CAD of Electronic Circuits, I

Modeling the Bipolar Transistor

Ian E. Getreu

Tektronix Laboratories, Tektronix, Inc.
Beaverton, Oregon, U.S.A.

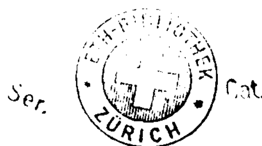
ELSEVIER SCIENTIFIC PUBLISHING COMPANY
Amsterdam – Oxford – New York 1978

ELSEVIER SCIENTIFIC PUBLISHING COMPANY
335 Jan van Galenstraat
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the United States and Canada:

ELSEVIER NORTH-HOLLAND INC.
52, Vanderbilt Avenue
New York, N.Y. 10017

P 301 167 : 1



Library of Congress Cataloging in Publication Data

Getreu, Ian E
Modeling the bipolar transistor.

(Computer-aided design of electronic circuits ; v. 1)
Bibliography: p.
1. Bipolar transistors--Data processing. 2. Bipolar transistors--Mathematical models. 3. Transistor circuits--Design and construction--Data processing.
I. Title. II. Series.
TK7871.96.B55G47 621.3815'28 78-17599
ISBN 0-444-41722-2

ISBN 0-444-41722-2 (Vol. 1)
ISBN 0-444-41723-0 (Series)

Copyright © Tektronix Inc., Beaverton, Oregon 97077, 1976
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher, Elsevier Scientific Publishing Company, P.O. Box 330, 1000 AH Amsterdam, The Netherlands

Printed in The Netherlands

Acknowledgments

The author is indebted to many people who assisted in the preparation of this book. He is grateful to his colleagues at Tektronix and several people at other institutions for their valuable feedback from earlier versions. The assistance of Fred Severson (Tektronix), Professors Robert Dutton (Stanford University), Don Pederson and Robert Meyer (University of California, Berkeley) and Dr. Richard Vaughan (University of New South Wales, Australia) in lengthy discussions and review of the manuscripts is greatly appreciated. The comments and advice rendered by George Wilson, Carl Battjes, David Hannaford, Dr. Binoy Rosario, Dr. Jim Smith, Dick Hung and Jack Millay (Tektronix), Drs. Don Scharfetter and Larry Nagel (Bell Laboratories) and Johan Brinch (University of New South Wales, Australia) are also very gratefully acknowledged. The contributions of several people to various sections of the book in the measurement section are appreciated and are acknowledged in the appropriate places. The assistance of Joyce Lekas, Fred Severson, Dr. Robert Nordstrom, Ted Niimi, Jeanne Galick, Doreen Weaver, Diana Clark and Jane West in the preparation of the manuscripts and drawing of the figures has been invaluable.

Some of the material in Section I, Introduction; Section II, Theoretical Derivation of the Models, 2.1 and 2.2; and Section III, Parameter Measurements 3.1, 3.2 and 3.3 appeared in a series of articles in Electronics, September 19, October 31, and November 14, 1974, Copyright © McGraw-Hill, Inc.

led

Contents

SECTION 1 - INTRODUCTION

Usefulness of Computer-aided Design 1

Objective 2

Structure 2

Notation and Orientation 3

Input and Model Parameters 4

Dependence of Models on the Physics

of the Device 6

Model Notation 6

Which Model to Use? 7

SECTION 2 - THEORETICAL DERIVATION OF THE MODELS

2.1 Introduction 9

2.2 The EM_1 Model 10

2.2.1 The Injection Version 12

2.2.2 The Transport Version 15

2.2.3 An Alternative Form -- the

Nonlinear Hybrid-π 18

2.2.4 Temperature Variation 20

2.2.5 Summary and Additional Comments 21

2.3 The EM_2 Model 24

2.3.1 Improved dc Characterization 24

a) r'_c 24

b) r'_e 27

c) r'_b 28

2.3.2 Charge-storage Effects 28

a) Junction Capacitors 29

b) Diffusion Capacitors 33

c) Substrate Capacitor 38

2.3.3 Small-signal (Linearized) EM_2 Model 39

2.3.4 Summary 40

2.4 The EM_3 Model 43

2.4.1 An Improved dc Model at a Given Temperature 44

a) Basewidth Modulation 44

b) β_{dc} Variation with Current 48

i) Region II: Mid Currents 50

ii) Region I: Low Currents 50

iii) Region III: High Currents 55

iv) The Full Picture: Regions I, II and III 57

v) Effect of Ohmic Resistances 57

vi) β_F Versus I_C Input Parameters in SLIC and SINC . . . 58

vii) $\beta_{F_{ac}}$ and Its Variation with Current 61

2.4.2 An Improved Charge-storage Model at a Given Temperature 61

a) Improved C_{jC} - r_b' Model 61

b) Variation of τ_F with Current 63

2.4.3 An Improved Variation with Operating Temperature 65

a) Physics-based Temperature Variation 65

i) τ_F 65

ii) C_{jC} and C_{jE} 66

b) Temperature-dependent Parameters that Require Extra Input Parameters . . 67

2.4.4 Small-signal (Linearized) EM_3 Model 67

2.4.5 Summary and Conclusions 68

2.5 The GP Model 69

2.5.1 Introduction 69

2.5.2 The Physical Definition of I_S 74

The Q_B Concept 81

2.5.3 The Components of Q_B 83

a) Mathematical Derivation of Components of Q_B 84

b) Physical Significance of

Components of Q_B 86

i) Q_{BO} 87

ii) Q_E 87

iii) Q_C 88

iv) Q_F 88

v) Q_R 89

vi) Effect of Operating Regions 90

vii) Summary 90

2.5.4 Evaluation of q_b 90

a) Component q_e 92

b) Component q_c 96

c) Component q_f 98

d) Component q_r 99

e) Effect of Depletion

Approximation 99

i) Q_B 100

ii) I_{SS} 100

iii) Q_E, Q_C 100

iv) Q_F, Q_R 101

f) Solution for q_b 103

g) High-level Injection

Solution 105

i) Verification of High-Current

Solution 105

ii) Simplification of q_2 106

h) Final Solution 108
i) Comparison with Gummel-Poon Derivation 109
2.5.5 Base-widening Effects 110
a) dc Characteristics 111
b) ac Characteristics 112
i) Effect of Base-widening on τ_B 112
ii) Emitter Delay (τ_1) Versus I_C 113
iii) τ_F Versus I_C 116
2.5.6 Comparison of the GP Model with the EM_3 Model 116
a) Basewidth Modulation 117
b) High-level Injection 118
c) β Versus I 119
d) τ_F Versus I_C 121
2.5.7 Small-signal (Linearized) GP Model 121
2.5.8 Summary 121
a) dc Model 122
b) Charge-storage Model 123
c) Temperature Variation Model 123
2.6 Limitations of the Models 124
a) Three-dimensional Effects 124
b) Breakdown 125
c) Saturation 126

SECTION 3 - PARAMETER MEASUREMENTS

3.1 Introduction 127

3.2 EM_1 Model Parameter Measurements 129

β_F 130

β_R 132

I_S 133

T_{nom} 136

E_g 137

3.3 EM_2 Model Parameter Measurements 138

r'_e 140

r'_c 144

a) Definitions of Two Limiting Values of r'_c 144

b) Measurement of $r'_{c_{normal}}$ and $r'_{c_{sat}}$ 146

r'_b 151

a) Introduction 151

b) Comparison of Measurement Techniques for r'_b 151

c) Small-signal Measurements 152

d) Pulse Measurement Techniques 158

e) Noise Measurement Technique 164

C_{jO}, ϕ and m 165

τ_F (or f_T at I_C, V_{CE}) 169

a) Determination of τ_F from f_T 170

b) Determination of f_T 171

τ_R (or τ_{SAT}) 176

C_{SUB} or C_{CS} 179

3.4 EM_3 Model Parameter Measurements 180

V_A 182

β_{FM}, C_2, n_{EL} and θ (or β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE}) 188

β_{RM}, C_4, n_{CL} and θ_R 200

RATIO 201

$\frac{L_E}{W}$, I_{CO} 202

TC_1, TC_2 204

3.5 GP Model Parameter Measurements 205

I_{SS} 206

V_B 207

I_K 213

I_{KR} 215

B 216

APPENDIX 1 -

A Comparison of the Transport Notation With the Injection Notation 219

APPENDIX 2 -

EM_3 Model Basewidth Modulation Analysis 223

APPENDIX 3 -

Derivation of the Five Components of Q_B in the GP Model 229

APPENDIX 4 -

The Accuracy of the EM_3 and GP Basewidth Modulation Models 233

APPENDIX 5 -

The Small-Signal, Linearized EM_3 and GP Models 239

APPENDIX 6 -

Input Parameters Cross-Reference for SLIC, SINC and SPICE 246

REFERENCES 255

Usefulness of Computer-aided Design

It has almost become a cliché to comment on the rapid growth of the computer as a tool in the design of circuits, both discrete and integrated. Although the computer is often thought of only in terms of a cheap and fast breadboard (or dry lab), computer-aided design (CAD) enables the circuit designer to do things which are not possible with other techniques. Using the computer, he can:

- Observe waveforms and frequency responses of voltages and currents without loading the circuit as a probe would in an actual circuit.
- Predict the performance of an IC at high frequencies, without the parasitics a breadboard introduces.
- Use ideal devices selectively, such as one with an infinite bandwidth or a very large gain, to isolate the effects of various device parameters on the circuit performance or to do futuristic "blue sky" analyses.
- Feed into a circuit ideal waveforms, such as extremely fast pulses or a mixture of pulses and sinusoids.
- Separate out dc circuitry in order to understand the basic part of the circuit.
- Open a feedback loop without disturbing the dc levels.
- Determine the poles and zeros of a transfer function for even large circuits.
- Do noise, sensitivity, worstcase and statistical analyses.

Objective

The circuit designer today has programs available which allow him to do a wide variety of analyses. However, the different programs with their differing input formats, rules, notations and device models can be very confusing and discouraging to the inexperienced user. Fortunately, the program input formats and rules are normally well documented, so if mistakes are made they are relatively easily detected (either by the computer, the program or the user). The biggest problem, however, lies in the lack of standardization of the notation and device models being used and the measurement of the model parameters.

This book is aimed at reducing the modeling confusion by systematically describing how to model the bipolar junction transistor (BJT). Emphasis is on the nonlinear, large-signal models used in nonlinear dc and transient analyses, since this is where most modeling problems are encountered. The linear model (the well-known hybrid-π model) which is already fairly well documented and easy to work with, is simply a linearized version of the nonlinear model presented. This linearization process is also described.

Structure

The book ties together recent developments in the modeling of the BJT with the established models in such a way that the reader can use most, if not all, of the circuit analysis programs currently available. To this end, two aspects of the modeling are covered.

First, in Section 2 of the book, the effects that are modeled are explained in terms of the physics of the device so that the reader obtains an understanding of the assumptions and limitations

of the model and the ability to handle the different notations that he will encounter. The descriptions start with the simplest model and build up to the most sophisticated model.

The second modeling aspect is the measurement of the device parameters. Techniques for obtaining these parameters from terminal measurements are given in Section 3.

Notation and Orientation

The model parameters and their notations are oriented towards three computer programs, SLIC[1] (up to Version I), SINC[2] (up to Version D) and SPICE[3] (Version 1), all readily available from the University of California at Berkeley.* These programs are excellent vehicles for this book because (i) their built-in transistor models cover the full range from very simple to very sophisticated, (ii) the strong dependence of the models on the physics of the device leads to concise input requirements, (iii) the programs are relatively widely accepted, and (iv) they are readily available.

The use of these programs as examples does not restrict the applicability of this book. The parameters described can normally be readily transformed for other programs in which models are based on the Ebers-Moll model[4] (in the dc and large-signal, nonlinear case), the hybrid-π model[5] (in the small-signal, linear case), or are user-defined. A cross-reference for the SLIC, SINC and SPICE model parameters with those described herein is given at the end (Appendix 6).

---

*These programs can be obtained for a nominal handling charge by contacting Professor D. O. Pederson, Electronic Research Laboratory, University of California at Berkeley, California, 94720, U.S.A. All three programs (SLIC, SINC and SPICE) are written in FORTRAN IV for a CDC 6000 Series computer. Versions of the programs for other computers are also available from other institutions.

Input and Model Parameters

To specify a transistor model completely, a program requires three types of information: fundamental constants, operating conditions and model parameters. Only one type (the last) is specified by the user.

The fundamental physical constants, such as Boltzmann's constant (k) and electronic charge (q), are normally defined inside the program. The operating conditions define the circumstances under which the model equations are to be used. In a nodal analysis program, for example, the operating conditions are normally the transistor's bias voltages, say V_{BE} and V_{BC}. These bias voltages are determined internally as the computer iterates to the solution. That is, the program assumes a set of bias voltages, solves the equations and then selects new and better values until it converges to an adequate solution. This is all done internally. In this book it will be assumed that the operating conditions consist of not only the bias voltages V_{BE} and V_{BC}, but also the temperature T at which the analysis is to be performed. The value of T is normally required as an input to the programs, and it will be assumed throughout that T has been specified.

The third type of information required is the set of model parameters for each different device in the circuit. The meaning and measurement of the model parameters is the subject of this book. Examples of model parameters are β_F (= $\frac{I_C}{I_B}$ in the normal, active region) and τ_F (the total transit time in the normal, active region). The values of the model parameters must be supplied by the user in a manner predetermined by the program. Some programs are very flexible and allow some model parameters to be specified indirectly. For example, τ_F is normally determined via the measurement of f_T, the unity gain bandwidth. In SLIC and SINC, the user has the option of specifying either τ_F or the f_T value at a bias condition. As explained in detail later, if f_T information

is provided, the programs determine the value of τ_F internally, taking into account the effects of junction capacitances, base-width modulation and collector ohmic resistance. Note that in this particular case τ_F is the model parameter, yet f_T is the input. Therefore, a distinction must be made between the model parameters and the program's input parameters. This distiction is illustrated in Fig. 1.1. The model parameters are those parameters used in the model equations to describe the device for a given set of operating conditions. The input parameters are the data required by the program to specify the model parameters. Some or all of the input parameters may be model parameters, depending on the program.

The formal distinction between model and input parameters, though appearing at first to be rather pedantic, is in fact very important. It helps to maintain the proper perspective. In the above example, the fact that τ_F is the model parameter underscores its importance. If τ_F were measured directly (rather than via f_T), then there would be no need to measure f_T. The measurement of f_T arises only in that it is a means of determining τ_F.

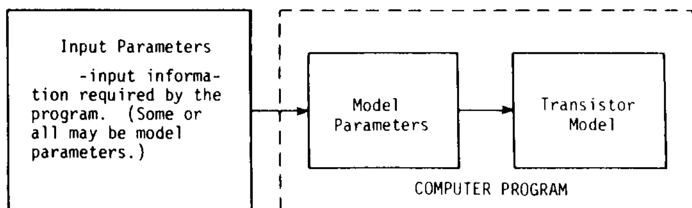


Fig. 1.1. The distinction between input parameters and model parameters.

Dependence of Models on the Physics of the Device

In general, a model based on the physics of the device permits a more thorough understanding and fewer input and model parameters. Fo· example, some programs require a table of values to describe thє variation of β_F with collector current. The physics-based model described in this book, however, requires only four model parameters to completely specify β_F versus I_C over the whole range. As well, in the process of determining these four parameters, an understanding is obtained of the reasons for the variation of β_F with I_C. For these reasons, the development of the models in this book emphasizes an understanding of the physics of the device.

Model Notation

The nonlinear models are based on the Ebers-Moll model. The original Ebers-Moll model[4] has been modified by many people to include effects such as charge storage, β variation with current, and basewidth modulation, among others. As a result, there are many interpretations of the phrase, "modified Ebers-Moll model." The modifications used here have three different levels of complexity. The notation, first put forward by Pederson,[6] is as follows:

EM_1 is the original Ebers-Moll model.[4] It is a nonlinear dc model only.

EM_2 is the next level of complexity. With the EM_1 model as its basis, it provides a first-order model of the nonlinear charge-storage effects and ohmic resistance.

EM_3 is the third level of complexity. It includes such second-order effects as basewidth modulation, β and τ_F variations with current, a better representation of the distributed collector-base junction capacitance and an improved temperature dependence.

GP is the Gummel-Poon model as implemented in the program SPICE.[3] It differs from the Integral

Charge-Control model by Gummel and Poon[7] mainly in the input parameters required and the absence of base push-out modeling. With respect to the effects that are modeled in both the EM_3 and GP models, the two models are basically equivalent.[6]

None of the computer programs describe the models with this terminology; it was chosen merely as a convenience. For example, a program may allow the user to specify some parameters from EM_1, some from EM_2 and some from EM_3. In all the Berkeley programs any parameter not specified by the user has a default value. If no parameters are specified, the resulting model is the EM_1 model. For example, if no capacitance parameters are specified, all capacitors are defaulted to zero, since the EM_1 model is a dc model.

Which Model to Use?

Use the simplest model that will do the job. This saves modeling effort and computer time, and the results are easier to understand. For example, if a transistor is current-fed, the external base resistance is not important and no effort should be made to determine its value.

The usefulness of the different models can be summarized as follows:

EM_1 model:	This model is very useful for first-order dc analyses.
EM_2 model:	For most applications (especially with digital circuits) the EM_2 model represents a good compromise between accuracy, ease of modeling, speed and understandable results.[8] It is the model most often used.
EM_3 and GP models:	These models should only be used when the extra accuracy they afford is necessary.

2.1 Introduction

The structure of this section, which presents a systematic, theoretical derivation of the BJT model, is as follows:

Section 2.2	The EM_1 model
Section 2.3	The EM_2 model
Section 2.4	The EM_3 model
Section 2.5	The GP model
Section 2.6	Limitations of the models

In the derivations of the models, emphasis is placed on an understanding of the effect being modeled and an explanation of the parameters required. The techniques for measuring these parameters are then given in Section 3.

2.2 The EM_1 Model

The EM_1 model is basically the simple, nonlinear model described by Ebers and Moll in 1954.[4] It is a dc model in that there is no characterization of charge storage in the device. The Ebers-Moll model is valid for all regions of operation: saturation, inverse, normal and off. These four regions are defined by the bias voltage on the junctions as illustrated in Fig. 2.1.

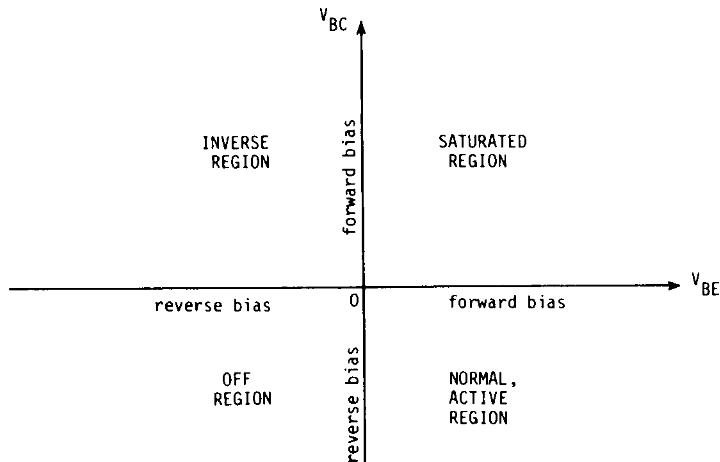


Fig. 2.1. The four regions of operation of the bipolar transistor.

The model parameters that fully describe the EM_1 model are defined and described in this section; they are β_F, β_R, I_S, T_{nom} and E_g.

Virtually all dc and large-signal, nonlinear models are based on the EM_1 model. Currently there are two popular versions of the EM_1 model: the injection version (Fig. 2.2a) and the transport version (Fig. 2.2b). These two versions, both drawn for an npn transistor, are mathematically identical. It is not immediately

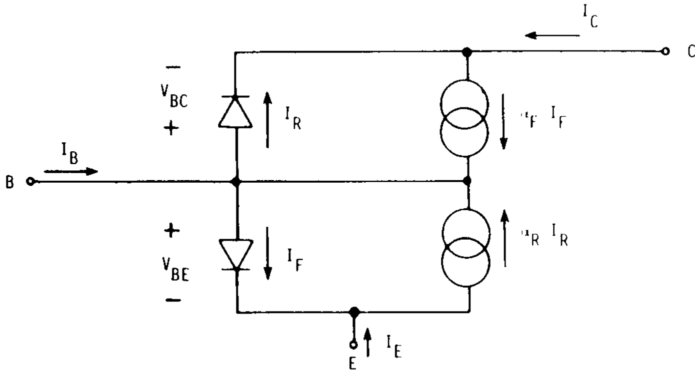


Fig. 2.2a. Injection version of EM_1 model.

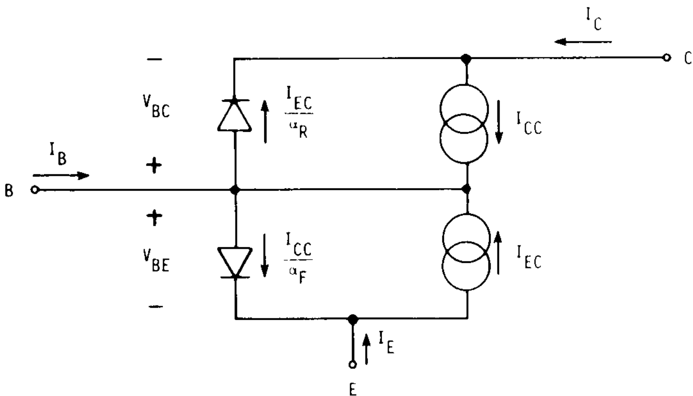


Fig. 2.2b. Transport version of EM_1 model.

apparent why one should be preferred over the other nor why any time or effort should be devoted to the distinction between the two. However, there are valid reasons why one (the transport version) is preferred for computer simulations.[9] These reasons only become obvious when higher levels of model complexity are

considered.* An explanation of both versions is given next, starting with the injection version.

2.2.1 The Injection Version

The injection version is the original and better-known EM_1 model. Its reference currents (those currents in terms of which all other currents are expressed) are I_F and I_R, the currents through the diodes. The reference forward diode current, I_F is given by:**

$$I_F = I_{ES}\left[e^{\frac{qV_{BE}}{kT}} - 1\right] \tag{2.1}$$

where I_{ES} is the emitter-base saturation current, V_{BE} is the base-emitter voltage, q is electron charge, k is Boltzmann's constant, and T is temperature. The reference reverse diode current, I_R, is given by:

$$I_R = I_{CS}\left[e^{\frac{qV_{BC}}{kT}} - 1\right] \tag{2.2}$$

where I_{CS} is the collector-base saturation current, and V_{BC} is the base-collector voltage. The collector terminal current can now be expressed in terms of I_F and I_R:

---

*The reasons for the preference of the transport model are (i) the reference currents are actually ideal over many decades of currents, (ii) the complete specification of both reference currents for given voltages by one fundamental constant I_S, and (iii) the ease of description of the diffusion capacitances. A more detailed comparison of these two versions of the EM_1 model is given in Appendix 1.

**In this book, all terminal currents are defined as positive when flowing into the device and the voltage V_{xy} is defined as positive when V_x is equal to or greater than V_y.

$$I_C = \alpha_F I_F - I_R \tag{2.3}$$

where α_F is the large-signal forward current gain of a common-base transistor. Similarly, the base terminal current can be written as:

$$I_B = (1 - \alpha_F)\, I_F + (1 - \alpha_R)\, I_R \tag{2.4}$$

where α_R is the large-signal reverse current gain of a common-base transistor. And lastly, the emitter terminal current becomes:

$$I_E = -I_F + \alpha_R I_R \tag{2.5}$$

The mathematical derivation of these equations is given in most transistor physics text books,[10] but a simple intuitive feel for the model can be obtained by inspection.

The diodes represent the transistor's base-emitter and base-collector junctions. I_F is the current that would flow across the base-emitter junction for a given base-emitter voltage, V_{BE}, if the collector region were replaced by an ohmic contact without disturbing the base. I_{ES} is the saturation current of this junction. The value of I_F for a given V_{BE} is given by Eq. (2.1). Similarly, Eq. (2.2) describes the collector-base junction if the emitter were replaced by an ohmic contact without disturbing the base, and I_{CS} is the saturation current of the collector-base junction.

The two diodes alone, back-to-back, do not fully represent the transistor. Coupling between the junctions is provided physically by the very narrow base region and is modeled by the two current-dependent current sources.

Consider the transistor to be biased in the normal, active region (B-E forward-biased and B-C reverse-biased). Then, the collector-base diode can be approximated by an open circuit, and the model reduces to the $\alpha_F I_F$ current generator and the base-emitter diode. I_F represents the total current flowing across the base-emitter junction while α_F is the fraction of that current that is collected at the base-collector junction.

Similarly, when the transistor is operated in its inverse mode (B-E reverse-biased, B-C forward-biased), α_R is the fraction of the total current that is flowing across the collector-base junction that is collected at the emitter-base junction.*

From Eqs. (2.1) through (2.5), four parameters are required to describe the injection version of the EM_1 model at one temperature: I_{ES}, I_{CS}, α_F and α_R. The number of parameters is reduced by one when the reciprocity property is applied. This property is defined by:

$$\alpha_F I_{ES} = \alpha_R I_{CS} \triangleq I_S \tag{2.6}$$

where I_S is called the transistor saturation current.** Reciprocity, which is experimentally observed,[12] is easily proven for the very simple case of a one-dimensional, constant base-doping transistor under low-level injection.[13] It has also been proven for the general case under low-level injection by an application of Green's theorem.[4]

The constants α_F and α_R are related to the large-signal forward and reverse current gains of a common-emitter transistor (β_F and β_R, respectively) by the well-known expressions:

---

*Since, for reasonable α values, the reference currents, I_F and I_R, represent the currents injected into the base region, the name "injection version" is used.

**The physical interpretation of I_S is as follows. A p-n junction saturation current consists of two terms, one from an analysis of each neutral region. For example, for the constant-doping, short-base diode, the saturation current is given by[11]

$$I_{SAT} = \frac{qAD_p p_{no}}{L_p} + \frac{qAD_n n_{po}}{W_B} \quad \text{for } W_B \ll L_n$$

$\alpha_F I_{ES}$ is the portion of the emitter-base saturation current (I_{ES}) that arises from the analysis of the base region. Similarly, $\alpha_R I_{CS}$ is the portion of the collector-base saturation current (I_{CS}) that arises from the analysis of the base region. Reciprocity, Eq. (2.6), then simply means that this analysis of the base region is the same for both I_{ES} and I_{CS} and that I_S is this common portion of both saturation currents.

$$\beta_F = \frac{\alpha_F}{1-\alpha_F} \tag{2.7}$$

$$\beta_R = \frac{\alpha_R}{1-\alpha_R} \tag{2.8}$$

Thus, only three model parameters are needed at one temperature. Those normally used are β_F, β_R, and I_S. All the other model parameters (I_{ES}, I_{CS}, α_F, α_R) can be obtained if necessary from these three model parameters.

2.2.2 The Transport Version

The transport version of the EM_1 model differs from the injection version only in the choice of the reference currents.

In the transport version the reference currents, I_{CC} and I_{EC}, are those flowing through the model's current sources. They represent those currents that are collected.* The reference collector source current can be written as:

$$I_{CC} = I_S \left[e^{\frac{qV_{BE}}{kT}} - 1\right] \tag{2.9}$$

and the reference emitter source current is:

$$I_{EC} = I_S \left[e^{\frac{qV_{BC}}{kT}} - 1\right] \tag{2.10}$$

---

*Or transported across the base, hence the name "transport version."

These two reference currents can then be used to express the transistor's terminal currents:

$$I_C = I_{CC} + \left[-\frac{1}{\alpha_R}\right] I_{EC} \tag{2.11}$$

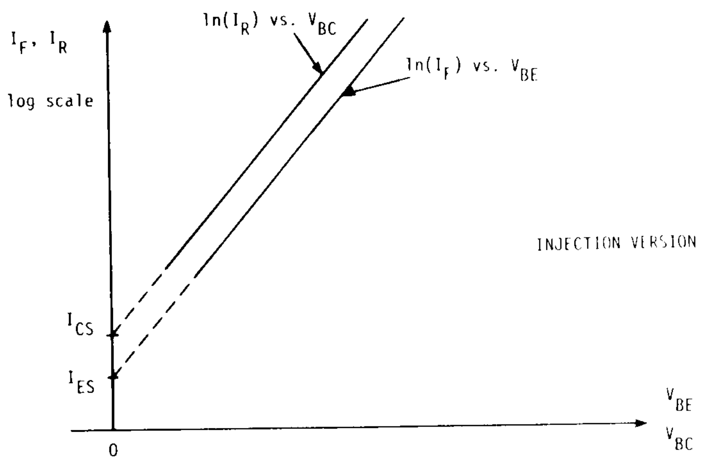
$$I_B = \left[\frac{1}{\alpha_F} - 1\right] I_{CC} + \left[\frac{1}{\alpha_R} - 1\right] I_{EC} \tag{2.12}$$

$$I_E = \left[-\frac{1}{\alpha_F}\right] I_{CC} + I_{EC} \tag{2.13}$$

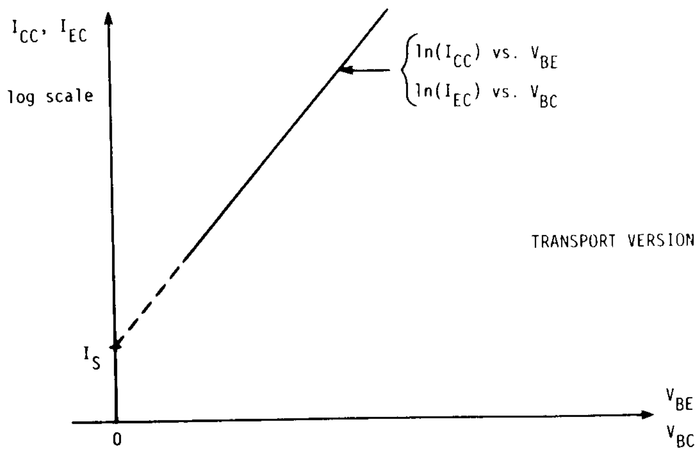
Mathematically, Eqs. (2.9) through (2.13) are identical to Eqs. (2.1) through (2.6). Note that the dependence of the reference currents on the junction voltages is the same for both I_{CC} and I_{EC}. That is:

$$I = I_S \left[e^{\frac{qV}{kT}} - 1\right] \tag{2.14}$$

where I is the reference current and V the appropriate junction voltage. Figure 2.3 shows the variation of the reference currents for both versions as a function of the appropriate junction voltages. The injection version curves are shown in Fig. 2.3a, while curves for the transport version are shown in Fig. 2.3b. Note that for the transport version (Fig. 2.3b), both lines are identical. That is, the variation of both reference currents with junction voltage is described by one fundamental constant, I_S. For given values of V_{BE} and V_{BC}, then, both reference currents for the transport version can be completely determined if I_S is known. For the injection version, however, two constants (I_{ES} and I_{CS}) are needed to obtain the reference currents. This, together with the fact that the transport reference currents are in practice linear over many decades on a semi-log plot, is the key to the preference of the transport version over the injection version, as described in Appendix 1.



(a)



(b)

Fig. 2.3. The variation of reference currents with junction voltages for (a) injection version and (b) transport of EM_1 model.

The difference between the injection and transport versions of the EM_1 model only involves a change in notation, not a change in the form of the model. The transport version's notation will be used hereafter.

2.2.3 An Alternative Form - the Nonlinear Hybrid-π (14)

At this point, a change in the model form is made. The change, as shown in Fig. 2.4, consists of replacing the transport model's two reference current sources with a single current source between the collector and the emitter. To do this, the equations for the diode saturation currents must be changed appropriately. As well, the equation for the single reference current source, I_{CT}, must be defined. The diode currents (Fig. 2.4) become:

$$\frac{I_{CC}}{\beta_F} = \frac{I_S}{\beta_F}\left[e^{\frac{qV_{BE}}{kT}} - 1\right] \tag{2.15}$$

and:

$$\frac{I_{EC}}{\beta_R} = \frac{I_S}{\beta_R}\left[e^{\frac{qV_{BC}}{kT}} - 1\right] \tag{2.16}$$

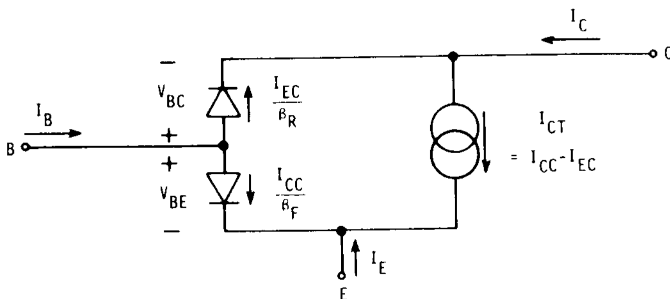


Fig. 2.4. The nonlinear hybrid-π version of the EM_1 model.

and the generator current, I_{CT} is given by

$$I_{CT} = I_{CC} - I_{EC}$$
$$= I_S \left\{ \left[e^{\frac{qV_{BE}}{kT}} - 1 \right] - \left[e^{\frac{qV_{BC}}{kT}} - 1 \right] \right\} \quad (2.17)$$

The model's terminal currents can now be written as:

$$I_C = (I_{CC} - I_{EC}) - \left[\frac{I_{EC}}{\beta_R} \right] \quad (2.18)$$

which is equivalent to Eq. (2.11), and as

$$I_B = \left[\frac{I_{CC}}{\beta_F} \right] + \left[\frac{I_{EC}}{\beta_R} \right] \quad (2.19)$$

which is equivalent to Eq. (2.12), and as

$$I_E = - \left[\frac{I_{CC}}{\beta_F} \right] - (I_{CC} - I_{EC}) \quad (2.20)$$

which is equivalent to Eq. (2.13).

There is a good reason for making this alteration. With the model form of Fig. 2.4, the linearized small-signal equivalent circuit of a transistor operating in its forward, active region reduces to that of the well-known linear small-signal hybrid-π model. As shown in Fig. 2.5, the I_{CT} current generator becomes the g_m (transconductance) current generator. The forward-biased diode between the base and the emitter terminals becomes the r_π resistor $\left(r_\pi = \frac{\beta_F}{g_m} \right)$. And the reverse-biased diode between the base and the collector terminals becomes the r_μ resistor, which is often assumed to be an open circuit.

Because of this easy transition from a nonlinear model to a linear model, the equivalent circuit of Fig. 2.4 is called "the

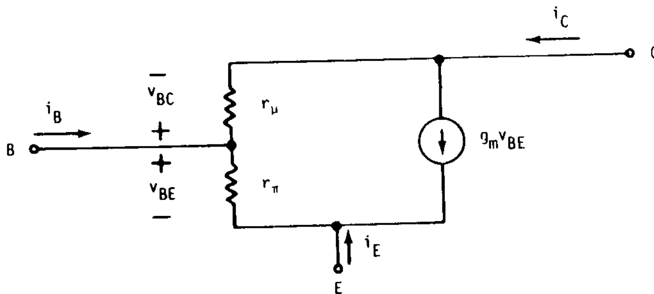


Fig. 2.5 A linearized version of Fig. 2.4 for operation in the forward, active region. Since the EM_1 model is a dc model, this model is the dc portion of the well-known linear hybrid-π model.

nonlinear hybrid-π model." Its similarity in form to the linear hybrid-π model means that a computer program can perform the small-signal hybrid-π linearization with very little effort.* Note, however, that a physical "feel" for the model is decreased since the diodes in the nonlinear hybrid-π model no longer represent the actual transistor pn junctions. The diode currents now actually represent components of the base current. However, the change in model form does not affect the parameters required to specify the first-level Ebers-Moll model at one temperature -- these parameters are still β_F, β_R, and I_S.

2.2.4 Temperature Variation

At the EM_1 level, β_F and β_R are both regarded as constants, independent of current, voltage and temperature. This assumption is maintained until the EM_3 level. The only parameter which is

*A second reason for using the nonlinear hybrid-π model of Fig. 2.4 lies in the modeling of the low-current variation of β_{dc}. This will be brought out later in the EM_3 section.

assumed to change with temperature is I_S and is of the form:(15)*

$$I_S(T) = I_S(T_{nom}) \left[\frac{T}{T_{nom}}\right]^3 e^{\frac{-E_g}{k}\left[\frac{1}{T} - \frac{1}{T_{nom}}\right]} \qquad (2.21)$$

where T is the analysis temperature in °K, T_{nom} is the nominal temperature in °K at which the device data is taken and E_g is the effective energy gap in electronvolts of the semiconductor material. Two more model parameters, T_{nom} and E_g, are needed to account for the variation of I_S with temperature assumed in the computer programs. (Remember, the temperature T is regarded as an operating condition and not a model parameter.) In some programs (SLIC and SINC for example) E_g is internally fixed at a value for silicon. In other programs (SPICE for example) E_g is a model parameter that the user can specify.

2.2.5 Summary And Additional Comments

To summarize the important aspects of the EM_1 model:

- For computer simulations, the transport notation is a better choice than the injection notation.

- At a given temperature, only three model parameters, β_F, β_R, and I_S, are needed to specify the model. Two additional parameters, T_{nom} and E_g, are required to model the variation of saturation current I_S with temperature. Because the EM_1 model is a dc model, β_F and β_R are dc parameters, not ac parameters. At the EM_1 level the distinction between ac and dc beta values is purely academic. However, at the higher-level models where β_{dc} versus I is modeled, the distinction becomes important.

---

*Neglecting the temperature dependence of diffusivity, $kT\mu(T)/q$. See page 137 for more details.

- A transformation from the transport model form (Fig. 2.2b) to the nonlinear hybrid-π model form (Fig. 2.4) makes the computer linearization to the small-signal linear hybrid-π model (Fig. 2.5) simpler. The nonlinear hybrid-π model is the basis for the other models. Elements will be added to it and some of its parameters will be redefined later.

All the model diagrams given here are for an npn transistor. For a pnp transistor, the voltage and current polarities must be changed appropriately. In most computer programs, the model parameter values are always considered to be positive, and the appropriate sign changes are implemented internally by the program.

The collector characteristics of the EM_1 model are shown in Fig. 2.6 as they would appear on a curve tracer. (Note that, even for this model, there is an inherent saturation voltage, $V_{CE(sat)}$.)

Although it is very simple in form and requires, at most, five parameters, the EM_1 model is quite accurate. It is useful, not only for a dc characterization of the bipolar junction transistor, but also as an "ideal" transistor. This "ideal" transistor can be used in CAD in a variety of ways as outlined in the introduction in Section 1.

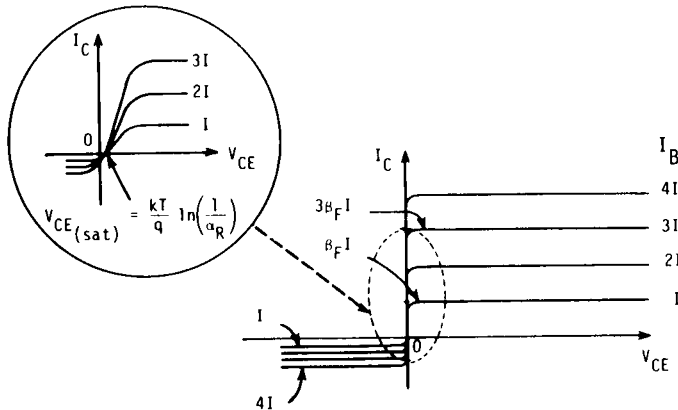


Fig. 2.6. Curve tracer collector characteristics of the EM_1 model.

The limitations of the EM_1 model lie mainly in its neglect of transistor charge storage (i.e., no diffusion or junction capacitance) and ohmic resistances to the terminals. These effects are included in a first-order manner in the EM_2 model described in the next section.

2.3 The EM_2 Model

The EM_2 model provides a first-order modeling of the charge storage effects which permits the realization of finite frequency and time responses. It also provides a more accurate dc representation of the device.* Eight additional components are introduced, requiring twelve more model parameters to describe them.

Figure 2.7 shows the progression from the EM_1 model to the EM_2 model for an npn transistor. The complete EM_2 model is given in Fig. 2.7e. Three ohmic bulk resistors (r_c', r_e' and r_b'), two diffusion capacitors (C_{DE} and C_{DC}) and three junction capacitors (C_{jE}, C_{jC}, and C_{SUB}) are added to the first-level model, which is simply the two diodes and the current generator of the non-linear hybrid-π model form.

2.3.1 Improved dc Characterization

The inclusion of three constant resistors (r_c', r_e' and r_b') improves the dc characterization. They represent the transistor's ohmic resistances from its active region to its collector, emitter and base terminals, respectively. These resistors are included in the model as shown in Fig. 2.7b. The internal nodes of these resistances at the active region are denoted by the letters C', E' and B' in the model diagram. The voltages used in describing the two ideal diodes and the current source are the internal voltages.

a) r_c'

The effect of the collector resistance (r_c') is seen in Fig. 2.8 in which collector characteristics of the EM_2 model (solid lines) are compared to collector characteristics of the EM_1

---

*The improvement to the dc representation will also obviously affect the ac and transient responses of the model.

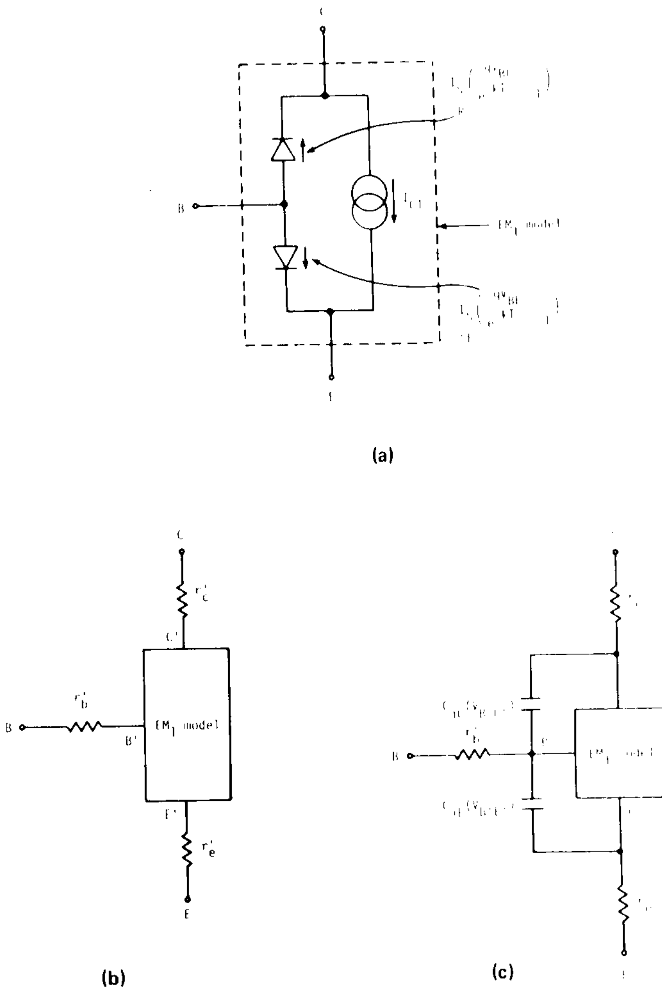


Fig. 2.7. The progression from EM_1 to EM_2 model for an npn transistor. (a) EM_1 model (nonlinear hybrid-π); (b) addition of the three ohmic resistors; (c) addition of the nonlinear junction capacitors.

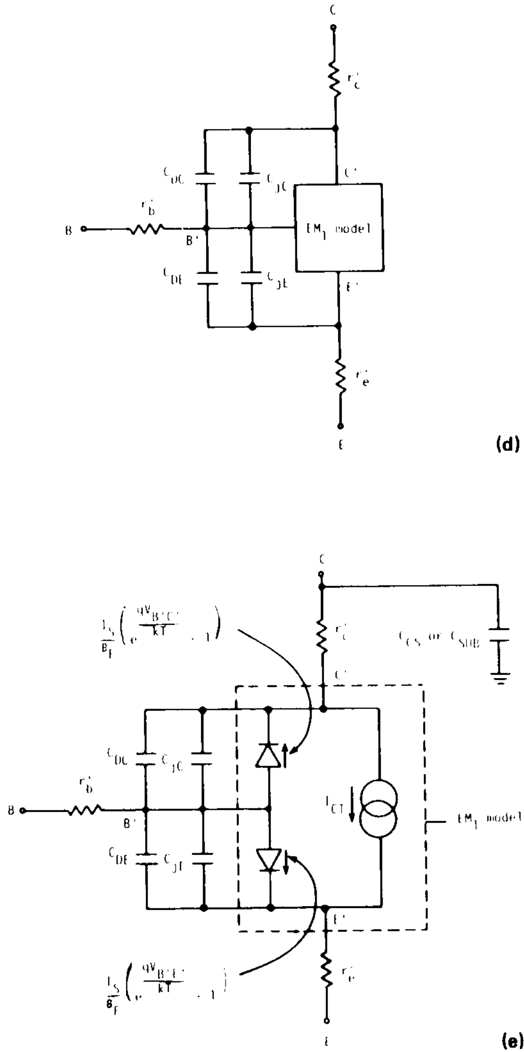


Fig. 2.7. The progression from EM_1 to EM_2 model for an npn transistor. (d) addition of the nonlinear diffusion capacitors; (e) complete EM_2 model.

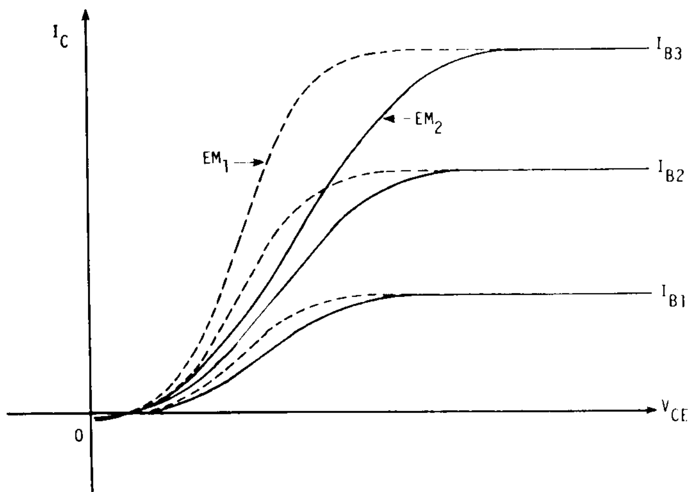


Fig. 2.8. The effect of r_c' on the I_C versus V_{CE} characteristics. Dashed lines represent the EM_1 model ($r_c' = 0$). Solid lines represent the EM_2 model.

model (dashed lines). Resistance r_c' decreases the slope of the curves in the saturated region for low collector-emitter voltages.

In the EM_2 model, r_c' is assumed to be constant. In an actual device, however, it will be a function of collector current and base-collector voltage. Therefore, the biggest problem in obtaining r_c' is not how to measure it, but which value to use. A more detailed description of this problem is given in the measurement section on r_c'.

b) r_e'

The emitter is the most heavily doped region in most present-day transistors in order to produce a high emitter injection

efficiency[16] and therefore a high β_F. For this reason, the dominant component of <u>emitter resistance</u> (r_e') is normally the contact resistance (usually on the order of 1 Ω). r_e', which is often neglected, can normally be assumed to have a small, constant value. Its main effect is a reduction in the voltage seen by the emitter-base junction by a factor of $r_e' I_E$.

In this effect on V_{BE}, resistance r_e' is equivalent to a base resistance of $(1 + \beta_F) r_e'$. Therefore r_e' affects the collector current as well as the base current, as shown in Fig. 2.9. This effect can be significant and r_e' can cause substantial errors in the determination of r_b'.[17] Resistance r_e' can also seriously affect the collector characteristics in the saturation region if the transistor has a low r_c' value.

c) r_b'

<u>Base resistance</u> r_b' is an important model parameter. Its greatest impact is normally its effect on the small-signal and transient responses. It is also one of the most difficult parameters to measure accurately, partly because of its strong dependence on operating point,[17,18] (due to crowding) and partly because of the error introduced by the small, but finite value of r_e'.[17] In the EM_2 model, r_b' is assumed to be constant. The dc effect of r_b' is seen on the $\ln(I_C)$ and $\ln(I_B)$ versus V_{BE} curve obtained from the EM_2 model, as illustrated in Fig. 2.9.

2.3.2 Charge-Storage Effects

Charge storage in the bipolar junction transistor is modeled by the introduction of three types of capacitors: two nonlinear junction capacitors, two nonlinear diffusion capacitors and a constant substrate capacitor.

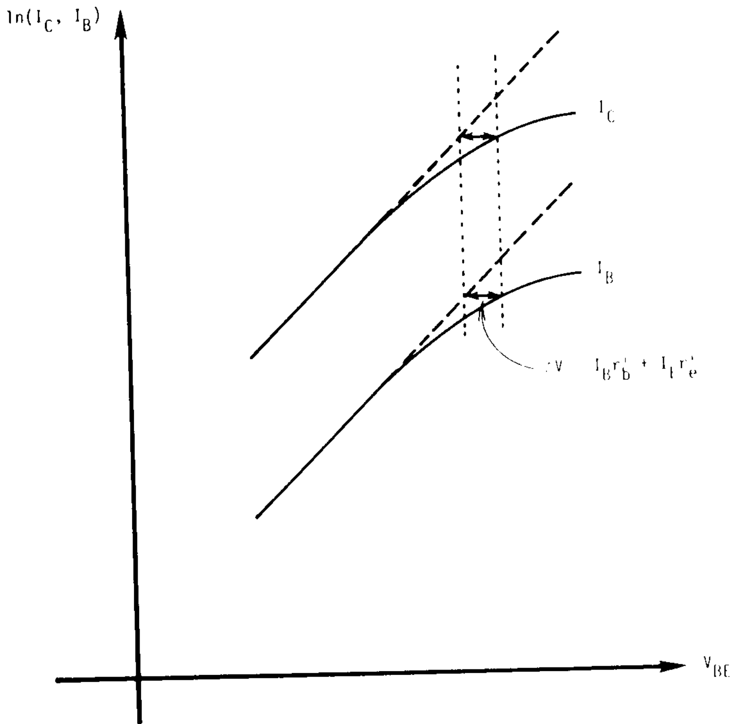


Fig. 2.9. The effect of r_b' and r_e' on the $\ln(I_C)$ and $\ln(I_B)$ versus V_{BE} characteristics of the EM_2 model. (Note that the EM_2 model neglects other high-level effects which are treated in higher-order models.)

a) Junction capacitors

The two junction capacitors (sometimes called transition capacitors) model the incremental fixed charges stored in the transistor's space-charge layers for incremental changes in the associated junction voltages. These capacitors, denoted by C_{jE} for the base-emitter junction and C_{jC} for the base-collector junction are

included in the model as shown in Fig. 2.7c. Each junction capacitance is a nonlinear function of the voltage across the junction with which it is associated.

