges, the turn-off pulse current decreases, the magnitude of the incremental negative resistance of T_2 increases, and in time the transistor regeneratively turns off. When the turn-off pulse is terminated,

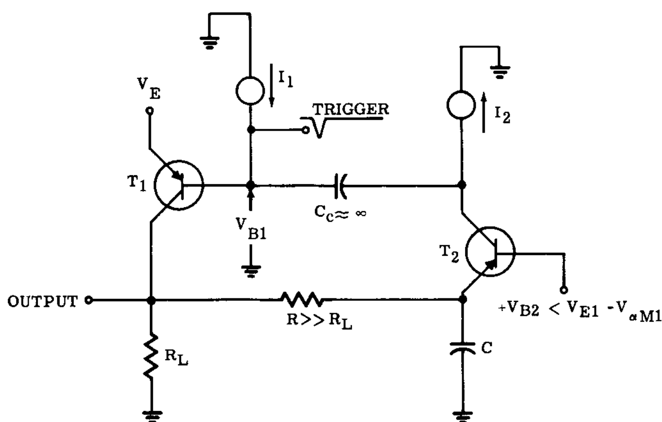


Figure 9-10 Basic Circuit of a High-Current Pulse Generator

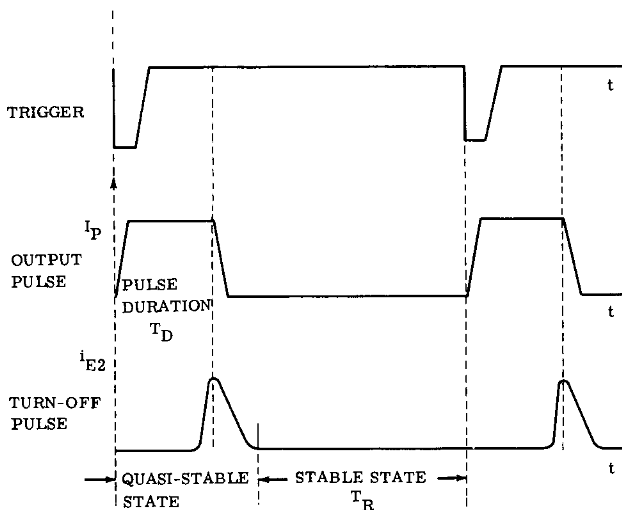


Figure 9-11 Pulse Waveforms for High-Current Pulse Generator

both transistors are in the non-conducting state, the circuit is in the stable state, and one cycle of operation has been completed.

The pulse duration is the time required by C to charge to V_{B2} volts plus the time required by the turn-off pulse to initiate turn-off of T_1. The latter of these two times is generally small compared to the first, and, therefore, the pulse duration, denoted T_D can be calculated as the charging time of C from zero to V_{B2} volts. The initial voltage is zero and the voltage C is charging to $V_{E1} - V\alpha_{M1}$. Therefore,

$$T_D = RC \ln \frac{V_{E1} - V\alpha_{M1}}{V_{E1} - V\alpha_{M1} - V_{B2}} \qquad (9\text{-}17)$$

A Practical Circuit Configuration: A practical circuit configuration is shown in Figure 9-12. The voltage sources E_1 and E_2 with associated resistors R_1, R_2, R_3, and R_4 are considered to be constant current sources of I_1, I_2, I_3, and I_4 respectively. During the stable state the reverse emitter-base junction bias for T_1 is provided by the forward voltage across D_1, due to bias current I_1. The positive voltage bias for the base of T_2 is established by biasing zener diode D_Z in its zener region with current $(I_4 - I_3)$. Capacitor C_1 is sufficiently large so that it appears to be a battery of V_{E1} volts during the generation of the output pulse.

The purpose of inductor L is to prevent the turn-off pulse from being shunted to the emitter node of T_1, and D_2 is included to prevent excessive base current in T_2 if its base voltage should become negative during the generation of the turn-off pulse. It is desirable to isolate the trigger source impedance from the circuit and this is accomplished by components D_t, R_5, and R_6. For biasing purposes, the T_2 base voltage, V_{B2}, which is the zener voltage of D_Z plus the forward drop across D_2, must be between zero and $(V_{E1}-V\alpha_{M1})$ volts. The zener diode D_Z and its bias current $(I_4 - I_3)$ are chosen to provide this bias.

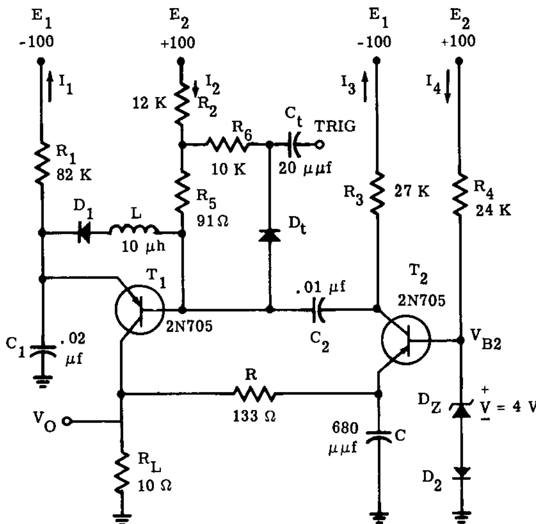


Figure 9-12 High Current Pulse Generator

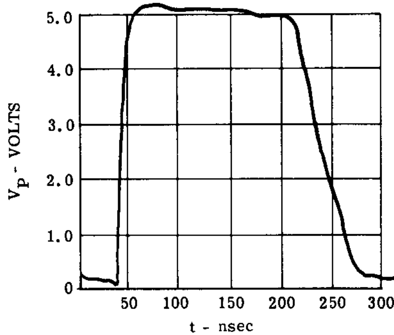


Figure 9-13 Waveform for Circuit of Figure 9-12

The output waveform is shown in Figure 9-13.

Pulse Amplitude, Duration and Maximum Repetition Rate: The maximum attainable pulse amplitude is determined by parameters V_B and $V\alpha_M$ of transistor T_1. If the desired pulse current is of such a magnitude that the increase of $V\alpha_M$ due to bulk resistance and decrease of α at high current is significant, the high current behavior of $V\alpha_M$ must be known to be able to predict the amplitude of the output pulse. The voltage and current amplitudes of the output pulse, denoted V_P and I_P, are:

$$V_P = V_{E1} - V\alpha_{M1} \qquad I_P = \frac{V_P}{R_L}$$

Neglecting the transition time of the output pulse and the rise time of the turn-off pulse, and assuming R_L is small compared to R, the pulse duration is:

$$T_D = RC \ln \frac{V_P}{V_P - V_{B2}}$$

It is seen that the pulse duration is independent of transistor parameters and can be controlled by varying R, C, V_P, or V_{B2}.

There are two interrelated constraints upon the maximum repetition rate at which the circuit can be operated. The first is a basic limitation imposed by the power dissipation capabilities of T_1, and the second is due to the problem of recharging capacitors C_1 and C_2. The rate at which these capacitors can be recharged is also limited by the power dissipation capabilities of T_1. The power dissipation rating of T_1 establishes an upper limit which the designer can attempt to approach by optimizing capacitor recovery time.

During the recovery period, the charge removed from capacitors C_1 and C_2 by the output and turn-off pulses respectively must be replaced. For the specific circuit shown in Figure 9-12 with a 0.5 amp., 0.1μsec output pulse the recovery time limits the maximum repetition rate to between 40 kc and 140 kc, depending on the amount of sacrifice in output pulse amplitude that can be tolerated. Without regard to recovery time, the power dissipation limitation for the same circuit limits the maximum repetition rate to 215 kc.

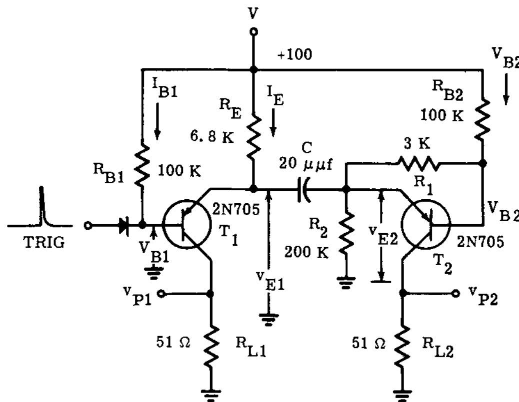


Figure 9-14 Ultra-Fast-Rise Monostable Circuit

A FAST-RISE MONOSTABLE CIRCUIT: When a resistance load is used to determine the peak pulse current, as is the case in the two previous circuits, a relatively long rise-time results because the multiplication factor M is less than 2 during approximately 80 percent of the transient. Large values of M are necessary to produce a rapid transient, and the two-transistor configuration shown in Figure 9-14 is arranged so that during transient periods there is always one transistor operating with high M. The principal function of the transistors is to act as high-speed switches, switching a fixed current between two load resistors R_{L1} and R_{L2}.

Circuit Operation: Supply voltage V and resistors R_{B1}, R_2, R_{B2} and R_E are all large, thus maintaining constant base currents to T_1 and T_2 and a constant current I_E in resistor R_E. For simplicity assume that T_1 and T_2 are identical, and that R_{L1} and R_{L2} are very small so that any voltage appearing across them is negligible. In the stable state, T_1 is conducting emitter current; the emitter and base of T_1 are at $V\alpha_M$ volts above ground. The reverse base current provided to T_2 by R_{B2} causes the collector junction of T_2 to be reverse biased to approximately its breakdown voltage V_B; thus the base of T_2 is V_B volts above ground. R_1 and R_2 are chosen so that the emitter of T_2 is a few tenths of a volt reverse biased.

Assume that a trigger is now applied to the base T_1. The emitter voltage of T_1 rises, and this transient is coupled to T_2 by C. If the trigger is larger than the few tenths of a volt of reverse bias on the emitter of T_2, T_2 conducts emitter current. Since the collector of T_2 is operating near V_B, the multiplication factor M is large and a rapid build-up of current in T_2 occurs. However, the current I_E remains essentially constant in R_E, and thus, as current begins to build up in T_2 it must begin to decrease in T_1. The transient process proceeds very rapidly, and all of the current I_E is switched from T_1 to T_2.

The transient process described above is sufficiently rapid that very little change in voltage occurs across C. As the current builds up in T_2, the emitter-collector voltage of T_2 decreases, and the emitter of T_1 is now $V\alpha_M$ volts above ground. The emitter of T_1 is thus reverse biased by $(V_B - V\alpha_M)$volts. As I_E flows

through C into T_2 the voltage across C increases, and when the emitter voltage of T_1 reaches V_B, injection begins. The M of T_1 is large, and a regenerative build-up of current in T_1 occurs. The current I_E is switched from T_2 back to T_1 and the pulse ends.

Circuit Analysis: The capacitor is charged by the constant current I_E. The pulse time T_D is the time for I_E to change the voltage on C by $(V_B - V\alpha_M)$ volts. Thus, by equating the charges

$$T_D I_E = C(V_B - V\alpha_M)$$

the pulse duration is given by:

$$T_D = (C/I_E)\ (V_B - V\alpha_M) \tag{9-18}$$

Sketches of the waveforms at various points of interest are shown in Figure 9-15, and a plot of the output waveform from a sampling oscilloscope is shown

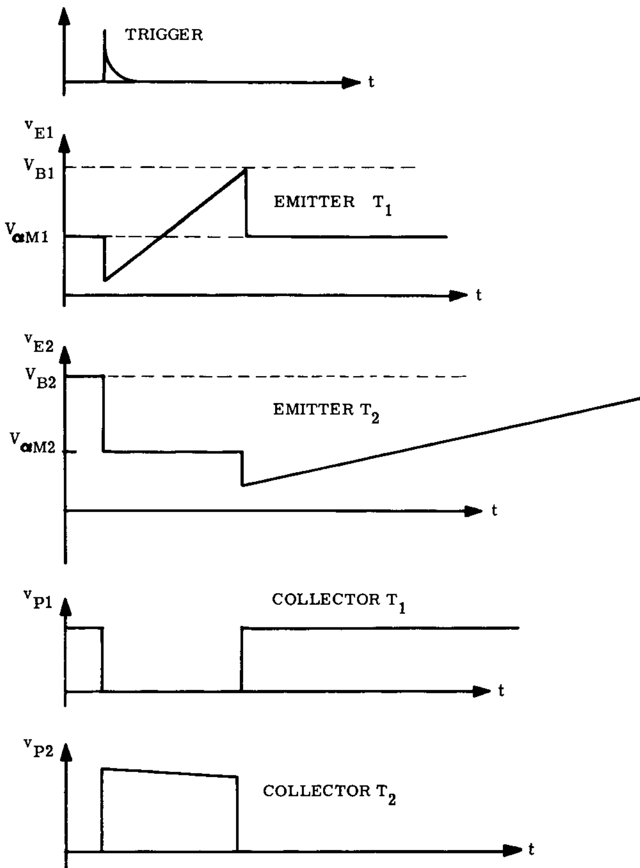


Figure 9-15 Waveforms for the Ultra-Fast Monostable Circuit

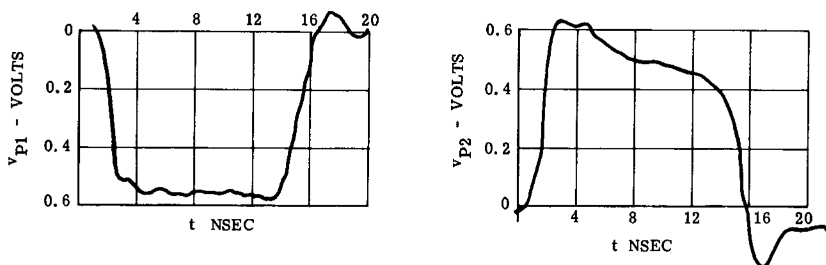


Figure 9-16 Waveforms for the Ultra-Fast Monostable Circuit

in Figure 9-16. Pulses having rise-times of 1nS, fall-times of 2 nS and durations of 12 nS to several microseconds have been obtained. A circuit designed to produce 12 nS pulses has been operated satisfactorily at a repetition rate of 20 Mc.

There are two principal requirements which must be fulfilled in the design of the circuit:

(1) In the stable operating condition, all of the current I_E flows in T_1; the collector voltage of T_1 is $V\alpha_M$. To prevent the transistor dissipation (P_D) from being exceeded, it is necessary that

$$V\alpha_M I_E < P_{D\ max}.$$

For the type 2N705 transistor, this means that the current I_E is limited to a value less than 20mA.

(2) For switching to occur as rapidly as possible, it is necessary that the load resistors be very small. This requirement can be expressed as

$$I_E R_L \ll (V_B - V\alpha_M)$$

In some instances, excessive charge storage in T_1 causes a spike to appear on the leading edge of the positive pulse. This can be eliminated by inserting a small resistance in series with C.

1. S. L. Miller, "Avalanche Breakdown in Germanium", *Physical Review* vol. 99, 1955 pp 1234-1241.
2. D. J. Hamilton, P. G. Griffith, D. C. Latham, "Avalanche Operation of Mesa Transistors" *Engineering Research Laboratories,* College of Engineering, University of Arizona, Tucson, Arizona.
3. D. J. Hamilton "Maximum Avalanche Multiplication in P-N Junctions", *Proceedings of the IRE,* vol 48, Oct 1960, pp 1787-1788.
4. R. B. Seeds "Triggering of Avalanche Transistor Pulse Circuits", Technical Report No. 1653-1, August 5, 1960, *Solid-State Electronics Laboratory,* Stanford Electronics Laboratories, Stanford University, Stanford, California.
5. D. J. Hamilton, F. H. Shaver, P. G. Griffith, "Avalanche Transistor Circuits for Generating Rectangular Pulses," *Electronic Engineering,* December, 1962.
6. F. H. Shaver, "A Fast-Rising, High-Current Pulse Generator," M. S. thesis, Dept. of Electrical Engineering, The University of Arizona, Tucson, Arizona, 1961.

FURTHER READING

AVALANCHE BREAKDOWN

K. F. McKay and K. B. McAfee, "Electron Multiplication in Silicon and Germanium," Phys. Rev., vol. 91, September 1, 1953, pp. 1079-1084.

S. L. Miller and J. J. Ebers, "Alloy Junction Avalanche Transistors," Bell System Technical Journal, vol. 34, September 1955, pp. 883-902.

W. Shockley and J. Gibbons, "Current Build-up in Semiconductor Devices," Proc. IRE, vol. 6, December 1958, pp. 1947-1949.

MODELS AND CIRCUITS

G. B. B. Chaplin, "A Method of Designing Transistor avalanche Circuits with Applications to a Sensitive Transistor Oscilloscope", paper presented at the 1958 IRE-AIEE Solid State Circuits Conference, Philadelphia, Penn., February 1958.

D. S. Gage, "A Study of Avalanche Transistors," Tech. Report No. 36, Stanford Electronics Laboratories, Stanford University, June 13, 1958.

SPECIFICATIONS and TRANSISTOR SELECTION INFORMATION

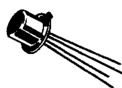
- **Condensed Transistor Specifications**
- **Characteristic Curves for Device Comparison**
- **Complete Data Sheets:**

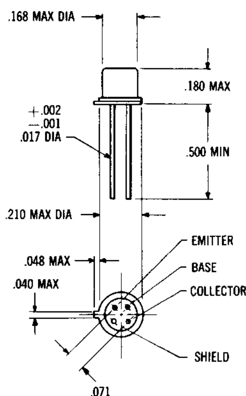
  2N964A

  2N2256 thru 2N2259

  2N2501

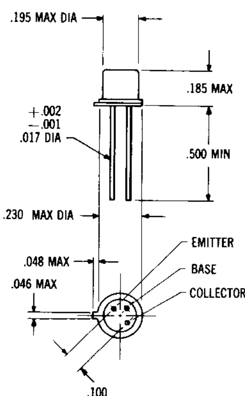
  2N2217 thru 2N2222





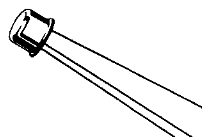
CASE 21

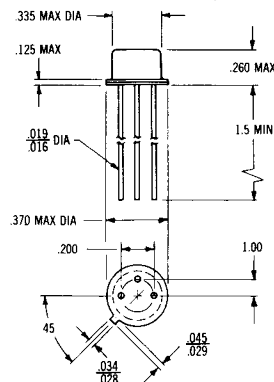
TO-17
PACKAGE



CASE 22

TO-18
PACKAGE





CASE 31

TO-5
PACKAGE

SILICON EPITAXIAL PASSIVATED NPN

		MAXIMUM RATINGS				ELECTRICAL CHARACTERISTICS @ 25°C				
Type	Case	P_D Ambient mW	T_J °C	V_{CB} volts	V_{EB} volts	typ $V_{CE(sat)}$ ($I_C/I_B = 10$) volts	@ I_C mA	h_{FE} min/max	@ I_C mA	f_T typ mc
2N706	22	300	175	25	3	0.2	10	20/—	10	400
2N706A	22	300	175	25	5	0.2	10	20/60	10	400
2N706B	22	300	175	25	5	0.2	10	20/60	10	400
2N706M(JAN)	22	300	175	25	5	0.2	10	30/120	10	400
2N708	22	360	200	40	5	0.2	10	30/120	10	450
2N744	22	300	200	20	5	0.35 (170°C)	10	40/120	10	450
2N753	22	300	175	25	5	0.18	10	40/120	10	400
2N834	22	300	175	40	5	0.15	10	25/—	10	500
2N835	22	300	175	25	3	0.18	10	20/—	10	450
2N914	22	360	200	40	5	0.4	200	30/120	10	500
2N2481	22	360	200	40	5	0.17	10	40/120	10	450
2N2501	22	360	200	40	6	0.18	10	50/150	10	500
2N697	31	600	175	60	5	0.3	150	40/120	150	300
2N1420	31	600	175	60	5	0.3	150	100/300	150	300
2N2217	31	800	175	60	5	0.2	150	20/60	150	400
2N2218	31	800	175	60	5	0.2	150	40/120	150	400
2N2219	31	800	175	60	5	0.2	150	100/300	150	400
2N2220	22	500	175	60	5	0.2	150	20/60	150	400
2N2221	22	500	175	60	5	0.2	150	40/120	150	400
2N2222	22	500	175	60	5	0.2	150	100/300	150	400
2N2537	31	800	200	60	5	0.35	150	50/150	150	400
2N2538	31	800	200	60	5	0.35	150	100/300	150	400
2N2539	22	500	200	60	5	0.35	150	50/150	150	400
2N2540	22	500	200	60	5	0.35	150	100/300	150	400
						$V_{EC(sat)}$ & V_{off} mV mV	@ I_B μA			
2N2330	31	800	175	30	5	1.0 0.3	200	50/—	10	200
2N2331	22	500	175	30	5	1.0 0.3	200	50/—	10	200

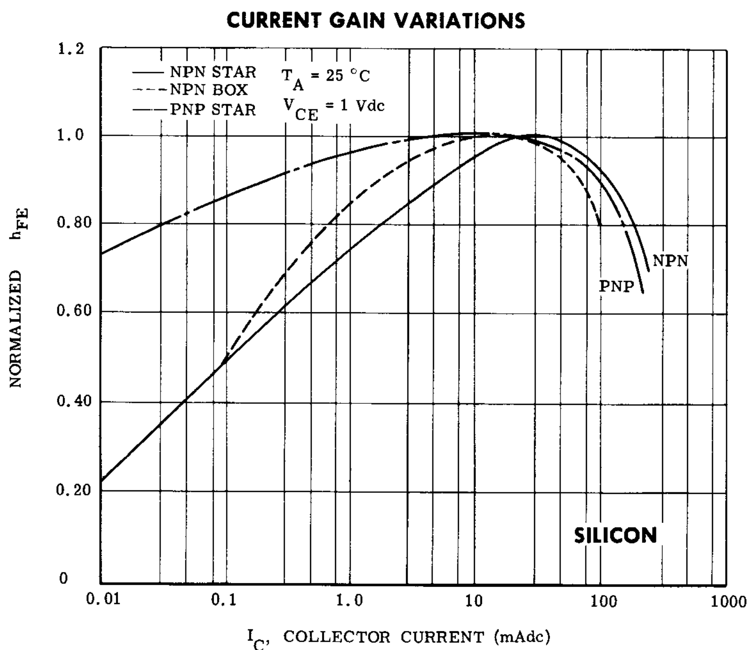
SILICON EPITAXIAL "BAND-GUARD" PNP

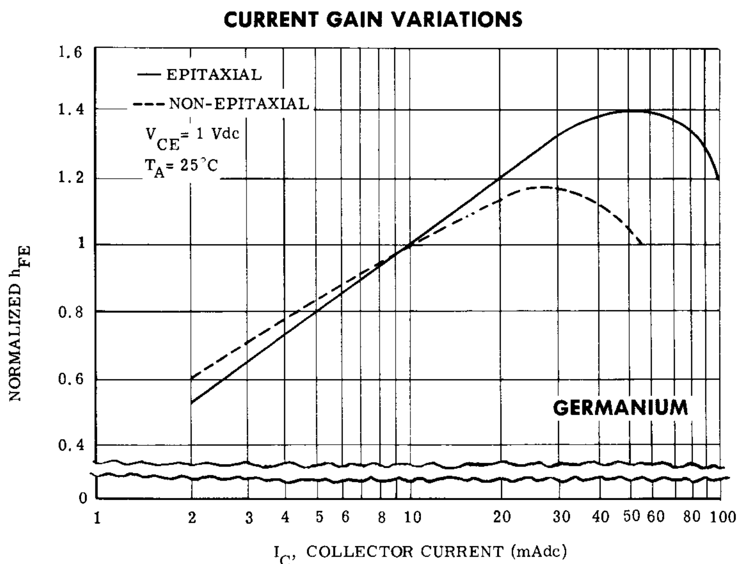
		MAXIMUM RATINGS				ELECTRICAL CHARACTERISTICS @ 25°C				
						typ $V_{CE(sat)}$ @ I_C ($I_C/I_B = 10$)		h_{FE} @ I_C		f_T min
Type	Case	P_D Ambient mW	T_J °C	V_{CB} volts	V_{EB} volts	volts	mA	min/max	mA	mc
2N722	22	400	175	50	5	1.5	150	30/90	150	60
2N1132	31	600	175	50	5	1.5	150	30/90	150	60
2N1132A	31	600	175	60	5	1.5	150	30/90	150	60
2N1132B	31	600	175	70	6	1.5	150	30/90	150	60
2N2800	31	800	200	50	5			30/90	150	120
2N2801	31	800	200	50	5	1.2 max	500	75/225	150	120
2N2837	22	500	200	50	5	0.4 max	150	30/90	150	120
2N2838	22	500	200	50	5			75/225	150	120
2N2904	31	600	200	60	5			40/120	150	200
2N2904A	31	600	200	60	5			40/120	150	200
2N2905	31	600	200	60	5			100/300	150	200
2N2905A	31	600	200	60	5	0.4 max	150	100/300	150	200
2N2906	22	400	200	60	5	1.6 max	500	40/120	150	200
2N2906A	22	400	200	60	5			40/120	150	200
2N2907	22	400	200	60	5			100/300	150	200
2N2907A	22	400	200	60	5			100/300	150	200
COMPLEMENTARY MESA SWITCHING						V_T @ I_C				
						volts	μA	typ		typ
2N2256 Si	22	300	175	7	1	0.5	200	30	10	320
2N2257 Si	22	300	175	7	1	0.5	200	50	10	320
2N2258 Ge	22	150	100	7	1	0.1	200	30	10	320
2N2259 Ge	22	150	100	7	1	0.1	200	50	10	320
GERMANIUM MESA PNP						typ $V_{CE(sat)}$ @ I_C ($I_C/I_B = 10$)				
						volts	mA			
2N695	21	75	100	15	3.5	0.18	10	40	10	250
2N705	22	150	100	15	3.5	0.18	10	40	10	325
2N705 USN	22	150	100	15	3.5	0.18	10	40	10	325
2N710	22	150	100	15	2.0	0.2	10	40	10	325
2N711	22	150	100	12	1.0	0.2	10	30	10	300
2N711A	22	150	100	15	1.5	0.2	10	40	10	320
2N711B	22	150	100	18	2.0	0.18	10	50	10	320
2N827	22	150	100	20	4.0	0.16	10*	150	10	350
2N828†	22	150	100	15	2.5	0.12	10	40	10	400
2N828A†	22	150	100	15	2.5	0.35	150	40	10	400
2N829†	22	150	100	15	2.5	0.38	150	80	10	400
2N838†	22	150	100	30	2.5	0.1	10*	70	10	450
2N960†	22	150	100	15	2.5	0.13	10	40	10	460
2N961†	22	150	100	12	2.0	0.13	10	40	10	460
2N962†	22	150	100	12	1.25	0.13	10	40	10	460
2N963†	22	150	100	12	2.0	0.13	10	40	10	460
2N964†	22	150	100	15	2.5	0.11	10	70	10	460
2N964A†	22	150	100	15	2.5	0.1	10	80	10	460
2N965†	22	150	100	12	2.0	0.11	10	70	10	460
2N966†	22	150	100	12	1.25	0.11	10	70	10	460
2N967†	22	150	100	12	2.0	0.11	10	70	10	460
2N968	22	150	100	15	2.5	0.19	10	35	10	320
2N969	22	150	100	12	2.0	0.19	10	35	10	320
2N970	22	150	100	12	1.25	0.19	10	35	10	320
2N971	22	150	100	7	1.25	0.19	10	35	10	320
2N972	22	150	100	15	2.5	0.19	10	75	10	320
2N973	22	150	100	12	2.0	0.19	10	75	10	320
2N974	22	150	100	12	1.25	0.19	10	75	10	320
2N975	22	150	100	7	1.25	0.19	10	75	10	320
2N1204†	31	250	100	20	4.0	0.25	200	30	400	400
2N1494†	25**	300	100	20	4.0	0.24	200	30	400	400
2N1495†	31	250	100	40	4.0	0.24	200	40	200	400
2N1496†	25**	300	100	40	4.0	0.24	200	40	200	400
2N2381†	31	300	100	30	4.0	0.25	200	45	200	400
2N2382†	31	300	100	45	4.0	0.25	200	45	200	400
2N2635†	22	150	100	30	2.5	0.20	50‡	100	50	300
2N2955†	22	150	100	40	3.5	0.20	50	43	50	350
2N2956†	22	150	100	40	3.5	0.16	50	76	50	375
2N2957†	22	150	100	40	3.5	0.13	50	130	50	400

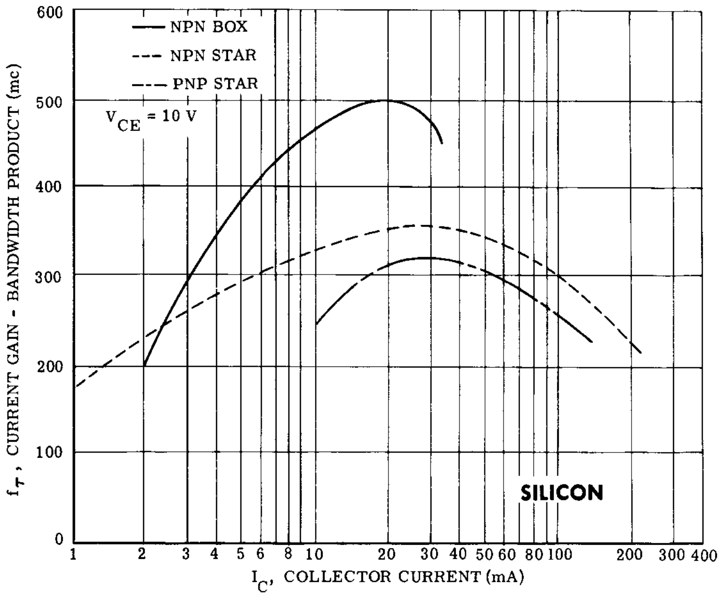
* $\frac{I_C}{I_B} = 3$　　‡ $\frac{I_C}{I_B} = 20$

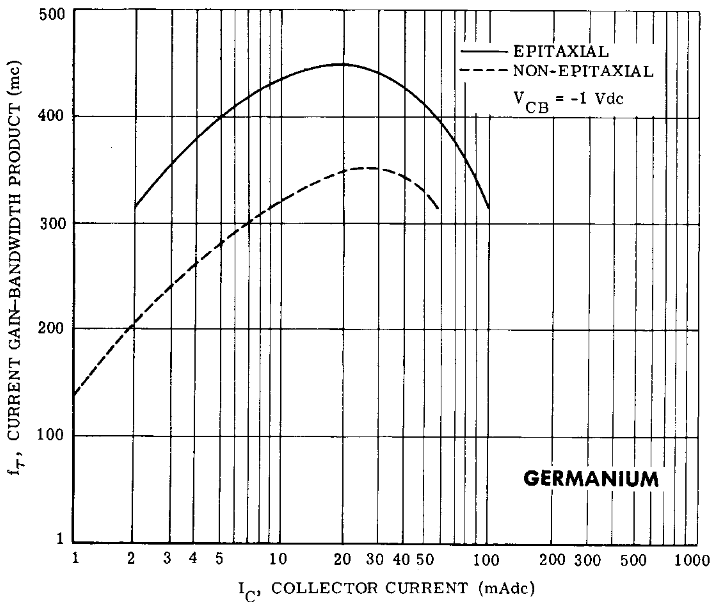
**Stud Mounted TO-5　　†Epitaxial

These curves illustrate the characteristic variations of device parameters due to process and geometry.