The normal, simple analysis of C_{jC} and C_{jE} makes the depletion approximation (which assumes that at the junction of interest the space-charge layer is depleted of carriers). Then, for a step (or abrupt) junction and for a linear (or graded) junction, the variation of the emitter junction capacitance with base-emitter junction voltage can be written for an npn transistor*[19] as

$$C_{jE}\,(V_{B'E'}) = \frac{C_{jE0}}{\left(1 - \frac{V_{B'E'}}{\phi_E}\right)^{m_E}} \tag{2.22}$$

where C_{jE0} is the value of the emitter-base junction capacitance at $V_{B'E'} = 0$, ϕ_E is the emitter-base barrier potential, and m_E is the emitter-base capacitance gradient factor. Likewise, the variation of the collector junction capacitance with the base-collector junction voltage is given for an npn transistor by:

$$C_{jC}\,(V_{B'C'}) = \frac{C_{jC0}}{\left(1 - \frac{V_{B'C'}}{\phi_C}\right)^{m_C}} \tag{2.23}$$

where C_{jC0} is the value of the collector-base junction capacitance at $V_{B'C'} = 0$, ϕ_C is the collector-base barrier potential, and m_C is the collector-base capacitance gradient factor.

For a step junction, m = 0.5; for a linear junction m = 0.333. Since most practical junctions lie between a step and a linear junction, the above equations are assumed to be general and to apply for all junctions with a gradient factor between 0.333 and 0.5. This is therefore an empirical fit.

---

*For a pnp transistor, since V_{xy} is positive when $V_x \geq V_y$, the minus signs in Eqs. (2.22) and (2.23) become plus signs.

The parameters ϕ_E and ϕ_C are called the transistor's built-in barrier potentials for the emitter-base and collector-base junctions, respectively. There can be some confusion about these built-in barrier potentials (which are typically of the order of 0.7 V from capacitance-voltage data). It may appear at first that since a typical value of V_{BE} at mA current levels is 0.7 V to 0.8 V, the junction applied voltage can be greater than the built-in barrier potential. This does not follow, for two reasons. First, the externally-applied base-emitter voltage, V_{BE}, should not be confused with the *internal* junction applied voltage, $V_{B'E'}$. At high current levels, the finite values of r_b' and r_e' in practical transistors can cause $V_{B'E'}$ to be significantly less than V_{BE}. Second, there are actually *two* built-in barrier potentials:* the well-known barrier potential which sets up the drift component of current to oppose the diffusion component at the junction[20] (which equals $\frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right)$ for the simple, constant-doping, abrupt-junction case and is typically of the order of 1 V) and the barrier potential used in the above capacitance formula[21] (which is typically of the order of 0.7 V). It is the *former* barrier potential that is greater than $V_{B'E'}$. Because of this condition the former barrier potential does not appear in any of the equations. *The user, however, need not concern himself with the details of this distinction between the two types of barrier potentials.* In the model described, ϕ is used only to compute the junction capacitance and therefore the value obtained for ϕ from C-V data is (by definition) the appropriate value to use.**

The above functional dependencies of C_{jE} and C_{jC} (Eqs. (2.22) and (2.23)) are built into the EM_2 model. To specify the junction

---

*An even further confusion results sometimes from mistaking the barrier potential for the transistor's energy gap, E_g.

**Actually, ϕ may also be used in the EM_3 and GP models to describe the variation of basewidth with temperature and high current levels, but normally it is only used for computing C_j.

capacitances completely, three parameters (C_{j0}, ϕ and m) are required for each. Although most experimental data can be forced to fit Eqs. (2.22) and (2.23), the reduction of the measured capacitance as a function of voltage into these three parameters is not trivial if any stray capacitances are present. Techniques for reducing the data are presented in the measurement section. Some programs allow the user to specify each junction capacitance at a value of its associated junction voltage, instead of requiring the C_{j0} value.

Figure 2.10 shows three plots of the variation of the junction capacitance as a function of voltage. The dashed line represents Eqs. (2.22) and (2.23). Under forward bias these equations predict infinite capacitance when the internal junction voltage equals the built-in voltage. Chawla and Gummel[22] have shown that under forward bias the depletion approximation is no longer valid and that Eqs. (2.22) and (2.23) no longer apply. The solid line in Fig. 2.10 shows the (non-infinite) variation of the junction capacitance obtained by Chawla and Gummel. An expression requiring four parameters has been fitted to this curve by Poon and Gummel.[23]

The third curve in Fig. 2.10 represents the straight line approximation made by the Berkeley programs SLIC and SINC for $V > \phi/2$.* The equation for this straight line, obtained by matching slopes at $\phi/2$, is given by:

$$C_j\left(V \geq \frac{\phi}{2}\right) = 2^m C_{j0}\left[2m\frac{V}{\phi} + (1-m)\right] \tag{2.24}$$

where the junction subscripts have been omitted and V must be greater than or equal to $\phi/2$. This approximation, while avoiding the infinite capacitance, is not as accurate as the Chawla-Gummel curve. However, it is acceptable because under forward bias the diffusion capacitances, described next, are dominant and inherently

*The other Berkeley program, SPICE, uses a similar straight line approximation, but about V = 0.[3] That is, $C_j(V \geq 0) = C_{j0}(1 + m\frac{V}{\phi})$.

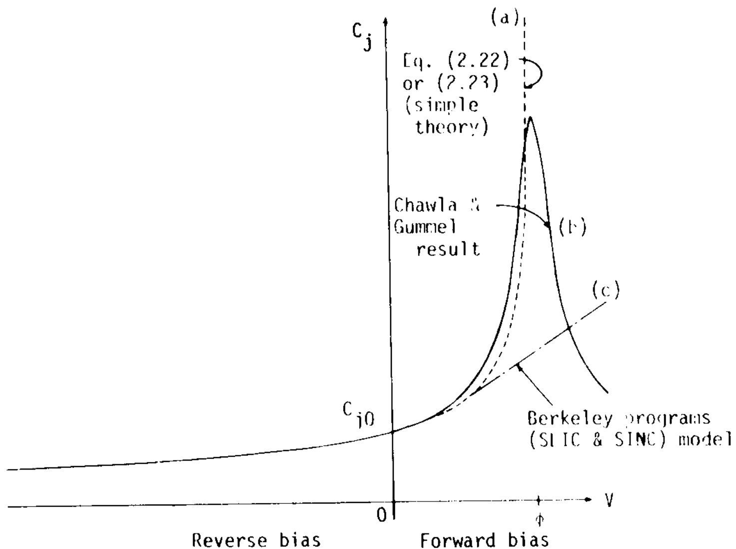


Fig. 2.10. A plot of three equations describing the variation of junction capacitance with voltage.

include the effect of the mobile charges in the space-charge layers. Besides needing one less parameter than the Poon-Gummel equation, the Berkeley model is accurate in the reverse-bias region where it is most important.

b) Diffusion capacitors

The diffusion capacitances model the charge associated with the mobile carriers in the transistor. This charge is divided into two components: one associated with the reference collector source current (I_{CC}) and the other with the reference emitter source current (I_{EC}). Each component is represented by a capacitor.

To evaluate the diffusion capacitance associated with I_{CC}, the total mobile charge associated with I_{CC} must be considered. Therefore, the base-emitter junction is assumed to be forward-biased and $V_{BC} = 0$. Figure 2.11a shows the minority-carrier concentrations

for the simplified one-dimensional case of constant base doping, negligible base recombination, and low-level injection.* The total mobile charge associated with I_{CC} (Q_{DE}) can be written as the sum of the individual minority charges:

$$Q_{DE} = Q_1 + Q_2 + Q_3 + Q_4 \tag{2.25}$$

where Q_1 is the mobile minority charge stored in the neutral emitter region, Q_2 is the mobile minority charge in the emitter-base space-charge region associated with I_{CC} (normally considered to be zero),** Q_3 is the minority mobile charge stored in the neutral base region and Q_4 is the mobile minority charge in the collector-base space-charge region associated with I_{CC}. Because of charge neutrality there will be identical majority charges stored in the neutral regions. However, to determine diffusion capacitance only one (minority or majority) needs to be considered.

Charge Q_2 is normally considered to be zero, so that the total mobile charge associated with I_{CC} can be expressed as:

$$\begin{aligned} Q_{DE} &= Q_1 + Q_2 + Q_3 + Q_4 \\ &= \tau_1 I_{CC} + \tau_{EB_{SCL}} I_{CC} + \tau_B I_{CC} + \tau_{CB_{SCL}} I_{CC} \\ &\approx (\tau_1 + \tau_B + \tau_{CB_{SCL}}) I_{CC} \end{aligned}$$

$$Q_{DE} \triangleq \tau_F I_{CC} \tag{2.26}$$

where τ_1 is the emitter delay,***[25] $\tau_{EB_{SCL}}$ is the emitter-base space-charge layer transit time, τ_B is the base transit time

*The analysis is not restricted to this simplified case. This case was chosen for ease of presentation only.

**For high-frequency devices, recent computer studies (by R. Mertens, University of Leuven, Belgium) have shown that Q_2 may not be negligible. This has also been concluded by Kerr and Berz.[24]

***The emitter delay, τ_1, is normally given the symbol τ_E. However, τ_1 is used here in order to avoid any confusion later in the GP model.

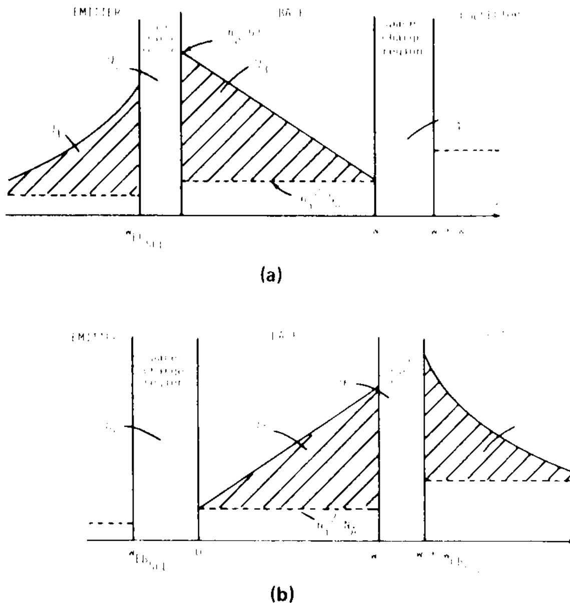


Fig. 2.11. The mobile minority charge associated with (a) I_{CC} and (b) I_{EC} (not to scale).

(= $\frac{W^2}{2D_n}$ for the constant-doping case),(26) $\tau_{CB_{SCL}}$ is the base-collector space-charge layer transit time $\left(= \frac{W_{CB_{SCL}}}{2v_{scat\ lim}}\right)$,(27) τ_F is the total forward transit time (here assumed to be a constant) and D_n and $v_{scat\ lim}$ are the diffusivity and scatter-limited velocity for the base minority carriers, respectively.*

---

*Note that Eq. (2.26) is a large-signal, or dc, definition of τ_F. The normal small-signal, or ac, definition is

$$\tau_{F_{ac}} \triangleq \frac{dQ_{DE}}{dI_{CC}}$$

For the EM_2 model where τ_F is assumed constant.

$$\tau_{F_{dc}} = \tau_{F_{ac}} = \tau_F$$

Similarly for τ_R and Eq. (2.28). In the EM_3 and GP models, where τ_F is not constant, the distinction between $\tau_{F_{dc}}$ and $\tau_{F_{ac}}$ is necessary.

A similar analysis of the total mobile charge associated with I_{EC} (see Fig. 2.11b) gives:

$$Q_{DC} = Q_5 + Q_6 + Q_7 + Q_8 \tag{2.27}$$

where Q_5 is the mobile minority charge stored in the neutral collector region, Q_6 is the mobile minority charge in the collector-base space-charge region associated with I_{EC}, Q_7 is the minority mobile charge stored in the neutral base region, and Q_8 is the mobile minority charge in the emitter-base space-charge region associated with I_{EC}. If charge Q_6 is assumed to be zero, then:

$$Q_{DC} = \tau_C I_{EC} + \tau_{CB_{SCL}} I_{EC} + \tau_{BR} I_{EC} + \tau_{EB_{SCL}} I_{EC}$$

$$\simeq (\tau_C + \tau_{BR} + \tau_{EB_{SCL}}) I_{EC}$$

$$Q_{DC} \triangleq \tau_R I_{EC} \tag{2.28}$$

where τ_C is the collector delay, τ_{BR} is the reverse base transit time, and τ_R is the total reverse transit time (also assumed to be constant here).

For the saturated mode (i.e., V_{BE} and V_{BC} both forward-biased), both Q_{DE} and Q_{DC} are assumed to occur independently and the total minority charge stored in the transistor is the sum of components Q_1 through Q_8. The two charges, Q_{CE} and Q_{DC}, are modeled by two <u>nonlinear</u> capacitors C_{DE} and C_{DC} respectively (see Fig. 2.7d) given by:

$$C_{DE} \triangleq \frac{Q_{DE}}{V_{B'E'}} = \frac{\tau_F I_{CC}}{V_{B'E'}} \tag{2.29}$$

$$C_{DC} \triangleq \frac{Q_{DC}}{V_{B'C'}} = \frac{\tau_R I_{EC}}{V_{B'C'}} \tag{2.30}$$

For small-signal analyses, C_{DE} is linearized to:*

$$C_{DE_{small\ signal}} \triangleq \left. \frac{dQ_{DE}}{dV_{B'E'}} \right|_{V_{B'C'}=0} = g_{mF}\,\tau_F \qquad (2.31)$$

where g_{mF} is the transistor's forward transconductance:

$$g_{mF} \triangleq \left. \frac{dI_{CC}}{dV_{B'E'}} \right|_{V_{B'C'}=0} = \frac{qI_{CC}}{kT} \qquad (2.32)$$

Similarly, for small-signal analyses, capacitance C_{DC} is linearized to:

$$C_{DC_{small\ signal}} \triangleq \left. \frac{dQ_{DC}}{dV_{B'C'}} \right|_{V_{B'E'}=0} = g_{mR}\,\tau_R \qquad (2.33)$$

where g_{mR} is the transistor's reverse transconductance:

$$g_{mR} \triangleq \left. \frac{dI_{EC}}{dV_{B'C'}} \right|_{V_{B'E'}=0} = \frac{qI_{EC}}{kT} \qquad (2.34)$$

The position of these two diffusion capacitors in the model can be justified by considering the voltages that influence the charges -- $V_{B'E'}$ for Q_{DE} and $V_{B'C'}$ for Q_{DC}.

In present-day transistors, the contribution of the base-region terms Q_3 and Q_7 is not as significant as it used to be. For example, in the total forward transit time τ_F, emitter delay τ_1 can be as great as or greater than base transit time τ_B;(25) while in the total reverse transit time τ_R, collector delay τ_C is invariably the dominant component.

---

*This linearization is performed inside the computer program so the user is only required to specify τ_F. By specifying τ_F, the user models Q_{DE} in either a non-linear mode (for transient analyses) or a linear mode (for ac analyses) since τ_F is assumed to be constant (i.e., $\tau_{F_{dc}} = \tau_{F_{ac}}$).

Input parameters

Although τ_F and τ_R are the only extra model parameters needed to describe C_{DE} and C_{DC}, some programs enable the user to specify others, which are more easily measured.

For example, the total forward transit time τ_F can be specified in terms of the transistor's unity gain bandwidth, f_T, at a given collector current and collector-emitter voltage. From the operating point at which f_T is measured, the program first calculates V_{BE} and V_{BC}, then C_{jE} and C_{jC}. τ_F can then be determined from:[1]

$$\tau_F = \frac{1}{2\pi f_T} - \frac{kT}{qI_C}\left[C_{jE} + C_{jC}\left(1 + \frac{qI_C}{kT} r'_c\right)\right] \tag{2.35}$$

This formula is modified in the EM_3 model to take basewidth modulation effects into account. Note that f_T is <u>not</u> a model parameter, but only a means to obtain τ_F.

Similarly the total reverse transit time τ_R can be specified in the form of a measured parameter, τ_{SAT}. The parameter τ_{SAT} is defined as the saturation time constant and is related to the saturation delay time,[28] as shown in the measurement section. τ_R is obtained from τ_{SAT} by:[29]

$$\tau_R = \frac{\left(1 - \alpha_F\alpha_R\right)}{\alpha_R}\left(\tau_{SAT}\right) - \frac{\alpha_F\,\tau_F}{\alpha_R} \tag{2.36}$$

c) Substrate capacitor

The substrate capacitance (C_{CS} or C_{SUB}) can be important in integrated circuits. Although it is actually a junction capacitance in the way it varies with the epitaxial layer-substrate potential, it is modeled here as a constant-value capacitor. This representation is adequate for most cases since the epitaxial layer-substrate junction is reverse-biased for isolation purposes. (To include the

variation of C_{SUB} with epitaxial layer-substrate voltage, a separate diode or transistor can be included in the circuit schematic,[30] as described in the measurement description for C_{SUB} in Section 3.)

The placement of C_{SUB} in the model in Fig. 2.7e is shown for an npn transistor. However, for a pnp transistor, C_{SUB} may not be connected to the collector. Instead, for a lateral pnp device, C_{SUB} is connected between the base and the substrate; while for a substrate pnp device, C_{SUB} is set to zero since it is already modeled in the C_{jC} capacitance. (In SLIC these different connections are automatically made, whereas in the other Berkeley programs, C_{SUB} is connected to the collector for both npn and pnp transistors.)

2.3.3 Small-Signal (Linearized) EM_2 Model

A linearized version of Fig. 2.7e is used for ac linear analyses. The elements for this linear hybrid-π model, shown in Fig. 2.12, are given by:

$$r_\pi = \frac{\beta_F}{g_{mF}} \tag{2.37}$$

$$r_\mu = \frac{\beta_R}{g_{mR}} \tag{2.38}$$

$$C_\pi = g_{mF}\,\tau_F + C_{jE}(V_{B'E'}) \tag{2.39}$$

$$C_\mu = g_{mR}\,\tau_R + C_{jC}(V_{B'C'}) \tag{2.40}$$

In the normal region of operation, reverse transconductance g_{mR} is essentially zero, so that resistance r_μ can be regarded as infinite and capacitance $C_\mu \approx C_{jC}(V_{B'C'})$. The resultant model is the well-known linear hybrid-π model.[5]

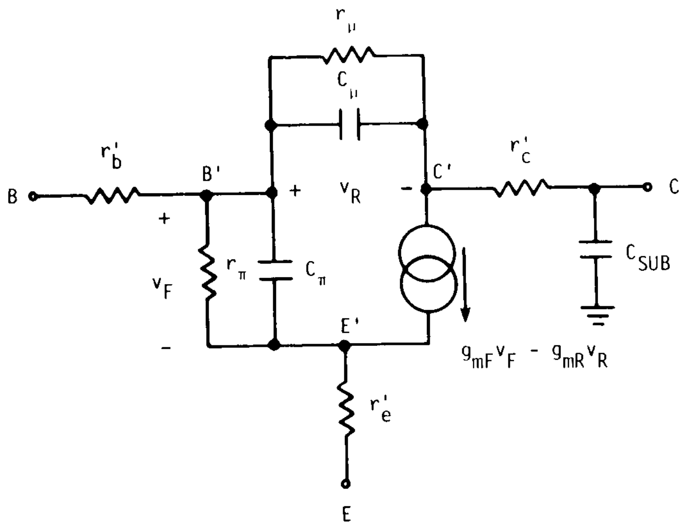


Fig. 2.12. The linearized version of the EM_2 model of Fig. 2.7e.

2.3.4 Summary

The complete EM_2 model requires eight extra components (three constant resistances, four nonlinear capacitors, and one constant capacitor) to be added to the EM_1 model. The five capacitors provide a first-order model of the charge storage in the transistor. As well, the three resistors give an improved dc representation over that provided by the EM_1 model.

To characterize these eight components, a total of 12 extra model parameters are required. These are:

The three resistors are described directly:

r_c' The (assumed constant) collector ohmic resistance

r_e' The (assumed constant) emitter ohmic resistance

r_b' The (assumed constant) base ohmic resistance

For the emitter junction capacitor, C_{jE}, three more model parameters are needed:

C_{jE0} The emitter-base junction capacitance at $V_{BE} = 0$ (or C_{jE} at a given V_{BE})

ϕ_E The emitter-base barrier potential

m_E The emitter-base capacitance gradient factor

For the collector junction capacitor, C_{jC}, a similar set of three model parameters is needed:

C_{jC0} The collector-base junction capacitance at $V_{BC} = 0$ (or C_{jC} at a given V_{BC})

ϕ_C The collector-base barrier potential

m_C The collector-base capacitance gradient factor

For the emitter diffusion capacitor, C_{DE}, only one model parameter is required:

τ_F The (assumed constant) total forward transit time (or the unity-gain bandwidth, f_T, at a given I_C and V_{CE}, from which τ_F is obtained).

For the collector diffusion capacitor, C_{DC}, one additional model parameter is required:

τ_R The (assumed constant) total reverse transit time (or the saturation time constant, τ_{SAT}, from which τ_R is obtained).

Lastly, the substrate capacitor, C_{SUB}, is specified directly:

C_{SUB} The (assumed constant) capacitance between the substrate and the collector for npn transistors, or between the substrate and the base for lateral pnp transistors.

In computer programs where the default model is the EM_1 model, parameters r'_c, r'_e, r'_b, C_{jEO}, C_{jCO}, τ_F, τ_R, and C_{SUB} all default to zero, if not otherwise specified. The default values for ϕ_E, ϕ_C, m_C and m_E need not be zero.

It should be emphasized that the EM_2 model is adequate for the majority of cases, especially for analyzing digital circuits. However, there are still some limitations, including the absence of such dc effects as basewidth modulation and the variation of β with current level. Both of these second-order effects are accounted for in the EM_3 and GP models which are described in the next two sections.

The collector-base junction capacitance is actually distributed across the base resistance. The use of a single capacitor in the EM_2 model is a first-order representation of this distributed capacitor. An improved, second-order model is given next in the EM_3 model. Other (second-order) improvements in the EM_3 model are a more complete treatment of the effects of temperature and the variation of τ_F with collector current.

2.4 The EM_3 Model

The EM_3 model is the third level of complexity in the non-linear modeling of the bipolar junction transistor. While the EM_1 is a simple dc model, and the EM_2 model contains a first-order representation of charge-storage effects and ohmic resistances, the EM_3 model is concerned with second-order improvements in the dc aspects of the EM_2 model, charge-storage modeling and temperature performance. The EM_3 model adds the following features:

- base-width modulation and variation of β with current and voltage.
- the ability to split the collector-base junction capacitance across r_b'.
- the rise of τ_F at high currents.
- variation of device parameters with temperature.

These effects are mainly incorporated by modifying existing equations. Added components are two diodes and a junction capacitor. Three extra model parameters are required for the variation of forward current gain with collector current (β_F vs I_C), three for the variation of inverse current gain with emitter current (β_R vs I_E), one for basewidth modulation, one for the junction capacitance split, two for the variation of τ_F with current level and six for the temperature variation.

The dc improvements described in this section are basically modeled equivalently in the GP model (which is the subject of the next section). In this EM_3 section, emphasis is placed mainly on a description of the effect being modeled and a brief justification for the method of incorporating it into the model. In the GP model, a more detailed derivation of the underlying physics is given. The EM_3 model[31] is essentially available in the computer programs SLIC[1] and SINC.[2] Although its dc treatment is not quite as complete as that of the Gummel-Poon model, its input parameters are easier to understand and determine.

2.4.1 An Improved dc Model at a Given Temperature

The dc model is improved by incorporating two second-order effects: basewidth modulation and the variation of β with operating current.

a) Basewidth modulation

Basewidth modulation (the so-called "Early Effect")[32] describes the change in basewidth that results from a change in the collector-base junction voltage. In the normal, active region the emitter-base junction is forward biased and the collector-base junction is reverse biased. The width of the space-charge layer of a p-n junction is a strong function of the applied potential. Large variations in V_{BC}, for example, may cause the collector-base space-charge layer to vary significantly. This, in turn, changes the normally thin basewidth.

The total effect of basewidth modulation on the device characteristics in the normal, active region is a modification (as a function of V_{BC}) of

- I_S (and thereby the collector current)
- β_F
- τ_F

These three model parameters are affected because of their strong dependence on the basewidth, W.

Only one extra parameter, the Early voltage, V_A, is used to model basewidth modulation in the forward, active region. Figure 2.13 shows the effect of basewidth modulation on the variation of collector current (I_C) with collector-emitter voltage (I_C vs V_{CE}) -- a non-zero slope in the normal, active region. (The dashed lines illustrate the zero slope obtained with the EM_2 model, where basewidth modulation is not included.)

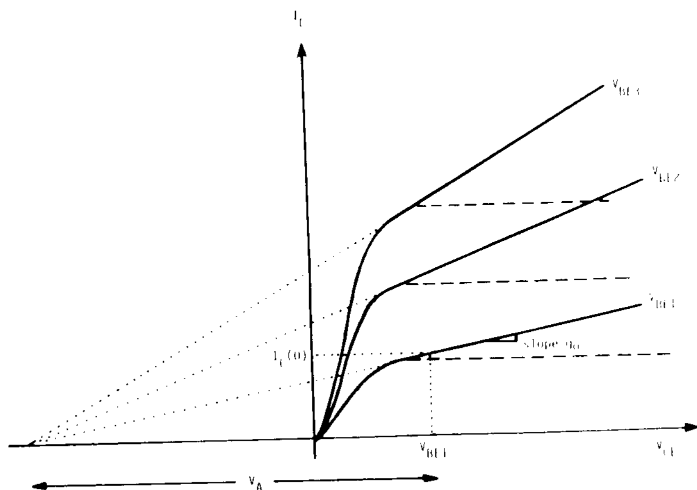


Fig. 2.13. The effect of basewidth modulation on the I_C versus V_{CE} characteristics. The dashed lines represent the characteristics obtained with the EM_2 model. (Not to scale.)

The dependence of the above three model parameters on V_{BC} (via the basewidth modulation phenomenon) has been derived by McCalla[31] and is given in Appendix 2. The analysis, which assumes that the transistor is operated in the linear region, first determines the effect of basewidth modulation on the basewidth and then on the three basewidth-related parameters.* The results of the analysis are:

$$W(V_{BC}) = W(0)\left(1 + \frac{V_{BC}}{V_A}\right) \tag{2.41}$$

*A more detailed consideration and justification of the assumptions made in the analysis is given in Appendices 2 and 4.

$$I_S(V_{BC}) = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)} \approx I_S(0)\left(1 - \frac{V_{BC}}{V_A}\right) \tag{2.42}$$

$$\beta_F(V_{BC}) = \frac{\beta_F(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)} \approx \beta_F(0)\left(1 - \frac{V_{BC}}{V_A}\right) \tag{2.43}$$

$$\tau_B(V_{BC}) = \tau_B(0)\left(1 + \frac{V_{BC}}{V_A}\right)^2 \tag{2.44}$$

where V_A is defined, for an npn transistor,* as

$$V_A \triangleq \left[\frac{1}{W(0)} \cdot \frac{dW}{dV_{BC}}\bigg|_{V_{BC} = 0}\right]^{-1} \tag{2.45}$$

Equation (2.41) describes the (assumed linear) variation of the basewidth with V_{BC}. Equations (2.42), (2.43) and (2.44), which give the variation of the three model parameters with V_{BC}, follow directly from Eq. (2.41) and the assumptions that the constant-base doping relationships $I_S \propto \frac{1}{W}$,[33] $\beta_F \propto \frac{1}{W}$[34] and $\tau_B \propto W^2$[26] are also approximately valid in general. An expression similar to Eq. (2.44) can be obtained for τ_{BR}, but since τ_{BR} is normally only a very small component of τ_R, it has not been included.

The second forms of Eqs. (2.42) and (2.43) are preferred computationally since the first forms become infinite at $V_{CB} = V_A$. ($V_{CB} = V_A$ may not necessarily be the correct value but could be a temporary value while the computer is iterating to the solution.) The second form assumes that $|V_{BC}| << V_A$. Since V_A is typically on the order of 50 V and basewidth modulation is itself a second-order effect, the analysis can give acceptable accuracy even when the value of V_{BC} is close to that of V_A.

---

*For a pnp transistor a minus sign should be introduced or V_{EC} changed to V_{CB}.

V_A has no physical counterpart in the circuit model; only a mathematical effect whereby existing equations are modified. (This process of altering equations or parameters without altering the form of the equivalent circuit will be observed for other effects in the EM_3 and GP models.) With the exception of r_B, which is described later, the total effect of basewidth modulation is accounted for if I_S and β_F are modified as in Eqs. (2.42) and (2.43). The expressions for I_{CT} and I_B (Eqs. (2.17) and (2.19), respectively) then become:

$$I_{CT} = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)} \left[\left(e^{\frac{qV_{BE}}{kT}} - 1 \right) - \left(e^{\frac{qV_{BC}}{kT}} - 1 \right) \right] \qquad (2.46)$$

$$I_B = \frac{I_S(0)}{\beta_F(0)} \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) + \frac{I_S(0)}{\beta_R} \left(e^{\frac{qV_{BC}}{kT}} - 1 \right) \qquad (2.47)$$

In the first term of Eq. (2.47) the similar dependence of I_S and β_F on V_{BC} (since both are assumed to be $\propto \frac{1}{W}$) results in a cancellation. Therefore, in the normal, active region where the second term is negligible, I_B is independent of V_{BC}. The V_{BC} variation in β_F is achieved by keeping I_B constant and modifying I_C. This introduces a very important concept which will be illustrated again later. The correct variation of β_F (with V_{BC} as here, or I_C, as later) is not obtained by varying β_F but by modeling correctly the expressions for I_C and/or I_B. A similar analysis for basewidth modulation when the device is operated in the inverse mode is included in the GP model.

The Early voltage V_A can be obtained directly from the I_C vs V_{CE} characteristics. This can be shown mathematically. The slope of these characteristics in the normal, active region, g_o, is obtained from Eq. (2.46) by first dropping the (negligible) second term and then differentiating with respect to V_{BC} (V_{BE} assumed

constant). The result, derived in Appendix 2, is:

$$g_o \triangleq \left. \frac{dI_C}{dV_{CE}} \right|_{V_{BE} = \text{const.}} \simeq \frac{I_C(0)}{V_A} \tag{2.48}$$

(In SLIC and SINC, $r_o(=1/g_o)$ and $I_C(0)$ are acceptable input parameters, from which V_A is obtained by Eq. (2.48).)

The geometrical interpretation of Eq. (2.48) shows that V_A is obtained from the intercept of the extrapolated slope on the V_{CE} axis (as shown for the V_{BE1} curve in Fig. 2.13). For example, a slope of $(50\ k\Omega)^{-1}$ at $I_C(0) = 1$ mA gives, from Eq. (2.48), $V_A = 50$ V. A more detailed description of the geometrical interpretation is given in Appendix 2.

b) β_{dc} variation with current

In general there are three regions of interest in the variation of β_{dc} with current. Figure 2.14 shows a typical variation of β_F with I_C. Region I is the low-current region in which β_F increases with I_C. Region II is the mid-current region in which β_F is constant $(\triangleq \beta_{FM})$. Region III is the high-current region in which β_F drops as the current is increased. Before analysing these regions, several points should be noted about Fig. 2.14.

Variation with V_{BC}. The curve in Fig. 2.14 is drawn for constant V_{BC}, in this case $V_{BC} = 0$. The variation of β_F with V_{BC} has just been covered. In the following it will be assumed that all data corresponds to $V_{BC} = 0$. Non-zero V_{BC} data can be reduced appropriately by the application of Eq. (2.43).

β_R Variation. For simplicity, most of the following analysis considers only the variation of β_F with I_C. A similar analysis can be performed for the variation of β_R with I_E at constant V_{BE}. Results for β_R variations are given where appropriate.

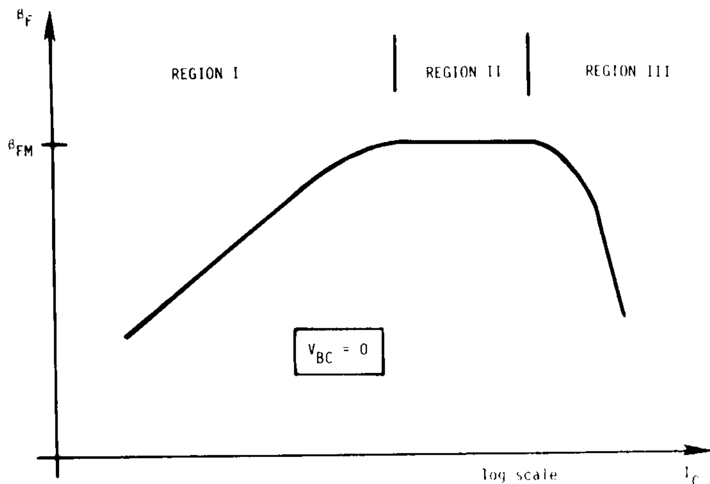


Fig. 2.14. Typical variation of β_F with I_C illustrating the three regions of interest. (Not to scale.)

Usefulness of Fig. 2.14. It will be shown that Region I is governed by additional components of I_B, while Region III results from a change in I_C. This information is not evident from Fig. 2.14. Therefore, an alternative form of presenting the above information is used: one in which more information and a clearer understanding of the device operation is obtained. This alternate form is a plot of $\ln(I_C)$ and $\ln(I_B)$ as a function of $V_{B'E'}$, as shown in Fig. 2.15.* Because of the logarithmic nature of the vertical axis, β_F is obtained directly from the plot as the distance between the I_C and I_B curves. Not only is it evident from this plot what causes the variation of β_F with I_C, but all the model parameters needed to characterize this variation can be obtained directly from it. <u>One of the most important concepts in the modeling of</u>

*Acceptable alternatives to using the natural logarithim is to use $\log_{10}$ or more simply to use semi-log graph paper. However, the slopes of the lines described here will assume graphs of ln(I) vs V.

transistors is this preference of the data in the form of Fig. 2.15, over that of Fig. 2.14. Data in the form of Fig. 2.15 will be used through the following analysis. It is still assumed that $V_{BC} = 0$ for all points.

Existence of Region II. For some transistors there may not appear to be a region in which β_F is constant. For these transistors Regions I and III have simply overlapped. The analysis and subdivision into the three regions is still valid, since the model parameters can still be obtained from Fig. 2.15 even when Region II in the β_F vs I_C curve does not exist.

Ohmic resistances. The analysis of the variation of β_F with I_C now proceeds, by region. In this analysis, it will first be assumed that the ohmic resistances r_b', r_e' and r_c' are all zero such that $V_{BE} = V_{B'E'}$ and $V_{BC} = V_{B'C'}$. The effects of this assumption will be examined at the end of this analysis.

i) Region II: mid-currents

In this region, the EM_1 model holds; the β_F used in the EM_1 model applies only to Region II and is now called β_{FM}. The two currents in this region (for $V_{BC} = 0$) are given by:

$$\left.\begin{aligned} I_C &= I_S(0)\left(e^{\frac{qV_{BE}}{kT}} - 1\right) && (2.49) \\ I_B &= \frac{I_S(0)}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) && (2.50) \end{aligned}\right\} \quad V_{BC} = 0$$

From Fig. 2.15, values of $\beta_{FM}(0)$ and $I_S(0)$ can be obtained directly, as explained in the measurement section.

ii) Region I: low currents

The drop in β_F at low currents is caused by extra components of I_B that until now have been ignored. For the normal, active region

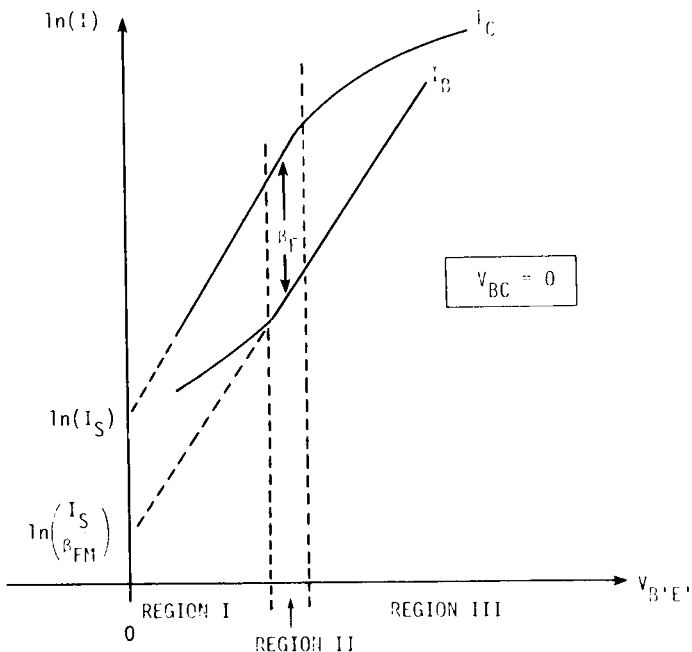


Fig. 2.15. **A plot of $\ln(I_C)$ and $\ln(I_B)$ as a function of $V_{B'E'}$ which gives a clearer understanding of Regions I, II, and III.**

with $V_{BC} = 0$ there are three extra components which are caused by[35]

- the recombination of carriers at the surface,
- the recombination of carriers in the emitter-base space-charge layer, and
- the formation of emitter-base surface channels.

All three components have a similar variation with base-emitter voltage, V_{BE}:[36]

$$I_{B(surface)} = I_{S(surface)}\left(e^{\frac{qV_{BE}}{2kT}} - 1\right) \quad (2.51)$$

$$I_{B(EB\ scl)} = I_{S(EB\ scl)}\left(e^{\frac{qV_{BE}}{2kT}} - 1\right) \quad (2.52)$$

$$I_{B(channel)} = I_{S(channel)}\left(e^{\frac{qV_{BE}}{4kT}} - 1\right) \quad (2.53)$$

These three components should each be added to the base current, Eq. (2.50). Fortunately, a simplification can be made. A composite current can be made of all three extra components which has the form:

$$I_{B(composite)} \triangleq I_{B(surface)} + I_{B(EB\ scl)} + I_{B(channel)}$$

$$\simeq I_{S(composite)}\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right) \quad (2.54)$$

where n_{EL} is called "the low-current, forward region emission coefficient" and lies between 1 and 4.*

For most cases, a fit to Eq. (2.54) can be made with reasonable accuracy.** Therefore, at $V_{BC} = 0$, the base current is approximated by

$$I_B = \frac{I_S(0)}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) + C_2 I_S(0)\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right) \quad (2.55)$$

---

* In SPICE,[3] n_{EL} is called NE. The change in notation has been made here to emphasize its applicability to low currents only.

** Since channeling and surface recombination can both be made small with careful processing, the dominant component is normally the recombination in the emitter-base space charge layer and n_{EL} is normally close to 2.

where the term $I_{S(composite)}$ in Eq. (2.54) has been replaced by $C_2 I_S(0)$ (i.e., it has simply been normalized to $I_S(0)$). The two additional model parameters are C_2 and n_{EL}.

Inverse region model at low currents. When the base-collector junction is forward biased, there will generally be three similar additional components of I_B at low current levels -- surface recombination, collector-base space-charge layer recombination and collector-base channeling. In a similar way they can be lumped together into a composite component that depends on V_{BC}. The expression for I_B in general then becomes

$$I_B = \frac{I_S(0)}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) + C_2 I_S(0)\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right)$$

$$+ \frac{I_S(0)}{\beta_{RM}}\left(e^{\frac{qV_{BC}}{kT}} - 1\right) + C_4 I_S(0)\left(e^{\frac{qV_{BC}}{n_{CL}kT}} - 1\right) \qquad (2.56)$$

where the two extra model parameters n_{CL} (the low-current, inverse-region emission coefficient) and C_4 have been introduced.*

Effect on equivalent circuit. The additional components of base current I_B are included in the circuit model by means of two non-ideal diodes, as shown in Fig. 2.16. The circuit to which these diodes have been added is the EM_1 model, in the nonlinear hybrid-π form.** The EM_1 model was used only for the sake of simplicity

*Again, as for n_{EL}, the SPICE notation has been altered from NC to n_{CL}.

**The simple addition of the two non-ideal diodes points out the second advantage of the nonlinear hybrid-π form. (The first is its similarity to the linear hybrid-π model.) The two ideal diodes had previously been identified as the two ideal components of base current. The extra components of I_B were therefore simply added by a natural extension of the model. This can be carried further. For those rare cases where the composite component of Eq. (2.54) is not accurate enough, the equivalent circuit is very simply extended by adding more non-ideal diodes for each of the extra components of I_B. A non-ideal diode is defined here to be one with a non-unity emission coefficient.

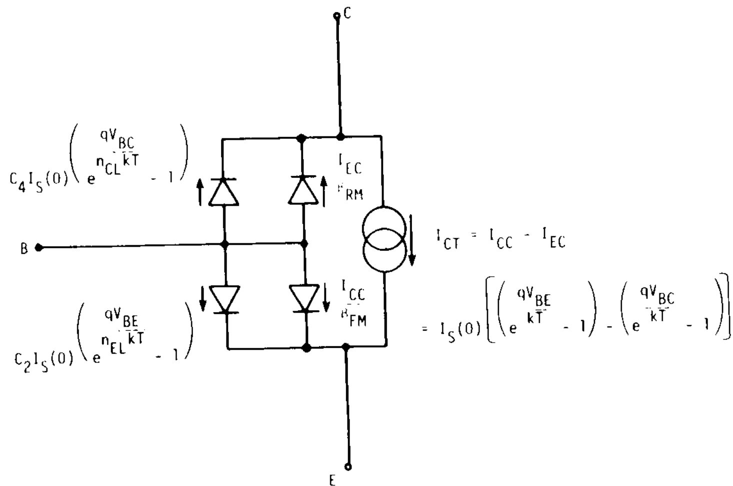


Fig. 2.16. **The inclusion of two diodes that model the extra components of I_B into the EM_I model (nonlinear hybrid-π form).**

in illustration: the capacitors of the EM_2 model have no effect in this dc analysis and the effect of the ohmic resistances r'_b, r'_c and r'_e are considered later. All these components can be added into the model, as was done before.

The plot of $\ln(I_B)$ vs $\frac{qV_{BE}}{kT}$ for $V_{BC} = 0$, shown in Fig. 2.17, illustrates the two components of base current I_B: the ideal component with the slope of 1 and the non-ideal component with the slope of $\frac{1}{n_{EL}}$. The extrapolation of these straight-line components to the line defined by $V_{BE} = 0$ gives the values of $C_2 I_S(0)$ and $\frac{I_S(0)}{\beta_{FM}(0)}$. A similar plot of $\ln(I_B)$ as a function of V_{BC} for inverse operation yields values for the model parameters C_4 and n_{CL}. A typical value for C_2 (and C_4) is 10^3, and a typical value for n_{EL} (and n_{CL}) is 2.

iii) Region III: high currents*

At high injection levels, the injection of minority carriers into the base region is significant with respect to the majority carrier concentration. Since space-charge neutrality is maintained in the base, the total majority carrier concentration is increased by the same amount as the total minority carrier concentration. The effect of the excess majority carriers on the collector current has been calculated by Webster(37) who showed that at high levels the collector current asymptotes to

$$I_{C(\text{high level})} \propto e^{\frac{qV_{BE}}{2kT}} \tag{2.57}$$

Equation (2.57) is incorporated into the EM_3 model by modifying the collector current expression for $V_{BC} = 0$, Eq. (2.49), to(31,1)

$$I_C(0) = \frac{I_S(0)}{\left(1 + \theta \; e^{\frac{qV_{BE}}{2kT}}\right)} \left(e^{\frac{qV_{BE}}{kT}} - 1\right) \tag{2.58}$$

where θ is the additional model parameter.

Equation (2.58) has the appropriate asymptotes (Eq. (2.49) at low currents and Eq. (2.57) at high currents).**

---

*There are two causes for the drop in β_F at high currents: saturation and the effect of high-level injection. Since saturation is modeled separately by r_C' (and the inherent saturation of the EM_1 model), this section concentrates only on the latter cause. It is therefore assumed that all data is obtained in the normal, active region.

**Note that Eq. (2.58) also provides a mathematical "definition" of high-level injection. That is, when

$$\theta \; e^{\frac{qV_{BE}}{2kT}} \gg 1$$

The extrapolation of the high-current asymptote of $\ln(I_C)$ vs $\frac{qV_{BE}}{kT}$ with $V_{BC} = 0$ (Fig. 2.17) to $V_{BE} = 0$ gives the value $I_S(0)/\theta$. The high-current asymptote has a slope half that of the low-current curve. The intersection of the two collector current asymptotes defines the point (I_K, V_K). This point will be used in the GP model.

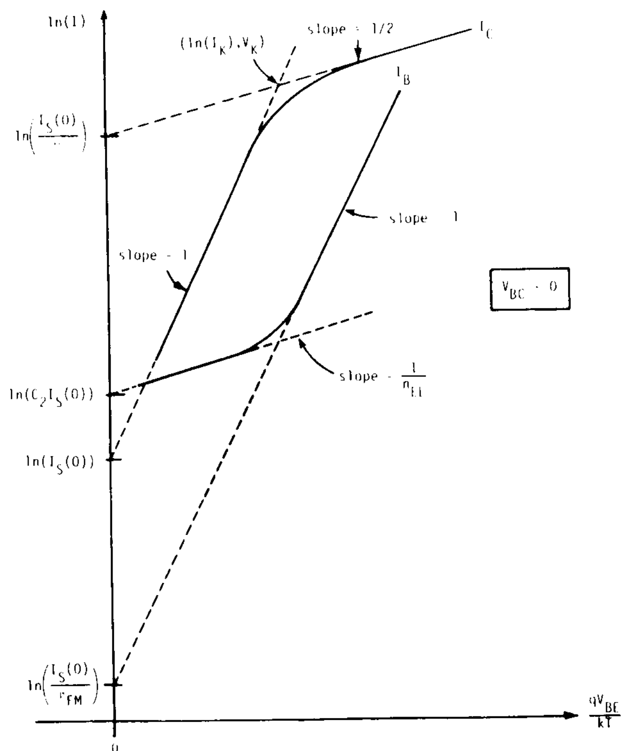


Fig. 2.17. The complete plot of $\ln(I_C)$ and $\ln(I_B)$ versus $\frac{qV_{BE}}{kT}$ for $V_{BC} = 0$. Note that here r_b' and r_e' are assumed to be zero.

iv) The full picture: regions I, II and III

The EM_3 model parameters required to describe completely the variation of β_F with I_C at $V_{BC} = 0$ are:

$$\begin{array}{ll} \beta_{FM} & \text{Region II} \\ \left.\begin{array}{l} C_2 \\ n_{EL} \end{array}\right\} & \text{Region I} \\ \theta & \text{Region III} \end{array}$$

The values of these four β_F vs I_C model parameters can all be obtained directly from Fig. 2.17. A similar set (β_{RM}, C_4, n_{CL}, θ_R) can be defined from the variation of β_R with I_E at $V_{BE} = 0$.

v) Effect of ohmic resistances

So far, the ohmic resistances r_c', r_b' and r_e' have been neglected. To first order, they do not affect the above analysis but they do affect the experimental data. The voltages in the above analyses should be the internal voltages $V_{B'E'}$ and $V_{B'C'}$. The externally measured value of V_{BC} should be corrected to $V_{B'C'}$ by subtracting $(I_C\, r_c' + I_B\, r_b')$; normally this has a small effect since V_{BC} only influences the basewidth modulation correction factor. The effect of r_b' and r_e' is an increase in V_{BE}, which can be seen from Fig. 2.18. To correct for the effects of r_b' and r_e', it is assumed that the slope of $\ln(I_B)$ vs $V_{B'E'}$ stays constant (the dashed line). Therefore, the value of $(I_B\, r_b' + I_E\, r_e')$ can be found from the distance marked A. The appropriate amount (marked B) is subtracted from the $\ln(I_C)$ curve and the resultant curve (the dashed I_C line) becomes that of Fig. 2.17 where the horizontal axis is now $V_{B'E'}$. This technique is explained in detail in the Measurement Section 3.4 (p. 196).

One additional complicating factor must be mentioned. The above analysis assumed constant junction temperature (T). At high current levels, care must be taken to ensure that excessive power dissipation does not heat the junction. To this end, high-current

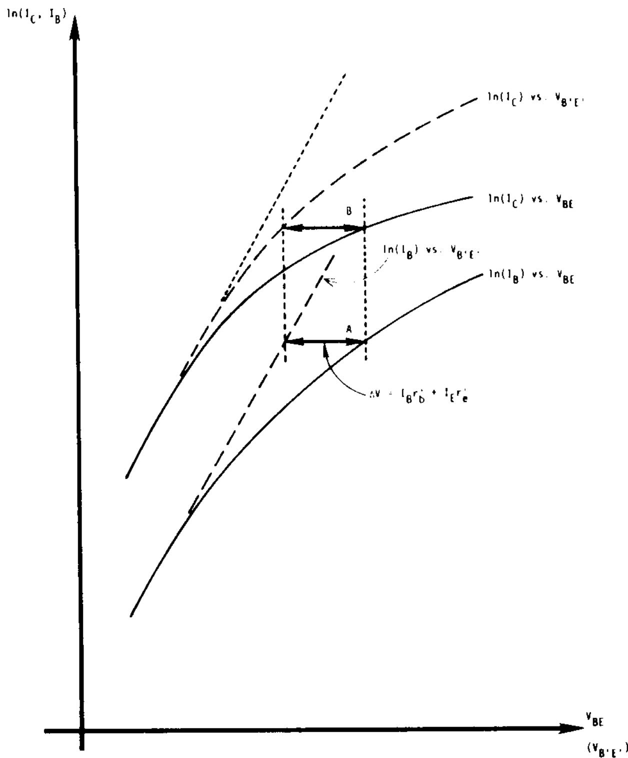


Fig. 2.18. The effect of r_b' and r_e' on the $\ln(I_C)$ and $\ln(I_B)$ versus V_{BE} characteristics.

measurements should be made in a pulsed mode and/or with V_{BC} as close to zero as possible (without entering saturation).

vi) β_F versus I_C input parameters in SLIC and SINC[31,1]

In SLIC and SINC, the <u>input</u> parameters needed to describe the variation of β_F with I_C are different from the above four <u>model</u> parameters (β_{FM}, C_2, n_{EL}, θ): the input parameters are designed to be more useful for the design engineer who deals with β_F vs I_C rather than $\ln(I_C)$ and $\ln(I_B)$ vs V_{BE}. To explain the input parameters it is first necessary to derive the equation for β_F vs I_C.