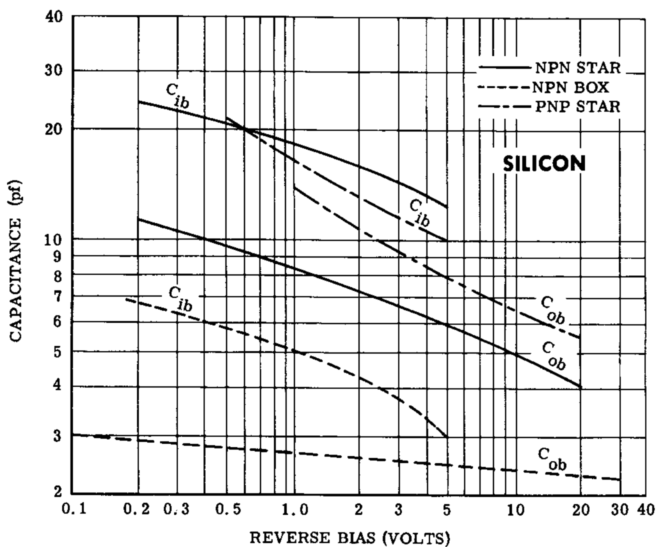




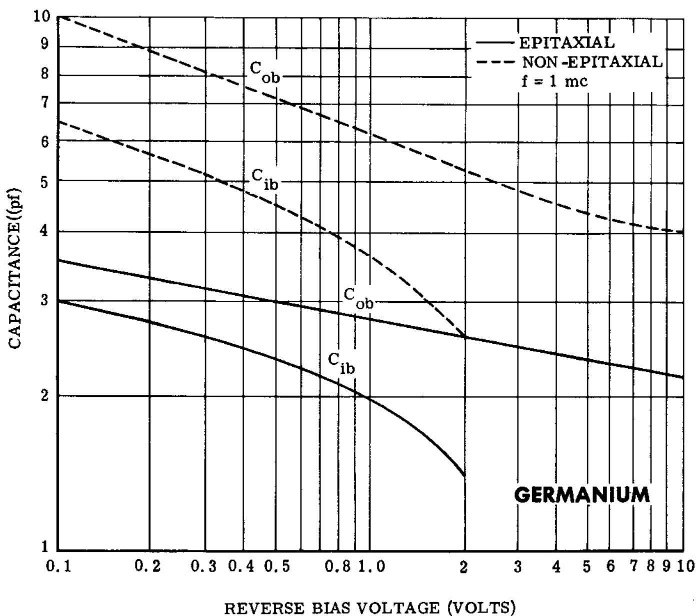

CURRENT GAIN — BANDWIDTH VARIATIONS
NPN BOX
NPN STAR
PNP STAR
V_{CE} = 10 V
SILICON
f_T, CURRENT GAIN - BANDWIDTH PRODUCT (mc)
600
500
400
300
200
100
0
1
2
3
4
5
6
8
10
20
30
40
50
60
80
100
200
300
400
I_C, COLLECTOR CURRENT (mA)


CURRENT GAIN — BANDWIDTH VARIATIONS
EPITAXIAL
NON-EPITAXIAL
V_{CB} = -1 Vdc
GERMANIUM
f_T, CURRENT GAIN-BANDWIDTH PRODUCT (mc)
500
400
300
200
100
1
1
2
3
4
5
10
20
30
40
50
100
200
300
500
1000
I_C, COLLECTOR CURRENT (mAdc)

CAPACITANCE VARIATIONS



CAPACITANCE VARIATIONS



DESIGNERS DATA FOR "WORST CASE" CONDITIONS

The Designers Data Sheet represents a new concept in transistor data sheets permitting the design engineer, in most cases, to design circuits entirely from information presented on the data sheets. In order to do this, the usual *typical* curves, which provided some guidance to the engineer, have been supplemented by *limit* curves which are directly applicable to "worst case" switching circuit design. Using the limit curves and relationships indicated by the equations given in the data sheet, the engineer can design switching circuits that are adequate for "worst case" component tolerances and environmental conditions.

MOTOROLA EPITAXIAL MESA SWITCHING TRANSISTOR

Germanium PNP Diffused Junction

2N964A

for ULTRA-HIGH-SPEED SWITCHING APPLICATIONS

- Low Total Control Charge — 75 Pico-coulombs Maximum @ $I_B = 1_mA$
- High f_T — 300 mc Minimum, 460 mc Typical
- High Current Gain — 40 Minimum From 10 to 100 mA

ABSOLUTE MAXIMUM RATINGS

Characteristic	Symbol	Rating	Unit
Collector-Base Voltage	V_{CB}	15	Vdc
Collector-Emitter Voltage	V_{CE}	15	Vdc
Emitter-Base Voltage	V_{EB}	2.5	Vdc
Collector Current	I_C	100	mAdc
Junction Temperature	T_J	100	°C
Storage Temperature	T_{stg}	−65 to + 100	°C
Total Device Dissipation at 25°C Case Temperature (Derate 4 mW/°C above 25°C)	P_D	300	mW
Total Device Dissipation at 25°C Ambient Temperature (Derate 2 mW/°C above 25°C)	P_D	150	mW

THERMAL RESISTANCE: θ_{JA} (air) $= 0.5\,°C/mW$

θ_{JC} (case) $= 0.25\,°C/mW$

THERMAL TIME CONSTANT: $\tau_{JC} = 10$ msec

2N964A

TABLE 1. ELECTRICAL CHARACTERISTICS (Registered with EIA as the 2N964A) (at 25°C unless otherwise noted)

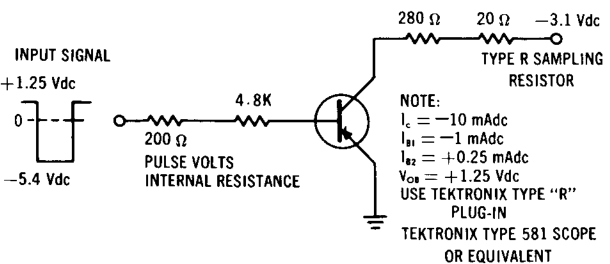
Characteristic	Figure No.	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS						
Collector-Base Breakdown Voltage ($I_C = -100\ \mu$Adc, $I_E = 0$)		BV_{CBO}	15	25	—	Vdc
Collector-Emitter Breakdown Voltage ($I_C = 10$ mAdc, $I_B = 0$)		BV_{CEO}	7	—	—	Vdc
Collector Latch-up Voltage	3	LV_{CEX}	11.5			Vdc
Emitter-Base Breakdown Voltage ($I_E = 100\ \mu$Adc, $I_C = 0$)		BV_{EBO}	2.5	—	—	Vdc
Collector-Emitter Cutoff Current ($V_{CE} = -15$ Vdc, $V_{EB} = 0$)		I_{CES}	—	—	100	μAdc
Collector Cutoff Current ($V_{CB} = -6$ Vdc, $I_E = 0$)		I_{CBO}	—	0.4	3	μAdc
Base Leakage Current	4	I_{BL}				μAdc
($V_{CE} = -6$ Vdc, $V_{OB} = +0.5$ Vdc)			—	—	4	
($V_{CE} = -6$ Vdc, $V_{OB} = +0.5$ Vdc, $T_j = 85°C$)			—	50	140	
ON CHARACTERISTICS						
Forward Current Transfer Ratio	8	h_{FE}				—
($I_C = -10$ mAdc, $V_{CE} = -0.3$ Vdc)			40	80	—	
($I_C = -10$ mAdc, $V_{CE} = -0.3$ Vdc, $T_j = -55°C$)			20	45	—	
($I_C = -50$ mAdc, $V_{CE} = -1$ Vdc)			48	105	—	
($I_C = -100$ mAdc, $V_{CE} = -1$ Vdc)			40	95	—	
($I_C = -100$ mAdc, $V_{CE} = -1$ Vdc, $T_j = 85°C$)			35	85	—	
Collector Saturation Voltage	5	$V_{CE(sat)}$				Vdc
($I_C = -10$ mAdc, $I_B = -1$ mAdc)			—	0.1	0.18	
($I_C = -50$ mAdc, $I_B = -5$ mAdc)			—	0.16	0.28	
($I_C = -100$ mAdc, $I_B = -10$ mAdc)			—	0.22	0.4	
Base-Emitter Voltage	6	V_{BE}				Vdc
($I_C = -10$ mAdc, $I_B = -1$ mAdc)			0.3	0.38	0.44	
($I_C = -50$ mAdc, $I_B = -5$ mAdc)			0.4	0.48	0.58	
($I_C = -100$ mAdc, $I_B = -10$ mAdc)			0.4	0.6	0.72	

TRANSIENT CHARACTERISTICS

Output Capacitance ($V_{CB} = -1$ Vdc, $I_E = 0$, f = 1 mc) ($V_{CB} = -10$ Vdc, $I_E = 0$, f = 1 mc)	11	C_{ob}	— —	2.7 2.2	5 4	pf
Input Capacitance ($V_{EB} = 1$ Vdc, $I_C = 0$, f = 100 kc)	11	C_{ib}	—	2	3.5	pf
Small Signal Forward Current Transfer Ratio ($I_C = -20$ mAdc, $V_{CE} = -1$ Vdc, f = 100 mc)		h_{fe}	3.0	4.6		—
Current-Gain — Bandwidth Product ($I_E = -20$ mAdc, $V_{CB} = -1$ Vdc)		f_T	300	460	—	mc
Delay Time Plus Rise Time ($I_C = -10$ mA) ($I_C = -100$ mA)	 1 2	$t_d + t_r$	— —	35 30	50 50	nsec
Storage Time Plus Fall Time ($I_C = -10$ mA) ($I_C = -100$ mA)	 1 2	$t_s + t_f$	— —	60 50	85 85	nsec —
Total Control Charge ($I_C = -10$ mA, $I_B = -1$ mA)	10	Q_T	—	50	75	pico-coulombs
Active Region Time Constant ($I_C = -10$ mA)	9	τ_A	—	0.6	1.5	nsec

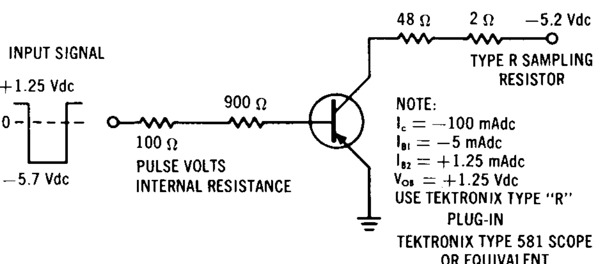
Use only the characteristics listed in Table 1 for incoming inspection testing.