The equation for β_F vs I_C at $V_{BC} = 0$ is built up from a consideration of each region. For convenience, β_F^{-1} will be found.

In Regions I and II, β_F^{-1} is given by:

$$\beta_F^{-1} \triangleq \frac{I_B}{I_C} = \frac{\dfrac{I_S(0)}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) + C_2 I_S(0)\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right)}{I_S(0)\left(e^{\frac{qV_{BE}}{kT}} - 1\right)}$$

$$= \frac{1}{\beta_{FM}(0)} + \left[C_2 I_S(0)^{(1-1/n_{EL})}\right] I_C^{(1/n_{EL}-1)}$$

$$\triangleq a_1 \quad + \quad a_2 \qquad I_C^{(1/n_{EL}-1)} \tag{2.59}$$

In Region III, β_F^{-1} is given by:

$$\beta_F^{-1} = \frac{I_B}{I_C} = \frac{\dfrac{I_{S(0)}}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}}\right)}{\dfrac{I_S(0)}{\theta}\left(e^{\frac{qV_{BE}}{2KT}}\right)}$$

$$= \frac{\theta^2}{\beta_{FM}(0)\, I_S(0)} I_C$$

$$\triangleq a_3 I_C \tag{2.60}$$

The combination of Eqs. (2.59) and (2.60) gives, for all regions:[38,31]

$$\beta_F^{-1} = a_1 + a_2 I_C^{(1/n_{EL}-1)} + a_3 I_C \tag{2.61}$$

where

$$a_1 \triangleq \beta_{FM}^{-1} \tag{2.62}$$

$$a_2 \triangleq C_2 I_S^{(1-1/n_{EL})} \tag{2.63}$$

$$a_3 \triangleq \frac{\theta^2}{\beta_{FM} I_S} \tag{2.64}$$

Equation (2.61) assumes no interaction between Regions I and III. From Eq. (2.61), the four input parameters could be a_1, a_2, a_3 and n_{EL}. However, one further transformation is made to the input parameters.

The input parameters required are: β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE}. β_{FMAX} is the maximum value of β_F, I_{CMAX} is the collector current at which β_{FMAX} occurs, β_{FLOW} is any value of β_F at a collector current less than I_{CMAX}, I_{CLOW} is the collector current at which β_{FLOW} is measured, BCEC equals $1/n_{EL}$ and V_{CE} is the value at which all this data is obtained.*

Basewidth modulation effects are first removed by multiplying the first four input parameters by $\left[1 + \left(\frac{V_{BC}}{V_A}\right)\right]$. Then, the following equations are solved for a_1, a_2 and a_3.[31]

$$\beta_{FMAX}^{-1} = a_1 + a_2\, I_{CMAX}^{(BCEC-1)} + a_3 I_{CMAX} \tag{2.65}$$

$$\beta_{FLOW}^{-1} = a_1 + a_2\, I_{CLOW}^{(BCEC-1)} + a_3 I_{CLOW} \tag{2.66}$$

$$I_{CMAX}^{(BCEC-2)} = \frac{a_3}{a_2}\left[1 - BCEC\right]^{-1} \tag{2.67}$$

where Eq. (2.67) is obtained by differentiating Eq. (2.61).

---

*In SINC and some versions of SLIC, BCEC is internally fixed at 0.5.

vii) β_{Fac} and its variation with current

β_{Fac} is defined as

$$\beta_{Fac}^{-1} \triangleq \frac{dI_B}{dI_C}$$

$$= \beta_F^{-1} + I_C \frac{d\beta_F^{-1}}{dI_C} \tag{2.68}$$

From Eq. (2.61),

$$\beta_{Fac}^{-1} = a_1 + \frac{a_2}{n_{EL}} I_C^{\left(\frac{1}{n_{EL}} - 1\right)} + 2\, a_3\, I_C \tag{2.69}$$

Therefore, a knowledge of the β_{Fdc} vs I_C model parameters also gives a complete description of β_{Fac} vs I_C.

2.4.2 An Improved Charge-Storage Model at a Given Temperature

The improvement in the charge-storage model is achieved by modeling two second-order effects: the ability to split the collector-base junction capacitance across r_b' and the variation of τ_F with collector current.

a) Improved C_{jC} - r_b' Model

Ideally, the collector-base junction capacitor should be modeled as a distributed capacitor across r_b' as shown in Fig. 2.19a. A first-order model of this distributed capacitor is used in the

EM_2 model in which all of C_{jC} lies on the inside of r_b'. An improved representation is available in the EM_3 model in that C_{jC} can be split up on either side of r_b', as shown in Fig. 2.19b. An extra model parameter, RATIO, is needed. RATIO lies between zero and unity.

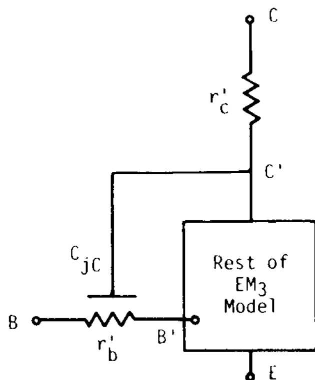


Fig. 2.19a. Diagram of C_{jC} distributed across r_b'.

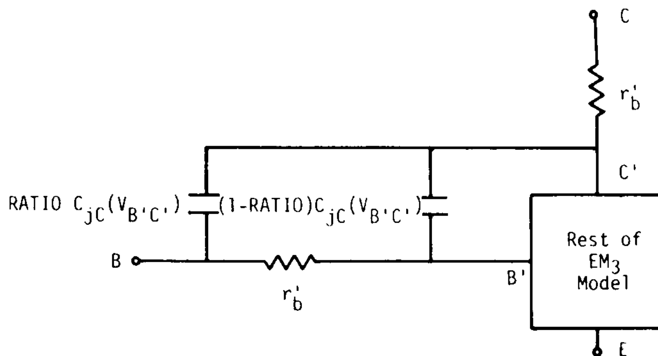


Fig. 2.19b. The EM_3 model of C_{jC} distributed across r_b'.

b) Variation of τ_F with current

The variation of τ_B with V_{BC} has already been covered in Section 2.4.1 on basewidth modulation (Eq. (2.44)). τ_B also varies with collector current.

There are three basic causes for the increase in τ_B with I_C at high currents: a reduction of the low-level aiding-field effect in drift transistors at high-level injection,[26] an effective base-widening described by Kirk[39] and a two-dimensional spreading effect described by van der Ziel and Agouridis.[40] There is still some controversy as to which of the base-widening effects dominate.[41,42,43] The true cause for the increase in τ_B at high currents probably lies in a combination of these effects ranging all the way from mostly one effect in one device to mostly the other effect in another device.

There are several techniques for modeling τ_F vs I_C. Gummel and Poon[7] use a multiplier B which requires 4 model parameters. In SLIC and SINC, it is assumed that the two-dimensional effect dominates. In the normal, active region, this results in the equation:[41]*

$$\tau_{F_{ac}}(I_C) \triangleq \frac{dQ_{DE}}{dI_{CC}} = \tau_{FL}(0)\left[1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2\left(\frac{I_{CC}}{I_{CO}} - 1\right)^2\right] \text{ for } I_C \geq I_{CO} \tag{2.70}$$

where $\tau_{FL}(0)$ is the low current value of the forward transit time (previously just called τ_F), L_E is the smallest width of the emitter, W is the basewidth, and I_{CO} is the current at which τ_F starts to rise. The variation of $\tau_{F_{ac}}$ with I_C as described by Eq. (2.70) is sketched in Fig. 2.20.** Two extra model parameters are

---

*Equation (2.70) actually applies to the τ_B component of τ_F. It is assumed that τ_F can also be fitted to this expression.

**In the normal, active region, $I_{CC} = I_C$.

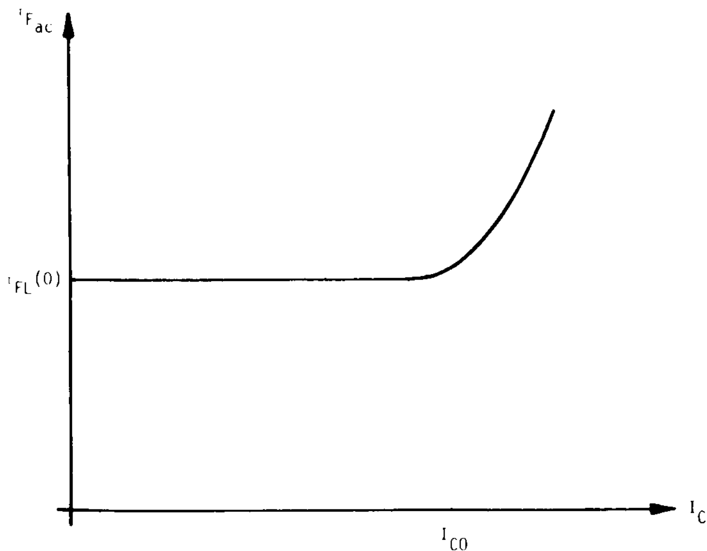


Fig. 2.20. The variation of τ_F with I_C as described by Eq. (2.70).

required: I_{CO} and the ratio $\left(\frac{L_E}{W}\right)$. These parameters are obtained by curve-fitting the experimental variation of $\tau_{F_{ac}}$ with I_C to Eq. (2.70) at high currents. The rise in $\tau_{F_{ac}}$ at high currents is observed as a drop in f_T at high currents.

The variation of $\tau_{F_{ac}}$ with current is caused by an effective increase in the basewidth. This will also affect the dc characteristics.* To take this variation into account, an appropriate change can be made in θ and θ_R in the EM_3 model (which are proportional to $\sqrt{\tau_B}$ and $\sqrt{\tau_{BR}}$, respectively).

---

*Since $\tau_{F_{dc}}$ is no longer equal to $\tau_{F_{ac}}$, it is given in this case by:

$$\tau_{F_{dc}} \triangleq \frac{Q_{DE}}{I_{CC}} = \frac{\int_0^{I_{CC}} \tau_{F_{ac}}\, dI_{CC}}{I_{CC}} = \tau_{FL}\left\{1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2 \left[\frac{1}{3}\left(\frac{I_{CC}}{I_{CO}}\right)^2 - \left(\frac{I_{CC}}{I_{CO}}\right) + 1 - \frac{1}{3}\left(\frac{I_{CO}}{I_{CC}}\right)\right]\right\}$$

for $I_{CC} \geq I_{CO}$.

2.4.3 An Improved Variation with Operating Temperature[31]

Thus far the only model parameter that has varied with temperature has been I_S, which was described in the EM_1 model. An improved temperature model is introduced here. Two types of improved temperature variations are considered: those that can be expressed in terms of the device physics and those that are modeled by an empirical data fit. Only the latter variations require extra model parameters.

a) Physics-based temperature variation

The model parameters whose temperature variation can be modeled by consideration of the device physics are τ_F, C_{jE} and C_{jC}.

(i) τ_F

McCalla[31] has shown that the temperature variation of the basewidth is given by

$$\frac{W(V_{BC},T)}{W(0,T)} = \frac{V_{BC}}{V_A} + \frac{\phi_C}{V_A}(T-T_{nom})\left(\gamma_T^{\varepsilon} - \gamma_T^{\phi_C}\right) \qquad (2.71)$$

where

ϕ_C is the collector-base barrier potential

T_{nom} is the temperature at which the model parameters were measured

T is the temperature at which the parameter values are to be calculated (the operating condition)

γ_T^{ε} is the sensitivity of the dielectric constant to temperature (typically 200 ppm/°C for silicon)

and $\gamma_T^{\phi_C}$ is the sensitivity of ϕ_C to temperature and is given by

$$\gamma_T^{\phi C} = \frac{1}{T_{nom}} - \frac{kT_{nom}}{q\phi_C}\left(\frac{3}{T_{nom}} + \frac{E_g}{kT^2_{nom}}\right) \qquad (2.72)$$

Equation (2.72) assumes the impurity concentrations on either side of the junction are approximately constant. Other assumptions involved in the derivation of Eq. (2.71) are: a Taylor expansion of W about $V_{BC} = 0$ and $T + T_{nom}$, $|V_{BC}| << V_A$ and the temperature correction is small. The temperature dependence of τ_F is then given by

$$\tau_F(T) = \tau_F(T_{nom}) \left[\frac{W(T)}{W(T_{nom})}\right]^2 \left(\frac{T}{T_{nom}}\right)^{1.5} \qquad (2.73)$$

where the W^2 term results from the dependence of τ_F on W^2 and the $T^{1.5}$ term results from the dependence of τ_F on D (i.e., $\tau_F \propto \frac{W^2}{D}$).* Equations (2.71), (2.72) and (2.73) give the temperature variation of τ_F.**

(ii) C_{jC} and C_{jE}

Equation (2.72) (and a similar one for $\gamma_T^{\phi E}$) is used in the determination of the temperature variation of C_{jC} (and C_{jE}). McCalla[31] has also shown that for each junction capacitor

$$C_j(T) = C_j(T_{nom}) \left[1 + (T-T_{nom})\, m \left(2\gamma_T^{\varepsilon} - \gamma_T^{\phi}\right)\right] \qquad (2.74)$$

All these temperature dependences are built into the program SLIC.

---

*Actually, the τ_B component of τ_F is the term that is proportional to W^2/D. It is assumed here that τ_F has the same temperature dependence.

**Note that it becomes necessary to have an accurate value for ϕ_C if the correct temperature variation of τ_F is required. Previously, ϕ_C was only used to compute C_{jC}.

b) Temperature-dependent parameters that require extra input parameters

The parameters described in this section require extra model parameters to model their temperature variation. The extra parameters are used to fit the general empirical relationship.

$$Par(T) = Par(T_{nom}) \left[1 + TC_1 (T-T_{nom}) + TC_2 \; (T-T_{nom})^2 \right] \qquad (2.75)$$

where Par is the parameter being varied

TC_1 is the first-order temperature coefficient

and TC_2 is the second-order temperature coefficient.

In SLIC and SINC Eq. (2.75) is applied to the parameters β_F, r_b' and r_c'. A theoretical expression for the temperature variation of these parameters (which model the effect on emitter injection efficiency, transport efficiency, crowding and variation in base conductivity) is not available. The extra model parameters are TC_1 and TC_2 for β_F, r_b' and r_c' (a total of 6 extra input parameters). The model parameters that have not been mentioned in this Section are assumed to be invariant with temperature.

2.4.4 Small-Signal (Linearized) EM_3 Model

The small-signal, linearized EM_3 model, which is described in Appendix 5, is very similar to that given for the EM_2 model in Section 2.3.3. There are only two basic differences:

- as a result of the C_{jC} - r_b' split, an extra capacitor (equal to RATIO C_{jC} $(V_{B'C'})$) is placed between nodes B and C', and
- the equations determining r_π, r_μ, C_π and C_μ are changed to include the effects incorporated in the EM_3 model.

2.4.5 Summary and Conclusions

The EM_3 model has improved the dc characteristics (basewidth modulation and β vs I), the charge-storage model (C_{jC} - r_b' split and τ_F vs I_C) and the variation with operating temperature.

Basewidth modulation is included by modifying I_S, β_F and τ_F. Thus these parameters must be specified at $V_{BC} = 0$. An Early voltage, V_A, is used to describe basewidth modulation by the collector-base junction (the Early effect). Whereas, in the normal, active region, the $\ln(I_C)$ vs V_{BE} curve is affected by basewidth modulation, the $\ln(I_B)$ vs V_{BE} curve is, to first order, unaffected.

β vs I is a result of the correct modeling of ln(I) vs V. The low-current drop in β is caused by extra components of I_B which can be described by four model parameters, C_2 and n_{EL} (for β_F) and C_4 and n_{CL} (for β_R). Two non-ideal diodes were added to the circuit model. The drop in β at high currents is caused by high-level injection effects in the base which can be described by two model parameters (θ for β_F and θ_R for β_R). A set of input parameters more familiar to the circuit designer is used in the EM_3 model. The usefulness of the ln(I) vs V curve was established: All of the model and input parameters are directly obtainable from it. Effects of ohmic resistances and change of temperatures must be considered.

Improved charge storage is obtained by splitting C_{jC} across r_b' (requiring one extra parameter, RATIO) and by including the variation of τ_F with I_C (the two-dimensional effect of van der Ziel and Agouridis,[40] available in the EM_3 model, requires two extra parameters, I_{CO} and $\frac{L_E}{W}$).

Improved temperature variation is obtained in the EM_3 model by the use of 6 parameters (3 sets of first- and second-order temperature coefficients for r_b', r_c' and β_F). The effects of temperature on τ_F, C_{jE} and C_{jC} are inherently included in the EM_3 model without the need for extra parameters.

2.5 The GP Model

2.5.1 Introduction

The Gummel-Poon (GP) model is the fourth and final nonlinear model of the BJT described in this book. It is based on the model formulated by Gummel and Poon in 1970.[7] Some modifications have been made, mainly with respect to terminology, the region of the transistor to which it is applied and the required model parameters.[44,3] The GP model, which is available in the computer program SPICE,[3] is described in some detail here. The GP model is almost entirely concerned with improvements in the dc characterization of the EM_3 model. These improvements, while appearing at first to be rather radical, are normally minor in effect. Thus the GP model is basically equivalent to the EM_3 model.[6]

The EM_3 model made improvements in three areas: dc performance (basewidth modulation and β versus I), ac performance (the C_{jC} - r_b' split and τ_F versus I_C) and variation with ambient temperature. These effects were treated separately and the model was altered piece by piece. The GP model in this chapter, covers only three effects (basewidth modulation, high-injection effects and τ_F versus I_C), but they are all treated together. This unified treatment provides a slightly more accurate and complete model than is provided by the EM_3 model. The GP model, however, is also more mathematical, less intuitive and less convenient (in the input parameters required). A comparison of the GP and EM_3 models is included at the end of this chapter. A tabular summary of the effects covered by the GP and EM_3 models is given in Table 2.1.

The reader may well ask why one should know about the GP model if the EM_3 model is simpler and almost equivalent. The answer lies in the understanding of the operation of the BJT that is obtained. A better appreciation of the limitations and assumptions involved in the models is obtained from the GP model. This will be illustrated further in this chapter.

Effect	EM_3	GP
β versus I - low currents	Yes	Yes
- high currents	Yes	Yes, in a unified manner
Basewidth Modulation	Yes	
τ_F versus I_C	Yes	
$r_b' - C_{jC}$ split	Yes	No
Temperature Variation	Yes	No

TABLE 2.1

To make this chapter as "appetizing" as possible, the mathematical derivations are kept to a minimum. Although the expressions may at times look formidable, their manipulations are relatively simple and emphasis is placed throughout on an appreciation of what the equations mean and the assumptions inherent in their use.

It should be pointed out that it is possible to make successful use of the GP model in computer programs without fully understanding the model. As long as the measurements of the GP parameters as described in this book are understood and used properly, the GP model can be used. This approach, however, is not desirable since an appreciation of the accuracy of the results and the limitations of the model is not obtained. As well, computing time may be wasted by using an unnecessarily complicated model.

The starting point for the GP model lies approximately halfway between the EM_2 and EM_3 models. It is assumed that the EM_2 model has been established and that the low-current drop in β is modeled by the inclusion of the two extra nonideal diodes. This assumed initial model is shown in Fig. 2.21. The improvements to the model of Fig. 2.21 afforded by the GP model will be incorporated by modifying existing equations (for I_{CT} and C_{DE} only) rather than by adding extra elements. Note that the expression for I_B (Eq. (2.56)) is unaffected by these modifications.

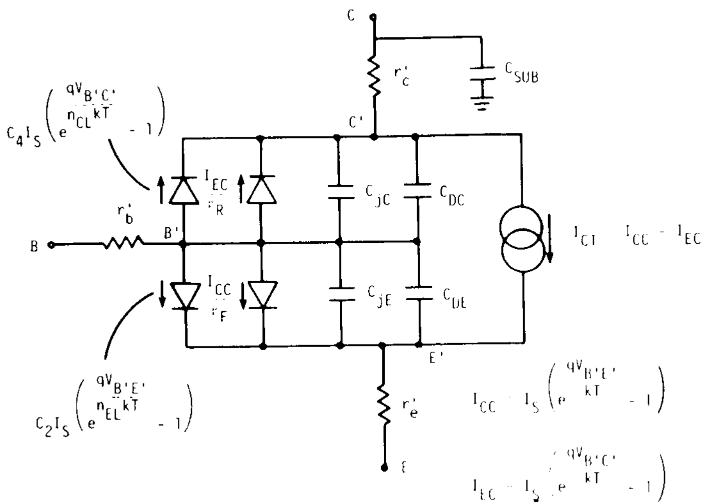


Fig. 2.21. The assumed "starting point" model for the GP derivation (shown for an npn transistor).

The effects included in the following derivations are:

1 - a complete description of basewidth modulation

2 - the effects of high-level injection

3 - base-widening effects which result in the variation of τ_F

Involved in the derivation is a new definition of I_S in terms of the internal physics of the transistor. Previously it had been defined (in the EM_1 model of Section 2.2) in terms of a terminal measurement which was a consequence of reciprocity (a property that is still assumed valid for the GP model).

It should be emphasized that the GP derivation, like all the previous derivations in this book, assumes a "one-dimensional"

transistor (as illustrated in Fig. 2.22). The extension of the GP model to the three-dimensional (real life) case is looked at briefly in the next chapter which is devoted to the limitations of the models described in this book.

The notation and model parameters described in this chapter are basically those employed by the program SPICE.[44,3] The extra model parameters required for the GP model, in addition to those for the EM_2 model are:

I_{SS}	(which replaces the I_S parameter)
C_2, C_4, n_{EL}, n_{CL}	(for low-current β)
V_A	(Early voltage)
V_B	(inverse Early voltage)
I_K	(knee current for ln I_C versus $V_{B'E'}$)
I_{KR}	(inverse knee current for ln I_E versus $V_{B'C'}$)

B and its model parameters (for τ_F versus I_C)

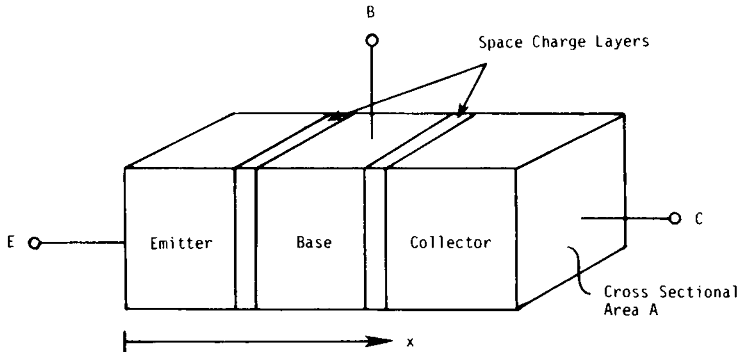


Fig. 2.22. Simplified illustration of the one-dimensional device assumed for the GP model.

The four parameters (C_2, C_4, n_{EL} and n_{CL}) were described in the previous chapter on the EM_3 model: they are included in this list for completeness only and will not be described further. The Early voltage, V_A, also described in the previous chapter, will be redefined for the GP model. I_{SS}, V_B, I_K, I_{KR} and B are the new model parameters;* B, the last one, itself requires several parameters to describe it.

The following derivation of the GP model proceeds in basically three major steps. First, (in Section 2.5.2), from an examination of the current density equations in the transistor a new expression for I_S is obtained. This new expression inherently incorporates the effects of basewidth modulation, high-level injection and τ_F vs. I_C. The second stage (covered in Section 2.5.3) consists of artificially separating out these effects so that they can be modeled and understood. The physical significance of this artificial separation is explained and then (in Section 2.5.4) each artificial component is modeled in terms of measurable parameters. Probably the most significant concept arising from the GP derivation is the new importance given to the majority carrier (hole for an npn) concentration profile in the base. The emergence of the majority carrier as important arises basically from the use of a simple mathematical "trick" in the derivation of I_S.

Finally, to place the following analysis in its proper perspective, it is mainly concerned with improvements to the modeling of second-order effects in the BJT. It must be remembered that the seemingly significant differences between the GP and EM_3 models are normally minor in their effect on the transistor's overall characteristics. The emphasis on the differences between the EM_3 and GP models are stressed here mainly as an aid to the understanding of the models and not as an indication of gross errors. If points are belabored (such as the accuracy of the basewidth modulation models)

*The parameter I_K was actually explained briefly in the EM_3 model but it will be treated in a fuller manner in this chapter.

they are done so not because they are important on a first-order or even, perhaps, a second-order basis but to justify the approach used.

2.5.2 The Physical Definition of I_S [7,44,3]

The derivation starts with the one-dimensional, dc equations for the electron current density, J_n, Eq. (2.76), and the hole current density, J_p, Eq. (2.77), in an npn transistor.[45]

$$J_n = q\mu_n n(x)\,\mathcal{E}(x) + qD_n \frac{dn(x)}{dx} \tag{2.76}$$

$$J_p = q\mu_p p(x)\,\mathcal{E}(x) - qD_p \frac{dn(x)}{dx} \tag{2.77}$$

where $\mathcal{E}(x)$ is the electric field, $n(x)$ is the free electron concentration and $p(x)$ is the hole concentration. No restriction is placed here, or later, on the variation of the carrier concentrations, so that the following analysis applies for any doping profile. Equations (2.76) and (2.77) apply for both high- and low-level injection.

At this point it is assumed that the hole current is zero. This is not exactly the case, but is normally a reasonable approximation. The approximation is justified by showing that there is no place where a large hole current could go.

(a) The base-emitter junction is normally designed for high emitter-injection efficiency (high emitter doping with respect to the base doping).[16] This means that the current injected into the emitter from the base is small even when the emitter-base junction is forward-biased.

(b) In the normal, active region the collector-base junction is reverse-biased and therefore no significant current flows across it from the base to the collector. In the inverse and saturated regions of operation, however, the

collector-base junction is forward-biased. It is assumed that for most cases of interest, the hole current that flows from the base to the collector is still small. Obviously, then, the following analysis will only be reasonably valid for the normal, active region and when the device is "weakly" saturated. For cases of strong saturation and inverse operation, it may not hold.*

If the hole current is assumed to be zero, Eq. (2.77) becomes:

$$J_p = q\mu_p \, p(x) \, \mathcal{E}(x) - qD_p \frac{dp(x)}{dx}$$

$$= 0 \tag{2.78}$$

which can be rearranged to solve for the field, $\mathcal{E}$.

$$\mathcal{E}(x) = \frac{D_p}{\mu_p} \cdot \frac{1}{p(x)} \cdot \frac{dp(x)}{dx} \tag{2.79}$$

Substitution of Eq. (2.79) into Eq. (2.76) gives

$$J_n = q\mu_n \left[\frac{D_p}{\mu_p} \frac{1}{p(x)} \frac{dp(x)}{dx} \right] n(x) + qD_n \frac{dn(x)}{dx} \tag{2.80}$$

or

$$p(x) \, J_n = qD_n \, n(x) \frac{dp(x)}{dx} + qD_n \, p(x) \frac{dn(x)}{dx} \tag{2.81}$$

where the Einstein relationship

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} \left(= \frac{kT}{q}\right)$$

has been used. The multiplication of Eq. (2.80) by p(x) is the mathematical "trick," which, with the following use of the differential product rule will result in the importance of the majority carrier concentration, p(x), in the base. The use of the product rule on Eq. (2.81) gives

*Actually, the assumption of $J_p = 0$ is better than was explained. If the J_p term is retained in the following derivation, it will be multiplied by $n(x)/p(x)$ which reduces its effect even further (since, except at high level injection, $n(x) << p(x)$).

$$p(x)\ J_n = qD_n \frac{d}{dx}\left[n(x)\ p(x)\right] \qquad (2.82)$$

Both sides of this equation are now integrated from x_E' to x_C' where x_E' is the position of the emitter side of the emitter-base space-charge layer and x_C' is the position of the collector side of the collector-base space-charge layer, as illustrated in Fig. 2.23. Figure 2.23 also defines x_E and x_C, the positions of the base sides of these space-charge layers.

Since current density J_n is constant for dc and independent of x (assuming negligible recombination in the base region*), it is taken out of the integral.

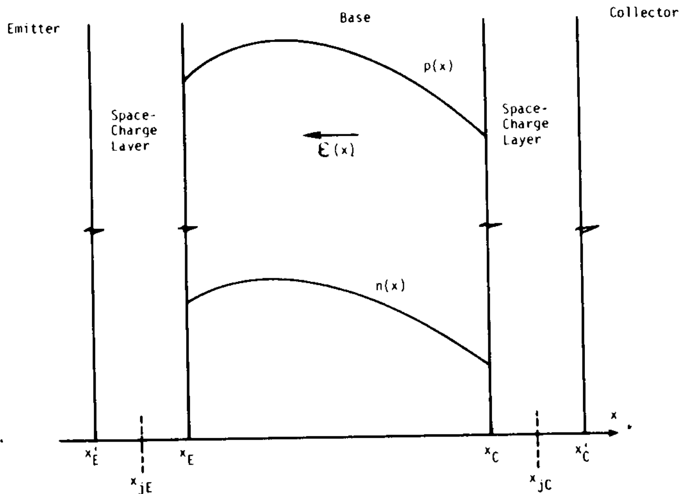


Fig. 2.23. Base profile for an npn transistor defining the integral limits x_E', x_C', x_E and x_C.

*Recombination in the space-charge layers is independently modeled by the two non-ideal diodes in Fig. 2.21, so is assumed to be zero here.

$$\therefore J_n \int_{x_E'}^{x_C'} p(x)\,dx = qD_n \int_{x_E'}^{x_C'} \frac{d}{dx}\left[n(x)p(x)\right]\,dx$$

$$= qD_n \left[n(x_C')p(x_C') - n(x_E')p(x_E')\right] \qquad (2.83)$$

$$\therefore J_n = \frac{qD_n \left[n(x_C')p(x_C') - n(x_E')p(x_E')\right]}{\int_{x_E'}^{x_C'} p(x)\,dx} \qquad (2.84)$$

The integration limits in the above (and following) equations need further explanation and justification. In their paper, Gummel and Poon[7] perform the integration over the entire transistor. That is, x_E' represented a point in the neutral emitter region near the emitter contact and x_C' represented a point in the neutral collector region near the collector contact. Their analysis then assumes that the minority carrier quasi-Fermi levels in the neutral emitter and collector regions stay constant. This assumption is implicit in the next step of the derivation. The approach taken here, however, is slightly different. The integration is actually performed from the outside edges of the two space-charge layers, -- that is, from the emitter side of the emitter-base space-charge layer, x_E', to the collector side of the collector-base space-charge layer, x_C'. As a result, no assumption is made about the minority carrier quasi-Fermi levels in the neutral emitter and collector regions in the next step. (In fact, the original derivation of the new expression for I_S, as presented by Gummel,[46] performed the integration from the outside of the space-charge layers and then assumed the result to be valid for the whole transistor). A comparison of this approach with that of Gummel and Poon is given later in this chapter.

A further simplification can be made to Eq. (2.84) by applying Boltzmann statistics to the pn products:

$$n(x_C')p(x_C') = n_i^2 \, e^{\frac{qV_{B'C'}}{kT}}$$

$$n(x_E')p(x_E') = n_i^2 \, e^{\frac{qV_{B'E'}}{kT}} \qquad (2.85)$$

Therefore

$$J_n = \frac{qD_n n_i^2 \left(e^{\frac{qV_{B'C'}}{kT}} - e^{\frac{qV_{B'E'}}{kT}} \right)}{\displaystyle\int_{x_E'}^{x_C'} p(x)\, dx} \qquad (2.86)$$

At this point, the depletion approximation is made. This assumes that there are no (or negligible) mobile carriers in a space-charge layer. That is, the field that is experienced by the carriers and the thickness of the space-charge layer are such that the carriers are transported "instantaneously" across it. This approximation, as has been pointed out in the EM_2 model (Section 2.3.2), is not valid for junctions under forward bias. The analysis that follows therefore takes the form of first assuming the depletion approximation to be true and then later fixes up the solution to take the mobile charges in the space-charge layers into account. In effect, the depletion approximation will only be applied for the thermal equilibrium condition. This is illustrated in Appendix 3, which gives an alternative, mathematical and more rigorous derivation of one aspect of the model. The treatment given in this chapter, however, while equivalent to that given in

Appendix 3, concentrates more on an understanding of the concepts involved. The depletion approximation is therefore made so that the introduction to the concepts is kept as simple as possible.

The application of the depletion approximation results in the limits of the integral in Eq. (2.86) being replaced by x_C and x_E and the integration being performed in the neutral base region only. The change in integration limits results, of course, from the assumption that p(x) is approximately zero inside the space-charge layers. Equation (2.86) can therefore be rewritten as:

$$I_n \simeq \frac{-qD_n A n_i^2 \left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)\right]}{\int_{x_E}^{x_C} p(x)\, dx} \qquad (2.87)$$

where A is the one-dimensional cross-sectional area (which converted current density to current) and the "-1" terms on the right hand side have been introduced for a later comparison.

I_n represents the total dc minority current in the positive x direction that results from minority carriers injected into the base at the emitter and/or the collector. It is represented in the model of Fig. 2.21 by the current generator I_{CT}. (The other components of the collector current in Fig. 2.21 are components of base current resulting from the injection of holes (for an npn) from the base towards the collector.) The equation for I_{CT}, previously obtained in the EM_1 model, is:

$$I_{CT(EM_1\ \text{model})} = I_{CC} - I_{EC}$$

$$= I_S\left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)\right] \qquad (2.88)$$

The sign of I_{CT} is opposite to that of I_n (Eq. (2.87)) since I_n has the opposite direction (out of the collector terminal) to that assumed for I_C (into the collector region from the terminal).

A direct comparison of Eqs. (2.87) and (2.88) yields a physical definition of I_S. However, care must be exercised in the comparison. In previous work, I_S has been considered as a fundamental constant of the device. Yet the integral in Eq. (2.87) is not constant in that under high-level injection p(x), the majority carrier concentration, is a function of the applied bias. To reconcile this difference, a new symbol, I_{SS}, is used in the GP model and is defined from Eq. (2.87) under low-level injection conditions only. At low-level injection, the combination of Eqs. (2.87) and (2.88) becomes:

$$I_{CT(\text{low level})} = \frac{qD_n n_i^2 A}{\int_{x_E}^{x_C} N_A(x)\,dx}\left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)\right] \tag{2.89}$$

where p(x) has been replaced by $N_A(x)$, since, at low current levels

$$p(x)_{\text{low level}} \simeq N_A(x) \text{ in the neutral base region, where } x_E \le x \le x_C \tag{2.90}$$

Before the definition of I_{SS} is made, though, more attention must be paid to the limits of the integration, x_E and x_C. Because of the variation of the space-charge layer widths with applied voltage, x_E and x_C are functions of the appropriate bias voltage (and, in fact, will be seen later to incorporate the effects of basewidth modulation). The fundamental constant, I_{SS}, is therefore defined at <u>zero</u> V_{BE} and V_{BC} as:

$$I_{SS} \triangleq \frac{qD_n n_i^2 A}{\int_{x_{E0}}^{x_{C0}} N_A(x)\, dx} \tag{2.91}$$

where x_{E0} and x_{C0} are the values of x_E and x_C when the applied junction voltages are zero. The fundamental nature of I_{SS} is seen immediately from Eq. (2.91) since it is uniquely determined once the base-doping profile is fixed. (As well, at zero bias voltages, Eq. (2.90) becomes almost exact.)

In the derivation of Eq. (2.91), the diffusion constant, D_n, has been assumed to be constant and independent of x. In practice, this assumption is not valid. The diffusion constant should be included in the denominator integral. Instead, D_n is interpreted in Eq. (2.91) as an effective diffusion constant in the base.

The Q_B concept [7,44]

The general expression for I_S (which is now a function of bias voltages) can be obtained in terms of the zero-bias constant I_{SS}. However, before this is done, it will be worthwhile to make a few definitions and to introduce some new concepts.

When multiplied by q and A, the integral in Eq. (2.87) represents the total majority charge in the neutral base region and is given the symbol Q_B:

$$Q_B \triangleq \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qAp(x)\, dx \tag{2.92}$$

The dependence of the integration limits on the junction voltages has been emphasized here.

The zero-bias majority base charge, Q_{BO}, is defined by:

$$Q_{BO} \triangleq \int_{x_{E0}}^{x_{C0}} qAN_A(x)\, dx \tag{2.93}$$

Finally, the normalized majority base charge, q_b, is defined as:

$$q_b \triangleq \frac{Q_B}{Q_{BO}} \tag{2.94}$$

In the following analysis, all charge normalizations are with respect to Q_{BO} and are represented by q with an appropriate subscript.

These above definitions (Eqs. (2.92) through (2.94)) can be used to find the new definition of I_S. If Eq. (2.87) is multiplied and divided by Q_{BO} (Eq. (2.93)) and I_n is replaced by $-I_{CT}$, then

$$I_{CT} = \frac{qD_n n_i^2 A}{\int_{x_E}^{x_C} p(x)dx} \cdot \left(\frac{qA \int_{x_{EO}}^{x_{CO}} N_A(x)dx}{qA \int_{x_{EO}}^{x_{CO}} N_A(x)dx} \right) \left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1 \right) \right]$$

The combination of the first term in the numerator with the second term in the denominator and using Eqs. (2.91), (2.92), (2.93) and (2.94) gives

$$I_{CT} = \frac{I_{SS}\, Q_{BO}}{Q_B} \left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1 \right) \right]$$

$$\boxed{I_{CT} = \frac{I_{SS}}{q_b} \left[\left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) - \left(e^{\frac{qV_{B'C'}}{kT}} - 1 \right) \right]} \tag{2.95}$$

Equation (2.95) is the new equation introduced by the GP model and the new concept introduced is the fundamental importance of the (normalized) majority charge in the base, q_b. The old saturation current I_S (which was assumed constant in the EM_1 model) has been replaced by the new term $\frac{I_{SS}}{q_b}$ where I_{SS} is the fundamental constant (defined at zero bias condition) and q_b is a variable that still needs to be determined. The rest of the GP derivation involves the determination of q_b as a function of the bias conditions in terms of measurable parameters.

2.5.3 The Components of Q_B (7,44,3)

Because of the general nature of the analysis, the function p(x) is not known and the integration cannot be performed here. Instead, Q_B is split into five components and each component is modeled separately. This separation of Q_B into its components is done in such a way that the performance of the integration is avoided and the dependence of the integration limits on junction voltage is handled. As well, a careful examination of the physical significance of each of these components improves the understanding of the device operation.

Before proceeding with the split of Q_B into its five components, the following points should be noted.

- Saturation. The components of Q_B are identified by considering the transistor to be in the saturation region since it will be seen that this is the only region where all five components are positive and non-zero. This region is therefore chosen for convenience in presentation only. The effect of operating the transistor in other regions will be obvious and will be explained.
- Superposition. In the saturation region, both junctions are forward-biased and minority carriers are injected into the base from both the collector and the emitter. From the preservation of charge neutrality in the base, the total concentration of the majority carriers increases by the same amount as the total increase in the minority carrier concentration. Superposition is assumed to hold, whereby it is assumed that the total excess majority carrier density is the sum of the excess majority carrier density due to each junction acting separately. It therefore follows that the total excess majority carrier concentration in the base is given by:

$$p'(x) = \left[p_F(x) - N_A(x)\right] + \left[p_R(x) - N_A(x)\right] \tag{2.96}$$

where $p_F(x)$ is the majority carrier concentration in the base if the collector-base junction has zero volts across it (i.e., $V_{B'C'} = 0$); and the emitter-base junction is forward-biased to the value under consideration (i.e., $V_{B'E'}$). Similarly, $p_R(x)$ is the majority carrier concentration in the base if the emitter-base junction has zero volts across it (i.e., $V_{B'E'} = 0$); and the collector-base junction is forward-biased to the value under consideration (i.e., $V_{B'C'}$). This superposition principle will be used in the following analysis of Q_B.

- Format. The split of Q_B into its components will be seen to be a relatively abstract, mathematical split. As a result, the format used is to first present this split mathematically, defining in this way each of the five components. Then, the physical significance of each of the components is given. This is then followed in the next section by a detailed evaluation of each component in terms of measurable parameters. The effect of the depletion approximation is considered and "removed" before the final solution for q_b is obtained. Finally, the results obtained here are compared with those obtained by Gummel and Poon and are shown to be virtually equivalent.

a) Mathematical derivation of components of Q_B

The derivation starts with the definition of Q_B, Eq. (2.92). The hole concentration, p(x), is separated into its equilibrium and excess components:

$$Q_B \triangleq \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qAp(x)\, dx$$

$$= \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qA\, N_A(x)\, dx + \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qAp'(x)\, dx \qquad (2.97)$$

The equilibrium component is split into three further components by the selection of appropriate limits:

$$Q_B = \int_{x_E(V_{B'E'})}^{x_{EO}} qA\ N_A(x)dx + \int_{x_{EO}}^{x_{CO}} qA\ N_A(x)dx + \int_{x_{CO}}^{x_C(V_{B'C'})} qA\ N_A(x)dx + \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qAp'(x)dx$$

$$\triangleq \quad Q_E \quad + \quad Q_{BO} \quad + \quad Q_C \quad + \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qAp'(x)dx \tag{2.98}$$

The component Q_{BO} has been previously defined (Eq. (2.93)). The components Q_E and Q_C are defined here by the first and third integrals respectively.

The last integral in Eq. (2.98) is split into two components by the application of superposition (Eq. (2.96))

$$\therefore Q_B = Q_E + Q_{BO} + Q_C + \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qA[p_F(x) - N_A(x)]dx + \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qA[p_R(x) - N_A(x)]dx$$

$$\triangleq Q_E + Q_{BO} + Q_C + Q_F + Q_R \tag{2.99}$$

where Q_F and Q_R are now defined by the two integrals respectively. Equation (2.99) defines the five components of Q_B,* which are illustrated in Fig. 2.24 and are now explained, physically.

---

*A similar derivation of Eq. (2.99) is given in Appendix 3 in which the depletion approximation is not assumed for the non-equilibrium component of p(x). The results obtained in Appendix 3 are consistent with those obtained later in this chapter.

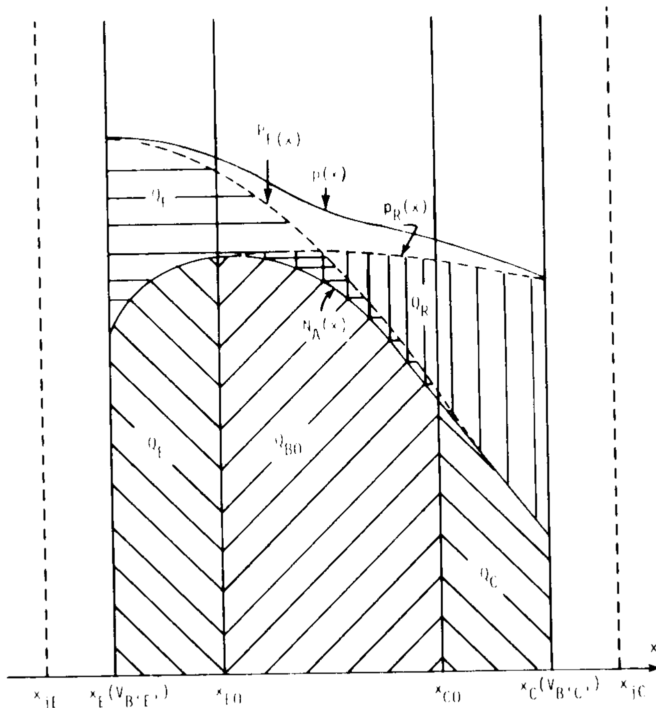


Fig. 2.24. The five components of Q_B for an npn transistor in the saturated region. (This bias condition corresponds to all positive, non-zero components.) $p_F(x)$ and $p_R(x)$ are the two concentrations assigned to satisfy the superposition principle defined by Eq. (2.96). The depletion approximation is assumed to hold here for all components of Q_B. (Not drawn to scale.)

b) Physical significance of components of Q_B

The components of Q_B have been defined from mathematical considerations. As a result, with the exception of Q_{BO}, they are artificial components in that they are not completely physically sensible on their own.

i) Q_{BO}

The component Q_{BO} has been previously defined (Eq. (2.93)) as the value of Q_B when the transistor is zero-biased ($V_{BE} = V_{BC} = 0$). It retains this definition. Under zero bias, x_E becomes x_{EO}, x_C becomes x_{CO} and $p(x)$ becomes $N_A(x)$. Q_{BO} can be considered as the "starting point" for Q_B: starting with Q_{BO}, the other components add the extra charges upon the application of biasing.

ii) Q_E

The component Q_E is defined by

$$Q_E \triangleq \int_{x_E(V_{B'E'})}^{x_{EO}} qA\ N_A(x)\ dx \qquad (2.100)$$

It represents the increase in the "equilibrium" component of Q_B when the application of a base-emitter voltage ($V_{B'E'}$) changes the position of the emitter-base space-charge layer edge from x_{EO} (the zero-bias position) to $x_E(V_{B'E'})$. For example, for a forward-biased base-emitter junction, Q_E is the charge due to N_A which is uncovered as the emitter-base depletion region retreats from its invasion into the neutral base region. Q_E is a mathematical entity, not a physical one. This can be understood by realizing that the application of $V_{B'E'}$ also causes excess charges, $p'(x)$, to be introduced into the base region and these are not included in Q_E.

It is therefore not possible to produce a condition whereby Q_E is obtained without also introducing these excess charges. However, under low-level injection conditions the excess charges are, by definition, insignificant compared with $N_A(x)$. Therefore, Q_E can be thought of, but is not defined as, the increased majority charge when $V_{B'E'}$ is applied if low-level injection is obtained.

iii) Q_C

The component Q_C is defined similarly by*

$$Q_C \triangleq \int_{x_{C0}}^{x_C(V_{B'C'})} qA\, N_A(x)\, dx \qquad (2.101)$$

It represents the increase in the "equilibrium" component of Q_B when the application of a base-collector voltage ($V_{B'C'}$) changes the position of the collector-base space-charge layer edge from x_{C0} (the zero-bias position) to $x_C(V_{B'C'})$. For a forward-biased collector-base junction, Q_C is the charge due to N_A which is uncovered as the collector-base depletion region retreats from its invasion into the neutral base region. As with Q_E, Q_C is a mathematical entity and not a physical one since the application of $V_{B'C'}$ also introduces excess charges into the base which are not included in Q_C. Similarly, Q_C can be thought of, but not defined as, the increased majority charge when $V_{B'C'}$ is applied if low-level injection is obtained.

iv) Q_F

The component Q_F was defined as

$$Q_F \triangleq \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qA\left[p_F(x) - N_A(x)\right] dx \qquad (2.102)$$

Q_F represents the additional excess majority charge in the actual, biased-transistor base when a base-emitter voltage is

---

*The distance variable, x, increases from left to right. Therefore

$$x_E(V_{B'E'}) \le x_{E0} \le x_{C0} \le x_C(V_{B'C'})$$

This inequality explains why the integration limits in Eqs. (2.100) and (2.101) are not in the same order.

applied (with $V_{B'C'}$ kept at zero). This component, Q_F, is also a mathematical entity and not a physical one. The excess charge only exists, of course, between x_E and x_C for the biased transistor. Therefore, although the upper integration limit in Eq. (2.102) is not the value of x_C that is physically consistent with the assumption of $V_{B'C'} = 0$, it is mathematically consistent. Q_F can be given a physical interpretation, though, if a slight approximation is made. Since $V_{B'C'} = 0$ certainly corresponds to low-level injection, $p_F(x) \approx N_A(x)$ near x_C and therefore the upper limit in Eq. (2.102) could be replaced, with little error, by x_{CO}. It then follows that Q_F can be thought of, though not defined as, the total excess majority charge in the base if the emitter-base is biased to the value under consideration and $V_{B'C'}$ is kept at zero.

v) Q_R

The component Q_R was defined as

$$Q_R \triangleq \int_{x_E(V_{B'E'})}^{x_C(V_{B'C'})} qA\left[p_R(x) - N_A(x)\right] dx \qquad (2.103)$$

Q_R therefore represents the additional excess majority charge in the actual, biased-transistor base when a base-collector voltage is applied (with $V_{B'E'}$ kept at zero). Like Q_F, Q_R is a mathematical (and not physical) entity since the lower integration limit, x_E, is not physically consistent with $V_{B'E'} = 0$. But, again like Q_F, since $V_{B'E'} = 0$ corresponds to low-level injection, Q_R can be thought of (with little associated error) as the total excess majority charge in the base if the collector-base is biased to the value under consideration and $V_{B'E'}$ is kept zero.

vi) Effect of operating regions

In the saturation region, all of the components of Q_B are positive since each models an increase in the majority charge in the base. In the normal, active region of operation, with the collector-base junction reverse biased, Q_C and Q_R are negative. Q_C is negative because $x_C(V_{B'C'})$ is closer to the emitter than x_{CO} and therefore majority charge is subtracted from the neutral base region. Q_R is negative because the reverse bias on the junction removes majority charge from the base. The negative nature of Q_C and Q_R is illustrated in Fig. 2.25, which shows the components of Q_B for an npn transistor in the normal, active region (V_{BE} positive and V_{BC} negative). For simplicity in drawing only, the simple, constant base-doping case is illustrated. Similarly, in the inverse region of operation, Q_E and Q_F will be the only negative components while in the off region, Q_{BO} will be the only positive component.

vii) Summary

Of the five components of Q_B, three (Q_{BO}, Q_E and Q_C) model the majority charge in the base due to the doping profile only and two (Q_F and Q_R) model the injected majority charges stored in the base. Obviously, then, at low injection levels, where the injected majority charges are negligible compared with the doping concentration, Q_{BO}, Q_E and Q_C are the dominant components of Q_B. Q_F and Q_R will only be important under high-level injection conditions.