NOTE: See Glossary of Terms on Page 6



10-mA (I_C) SWITCHING TIME TEST CIRCUIT

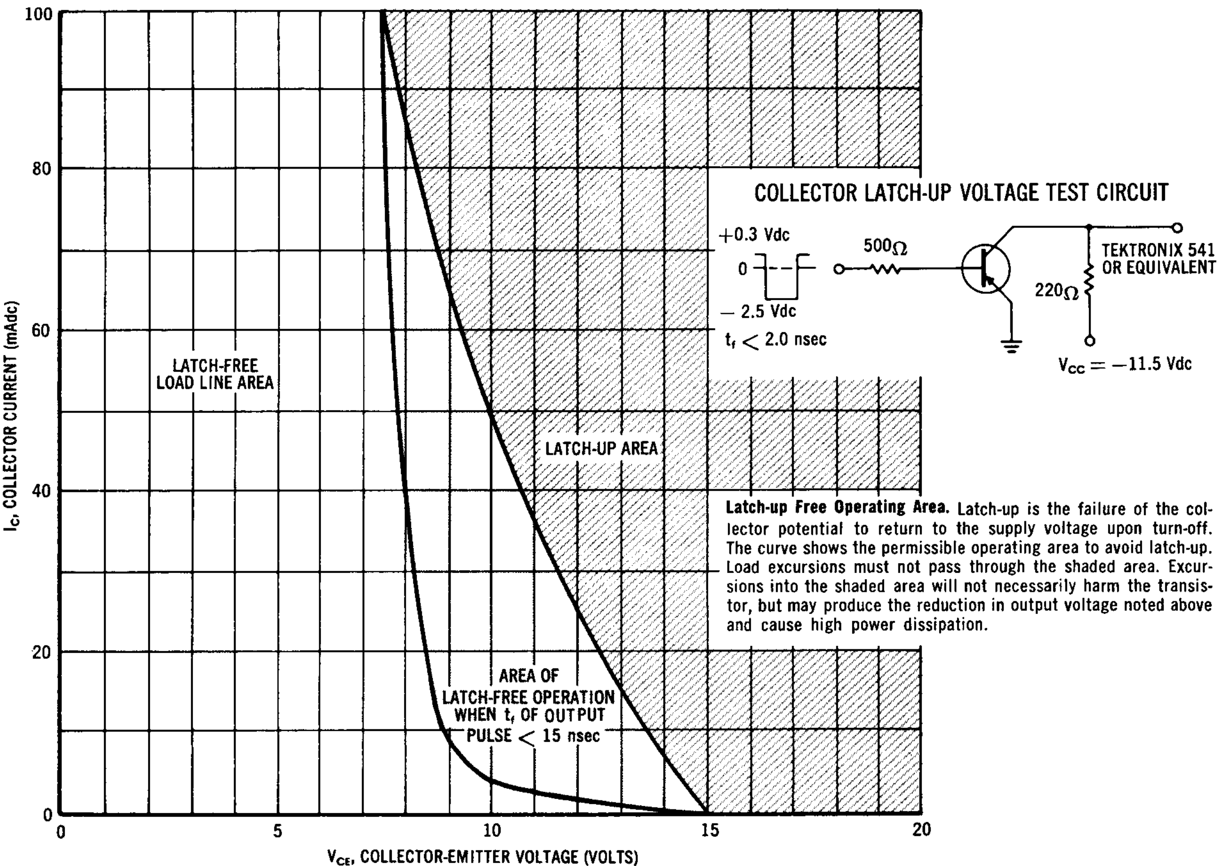
FIGURE 1

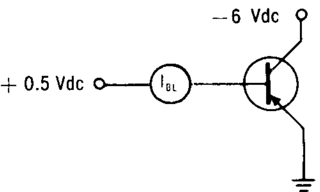


100-mA (I_C) SWITCHING TIME TEST CIRCUIT

FIGURE 2

FIGURE 3—AREA OF PERMISSIBLE LOAD LOCI

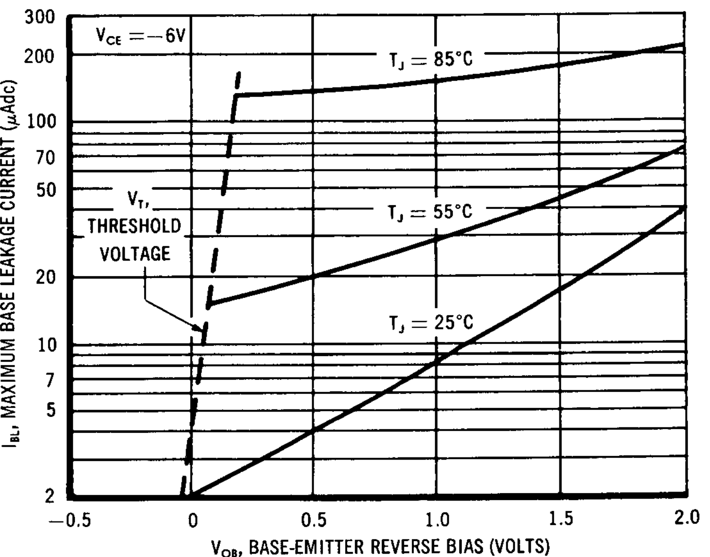




BASE LEAKAGE CURRENT TEST CIRCUIT

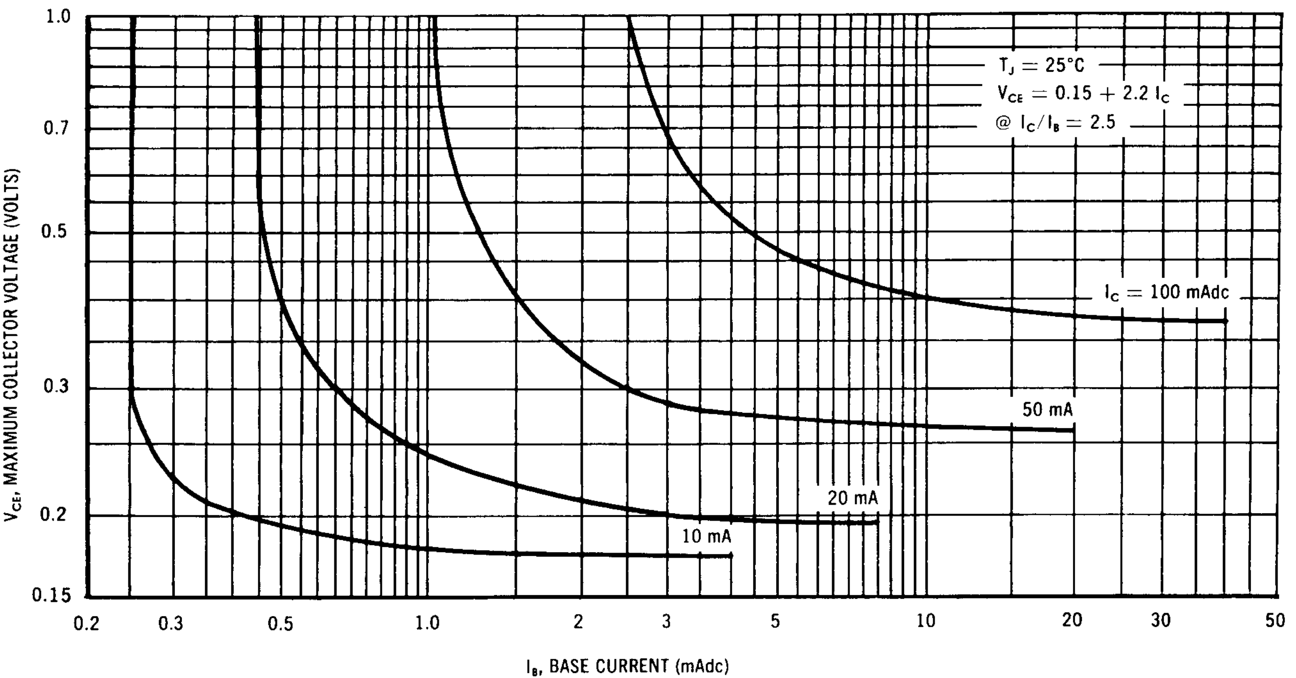
Base Leakage Current. I_{BL} is defined as base leakage current with both junctions reverse biased. I_c is always less than I_{BL} for $V_{OB} > V_T$. (V_{OB} is off condition base bias, V_T is base voltage at threshold of conduction.)

FIGURE 4—COMMON EMITTER DC LEAKAGE CHARACTERISTICS



NOTE: Limit Curves are based on periodic engineering evaluation. Production Tests are made at points indicated in Table 1.

FIGURE 5—COLLECTOR-EMITTER SATURATION VOLTAGE versus BASE CURRENT



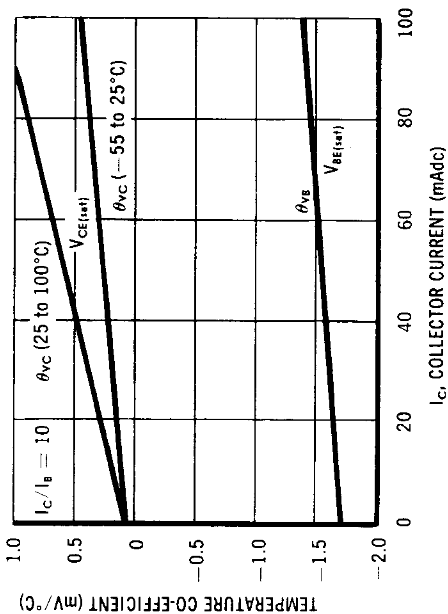


FIGURE 7—TEMPERATURE CO-EFFICIENTS

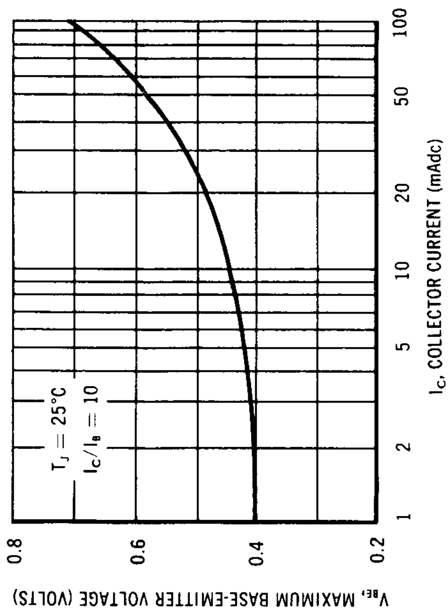


FIGURE 6—BASE-EMITTER VOLTAGE versus COLLECTOR CURRENT

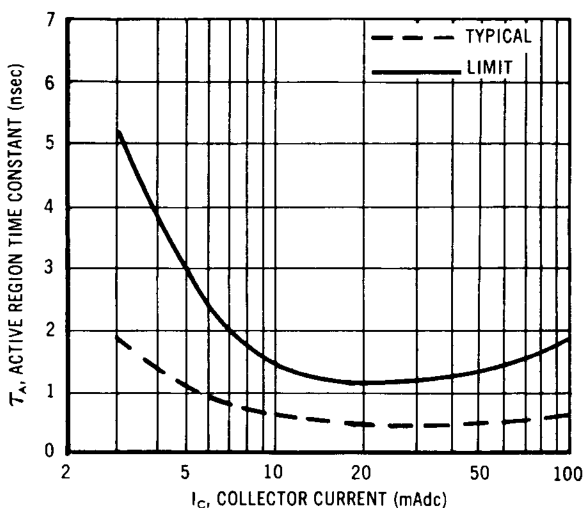
FIGURE 8—CURRENT GAIN CHARACTERISTICS

$V_{CE} = 1$ Volt

$T_J = 85°C$
$T_J = 25°C$
$T_J = 0°C$
$T_J = -55°C$

h_{FE}, MINIMUM CURRENT GAIN: 60, 50, 40, 30, 20, 10

I_C, COLLECTOR CURRENT (mAdc): 2, 5, 10, 20, 50, 100


9—ACTIVE REGION TIME CONSTANT
TYPICAL
LIMIT
τA, ACTIVE REGION TIME CONSTANT (nsec)
7
6
5
4
3
2
1
0
2
5
10
20
50
100
IC, COLLECTOR CURRENT (mAdc)

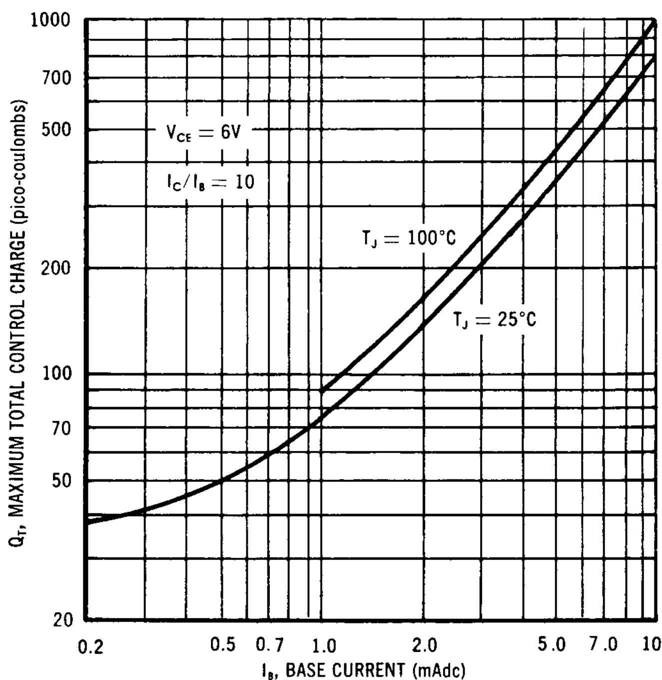
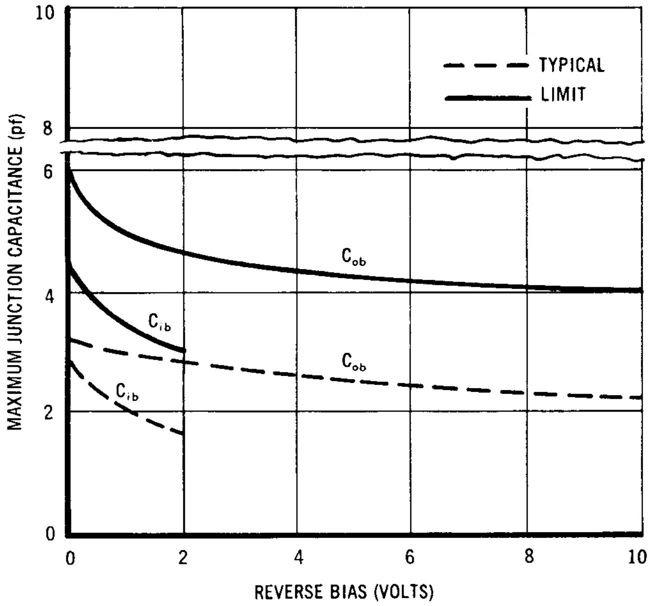


FIGURE 10—TOTAL CONTROL CHARGE
QT, MAXIMUM TOTAL CONTROL CHARGE (pico-coulombs)
1000
700
500
200
100
70
50
20
VCE = 6V
IC/IB = 10
TJ = 100°C
TJ = 25°C
0.2
0.5
0.7
1.0
2.0
5.0
7.0
10
IB, BASE CURRENT (mAdc)

FIGURE 11— JUNCTION CAPACITANCE VARIATIONS



MOTOROLA MESA COMPLEMENTARY SWITCHING TRANSISTORS

Silicon NPN

2N2256, 2N2257

Germanium PNP

2N2258, 22N259

for ULTRA HIGH SPEED NON-SATURATED SWITCHING

- Extremely Fast t_{on} — 3 nsec Typical
- Extremely Fast t_{off} — 4 nsec Typical
- High f_T — 250 mc Minimum

ABSOLUTE MAXIMUM RATINGS

Characteristic	Symbol	2N2256 2N2257	2N2258 2N2259	Unit
Collector-Base Voltage	V_{CB}	7	7	Vdc
Collector-Emitter Voltage	V_{CE}	7	7	Vdc
Emitter-Base Voltage	V_{EB}	1	1	Vdc
DC Collector Current	I_C	100	100	mAdc
Storage Temperature	T_{stg}	−65 to + 175	−65 to + 100	°C
Junction Temperature	T_J	+ 175	+ 100	°C
Device Dissipation at 25°C Case	P_D	1000	300	mW
Derating factor above 25°C		6.67	4	mW/°C
Device Dissipation at 25°C Ambient	P_D	300	150	mW
Derating factor above 25°C		2	2	mW/°C

TRANSISTOR SELECTION CHART

	TYPE		h_{FE} @ $I_C = 25$ mA	
TYPE	NPN	PNP	20	40
2N2256	X		X	
2N2257	X			X
2N2258		X	X	
2N2259		X		X

TABLE I — ELECTRICAL CHARACTERISTICS

(At 25°C unless otherwise noted — All voltages and currents are magnitudes only)

Characteristic	Symbol	Minimum	Typical	Maximum	Unit
Collector-Base Breakdown Voltage	BV_{CBO}				
$I_C = 100\mu Adc$ $I_E = 0$					
ALL TYPES		7	15	—	Vdc
Collector-Emitter Breakdown Voltage	BV_{CES}				
$I_C = 100\mu Adc$ $V_{EB} = 0$					
ALL TYPES		7	15	—	Vdc
Emitter-Base Breakdown Voltage	BV_{EBO}				
$I_E = 100\mu Adc$ $I_C = 0$					
ALL TYPES		1	—	—	Vdc
Collector Cutoff Current	I_{CBO}				
$V_{CB} = 6Vdc$ $I_E = 0$					
ALL TYPES		—	3	10	μAdc
Collector Cutoff Current	I_{CBO}				
$V_{CB} = 6Vdc$ $I_E = 0$ $T_A = 65°C$					
ALL TYPES		—	30	100	μAdc
DC Forward Current Transfer Ratio	h_{FE}				
$I_C = 10mAdc$ $V_{CE} = 1 Vdc$					
2N2256, 2N2258		17	30	—	
2N2257, 2N2259		40	50	—	
$I_C = 25mAdc$ $V_{CE} = 1 Vdc$					
2N2256, 2N2258		20	35	—	
2N2257, 2N2259		40	55	—	
Base-Emitter Voltage	V_{BE}				
$I_C = 10mAdc$ $V_{CE} = 1 Vdc$					
2N2256, 2N2257		—	0.70	0.8	Vdc
2N2258, 2N2259		—	0.35	0.5	Vdc
$I_C = 25mAdc$ $V_{CE} = 1 Vdc$					
2N2256, 2N2257		—	0.8	0.9	Vdc
2N2258, 2N2259		—	0.45	0.6	Vdc
Conduction Threshold Base-Emitter Voltage*	V_T				
$I_C = 200\mu A$ $V_{CE} = 6V$					
2N2256, 2N2257		0.5	—	—	Vdc
2N2258, 2N2259		0.1	—	—	
Collector Output Capacitance	C_{ob}				
$V_{CB} = 5Vdc$ $I_E = 0$ $f = 4mc$					
2N2256, 2N2257		—	4	5	pf
2N2258, 2N2259		—	4	8	pf
Current-Gain — Bandwidth Product	f_T				
$V_{CE} = 1$ V, $I_C = 10$ mA, 2N2256, 2N2257 (Si); $V_{CE} = 15$ V, $I_C = 10$ mA, 2N2258, 2N2259 (Ge)		250	320	—	mc
Turn-on Time	t_{on}				
2N2256, 2N2257 — See Fig. 1		—	3	7	nsec
2N2258, 2N2259 — See Fig. 2		—	4	8	
Turn-off Time	t_{off}				
2N2256, 2N2257 — See Fig. 1		—	4	7	nsec
2N2258, 2N2259 — See Fig. 2		—	3	7	
Base Resistance	r'_b				
$V_{CB} = 2V$ $I_E = 5mA$ $f = 300mc$					
2N2256, 2N2257		—	50	100	ohms
2N2258, 2N2259		—	75	125	

*Base to emitter forward bias voltage at which transistor will be at the threshold of conduction; i.e. that base to emitter voltage at which the collector current is less than or equal to the specified amount under a given collector to emitter voltage condition.

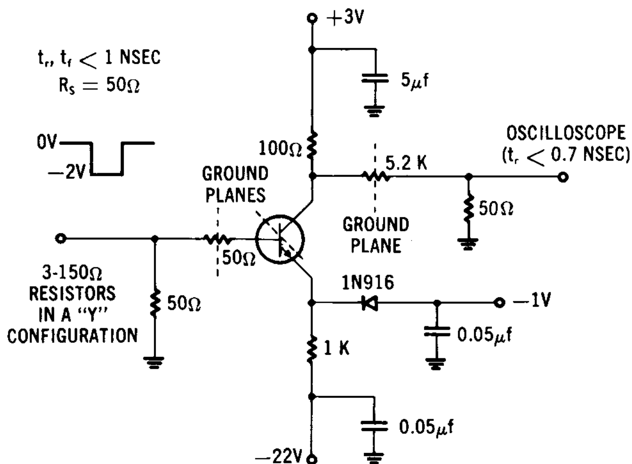


FIGURE 1 — NPN SWITCHING TIME TEST CIRCUIT

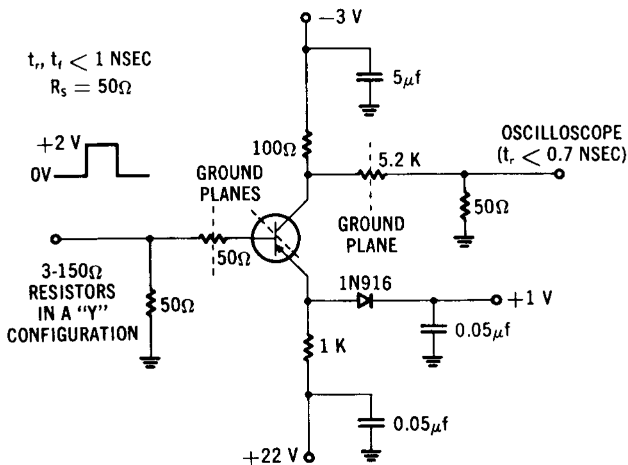


FIGURE 2 — PNP SWITCHING TIME TEST CIRCUIT

FIGURE 3 — CURRENT GAIN CHARACTERISTICS

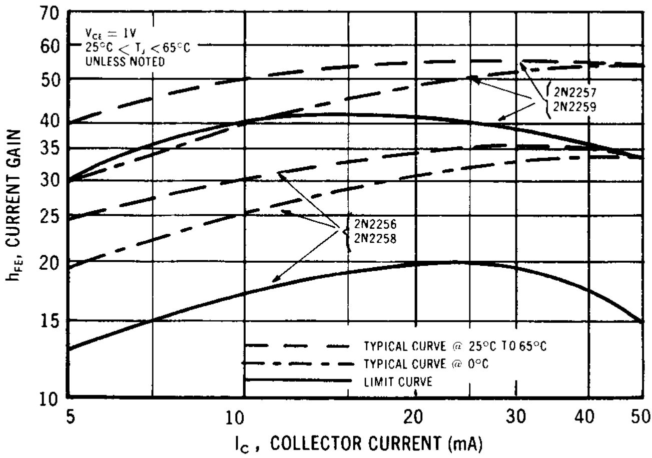


FIGURE 4 — COMMON EMITTER DC INPUT CHARACTERISTICS

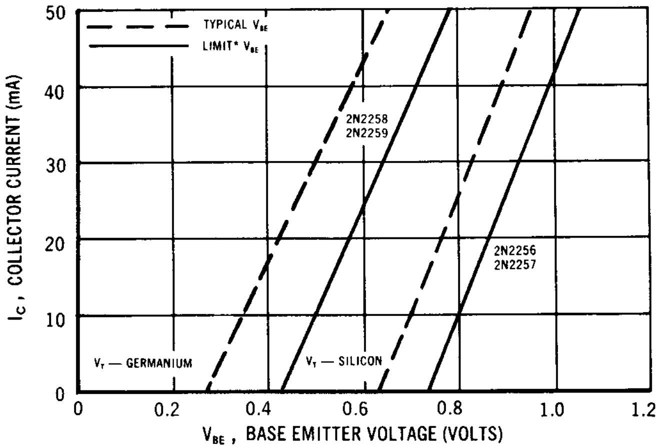


FIGURE 5 — COLLECTOR LEAKAGE CURRENT CHARACTERISTICS

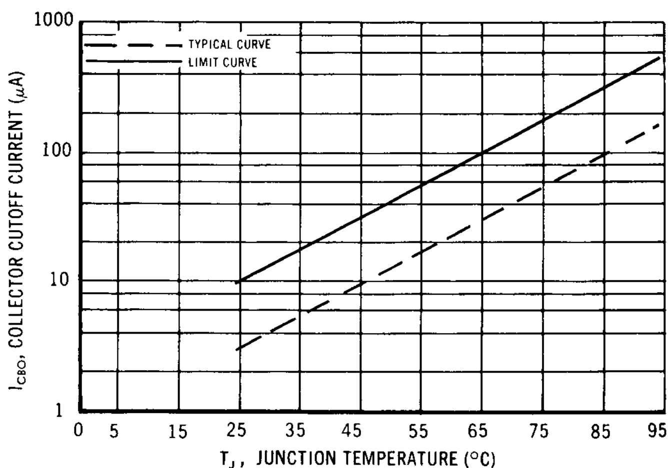


FIGURE 6 — GAIN-BANDWIDTH PRODUCT CHARACTERISTICS

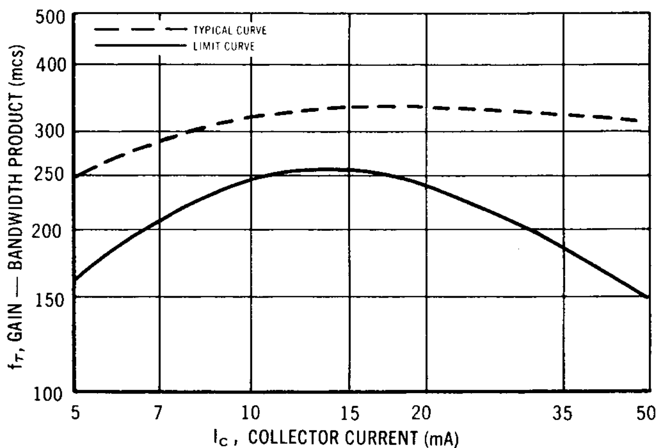


FIGURE 7 — VARIATION OF COLLECTOR CAPACITANCE

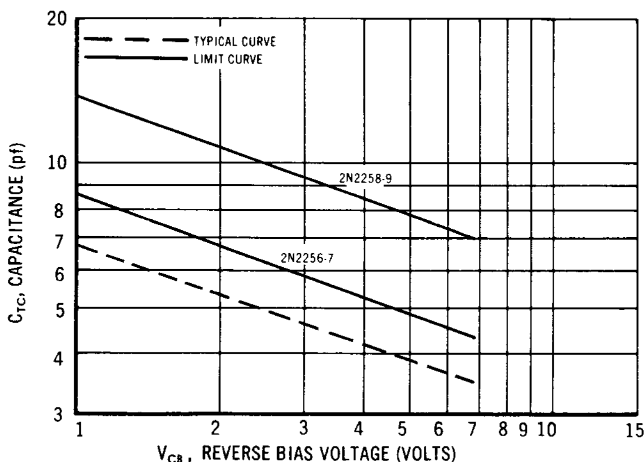
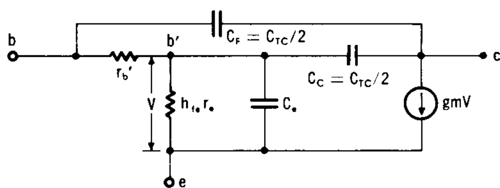


FIGURE 8 — GENERAL EQUIVALENT CIRCUIT (High Frequency, Small Signal)



Approximate values may be obtained or calculated from data as follows:

PARAMETER	OBTAINED	PARAMETER	OBTAINED
r'_b	Table 1	gm	$1/r_e$
h_{fe}	$2\ h_{FE}$ (approx)	C_{TC}	Fig. 7
h_{FE}	Fig. 3	$C_e + C_{TC}$	$\frac{1}{2\pi f_T r_e}$
r_e	$\frac{KT}{qI_E}\left(26\Omega\ @\ \begin{matrix} I_E = 1\text{ mA} \\ T = 25°C \end{matrix}\right)$	f_T	Fig. 6

MOTOROLA EPITAXIAL SWITCHING TRANSISTOR

Silicon NPN Diffused Junction

2N2501

for LOW-LEVEL LOGIC SWITCHING APPLICATIONS

- Low Total Control Charge
  $Q_T = 60$ pico-coulombs max @ $I_C = 10mA$, $I_B = 1mA$
- Guaranteed Active Region Time Constant
  $\tau_A = 2.5$ nsec maximum @ $I_C = 10$ mAdc
- Beta (h_{FE}) specified for a Wide Current Range from 100μ A to 500 mA
  $h_{FE} = 50$ min @ 10 mA

ABSOLUTE MAXIMUM RATINGS

Characteristic	Symbol	Rating	Unit
Collector-Base Voltage	V_{CBO}	40	Vdc
Collector-Emitter Voltage	V_{CEO}	20	Vdc
Emitter-Base Voltage	V_{EBO}	6	Vdc
Total Device Dissipation @ 25°C Ambient Temperature (Derate 2.06 mW/°C above 25°C)	P_D	0.36	Watts
Total Device Dissipation @ 25°C Case Temperature (Derate 6.9 mW/°C above 25°C)	P_D	1.2	Watts
Junction Temperature	T_J	+ 200	°C
Storage Temperature	T_{stg}	−65 to + 300	°C

SWITCHING CHARACTERISTICS

Characteristic	Conditions	Symbol	Maximum	Unit
Charge-Storage Time Constant (Fig. 1)	$I_C = 10$ mAdc $I_{B1} = I_{B2} = 10$ mAdc	τ_S	15	nsec
Total Control Charge (Figure 6)	$I_C = 10$ mAdc $I_{B1} = 1$ mAdc	Q_T	60	pico-coulombs
Active Region Time Constant (Fig. 3)	$I_C = 10$ mAdc	τ_A	2.5	nsec

ELECTRICAL CHARACTERISTICS (At 25°C unless otherwise stated)

Characteristic	Figure No.	Symbol	Minimum	Maximum	Unit
Collector-Base Breakdown Voltage		BV_{CBO}			Vdc
($I_C = 10\ \mu$Adc, $I_E = 0$)			40	—	
Collector-Emitter Breakdown Voltage		BV_{CEO}			Vdc
($I_C = 30$ mAdc, $I_B = 0$, Pulsed)			20	—	
Emitter-Base Breakdown Voltage		BV_{EBO}			Vdc
($I_E = 10\ \mu$Adc, $I_C = 0$)			6	—	
Collector Leakage Current		I_{CEX}			nAdc
($V_{CE} = 20$ Vdc, $V_{BE} = -3$ Vdc)			—	25	
Base Leakage Current	10	I_{BL}			
($V_{CE} = 20$ Vdc, $V_{BE} = -3$ Vdc)			—	25	nAdc
($V_{CE} = 20$ Vdc, $V_{BE} = -3$ Vdc, $T_A = 150°C$)			—	10	μAdc
DC Forward Current Transfer Ratio*	2	h_{FE}			—
($I_C = 100\ \mu$Adc, $V_{CE} = 1$ Vdc)			20	—	
($I_C = 1$ mAdc, $V_{CE} = 1$ Vdc)			30	—	
($I_C = 10$ mAdc, $V_{CE} = 1$ Vdc)			50	150	
($I_C = 10$ mAdc, $V_{CE} = 1$ Vdc, $T_A = -55°C$)			20	—	
($I_C = 50$ mAdc, $V_{CE} = 1$ Vdc)			40	—	
($I_C = 100$ mAdc, $V_{CE} = 1$ Vdc)			30	—	
($I_C = 500$ mAdc, $V_{CE} = 5$ Vdc)			10	—	
Collector-Emitter Saturation Voltage*	7	$V_{CE(sat)}$			Vdc
($I_C = 10$ mAdc, $I_B = 1$ mAdc)			—	0.2	
($I_C = 50$ mAdc, $I_B = 5$ mAdc)			—	0.3	
($I_C = 100$ mAdc, $I_B = 10$ mAdc)			—	0.4	
Base-Emitter Saturation Voltage*	8	$V_{BE(sat)}$			Vdc
($I_C = 10$ mAdc, $I_B = 1$ mAdc)			—	0.85	
($I_C = 50$ mAdc, $I_B = 5$ mAdc)			—	1.0	
$I_C = 100$ mAdc, $I_B = 10$ mAdc)			—	1.2	
Output Capacitance		C_{ob}			pf
($V_{CB} = 10$ Vdc, $I_E = 0$, f = 100 kc)			—	4	
Input Capacitance		C_{ib}			pf
($V_{EB} = 0.5$ Vdc, $I_C = 0$, f = 100 kc)			—	7	
Small Signal Forward Current Transfer Ratio		h_{fe}			—
($V_{CE} = 20$ Vdc, $I_C = 10$ mAdc, f = 100 mc)			3.5	—	
Current-Gain-Bandwith Product		f_T			mc
($V_{CE} = 20$ Vdc, $I_C = 10$ mAdc)			350	—	
Charge Storage Time Constant	1	τ_S			nsec
($I_C = I_{B1} = I_{B2} = 10$ mAdc)			—	15	
Total Control Charge	6	Q_T			pico-coulombs
($I_C = 10$ mAdc, $I_B = 1$ mAdc)			—	60	
Active Region Time Constant	3	τ_A			nsec
($I_C = 10$ mAdc)			—	2.5	