2.5.4 Evaluation of q_b

When the components of Q_B are normalized (with respect to Q_{BO}), Eq. (2.99) becomes:

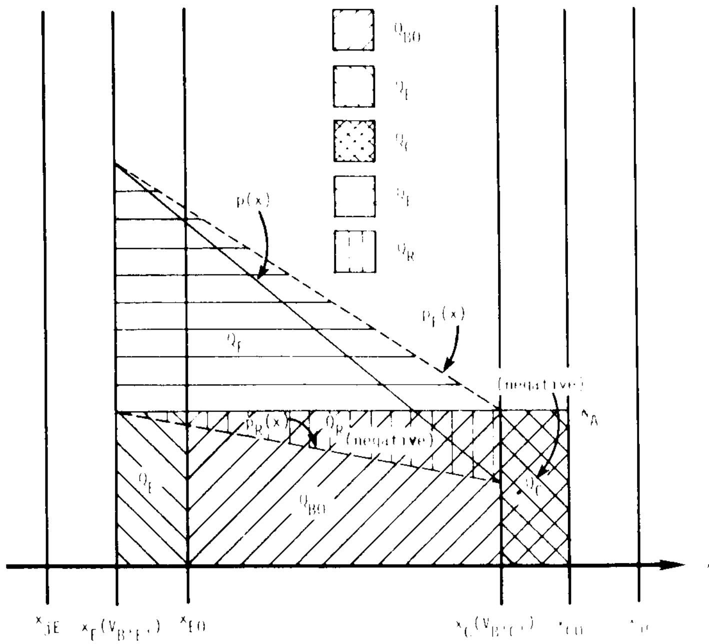


Fig. 2.25. The five components of Q_B for an npn transistor in the normal, active region ($V_{BE}+$, $V_{BC}-$), showing the negative nature of Q_C and Q_R. For simplicity, a constant doping is assumed in the base. The depletion approximation is assumed to hold here for <u>all</u> components of Q_B. (Not drawn to scale.)

$$q_b = 1 + q_e + q_c + q_f + q_r \tag{2.104}$$

where

$$q_e \triangleq \frac{Q_E}{Q_{BO}} \tag{2.105}$$

$$q_c \triangleq \frac{Q_C}{Q_{BO}} \tag{2.106}$$

$$q_f \triangleq \frac{Q_F}{Q_{BO}} \tag{2.107}$$

$$q_r \triangleq \frac{Q_R}{Q_{BO}} \tag{2.108}$$

Each of these components is treated separately below.

a) Component q_e

Q_E has been identified as the increased majority charge in the base when $V_{B'E'}$ is applied if low-level injection is obtained. It can therefore be related to the emitter-base junction capacitance, C_{jE}, by:

$$Q_E = \int_0^{V_{B'E'}} C_{jE}(V)\, dV \tag{2.109}$$

Therefore

$$q_e = \frac{1}{Q_{BO}} \int_0^{V_{B'E'}} C_{jE}(V)\, dV \tag{2.110}$$

If C_{jE} is assumed to be constant, indpendent of voltage (and equal to its average value, $\overline{C}_{jE}$)

$$q_e = \frac{\overline{C}_{jE}\, V_{B'E'}}{Q_{BO}} \triangleq \frac{V_{B'E'}}{V_B} \tag{2.111}$$

where V_B, the inverse Early voltage, is defined by

$$V_B \triangleq \frac{Q_{BO}}{\frac{1}{V_{B'E'}} \int_0^{V_{B'E'}} C_{jE}(V)\, dV} = \frac{Q_{BO}}{\bar{C}_{jE}} \quad (2.112)$$

In the following analysis, V_B will be assumed to be a constant, independent of $V_{B'E'}$ (which, in turn, assumes C_{jE} to be constant). The conditions under which this applies needs to be examined because of the apparently gross assumption of constant C_{jE}.* This is done in detail in Appendix 4, the conclusions of which are summarized as follows. The constant - V_B assumption has its greatest validity when the emitter-base junction is reverse biased since C_{jE} is then approximately constant. For a forward biased emitter-base junction, the integration in Eq. (2.112) should theoretically be performed since the constant - V_B assumption can result in a very large error in q_e.** However, this error is normally acceptable because q_e is usually not a dominant component of the total normalized base charge, q_b. That is, Q_E is normally much less than Q_{BO} and therefore q_e is normally much less than unity. This can be appreciated by remembering that q_e models the modulation of the base resulting from the variation of the emitter-base space-charge layer width. Since this effect is the inverse of the Early effect, it is here (facetiously) called the Late Effect.

Because the use of a constant V_B could give a large error in q_e when the emitter-base junction is forward biased, it is desirable to have a technique for determining the importance of q_e. One such technique is the measurement of the slope of the $\ln(I_C)$ versus $\frac{qV_{BE}}{kT}$ curve at low current levels. In all the previous models (which ignored the Late Effect) this slope was unity. The following

*This is described in detail here not because of its modeling importance but to justify the approach used.

**The integration would be relatively easy to perform since the functional dependence of C_{jE} on $V_{B'E'}$ is already modeled. This is described in more detail in Appendix 4.

analysis shows that with finite q_e this slope, though close to unity, can never be exactly unity.[7] Therefore, the deviation from a unity slope is a measure of the importance of q_e (and therefore a measure of how necessary an accurate modeling of it is needed).

Since the effect of q_e is desired, assume $q_c = q_r = q_f = 0$.* Therefore $q_b = 1 + q_e$ and the expression for I_C in the normal, active region is

$$I_C = \frac{I_{SS}}{(1 + q_e)} \left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) \tag{2.113}$$

Differentiation gives:

$$\frac{dI_C}{dV_{B'E'}} = \frac{I_{SS}}{(1+q_e)} e^{\frac{qV_{B'E'}}{kT}} \cdot \frac{q}{kT} + \frac{I_{SS}}{(1+q_e)^2} \left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) (-1) \cdot \frac{dQ_E}{dV_{B'E'}} \cdot \frac{1}{Q_{BO}}$$

$$\simeq \frac{qI_C}{kT} - \frac{I_C}{(1+q_e)} \frac{C_{jE}\,(V_{B'E'})}{Q_{BO}}$$

$$= \frac{qI_C}{kT} \left[1 - \frac{kT}{q} \frac{C_{jE}\,(V_{B'E'})}{(1+q_e)\, Q_{BO}} \right]$$

The slope of the $\ln(I_C)$ versus $\frac{qV_{B'E'}}{kT}$ characteristics (defined as $\frac{1}{n_E}$) is therefore given by

---

*This assumption is equivalent to assuming $V_{B'C'} = 0$ (for q_c and q_r) and low-level injection (for q_f), as will be seen later.

$$\frac{1}{n_E} \triangleq \frac{kT}{qI_C} \cdot \frac{dI_C}{dV_{B'E'}} \Bigg|_{V_{B'C'} = 0}$$

$$= 1 - \frac{kT}{q} \cdot \frac{C_{jE}(V_{B'E'})}{(1+q_e)\, Q_{BO}}$$

$$\therefore n_E = \frac{1}{\left[1 - \frac{kT}{q} \cdot \frac{C_{jE}(V_{B'E'})}{(1+q_e)\, Q_{BO}}\right]} \qquad (2.114)$$

When the constant - C_{jE} approximation is used, n_E (the emission coefficient) is modeled by:

$$n_E \simeq \frac{1}{\left[1 - \frac{\frac{kT}{q}}{V_B + V_{B'E'}}\right]} \qquad (2.115)$$

To summarize, q_e can, for most devices, be represented by $\frac{V_{B'E'}}{V_B}$ where V_B, the (assumed constant) inverse Early voltage, is defined as $\frac{Q_{BO}}{C_{jE}}$. The large absolute error in q_e that results from this simple representation when the emitter-base junction is forward biased is normally acceptable because q_e is usually a very small component of q_b. A more accurate representation of q_e (obtained by integrating C_{jE}) can be used for those cases where q_e is important (e.g., devices with low Q_{BO}). The main effect of q_e on the device characteristics in the normal, active region is an emission coefficient (n_E) greater than unity for the $\ln(I_C)$ versus V_{BE} curve at low and medium current ranges. This departure of the value of n_E from unity (which is normally so small as to be of the order of the experimental measurement error) is therefore a measure of the importance of both q_e and its accurate representation in this region of operation. That is, for slopes very close to unity, q_e can either be ignored or Eq. (2.111) used with a constant V_B. For slopes that depart significantly from unity, q_e may need to be modeled accurately.

b) Component q_c

Component q_c models the Early effect: the effect on q_b caused by the application of a collector-base bias at low current levels (or basewidth modulation by the variation of the collector-base space-charge layer width). In a derivation similar to that performed for q_e, q_c is given by:

$$q_c = \frac{1}{Q_{BO}} \int_0^{V_{B'C'}} C_{jC}(V)\, dV \qquad (2.116)$$

For the constant-C_{jC} approximation (with $C_{jC} = \bar{C}_{jC}$)

$$q_c = \frac{\bar{C}_{jC}\, V_{B'C'}}{Q_{BO}}$$

$$\triangleq \frac{V_{B'C'}}{V_A} \qquad (2.117)$$

where V_A, the Early voltage, has been redefined as

$$V_A \triangleq \frac{Q_{BO}}{\frac{1}{V_{B'C'}} \int_0^{V_{B'C'}} C_{jC}(V)\, dV} = \frac{Q_{BO}}{\bar{C}_{jC}} \qquad (2.118)$$

It is shown at the end of this chapter that this redefinition of V_A is consistent with the definition of V_A in the EM_3 model of the previous chapter.

As with q_e, the assumption of constant V_A (i.e., constant C_{jC}) needs to be examined further. This is done in Appendix 4 (for both q_e and q_c). The result of this examination for q_c is that when the collector-base junction is reverse biased (as in the normal, active region of operation) the assumption of constant C_{jC} is relatively

valid. However, when the collector-base junction is forward biased (inverse or saturation regions), a more accurate expression for q_c may be required, depending on the relative importance of q_c in the expression for q_b. This highlights one of the advantages of the GP model over the EM_3 model: if a more accurate model for base-width modulation is required, the solution is obvious in the GP model but not for the EM_3 model.

The effect of q_c on the device characteristics in the normal, active region of operation is a finite output conductance, g_o, in the common-emitter configuration. To observe this, assume that $q_e = q_f = q_r = 0$.* Therefore, the collector current is given (for $V_{BC} = 0$ and neglecting ohmic resistances) by:

$$I_C = \frac{I_{SS}}{(1 + q_c)} \cdot \left(e^{\frac{qV_{BE}}{kT}} - 1\right)$$

$$= \frac{I_{SS}}{\left(1 + \frac{V_{BC}}{V_A}\right)} \cdot \left(e^{\frac{qV_{BE}}{kT}} - 1\right) \qquad (2.119)$$

Since Eq. (2.119) is the same as that obtained in the EM_3 model for the incorporation of basewidth modulation, the result obtained in Appendix 2 (Eq. (A2.15)) is also valid. That is, for constant V_A:**

$$g_o = \left.\frac{dI_C}{dV_{CE}}\right|_{V_{BE} = \text{constant}} \simeq \frac{I_C(0)}{V_A} \qquad (2.120)$$

*This assumption is equivalent to assuming that $q_e \ll q_c$ and, as will be seen later, low-level injection (for q_f and q_r).

**A more realistic, non-constant slope, g_o, is obtained if the integral in Eq. (2.118) is used to model the variation of V_A with V_{BC}. For most cases, however, the constant-V_A approximation gives acceptable results.

c) Component q_f

The component q_f can be regarded as the (normalized) excess majority carrier concentration in the base when $V_{B'E'}$ is applied to the base-emitter junction. From charge neutrality, the total excess majority carrier equals the total excess minority carriers. Therefore, in determining q_f, use can be made of the well-established charge-control theory that describes the total excess minority carrier charge in the base. For an npn transistor:

$$Q_F = \int_{x_E}^{x_C} qA \left[p_F(x) - N_A(x) \right] dx$$

$$= \int_{x_E}^{x_C} qA \left[n_F(x) - \frac{n_i^2}{N_A(x)} \right] dx \tag{2.121}$$

The integral in Eq. (2.121) represents the total excess minority charge stored in the actual neutral base region with $V_{B'C'} = 0$ and $V_{B'E'}$ non-zero. In Section 2.3 (the EM_2 model), this charge has been given the symbol Q_3 (see Fig. 2.11) and is given (in Eq. (2.26)) by:*

$$Q_F = Q_3 = \tau_B \, I_{CC} \tag{2.122}$$

Therefore:

$$q_f = \frac{\tau_B I_{CC}}{Q_{BO}}$$

$$= \frac{\tau_B}{Q_{BO}} \cdot \frac{I_{SS}}{q_b} \left(e^{\frac{qV_{B'E'}}{kT}} - 1 \right) \tag{2.123}$$

---

*In the EM_2 model (and here), τ_B is assumed to be constant. Its variation with I_C is covered in a later section.

d) Component q_r

For q_r, a similar derivation to that performed for q_f yields:

$$Q_R = \int_{x_E}^{x_C} qA \left[p_R(x) - N_A(x)\right] dx$$

$$= \int_{x_E}^{x_C} qA \left[n_R(x) - \frac{n_i^2}{N_A(x)}\right] dx \qquad (2.124)$$

The integral in Eq. (2.124) represents the total excess minority charge stored in the neutral base region with $V_{B'E'} = 0$ and $V_{B'C'}$ non-zero. From Section 2.3 (Fig. 2.11 and Eq. (2.28))*

$$Q_R = Q_7 = \tau_{BR} I_{EC} \qquad (2.125)$$

Therefore

$$q_r = \frac{\tau_{BR}}{Q_{BO}} \cdot \frac{I_{SS}}{q_b} \left(e^{\frac{qV_{B'C'}}{kT}} - 1\right) \qquad (2.126)$$

Note that q_f and q_r both model high-level injection effects, in that at low current levels, $p(x) \approx N_A(x)$ and $q_f \approx q_r \approx 0$. Therefore, q_f and q_r only contribute significantly to q_b under high-level injection. Also note that Eqs. (2.123) and (2.126) contain q_b on the right-hand side.

e) Effect of depletion approximation

At this point in the analysis, the depletion approximation can be "removed" (or at least applied more accurately). The application

---

* τ_{BR} is assumed to be constant here.

of the depletion approximation allowed the above derivation to concentrate solely on the majority carriers in the neutral base region. As a result, the mobile carriers in the space-charge layers have been ignored. These carriers are now included in Q_F (fcr the emitter-base space charge layer) and Q_R (for the collector-bas space charge layer) by making the following definitions:

i) $\mathbf{Q_B}$. The Q_B term includes the space-charge layers

$$Q_B \triangleq \int_{x'_E (V_{B'E'})}^{x'_C (V_{B'C'})} qA\ p(x)\ dx \tag{2.127}$$

$$Q_{BO} \triangleq \int_{x'_{EO}}^{x'_{CO}} qA\ p_o(x)\ dx \simeq \int_{x_{EO}}^{x_{CO}} qA\ N_A(x)\ dx \tag{2.128}$$

where $p_o(x)$ is the equilibrium hole concentration in the neutral base and space-charge layer regions. The second form of Eq. (2.128), which assumes that the depletion approximation is valid for the equilibrium case, naturally agrees with the previous definition (Eq. (2.93)).

ii) $\mathbf{I_{SS}}$. The definition of I_{SS} is updated

$$I_{SS} \triangleq \frac{q\ D_n\ n_i^2\ A}{\int_{x'_{EO}}^{x'_{CO}} p_o(x)\ dx} \tag{2.129}$$

iii) $\mathbf{Q_E}$, $\mathbf{Q_C}$. Components Q_E and Q_C remain virtually unchanged

$$Q_E \triangleq \int_{x'_E}^{x_{EO}} qA\ p_o(x)\ dx \simeq \int_{x_E}^{x_{EO}} qA\ N_A(x)\ dx \tag{2.130}$$

$$Q_C \triangleq \int_{x_{CO}}^{x'_C} qA\ p_o(x)\ dx \approx \int_{x_{CO}}^{x_C} qA\ N_A(x)\ dx \qquad (2.131)$$

The second forms of these equations also assume that the depletion approximation is valid for the equilibrium case. Since these equations are unchanged, their solutions obtained previously are also unchanged.

iv) Q_F, Q_R. Components Q_F and Q_R are modified to include the mobile charges in the space-charge layers

$$Q_F \triangleq \int_{x'_E}^{x'_C} qA\left[p_F(x) - p_o(x)\right]\ dx \qquad (2.132)$$

$$= \int_{x'_E}^{x_E} qA\left[p_F(x) - p_o(x)\right]\ dx + \int_{x_E}^{x_C} qA\left[p_F(x) - p_o(x)\right]\ dx +$$

$$+ \int_{x_C}^{x'_C} qA\left[p_F(x) - p_o(x)\right]\ dx$$

$$\approx \int_{x'_E}^{x_E} qA\left[p_F(x) - p_o(x)\right]\ dx + \int_{x_E}^{x_C} qA\left[n_F(x) - n_o(x)\right]\ dx \qquad (2.133)$$

$$= \qquad Q'_2 \qquad + \qquad Q_3 \qquad (2.134)$$

$$= \qquad \tau'_{EB_{SCL}}\ I_{CC} \qquad + \qquad \tau_B\ I_{CC} \qquad (2.135)$$

$$\triangleq \qquad \tau^{\star}_{B_{dc}}\ I_{CC} \qquad (2.136)$$

where Eq. (2.133) assumes (in the absence of the Kirk effect, which will be treated later) that the third integral is zero (i.e., no mobile charges in collector-base space-charge layer) and that charge neutrality allows the replacement of p' by n' in the neutral base region; Eq. (2.134) identifies the first integral as charge Q_2' which is not necessarily the same as Q_2 in Fig. 2.11;* Eq. (2.135) expresses this charge in terms of an (assumed constant) transit time; and $\tau^*_{B_{dc}}$ is a modified base transit time, in which the "dc" subscript is included for emphasis. Note that the mobile charge in the emitter-base space-charge layer is represented by Q_2' and is simply included in the model by adding an extra term to Q_F which results in a modified value of τ_B.

Similarly, for Q_R

$$Q_R \triangleq \int_{x_E'}^{x_C'} qA \left[p_R(x) - p_o(x)\right] dx \qquad (2.137)$$

$$= \int_{x_E'}^{x_E} qA \left[p_R(x) - p_o(x)\right] dx + \int_{x_E}^{x_C} qA \left[p_R(x) - p_o(x)\right] dx + \int_{x_C}^{x_C'} qA \left[p_R(x) - p_o(x)\right] dx$$

$$\simeq \int_{x_E}^{x_C} qA \left[n_R(x) - n_o(x)\right] dx + \int_{x_C}^{x_C'} qA \left[p_R(x) - p_o(x)\right] dx \qquad (2.138)$$

$$= Q_7 + Q_6' \qquad (2.139)$$

$$= \tau_{BR} I_{EC} + \tau'_{CB_{SCL}} I_{EC} \qquad (2.140)$$

$$\triangleq \tau^*_{BR_{dc}} I_{EC} \qquad (2.141)$$

---

*Since Q_2 represents electrons (for an npn).

where Eq. (2.138) assumes the first integral to be zero, charge neutrality allows the replacement of p' by n' in the neutral base, the third integral is identified as Q_6' (which is again not necessarily equal to Q_6 in Fig. 2.11)* and is expressed in terms of an (assumed constant) transit time and $\tau^*_{BR_{dc}}$ is a modified reverse base transit time. Again, the mobile charges in the collector-base space-charge layer are simply included by adding an extra term to Q_R which results in a modified value of τ_{BR}.

Two further points should be made. First, since the mobile charges in the space-charge layers are included in Q_F and Q_R they should not be included again in Q_E and Q_C. Therefore the capacitances used in modeling Q_E and Q_C are the simple ones that assumed the depletion approximation and not the Chawla-Gummel[22] ones. The second point to be made is that $\tau^*_{B_{dc}}$ becomes τ_B when the emitter base junction is zero- or reverse-biased and $\tau^*_{BR_{dc}}$ becomes τ_{BR} when the collector-base junction is zero- or reverse-biased since for these cases the depletion approximation is valid.

The alternate, mathematically-complete derivation of Eq. (2.99) given in Appendix 3 (in which the carriers in the space-charge layers are included at the beginning and the depletion approximation successively applied) uses definitions for Q_B, Q_{BO}, Q_E, Q_C, Q_F and Q_R that are the same as those just given above (Eqs. (2.127) through (2.137)).

f) Solution for q_b

The complete solution for q_b is given from a compilation of Eqs. (2.104), (2.111), (2.117), (2.123), (2.126), (2.136) and (2.141):

$$q_b = 1 + \frac{V_{B'E'}}{V_B} + \frac{V_{B'C'}}{V_A} + \frac{\tau^*_{B_{dc}}}{Q_{BO}} \cdot I_{SS} \frac{\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right)}{q_b} + \frac{\tau^*_{BR_{dc}}}{Q_{BO}} I_{SS} \frac{\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)}{q_b} \tag{2.142}$$

---

*Since Q_6 represents electrons (for an npn).

To simplify the algebra, the following definitions can be made

$$q_b \triangleq q_1 + \frac{q_2}{q_b} \tag{2.143}$$

where

$$q_1 \triangleq 1 + q_e + q_c = 1 + \frac{V_{B'E'}}{V_B} + \frac{V_{B'C'}}{V_A} \tag{2.144}$$

and

$$q_2 \triangleq \frac{\tau^*_{Bdc}\, I_{SS}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) + \tau^*_{BRdc}\, I_{SS}\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)}{Q_{BO}} \tag{2.145}$$

Note that q_1 models basewidth modulation effects and q_2 models the effects of high-level injection.

Equation (2.143) results in a quadratic expression for q_b:

$$q_b^2 - q_b q_1 - q_2 = 0 \tag{2.146}$$

which gives:

$$q_b = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2} \tag{2.147}$$

where the negative solution has been ignored because q_b is greater than zero.

Equation (2.147) not only gives a solution for I_C at high injection levels, but also provides a definition for the term "high level injection." From Eq. (2.147), if $q_2 \ll \frac{q_1^2}{4}$, then $q_b \approx q_1$. This means that $q_f \approx q_r \approx 0$ and, as previously noted, this corresponds to low-level injection. However, if:

$$q_2 \gg \frac{q_1^2}{4} \tag{2.148}$$

then,

$$q_b \approx \sqrt{q_2} \tag{2.149}$$

The inequality (2.148) defines "high-level injection" while Eq. (2.149) is used to obtain I_C at high-level injection.

g) High-level injection solution

A consideration of the high-level injection case yields two benefits: a verification of the formula used in the EM_3 model (Eq. (2.57)) and a simplification of the above equation for q_2 (Eq. (2.145)) that gives τ^*_{Bdc}/Q_{BO} (and τ^*_{BRdc}/Q_{BO}) in terms of measurable quantities (I_K and I_{KR}).

i) Verification of high-current solution. In the normal, active region at high injection levels,

$$q_b \approx \sqrt{q_2}$$

To simplify the situation, consider the case of $V_{B'C'} = 0$ (i.e., $q_r = 0$). The extension of the result to non-zero q_r will be made later by analogy. For zero q_r

$$q_b = \sqrt{q_2} \approx \sqrt{\frac{\tau^*_{B_{dc}} I_{SS}}{Q_{BO}} e^{\frac{qV_{B'E'}}{kT}}}$$

$$= \sqrt{\frac{\tau^*_{B_{dc}} I_{SS}}{Q_{BO}}} e^{\frac{qV_{B'E'}}{2kT}} \tag{2.150}$$

Therefore, solving for I_C (= I_{CC} in this situation)

$$I_C = \frac{I_{SS}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right)}{\sqrt{\frac{\tau^*_{B_{dc}} I_{SS}}{Q_{BO}}} e^{\frac{qV_{B'E'}}{2kT}}} \approx \sqrt{\frac{Q_{BO} I_{SS}}{\tau^*_{B_{dc}}}} e^{\frac{qV_{B'E'}}{2kT}} \tag{2.151}$$

Therefore

$$I_C \propto e^{\frac{qV_{B'E'}}{2kT}} \tag{2.152}$$

which is the same result obtained by Webster[37] and used in the EM_3 model (Eq. (2.57)).

ii) Simplification of q_2. The simplification of the coefficient $\sqrt{\frac{Q_{BO} I_{SS}}{\tau^*_{B_{dc}}}}$ arises from a consideration of the $\ln(I_C)$ versus V_{BE} characteristics at the two extremes: high- and low-level injection. Figure 2.26 shows the variation of $\ln(I_C)$ as a function of $\frac{qV_{B'E'}}{kT}$. The low-current asymptote is given approximately by (for $q_e \simeq q_c \simeq 0$)

$$I_C \simeq I_{SS} \left(e^{\frac{qV_{B'E'}}{kT}} \right) \tag{2.153}$$

The high-current asymptote is given by Eq. (2.151). The intersection of those two asymptotes defines the knee current (I_K) and the knee voltage (V_K). From Eq. (2.151):

$$I_K = \sqrt{\frac{Q_{BO} \; I_{SS}}{\tau^*_{B_{dc}}}} \left(e^{\frac{qV_K}{2kT}} \right) \tag{2.154}$$

While from Eq. (2.153):

$$I_K = I_{SS} \left(e^{\frac{qV_K}{kT}} \right) \tag{2.155}$$

The solution of Eqs. (2.154) and (2.155) yields:

$$I_K = \frac{Q_{BO}}{\tau^*_{B_{dc}}} \tag{2.156}$$

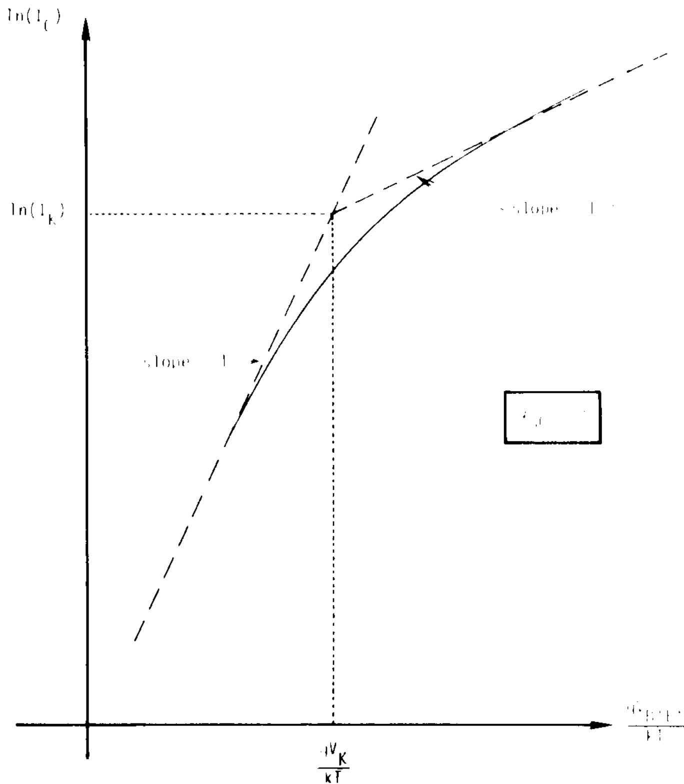


Fig. 2.26. The $\ln(I_C)$ versus $\frac{qV_{B'E'}}{kT}$ curve illustrating the high- and low-current asymptotes and the knee point (I_K, V_K).

A similar analysis for the inverse region defines I_{KR}, the knee current for $\ln(I_E)$ versus V_{BC} in the inverse region as:

$$I_{KR} = \frac{Q_{BO}}{\tau^{*}_{BR_{dc}}} \qquad (2.157)$$

h) Final solution

As a result of Eqs. (2.156) and (2.157), q_b can be finally written as:

$$q_b = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2} \qquad (2.158)$$

where
$$q_1 = 1 + \frac{V_{B'E'}}{V_B} + \frac{V_{B'C'}}{V_A} \qquad (2.159)$$

$$q_2 = \frac{I_{SS}}{I_K}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) + \frac{I_{SS}}{I_{KR}}\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right) \qquad (2.160)$$

The condition for high-level injection is:

$$q_2 >> \frac{{q_1}^2}{4} \qquad (2.161)$$

Notice that all parameters in the above expressions (V_B, V_A, I_{SS}, I_K and I_{KR}) are measurable from plots of $\ln(I_C)$ versus $V_{B'E'}$ in the normal, active region, $\ln(I_E)$ versus $V_{B'C'}$ in the inverse region, I_C versus V_{CE} and I_E versus V_{EC} characteristics. Once these parameters are known, q_1, q_2 and therefore, q_b can be easily determined by the above expressions.*

---

*Actually, the above derivation has ignored the variation of τ_{Bdc} and τ_{BRdc} due to basewidth modulation effects. If it is assumed that $\tau^{*}_{B} \approx \tau^{*}_{B}(0)\,\frac{W^2}{W(0)^2} = \tau^{*}_{B}(0)\left(1 + \frac{V_{B'E'}}{V_B} + \frac{V_{B'C'}}{V_A}\right)^2 = \tau^{*}_{B}(0)\,{q_1}^2$ and a similar expression holds for τ^{*}_{BR}, then Eq. 2.158 becomes $q_b = \left(\frac{q_1}{2}\right)\left[1 + \sqrt{1 + 4\hat{q}_2}\right]$ where $\hat{q}_2$ is defined as $\frac{I_{SS}}{I_K(0)}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) + \frac{I_{SS}}{I_{KR}(0)}\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right)$ and $I_K(0) = I_K\left(1 + \frac{V_K}{V_B}\right)$, $I_{KR}(0) = I_{KR}\left(1 + \frac{V_{KR}}{V_A}\right)$. However, for the normal situation of $q_1 \approx 1$, this change in q_b is insignificant.

i) Comparison with Gummel-Poon derivation

As indicated earlier, the above derivation of the GP model differs from that of Gummel and Poon(7) in the integration limits used. Whereas the above derivation integrated from the outside of the space-charge layers, Gummel and Poon integrated over virtually the entire transistor. The reason for this difference has been previously explained. Two basic differences result. They are: different definitions of I_{SS} and the knee currents (I_K and I_{KR}).

I_{SS} - In the definition of I_{SS} (Eq. (2.129)) the Gummel and Poon approach results in the denominator integration being performed over virtually the entire transistor. Since the base majority carriers are minority carriers in the neutral emitter and collector regions, this difference should have a negligible effect. This difference in definition is not important for the device characterization since I_{SS} is determined experimentally.

I_K, I_{KR} - In the equations for I_K and I_{KR} (Eqs. (2.156) and (2.157)), the terms τ^*_{Bdc} and τ^*_{BRdc} occur. If the integration is performed over the entire transistor, the definitions of τ^*_{Bdc} and τ^*_{BRdc} would include the emitter delay (τ_1) and the collector delay (τ_C), respectively (since Q_1 would be added to Q_F and Q_5 would be added to Q_R). Again, since I_K and I_{KR} are both experimentally determined, this difference in τ^*_{Bdc} and τ^*_{BRdc} is not important for the device characterization.

With the exception of the implementation of τ_F versus I_C, this completes the solution for the normalized charge q_b, the extra factor introduced by the GP model. Four more model parameters (I_{SS}, V_B, I_K and I_{KR}) have been introduced. The other model parameter, V_A, has been redefined. The variation of τ_F with I_C is treated next.

2.5.5 Base-Widening Effects

As pointed out in the EM_3 model, there are two effective base-widening effects (the one-dimensional effect described by Kirk[39] and the two-dimensional spreading effect described by van der Ziel and Agouridis[40]). And, as also pointed out in the EM_3 model, there is still some controversy as to which of the two base-widening effects dominate[41,42,43]

The effective base-widening at high currents is included by means of a multiplier, B, called the "base push-out factor."[7] B, which multiplies $\tau_{B_{dc}}$, has the property of being equal to unity at low currents and increases at high currents. That is,

$$\therefore \tau_{B_{dc}} = B\,\tau_{BL} \tag{2.162}$$

where τ_{BL} is the constant, low-current value of $\tau_{B_{dc}}$ and thus the dependence on I_{CC} is contained in B.

In the EM_3 model (which assumes that the van der Ziel and Agouridis effect is dominant), B is given by:*

$$B_{EM_3} = 1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2 \left[\frac{1}{3}\left(\frac{I_{CC}}{I_{CO}}\right)^2 - \left(\frac{I_{CC}}{I_{CO}}\right) + 1 - \frac{1}{3}\left(\frac{I_{CO}}{I_{CC}}\right)\right] \text{ for } I_{CC} \geq I_{CO}$$

$$= 1 \quad \text{for } I_{CC} \leq I_{CO} \tag{2.163}$$

In this equation, two model parameters are used: $\frac{L_E}{W}$ and I_{CO}. Alternatively, Gummel and Poon use the expression:[7]

$$B = \left\{1 + \frac{r_W}{\left[4\left(\frac{I_C}{I_K}\right)^2 + r_P\right]} \cdot \left[\frac{\sqrt{I_4^{\,2} + I_K^{\,2} r_P} - I_4}{I_K}\right]^{n_P}\right\}^2 \tag{2.164}$$

---

*Equation (2.163) strictly applies to the τ_B component of $\tau_{F_{ac}}$ only, but it is assumed in the EM_3 model that it can also be applied to $\tau_{F_{dc}}$.

where

$$I_4 = I_C + \left[\phi_C - V_{BC} - \frac{kT}{q} \ln\left(\frac{I_K}{I_{SS}}\right)\right] \frac{I_K}{V_{rp}} \tag{2.165}$$

The four model parameters are:

- r_w - the ratio of the width of the collector epitaxial region to the width of the metallurgical base ($x_{jC} - x_{jE}$ in Fig. 2.23)
- r_p - which determines the steepness of the variation of τ_B versus I_C at high current levels
- n_p - which determines the steepness of the drop of f_T versus I_C at high current levels
- V_{rp} - the resistive voltage drop across the collector when $I_C = I_K$.

The increase in the basewidth affects both the dc and ac characteristics.

a) dc characteristics

The increase in Q_B due to the effective increase in the basewidth is included in the Q_F term. Since I_K is inversely proportional to τ_B, q_2 is given by:

$$q_2 = \frac{B\ I_{SS}}{I_K}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) + \frac{I_{SS}}{I_{KR}}\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right) \tag{2.166}$$

The decrease in β_F at high currents due to base-widening effects[47] is therefore inherently included in Eq. (2.166). A similar push-out factor for inverse operation could also be defined.

b) ac characteristics

The analysis and modeling of the base-widening effect on the ac characteristics can become very complicated. A simplified analysis will be presented here, in which at first only the effect on the τ_B component of τ_F is considered. That is, it will be assumed that the other components of τ_F are either constant or negligible. This assumption may not be valid, especially for the emitter delay (τ_1) component.[25] Therefore, at the end of this analysis, the τ_1 component variation and then finally both τ_B and τ_1 together are also considered.

i) Effect of base-widening on τ_B. Since the parameter B above has been defined for $\tau_{B_{dc}}$, a similar multiplier (B_{ac}) can be obtained for $\tau_{B_{ac}}$

$$\tau_{B_{ac}} \triangleq \frac{d\, Q_3}{d\, I_{CC}}$$

$$= \frac{d\left(\tau_{B_{dc}}\, I_{CC}\right)}{d\, I_{CC}}$$

$$= \tau_{B_{dc}} + I_{CC}\left(\frac{d\, \tau_{B_{dc}}}{d\, I_{CC}}\right) \qquad (2.167)$$

If $\tau_{B_{dc}}$ is given by Eq. (2.162), Eq. (2.167) becomes

$$\tau_{B_{ac}} = B\, \tau_{BL} + \tau_{BL}\, I_{CC} \frac{dB}{d\, I_{CC}}$$

$$= \left[B + I_{CC} \frac{dB}{d\, I_{CC}}\right] \tau_{BL}$$

$$\triangleq B_{ac}\, \tau_{BL} \qquad (2.168)$$

In the EM_3 model, for example, B is given by Eq. (2.163) which gives:[41]

$$B_{ac_{(EM_3)}} = 1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2 \left(\frac{I_{CC}}{I_{CO}} - 1\right)^2 \text{ for } I_{CC} \geq I_{CO} \tag{2.169}$$

In the GP model, B is given by Eq. (2.164), which yields a very complicated expression for $B_{ac_{(GP)}}$.

The increased basewidth affects the nonlinear diffusion capacitance, C_{DE}. Therefore, neglecting other components of τ_F (i.e., assuming $Q_{DE} \simeq Q_3$ only)

$$C_{DE_{nonlinear}} \triangleq \frac{Q_{DE}}{V_{B'E'}}$$

$$\simeq \frac{Q_3}{V_{B'E'}} = \frac{B\tau_{BL}\ I_{CC}}{V_{B'E'}} \tag{2.170}$$

$$C_{DE_{small\ signal}} \triangleq \frac{dQ_{DE}}{dV_{B'E'}}$$

$$\simeq \frac{dQ_3}{dV_{B'E'}} = B_{ac}\ \tau_{BL}\ g_{mF} \tag{2.171}$$

ii) Emitter delay (τ_1) versus I_C. An inspection of Eqs. (2.170) and (2.171) would indicate that the entire rise in τ_F at high currents is caused by the effect of base-widening on τ_B. This is incorrect. A rise in τ_F at high currents will also occur from the effect on the τ_1 component of both high-level injection in the base

and base-widening. This is now illustrated, the rise in τ_1 due to high-level injection effects being examined first.

High-level injection in the base. Under the condition of low-level injection in the base, I_{CC} is given by (Eq. (2.153)):

$$I_{CC_{low}} \simeq I_{SS}\, e^{\frac{qV_{B'E'}}{kT}} \tag{2.172}$$

while for high-level injection in the base (Eqs. (2.151) and (2.156)):

$$I_{CC_{high}} = \sqrt{I_K\, I_{SS}}\; e^{\frac{qV_{B'E'}}{2kT}} \tag{2.173}$$

The charge stored in the emitter region, however, is assumed for both low- and high-level injection in the base to be given by

$$Q_1 = \text{const.}\; e^{\frac{qV_{B'E'}}{kT}} \tag{2.174}$$

since it is assumed that because of the normally high doping in the emitter, low-level injection conditions always apply in the emitter. Combining Eq. (2.174) with Eqs. (2.172) and (2.173) yields

$$Q_{1_{low}} = \frac{\text{const.}}{I_{SS}}\, I_{CC_{low}} \triangleq \tau_{1L}\, I_{CC_{low}} \tag{2.175}$$

$$Q_{1_{high}} = \frac{\text{const.}}{I_K I_{SS}}\, I^2_{CC_{high}} = \frac{\tau_{1L}}{I_K}\, I^2_{CC_{high}} \tag{2.176}$$

where τ_{1L} is the low-current, constant value of τ_1 and Eq. (2.175) was used in the second form of Eq. (2.176). The unusual form of Eq. (2.176) arises from the fact that although Q_1 follows the ideal exponential law, I_{CC} no longer does and Q_1 has been expressed in terms of this I_{CC}. The combination of Eqs. (2.175) and (2.176) give, in general,

$$Q_1 = \tau_{1L} I_{CC} (1 + \frac{I_{CC}}{I_K}) \tag{2.177}$$

From Eq. (2.177),

$$\tau_{1_{dc}} \triangleq \frac{Q_1}{I_{CC}} = \tau_{1L} (1 + \frac{I_{CC}}{I_K}) \tag{2.178}$$

$$\tau_{1_{ac}} \triangleq \frac{dQ_1}{dI_{CC}} = \tau_{1L} (1 + \frac{2I_{CC}}{I_K}) \tag{2.179}$$

These equations show that a rise in τ_1 occurs that is proportional to I_{CC} when high-level injection occurs in the base region. This rise in τ_1 is a consequence of the definition of τ_1 in terms of I_{CC}, which has a changing dependence on $V_{B'E'}$.

Base-widening. Base-widening does not affect Q_1 but it does affect I_{CC} (via I_K). Therefore Eq. (2.177) becomes

$$Q_1 = \tau_{1L} I_{CC} (1 + \frac{I_{CC} B}{I_K}) \tag{2.180}$$

and therefore:

$$\tau_{1_{dc}} \triangleq \frac{Q_1}{I_{CC}} = \tau_{1L} (1 + \frac{B I_{CC}}{I_K}) \tag{2.181}$$

$$\tau_{1_{ac}} \triangleq \frac{dQ_1}{dI_{CC}} = \tau_{1dc} + \tau_{1L} \cdot \frac{B_{ac} I_{CC}}{I_K} \tag{2.182}$$

iii) τ_F versus I_C. As indicated earlier, the effect of base-widening on τ_F can be very complicated. Considering the τ_B and τ_1 components only, the results obtained above, Eqs. (2.162), (2.168), (2.181) and (2.182), give

$$\tau_{F_{dc}} \simeq \tau_{B_{dc}} + \tau_{1_{dc}} = B\,\tau_{BL} + \left(1 + \frac{BI_{CC}}{I_K}\right)\tau_{1L} \qquad (2.183)$$

$$\tau_{F_{ac}} \simeq \tau_{B_{ac}} + \tau_{1_{ac}} = B_{ac}\,\tau_{BL} + \tau_{1_{dc}} + \frac{B_{ac}\,I_{CC}}{I_K}\,\tau_{1L} \qquad (2.184)$$

Since τ_B and τ_1 are not individually identified in the model, but are components of the model parameter, τ_F, the variation of τ_F with I_{CC} is not simply expressed. The model used in SPICE, for example, effectively assumes that $\tau_1 >> \tau_B$ and therefore[3]

$$\tau_{F_{dc(SPICE)}} = \tau_{FL}\left(1 + \frac{BI_{CC}}{I_K}\right) \qquad (B = 1,\ \text{fixed}) \qquad (2.185)$$

which is the same variation that is obtained for β_F^{-1} at high currents (as described in the next Section, Eq. (2.193)).

Note that the above analysis has neglected the following contributions to τ_F variations with I_C.

- the reduction of the low-level aiding field in drift transistors[26]
- the variation of $\tau_{EB_{SCL}}$
- the variation of $\tau_{CB_{SCL}}$

2.5.6 Comparison of the GP Model with the EM$_3$ Model

A comparison of the effects modeled by the EM$_3$ and GP models shows that the treatments are virtually equivalent with respect to

the effects that are included in both models. For the EM_3 model however, each effect is treated separately, while in the GP model all the effects are treated together in a cohesive, unified manner. As well, the GP model treats the inverse region in a fuller manner, covering the Late Effect with the inverse Early voltage, V_B. The GP model also incorporates the effect of base widening on the dc characteristics. For simplicity in presentation, the following comparison neglects ohmic resistances and therefore $V_{B'E'} = V_{BE}$ and $V_{B'C'} = V_{BC}$.

a) Basewidth modulation

Both the GP and EM_3 models include basewidth modulation due to V_{BC} in the same manner, but the GP model also includes basewidth modulation due to V_{BE} (the Late Effect).

In the EM_3 model, the effects of the collector-base space-charge layer width variation were included by means of the Early voltage, V_A. This voltage is used to modify both τ_F and I_S in the I_{CT} expression. (The variation in β_F follows from the variation in I_S.) The modified I_S is given by:

$$I_S(V_{BC})_{EM_3\ \text{model}} = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)} \simeq I_S(0)\left(1 - \frac{V_{BC}}{V_A}\right) \qquad (2.186)$$

The GP model provides a more complete treatment, since it also models the variation of the emitter-base space-charge layer width by means of the inverse Early voltage, V_B. The equivalent I_S in the GP model is:

$$I_S(V_{BE}, V_{BC})_{GP\ \text{model}} = \frac{I_{SS}}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)} \qquad (2.187)$$

Equations (2.186) and (2.187) are equal if the effect of the change in emitter-base space-charge layer width is ignored.

b) High-level injection

Both models have the same high-current asymptote for the $\ln(I_C)$ versus V_{BE} curve: a straight line with a slope of 1/2 (see Fig. 2.26). Only the model parameter describing the position of the line is different: θ for the EM_3 model, I_K for the GP model.

For the EM_3 model, at $V_{BC} = 0$:

$$I_{C(\text{high level})_{EM_3}} = \frac{I_S(0)}{\theta} \cdot e^{\frac{qV_{BE}}{2kT}} \quad (2.188)$$

while, for the GP model (neglecting basewidth modulation terms):

$$I_{C(\text{high level})_{GP}} = \sqrt{I_{SS}\ I_K} \cdot e^{\frac{qV_{BE}}{2kT}} \quad (2.189)$$

A comparison of Eqs. (2.188) and (2.189) yields

$$\theta = \frac{I_S(0)}{\sqrt{I_{SS}\ I_K}} = \sqrt{\frac{I_{SS}}{I_K}} \quad (2.190)$$

since $I_S(0) \simeq I_{SS}$. A similar comparison can be made for the inverse region.

c) β versus I

In both models, the variation of β_F with I_C is a consequence of the correct modeling of I_C and I_B. In the EM_3 model, the variation of β_F with I_C is given (for $V_{BC} = 0$) by:

$$\beta_F(0)_{EM_3} = \frac{1}{\dfrac{1}{\beta_{FM}(0)} + C_2\left(\dfrac{I_S(0)}{I_C}\right)^{(1-1/n_{EL})} + \dfrac{\theta^2 I_C}{\beta_{FM}(0)\, I_S(0)}} \qquad (2.191)$$

while for the GP model, Dowell[44] has shown that, for B = 1:

$$\beta_F(0)_{GP} = \frac{1}{\dfrac{1}{\beta_{FM}(0)} + C_2\, I_{SS}^{(1-1/n_{EL})}\left[\dfrac{I_C^{(2-n_{EL})}}{I_K} + I_C^{(1-n_{EL})}\right]^{1/n_{EL}} + \dfrac{I_C}{\beta_{FM}(0) I_K}}$$

(2.192)

At low current levels (obtained by putting $\theta = 0$ and $I_K = \infty$), both β_F expressions are identical. At high current levels, (obtained by setting $C_2 = 0$) both expressions show the same dependence on I_C:

$$\beta_F(0)^{-1} = \beta_{FM}(0)^{-1}\left(1 + \frac{I_C}{I_K}\right) \qquad (2.193)$$

The difference in the two expressions lies mainly in the transition region from medium to high injection levels. The GP formula is more accurate in that the derivation of EM_3 model Eq. (2.191) assumes that the high- and low-current regions are well separated and there is no interaction between them.

A similar comparison of expressions for β_R as a function of I_E can also be made.

A comparison of the improvements of the dc characterization provided by the EM_3 and GP models is given in tabular form in Table 2.2.

	EM_3 MODEL	GP MODEL
Basewidth Modulation		
Model parameters	V_A	V_A, V_B
I_S Formula (basewidth modulation only)	$I_S(V_{BC}) = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)}$	$I_S(V_{BE},V_{BC}) = \frac{I_{SS}}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)}$
β_F Versus I_C		
Model parameters	C_2, n_{EL}, β_{FM}, θ	C_2, n_{EL}, β_{FM}, I_K
Input parameters	β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, V_{CE}, BCEC	C_2, n_{EL}, β_{FM}, I_K
Low-current model ($V_{BC} = 0$)	$I_B = \frac{I_S(0)}{\beta_{FM}(0)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) + C_2 I_S(0)\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right)$ for $I_{SS} = I_S(0)$ i.e., $q_e = 0$	
High-current model ($V_{BC} = 0$)	$I_S = \frac{I_S(0)}{1 + \theta\, e^{\frac{qV_{BE}}{2kT}}}$	$I_S = \frac{I_{SS}}{1 + \sqrt{\frac{I_{SS}}{I_K}\left(e^{\frac{qV_{BE}}{kT}} - 1\right)}}$ for $q_e = 0$
Full I_S Definition	$I_S(V_{BE},V_{BC}) = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)\left(1 + \theta\, e^{\frac{qV_{BE}}{2kT}} + \theta_R\, e^{\frac{qV_{BC}}{2kT}}\right)}$	$I_S = \frac{I_{SS}}{q_b}$ $q_b = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2}$ $q_1 = 1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}$ $q_2 = \frac{BI_{SS}}{I_K}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) + \frac{I_{SS}}{I_{KR}}\left(e^{\frac{qV_{BC}}{kT}} - 1\right)$

Note: Ohmic resistances ignored $\therefore$ $V_{BE} = V_{B'E'}$, $V_{BC} = V_{B'C'}$.

TABLE 2.2

d) τ_F versus I_C

Both the GP and EM_3 models treat the variation of τ_{Fdc} with I_C by employing a multiplier, B. In both models an ac version of the multiplier (B_{ac}) is used to modify the value of the small-signal diffusion capacitor C_{DE}. However, in the GP model the dc characteristics are also modified by including the multiplier B in the expression for q_f (and therefore q_2).

In the implementation of the EM_3 and GP models in the Berkeley programs, the multiplier B is modeled as follows. In the EM_3 model (in SLIC and SINC), B models the van der Ziel and Agouridis effect by means of two model parameters. In the GP model in SPICE, B is not modeled (but an increase in τ_F with I_C has been built in which follows the variation of the τ_1 component and which has the same variation as β_F with I_C).[3]

2.5.7 Small-Signal (Linearized) GP Model

The small-signal, linearized GP model is described (with that of the EM_3 model) in Appendix 5. This model is also very similar to that given for the EM_2 model in Section 2.3.3, the only difference being the definition of r_π, r_μ, C_π and C_μ to include the effects incorporated in the GP model.

2.5.8. Summary

The improvements to the EM_2 model provided by the GP model are virtually all concerned with the dc aspects of the model. The GP dc model is basically equivalent in performance to the EM_3 with some (normally minor) differences. These differences can be summarized as follows:

a) dc model

The only difference between the EM_3 and GP models lies in the value of I_S. For the EM_3 model

$$I_S(V_{B'C'},V_{B'E'}) = \frac{I_S(0)}{\left(1 + \frac{V_{B'C'}}{V_A}\right)\left(1 + \theta\, e^{\frac{qV_{B'E'}}{2kT}}\right)} \qquad (2.194)$$

where the first bracketed term in the denominator models basewidth modulation by the variation of V_{BC} and the second bracketed term in the donominator models the effect of high-level injection.

For the GP model, a similar equation is used:

$$I_S(V_{B'C'},V_{B'E'})_{GP} = \frac{I_{SS}}{q_b} \qquad (2.195)$$

where

$$q_b = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2} \qquad (2.196)$$

$$q_1 = 1 + \frac{V_{B'E'}}{V_B} + \frac{V_{B'C'}}{V_A} \qquad (2.197)$$

$$q_2 = \frac{BI_{SS}}{I_K}\left(e^{\frac{qV_{B'E'}}{kT}} - 1\right) + \frac{I_{SS}}{I_{KR}}\left(e^{\frac{qV_{B'C'}}{kT}} - 1\right) \qquad (2.198)$$

The q_1 term models basewidth modulation by the variation of both V_{BC} and V_{BE}. The q_2 term models the effect of high-level injection. The B term in Eq. (2.198) models the effective increase in the basewidth at high currents.

It is shown in Section 2.5.6 that these two expressions for I_S are virtually equivalent with the GP model providing a slight

increase in accuracy. Another feature provided by the GP model is the different approach to the device that is used; namely, how the analysis of the majority carriers in the base can be used to generate the above equations. As well, the general and unified nature of the GP model gives a better indication than does the EM_3 model of what should be done if any of the underlying assumptions (such as large V_A or V_B, for example) are not valid.

b) Charge-storage model

Whereas the EM_3 model introduces both the $r_b' - C_{jC}$ split and an empirical fit to τ_F versus I_C, the GP model only introduces the latter (τ_F versus I_C).

c) Temperature variation model

The temperature variation aspect of the GP model is equivalent to that of the EM_1 model and in this respect is quite inferior to the EM_3 model.

Although the GP (and EM_3) models appear to be very complete and sophisticated, there are still some effects that have not even been considered. What these effects are and how they affect the validity of the models is described briefly in the next chapter.

2.6 Limitations of the Models

Probably the most significant limitations of all the models described here are a lack of modeling of three-dimensional effects, junction breakdown and an accurate saturation model.

a) Three-Dimensional Effects

All the models presented have assumed a one-dimensional device with a constant cross-sectional area which generally gives very good agreement with experimental results, indicating that the one-dimensional approach is often adequate. But neglect of three-dimensional effects ignores the crowding phenomenon (for dc and ac) and the consequent variation of the device parameters. The classical two-dimensional analysis of crowding and its effect on r_b' was performed by Hauser.[18] However, his analysis neglects the effects of the lateral base diffusion that results when crowding occurs (see Fig. 2.27) and localized base-widening effects.

The variation of r_b' is normally modeled by splitting it into two components;[5] the assumed-constant extrinsic component and the variable intrinsic component. Neglecting crowding, the effect of high-level injection on the intrinsic component of r_b' can be represented by modeling this component as a constant divided by q_b, the normalized base charge introduced in the GP model.[7,48] McBride[48] has included the effect of crowding (on r_b' and q_b) by the introduction of three extra model parameters and has shown that two-dimensional effects can be incorporated relatively well by splitting the transistor into two (non-identical) sections. Work is still being done on the three-dimensional effects and their influence on transistor models. In this respect, the use of computer programs such as SITCAP[25] (which predicts transistor characteristics from geometrical and process data) is proving invaluable.

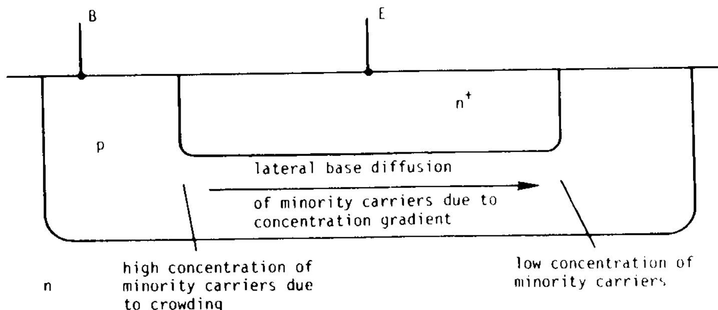


Fig. 2.27. **Lateral base diffusion that results from crowding in the base.**

b) Breakdown

The second important phenomenon that has been neglected here is breakdown. The problem of inclusion of breakdown has been solved.[49,50] Base-collector breakdown can be included very simply as a current-dependent current source between the collector and base terminals with a value of (M-1) I_C where M is the well-known breakdown multiplication factor and is given, to first order, by:[51]

$$M = \left[1 - \left(\frac{V_{BC}}{BV_{CBO}}\right)^n\right]^{-1} \tag{2.199}$$

The biggest problem with the inclusion of breakdown in the model lies not so much in its modeling, but rather with its effect on the convergence of the computer program. In most computer programs, the answer is obtained by an iterative procedure: a guess is made for the answer, an analysis is performed and the guess is improved. The effect of breakdown (which causes very large currents at some voltages near breakdown) on the convergence of these programs has not been thoroughly investigated.

c) Saturation

Another area in which the models presented can be inadequate is in the modeling of saturation; the dependence of r_c' on operating conditions and the phenomenon of "soft" saturation.[52,53]

For the great majority of users of CAD programs, the limitations of the EM_3 and GP models are perfectly acceptable. In fact, the EM_2 model (with its first-order modeling of charge-storage and ohmic resistances) is adequate for most analyses. It is rare that three-dimensional effects, breakdown and a more accurate saturation model is needed. More often than not, imaginative usage of the models and elements available in present-day computer programs can overcome many problems.*

*For example, the effects of crowding can be approximated by the connection of several "one-dimensional" transistors in parallel, each with its own constant base resistance.

3.1 Introduction

This section of the book describes measurement techniques that can be used to obtain the input parameters required for each level of complexity. The following points should be noted:

1. These measurement schemes are by no means meant to encompass all possible schemes; there are other schemes[54] some of which may be just as good, or even better. The measurement schemes described here have been tested and found to provide acceptable accuracy for computer simulations.
2. At least one page is devoted to each model parameter. It is defined, a typical value is given, and at least one measurement scheme is described.
3. The use of brand names indicates the type of equipment needed and is not meant to imply that only the specified brands can be used.

The structure of this section is as follows:

3.2 The EM_1 Model Parameters

3.3 The EM_2 Model Parameters

3.4 The EM_3 Model Parameters

3.5 The GP Model Parameters

The parameters described in each section are listed here:

EM_1: β_F

β_R

I_S

T_{nom}

E_g

EM_2: r'_e

r'_c

r'_b

C_{jo}, ϕ and m

τ_F (or f_T at I_C, V_{CE})

τ_R (or τ_{SAT})

C_{SUB} or C_{CS}

EM_3: V_A

β_{FM}, C_2, n_{EL} and θ (or β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE})

β_{RM}, C_4, n_{CL} and θ_R

RATIO

L_E/W, I_{CO}

TC_1, TC_2

GP: I_{SS}

V_B

I_K

I_{KR}

B

3.2 EM_1 Model Parameter Measurements

This section describes measurement techniques for obtaining the five parameters required for the EM_1 model. These parameters are:

- β_F, the forward common-emitter large-signal current gain,
- β_R, the inverse common-emitter large-signal current gain,
- I_S, the transistor saturation current (which is the common portion of the emitter-base and collector-base saturation currents),
- T_{nom}, the temperature at which the parameters are obtained.
- E_g, the energy gap of the transistor's semiconductor material.

Minimum equipment needed to obtain the EM_1 parameters is:

- a curve tracer or other setup for measuring dc characteristics
- a thermometer
- a calibrated temperature chamber

β_F

Definition

β_F is the ratio of the dc collector current to the dc base current when the transistor is in the normal, active region (i.e., with its base-emitter junction forward biased and its base-collector junction reverse biased).

Typical Value

Although β_F typically varies with collector current as shown in Fig. 3.1, for many applications the constant value assumed in the EM_1 and EM_2 models is adequate. Parameters needed to model β_F versus I_C are described in the EM_3 and GP sections.

Typical values of constant β_F may range from 5 for a high-current device to 100 for a small-signal device to 1000 for a super-β device.

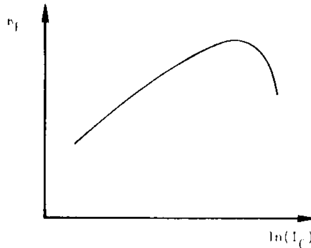


Fig. 3.1. Typical plot of β_F versus $\ln(I_C)$.

Measurement Scheme

The appropriate constant value of β_F can be determined from a curve tracer display of collector current versus collector-emitter voltage for a fixed base-current drive. β_F should be measured at the values of I_C and V_{CE} at which the transistor will be operated.

Note that the dc value should be used, not the ac value. For example, the $\beta_{F_{dc}}$ shown in Fig. 3.2 would be:*

$$\beta_{F_{dc}} = \frac{I_{C2}}{I_{B2}} \quad (3.1)$$

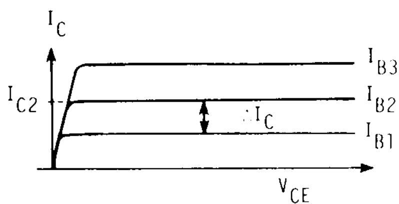


Fig. 3.2. Typical curve tracer characteristics of I_C versus V_{CE} for constant I_B.

The value of β_F can be obtained over a wide current range with modern curve tracer equipment. For example, values of base current as low as 5 nA are obtainable with the Tektronix 577 with 177 porch.

*The β_{Fac} would be:

$$\beta_{Fac} = \frac{\Delta I_C}{(I_{B2} - I_{B1})}$$

β_R

Definition

β_R, the inverse β, is the ratio of the dc emitter current to the dc base current when the collector-base junction is forward biased and the emitter-base junction is reverse biased.

Typical Value

A typical value of β_R is 1. It is usually assumed to be constant.

Measurement Scheme

β_R can be measured with the same technique as β_F but with the emitter and collector leads interchanged. With some curve tracers this can be accomplished simply by rotating the transistor test fixture 180°. As for β_F, the dc value should be determined, not the ac value.

I_S

Definition

I_S is the transistor saturation current. It is defined by the reciprocity relation:[4]

$$I_S \triangleq \alpha_F I_{ES} = \alpha_R I_{CS} \tag{3.2}$$

It can be measured from a consideration of I_C as a function of V_{BE} when the transistor is in the normal, active region (forward-biased base-emitter junction) with zero-biased base-collector junction. In this region, the collector current is given at low current levels by

$$I_C = I_S \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \tag{3.3}$$

Typical Value

I_S is directly proportional to the active emitter-base junction area and therefore can vary significantly from device to device. A typical value for an integrated circuit transistor is 10^{-16} amperes.

Measurement Scheme

To measure I_S, a curve tracer can be used to display collector current, I_C, versus collector-emitter voltage, V_{CE}, at a constant base-emitter voltage, V_{BE}. Figure 3.3 shows the characteristics obtained.*

I_S can be computed at a single value of V_{BE} or by plotting a graph of I_C as a function of V_{BE} with $V_{BC} = 0$. V_{BC} is set to zero as shown in Fig. 3.3 and not by shorting the base to the collector.

---

*Precaution: Because of the exponential dependence of I_C on V_{BE}, be careful not to set V_{BE} too large. Always start with small values and increase, never vice versa. A safer alternative would be to use a constant base-current drive and to measure V_{BE} externally (see p. 190).

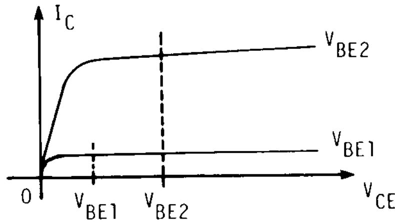


Fig. 3.3. Typical curve tracer I_C versus V_{CE} characteristics for constant V_{BE}.

Measurement at a Single Value of V_{BE}

Set V_{BE} equal to a convenient value* and measure I_C at $V_{CE} = V_{BE}$. I_S is then the measured value of I_C divided by the value of $\left(e^{\frac{qV_{BE}}{kT}}\right)$. Although this single-point measurement is simple and fast, it may not be accurate enough for many computer simulations. One is unaware of any deviations from the simple EM_1 theory.

Plotting the Graph

A more precise method is to obtain I_S from a plot of I_C versus V_{BE}. To do this, measure I_C from the characteristics of Fig. 3.3 at several points with $V_{CE} = V_{BE}$ so that $V_{BC} = 0$. Next, plot the natural logarithm (ln) of I_C as a function of V_{BE}, as in Fig. 3.4.** The value of I_S is then obtained by extrapolating the curve to $V_{BE} = 0$. The extrapolation can be done graphically or mathematically, the mathematical extrapolation being more accurate. At high-current levels the experimental curve deviates from a straight line due to the effects of high-level injection and/or ohmic resistance. Both high-level injection and ohmic-resistance effects are accounted for by the higher-order models.

*V_{BE} = 591.5 mV yields the convenient value of 10^{10} for the exponential divisor.

**Alternatively, simply use semi-log paper.

It is possible to use a curve tracer to display I_C as a function of V_{BE} directly. However, although the instrument accuracy is typically the same for both techniques, the use of the potentiometer removes the need for accurately setting the horizontal zero.

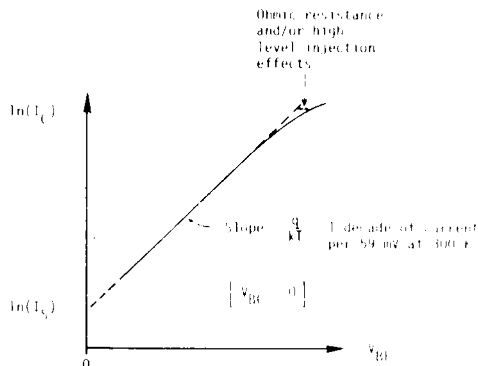


Fig. 3.4. Typical ln(I_C) versus V_{BE} characteristics. High-current departure from the straight line is explained in the EM_3 and GP models.

Tnom

Definition

T_{nom} is the temperature at which all the model parameters are obtained.

Typical Value

It is typically considered to be room temperature, about 27°C or 300°K.

Measurement Scheme

The simplest technique for measuring T_{nom} is by means of a thermometer placed near the transistor. As long as the power dissipation of the device is low enough to cause a negligible increase in the junction temperature, then the junction temperature is approximately the room temperature. All EM_1 model parameters can (and should) be obtained under low-power conditions (e.g., $V_{BC} = 0$). If thermal effects are suspected to be important, a higher-order model (such as the EM_3 model) should be used.

E_g

Definition

E_g is the effective energy gap of the semiconductor material.

Typical Value

E_g is typically 1.11 eV for silicon[57]

0.67 eV for germanium[57]

0.69 eV for Schottky-barrier diodes[3,57]

Measurement Scheme

E_g is used to model the variation of I_S with temperature and is therefore obtained by curve-fitting the model equation to the I_S versus T curve. The I_S versus T curve is determined from measurements of I_S (as described on pages 133 and 190) with the device in a controlled-temperature environment (for example, an oven)*.

The model equation given in Section 2 (Eq. 2.21) assumes that I_S is proportional to $T^3 \exp(-E_g/kT)$. This expression[15] neglects the temperature dependence of diffusivity ($=kT\mu(T)/q$).** In practice, the T-term before the exponential term has a power not necessarily equal to 3.[73,74] With the T^3 form forced (as in SPICE, Version 1) the only degree of freedom is E_g and it turns out that in normal temperature ranges 1.11 eV is a reasonable value for silicon. When the T-term power is also allowed to vary, typical values generated by curve fitting for silicon, are 1 to 4 for the power of T while for E_g (the extrapolated, zero-temperature energy gap) typical values range from 1.206 eV for low doping levels to approximately 1.11 eV for high doping levels.[73]

*Care must be taken in making these measurements that the power dissipated by the device does not raise the junction temperature significantly above the ambient temperature.

**This expression also neglects the temperature dependence of the bandgap narrowing caused by heavy doping.[73]

3.3 EM_2 Model Parameter Measurements

This section gives measurement techniques for obtaining the twelve additional parameters for the EM_2 model. The five EM_1 model parameters described in Section 3.2 are also part of the full EM_2 model.

The additional twelve parameters for the EM_2 model are:

r'_e, the emitter ohmic resistance,

r'_c, the collector ohmic resistance,

r'_b, the base ohmic resistance,

C_{jE0}, the emitter-base junction capacitance at $V_{BE} = 0$,

C_{jC0}, the collector-base junction capacitance at $V_{BC} = 0$,

ϕ_E, the emitter-base barrier potential

ϕ_C, the collector-base barrier potential,

m_E, the emitter-base capacitance gradient factor,

m_C, the collector-base capacitance gradient factor,

τ_F, the total forward transit time (which can be computed from the transistor's unity-gain bandwidth f_T),

τ_R, the total reverse transit time (which can be calculated from the saturation time constant τ_{SAT}),

C_{SUB}, the substrate capacitance

The best way to determine any of the twelve EM_2 model parameters is to duplicate, as nearly as possible, the operating conditions under which the analysis is to be run. Minimal equipment needed is:

- a curve tracer
- a capacitance bridge
- a power supply

- a pulse generator and a high speed oscilloscope
- a small-signal measurement system for determining f_T (such as an s-parameter system).

NOTE: Only one (general) set of junction capacitance parameters (C_{j0}, ϕ and m) is described here, since $C_{jE}(V_{B'E'})$ and $C_{jC}(V_{B'C'})$ are both obtained in the same way.

r'_e

Definition

r'_e is a constant resistor which models the resistance between the active emitter region and the emitter terminal.

Typical Value

A typical value of r'_e is approximately 1 ohm, of which the metal contact resistance is normally a significant portion.

Measurement Scheme

The value of r'_e can be obtained by observing the base current as a function of collector-emitter voltage, V_{CE}, for a transistor with an open-circuited collector.[58] The setup is shown in Fig. 3.5.

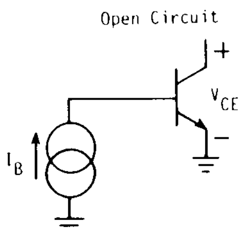


Fig. 3.5. Set-up used to measure r'_e.

The resultant graph is as shown in Fig. 3.6 and it can be readily shown* that the slope as marked is approximately $(r'_e)^{-1}$.

---

*To show that the slope is equal to $(r'_e)^{-1}$, it is only necessary to note that the transistor is saturated. Therefore, V_{CE} is given by the sum of the inherent saturation voltage of the device (see Fig. 2.6) and the drop across r'_e. That is, since $I_E = I_B$,

$$V_{CE} = \frac{kT}{q} \ln\left(\frac{1}{\alpha_R}\right) + I_B r'_e$$

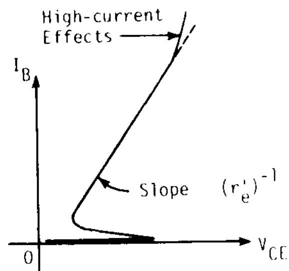


Fig. 3.6. Plot of I_B versus V_{CE} obtained from the r'_e measurement set-up.

The low-current "flyback" effect is caused by the decrease of the inverse beta, β_R, at low currents. (Sometimes the flyback effect is difficult to observe.) The slope should be determined as close as possible to the flyback region. At high currents, the trace of I_B versus V_{CE} departs from a straight line.

As an example of how the characteristics of Fig. 3.6 can be obtained directly from a curve tracer, the settings needed to obtain r'_e on a Tektronix 576 curve tracer are shown in Fig. 3.7a and in Fig. 3.7b for a Tektronix 577 curve tracer.

TEST SET-UP CHART TYPE 576

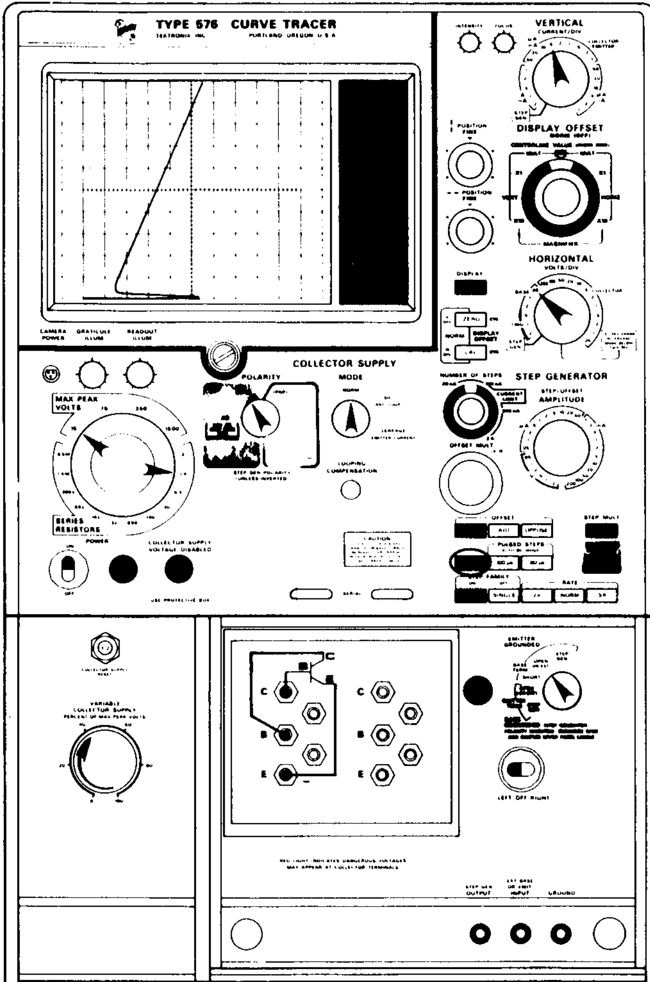


Fig. 3.7a. Settings on the Tektronix 576 Curve Tracer to determine r'_e. Note: It may be necessary to vary the settings of VERTICAL and HORIZONTAL to obtain the appropriate display. Settings shown for an npn. For a pnp, change POLARITY to PNP and press DISPLAY INVERT.

TEST SET-UP CHART 577-177-D1

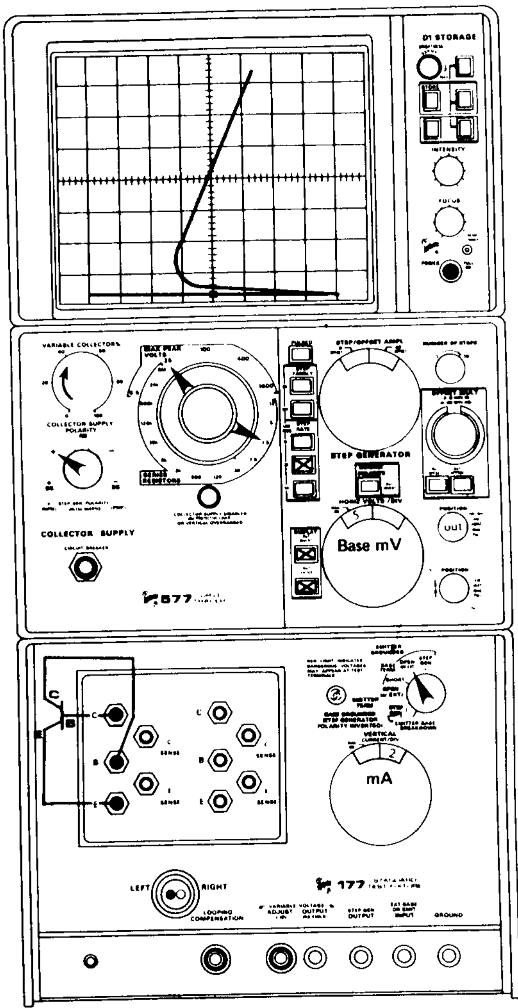


Fig. 3.7b. Settings on the Tektronix 577 Curve Tracer to determine r'_e. Note: It may be necessary to vary the settings of VERTICAL and HORIZONTAL VOLTS/DIV to obtain the appropriate display. Settings shown for an npn. For a pnp, change POLARITY to – and put DISPLAY in INVERT mode.

r'_c

r'_c

Definition

r'_c models the resistance between the transistor's active collector region and its collector terminal.

Typical Value

For a given device, r'_c actually varies with current level, but for the EM_2 model it is considered to be constant. The value of r'_c can vary significantly from device to device: from a few ohms for discrete and deep-collector integrated devices to hundreds of ohms for standard integrated devices.

Measurement Schemes

A big problem associated with measuring r'_c is which value to use. Selection of the r'_c value depends strongly on how the transistor is being used or which aspect of the device behavior needs to be modeled accurately. All following measurements can be made with a curve tracer.

a) Definitions of two limiting values of r'_c

Typical I_C versus V_{CE} transistor characteristics from a curve tracer are sketched in Fig. 3.8. The dashed lines labelled A and B represent the two limiting values of r'_c.

(i) Normal, Active Region Value ($r'_{c_{normal}}$)

The dashed line A is drawn through the "knee" of each curve, which is where the curve departs from the straight-line approximation of the normal, active region. The inverse of its slope ($r'_{c_{normal}}$) is the ohmic collector resistance when the device is in the normal, active mode and is not saturated.[52] The value of r'_c obtained from line A normally agrees well with the value computed from a knowledge of resistivity and geometry. The value of $r'_{c_{normal}}$ should be used when the device is in the normal, active region, when it is never saturated and when correct modeling of f_T and charge storage is required.

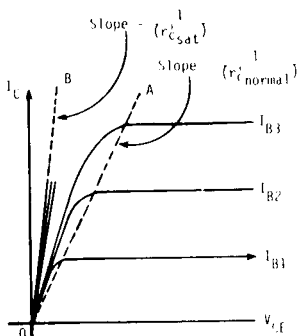


Fig. 3.8. Typical I_C versus V_{CE} characteristics at constant I_B, as seen on a curve tracer. Lines A and B show two limiting values of r_c'.

(ii) Saturation Region Value ($r'_{c_{sat}}$)

If the transistor is strongly saturated, the inverse of the slope of the dashed line B ($r'_{c_{sat}}$) provides the appropriate value of r_c'.

(iii) Both Saturation and Normal, Active Regions

When the transistor is to be modeled accurately in both its saturation and normal, active regions, an appropriate compromise should be made. Since the r_c' value is used in some programs to internally compute transit time τ_F from bandwidth f_T, it may be advisable to specify τ_F directly (if possible) whenever the r_c' value for the normal, active region ($r'_{c_{normal}}$) is not used. Note that because of the structure of the transistor model with r_c' constant, the value of V_{CE} is never less than $I_C r_c'$ and the computed characteristics can never lie to the left of the straight line with slope $1/r_c'$.

b) Measurement of $r'_{c\,normal}$ and $r'_{c\,sat}$

The values of $r'_{c_{normal}}$ and $r'_{c_{sat}}$ can be obtained directly from a curve tracer display of the I_C versus V_{CE} characteristics, in accordance with their above definitions. However, there can be some difficulties associated with this measurement. In the $r'_{c_{normal}}$ case, the "knees" are difficult to determine with any degree of precision. For the $r'_{c_{sat}}$ case, a correction must be applied to take into account the effects of r'_e and the finite slope of the EM_1 characteristics (i.e., $r'_c = 0$) in the saturation region (see Fig. 2.6). The corrected value of $r'_{c_{sat}}$ is given by[56]*

$$r'_{c_{sat}} = \frac{1}{\text{Slope B}} - \left(1 + \frac{1}{\beta_F}\right) r'_e - \frac{kT}{q}\left[\frac{1}{\beta_F I_B - I_C} + \frac{1}{(1+\beta_R)I_B + I_C}\right] \tag{3.4}$$

The following measurement technique,** still possible with a curve tracer, eliminates some of these disadvantages.

In general, V_{CE} has the following dependence on bias point,[4,56] when saturated.

$$V_{CE} = \frac{kT}{q} \ln\left[\frac{1 + \frac{I_C}{I_B}(1-\alpha_R)}{\alpha_R\left(1 - \frac{I_C}{\beta_F I_B}\right)}\right] + I_E\, r'_e + I_C\, r'_c \tag{3.5}$$

The above expression for $r'_{c_{sat}}$ (Eq. (3.4)) arose from a consideration of the first term (which leads to the finite slope of the EM_1 curves) as well as the r'_e term of Eq. (3.5). The following measurement avoids the correction for the first term of Eq. (3.5) by arranging to cancel out its effect.

---

*Since $r'_{c_{normal}}$ is usually much larger than $r'_{c_{sat}}$, the appropriate (small) correction, is not normally made for $r'_{c_{normal}}$.

**Measurement technique developed by Dr. B. A. Rosario, Tektronix, Inc.

The measurement involves a simple extension to the r'_e measurement scheme.* Instead of an open-circuited collector, a stepped collector current is used, as shown in Fig. 3.9a. The overall characteristics obtained are sketched in Fig. 3.9b. The $I_C = 0$ curve is simply the curve described in the r'_e measurement. The other curves correspond to constant-I_C curves. Region B (high values of I_B for a given I_C value) is used to obtain $r'_{c_{sat}}$ while Region A (low values of I_B near the knees) is used to obtain $r'_{c_{normal}}$. Figure 3.9c illustrates how r'_c is obtained from these curves in both regions. By determining ΔV for two points where $\frac{I_C}{I_B}$ is constant, the first term in the above expression for V_{CE} is constant** and therefore

$$\Delta V = (I_{C2} - I_{C1})r'_c + (I_{C2} + I_{B2} - I_{C1} - I_{B1})r'_e \qquad (3.6)$$

This expression which assumes that r'_e and r'_c are unchanged at both measurement points, can be used to obtain r'_c.***

Notes:

1. Equation (3.5) only applies for the device in the saturation mode. Therefore, Region A, which is used for determining $r'_{c_{normal}}$, must be as close to the knee as possible (corresponding to the weakly saturated case). The accuracy of this technique for determining $r'_{c_{normal}}$ is thus a function of the amount of saturation. Difficulties may be experienced with this measurement near the knees.

---

*It would appear at first that the r'_e scheme could be used for $r'_{c_{sat}}$ by simply interchanging the collector and emitter leads. However, because of the different carrier distributions that result, the effective junction areas must be taken into account. The r'_e measurement technique used this way will invariably give a value for $r'_{c_{sat}}$ that is too low.

**Assuming α_R and T are also constant.

***Before making these measurements it is advisable to first inspect the standard I_C versus V_{CE} characteristics (Fig. 3.8) to determine the I_C, I_B and V_{CE} values at which these measurements are to be made.

2. If the lines are vertical in the measurement region (i.e., r'_e small), ΔV can be determined at one value of I_B.
3. The above analysis assumes α_R, β_F, r'_c, r'_e and T are constant for the two points. Therefore small changes in I_C and I_B should be used.

Figure 3.9d shows the settings on the Tektronix 576 curve tracer needed to obtain these characteristics and Fig. 3.9e shows the corresponding settings for the Tektronix 577 curve tracer.

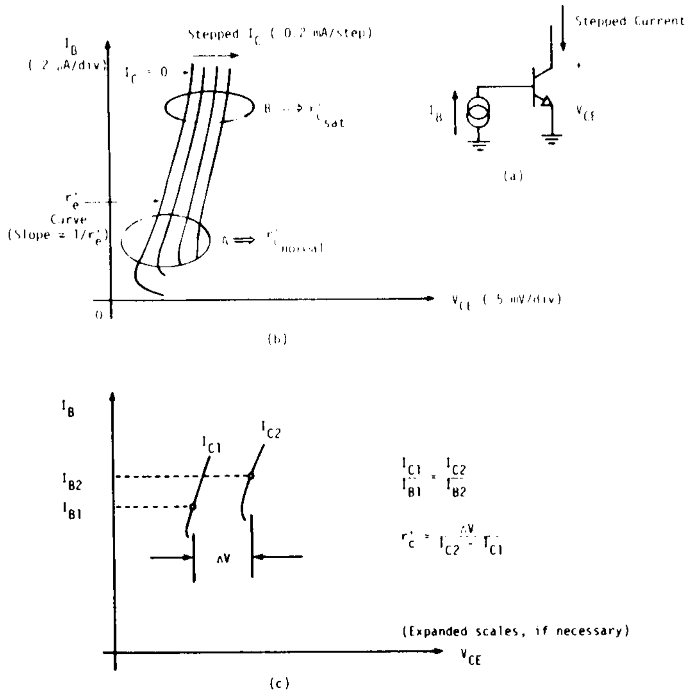


Fig. 3.9. r'_c measurement: a) the test circuit; b) the characteristics observed from the test circuit; c) an expanded section of curves defining ΔV.

TEST SET-UP CHART TYPE 576

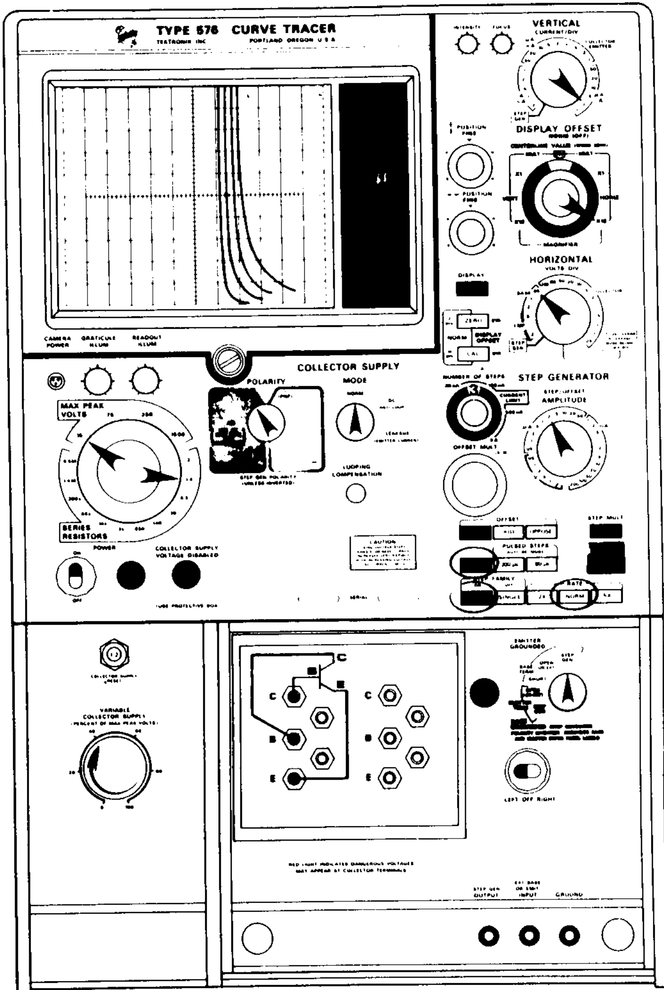


Fig. 3.9d. Settings on the Tektronix 576 Curve Tracer to obtain the characteristics of (b). (Note: It may be necessary to vary the settings of VERTICAL, HORIZONTAL, DISPLAY OFFSET and STEP GENERATOR to obtain the appropriate display.) Settings shown for an npn. For a pnp, change POLARITY to PNP and press DISPLAY INVERT.

TEST SET-UP CHART 577-177-D1

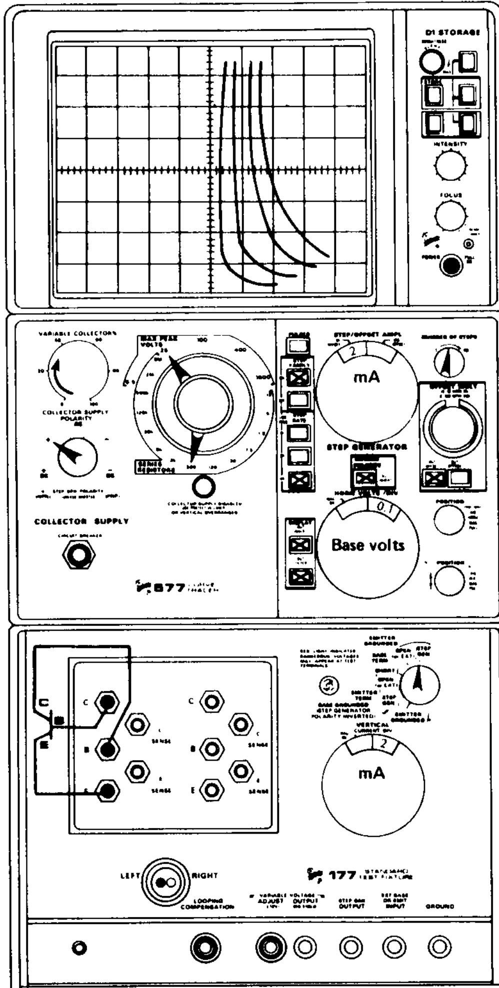


Fig. 3.9e. Settings on the Tektronix 577 Curve Tracer to obtain the characteristics of (b). (Note: It may be necessary to vary the settings of VERTICAL, HORIZONTAL VOLTS/DIV and STEP/OFFSET AMPL to obtain the appropriate display.) Settings shown for an npn. For a pnp, change POLARITY to – and put DISPLAY in INVERT mode.

r_b'

Definition

r_b' models the resistance between the active base region and the base terminal.

Typical Value

Values of r_b' can range from approximately 10 ohms (for microwave devices) to several kilohms (for lower frequency devices). r_b' varies with the operating condition but in the EM_2 model it is assumed to be constant.

Measurement Schemes

a) Introduction

Traditionally, r_b' has been a difficult parameter to measure, mainly because it is modeled as a lumped constant resistance although it is actually a distributed, variable resistance.[18] As a result, the value obtained for r_b' depends strongly on the measurement technique used as well as the transistor's operating conditions. Several measurement techniques are described here.

b) Comparison of measurement techniques for r_b'

r_b' should be determined by the method closest to the operating conditions under study. If r_b' is being measured to ascertain its effect on noise performance, the noise measurement technique should be used. Similarly, if the transistor is to be used in a switching application, the pulse measurement techniques may provide the most appropriate value. For small-signal analyses four measurement techniques are described: the input impedance circle method, the phase-cancellation technique, the two-port network method and the h-y ratio technique. A comparison of the advantages and disadvantages of these four methods is given in Section c), the small-signal measurement section.

For dc analyses, it may be possible to obtain r_b' from a plot of $\ln(I_C)$ and $\ln(I_B)$ versus V_{BE} (see Fig. 2.18). However, since this procedure involves subtracting two large numbers, substantial errors can be introduced. In fact, it is not uncommon to obtain negative values for r_b' by this method.

c) Small-signal measurements

(i) The Input Impedance Circle Method[59,17]*

This measurement assumes the small-signal linear hybrid-π transistor model. As the signal frequency is varied, with the ac collector voltage kept at zero, the input impedance into the base-emitter junction is plotted on the complex impedance plane.

The locus of points forms a semi-circle as shown in Fig. 3.10. The right intercept of the circle with the real axis occurs at zero frequency (dc). The impedance value (for zero r_e') is the sum of base resistance r_b' and resistance r_π. The impedance value of the

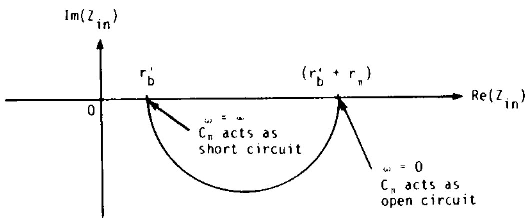


Fig. 3.10. The locus of the common-emitter input imedance as a function of frequency for the simple hybrid-π model.

*This section contributed by Prof. H. Ablin, University of Nebraska.

left intercept which occurs at infinite frequency is (for zero r_e') r_b' alone.*

The accuracy of this measurement depends on the value of the collector current. At a low collector current the value of r_π will be large, resulting in a large semicircle. When the collector current is high the value of r_π will be small, giving a small semicircle, and permitting the left intercept to be determined more accurately.

At high frequencies, the linear hybrid-π model is no longer accurate. The distributed nature of the transistor and such parasitic elements as lead capacitance cause the measured points to deviate from the predicted semicircle (Fig. 3.11). When this happens, the semicircle construction is based on the measured points obtained at low frequencies.

Corrections for the effects of parasitic capacitances and r_e' are available and should be applied.[17]

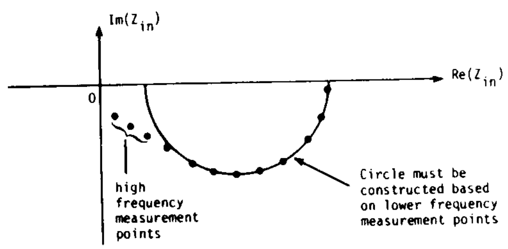


Fig. 3.11. Typical plot of experimental values of the common-emitter input impedance.

*The effect of a finite r_e' is to add r_e' to the $\omega = \infty$ intercept and $(1 + \beta)r_e'$ to the $\omega = 0$ intercept.[17]

Equipment needed for this measurement scheme is either:

- an RX Meter such as the Boonton Model 250A with an appropriate test jig,

or - a variable-frequency admittance bridge such as the Wayne-Kerr Model 801B,

or - an s-parameter measurement setup such as the General Radio Model 1710 Network Analyzer. (With this equipment, data must first be converted into input impedance and must be able to be taken at low frequencies so that the semicircle can be fitted.)

(ii) Phase-cancellation Technique(17)

The transistor is connected in its common-base configuration and an admittance bridge (such as the Wayne-Kerr 801B) or an s-parameter system is used to measure the real and imaginary parts of the device's input impedance across the base-emitter junction. At any frequency between $\frac{f_T}{3}$ and $\frac{f_T}{\beta}$, the collector current is varied until the reactive part of the input impedance goes to zero. At this value of collector current, it can be shown that(17)

$$r_e' + r_b' = g_{mF}{}^{-1} = \frac{kT}{qI_C} \tag{3.7}$$

This technique cannot be used with devices that have low values of β (such as lateral pnp transistors).

(iii) Two-port Network Measurement Technique*

This technique has been successfully used for simulating the transistor in the frequency domain. It has the advantages of using explicit formulae for parameter determination and with it parameter error calculations can be made (i.e., the two-port parameters of a lumped model can be evaluated at one frequency and compared to measured parameters). The technique, which can also be used for determining r_b', makes use of a simplified (and slightly

---

*This section contributed by T. Hallen and C. Battjes, Tektronix, Inc.

modified) version of the linearized EM_2 model of Fig. 2.12. In this model, shown in Fig. 3.12, the C_μ capacitor is connected to the external base node, which is more appropriate for modern, stripe-geometry integrated transistors.

The base resistance, r_b', is obtained directly from the four y parameters for the device connected in the common-emitter configuration. The particular form of the model assumed in Fig. 3.12 results in a closed-form solution for r_b' at one frequency. The analysis of the y parameters yields:*

$$r_b' = \frac{1}{g_{mF}} \left\{ \text{Real} \left(\frac{y_{ie} + y_{re}}{g_{mF} + y_{re} - y_{fe}} \right) \right\}^{-1} \qquad (3.8)$$

where, assuming low-level injection conditions,

$$g_{mF} = \frac{qI_C}{kT} \qquad (3.9)$$

Since the model used in Fig. 3.12 does not include the emitter ohmic resistance, r_e', as a separate entity, a more accurate value of r_b' is obtained if r_e' is accounted for. This can be done by replacing g_{mF} with g_{mEFF} in Eq. (3.8) where

$$g_{mEFF} = \left[\frac{1}{g_{mF}} + \frac{\beta_{F_{ac}} + 1}{\beta_{F_{ac}}} r_e' \right]^{-1} \qquad (3.10)$$

*The analysis of the model assumes a complex admittance for the r_b', C_μ and C_{SUB} branches. Therefore the analysis has four complex unknowns (Y_b', Y_π, Y_μ and Y_{SUB}) and four complex measurements (y_{ie}, y_{re}, y_{fe} and y_{oe}). It is, of course, assumed that at the operating point g_{mF} is known from Eq. (3.9). The value of r_π obtained from this technique can sometimes be negative. This is the way the model matches the phenomenon of excess phase shift in the transistor.

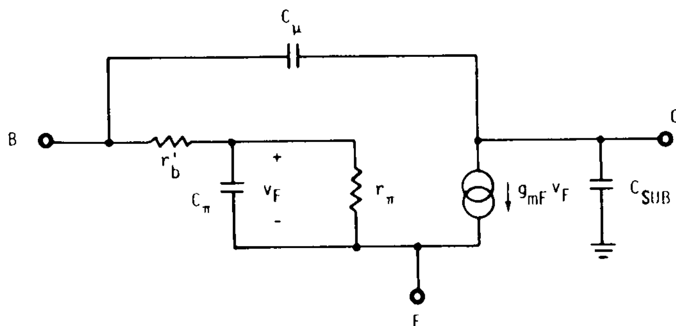


Fig. 3.12. Model used for the two-port measurement of r'_b. Note that C_μ is connected to the external base node. This is more appropriate for modern, stripe-geometry integrated transistors.

The determination of r'_b is basically enhanced by using a high measurement frequency. At high frequencies the input impedance is no longer dominated by the r_π and C_π and the transadmittance departs significantly from the ideal. In effect r'_b has been "revealed" since it is responsible for this behaviour. Modern s-parameter measuring or network analyzer equipment can be used to measure high-frequency two-port parameters. At low frequencies this measurement technique loses accuracy since r'_b tends to be found as a small difference between large quantities. Typically the measurement frequency for this two-port r'_b measurement technique should be between f_β and f_T (say 0.1 f_T for $\beta \sim 100$).*

(iv) The h-y Ratio Technique(70)

This measurement makes use of the fact that for the linear, EM_2 hybrid-π model, r'_b is given by(70)

$$r'_b = \frac{h_{f_e}(0) \cdot f_\beta}{y_{f_b}(0) \cdot f_y} = \frac{f_T}{y_{f_b}(0) \cdot f_y} \qquad (3.11)$$

*In this frequency range (between f_β and f_T), any emitter lead inductance, L_e, will tend to appear as an extra component of r'_b (approximately equal to $2\pi f_T L_e$).

where $h_{fe}(0)$ is the low-frequency value of h_f in the common-emitter configuration ($= \beta_{F_{ac}}$), f_β is the -3 dB frequency for h_{fe}, $y_{fb}(0)$ is the low-frequency value of y_f in the common-base configuration, f_y is the -3 dB frequency for y_{fb} and a single, dominant-pole response has been assumed for both h_{fe} and y_{fb}. All the terms on the right-hand side of Eq. (3.11) are measurable. The numerator is simply f_T and its measurement is described in the $\tau_F(f_T)$ measurement. The denominator, like the numerator, can be obtained from a network analyzer or s-parameter measurements. The advantage of this technique is that Eq. (3.11) is relatively unaffected by finite values of r_e' and r_c' and that the measurement of y_{fb} is relatively insensitive to stray capacitances.

(v) Comparison of the Small-signal Measurements

All four measurement techniques have their advantages, causes of inaccuracies, limitations and disadvantages. The input impedance circle method is accurate but involves the most work (especially when corrections are made to account for r_e' and parasitic capacitances).[17] It loses accuracy at low collector currents when the diameter of the circle is large. The phase-cancellation technique is quick and is relatively unaffected by r_e' (since it adds directly to r_b' rather than being first multiplied by β). However, it can only be used for devices with $\beta_{F_{ac}}$ greater than approximately 10 (to satisfy the frequency criterion) and applies at only one non-selectable value of collector current (that is usually low). A more detailed comparison of these two techniques is given by Sansen and Meyer.[17] The two-port network measurement technique is also simple, being performed at one frequency. For greatest accuracy, the frequency of the measurement should be between f_β and f_T (so that r_b' is not found from the subtraction of two large numbers). The h-y ratio method is a technique that is quite insensitive to r_e'. It requires the determination of the full frequency response of the two measurements to verify the assumption of a dominant pole for each case. One of the measurements, f_T, is used in the determination of another EM_2 model parameter (τ_F) while the other measurement is relatively insensitive to stray capacitances.

d) Pulse measurement techniques

(i) Low-Frequency Pulse Measurement Method[60]

The test circuit used for this technique is shown in Fig. 3.13 (with an npn transistor).

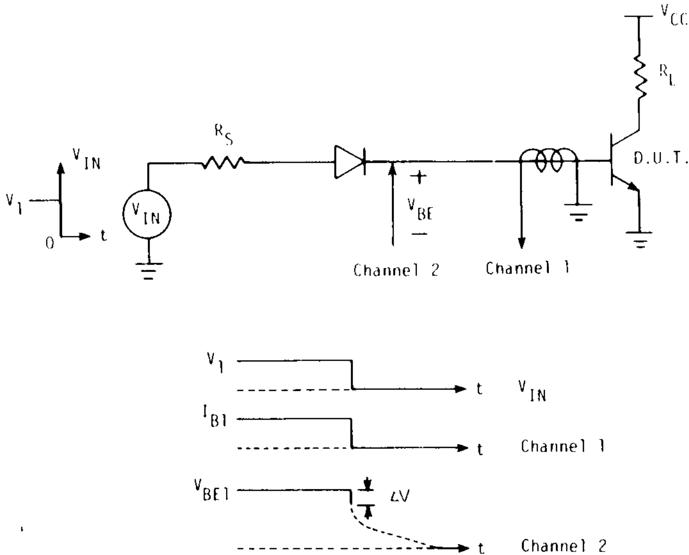


Fig. 3.13. Measurement set-up to determine r_b' by the pulse method. Waveforms are those observed during test.

A current pulse applied to the base of the transistor through a fast switching diode causes the device to turn off. The voltage across r_b' "instantaneously" drops to zero while the base-emitter capacitance keeps the internal junction potential ($V_{B'E'}$) constant. Resistance r_b' can then be determined from the display on a dual-channel oscilloscope by

$$r_b' = \frac{\Delta V}{I_{B1}} \tag{3.12}$$

This technique does not necessarily work well for all transistors. When the external component of the base resistance r_b' is small with respect to the internal component, the ΔV drop is not readily observable. However, other useful information can be gained from this technique. When the oscilloscope time-per-division scale is reduced to the point where ΔV no longer appears to be vertical, the simple constant-resistance model for r_b' is no longer valid -- giving some indication of the switching times at which the EM_2 model for r_b' is inadequate.

Equipment required for the pulse measurement technique includes:

- a pulse generator such as the Tektronix Model 109
- a current probe such as the Tektronix Model P6042 or Model CT-1
- a dual-channel oscilloscope such as the Tektronix 500 Series or 7000 Series
- a power supply
- a fast diode such as the FD 7003 (whose reverse recovery time is less than the expected decay time).

If supply voltage V_{CC} and biasing resistors R_L and R_B are chosen so that the transistor is saturated, greater sensitivity can be obtained from the current probe. However, the value of r_b' can be significantly different from that obtained when the transistor is in its normal, active region.

(ii) The Time-Domain Reflectometry Method*

For a high-frequency transistor, the capacitive loading of the oscilloscope at the base can affect the accuracy of the above pulse measurement method. For such a device, a time-domain reflectometry system** shown in Fig. 3.14 can be used to measure r_b'.

The technique consists of observing the pulse reflection from the base of a transistor terminating a coaxial transmission line. The pulse reflection contains information about the driving-point impedance at the base of the transistor. If one restricts the input circuit model complexity to that of a traditional hybrid-π transistor model it is fairly easy to solve for r_b'. Since the desired information is contained in the first fast portion of the reflection, required equipment is a fast pulse generator, such as the Tektronix S-52 pulse generator, and a sampling head such as an S-6 head installed in a 7S12 plug-in for the Tektronix 7000 Series oscilloscope.

The test circuit is shown in Fig. 3.14a. The amount of reflection (K) and its polarity give the magnitude of r_b'. To obtain valid results, high-frequency effects must be carefully considered in layout and bypassing. For example, to minimize stray inductance, leadless capacitors and short connections are essential.

Figure 3.14b shows various waveforms which may be observed with this setup. The reflection from the line termination reaches the sampler input at time t_1. If a resistor of value R_0 is substituted for the test transistor (first waveform) a smooth step response is obtained (no reflection will be observed).

The next three waveforms (solid lines) are ideal waveforms showing the reflections for the transistor input circuit shown in Fig. 3.14c. In each case the discontinuity is the reflection and the fractional discontinuity is the reflection coefficient of an r_b' termination of an R_0 transmission line. That is:

---

*This section contributed by C. Battjes, Tektronix, Inc.