*Pulse Test: Pulse width $\leqq 300\ \mu$sec, duty cycle $\leqq 2\%$

FIGURE 1 — CHARGE STORAGE TIME CONSTANT TEST CIRCUIT

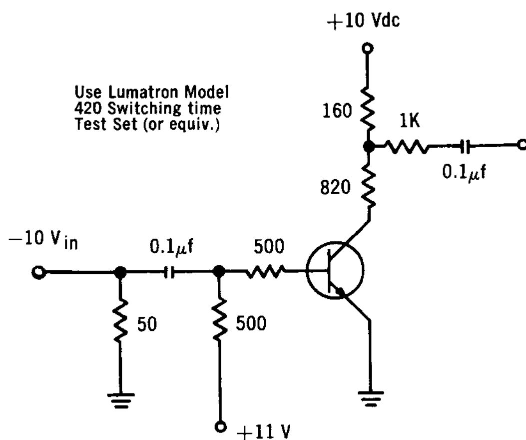


FIGURE 2 — NORMALIZED CURRENT GAIN CHARACTERISTICS

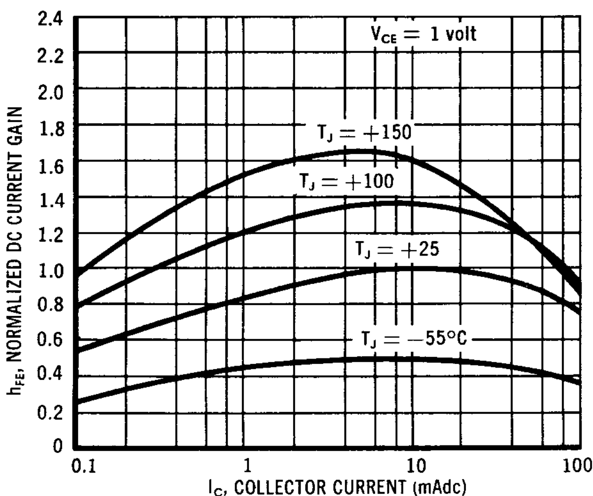
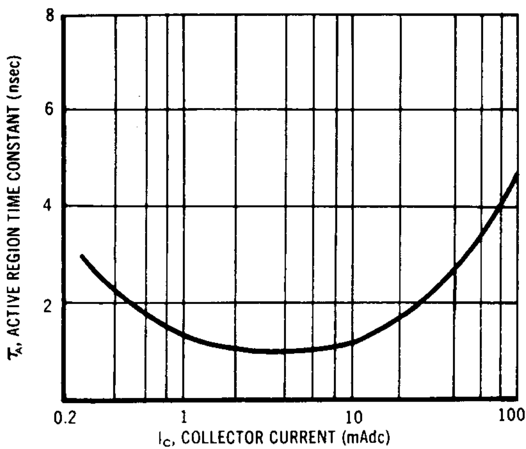


FIGURE 3 — ACTIVE REGION TIME CONSTANT AND TEST CIRCUIT



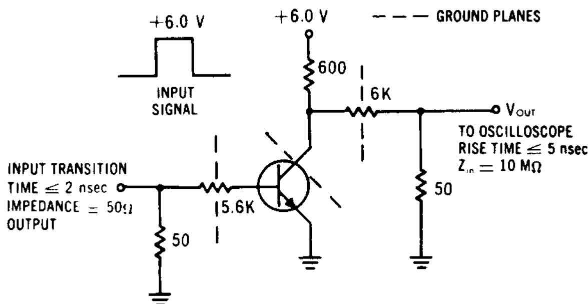
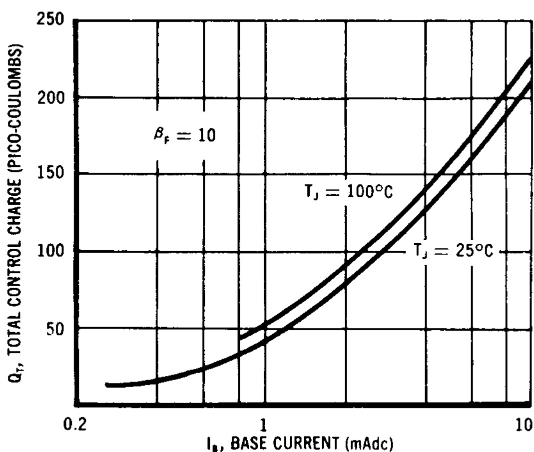


FIGURE 6 — TOTAL CONTROL CHARGE AND TEST CIRCUIT



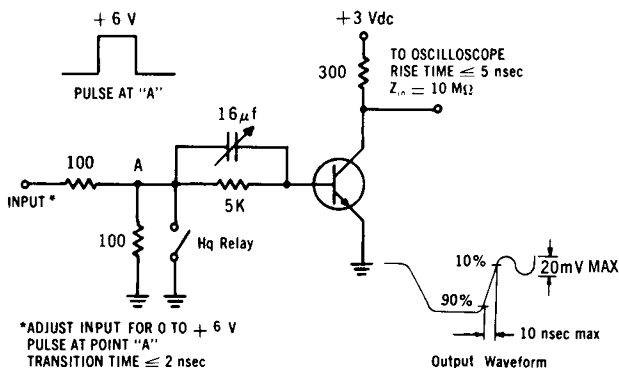


FIGURE 7 — COLLECTOR-EMITTER SATURATION VOLTAGES versus BASE CURRENT

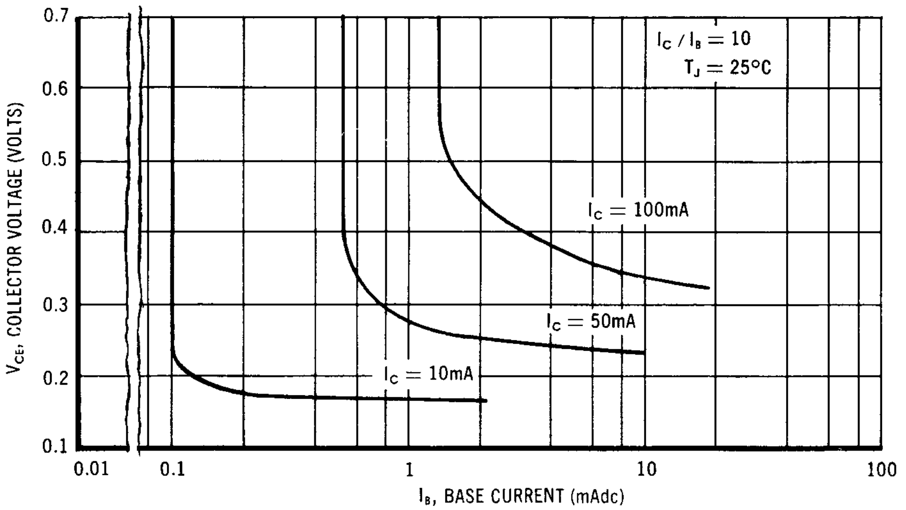


FIGURE 8 — BASE-EMITTER VOLTAGE versus COLLECTOR CURRENT

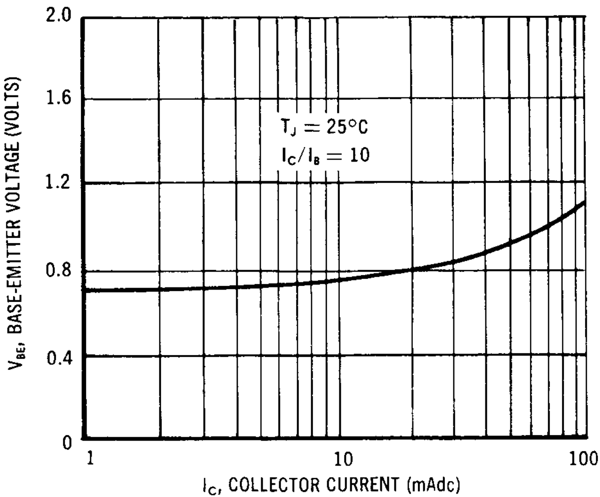


FIGURE 9 — TEMPERATURE CO-EFFICIENTS

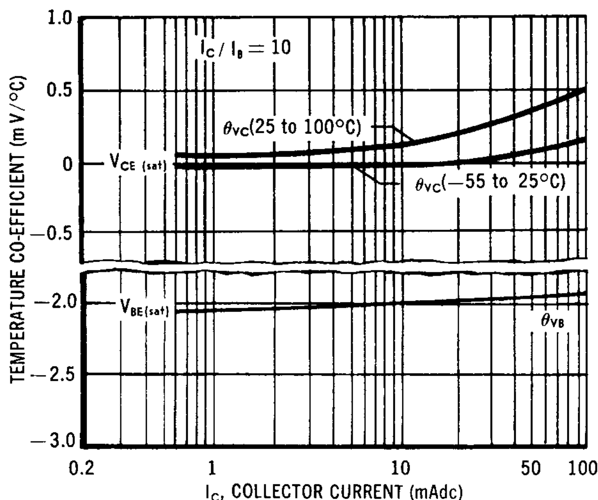
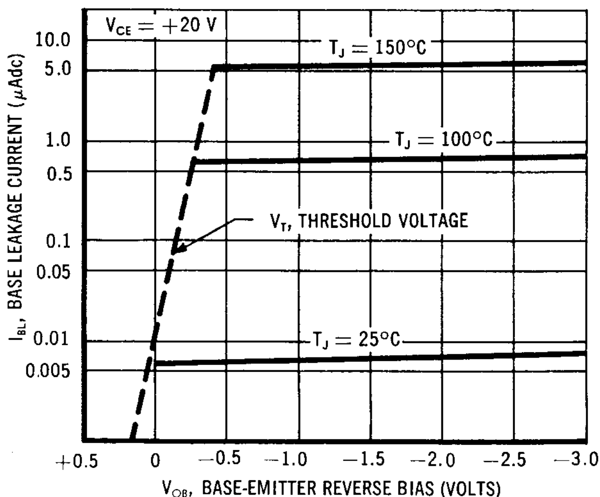
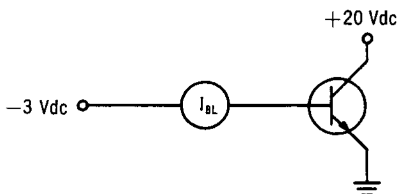


FIGURE 10 — COMMON EMITTER DC LEAKAGE CHARACTERISTICS AND TEST CIRCUIT





Base Leakage Current. I_{BL} is defined as base leakage current with both junctions reverse biased. I_C is always less than I_{BL} for $V_{OB} > V_T$. (V_{OB} is off condition base bias, V_T is base voltage at threshold of conduction.)

MOTOROLA EPITAXIAL STAR TRANSISTORS

Silicon NPN Double Diffused

2N2217 thru 2N2222

Optimized geometry for HIGH-SPEED SWITCHING CIRCUITS and DC to UHF AMPLIFIER APPLICATIONS

ABSOLUTE MAXIMUM RATINGS

Characteristic	Symbol	2N2217-19 (TO-5)	2N2220-22 (TO-18)	Unit
Collector-Base Voltage	V_{CBO}	60	60	Vdc
Collector-Emitter Voltage	V_{CEO}	30	30	Vdc
Emitter-Base Voltage	V_{EBO}	5	5	Vdc
Total Device Dissipation at 25°C Case Temperature Derating Factor above 25°C	P_D	3 20	1.8 12	Watts mW/°C
Total Device Dissipation at 25°C Ambient Temperature Derating Factor Above 25°C	P_D	0.8 5.33	0.5 3.33	Watts mW/°C
Junction Temperature	T_J	−65 to + 175		°C
Storage Temperature	T_{stg}	−65 to + 300		°C

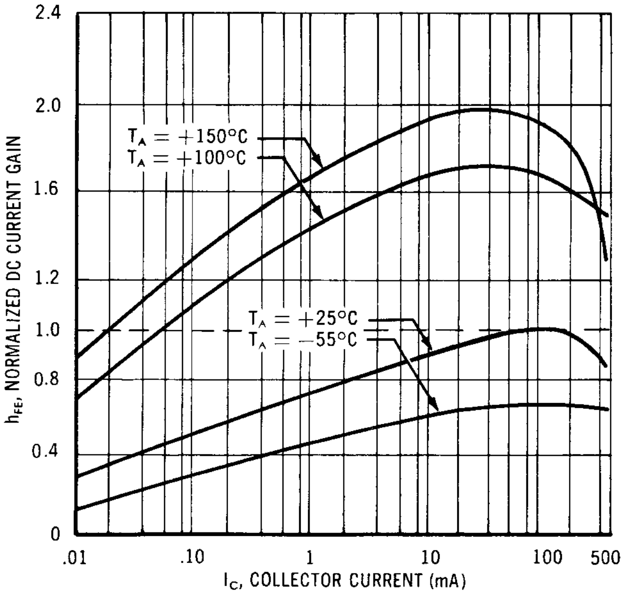
ELECTRICAL CHARACTERISTICS (At 25°C unless otherwise noted)

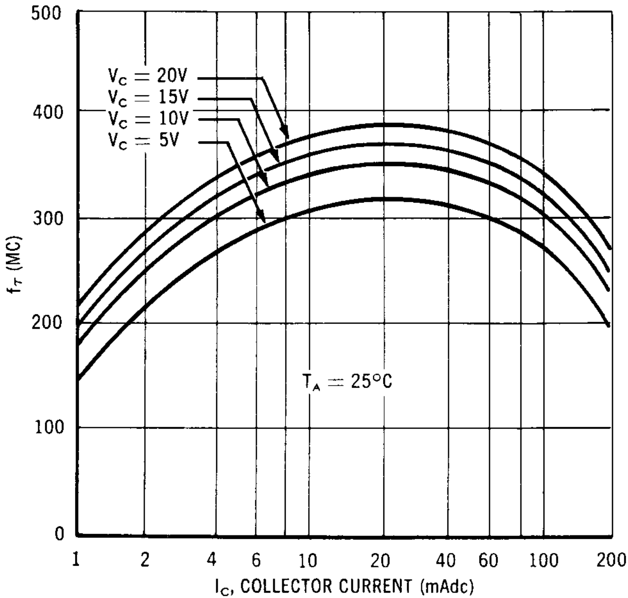
Characteristic	Symbol	Min.	Typ.	Max.	Unit
Collector Cutoff Current (V_{CB} =50 Vdc, I_E = 0)	I_{CBO}	—	.001	.01	μAdc
Collector Cutoff Current (V_{CB} = 50 Vdc, T_A = 150°C)	I_{CBO}	—	—	10	μAdc
Collector-Base Breakdown Voltage (I_C = 10 μAdc, I_E = 0)	BV_{CBO}	60	90	—	Vdc
Collector-Emitter Breakdown Voltage (I_C = 10 mAdc, I_B = 0)	BV_{CEO}	30	45	—	Vdc