**General time-domain reflectometry techniques are described by Strickland.[61]

$$K = \frac{(r_b' - R_0)}{(r_b' + R_0)} \tag{3.13}$$

from which it follows that:

$$r_b' = \frac{(1 + K)}{(1 - K)} R_0 \tag{3.14}$$

In practice the waveforms will be slightly more complicated. Figure 3.14d is a better representation of the transistor input and the complication of the circuit modifies the reflection picture in Fig. 3.14b. The dotted lines on Fig. 3.14b show, in an exaggerated manner, the effects of the added element, C_{ext}, which models any stray capacitance as well as that due to the portion of C_{jC} that is really connected to the outside of r_b' (see explanation of RATIO in EM_3 model). The basic response, as represented by the solid line waveform can be visualized by smoothing (i.e., stripping off the spike due to C_{ext}).

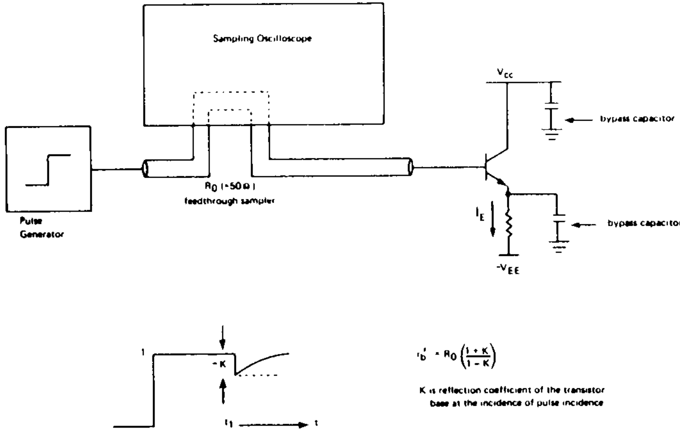


Fig. 3.14. (a) Measurement set-up for time-domain reflectrometry measurement of r_b'.

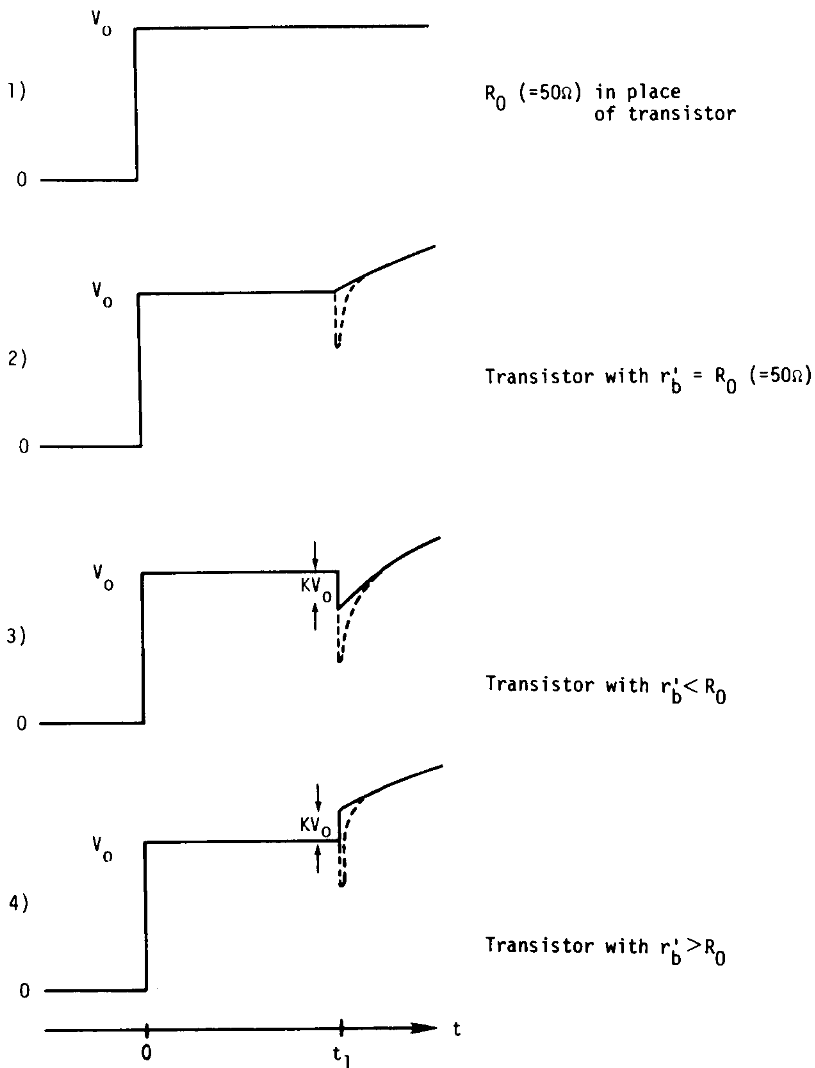
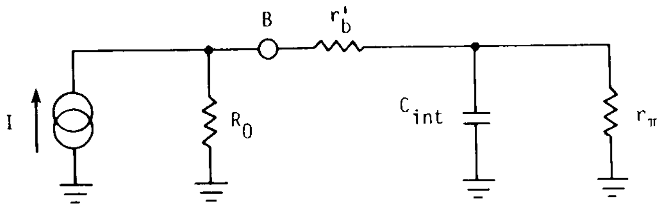
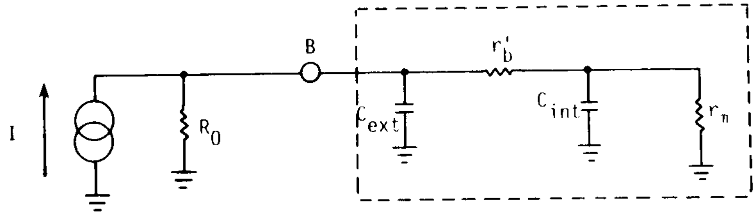


Fig. 3.14. (b) Observed time-domain reflectometry r_b' measurement waveforms for conditions as stated.

r'_b



(c)



(d)

Fig. 3.14. (c) Transistor input circuit used to generate waveforms in (b); (d) Improved equivalent circuit assumed for transistor.

e) Noise measurement technique[62,63]*

The use of noise measurements to determine base resistance presents a number of problems to anyone unfamiliar with noise work, and probably should not be attempted by a beginner. The measurements require the use of very-high-gain amplifiers whose gain is stable with time, as well as extensive shielding to prevent excessive rf interference and 60 Hz pick-up. Commercial equipment for this purpose is comparatively expensive.

Once a noise measurement system is set up, the method can be quite convenient.[56] The transistor is inserted into the apparatus and a single meter reading allows fast estimation of the base resistance. If the flicker noise is assumed to be negligible, r_b' can be estimated as:

$$r_b' = \frac{(\overline{v_i^2})}{(4kT\Delta f)} - \left(\frac{1}{2g_{mF}}\right) \tag{3.15}$$

where Δf is the bandwidth of the measurement, g_{mF} is calculated from the known collector current, and quantity $\overline{v_i^2}$ is the transistor's equivalent input mean-square noise voltage. The magnitude of $\overline{v_i^2}$ is determined from:

$$\overline{v_i^2} = \frac{\overline{v_0^2}}{G^2} \tag{3.16}$$

where $\overline{v_0^2}$ is the measured output mean square noise voltage from the test system and G is the voltage gain from the test device input to the system output. This measurement is performed with an ac short circuit between the transistor's base and emitter. Also, $\overline{v_0^2}$ must be measured on a true-rms-reading voltmeter.

---

*This section contributed by Prof. R. G. Meyer, University of California, Berkeley.

C_{jO}, Φ and m

Definition

C_{jO}, ϕ and m, are the three parameters that describe the junction capacitance due to the fixed charge in a junction depletion region. In the EM_2 model, when the appropriate junction voltage (V) is less than or equal to $\frac{\phi}{2}$, the junction capacitance is modeled by[19]

$$C_j(V) = \frac{C_{j0}}{\left(1 - \frac{V}{\phi}\right)^m} \tag{3.17}$$

where C_{j0} is the value of C_j at V = 0, ϕ is the built-in barrier potential, and m is the capacitance gradient factor.

For the emitter-base junction the subscript E is added. The junction capacitance, C_{jE}, is a function of the internal base-emitter voltage ($V_{B'E'}$) and the parameters are C_{jEO}, ϕ_E and m_E. Similarly, for the collector-base junction, $C_{jC}(V_{B'C'})$ is expressed in terms of C_{jCO}, ϕ_C and m_C.

Independent of the type of device (npn or pnp), in the Berkeley programs V is positive if the junction is forward-biased and negative if the junction is reverse-biased.

Typical Value

Typically C_j varies with V as shown in Fig. 3.15 for V less than or equal to $\frac{\phi}{2}$. C_{j0} varies from device to device, but is typically of the order of 0.3 pF/mil^2 of junction area. Barrier potential ϕ is usually about 0.5 V to 0.7 V and gradient factor m is assumed to be between 0.333 and 0.5, the graded junction and abrupt junction values, respectively.[19]

Measurement Scheme

Both junction capacitances can be obtained as a function of voltage by means of a bridge such as the Boonton Model 75. The two junction contacts are connected to the bridge and the third contact

C_{J0}, ϕ, m

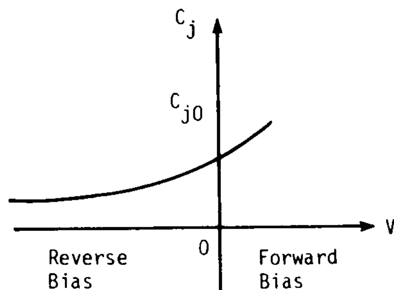


Fig. 3.15. A typical variation of C_j as a function of V.

is left open.* For example, for C_{jE}, the emitter and base leads are connected to the bridge and the collector contact is left open. The measurement frequency is normally low enough so that it is assumed that the ohmic resistances have a negligible effect.

A complicating factor is the extra capacitance, C_k, mainly caused by pin capacitance, stray capacitance, and pad capacitance. C_k is normally assumed to be constant. The capacitance that is measured by the bridge is:

$$C_{meas} = \frac{C_{j0}}{\left(1 - \frac{V}{\phi}\right)^m} + C_k \tag{3.18}$$

C_k can be determined in four ways: by an estimate (approximately 0.4 to 0.7 pF), by measurement with a dummy can, by a computer parameter optimization procedure or by graphical techniques.

The dummy can technique is the most accurate method. It requires an identical device can with its metal run disconnected

*If the bridge (such as the Boonton 75) can perform a three-terminal measurement, in which stray capacitances between the two test pins and the third pin are ignored, the third contact can be connected to this third pin rather than left open. Care must then be taken when dc voltages are applied (e.g., the base should not be connected to a voltage-varying terminal such as HI on the Boonton 75).

(scratched off) at the emitter or collector (but not the base). This dummy package can either be used to zero the capacitance bridge or its capacitance can be measured separately and the measured value subtracted from the bridge measurements.

The use of an optimization algorithm on a computer or calculator is fast and convenient once the algorithm has been written and tested. However, as with the graphical techniques described below, the solution is often not unique; several sets of solutions can be obtained depending on the initial estimates and the methods used. Since the parameters C_{j0}, ϕ and m are used in the programs *only* to recreate the junction capacitances,* *any set of positive values for these parameters is acceptable*.

A method of reducing the data by graphical techniques is to make an initial guess for ϕ and C_k, and then plot the resultant value of $(C_{meas} - C_k)$ as a function of $(\phi - V)$ on log-log graph paper. If a straight line (with a slope between -0.5 and -0.333) results, the chosen values are assumed to be correct. If the plotted line is not straight a second guess is made for C_k and/or ϕ and the plot redone.** This process continues until the appropriate straight line is obtained (line "a" in Fig. 3.16). Since the slope of the straight line is equal to -m, the values of ϕ, m, C_k and C_{j0} can be determined from this plot.***

An alternative graphical technique is to plot $(C_{meas} - C_k)^{-1/m}$ as a function of V. When a straight line is obtained, ϕ is determined by extrapolating the line to the V axis.

---

*Except when accurate variation of τ_F with temperature (EM_3 model) and I_C (GP model) is required. For these cases, ϕ_C is used.

**If the curve is concave (line "b" in Fig. 3.16) either decrease ϕ or increase C_k or both.

***It may sometimes be necessary to accept a value of m outside the 0.5 to 0.333 range in order to get a good fit. This can arise, for example, when the junction appears to be linear for smaller values of V and a step for larger values of V. For an integrated circuit, only part of the C_k thus found is a parasitic.

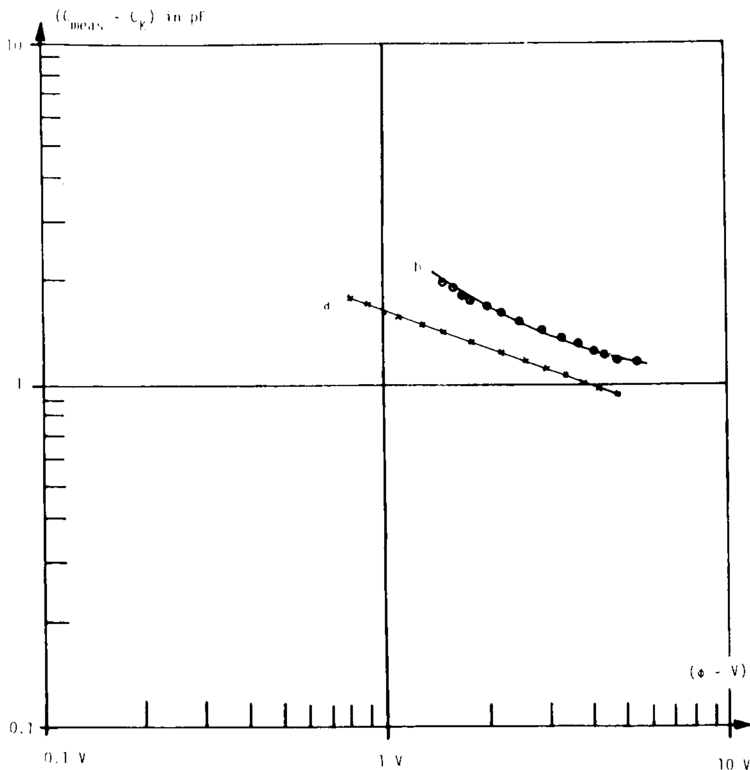


Fig. 3.16. Typical plots of (C_{meas} - C_k) versus (ϕ - V).
(a) Straight line results for "correct" values of C_k and ϕ.
(b) The concave curve resulting from incorrect values of C_k and ϕ.

τ_F (or f_T at I_C, V_{CE})

Definition

Parameter τ_F, the total forward transit time, is used for modeling the excess charge stored in the transistor when its emitter-base junction is forward biased and $V_{BC} = 0$. It is needed to calculate the transistor's emitter diffusion capacitance.

Typical Value

Typically, τ_F varies with I_C as shown in Fig. 3.17, but for the second-level model τ_F is assumed to be constant (i.e., $\tau_{F_{ac}} = \tau_{F_{dc}} = \tau_F$). The high-current increase in τ_F is modeled in the EM_3 and GP model. Values of τ_F generally range from 0.3 nanosecond for a standard, IC, npn transistor to 80 picoseconds for a high-frequency ($f_T \approx 2$ GHz) device.

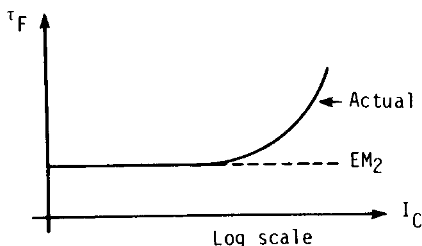


Fig. 3.17. Variation of τ_F with I_C. Dashed line represents the value of τ_F in EM_2 model.

Measurement Scheme

Generally, τ_F is determined from f_T,* the transistor's unity-gain bandwidth, which is defined as the frequency at which the

---

*An interesting technique for measuring τ_F directly (rather than via f_T) has been put forward by Cohen and Zakarevicius.(64) The main problem with this technique, however, seems to be not the measurement but the reduction of the data into τ_F.

common-emitter, zero-load, small-signal current gain extrapolates to unity.(65)* In some computer programs, the user has the option of either entering τ_F directly or f_T (with appropriate operating-point data). In the latter case, the program automatically converts the f_T data to τ_F. Otherwise, the conversion to τ_F must be performed by the user. The determination of τ_F from f_T is described first and is followed by a description of the measurement of f_T.

a) Determination of τ_F from f_T

Parameter f_T also varies with the operating point, as well as from device to device. A typical variation of f_T with $\ln(I_C)$ is sketched in Fig. 3.18. For discrete devices and integrated npn transistors the peak f_T is generally of the order of 600 MHz to 2 GHz. For integrated pnp transistors, the peak f_T is usually 10 MHz for a substrate pnp device and 1 MHz for a lateral pnp device. The drop in f_T at high currents is caused by the increase in τ_F at high currents. The drop in f_T at low currents is caused by junction capacitances C_{jE} and C_{jC}. Since these two capacitances are modeled separately, the drop in f_T at low currents is inherently included in the EM_2 model.

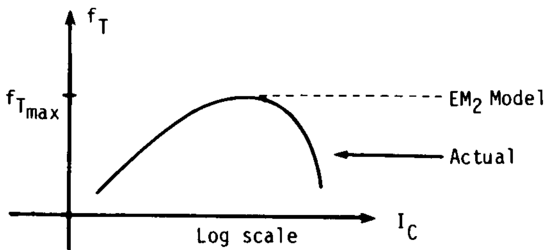


Fig. 3.18. Typical variation of f_T with I_C.

*The definition of f_T normally assumes a single-pole function for β versus f. Although this is a good approximation for the actual variation of β with f, the actual curve may not pass through the point (f_T, 1). Therefore, the word "extrapolate" is used in the definition here.

In the region where f_T is constant, τ_F is given by:

$$\tau_F = \left(\frac{1}{2\pi f_{T_{max}}}\right) - C_{jC} r_c' \qquad (3.19)$$

where $f_{T_{max}}$ is the peak value of f_T and r_c' is the $r_{c_{normal}}'$ value (see r_c' measurement). When there is no constant-f_T region, τ_F is obtained by plotting $\frac{1}{f_T}$ as a function of $\frac{1}{I_C}$, as shown in Fig. 3.19. The resultant curve can then be extrapolated to obtain τ_F. The intercept (noted in Fig. 3.19 by $\frac{1}{f_A}$) of the extrapolated straight line at $\frac{1}{I_C} = 0$ is related to τ_F by:

$$\tau_F = \frac{1}{2\pi}\left(\frac{1}{f_A}\right) - C_{jC}(V_{B'C'})\, r_c' \qquad (3.20)$$

In the programs where τ_F is computed internally from f_T, the computed value of τ_F may be given in the output (e.g., SLIC and SINC). In this case, if τ_F is still required (say for another program), the programs themselves can be used to perform the conversion from f_T to τ_F by making a prior run with a very simple circuit. The value of τ_F obtained from the programs can be very accurate if consistent data (C_j, r_c' etc.) is used since the programs compute V_{BE}, and V_{BC} and take junction capacitance and basewidth modulation effects into account.*

b) Determination of f_T

The unity-gain bandwidth can be measured with a small-signal measurement set-up or an s-parameter measurement system.

(i) Small-signal Measurement Set-up

A simplified version of the small-signal measurement circuit is drawn in Fig. 3.20. Equipment needed is a power supply and

*The complete formula used by the programs SLIC and SINC to find τ_F from f_T data is:

$$\tau_F(0) = \left\{\left(\frac{1}{2\pi f_T}\right) - \left[C_{jE} + C_{jC}\left(1 + \frac{q\, I_C\, r_c'}{kT}\right)\right]\frac{kT}{qI_C}\right\}\left[1 - \frac{V_{BC}}{V_A}\right]^2$$

where V_A is an EM_3 parameter that models basewidth modulation.

$\tau_F(f_T)$

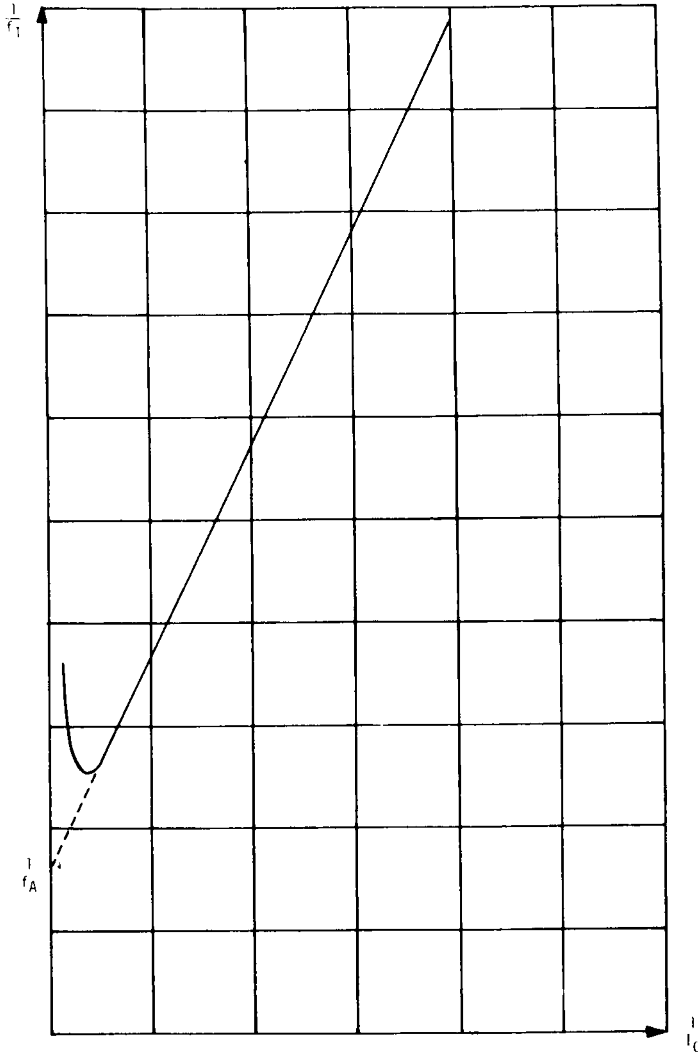


Fig. 3.19. A typical plot of $(1/f_T)$ as a function of $(1/I_C)$.

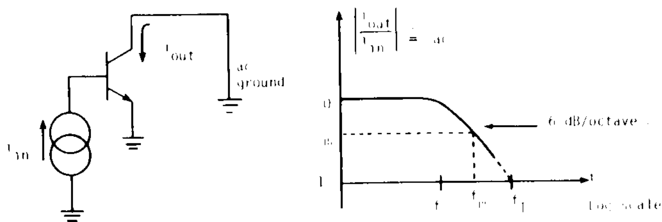


Fig. 3.20 The circuit and frequency variation of the short-circuit current gain used to define f_T.

a small-signal source and detector (such as a vecter voltmeter or an oscillator and an oscilloscope).

At the desired bias point, β_0 and its -3 dB frequency (f_β) are measured. f_T is then the product of these two values:

$$f_T = \beta_0 \times f_\beta \tag{3.21}$$

Alternatively, another beta and frequency value can be measured to determine f_T. For example, at any frequency (f_m) between $3\,f_\beta$ and $\frac{f_T}{3}$, the beta value at that frequency (β_m) is measured. Then:(66)

$$f_T = \beta_m \times f_m \tag{3.22}$$

It is advisable to make measurements over a range of frequencies to verify that f_m lies in the 6 dB/octave region.*

For measurement frequencies greater than ~500 MHz true short-circuit and open-circuit conditions are difficult to realize and s-parameter measurements are preferable to measurements of y or h parameters.

*Some commercial equipment (f_T meters) exist that measure β_m at a specific frequency, f_m. The user must verify, however, that f_m lies in the 6 dB/octave region.

If, in the small-signal set-up the collector ac ground is not perfect (that is, there is a finite ac load resistance, R_{load}) a correction must be applied:

$$f_T = \frac{1}{\left(\frac{1}{f_{T\ meas}}\right) - 2\pi C_{jC} R_{load}} \tag{3.23}$$

Equation (3.23) leads to an alternative f_T measurement. Measurements of the -3 dB frequency of current gain are obtained for different values of R_{load} (obtained by <u>ac</u> coupling in different load resistors). Extrapolation of the $\left(\frac{1}{f_{3\ dB}}\right)$ versus R_{load} straight line that results to the R_{load} = 0 point then gives $\frac{1}{f_\beta}$. Equation (3.21) can then be used to find f_T.[67]

(ii) <u>f_T from s-parameter Data*</u>

The hybrid-π parameter f_T is a high-frequency parameter. The s-parameters are best suited to obtain an accurate value for f_T since they do not require a high frequency short or open circuit but merely a termination into the characteristic impedance Z_0.

All hybrid-π parameters can be obtained from s-parameter data by straight matrix conversion from s-parameters to h-parameters. This method has two disadvantages: (1) a considerable computational effort is required; (2) the accuracy is not readily known. The following procedure to obtain f_T is simple, fast and yields accurate results.

In the hybrid-π model the small-signal common-emitter current gain β_{ac} assumes a -20 dB/decade roll-off down to f_T. This is expressed by

$$\beta_{ac} = \frac{\beta_0}{1 + j\frac{f}{f_\beta}} \tag{3.24}$$

and

$$f_T = \beta_0 \times f_\beta$$

*This section contributed by Dr. W. M. Sansen, University of Leuven, Belgium.

In the common-collector configuration with the base terminal as input, the value of $(\beta_{ac} + 1)$ is given by s_{21}/s_{12} for the normal condition of $h_{re} << 1$. As a consequence, the measurement of $|s_{21}/s_{12}|$ as a function of frequency gives the values of β_0, f_β and thus of f_T. Obviously β_0 can also be obtained by other means. This provides a good verification of the results. The frequency of measurement has to be extended only up to f_β. The common-collector configuration is chosen to avoid the Miller effect. If the s-parameter measurement system does not accept the packaged device for a common-collector measurement, the common-emitter measurement should be used.*

---

*β_{ac} is obtained from common-emitter s-parameters via the formula

$$\beta_{ac} = \frac{-2s_{21}}{(1-s_{11})(1+s_{22}) + s_{12}s_{21}}$$

τ_R (or τ_{SAT})

Definition

Parameter τ_R is the total reverse transit time and is used to model excess charge stored in the transistor when its collector-base junction is forward-biased and $V_{BE} = 0$. It is needed to calculate the transistor's collector diffusion capacitance.

Typical Value

Typical values of τ_R range from 1 to 20 nsec.

Measurement Scheme

If β_R is significantly greater than 1 the value of τ_R could be obtained in a similar manner to τ_F but with the emitter and collector terminals swapped. In most cases, however, β_R is less than or just greater than unity and a different measurement technique must be used.

The simplest method of obtaining τ_R is to compute it from the measured value of τ_{SAT}, the saturation delay time constant.[29] These two parameters are related by:

$$\tau_R = \tau_{SAT}\left(\frac{1-\alpha_F\alpha_R}{\alpha_R}\right) - \left(\frac{\alpha_F}{\alpha_R}\right)\tau_F \tag{3.25}$$

In some programs, this calculation is performed internally, and τ_{SAT} is the parameter that can be specified. Typical values of τ_{SAT} range from 2 to 40 nsec.

Determination of τ_{SAT}

τ_{SAT} determines how long the transistor takes to come out of saturation. It is determined by a simple pulse measurement of the transistor's saturation delay time, t_s (which is defined in Fig. 3.21a). The saturation delay time is given by:[28,55]

$$t_s = \tau_{SAT}\ \ln\left[\frac{I_{BF} + I_{BR}}{\left(\frac{I_{CF}}{\beta_F}\right) + I_{BR}}\right] \tag{3.26}$$

where I_{BF} is the forward base current, I_{BR} is the reverse base current, and I_{CF} is the forward collector current.

The circuit used to measure t_s (and thus τ_{SAT}) is shown in Fig. 3.21b.* Base currents should be measured, as shown, not calculated since V_{BE} varies significantly.

Equipment required for this measurement is:

- a fast pulse generator such as the Tektronix 109,
- two current probes, such as the Tekronix CT-1,
- a fast oscilloscope, such as the Tektronix 500 or 7000 Series with a dual-trace plug-in,
- a 90-nsec, 50-Ω delay line may be needed for pretriggering.

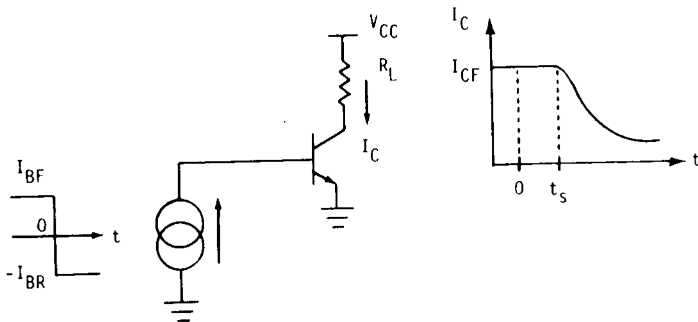
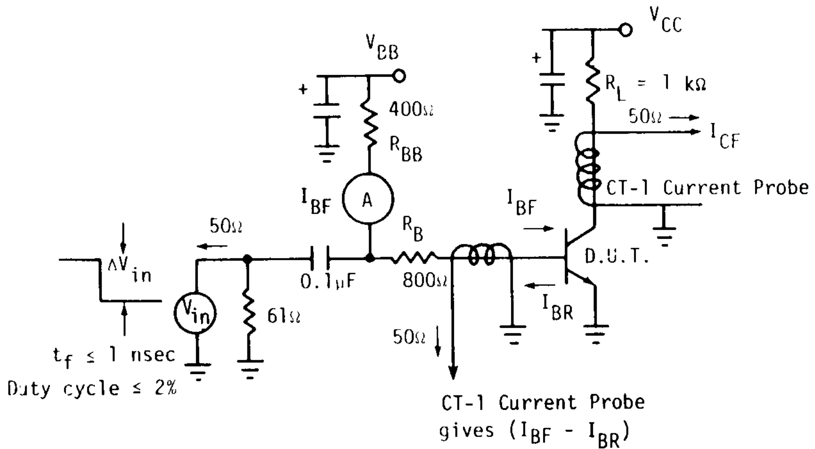
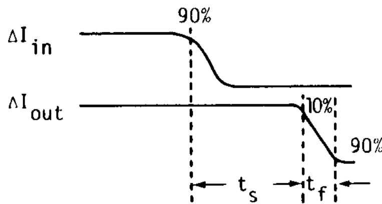


Fig. 3.21a. Circuit and waveforms defining t_s, the saturation delay.

*Circuit from D. Tzu-Tien Hung and F. Severson, Tektronix, Inc.





Design Constraints

V_{CC} sets I_{CF}

V_{BB} sets I_{BF}

ΔV_{in} sets I_{BR}

Fig. 3.21b. Experimental set-up used to measure τ_{SAT}.

C_{SUB} or C_{CS}

Definition

C_{SUB} or C_{CS} is the epitaxial layer-substrate capacitance (and is only present with I.C.'s).

Typical Value

C_{SUB} or C_{CS} is mainly important for integrated npn transistors and lateral pnp transistors. For npn devices, C_{SUB} is represented as a constant capacitance, typically 1 to 2 pF, from the collector terminal to ground. (Note that in SLIC, it is from the internal base node, B', to ground for pnp transistors.) Ideally, C_{SUB} should be modeled by a junction capacitance distributed across r_c' (with its dependence on the epitaxial layer-substrate voltage).

Measurement Scheme

C_{SUB} can be measured directly on a capacitance bridge such as the Boonton Model 75 at the bias voltage to be used in the analysis. If the bias voltage will change drastically, an averaging process should be used. Alternately, a separate reverse-biased diode (with its built-in junction capacitor) or a parasitic transistor* can be added to the circuit description[30] as shown in Fig. 3.22.

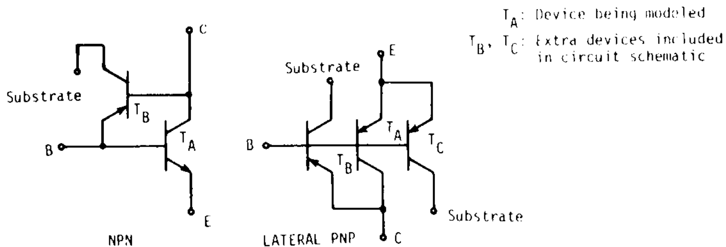


Fig. 3.22. Inclusion of parasitic devices for npn and lateral pnp IC transistors.

*If the parasitic transistor is used, care must be taken with the junction saturation currents. For example, for the npn transistor of Fig. 3.22, the collector-base junction of T_A and the emitter-base junction of T_B are the same junction. It should not be modeled twice.

3.4 EM_3 Model Parameter Measurements

Techniques for measuring the extra parameters for the EM_3 model are described in this section. The five EM_1 model parameters described in Section 3.2 and the twelve model parameters in Section 3.3 are also part of the full EM_3 model.

The additional EM_3 model parameters described here are:

V_A - the Early voltage (which can be computed from g_o, the slope of the I_C versus V_{CE} characteristics in the normal, active region).

β_{FM}, C_2, n_{EL} and θ - the β_F versus I_C model parameters (which can be obtained from β_F versus I_C <u>input</u> parameters β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE}).

β_{RM}, C_4, n_{CL} and θ_R - the β_R versus I_C model parameters.

RATIO - the ratio of the external component of C_{jC} to the internal component in the split C_{jC} - r'_b model.

$\frac{L_E}{W}$, I_{CO} - the ratio of the smallest emitter width to the basewidth and the collector current at which τ_F starts to rise, respectively (which are both used to model a van der Ziel and Agouridis style increase of τ_F with I_C).

TC_1, TC_2 - the first- and second-order temperature coefficients of the temperature-dependent model parameters r'_b, r'_c and β_F.

With the exception of RATIO, all the above parameters can be determined from terminal measurements. The parameter RATIO is determined from a knowledge of the device surface geometry.

Minimal equipment needed for the terminal measurements is:

- a curve tracer

- a small-signal measurement system for determining f_T
- a calibrated temperature chamber

V_A

Definition

V_A is the Early voltage which models the effect on the transistor characteristics of basewidth modulation due to variations in the collector-base space-charge layer. It is always a positive number.

Typical Value

Typical values of V_A are 50 to 70 V for npn transistors, 100 V for lateral pnp transistors and of the order of 1 to 10 V for devices with extremely thin basewidths (such as punch-through devices or super-β transistors).

Measurement Schemes

All the measurement techniques described here assume that the variation of the emitter-base space-charge layer width has a negligible effect on the transistor's characteristics in the normal, active region. The test for this assumption and what to do if it is violated are given at the end of the descriptions of these measurement techniques under the heading "Accuracy of Measurement Techniques."

Three techniques for determining V_A are described here, the simpler ones first. All the measurements should be made at low or medium current levels (i.e., before β_F decreases at high currents or $I_C < I_K$).

(i) Extrapolation of the Output Characteristics in the Common-emitter Configuration

The simplest method of obtaining V_A (though one prone to significant measurement error) is the extrapolation of the I_C versus V_{CE} characteristics when the transistor is in the common-emitter configuration. As shown in Fig. 3.23, the Early voltage is approximately the value of the voltage on the $-V_{CE}$ axis where the extrapolated output characteristics meet. This extrapolation can be performed graphically from either a photograph of the characteristics or directly from the curve tracer display. However, the experimental error associated with this approach can be very high.

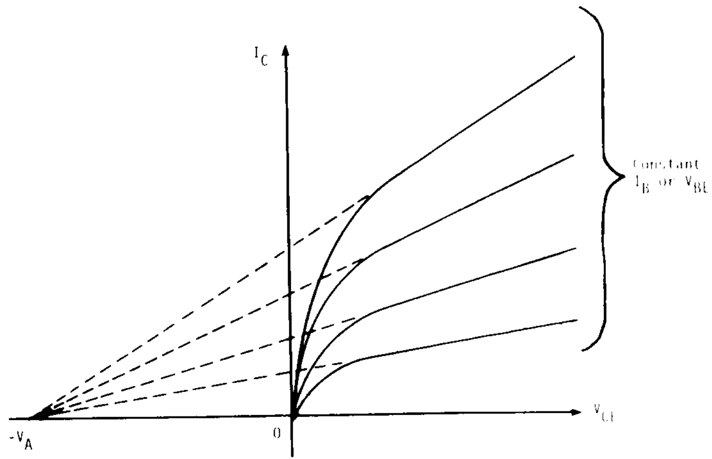


Fig. 3.23. Approximate definition of V_A from the I_C-versus V_{CE} characteristics (not to scale).

(ii) Slope of the Output Characteristics, g_o

A geometrically equivalent technique for determining V_A to that above is illustrated in Fig. 3.24. The slope of the output characteristics, g_o, is measured at the point $V_{CE} = V_{BE}$ ($\approx$ 0.6 V for silicon) and any corresponding value of I_C as long as the point on the curve is in the normal, active, linear region (and not in the saturated region). V_{BE} can be either estimated or obtained by using a voltage drive on the base. The slope g_o equals r_o^{-1}. Since the point chosen corresponds to $V_{BC} = 0$, the collector current at which the slope is measured is called $I_C(0)$. The Early voltage is then given by

$$V_A = \frac{I_C(0)}{g_o} = r_o \, I_C(0) \tag{3.27}$$

V_A

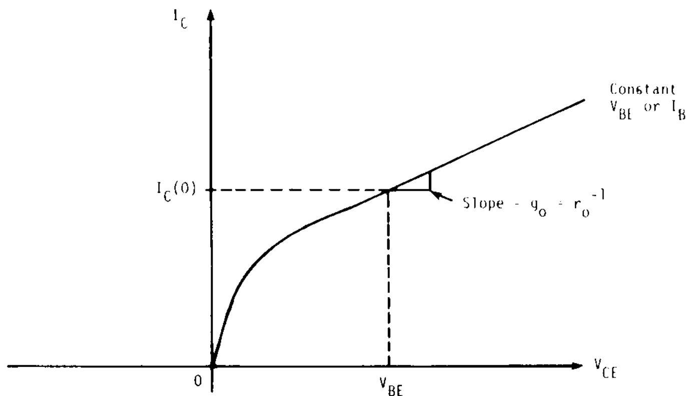


Fig. 3.24. The determination of r_o from the I_C versus V_{CE} characteristics (not to scale).

The accuracy of this technique for measuring V_A is not very high because of the large error that can result in the measurement of the slope, g_o.

In some programs (such as SLIC), r_o and $I_C(0)$ are acceptable input parameters and V_A is automatically computed from their product. It is also possible in SLIC to use the r_o data to compute I_S. This can be done if:

- a voltage drive, V_{BE}, is used on the base,
- both V_{BE} and V_{CE} are specified as part of the r_o data (as well as I_C)
- I_S is not specified separately.

The program then automatically calculates I_S from the formula(31)

$$I_S = \frac{I_C \, e^{-\frac{qV_{BE}}{kT}}}{\left[1 + \frac{V_{BE} - V_{CE}}{r_o \, I_C}\right]} \tag{3.28}$$

where I_C, V_{BE} and V_{CE} are the values fed in with r_o. The accuracy of this technique for determining I_S, however, is not high because the measurement is made at only one point.

(iii) $\ln(I_C)$ Versus V_{BE} Curves

The third technique for finding V_A is the determination of the $\ln(I_C)$ versus V_{BE} characteristic at two different values of V_{BC}. As shown in Fig. 3.25, this theoretically results in two parallel lines. The Early voltage is then determined from the ratio of the (extrapolated) I_S values or the ratio of the value of two I_C values at the same V_{BE}.

$$\frac{I_S(V_{BC1})}{I_S(V_{BC2})} = \left.\frac{I_C(V_{BC1})}{I_C(V_{BC2})}\right|_{\substack{\text{Same}\\ V_{BE}}} = \frac{\left(1 + \frac{V_{BC2}}{V_A}\right)}{\left(1 + \frac{V_{BC1}}{V_A}\right)} . \tag{3.29}$$

Equation (3.29), which is used to find V_A, can be simplified if one of the V_{BC} values is zero. To simplify or eliminate the need for correcting for finite r_b', r_e' and r_c', the curves should be measured at currents as low as possible and with V_{BC} as close to zero as practical.

(iv) Accuracy of Measurement Techniques

The third technique is more accurate than the previous two since it need not involve extrapolation or the measurement of a slope. If the $\ln(I_C)$ versus V_{BE} graph is obtained from a curve tracer, this method only requires two points on the graph to be measured and is effectively equivalent to the previous methods. If the graph is obtained by a setup which cannot make the measurement of two I_C values at the same V_{BE} (such as a setup which sets I_C and measures V_{BE}), several points are necessary for the graph to be drawn. In this case the accuracy of this technique is greater than the previous two techniques.

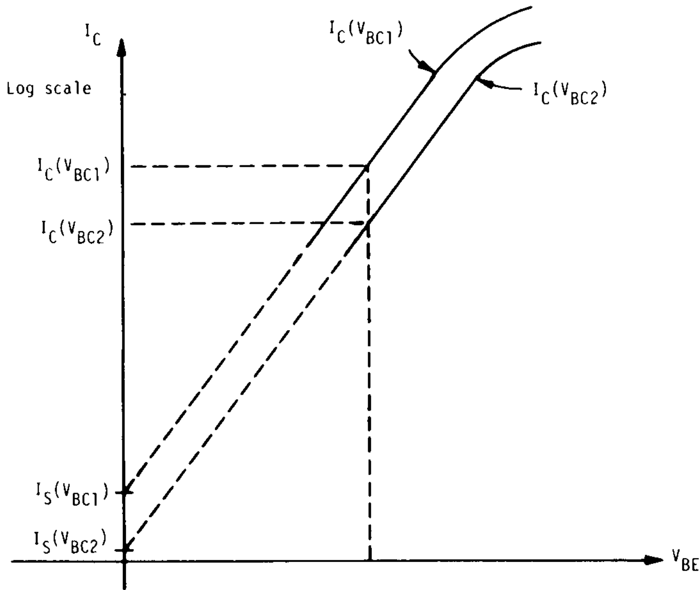


Fig. 3.25. Plots of $\ln(I_C)$ versus V_{BE} at two different values of V_{BC}, illustrating the third technique for determining V_A.

When V_A is large, all of the above techniques for finding V_A become very inaccurate. However, this is acceptable because the larger V_A is, the less important the effect of basewidth modulation becomes and therefore the need to accurately determine V_A diminishes.

As is seen from the GP analysis, V_A is not a true constant. Variations can be seen, especially for devices with low values of V_A. For these cases, an appropriate average should be used (for the slopes over a range of V_{CE} values in the first two measurement schemes and for V_A directly in the third measurement). Also note that a variation in the slope at high values of V_{CE} can be caused by the breakdown phenomenon, which is not included in this model and therefore its influence on the slope should not be included.

All of the above techniques assume that basewidth modulation due to variation in the width of the emitter-base space charge layer is negligible -- normally a valid assumption. To test for this, simply repeat the above measurements at different V_{BE} values. That is, for the first two measurement schemes, determine V_A for at least two constant-V_{BE} curves. For the third measurement scheme, determine V_A from the same $I_C(V_{BC1})$ and $I_C(V_{BC2})$ curves but at a different value of V_{BE}.* If the change in V_{BE} significantly affects the value of V_A obtained from any of these schemes, the above assumption is invalid. In this case, refer to the V_B-measurement scheme (p. 201) for the appropriate method of determining V_A.

*This test is equivalent to testing if the slope of the $\ln(I_C)$ versus $\frac{qV_{BE}}{kT}$ curve at low and medium current levels differs significantly from unity.

β_{FM}, C_2, n_{EL} and θ (or β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE})

Definition

The first four parameters define the variation of β_F with I_C. They are defined in terms of the plots of ln (I_C, I_B) versus $V_{B'E'}$ for $V_{B'C'} = 0$ (Fig. 3.26).

β_{FM} defines the magnitude of the ideal component of I_B at a given value of V_{BE}:

$$I_{B(ideal)}\left[= \frac{I_S(0)}{\beta_{FM}}\left(e^{\frac{qV_{BE}}{kT}} - 1\right)\right] \qquad (3.30)$$

C_2 (magnitude) and n_{EL} (variation with V_{BE}) describe the non-ideal component of I_B which is dominant at low currents:

$$I_{B(non\text{-}ideal)} = C_2 I_S(0)\left[\left(e^{\frac{qV_{BE}}{n_{EL}kT}} - 1\right)\right] \qquad (3.31)$$

This component of I_B is responsible for the drop in β_F at low currents.

The fourth parameter, θ, models the effect on I_C (and therefore the drop in β_F) at high currents due to high-level injection. It describes the high-current asymptote of the ln(I_C) versus V_{BE} graph as

$$I_{C(high\ level)} = \frac{I_S(0)}{\theta}\, e^{\frac{qV_{BE}}{2kT}} \qquad (3.32)$$

Typical Values

Typical values of β_{FM} range from the order of 10 for an I.C. lateral pnp or a power transistor, through about 100 for a small-signal npn transistor, to the order of 1000 for a super-β npn transistor.

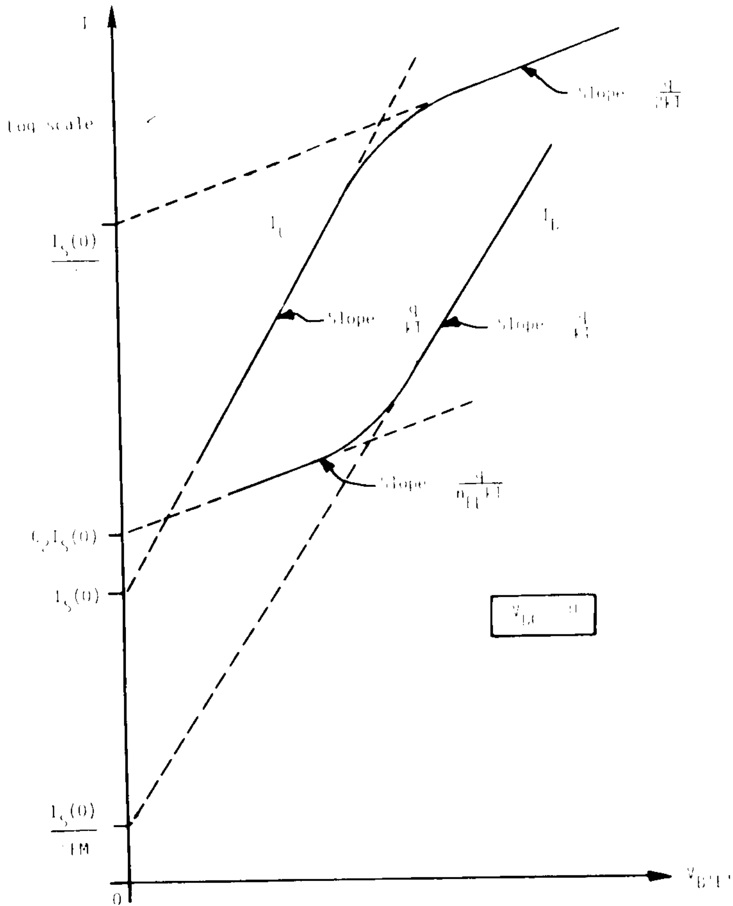


Fig. 3.26. Plots of $\ln(I_C)$ and $\ln(I_B)$ versus $V_{B'E'}$ (for $V_{BC} = 0$) illustrating the definitions of β_{FM}, C_2, n_{EL} and I_K (not to scale).

$\beta_F(I_C)$

The value of C_2 is typically on the order of 100 to 1000.

n_{EL} is typically between 2 and 4 (most often 2).

θ is normally about 10^{-7} to 10^{-6}.

Measurement Scheme

Two methods for determining these model parameters are described. The first method obtains their values directly. The second method takes advantage of a simpler user-oriented set of input parameters. The input parameters for the latter method (which are used in the programs SLIC and SINC) are used to indirectly determine the above model parameters.

(i) Direct Measurement Technique

The values of β_{FM}, C_2, n_{EL} and θ can be very simply obtained, directly from a plot of ln (I_C, I_B) versus $V_{B'E'}$, as shown in Fig. 3.26. There are two stages involved in the generation of Fig. 3.26: measurement of $\ln(I_C)$ and $\ln(I_B)$ as a function of V_{BE} and the reduction of V_{BE} to $V_{B'E'}$. These stages are described here.

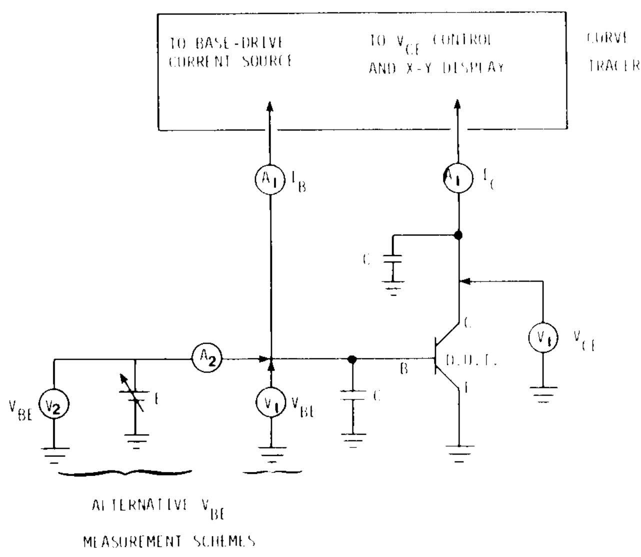
Measurement of $\ln(I_C)$ and $\ln(I_B)$ Versus V_{BE}*

The $\ln(I)$ versus V_{BE} curves can either be obtained from a curve tracer by determining I_C at a given V_{BE} and β_F versus I_C, from a specially-designed test set-up or from a combination of these such as the curve tracer set-up shown in Fig. 3.27a. The curve tracer is operated in the dc mode with external dc voltmeters and ammeters added to extend the accuracy of the measurements when necessary. The need for static operation arises from the effects of stray and junction capacitances which severely distort the I_C versus V_{CE} and I_C versus V_{BE} displays at low current levels.

---

* This measurement section contributed by Dr. J. Smith, Tektronix, Inc.

The curve tracer's continuously-variable, base-drive current source can provide precision dc currents as low as 5 nA and is used to precisely establish a dc base current. The V_{CE} control provides the resolution necessary for setting $V_{BC} = V_{BE} - V_{CE} = 0$. The CRT displays of I_C versus V_{CE} and I_C versus V_{BE}, when used in conjunction with external voltmeters and ammeters, permits monitoring of the voltages and currents with precision instruments while at the same time monitoring the analog display of the operating point for such effects as noise, pick-up and the influence of the external meters.