ELECTRICAL CHARACTERISTICS (At 25°C unless otherwise stated)

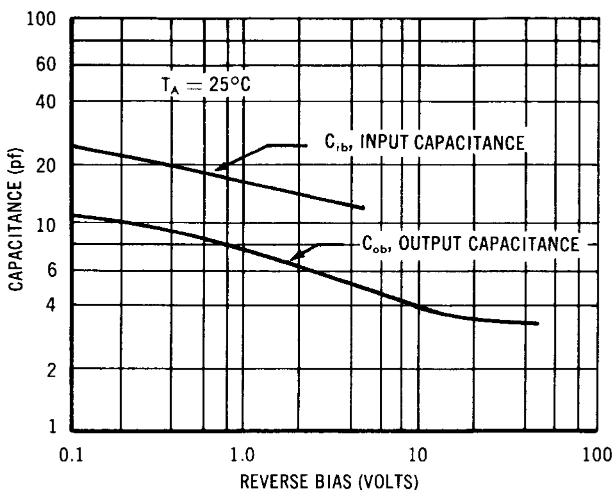
Characteristic		Symbol	Min	Typ	Max	Unit
Emitter-Base Breakdown Voltage		BV_{EBO}				Vdc
($I_E = 10\ \mu$Adc, $I_C = 0$)			5	—	—	
Collector Saturation Voltage		$V_{CE(sat)}$*				
($I_C = 150$ mAdc, $I_B = 15$ mAdc)	All Types		—	0.24	0.4	Vdc
($I_C = 500$ mAdc, $I_B = 50$ mAdc)	2N2218, 2N2219 2N2221, 2N2222		—	0.8	1.6	Vdc
Base-Emitter Saturation Voltage		$V_{BE(sat)}$*				
($I_C = 150$ mAdc, $I_B = 15$ mAdc)	All Types		—	1.0	1.3	Vdc
($I_C = 500$ mAdc, $I_B = 50$ mAdc)	2N2218, 2N2219 2N2221, 2N2222		—	1.5	2.6	Vdc
DC Forward Current Transfer Ratio		h_{FE}				
($I_C = 0.1$ mAdc, $V_{CE} = 10$ Vdc)	2N2218, 2N2221		20	—	—	—
	2N2219, 2N2222		35	—	—	—
($I_C = 1.0$ mAdc, $V_{CE} = 10$ Vdc)	2N2217, 2N2220		12	—	—	—
	2N2218, 2N2221		25	—	—	—
	2N2219, 2N2222		50	—	—	—
($I_C = 10$ mAdc, $V_{CE} = 10$ Vdc)	2N2217, 2N2220		17	—	—	—
	2N2218, 2N2221		35	—	—	—
	2N2219, 2N2222		75	—	—	—
($I_C = 150$ mAdc, $V_{CE} = 10$ Vdc) *	2N2217, 2N2220		20	—	60	—
	2N2218, 2N2221		40	—	120	—
	2N2219, 2N2222		100	—	300	—
($I_C = 500$ mAdc, $V_{CE} = 10$ Vdc) *	2N2218, 2N2221		20	—	—	—
	2N2219, 2N2222		30	—	—	—
Output Capacitance		C_{ob}				pf
$V_{CB} = 10$ Vdc, $I_E = 0$, f = 100KC			—	4	8	
Input Capacitance		C_{ib}				pf
$V_{EB} = 0.5$ Vdc, $I_C = 0$, f = 100KC			—	20	—	
Small Signal Forward Current Transfer Ratio		h_{fe}				—
($V_{CE} = 20$ Vdc, $I_C =$ mAdc, f = 100mc)			2.5	4.0	—	
Current Gain — Bandwidth Product		f_T				mc
($I_C = 20$ mAdc, $V_{CE} = 20$ Vdc)			250	400	—	
Turn-on Time		t_{on}	—	26	—	nsec
Turn-off Time		t_{off}	—	68	—	nsec
Total Switching Time		t_{total}	—	12	—	nsec

**Pulse Test:*
Pulse width $\leqq$ 300 μsec
Duty Cycle $\leqq$ 2%

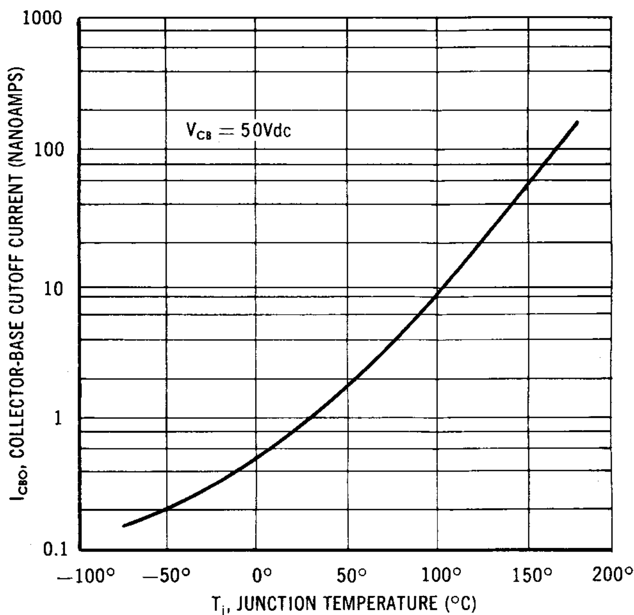

NORMALIZED CURRENT GAIN versus COLLECTOR CURRENT
2.4
2.0
1.6
1.2
1.0
0.8
0.4
0
T_A = +150°C
T_A = +100°C
T_A = +25°C
T_A = −55°C
h_FE, NORMALIZED DC CURRENT GAIN
.01
.10
1
10
100
500
I_C, COLLECTOR CURRENT (mA)


CURRENT GAIN-BANDWIDTH PRODUCT
versus
COLLECTOR CURRENT
500
400
300
200
100
0
V_C = 20V
V_C = 15V
V_C = 10V
V_C = 5V
T_A = 25°C
f_T (MC)
1
2
4
6
10
20
40
60
100
200
I_C, COLLECTOR CURRENT (mAdc)

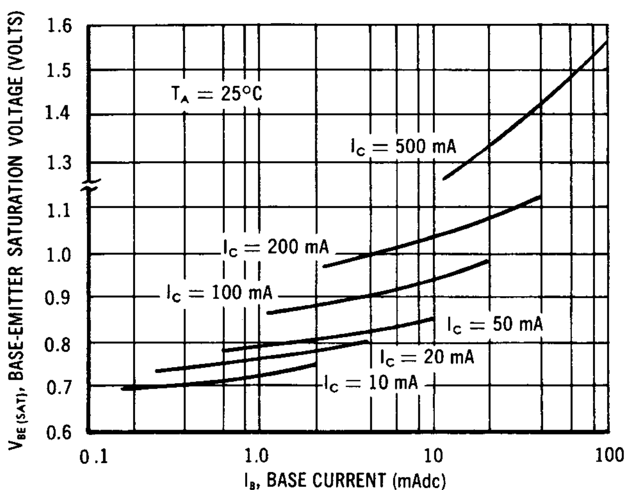
OUTPUT CAPACITANCE versus COLLECTOR-BASE VOLTAGE and INPUT CAPACITANCE versus EMITTER-BASE VOLTAGE



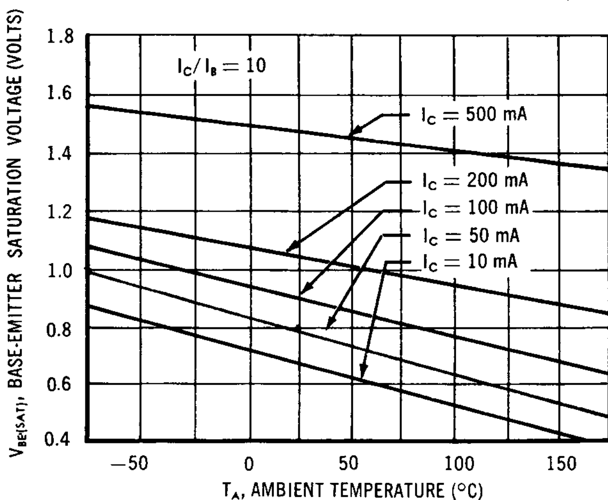
I_{CBO} versus JUNCTION TEMPERATURE

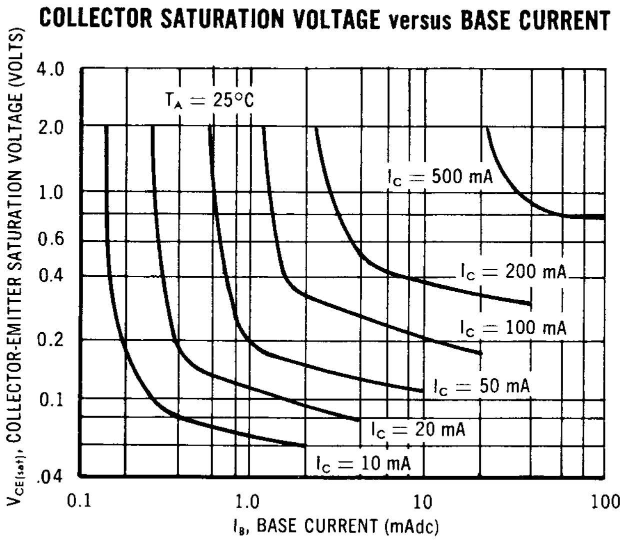


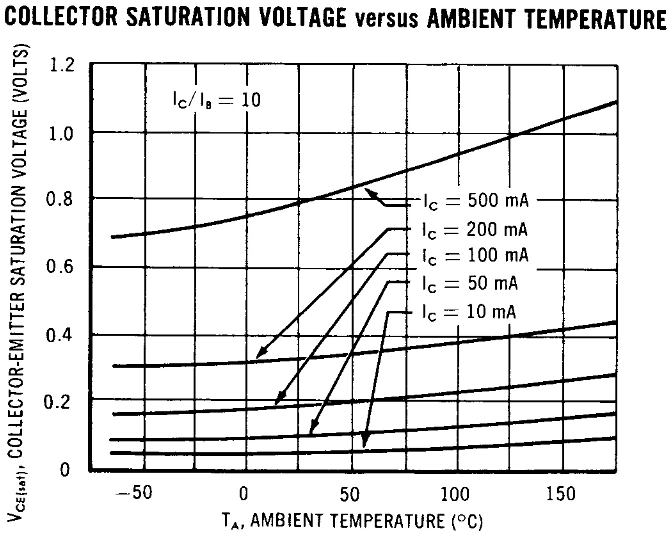
BASE SATURATION VOLTAGE versus BASE CURRENT



BASE SATURATION VOLTAGE versus AMBIENT TEMPERATURE

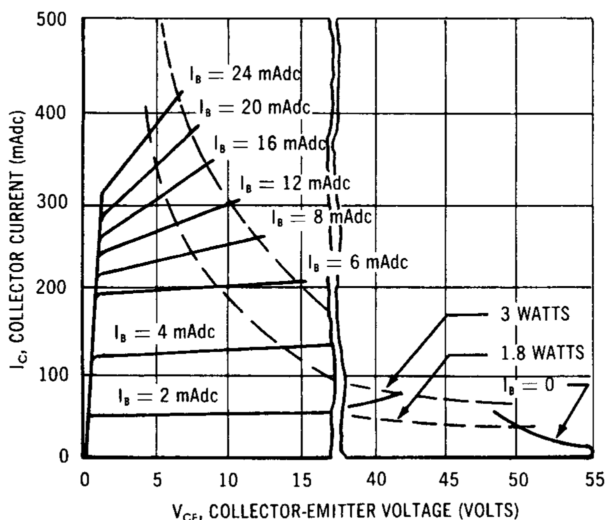


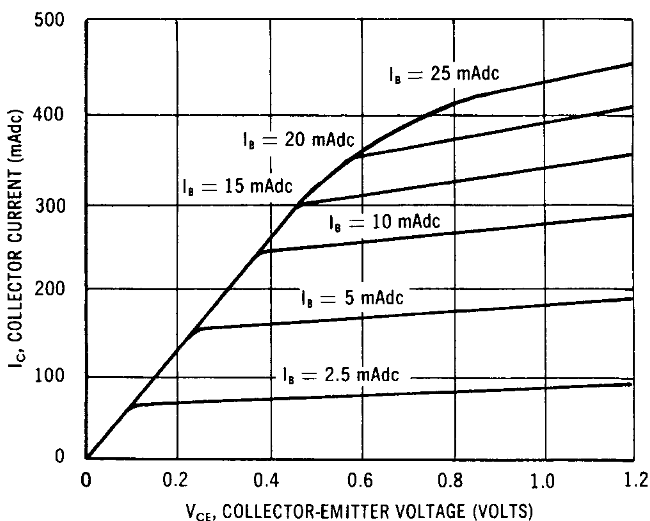

COLLECTOR SATURATION VOLTAGE versus BASE CURRENT
$V_{CE(sat)}$, COLLECTOR-EMITTER SATURATION VOLTAGE (VOLTS)
4.0
2.0
1.0
0.6
0.4
0.2
0.1
.04
0.1
1.0
10
100
I_B, BASE CURRENT (mAdc)
$T_A = 25°C$
$I_C = 500$ mA
$I_C = 200$ mA
$I_C = 100$ mA
$I_C = 50$ mA
$I_C = 20$ mA
$I_C = 10$ mA


COLLECTOR SATURATION VOLTAGE versus AMBIENT TEMPERATURE
$V_{CE(sat)}$, COLLECTOR-EMITTER SATURATION VOLTAGE (VOLTS)
1.2
1.0
0.8
0.6
0.4
0.2
0
−50
0
50
100
150
T_A, AMBIENT TEMPERATURE (°C)
$I_C/I_B = 10$
$I_C = 500$ mA
$I_C = 200$ mA
$I_C = 100$ mA
$I_C = 50$ mA
$I_C = 10$ mA

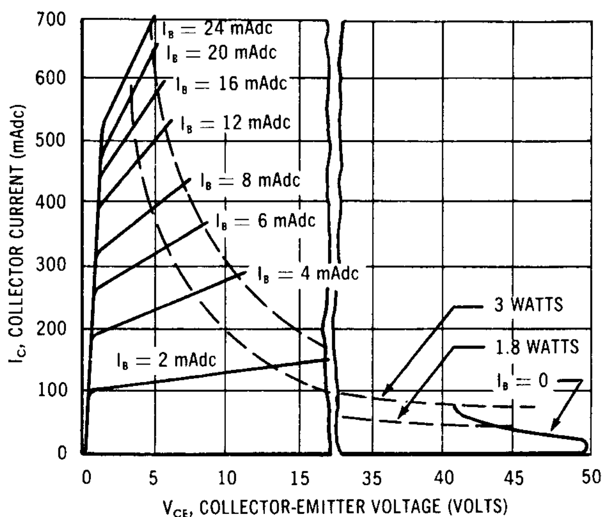
COLLECTOR CHARACTERISTICS COMMON EMITTER CONFIGURATION $T_A = 25°C$

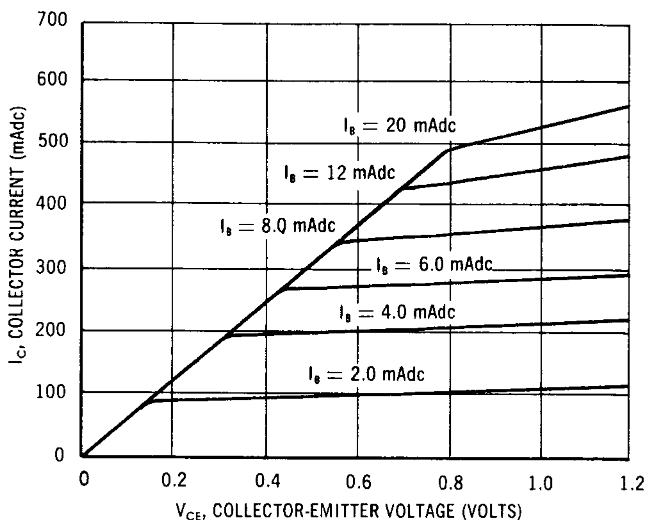
2N2217 and 2N2220



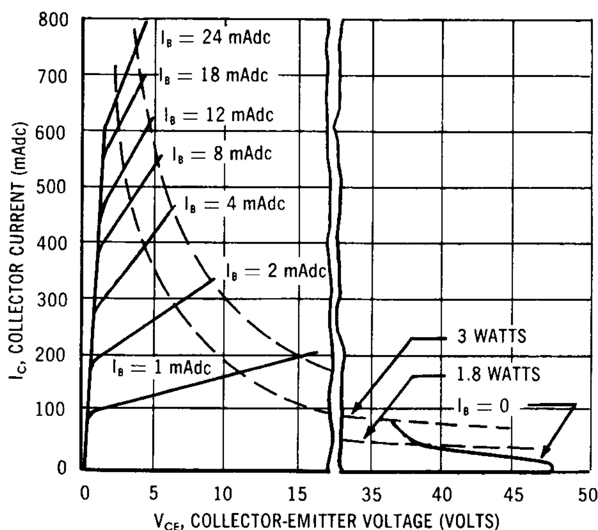


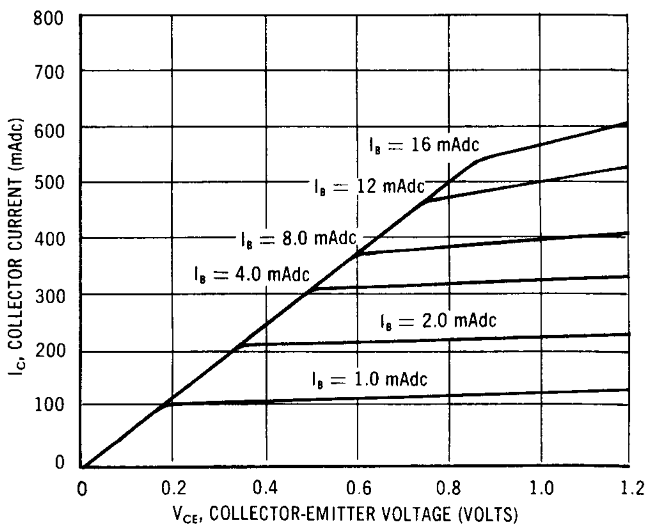
2N2218 and 2N2221





2N2219 and 2N2222





APPENDIX I
THE EBERS & MOLL EQUATIONS

Ebers and Moll found by analyzing carrier flow based upon the diffusion equations, that the voltage-current relationship for a transistor could be described as

$$I_E = -\frac{I_{EO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_E}{kT}} - 1\right) + \frac{\alpha_I I_{CO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_C}{kT}} - 1\right) \quad (1)$$

$$I_C = \frac{\alpha_N I_{EO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_E}{kT}} - 1\right) - \frac{I_{CO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_C}{kT}} - 1\right) \quad (2)$$

With the collector junction reverse biased ϕ_C is always negative which makes $e^{\frac{q\phi_C}{kT}} \approx 0$, thus, the equations reduce to

$$I_E = -\frac{I_{EO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_E}{kT}} - 1\right) - \frac{\alpha_I I_{CO}}{1-\alpha_N\alpha_I} \quad (3)$$

$$I_C = \frac{\alpha_N I_{EO}}{1-\alpha_N\alpha_I}\left(e^{\frac{q\phi_E}{kT}} - 1\right) + \frac{I_{CO}}{1-\alpha_N\alpha_I} \quad (4)$$

It is convenient to have all currents and voltages expressed in terms of magnitudes only so the equations will apply to either NPN or PNP types without the necessity of making sign corrections. Thus for an NPN type, the actual currents which flow in the cutoff region (both junctions reverse biased) are indicated on Figure A, and the equations become

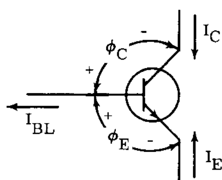


Figure A — Sign Convention Used for Cut-off Condition

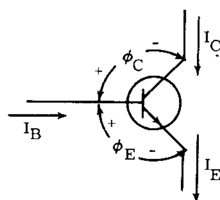


Figure B — Sign Convention Used for On Condition

$$I_E = \frac{I_{EO} - \alpha_I I_{CO}}{1-\alpha_N\alpha_I} \quad (5)$$

$$I_C = \frac{-\alpha_N I_{EO} + I_{CO}}{1-\alpha_N\alpha_I} \quad (6)$$

Ebers and Moll also derive the following useful equation

$$\alpha_N I_{EO} = \alpha_I I_{CO} \quad (7)$$

which can be used to simplify equations 5 and 6.

$$I_E = \frac{I_{EO}(1 - \alpha_N)}{1 - \alpha_I \alpha_N} \tag{8}$$

$$I_C = \frac{I_{CO}(1 - \alpha_I)}{1 - \alpha_N \alpha_I} \tag{9}$$

Using equations 3 and 4, the various equations given in Chapter 3 can be developed.

With both junctions forward biased, the direction of I_E and I_B have reversed from that shown in Figure A. Figure B now applies. Solving the general equations for voltages in terms of the currents and changing the sign of I_E:

$$\phi_E = \frac{kT}{q} \ln \left[1 + \frac{I_E - \alpha_I I_C}{I_{EO}} \right] \tag{10}$$

$$\phi_C = \frac{kT}{q} \ln \left[1 + \frac{\alpha_N I_E - I_C}{I_{CO}} \right] \tag{11}$$

In the Ebers & Moll analysis, the transistor model is assumed to be a homogeneous base transistor. Although no restriction is placed upon transistor geometry, voltage gradients due to I_R drops caused by the finite resistivity are neglected. In effect, this approximation restricts the geometry to devices having junctions that are parallel as found in alloy transistors. Hamilton[1] has shown that the graded base or drift transistor can be represented as a uniform base transistor having a narrower base width for forward current than for inverse current, and that the Ebers and Moll equations apply. Thus, the Ebers and Moll equations apply in practice to most alloy types of transistors. The drop due to the resistivity of the bulk material can be accounted for by adding resistance terms as illustrated in Chapter 4.

It is informative to consider the geometry effects in mesa type transistors which do not have plane parallel junctions, as illustrated in Figure C.

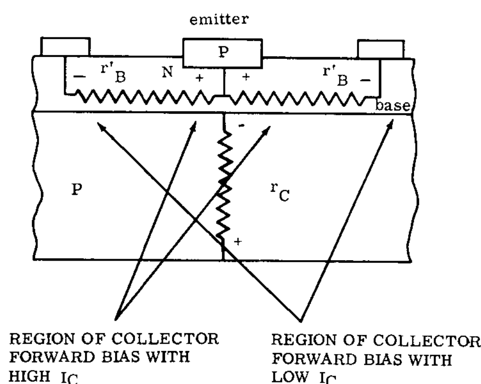


Figure C — Mesa Transistor Geometry

1. Hamilton, D. J., P. G. Griffith and D. C. Latham, "Avalanche Operation of Mesa Transistors," *Engineering Research Laboratories*, College of Engineering, University of Arizona, Tucson, Arizona.

When the base current is large, the drop across r'_B can be considerable. Assume the collector current is low so that the drop across r_c is negligible. In this case, holes emitted by the collector would be attracted to the base region around the base contact. The active region of the transistor under the emitter would not serve as a place for stored charge, therefore, the transistor proper would not be in saturation. This situation is depicted by the equivalent circuit of Figure D. The diode represents the part of the transistor which is in saturation. For this circuit, collector-emitter voltage can be written as

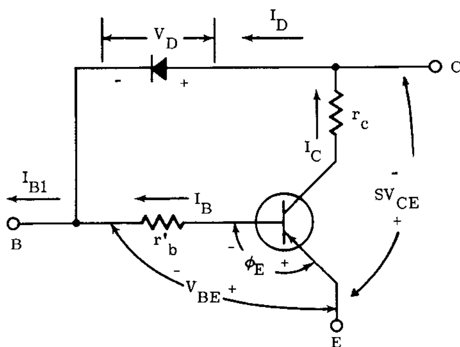


Figure D — Mesa Transistor Equivalent Circuit

$$V_{CE} = V_{BE} - V_D. \tag{12}$$

Expressing V_D by the ideal diode equation and neglecting bulk drops in the emitter and collector, equation 12 can be written as

$$V_{CE} = \phi_E + I_B r'_B - \frac{kT}{q} \ln \left(\frac{I_F}{I_{CD}} \right) \tag{13}$$

where ϕ_E is described by equation 4-10 since the transistor is not saturated. The diode current equals the excess base current $I_{B1} - I_C/\beta$ and I_E may be expressed as $(\beta + 1)\, I_B$. Using these expressions and substituting equation 4-10 into equation 13

$$SV_{CE} = \frac{kT}{q} \ln \left[\frac{(\beta_o + 1)\, I_B (1 - \alpha_I \alpha_N)}{I_{ED}} \right] + \frac{I_C}{\beta}\, r'_B - \ln \left(\frac{I_B - I_C/\beta_o}{I_{CD}} \right) \tag{14}$$

Using the identities $\alpha_N I_{ED} = \alpha_I I_{CD}$, assuming $\beta_o \gg 1$ and $\beta_F = I_C/I_B$, and simplifying

$$SV_{CE} = \frac{kT}{q} \ln \left[\frac{\alpha_N}{\alpha_I} (1 - \alpha_N \alpha_I) \frac{1}{(1/\beta_F - 1/\beta_o)} \right] + \frac{I_C\, r'_B}{\beta_o} \tag{15}$$

The differences between equation 15 and the Ebers and Moll voltage, ϕ_{CE}, are not readily apparent except for the addition of the $\frac{I_C r'_B}{\beta_o}$ term. Neglecting this

term, the remainder of the equation can be plotted as a family of curves depicting ϕ_{CE} for a mesa transistor, as shown in Figure E. It is seen that these curves resemble those plotted from the Ebers and Moll equations as shown in Figure 4-13.

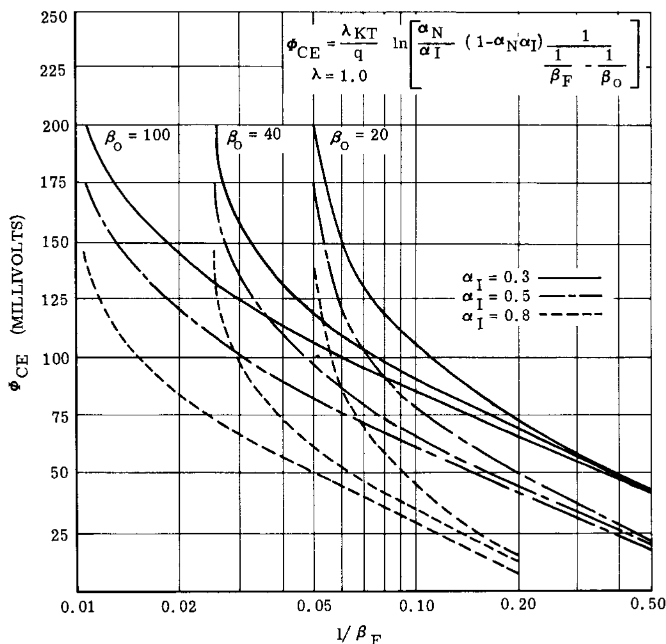


Figure E — Family of Curves Depicting Φ_{CE} for a Mesa Transistor

A comparison of the two ϕ_{CE} terms is shown in Figure F for the case where $\alpha_I = 0.3$, a typical value for a mesa transistor. Curves drawn for corresponding values of β_0 and β_F differ by only a few millivolts until β_F becomes extremely small. Such small differences would never be seen in practice since bulk resistances normally would contribute toward a major portion of the measured SV_{CE}, particularly when β_F is low.

In a mesa type transistor, as I_C is raised, the drop across r_C would bias the area under the emitter more negative which would attract more injected collector carriers to this region. Saturation voltage in this case would be more nearly described by the Ebers - Moll equation.

Therefore, it may be concluded that use of the Ebers - Moll equation plus the correction for device resistivity can be used to describe saturation behavior for any junction transistor.

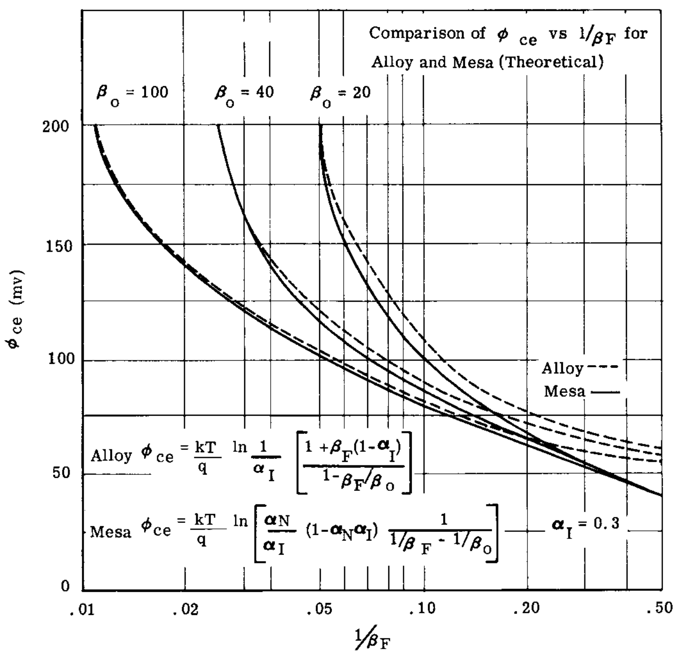


Figure F — Comparison of Φ_{CE} for Alloy and Mesa Transistors

NOTES

MOTOROLA Semiconductor Products Inc.

P.O. BOX 955 • PHOENIX 1, ARIZONA
A SUBSIDIARY OF MOTOROLA INC.