A_1: PRECISION AMMETER
A_2: NULL-INDICATING AMMETER
V_1: HIGH RESOLUTION, HIGH INPUT IMPEDANCE VOLTMETER
V_2: HIGH RESOLUTION VOLTMETER (CAN BE LOW INPUT IMPEDANCE)
E: HIGH RESOLUTION VARIABLE VOLTAGE SOURCE
C: LOW LEAKAGE BY-PASS CAPACITOR

Fig. 3.27a. Test set-up for static measurement of I_C, I_B, and V_{BE} = 0 at low current levels.

$\beta_F(I_C)$

The measurement of V_{BE} requires high resolution over the range of 0.4 to 1 V.* As shown in the test set-up figure, the measurement of V_{BE} can either be made with a high-input-impedance voltmeter (such as a Keithley Model 616 Electrometer) or with a low-leakage, null-indicating ammeter (such as the Hewlett-Packard Mo el 419A DC Micro Volt-Ammeter) in conjunction with a high-resolution voltage source and a low-input-impedance (10 MΩ) voltmeter (such as the Tektronix DM501 Digital Multimeter). The same instruments can be used to measure both V_{BE} and V_{CE}.

The external precision ammeter (for example, the Keithley or Hewlett-Packard instruments mentioned above) can be used to measure I_B and I_C when greater accuracy than that offered by the curve tracer is desired.**

Any error in operating point, noise or pick-up which may be present in the test set-up is displayed on the CRT. The presentation of a single, clean spot indicates a noise-free, stable operating point and permits a high degree of confidence that good measurements were obtained. Noise and pick-up can usually be removed by careful cabling and the use of the low-leakage bypass capacitors.*** Care must be taken to insure that large-valued bypass capacitors have charged to their final states before measurements are recorded.

The set-up for the Tektronix type 576 curve tracer is shown in Fig. 3.27b. The vertical channel, I_C, has a maximum sensitivity of

---

* High resolution is required in the measurement of V_{BE} since at T = 300°K an error of only 1 mV in V_{BE} corresponds to an error of approximately 4% in the base currer

---

** The accuracy of the base-drive current source should prove adequate for most measurements of I_B. However, an external meter may sometimes be needed for I_C.

*** In some cases it may be necessary to employ a constant voltage transformer (such as the Sola Electric type CVS) to remove the effects of power line fluctuations and/or an isolation transformer (such as the Topaz Electronics Ultra-Isolation transformer) to remove power line noise.

TEST SET-UP CHART TYPE 576

LOCUS OF I_C VS V_{BE}

OPERATING POINT

LOCUS OF I_C VS V_{CE}

V_{BE}

V_{CE}

Fig. 3.27b. Test set-up for low-current, static measurement of I_C, I_B, V_{CE}, and V_{BE}. The horizontal channel can be used to measure both V_{BE} and V_{CE}. The external meters offer increased accuracy and are optional. For a pnp, change POLARITY and press DISPLAY INVERT.

$\beta_F(I_C)$

1 μA per division except in the leakage (emitter current) mode of operation which offers a maximum sensitivity of 1 nA per division.*

The Tektronix type 577 curve tracer offers increased sensitivity of 0.2 nA per division in the vertical channel with improved noise performance. The set-up for the type 577 curve tracer is shown in Fig. 3.27c.

With either the 576 or 577 curve tracer the sequence of steps for making the measurement is as follows:

1. Base drive:
   - -Select SINGLE-STEP, OFFSET-MULTIPLIER operation.
   - -Using the STEP/OFFSET-AMPLITUDE selector and the OFFSET-MULTIPLIER adjust the base current to the desired dc value.
2. Collector drive:
   - -Select the DC COLLECTOR SUPPLY MODE of operation.
   - -Select COLLECTOR VOLTS with the HORIZONTAL VOLTS/DIVISION selector
   - -Rotate the VARIABLE COLLECTOR SUPPLY control in order to generate the I_C versus V_{CE} locus shown in Figs. 3.27b or 3.27c. Set V_{CE} to the approximate expected value of V_{BE}.
3. Switch the HORIZONTAL VOLTS/DIVISION selector alternately between COLLECTOR VOLTS and BASE VOLTS. Adjust the VARIABLE COLLECTOR SUPPLY such that the operating point, as indicated by the single spot on the CRT display, does not move when the HORIZONTAL VOLTS/ DIVISION selector is switched from COLLECTOR VOLTS to BASE VOLTS. A stationary spot indicates that $V_{CE} = V_{BE}$ which establishes the desired condition, $V_{BC} = 0$. Note that the same scale factor must be used for both the COLLECTOR VOLTS and BASE VOLTS.

If external meters are not used, I_C, I_B and V_{BE} are measured as follows: 1) Collector current, I_C, is read from the vertical axis of the CRT display. 2) Base current, I_B, is read from the STEP/OFFSET AMPLITUDE selector and the OFFSET MULTIPLIER and 3) Base-emitter voltage, V_{BE}, is read from the horizontal axis of the CRT display.

---

* Use of the leakage (emitter current) mode requires the application of correction factors which are explained in the type 576 curve tracer instruction manual (p. 2-18).

TEST SET-UP CHART 577-177-D1

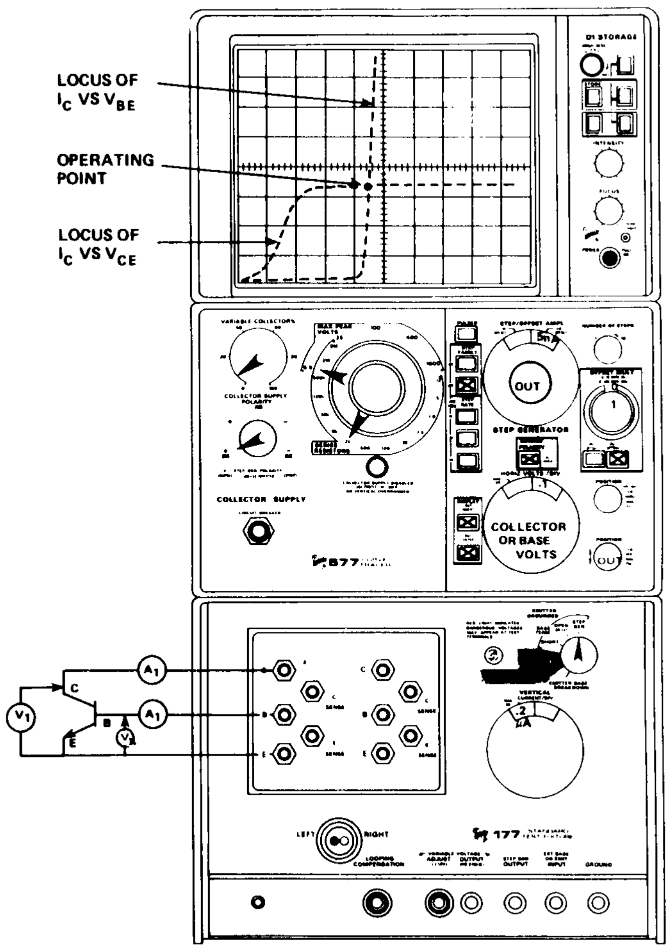


Fig. 3.27c. Test set-up for low-current, static measurement of I_C, I_B, V_{CE}, and V_{BE}. The horizontal channel can be used to measure both V_{BE} and V_{CE}. The external meters offer increased accuracy and are optional. For a pnp, change POLARITY to –DC and put display in INVERT mode.

When BASE VOLTS is selected with the HORIZONTAL VOLTS/DIVISION selector the I_C versus V_{BE} locus shown in Figs. 3.27b or 3.27c can be generated by sweeping the base current I_B with the OFFSET MULTIPLIER potentiometer.

The generation of the I_C versus V_{CE} and I_C versus V_{BE} loci is facilitated by the use of the storage display option of the model 577 curve tracer.

The more familiar pulsed-mode of operation for the curve tracer can be used in the higher-current range and is perhaps desirable there in order to avoid self-heating effects.*

Reduction of V_{BE} to $V_{B'E'}$

Corrections must be made to the experimental curves to allow for the voltage drop across the ohmic resistances r_e' and r_b'.** A typical correction is shown in Fig. 3.28.

The correction is made as follows:

(1) Extrapolate the ideal I_B graph as a straight line with a slope of $\frac{q}{kT}$ to the high-current region. Assume this extrapolated line (shown dotted in Fig. 3.28) is the plot of $\ln(I_B)$ versus $V_{B'E'}$.

(2) Assume that the horizontal distance from the extrapolated ideal I_B line to the measured I_B curve is due to $I_B r_b' + I_E r_e'$. Then, subtract this amount from the I_C graph. To accomplish this, draw the two vertical lines marked "c" and "d" on Fig. 3.28. From the intersection of the I_C graph and line c (point "e"), draw a horizontal line until it intersects line d. This intersection (point "f"), is a point on the $\ln(I_C)$ versus $V_{B'E'}$ graph.

---

* In the pulsed mode, measurement of V_{BE} with an external voltmeter will produce erroneous results.

** A correction should also be made for r_c' and r_b' in selecting V_{BC} such that $V_{B'C'} = 0$. For reasonable values of V_A, this correction is normally negligible

Fig. 3.28. The correction applied to the experimental graphs of $\ln(I_C, I_B)$ versus V_{BE} to obtain Fig. 3.26.

(3) Repeat this procedure at several points until the full $\ln(I_C)$ versus $V_{B'E'}$ graph can be drawn (shown as the dotted I_C line in Fig. 3.28).

Sometimes this correction for $(r_b' I_B + r_e' I_E)$ can give a "corrected" I_C which lies to the left of the ideal I_C curve over some range. This is probably caused (in modern, planar structures) by the combination of the effects of the crowding phenomenon[18] and lateral injection from the emitter sidewall. One technique for handling the situation is to force the corrected I_C curve always to lie either on or to the right of the ideal line. This can be done by back-correcting by ΔV, the amount the "corrected" curve lies to the left of the ideal line. After ΔV reaches a maximum value, ΔV_m, the back-correction equals ΔV_m. The resultant error in the I_B graph is then ignored.

$\beta_F(I_C)$

It is not always possible to obtain β_{FM}, C_2, n_{EL} and θ directly from a plot of β_F versus I_C because region II, the constant-β_F region, is not always present. Although β_{FM} cannot be found for this case from β_F versus I_C, it can be found with the technique described above.

(ii) Indirect Measurement Technique

This technique makes use of simpler input parameters to determine β_{FM}, C_2, n_{EL} and θ. The simpler input parameters (as used in SLIC and SINC) are:

β_{FMAX} - the maximum value of β_F

I_{CMAX} - the collector current at which β_{FMAX} occurs (may be called I_{CM})

β_{FLOW} - any value of β_F at a current less than I_{CMAX}

I_{CLOW} - the collector current at which β_{FLOW} occurs (may be called I_{CL})

BCEC - $1/n_{EL}$ or 1 minus the slope of the $\ln(\beta_F)$ versus $\ln(I_C)$ curve at low currents, as show in Fig. 3.29. In SINC and some versions of SLIC, BCEC is fixed internally at 0.5.

V_{CE} - the value of the collector-emitter voltage at which all the above four parameters were measured.

Several points should be made about these input parameters:

(1) All parameters can be obtained directly from a plot of β_F versus I_C at constant V_{CE} as illustrated in Fig. 3.29.

(2) The β_F versus I_C curve can be obtained from a curve tracer, as described for the EM_1 model.

(3) The above β_F curve refers to dc values of β not ac values.

(4) Typical values of (β_{FMAX}, I_{CMAX}) are of the order of (100, 1 mA) for discrete and integrated npn transistors, (30, 100 μA) for an integrated substrate pnp transistor and (20, 30 μA) for an integrated lateral pnp transistor. Since (β_{FLOW}, I_{CLOW}) is any point below (β_{FMAX}, I_{CMAX}) there can be no typical value.

(5) The conversion from the input parameters (β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE}) to the model parameters (β_{FM}, C_2, n_{EL} and θ) is automatically performed by the programs.

(6) If a constant value of β_F is used (as in the EM_1 and EM_2 models), β_{FMAX} is the constant value.

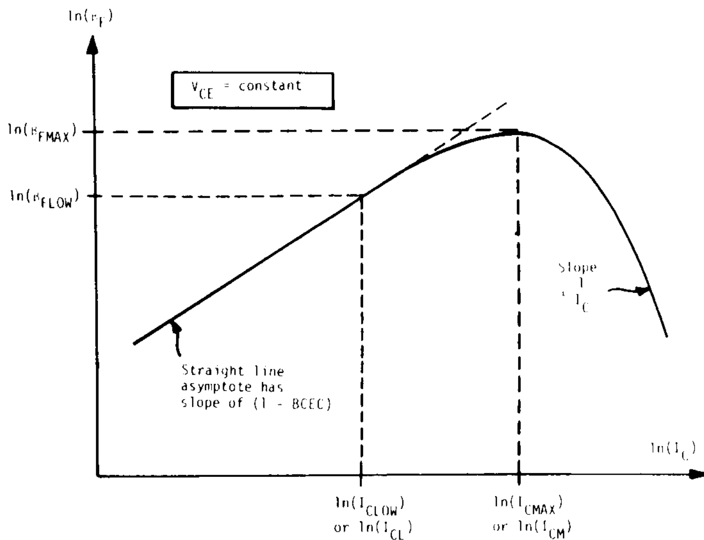


Fig. 3.29. A plot of ln(β_F) versus ln(I_C) which illustrates the definitions of β_{FMAX}, I_{CMAX}, β_{FLOW}, I_{CLOW}, BCEC and V_{CE}.

β_{RM}, C_4, n_{CL} and θ_R

Definition

These four parameters define the variation of β_R with I_E. They are analagous to β_{FM}, C_2, n_{EL} and θ, respectively, with V_{BE} replaced by V_{BC}, I_C replaced by I_E, V_{BC} replaced by V_{BE} and β_F replaced by β_R.

Typical Values

Typical values are the same as for the β_F parameters except that β_{RM} is normally about 0.1 to 10.

Measurement Scheme

These β_R versus I_E parameters are obtained by the same direct method (Method (a)) as is used with the β_F versus I_C parameters, except that the emitter and collector terminals are interchanged.

RATIO

Definition

RATIO models the split of the collector-base junction capacitor C_{jC} across the base resistor r_b'. The value of the capacitance to the outside of r_b' (that is to node B) is RATIO x C_{jC}. The value of the capacitance to the inside of r_b' (to node B') is (1-RATIO) C_{jC}.

Typical Value

RATIO must lie between 0 and 1. It is typically on the order of 0.8.

Measurement Scheme

RATIO is a difficult parameter to determine from terminal measurements.* It can be found with relative ease if the geometry of the device is known, since it is equal to $\left(1 - \frac{A_E}{A_B}\right)$ where A_E is the area of the emitter and A_B is the total area of the base, including the emitter area. For a discrete device it may be necessary to remove the encapsulant to determine the geometry of the device -- a destructive process.

---

*It may be possible to obtain RATIO via terminal measurements by measuring the product $r_b'(1\text{-RATIO})C_{jC}$. This product, given the small-signal EM_3 model of Fig. A5.1, is simply $|h_{rb}| / \omega_{meas}$ where $|h_{rb}|$ is the magnitude of the common-base value of h_r (which is determinable from s-parameter measurements, a network analyzer or a special test jig) and ω_{meas} is the measurement frequency[(72)] (which must be in the frequency range below f_T over which $|h_{rb}|$ is rising at 20 dB/decade for the effects of basewidth modulation to be negligible). Note, however, that $|h_{rb}|$ is normally <<1, that this technique requires an accurate small-signal value of r_b' and that the effect of an external base resistance, which may be significant, is neglected.

$\frac{L_E}{W}$, I_{CO}

Definition

$\frac{L_E}{W}$ and I_{CO} model the rise of τ_F with I_C, assuming a van der Ziel and Agouridis mechanism.[40] L_E is the smallest emitter width, W is the basewidth and I_{CO} is the current at which τ_F starts to rise.

Typical Value

Typical values of L_E and W are 10 μ (0.4 mil) and 0.5 μ, respectively. Therefore $\frac{L_E}{W}$ is typically on the order of 20. I_{CO} is typically on the order of 0.1 mA to 100 mA.

Measurement Scheme

A plot of $\tau_{F_{ac}}$ versus I_C is obtained by applying the (small-signal) τ_F-measurement techniques outlined in Section 3.3 at several collector currents. The rise in $\tau_{F_{ac}}$ at high currents is fitted to the formula[41]

$$\tau_{F_{ac}}(I_C) = \tau_{FL}(0)\left[1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2\left(\frac{I_C}{I_{CO}} - 1\right)^2\right] \text{for } I_C \geq I_{CO} \quad (3.33)$$

Fig. 3.30 shows the variation of τ_F with I_C that results from this expression.

Since this formula does not necessarily apply for all devices, it may not be possible to fit this expression for every graph of $\tau_{F_{ac}}$ versus I_C. Two points on the curve could be taken and the two simultaneous equations solved for $\frac{L_E}{W}$ and I_{CO}. Alternately, a curve-fitting computer algorithm could be used.

$\frac{L_E}{W}$, I_{CO}

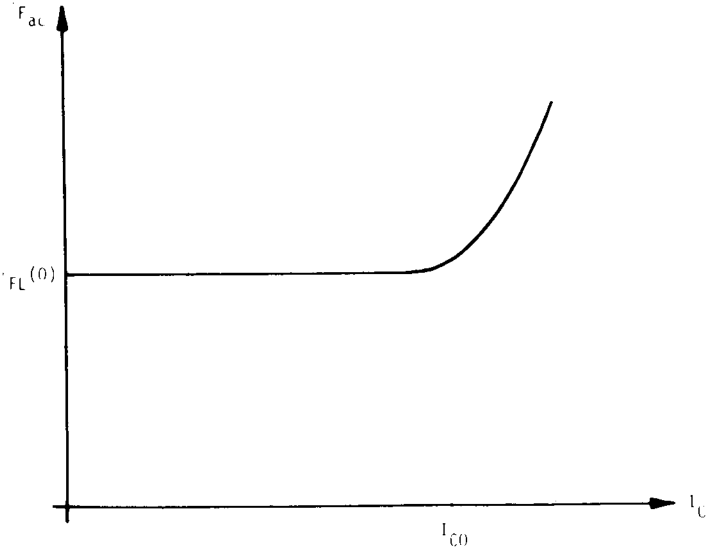


Fig. 3.3Q. The shape of the variation of τ_F with I_C that is modeled by the van der Ziel and Agouridis expression of Eq. (3.33).

TC_1, TC_2

Definition

TC_1 and TC_2 are first- and second-order temperature coefficients. A set of each is used for r_b', r_c' and β_F. The coefficients are defined by:

$$Par(T) = Par(T_{nom})\left[1 + TC_1(T-T_{nom}) + TC_2(T-T_{nom})^2\right] \qquad (3.34)$$

where Par is the parameter being considered, T_{nom} is the nominal temperature at which the parameter was measured and T is the temperature at which the analysis is to be performed.

Typical Values[1,31]

Values of TC_1 and TC_2 vary from one parameter to the other. Typical values of TC_1 and TC_2 for β_F are 6.67×10^{-3} $(°K)^{-1}$ and -3.6×10^{-6} $(°K)^{-2}$, respectively. For r_c' and r_b' typical values are 2×10^{-3} $(°K)^{-1}$ and 8×10^{-6} $(°K)^{-2}$, respectively.

Measurement Scheme

The temperature variation of each parameter is obtained by placing the device in a controlled-temperature environment (for example, an oven) and making the appropriate measurements.* The observed temperature variation is then fitted to the above expression. TC_1 represents a linear variation with temperature and TC_2 represents a nonlinear (square law) variation with temperature.

---

*Care must be taken in making these measurements that the power dissipated by the device does not raise the junction temperature significantly above ambient temperature. As well, for measurements of r_b' and r_c' the resistance of the leads between the device and the measurement equipment needs to be minimized and/or removed from the result. For r_b' the inductance of the leads may also need to be minimized and/or removed from the result.

3.5 GP Model Parameter Measurement

With the EM_2 model as the starting point, the GP requires the following extra model parameters:

I_{SS}	- the saturation current which <u>replaces</u> I_S.
V_A	- the Early voltage
V_B	- the inverse Early voltage
β_{FM}, C_2, n_{EL} and I_K	- the β_F versus I_C model parameters
β_{RM}, C_4, n_{CL} and I_{KR}	- the β_R versus I_E model parameters
B	- the base push-out factor, which itself requires model parameters to describe it.

Of these model parameters, measurements have been described in the EM_3 model (Section 3.4) for all the model parameters except:

- I_{SS}
- V_B
- I_K
- I_{KR}
- B

The determination of these five model parameters from terminal measurements is described in this section. Minimal equipment needed is:

- a curve tracer
- a small-signal measurement system for determining f_T

I_{SS}

Definition

I_{SS} is defined in the GP model from considerations of the internal physics of the device (for an npn transistor) as:

$$I_{SS} \triangleq \frac{q\, D_n\, n_i^2\, A}{\int_{x'_{EO}}^{x'_{CO}} p_o(x)dx} \tag{3.35}$$

In terms of terminal measurements, I_{SS} is equal to $I_S(0)$, the value of I_S obtained from extrapolation of the I_C vs V_{BE} to $V_{BE} = 0$ (and with $V_{BC} = 0$).

Typical Value

Since $I_{SS} = I_S(0)$, it has the same typical value: of the order of 10^{-16} A for an integrated circuit transistor. As seen from Eq. (3.35), I_{SS} is proportional to the emitter area, A.

Measurement Scheme

I_{SS}, like I_S, is obtained from a plot of $\ln(I_C)$ as a function of $\frac{qV_{BE}}{kT}$ with $V_{BC} = 0$.* The value of I_{SS} is obtained by extrapolation of the straight line at low currents to $V_{BE} = 0$. Alternately, I_{SS} could be obtained from several points on the graph by fitting these points to the expression

$$I_C = \frac{I_{SS}}{\left(1 + \frac{V_{BE}}{V_B}\right)}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) \tag{3.36}$$

This may be necessary when V_B is small (<~5 V). Since these measurements are performed at low currents, the correction to obtain $V_{B'E'}$ from V_{BE} should not be necessary.

---

* See page 190 for measurement technique.

V_B

Definition

V_B is the inverse Early voltage which models the effect of basewidth modulation due to emitter-base space-charge layer width variations. It is always a positive number.

Typical Value

V_B is typically on the order of 10 to 50 V.

Measurement Scheme

There is a complicating factor in measuring V_B in an analogous way to that used for V_A (Section 3-4). The measurement of V_A assumed that the variation of the width of the emitter-base space-charge layer had a negligible effect on the transistor's characteristics in the normal, active region. That is, that $\left(\frac{V_{BE}}{V_B}\right)$ is much less than unity. The equivalent assumption that would be necessary if the V_A-measurement schemes are to be used directly for V_B is that $\left(\frac{V_{BC}}{V_A}\right)$ is much less than unity. Two measurement techniques for measuring V_B are given here, both assuming that $\left(\frac{V_{BC}}{V_A}\right)$ is not negligible compared with unity and therefore both also giving a value for V_A.

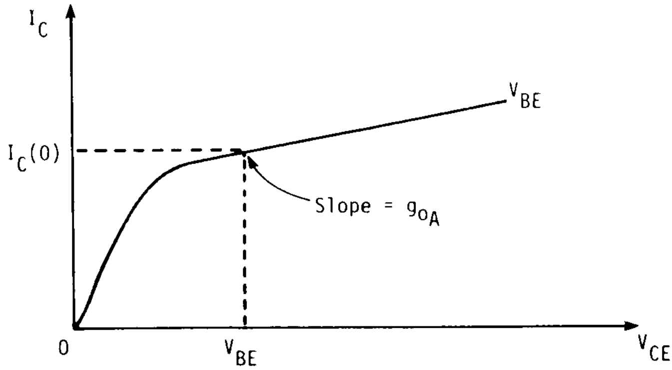
(i) Measurement from output characteristics

When the transistor is operated in the normal, active region in the common-emitter mode (with V_{BE} kept constant) its output characteristics are as sketched in Fig. 3.31a. The slope of the characteristics at $V_{CE} = V_{BE}$ (i.e., selected so that $V_{BC} = 0$) is designated g_{o_A} and is given by:*

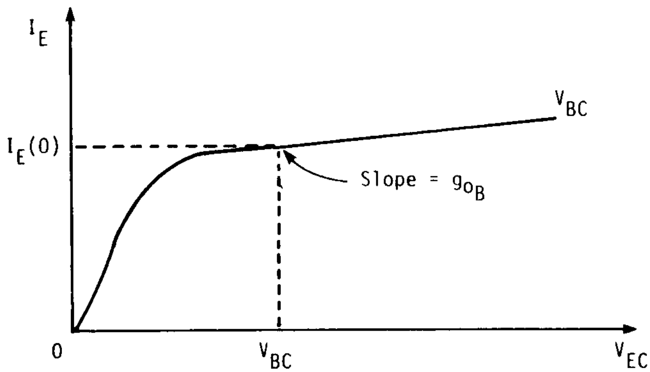
$$g_{o_A} \triangleq \left.\frac{dI_C}{dV_{CE}}\right|_{\substack{V_{BE}\ \text{constant} \\ V_{BC} = 0}}$$

*Note that the measurements must be made at a point where the device is not saturated and before β_F starts decreasing due to high-current effects. These conditions are implied in the expression for I_C in Eq. (3.37).

V_B



(a)



(b)

Fig. 3.31. The output characteristics in the normal, active region, (a), and the inverse region, (b), used to define g_{oA} and g_{oB} and to obtain V_A and V_B. Both curves are generated under low-level injection conditions.

V_B

$$\simeq - \frac{d}{dV_{BC}} \left[\frac{I_{SS}\left(e^{\frac{qV_{BE}}{kT}} - 1\right)}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)} \right]_{\substack{V_{BE} = \text{constant} \\ V_{BC} = 0}}$$

$$= \frac{I_C(0)}{V_A\left(1 + \frac{V_{BE}}{V_B}\right)} \tag{3.37}$$

(Note that when $V_{BE} << V_B$, Eq. (3.37) reduces to that used for determining V_A in Section 3.4.)

Similar output characteristics for the transistor in the inverse region (obtained by swapping emitter and collector leads) are given in Fig. 3.31b. The slope of these characteristics at $V_{EC} = V_{BC}$ (i.e., selected sc that $V_{BE} = 0$) is designated g_{o_B} and is given by:

$$g_{o_B} \triangleq \left. \frac{dI_E}{dV_{EC}} \right|_{V_{BC} \text{ constant}}$$

$$\simeq - \frac{d}{dV_{BE}} \left[\frac{I_{SS}\left(e^{\frac{qV_{BC}}{kT}} - 1\right)}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)} \right]_{\substack{V_{BC} \text{ constant} \\ V_{BE} = 0}}$$

$$= \frac{I_E(0)}{V_B\left(1 + \frac{V_{BC}}{V_A}\right)} \tag{3.38}$$

Equations (3.37) and (3.38) can be solved to find V_A and V_B;

V_B

$$
\left.
\begin{aligned}
V_A &= \frac{I_C(0)\; I_E(0) - g_{o_A}\, g_{o_B}\, V_{BE}\, V_{BC}}{g_{o_A}\, I_E(0) + g_{o_A}\, g_{o_B}\, V_{BE}} \\
V_B &= \frac{I_C(0)\; I_E(0) - g_{o_A}\, g_{o_B}\, V_{BE}\, V_{BC}}{g_{o_B}\, I_C(0) + g_{o_A}\, g_{o_B}\, V_{BC}}
\end{aligned}
\right\} \tag{3.39}
$$

(ii) <u>Measurement from ln(I) versus V</u>

Figure 3.32 shows the $\ln(I_C)$ versus V_{BE} characteristics in the normal, active region and the $\ln(I_E)$ versus V_{BC} characteristics in the inverse region in parts (a) and (b), respectively. From the first graph, since

$$
I_C = \frac{I_{SS}\left(e^{\frac{qV_{BE}}{kT}} - 1\right)}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)}
$$

it follows that

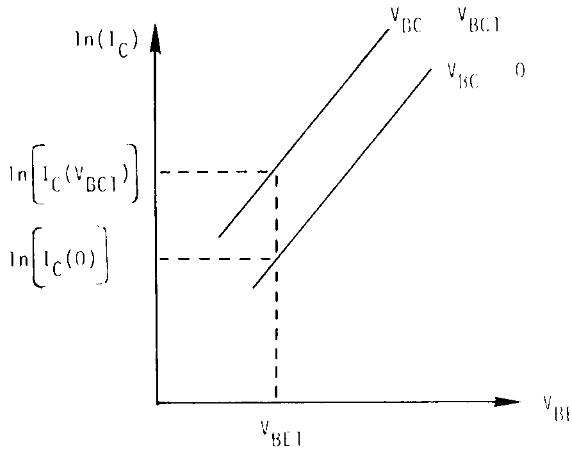
$$
\frac{I_C(0)}{I_C(V_{BC1})} = \frac{1 + \frac{V_{BE1}}{V_B} + \frac{V_{BC1}}{V_A}}{1 + \frac{V_{BC1}}{V_B}} \tag{3.40}
$$

Similarly from part (b), since

$$
I_E = \frac{I_{SS}\left(e^{\frac{qV_{BC}}{kT}} - 1\right)}{\left(1 + \frac{V_{BE}}{V_B} + \frac{V_{BC}}{V_A}\right)}
$$

then

V_B



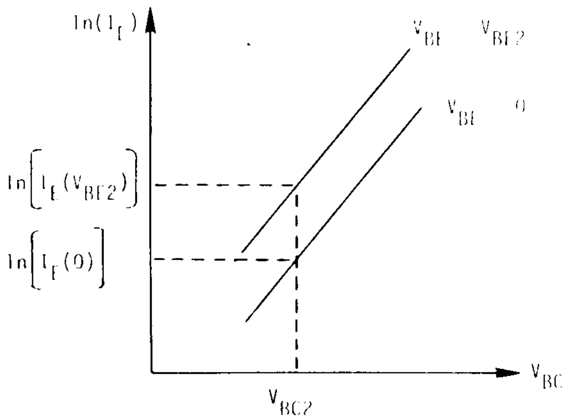


Fig. 3.32. The ln(I_C) versus V_{BE} characteristics in the normal, active region, (a), and the ln(I_E) versus V_{BC} characteristics in the inverse region, (b), from which V_A and V_B can be determined.

V_B

$$\frac{I_E(0)}{I_E(V_{BE2})} = \frac{1 + \frac{V_{BE2}}{V_B} + \frac{V_{BC2}}{V_A}}{1 + \frac{V_{BC2}}{V_A}} \qquad (3.41)$$

and Eqs. (3.40) and (3.41) can be solved for V_A and V_B. Since this technique does not involve the estimation of a slope it can be more accurate than the previous measurement. This is explained in more detail in the V_A-measurement description in Section 3.4.

A measure of the importance of V_B in the normal, active region can be obtained from a measurement of the slope of the $\ln(I_C)$ versus $\frac{qV_{BE}}{kT}$ curve (with $V_{BC} = 0$). If this slope departs significantly from unity, V_B can be important and may need better modeling. The slope of the $\ln(I_C)$ versus $\frac{qV_{BE}}{kT}$ curve in the normal region at low currents is given approximately by: (see Eq. (2.115)):

$$\frac{1}{n_E} = 1 - \frac{\frac{kT}{q}}{V_B + V_{BE}} \simeq 1 - \frac{kT}{qV_B} \qquad (3.42)$$

The measurement of this slope could be used to determine the appropriate value of V_B for the normal, active region, although the measurement could be inaccurate if the slope is near unity (which is usually the case).*

---

*For large departures of n_E from unity, Eq. (3.42) is too inaccurate. The more accurate, but less useful, Eq. (2.114) should be used.

I_K

Definition

I_K is the "knee current." It models the drop in β_F at high collector currents due to high-level injection. It is the (current) intersection of the low-current and high-current asymptotes of the $\ln(I_C)$ versus $V_{B'E'}$ graph, as shown in Fig. 3.33.

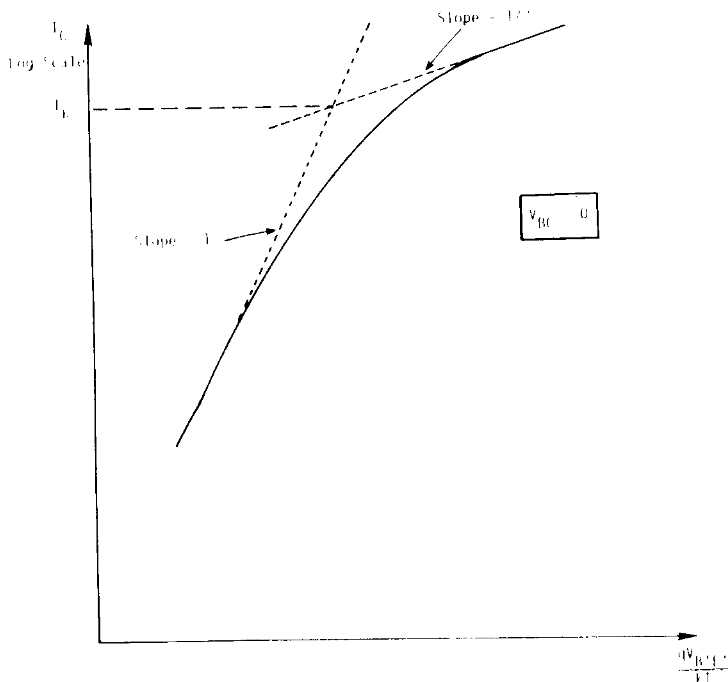


Fig. 3.33. The definition of I_K from the $\ln(I_C)$ versus $\frac{qV_{B'E'}}{kT}$ curve.

I_K

Typical Value

I_K typically ranges from approximately 0.1 mA to 10 mA.

Measurement Scheme

I_K is best determined from the plot of $\ln(I_C)$ versus $V_{B'E'}$, as shown in Fig. 3.33. Involved in the determination of I_K by this technique is the transformation of the horizontal axis from V_{BE} to $V_{B'E'}$ (as explained in the description of β_{FM}, C_2, n_{EL} and θ in Section 3.4) and the accurate determination of the two asymptotes.

An alternative method for determining I_K is to fit the experimentally observed β_F versus I_C curve (for $V_{BC} = 0$) to Eq. (2.192).

Note that both techniques for determining I_K are compatible with the determination of the other β_F versus I_C model parameters (β_{FM}, C_2 and n_{EL}). In fact, I_K would not be determined separately but at the same time as these other three model parameters.

I_{KR}

Definition

I_{KR} is the "inverse knee current." It models the drop in β_R at high emitter currents due to high-level injection. It is the (current) intersection of the low-current and high-current asymptotes of the $\ln(I_E)$ versus $V_{B'C'}$ graph (in the inverse mode of operation).

Typical Value

I_{KR} typically is of the order of 1 mA.

Measurement Scheme

I_{KR} is determined in exactly the same way as I_K from a plot of $\ln(I_E)$ versus $V_{B'C'}$ in the inverse mode of operation. As with I_K, I_{KR} is normally determined at the same time as the other β_R versus I_E model parameters -- β_{RM}, C_4 and n_{CL}.

B

Definition

B is the "base push-out factor."[7] It is a multiplier that models the effective increase in the basewidth at high current levels. It therefore has the property of being equal to unity at low currents and greater than unity at high currents.

Typical Value

Since B is a function of I_C and requires several model parameters to describe it, there is no typical value.

Measurement Scheme

The effective increase in the basewidth at high currents modeled by B causes an extra drop in β_F at high currents over that modeled by I_K. The variation of B with I_C could therefore be obtained by comparing the actual β_F versus I_C at high currents with that computed with I_K. This is done by considering the expression for β_F versus I_C with and without the multiplier B. For B = 1, and $C_2 = 0$ (i.e., ignoring low-current effects), Eq. (2.192) gives:

$$\beta_F(0)\Big|_{B=1}^{-1} = \frac{1}{\beta_{FM}(0)} + \frac{I_C}{\beta_{FM}(0)\, I_K} \tag{3.43}$$

For non-unity B,

$$\beta_F(0)\Big|_{B\neq 1}^{-1} = \frac{1}{\beta_{FM}(0)} + \frac{I_C \cdot B}{\beta_{FM}(0)\, I_K} \tag{3.44}$$

Therefore,

$$B = \frac{I_K}{I_C}\left[\frac{\beta_{FM}(0)}{\beta_F(0)\Big|_{meas}} - 1\right] \tag{3.45}$$

where it is assumed that the non-unity B case corresponds to the measured case.*

*To reduce the effects of heating on β_F at high currents, measurements should be taken at the minimum V_{BC} that still results in normal, active operation and in a pulsed mode.

From the generated curve of B versus I_C the appropriate parameters used to model B could be obtained by a curve-fitting procedure. Since the GP model as implemented in SPICE does not model B, the determination of its model parameters is not described here.*

The value of B is obviously a strong function of the value of I_K used. For each I_K chosen there will be an appropriate B versus I_C curve. Therefore, if the generated B curve is not fittable to the expression used for B, another value of I_K could be chosen until the curve is adequately fitted.** The value of I_K chosen may be important if the mid-current, constant-β region does not exist and an accurate fit for the transition from high to low currents is required. The value of I_K (and B) will also be important for the correct modeling of τ_F versus I_C.

It would appear that since B also results in the variation of $\tau_{F_{ac}}$ with I_C, the value of B (and I_K) could be obtained from the measurement of $\tau_{F_{ac}}$ versus I_C. However, the above method of generating B from β_F data would appear to have the following advantages over the use of $\tau_{F_{ac}}$ data.

(a) Since β_F is a dc measurement it is easier to determine than $\tau_{F_{ac}}$ (which is an ac measurement)

(b) β_F is measured relatively directly whereas τ_F is usually measured via f_T. Therefore the β_F measurements are normally more accurate.

(c) The expression for β_F versus I_C is considerably simpler than that for $\tau_{F_{ac}}$, especially if neither τ_B nor the emitter delay, τ_1, is dominant.

---

* In SPICE (Version 1) an increase in τ_F with I_C has been built in that follows the variation of the emitter delay component(3) (see Eq. (2.185)).

** β and τ_F are also functions of crowding(48) and V_{BC}(71). Therefore it may be difficult to fit B to an expression that only considers effective base widening.

APPENDIX - 1

A COMPARISON OF THE TRANSPORT NOTATION WITH THE INJECTION NOTATION[9]

This Appendix compares the transport and injection notations, described in the EM_1 model (Section 2.2), and shows that for CAD the transport notation is preferred. The reasons for the preference of this notation lie in the consideration of both β at low currents and the description of the diffusion capacitance. The variation of β at low currents is considered first.

a) β_F at low currents

Figure A1.1 shows the variation of the reference currents for both notations as a function of the appropriate junction voltages at low and medium currents. The injection notation curves are shown in Fig. A1.1a, while the curves for the transport notation are shown in Fig. A1.1b.

In the injection notation, the equations describing the curves in Fig. A1.1a are:

$$I_F = \frac{I_S}{\alpha_F}\left(e^{\frac{qV_{BE}}{kT}} - 1\right) \tag{A1.1}$$

$$I_R = \frac{I_S}{\alpha_R}\left(e^{\frac{qV_{BC}}{kT}} - 1\right) \tag{A1.2}$$

In this comparison, the ohmic resistors r'_b, r'_e and r'_c are ignored. The saturation current, I_S, is shown in the GP model description (Section 2.5) to be a fundamental constant of the transistor when operating under low-level injection.[7]* Since I_S is constant, the

*In the GP model, I_S is replaced by I_{SS}. In this Appendix, the effects of base-width modulation due to emitter-base space-charge variations are ignored.

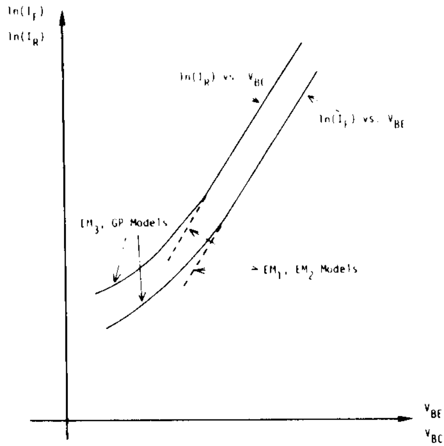


Fig. A1.1a. Variation of the injection notation reference currents.

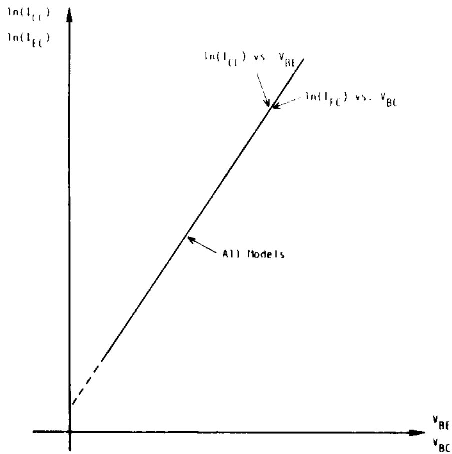


Fig. A1.1b. Variation of the transport notation reference currents.

coefficients $\frac{I_S}{\alpha_F}$ and $\frac{I_S}{\alpha_R}$ increase as the currents are decreased (because of the drop in the β's and therefore the drop in the α's). This increase in the coefficients explains the deviations from the straight lines at low currents shown in Fig. A1.1a.

For the transport notation, even for non-constant α, not only are the lines of Fig. A1.1b straight but they are identical. The equations describing the curves are:

$$I_{CC} = I_S\left(e^{\frac{qV_{BE}}{kT}} - 1\right) \tag{A1.3}$$

$$I_{EC} = I_S\left(e^{\frac{qV_{BC}}{kT}} - 1\right) \tag{A1.4}$$

The variation of both reference currents with the appropriate junction voltage is described by the one, fundamental constant, I_S, as evidenced by Eqs. (A1.3) and (A1.4). This represents the main advantage of the transport notation -- the fact that reference currents (those currents in terms of which all other elements in the model are expressed) are ideal and defined by one fundamental constant. Another way of describing the same result is that in the transport notation all the non-ideality in β at low currents is in the base currents (where it belongs).

b) Diffusion capacitances [9]

The diffusion capacitances introduced in the EM_2 model characterize the total mobile charges in the transistor. The minority mobile charges stored in the base region (Q_3 and Q_7) are given by:

$$Q_3 = \tau_{EB}\, I_F \quad \text{(injection notation)} \tag{A1.5}$$

$$= \tau_B\, I_{CC} \quad \text{(transport notation)} \tag{A1.6}$$

$$Q_7 = \tau_{CB}\, I_R \quad \text{(injection notation)} \tag{A1.7}$$

$$= \tau_{BR}\, I_{EC} \quad \text{(transport notation)} \tag{A1.8}$$

The difference between the two notations lies in the difference between τ_{EB} and τ_B (and between τ_{CB} and τ_{BR}). The injection parameters, τ_{EB} and τ_{CB}, describe <u>both</u> base transit time and base recombination. This can be seen from equations (A1.5) and (A1.7) and Fig. A1.1a. At low currents, I_F and I_R deviate from the straight line due to the drop in α_F and α_R, respectively. This must be compensated for by the parameters τ_{EB} and τ_{CB} which are therefore not constant at low currents. The transport parameters, τ_B and τ_{BR}, <u>are</u> constant at low currents, however. The variations of τ_{EB} and τ_{CB} with current can cause complications when the above equations ((A1.5) through (A1.8)) are differentiated to obtain the base region components of the small-signal diffusion capacitance values.

APPENDIX - 2

EM_3 MODEL BASEWIDTH MODULATION ANALYSIS[31]

The following analysis of basewidth modulation proceeds in five stages: first, the assumptions and the general philosophy are explained; an expression for the variation of the basewidth is obtained; the Early voltage, V_A, is then defined mathematically; the three model parameters that are a strong function of the basewidth are modified appropriately; finally a geometrical interpretation of V_A is given.

a) Assumptions made in analysis

In the following analysis the transistor is assumed to be operating in a linear mode. As well, many results taken from a simple, constant-doping analysis are assumed to also hold in the general case. All these assumptions, which appear to be not very accurate, are justifiable when one remembers that basewidth modulation is itself normally a second-order effect. Therefore a first-order analysis of it is acceptable.

b) Basewidth variation with V_{BC}

The basewidth, W, is a function of V_{BC}

$$W = f(V_{BC}) \tag{A2.1}$$

Assuming a linear mode of operation, a Taylor series expansion can be made about $V_{BC} = 0$. Therefore, neglecting the second-order terms.

$$W(V_{BC}) \simeq W(0) + V_{BC} \left. \frac{dW}{dV_{BC}} \right|_{V_{BC} = 0} \tag{A2.2}$$

$$\frac{W(V_{BC})}{W(0)} \simeq 1 + \frac{V_{BC}}{W(0)} \cdot \left. \frac{dW}{dV_{BC}} \right|_{V_{BC} = 0} \tag{A2.3}$$

Basically, Eq. A2.3 assumes a linear variation of W as a function of V_{BC}.

c) Definition of V_A

The Early voltage, V_A, is defined as

$$V_A \triangleq \left[\frac{1}{W(0)} \frac{dW}{dV_{BC}} \bigg|_{V_{BC} = 0} \right]^{-1} \quad \text{(npn transistor) (A2.4)}$$

$$\triangleq \left[- \frac{1}{W(0)} \frac{dW}{dV_{BC}} \bigg|_{V_{BC} = 0} \right]^{-1} \quad \text{(pnp transistor) (A2.5)}$$

The difference between the definitions for the npn and pnp transistor lies only in the sign of V_{BC}. For an npn transistor in the normal, active region, V_{BC} is negative (i.e., reverse-bias). An increase in V_{BC} (a decrease in the reverse bias) results in an increase in the basewidth and the derivative in Eq. (A2.4) is positive. The minus sign in Eq. (A2.5) preserves the positive nature of V_A for a pnp transistor. The following derivation assumes an npn transistor. For a pnp transistor, V_{BC} should be changed to V_{CB} (= - V_{BC}). The basewidth as a function of V_A is given by

$$\frac{W(V_{BC})}{W(0)} = 1 + \frac{V_{BC}}{V_A} \qquad \text{(npn transistor)} \qquad \text{(A2.6)}$$

A typical value for V_A is 50 V.

d) Modification of model parameters

Three model parameters have a strong dependence on W: I_S, β_F and τ_B.

(i) I_S

The saturation current, I_S, is inversely proportional to W for constant base doping.[33] Assuming that approximately the same dependence on W holds for transistors with non-constant doping

$$I_S \propto \frac{1}{W} \tag{A2.7}$$

and

$$I_S(V_{BC}) = I_S(0)\,\frac{W(0)}{W(V_{BC})}$$

$$\therefore \quad \boxed{I_S(V_{BC}) = \frac{I_S(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)}} \tag{A2.8}$$

Computationally, Eq. (A2.8) becomes infinite at $V_{CB} = V_A$. To overcome this problem, the binomial expansion can be used. Assuming that $|V_{BC}| << V_A$,

$$\boxed{I_S(V_{BC}) \approx I_S(0)\left(1 - \frac{V_{BC}}{V_A}\right)} \tag{A2.9}$$

(ii) $\underline{\beta_F}$

If it is assumed that most of the base current is due to the injection of carriers from the base into the emitter (i.e., the emitter injection efficiency component), then it can be shown that for constant base doping[34]

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F}$$

$$= \frac{D_n}{D_p} \cdot \frac{n_p}{p_n} \cdot \frac{L_p}{W} \propto \frac{1}{W} \tag{A2.10}$$

Therefore, as with I_S

$$\boxed{\beta_F(V_{BC}) = \frac{\beta_F(0)}{\left(1 + \frac{V_{BC}}{V_A}\right)} \approx \beta_F(0)\left[1 - \frac{V_{BC}}{V_A}\right]} \tag{A2.11}$$

The assumptions involved in this equation are negligible base recombination (i.e., all base current is injected into the emitter), the validity of Eq. (A2.10) for non-constant doping and, for the second form of Eq. (A2.11), $|V_{BC}| \ll V_A$.

(iii) $\underline{\tau_B}$

The base transit time, τ_B, is nominally proportional to the square of the basewidth (for low-level injection and constant base doping),[26] i.e.,

$$\tau_B = \frac{W^2}{2D_n} \propto W^2 \tag{A2.12}$$

Therefore, assuming the same dependence in general[26]

$$\tau_B(V_{BC}) = \tau_B(0)\left[\frac{W(V_{BC})}{W(0)}\right]^2 \tag{A2.13}$$

$$\boxed{\tau_B(V_{BC}) = \tau_B(0)\left(1 + \frac{V_{BC}}{V_A}\right)^2} \tag{A2.14}$$

The reverse base transit time, τ_{BR}, will have a similar dependence on V_{BC}. However, since τ_{BR} is normally only a small component of τ_R, the dependence of τ_{BR} on V_{BC} is neglected.

e) Geometrical interpretation of V_A

The Early voltage, V_A, is measured from the slope of the I_C versus V_{CE} characteristics. The slope, in the normal, active region, g_o, is given (assuming constant V_{BE}) by:

$$g_o = \left.\frac{dI_C}{dV_{CE}}\right|_{V_{BE} = \text{const.}} = -\left.\frac{dI_C}{dV_{BC}}\right|_{V_{BE} = \text{const.}}$$

$$= - \frac{d}{dV_{BC}} \left[I_S(V_{BC}) \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \right]_{V_{BE} = \text{const.}}$$

$$- \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \cdot \frac{d}{dV_{BC}} \left[I_S(0) \left(1 - \frac{V_{BC}}{V_A} \right) \right]$$

$$= + \frac{I_S(0) \left(e^{\frac{qV_{BE}}{kT}} - 1 \right)}{V_A}$$

$$= \frac{I_C(0)}{V_A} \qquad \text{(A2.15)}$$

The geometric interpretation of Eq. (A2.15) shows that V_A is obtained from the intercept of the extrapolated slope on the V_{CE} axis ($-V_A'$), as shown in Fig. A2.1.

$$\therefore \; V_A = V_A' + V_{BE} \qquad \text{(A2.16)}$$

For typical values of V_A (of the order of 50 V), Eq. (A2.16) can be approximated by

$$V_A \simeq V_A' \qquad \text{(A2.17)}$$

The preceding analysis of basewidth modulation applied to the constant-V_{BE} curves. It can also be extended to the constant-I_B curves. The base current is given by

$$I_B = \frac{I_S}{\beta_F} \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \qquad \text{(A2.18)}$$

Since, from Eqs. (A2.7) and (A2.10), both I_S and β_F are inversely proportional to W, their ratio is independent of W (to first order). It follows then that the constant-V_{BE} curve (for any V_{BC}) is a constant - I_B curve since I_B is independent of V_{BC} and for one value of V_{BE} there is one value of I_B. Therefore, the above analysis is, to first order, also valid for constant - I_B curves.

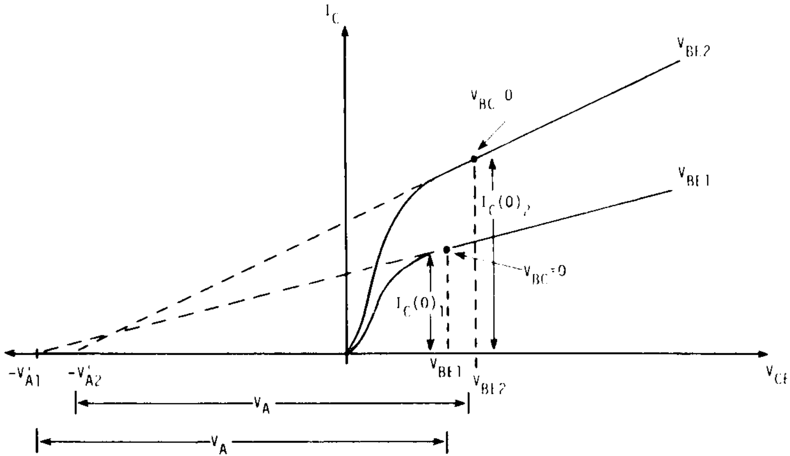


Fig. A2.1. The geometrical interpretation of V_A (not drawn to scale).

APPENDIX - 3

DERIVATION OF THE FIVE COMPONENTS OF Q_B IN THE GP MODEL

In this derivation of the five components of Q_B, the mobile charge in the space-charge layers for non-equilibrium conditions will be included from the beginning. The depletion approximation will only be applied for the thermal equilibrium case (i.e., when $p(x) = p_o(x)$). That is, it will be assumed that

$$\int_{x_E'}^{x_E} qAp_o(x)\,dx = 0 \tag{A3.1}$$

$$\int_{x_C}^{x_C'} qAp_o(x)\,dx = 0 \tag{A3.2}$$

The starting point for the derivation is the definition of Q_B that includes the space-charge layers (Eq. (2.127))

$$Q_B \triangleq \int_{x_E'}^{x_C'} qAp(x)dx \tag{A3.3}$$

The first step is to separate $p(x)$ into two components: $p_o(x)$ and $p'(x)$

$$\therefore\ Q_B = \int_{x_E'}^{x_C'} qAp_o(x)dx + \int_{x_E'}^{x_C'} qAp'(x)dx \tag{A3.4}$$

where $p_o(x)$ is the equilibrium hole concentration and $p'(x)$ is the excess hole concentration.

The first integral is now split into three components

$$\therefore Q_B = \int_{x_E'}^{x_{E0}} qAp_o(x)dx + \int_{x_{E0}}^{x_{C0}} qAp_o(x)dx + \int_{x_{C0}}^{x_C'} qAp_o(x)dx + \int_{x_E'}^{x_C'} qAp'(x)dx$$

$$\triangleq \quad Q_E \quad + \quad Q_{B0} \quad + \quad Q_C \quad + \int_{x_E'}^{x_C'} qAp'(x)dx$$

(A3.5)

Superposition (Eq. (2.96)) is used to simplify the remaining integral by splitting $p'(x)$ into two components: $p_F(x)$ and $P_R(x)$.

$$\therefore Q_B = Q_E + Q_{B0} + Q_C + \int_{x_E'}^{x_C'} qA\left[p_F(x)-p_o(x)\right] dx + \int_{x_E'}^{x_C'} qA\left[p_R(x)-p_o(x)\right] dx$$

$$\triangleq Q_E + Q_{B0} + Q_C + Q_F + Q_R \qquad \text{(A3.6)}$$

The definitions of Q_E and Q_C can be simplified by applying the depletion approximation.

$$\therefore Q_E \triangleq \int_{x_E'}^{x_{E0}} qAp_o(x)dx$$

$$= \int_{x_E'}^{x_E} qAp_o(x)dx + \int_{x_E}^{x_{E0}} qAp_o(x)dx \approx \int_{x_E}^{x_{E0}} qAp_o(x)dx \qquad \text{(A3.7)}$$

$$Q_C \triangleq \int_{x_{C0}}^{x'_C} qAp_o(x)dx$$

$$= \int_{x_{C0}}^{x_C} qAp_o(x)dx + \int_{x_C}^{x'_C} qAp_o(x)dx$$

$$\approx \int_{x_{C0}}^{x_C} qAp_o(x)dx \qquad \text{(A3.8)}$$

The definition of Q_{B0} above (Eq. (A3.5)) can be "expanded" by the application of the depletion approximation. Q_{B0} can be alternately defined as the value of Q_B under zero bias. The application of the depletion approximation yields:

$$Q_{B0} \triangleq \int_{x'_{E0}}^{x'_{C0}} qAp_o(x)dx$$

$$= \int_{x'_{E0}}^{x_{E0}} qAp_o(x)dx + \int_{x_{E0}}^{x_{C0}} qAp_o(x)dx + \int_{x_{C0}}^{x'_{C0}} qAp_o(x)dx$$

$$\approx \int_{x_{E0}}^{x_{C0}} qAp_o(x)dx \qquad \text{(A3.9)}$$

which is the definition used in Eq. (A3.5).

The above equations for Q_E, Q_{BO}, Q_C, Q_F and Q_R are consistent with those given in Eqs. (2.128) through (2.137) when $p_o(x)$ is replaced by $N_A(x)$ in the neutral base region.

APPENDIX - 4

THE ACCURACY OF THE EM_3 AND GP BASEWIDTH MODULATION MODELS

In the derivation of both the EM_3 and GP basewidth modulation models, assumptions were made that appear to be rather inaccurate. The merits of these assumptions are described here.

a) EM_3 Model

The derivation of the EM_3 basewidth modulation model is given in Appendix 2. The major assumptions in the derivation are: (i) the neglect of second-order and higher terms in the Taylor series expansion of the basewidth as a function of V_{BC} and (ii) the extension of the constant base-doping results for I_S, β_F and τ_B to the general, non-constant doping case.

The validity of these assumptions is difficult to assess theoretically. Their main justification lies in realizing that basewidth modulation is itself a second-order effect. Therefore, only a first-order model of it is required, since any second-order effects in the modeling of basewidth modulation are third-order effects in the overall performance. A second, perhaps more convincing, argument is the fact that the simple EM_3 model for basewidth modulation has been used successfully in many applications. For those interested in a more accurate model for basewidth modulation, the following analysis of the GP model gives the details on how it is obtained.

b) GP Model

The EM_3 model made the linear assumption at the beginning of the analysis. The GP model, however, makes its major assumption (constant junction capacitance) at the end of the analysis. Therefore, unlike the EM_3 model, the GP model shows what needs to be done if a more accurate model is required. This is illustrated here.

In the GP model, basewidth modulation is modeled by q_c (the Early effect) and q_e (the Late effect). The expressions for q_c and q_e are

$$q_c = \frac{\int_0^{V_{B'C'}} C_{jC}(V)dV}{Q_{BO}} \qquad (A4.1)$$

$$q_e = \frac{\int_0^{V_{B'E'}} C_{jE}(V)dV}{Q_{BO}} \qquad (A4.2)$$

In both cases, the integral is replaced by an average capacitance which is assumed to be constant.
Therefore

$$q_c = \frac{\overline{C}_{jC} V_{B'C'}}{Q_{BO}} \triangleq \frac{V_{B'C'}}{V_A} \qquad (A4.3)$$

$$q_e = \frac{\overline{C}_{jE} V_{B'E'}}{Q_{BO}} \triangleq \frac{V_{B'E'}}{V_A} \qquad (A4.4)$$

where V_A(the Early voltage) and V_B (the inverse Early voltage), given by:

$$V_A \triangleq \frac{Q_{BO}}{\frac{1}{V_{B'C'}} \int_0^{V_{B'C'}} C_{jC}(V)dV} \qquad (A4.5)$$

$$V_B \wedge \frac{Q_{BO}}{\frac{1}{V_{B'E'}} \int_0^{V_{B'E'}} C_{jE}(V)dV} \tag{A4.6}$$

are assumed to be constant and independent of junction voltage. Normally, the assumption of constant V_A and V_B is acceptable because basewidth modulation is a second-order effect (i.e., V_A and V_B are large). The accuracy of the assumption of constant V_A and V_B is actually a function of the ranges of voltages over which they are applied and of how they are measured. If they are measured for one particular applied junction voltage they will, of course, be most accurate at that voltage. However, if they are averaged over a range of junction voltages, the accuracy over that range may be improved. This point is illustrated now by assuming that V_A and V_B have been measured at zero applied junction voltage (a convenient standard applied voltage) and by looking at the accuracy of the constant-V_A and constant-V_B assumption under reverse and forward bias. Therefore, it will be assumed that

$$V_{A_{meas}} = \frac{Q_{BO}}{C_{jCO}} \tag{A4.5}$$

$$V_{B_{meas}} = \frac{Q_{BO}}{C_{jEO}} \tag{A4.6}$$

and that q_c and q_e are modeled as

$$q_c = \frac{V_{B'C'}}{V_{A_{meas}}} = \frac{C_{jCO}\ V_{B'C'}}{Q_{BO}} \tag{A4.7}$$

$$q_e = \frac{V_{B'E'}}{V_{B_{meas}}} = \frac{C_{jEO}\ V_{B'E'}}{Q_{BO}} \tag{A4.8}$$

The effects of this assumption are shown in Fig. A4.1, which is drawn for both the reverse-bias case, (a), and the forward-bias case, (b). As illustrated in the figure, the integrals in Eqs. (A4.1) and (A4.2) represent the area under the actual C-V curve while Eqs. (A4.7) and (A4.8) approximate this area by the rectangles shown. Since Fig. A4.1 refers to <u>both</u> Q_C and Q_E, the subscripts C and E have been dropped.

(i) <u>Reverse Bias Case</u>

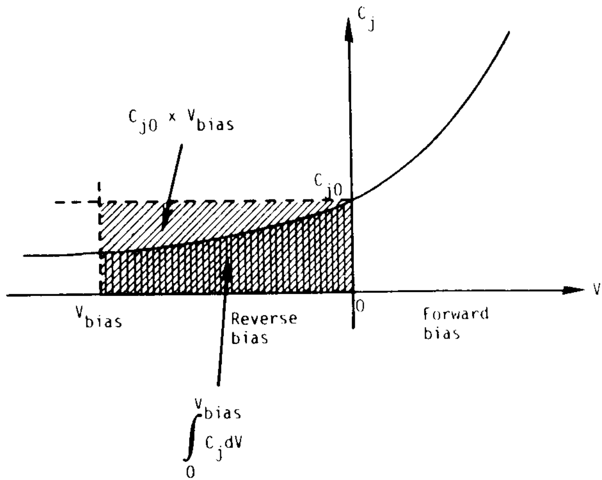
As can be seen from Fig. A4.1a, the constant-capacitance approximation is relatively valid when the junction is reverse biased. A better agreement would be obtained in this case if the measurement of the appropriate Early voltage was made at some voltage between 0 and V_{bias}. However, this would require a prior knowledge of V_{bias}. Without this information, the use of Eqs. (A4.7) and (A4.8) appears to be relatively reasonable.

(ii) <u>Forward Bias Case</u>

Obviously, as seen from Fig. A4.1b, the constant-capacitance approximation can cause gross errors when the junction is forward biased. As with the reverse-bias case, a better value of V_A or V_B would be obtained if the measurement was made at some voltage between 0 and V_{bias}. The justification for still retaining Eqs. (A4.7) and (A4.8) lies in the assumption that under forward bias the value of q_e or q_c is very much less than unity. If this is not so, the integration should be performed, as shown below.

In the EM_2 model, the variation of C_j with V was assumed to have the form[1] (in the SLIC and SINC programs)

$$C_j(V) = \begin{cases} \dfrac{C_{j0}}{\left(1 - \dfrac{V}{\phi}\right)^m} & \text{for } V \leq \dfrac{\phi}{2} \\[2ex] = 2^m C_{j0}\left[2m\dfrac{V}{\phi} + (1-m)\right] & \text{for } V \geq \dfrac{\phi}{2} \end{cases} \qquad (A4.9)$$



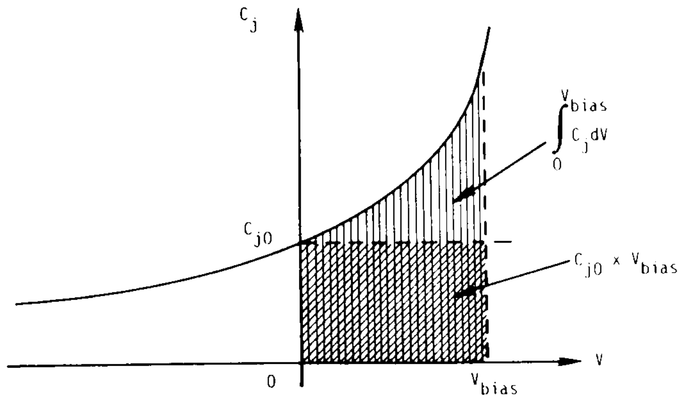


Fig. A4.1. The difference between the accurate evaluation of

$$\int_0^{V_{bias}} C_j dV \text{ (represented by the vertical lined area)}$$

and the C_{jO}, V_{bias} approximation (represented by the diagonal lined area) for the reverse bias case (a) and the forward bias case (b).

Integration yields for q_e and q_c

$$
\left.
\begin{aligned}
q &= \frac{\phi}{V_{meas}} \cdot \frac{1}{(1-m)} \cdot \left[1 - \left(1 - \frac{V}{\phi}\right)^{1-m}\right] \quad \text{for } V \le \frac{\phi}{2} \\
&= \frac{\phi}{V_{meas}} \cdot \frac{1}{(1-m)} \cdot \left[1 - \left(\frac{1}{2}\right)^{1-m}\right] \\
&+ \frac{2^m}{V_{meas}} \left(V - \frac{\phi}{2}\right)\left[m \frac{V}{\phi} + 1 - \frac{m}{2}\right] \qquad \text{for } V \ge \frac{\phi}{2}
\end{aligned}
\right\} \qquad \text{(A4.10)}
$$

The more accurate Poon-Gummel equation[23] for $C_j(V)$ is not used since the effect of excess mobile carriers in the space-charge layers are not included in q_e or q_c (but in q_f and q_r).

APPENDIX- 5

THE SMALL-SIGNAL, LINEARIZED EM_3 AND GP MODELS

a) The EM_3 Model

The small-signal EM_3 model is drawn in Fig. A5.1 for an npn transistor. (For a pnp transistor, the position of C_{SUB} in the model may be altered.) The only difference in form between this model and that used for the EM_2 model (Fig. 2.12) is the inclusion of the capacitor RATIO $C_{jC}(V_{B'C'})$ between nodes B and C'. This has resulted from the split of C_{jC} across r_b'. The only other differences between the EM_2 and EM_3 small-signal models are in the determination of r_π, r_μ, C_π and C_μ and the replacement of g_{mF} and g_{mR} by $g_{m\pi}$ and $g_{m\mu}$, respectively. The EM_2 model equations are given in Section 2.3.3. The EM_3 model definitions and equations are given here

$$r_\pi \triangleq \frac{\beta_{F_{ac}}}{g_{m\pi}} \tag{A5.1}$$

$$r_\mu \triangleq \frac{\beta_{R_{ac}}}{g_{m\mu}} \tag{A5.2}$$

$$C_\pi \triangleq g_{m\pi}\, \tau_{F_{ac}} + C_{jE}(V_{B'E'}) \tag{A5.3}$$

$$C_\mu \triangleq g_{m\mu}\, \tau_{R_{ac}} + (1\text{-RATIO})\; C_{jC}(V_{B'C'}) \tag{A5.4}$$

where

$$g_{m\pi} \triangleq \left.\frac{dI_{CT}}{dV_{B'E'}}\right|_{\Delta V_{B'C'}=0} \tag{A5.5}$$

$$\approx \frac{qI_{CC}}{kT}\left[1-\frac{\theta}{2}\,\frac{I_{CT}}{I_S(0)}\left(1+\frac{V_{B'C'}}{V_A}\right)e^{-\frac{qV_{B'E'}}{2kT}}\right] \tag{A5.6}$$

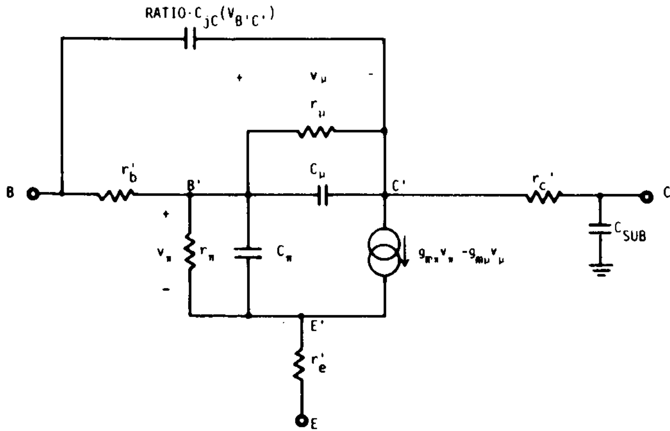


Fig. A5.1. The linearized EM_3 model for an npn transistor.

$$g_{m\mu} \triangleq \left.\frac{d(-I_{CT})}{dV_{B'C'}}\right|_{\Delta V_{B'E'}=0} \tag{A5.7}$$

$$\simeq \frac{qI_{EC}}{kT}\left[1+\frac{\theta_R}{2}\,\frac{I_{CT}}{I_S(0)}\left(1+\frac{V_{B'C'}}{V_A}\right)e^{-\frac{qV_{B'C'}}{2kT}}+\frac{I_{CT}}{I_{EC}}\,\frac{\frac{kT}{q}}{V_A+V_{B'C'}}\right] \tag{A5.8}$$

$$\beta_{F_{ac}}^{-1} \triangleq \frac{g_\pi}{g_{m\pi}} = \left.\frac{1}{g_{m\pi}}\,\frac{dI_B}{dV_{B'E'}}\right|_{\Delta V_{B'C'}=0} \tag{A5.9}$$

$$= \frac{qI_S(0)}{g_{m\pi}kT}\left[\frac{1}{\beta_{FM}}\,e^{\frac{qV_{B'E'}}{kT}}+\frac{C_2}{n_{EL}}\,e^{\frac{qV_{B'E'}}{n_{EL}kT}}\right] \tag{A5.10}$$

$$= \beta_{FM}^{-1} + \frac{C_2}{n_{EL}}\left(\frac{I_C}{I_S(0)}\right)^{\left(\frac{1}{n_{EL}}-1\right)} + \frac{2\theta^2}{\beta_{FM}I_S(0)}\,I_C \quad \text{at } V_{B'C'}=0 \tag{A5.11}$$

$$\beta_{R_{ac}}^{-1} \triangleq \frac{g_{\mu}}{g_{m\mu}} = \frac{1}{g_{m\mu}} \cdot \frac{dI_B}{dV_{B'C'}}\Bigg|_{\Delta V_{B'E'}=0} \tag{A5.12}$$

$$= \frac{qI_S(0)}{g_{m\mu}kT}\left[\frac{1}{\beta_{RM}} e^{\frac{qV_{B'C'}}{kT}} + \frac{C_4}{n_{CL}} e^{\frac{qV_{B'C'}}{n_{CL}kT}}\right] \tag{A5.13}$$

$$\tau_{F_{ac}} = \tau_{FL}\left[1 + \frac{1}{4}\left(\frac{L_E}{W}\right)^2\left(\frac{I_{CC}}{I_{CO}} - 1\right)^2\right] \tag{A5.14}$$

$$\tau_{R_{ac}} = \tau_{R_{dc}} = \tau_R \tag{A5.15}$$

All other parameters are as described in the EM_3 model (Section 2.4).

The equations for $g_{m\pi}$ and $g_{m\mu}$ (Eqs. (A5.6) and (A5.8)) illustrate the effect of high-level injection. Under low-level injection conditions and the forward, active region for Eq. (A5.6) and the inverse region for Eq. (A5.8), the second term in each equation is negligible and the equations reduce to the well-known EM_2 equations. Under high-level conditions and the just mentioned operating regions the term in the brackets becomes 0.5 (a result that is easily obtained from a consideration of the high-level injection asymptotes, such Eq. (2.57) for I_{CC}).

The expression for $\beta_{F_{ac}}$ (and the similar one for $\beta_{R_{ac}}$) is directly obtainable from the formula for I_B (Eq. (2.56)) and illustrates the effects of the non-ideal components of I_B and high-level injection. If these effects are not present (i.e., $C_2 = C_4 = 0$, $\theta = \theta_R = \infty$), $\beta_{F_{ac}} = \beta_{F_{dc}} = \beta_{FM}$ and $\beta_{R_{ac}} = \beta_{R_{dc}} = \beta_{RM}$.

The variation of τ_F with I_C due to the van der Ziel and Agouridis effect[40] is included in the formula for $\tau_{F_{ac}}$. τ_R is assumed here to be a constant.

The effect of basewidth modulation on the small-signal parameters is included but hidden in the above equations. Basewidth modulation is inherently contained in the terms I_{CC}, I_{EC}, I_C, I_E, I_S, β_{FM} and τ_{FL}. Therefore, unlike in the familiar hybrid-π model,[5] there is no need for r_o nor does r_μ here contain the effects of basewidth modulation (corresponding to an assumption of negligible base recombination -- see Eq. (A5.33)).

b) The GP Model

The small-signal GP model is drawn in Fig. A5.2 for an npn transistor. It is identical, in form, to that used for the EM_2 model (Fig. 2.12). The formulae for r_π, r_μ, C_π and C_μ, given here, are different and g_{mF} and g_{mR} are replaced by $g_{m\pi}$ and $g_{m\mu}$, respectively.

$$r_\pi \triangleq \frac{\beta_{F_{ac}}}{g_{m\pi}} \qquad (A5.16)$$

$$r_\mu \triangleq \frac{\beta_{R_{ac}}}{g_{m\mu}} \qquad (A5.17)$$

$$C_\pi = g_{m\pi}\, \tau_{F_{ac}} + C_{jE}(V_{B'E'}) \qquad (A5.18)$$

$$C_\mu = g_{m\mu}\, \tau_{R_{ac}} + C_{jC}(V_{B'C'}) \qquad (A5.19)$$

where

$$g_{m\pi} \triangleq \left.\frac{dI_{CT}}{dV_{B'E'}}\right|_{\Delta V_{B'C'}=0} \qquad (A5.20)$$

$$\simeq \frac{qI_{CC}}{kT} - \frac{I_{CT}}{q_b V_B}\left[\frac{q_b + \frac{I_{SS}}{I_K} \cdot \frac{qV_B}{kT} \cdot e^{\frac{qV_{B'E'}}{kT}}}{2q_b - q_1}\right] \quad \text{for } B=1 \qquad (A5.21)$$

$$g_{m\mu} \triangleq \left.\frac{d(-I_{CT})}{dV_{B'C'}}\right|_{\Delta V_{B'E'}=0} \qquad (A5.22)$$

$$\simeq \frac{qI_{EC}}{kT} + \frac{I_{CT}}{q_b V_A}\left[\frac{q_b + \frac{I_{SS}}{I_{KR}} \cdot \frac{qV_A}{kT}\, e^{\frac{qV_{B'C'}}{kT}}}{2q_b - q_1}\right] \qquad (A5.23)$$

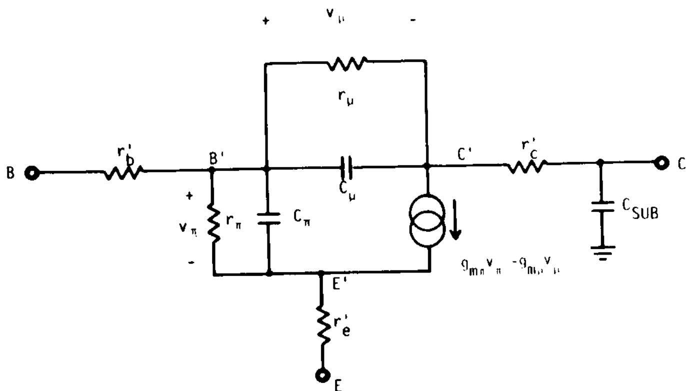


Fig. A5.2. The linearized GP model for an npn transistor.

$$\beta_{F_{ac}}^{-1} \triangleq \frac{g_\pi}{g_{m\pi}} = \frac{1}{g_{m\pi}} \left. \frac{dI_B}{dV_{B'E'}} \right|_{\Delta V_{B'C'}=0}$$

$$= \frac{qI_{SS}}{g_{m\pi}kT} \left[\frac{1}{\beta_{FM}} e^{\frac{qV_{B'E'}}{kT}} + \frac{C_2}{n_{EL}} e^{\frac{qV_{B'E'}}{n_{EL}kT}} \right] \tag{A5.24}$$

For $V_{B'C'}=0$ and B=1 it can be shown from Eq. (2.192) that

$$\beta_{F_{ac}}^{-1} = \frac{dI_B}{dI_C} = 2\beta_F^{-1} - \beta_{FM}^{-1}$$

$$+ \frac{C_2}{n_{EL}} \frac{I_{SS}}{I_K}^{\left(1-\frac{1}{n_{EL}}\right)} \cdot I_C^{\left(2-n_{EL}\right)} \left[\frac{I_C}{I_K}^{\left(2-n_{EL}\right)} + I_C^{\left(1-n_{EL}\right)} \right]^{\left(\frac{1}{n_{EL}}-1\right)}$$

$$\beta_{R_{ac}}^{-1} \triangleq \frac{g_\mu}{g_{m\mu}} = \frac{1}{g_{m\mu}} \left. \frac{dI_B}{dV_{B'C'}} \right|_{\Delta V_{B'E'}=0} \qquad \text{(A5.25)}$$

$$= \frac{qI_{SS}}{g_{m\mu}kT} \left[\frac{1}{\beta_{RM}} e^{\frac{qV_{B'C'}}{kT}} + \frac{C_4}{n_{CL}} e^{\frac{qV_{B'C'}}{n_{CL}kT}} \right] \qquad \text{(A5.26)}$$

$$\tau_{F_{ac}} \triangleq \frac{dQ_{DE}}{dI_{CC}} \qquad \text{(A5.27)}$$

$$\simeq B_{ac}\,\tau_{BL} + \tau_{1_{dc}} + \frac{B_{ac}\,I_{CC}}{I_K}\,\tau_{1L} \;\text{(Ignoring other components of } \tau_F\text{)} \qquad \text{(A5.28)}$$

$$\tau_{R_{ac}} \triangleq \frac{dI_B}{dI_C} \qquad \text{(A5.29)}$$

$$= \tau_{R_{dc}} = \tau_R \qquad \text{(A5.30)}$$

and all other parameters are as described in the GP model (Section 2.5).

The above equations are not as simple, and therefore not as intuitive, as those for the EM_3 model. However, they are basically equivalent.

The low-level injection solution for $g_{m\pi}$ is obtained by noting that for this case $q_2 \simeq 0$ and $q_b \simeq q_1$. Therefore

$$g_{m\pi_{\text{low-level}}} \simeq \frac{q\,I_{CC}}{kT} - \frac{I_{CT}}{q_b V_B} \qquad \text{(A5.31)}$$

which, for $V_B = \infty$ (as assumed in the EM_2 and EM_3 models), reduces to the EM_2 expression. A similar expression results for g_{m_μ} under low-level injection. At high current levels, putting $q_b \approx \sqrt{q_2}$, $q_1 \cdot q_2$ yields for ($V_{B'C'}=0$ and $V_B=\infty$)

$$g_{m_\pi \text{high-level}} \approx \frac{q\ I_{CC}}{2\ kT} \qquad \text{(A5.32)}$$

which is the same as obtained for the EM_3 model. Again, a similar expression results for g_{m_μ}. The formula given for g_{m_π} in Eq. (A5.21) assumes that B = 1 since a general expression for B has not been determined. Eq. (A5.23) (and Eq. (A5.26)) implies that B for the inverse region of operation is also unity.

The expression for $\beta_{F_{ac}}$ (which also assumes B = 1) reduces to a constant ($=\beta_{FM}$) when the non-ideal component of I_B is neglected (i.e., $C_2 = 0$) and when high-level injection effects are ignored (i.e., $I_K = \infty$). However, the behaviour of $\beta_{F_{ac}}$ as a function of I_{CC} is not readily inferred from Eq. (A5.24). Similarly for $\beta_{R_{ac}}$ and Eq. (A5.26).

Equation (A5.28) ignores the variation of the other components of τ_F ($\tau_{EB_{SCL}}$ and $\tau_{CB_{SCL}}$) while Eq. (A5.30) indicates that, as assumed for the GP model here, τ_R is a constant.

In the above equations, the effect of basewidth modulation and high-level injection are buried inside some terms (such as I_{CC}, I_{EC}, I_C, I_E, β_F, β_R, τ_{BL}, τ_{1L} and τ_R). Therefore there is no need for the (basewidth modulation) component r_o of the well-known hybrid-π model.[5] Also, since the GP model inherently assumes zero recombination in the base region, r_μ here only models the saturation current of a reverse-biased junction when the transistor is biased in the normal, active region. That r_μ consequently should not include basewidth modulation effects* can be seen from the definition of $g_\mu(= 1/r_\mu)$

$$g_\mu \triangleq \left.\frac{\Delta I_B}{\Delta V_{BC}}\right|_{\Delta V_{BE}=0} = \left.\frac{\Delta I_B}{\Delta I_C}\right|_{\Delta V_{BE}=0} \cdot \left.\frac{\Delta I_C}{\Delta V_{CE}}\right|_{\Delta V_{BE}=0} = 0 \cdot g_o \qquad \text{(A5.33)}$$

---

* Dr. G. A. Rigby, University of California, Berkeley, 1968, private communication.

APPENDIX - 6

INPUT PARAMETERS CROSS-REFERENCE FOR SLIC, SINC AND SPICE

This Appendix contains a cross-reference for the programs

SLIC (up to Version I)

SINC (up to Version D)

SPICE (up to Version 1Q)

This cross-reference gives the following information for each input parameter for these programs:

i) the notation used for this parameter in this book,

ii) the model level (EM_1, EM_2, EM_3 or GP) at which this parameter is incorporated,

iii) the page where the parameter is described in the theoretical description of the model, (Section 2),

iv) the page where the measurement of this parameter is described, (Section 3).

SLIC

PARAMETER	SYMBOL USED HERE	MODEL LEVEL	THEORY	MEASUREMENT
	(Note 1)			
BF = BFMAX	β_F or β_{FMAX}	EM_1, EM_3	15, 58	130, 188
ICMAX (or ICM)	I_{CMAX}	EM_3	58	188
BFLOW	β_{FLOW}	EM_3	58	188
ICLOW (or ICL)	I_{CLOW}	EM_3	58	188
BCEC	BCEC	EM_3	58	188
VCE	V_{CE}	EM_3	58	188
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
BR	β_R	EM_1	15	132
RO	r_o $(=1/g_o)$	EM_3	48	182
IC	I_C	EM_3	48	182
VBE	V_{BE}	(Note 2)	(Note 2)	(Note 2)
VCE	V_{CE}	(Note 2)	(Note 2)	(Note 2)
RB	r_b'	EM_2	28	151
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
RC	r_c'	EM_2	24	144
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
FT	f_T	EM_2	38	169
IC	I_C	EM_2	38	169
VCE	V_{CE}	EM_2	38	169

PARAMETER	SYMBOL USED HERE	MODEL LEVEL	THEORY	MEASUREMENT
LE/WB	L_E/W	EM_3	63	202
ICO	I_{CO}	EM_3	63	202
TSAT	τ_{SAT}	EM_2	38	176
CJE	C_{jE}	EM_2	29	165
VBE	V_{BE}	EM_2	29	165
PHIE	ϕ_E	EM_2	29	165
NE	m_E	EM_2	29	165
CJC	C_{jC}	EM_2	29	165
VBC	V_{BC}	EM_2	29	165
PHIC	ϕ_C	EM_2	29	165
NC	m_C	EM_2	29	165
RATIO	RATIO	EM_3	61	201
CSUB	C_{SUB}	EM_2	38	179
TEMP	T_{nom}	EM_1	21	136
ISS (Notes 2,4)	$I_S(0)$	EM_1	14	133
TF (Note 3)	$\tau_F(0)$	EM_2	33	169
VA (Note 3)	V_A	EM_3	44	182
Derived Parameters: BCC1	a_1	EM_3	59	
BCC2	a_2	EM_3	59	
BCC3	a_3	EM_3	59	
ETA	$\eta = kT/qV_A$ (Note 5)	EM_3	(Note 5)	
TAUF	$\tau_F(0)$	EM_2	33	
CJEO	C_{jEO}	EM_2	29	
CJCO	C_{jCO}	EM_2	29	

Note 1:

For constant β_F, only the BFMAX value is specified. This is the EM_1 model value. For $\beta_F(I_C)$, BFMAX is as described in the EM_3 model.

Note 2:

SLIC allows the user to specify the value of V_{BE} and V_{CE} as well as I_C at which r_o was measured (see p. 184). If this is done, $I_S(0)$ will be calculated from this data according to[31]

$$I_S(0) = I_C \left(1 - \frac{V_{BC}}{V_A}\right)\left(e^{\frac{qV_{BE}}{kT}} - 1\right)$$

The input parameter I_{SS} is then not required.

Note 3:

TF and VA are alternate input parameters. They can be specified instead of FT(IC,VCE) and RO(IC), respectively.

Note 4:

For the case of no basewidth modulation, ISS is I_S, as described in the EM_1 model. For finite basewidth modulation, ISS corresponds to $I_S(0)$ in the EM_3 model (p. 47).

Note 5:

η, which is simply another method of specifying V_A,[5] is not mentioned in the EM_3 description.

SINC

PARAMETER	SYMBOL USED HERE (Note 1)	MODEL LEVEL	THEORY	MEASUREMENT
BF = BFMAX	β_F or β_{FMAX}	EM_1, EM_3	15, 58	130, 188
ICMAX	I_{CMAX}	EM_3	58	188
BF	β_{FLOW}	EM_3	58	188
IC	I_{CLOW}	EM_3	58	188
VCE	V_{CE}	EM_3	58	188
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
BR	β_R	EM_1	15	132
ISS (Note 2)	$I_S(0)$	EM_1	14	133
RB	r'_b	EM_2	28	151
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
RC	r'_c	EM_2	24	144
TC1	TC_1	EM_3	67	204
TC2	TC_2	EM_3	67	204
RO	r_o $(=1/g_o)$	EM_3	48	182
IC	I_C	EM_3	48	182
FT	f_T	EM_2	38	169
IC	I_C	EM_2	38	169
VCE	V_{CE}	EM_2	38	169
TSAT	τ_{SAT}	EM_2	38	176
CJE	C_{jE}	EM_2	29	165
VBE	V_{BE}	EM_2	29	165
PHIE	ϕ_E	EM_2	29	165
NE	m_E	EM_2	29	165

PARAMETER	SYMBOL USED HERE	MODEL LEVEL	THEORY	MEASUREMENT
CJC	C_{jC}	EM_2	29	165
VBC	V_{BC}	EM_2	29	165
PHIC	ϕ_C	EM_2	29	165
NC	m_C	EM_2	29	165
RATIO	RATIO	EM_3	61	201
CSUB	C_{SUB}	EM_2	38	179
TEMP	T_{nom}	EM_1	21	136

Note 1:

For constant β_F, only the BFMAX value is specified. This is the EM_1 model value. For $\beta_F(I_C)$, BFMAX is as described in the EM_3 model.

Note 2:

For the case of no basewidth modulation, ISS is I_S, as described in the EM_1 model. For finite basewidth modulation, I_{SS} corresponds to $I_S(0)$ in the EM_3 model (p. 47).

SPICE

SPICE contains two BJT models:

a) Ebers-Moll Model: a simpler model that is virtually the EM_2 model;

b) Gummel-Poon Model: the more complex GP model.

Each of these is described separately, below.

EBERS-MOLL MODEL

PARAMETER	SYMBOL USED HERE	MODEL LEVEL	THEORY	MEASUREMENT
BF	β_F	EM_1	15	130
BR	β_R	EM_1	15	132
RB	r'_b	EM_2	28	151
RC	r'_c	EM_2	24	144
RE	r'_e	EM_2	27	140
CCS	C_{CS}	EM_2	38	179
TF	τ_F	EM_2	33	169
TR	τ_R	EM_2	33	176
CJE	C_{jEO}	EM_2	29	165
CJC	C_{jCO}	EM_2	29	165
IS	I_S	EM_1	14	133
PE	ϕ_E	EM_2	29	165
PC	ϕ_C	EM_2	29	165
VA	V_A	EM_3,GP	44, 96	182
EG	E_g	EM_1	21	137

GUMMEL-POON MODEL

PARAMETER	SYMBOL USED HERE	MODEL LEVEL	THEORY	MEASUREMENT
BFM	$\beta_{FM}(0)$	EM_3	50	188
BRM	$\beta_{RM}(0)$	EM_3	53	200
RB	r'_b	EM_2	28	151
RC	r'_c	EM_2	24	144
RE	r'_e	EM_2	27	140
CCS	C_{CS}	EM_2	38	179
TF	τ_F	EM_2	33	169
TR	τ_R	EM_2	33	176
CJE	C_{jEO}	EM_2	29	165
CJC	C_{jCO}	EM_2	29	165
IS	I_{SS}	GP	81	206
VA	V_A	GP, EM_3	96, 44	182
VB	V_B	GP	92	207
C2	C_2	EM_3	50	188
IK	I_K	GP	106	213
NE	n_{EL}	EM_3	50	188
C4	C_4	EM_3	53	200
IKR	I_{KR}	GP	106	215
NC	n_{CL}	EM_3	53	200
PE	ϕ_E	EM_2	29	165
ME	m_E	EM_2	29	165
PC	ϕ_C	EM_2	29	165
MC	m_C	EM_2	29	165
EG	E_g	EM_1	21	137

REFERENCES

(1) T. E. Idleman, F. S. Jenkins, W. J. McCalla and D. O. Pederson, "SLIC - A Simulator for Linear Integrated Circuits," IEEE J. Solid-State Circuits, Vol. SC-6, pp. 188-203, August 1971.

(2) S. P. Fan, D. O. Pederson, University of California, Berkeley. SINC, a Nonlinear DC and Transient Analysis Program, has been developed at the University of California, Berkeley. It is a descendant of TIME.(68)

(3) L. W. Nagel and D. O. Pederson, "Simulation Program With Integrated Circuit Emphasis (SPICE)," 16th Midwest Symposium on Circuit Theory, Waterloo, Ontario, April 12, 1973. Available as Electronics Research Laboratory Report No. ERL-M383, University of California, Berkeley, 12 April 1973. SPICE is a descendant of CANCER.(69) A detailed description of SPICE is also available in L. W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits," Electronics Research Laboratory Report No. ERL-M520, University of California, Berkeley, 9 May 1975.

(4) J. J. Ebers and J. L. Moll, "Large-Signal Behavior of Junction Transistors," Proc. IRE, Vol. 42, pp. 1761-1772, December 1954.

(5) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, Ch. 8, J. Wiley, 1964.

(6) D. O. Pederson, "Recent Developments in Electronic Circuit Simulators," International Symposium on Circuit Theory, Los Angeles, pp. 105-111, April 18-21, 1972.

(7) H. K. Gummel and H. C. Poon, "An Integral Charge Control Model of Bipolar Transistors," Bell Syst. Tech. J., Vol. 49, pp. 827-852, May 1970.

(8) D. A. Hodges and D. O. Pederson, "The Here and Now of Computer-Aided Circuit Design," ISSCC Digest of Technical Papers, pp. 38-39, February 1974. In this paper the authors describe the 11 bipolar transistor parameters that give adequate results in the majority of cases. Of these 11 parameters, 3 are EM_1 parameters, 7 are EM_2 parameters and 1 is an EM_3 (or GP) parameter.

(9) J. Logan, "Characterization and Modeling for Statistical Design," Bell Syst. Tech. J., Vol. 50, pp. 1105-1147, April 1971.

(10) J. J. Ebers and J. L. Moll,[4] or standard transistor texts such as P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, Ch. 9, J. Wiley, 1964.

(11) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, p. 43 and 74, J. Wiley, 1964.

(12, B. L. Hart, "Direct Verification of the Ebers-Moll Reciprocity Condition," Int. J. Electronics, Vol. 31, pp. 293-295, 1971.

(13) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, p. 181, J. Wiley, 1964.

(14) W. J. McCalla and W. G. Howard, Jr., "BIAS-3-A Program for the Nonlinear DC Analysis of Bipolar Transistor Circuits," IEEE J. Solid-State Circuits, Vol. SC-6, pp. 14-19, February 1971.

(15) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, p. 48, J. Wiley, 1964.

(16) A. B. Phillips, "Transistor Engineering," Section 7.5, McGraw-Hill, 1962.

(17) W. M. C. Sansen and R. G. Meyer, "Characterization and Measurement of the Base and Emitter Resistances of Bipolar Transistors," IEEE J. Solid-State Circuits, Vol. SC-7, pp. 492-498, December 1972.

(18) J. R. Hauser, "The Effects of Distributed Base Potential on Emitter-Current Injection Density and Effective Base Resistance for Stripe Transistor Geometries," IEEE Trans. Electron Devices, Vol. ED-11, pp. 238-242, May 1964.

(19) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC Vol. 2, Section 5.4, J. Wiley, 1964.

(20) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC Vol. 2, Section 2.3, J. Wiley, 1964. In this text, ϕ is given the symbol ψ_0 and is called the contact potential.

(21) W. Nuyts and R. J. Van Overstraeten, "Numerical Calculations of the Capacitance of Linearly Graded Si p-n Junctions," Electronics Letters, Vol. 5, pp. 54-55, 6th Feb. 1969 (errata on p. 174). See also Chawla and Gummel[22] who call the

barrier potential for capacitance, the "offset voltage." The effects on this potential of heavy doping is treated in R. J. Van Overstraeten, H. J. DeMan and R. P. Mertens, "Transport Equations in Heavy Doped Silicon," IEEE Trans. Electron Devices, Vol. ED-20, pp. 290-298, March 1973.

(22) B. R. Chawla and H. K. Gummel, "Transition Region Capacitance of Diffused p-n Junctions," IEEE Trans. Electron Devices, Vol. ED-18, pp. 178-195, March 1971.

(23) H. C. Poon and H. K. Gummel, "Modeling of Emitter Capacitance," Proc. IEEE (Lett.), Vol. 57, pp. 2181-2182, December 1969.

(24) J. A. Kerr and F. Berz, "The Effect of Emitter Doping Gradient on f_T in Microwave Bipolar Transistors," IEEE Trans. Electron Devices, Vol. ED-22, pp. 15-20, January 1975.

(25) H. J. DeMan and R. Mertens, "SITCAP-A Simulator of Bipolar Transistors for Computer-Aided Circuit Analysis Programs," ISSCC Digest of Technical Papers, pp. 104-105, February 1973.

(26) J. Lindmayer and C. Wrigley, "The High-Injection-Level Operation of Drift Transistors," Solid-State Electronics, Vol. 2, pp. 79-84, 1961.

(27) J. M. Early, "PNIP and NPIN Junction Transistor Triodes," Bell System Tech. J., Vol. 33, pp. 517-533, May 1954.

(28) J. L. Moll, "Large-signal Transient Response of Junction Transistors," Proc. IRE, Vol. 42, pp. 1773-1784, December 1954. (Note that the formula for τ_{SAT} here is not the same as in Reference 29. The charge-control theory used in Reference 29 is more accurate -- see, for example, H. V. Allen, "Analysis of a 6 MHz Oscillator Circuit for Ultrasonic Applications," contained in R. W. Dutton, "Techniques and Applications of Computer-Aided Circuit Simulation for Integrated Circuit and System Design, Part II: CAD Applications," Stanford Electronics Laboratories Technical Report, SEL-74-017, Stanford University, California, May 1974.)

(29) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, p. 228, J. Wiley, 1964. (τ_{SAT} corresponds here to τ_{SL}).

(30) M. J. Callahan, "Models for the Lateral P-N-P Transistor Including Substrate Interaction," IEEE Trans. Electron Devices, Vol. ED-19, pp. 122-123, January 1972.

(31) W. J. McCalla, "Computer-Aided Design of Integrated Bandpass Amplifiers," University of California, Berkeley, Ph.D. Dissertation, June 1972. Some typical temperature coefficients and SLIC input parameters are given in W. J. McCalla, "An Integrated IF Amplifier," IEEE J. Solid-State Circuits, Vol. SC-8, pp. 440-447, December 1973.

(32) J. M. Early, "Effects of Space-Charge Layer Widening in Junction Transistors," Proc. IRE, Vol. 40, pp. 1401-1406, November 1952.

(33) P. E. Gray, D. DeWitt, A. R. Boothroyd and J. F. Gibbons, "Physical Electronics and Circuit Models of Transistors," SEEC, Vol. 2, Section 9.1, J. Wiley, 1964.

(34) A. B. Phillips, "Transistor Engineering," Section 9.2, McGraw-Hill, 1962.

(35) C. A. Bittman, G. H. Wilson, R. J. Whittier and R. K. Waits, "Technology for the Design of Low-Power Circuits," IEEE J. Solid-State Circuits, Vol. SC-5, pp. 29-37, February 1970.

(36) C. T. Sah, "Effect of Surface Recombination and Channel on p-n Junction and Transistor Characteristics, "IRE Trans. Electron Devices, Vol. ED-9, pp. 94-108, January 1962.

(37) W. M. Webster, "On the Variation of Junction-Transistor Current-Amplification Factor with Emitter Current," Proc. IRE, Vol. 42, pp. 914-920, June 1954.

(38) J. R. Hauser, "Bipolar Transistors," in Fundamentals of Silicon Integrated Device Technology, Vol. 2, R. M. Berger and R. P. Donovan, Eds., Prentice-Hall, p. 132, 1968.

(39) C. T. Kirk, "A Theory of Transistor Cutoff Frequency (f_T) Falloff at High Current Densities," IRE Trans. Electron Devices, Vol. ED-9, pp. 164-174, March 1962. See also, H. C. Poon, H. K. Gummel and D. L. Scharfetter, "High Injection in Epitaxial Transistors," IEEE Trans. Electron Devices, Vol. ED-16, pp. 455-457, May 1969.

(40) A. van der Ziel and D. Agouridis, "The Cutoff Frequency Falloff in UHF Transistors at High Currents," Proc. IEEE (Lett.), Vol. 54, pp. 411-412, March 1966.

(41) R. J. Whittier and D. A. Tremere, "Current Gain and Cutoff Frequency Falloff at High Currents," IEEE Trans. Electron Devices, Vol. ED-16, pp. 39-57, January 1969.

(42) D. L. Bowler and F. A. Lindholm, "High Current Regimes in Transistor Collector Regions," IEEE Trans. Electron Devices, Vol. ED-20, pp. 257-263, March 1973.

(43) R. Kumar and L. P. Hunter, "Collector Capacitance and High-Level Injection Effects in Bipolar Transistors," IEEE Trans. Electron Devices, Vol. ED-22, pp. 51-60, February 1975.

(44) R. I. Dowell, "Automated Biasing of Integrated Circuits," University of California, Berkeley, Ph.D. Dissertation, April 1972.

(45) A. B. Phillips, "Transistor Engineering," p. 177, McGraw-hill, 1962.

(46) H. K. Gummel, "A Charge Control Relation for Bipolar Transistors," Bell Syst. Tech. J., Vol. 49, pp. 115-120, January 1970.

(47) G. Rey and J. P. Bailbe, "Some Aspects of Current Gain Variations in Bipolar Transistors," Solid-State Electronics, Vol. 17, pp. 1045-1057, October 1974.

(48) L. C. McBride, "Large-Signal Modeling of Bipolar Transistors for Computer-Aided Circuit Analysis," Stanford Electronics Laboratories, Technical Report 4825-6, August 1971.

(49) H. C. Poon and J. C. Meckwood, "Modeling of Avalanche Effect in Integral Charge Control Model," IEEE Trans. Electron Devices, Vol. ED-19, pp. 90-97, January 1972.

(50) R. W. Dutton, "Bipolar Transistor Modeling of Avalanche Generation for Computer Circuit Simulation," IEEE Trans. Electron Devices, Vol. ED-22, pp. 334-338, June 1975.

(51) R. D. Thornton, D. DeWitt, E. R. Chenette and P. E. Gray, "Characteristics and Limitations of Transistors," SEEC Vol. 4, p. 46, J. Wiley, 1966.

(52) L. A. Hahn, "The Effect of Collector Resistance Upon the High Current Capability of n-p-ν-n Transistors," IEEE Trans. Electron Devices, Vol. ED-16, pp. 654-656, July 1969.

(53) H. C. deGraaff, "Collector Models for Bipolar Transistors," Solid-State Electronics, Vol. 16, pp. 587-600, May 1973.

(54) See, for example,
M. B. Das and A. R. Boothroyd, "Determination of Physical Parameters of Diffusion and Drift Transistors," IRE Trans. Electron Devices, Vol. ED-8, pp. 15-30, January 1961.
R. D. Thornton, et al, "Handbook of Basic Transistor Circuits

and Measurements," SEEC, Vol. 7, Part Two, J. Wiley, 1966.
K. G. Ashar, H. N. Ghosh, A. W. Aldridge and L. J. Patterson[55]
R. Paul, "Measurement of Transistor Parameters," London Iliffe Books Ltd., London, 1969.
C. D. Root, "Circus Means Versatility as a CAD Program," Electronics, Vol. 43, pp. 86-96, February 2, 1970. (Also available in "Computer-aided Design," a collection of reprints from Electronics, McGraw-Hill, New York, 1971.)
J. Mar, "A Time-domain Method of Measuring Transistor Parameters," IEEE J. Solid-State Circuits, Vol. SC-6, pp. 223-226, August 1971.
C. L. Wilson, "Data Acquisition and Reduction Techniques for Device Modeling," Ch. 7 in "Semiconductor Device Modeling for Computer-Aided Design," Eds. G. A. Herskowitz and R. B. Schilling, McGraw-Hill, 1972.
H. E. Mussman[56]

(55) K. G. Ashar, H. N. Ghosh, A. W. Aldridge and L. J. Patterson, "Transient Analysis and Device Characterization of ACP Circuits," IBM J. Res. and Dev., Vol. 7, pp. 207-223, July 1963.

(56) H. E. Mussman, "Characterization and Model Parameter Determinations for Program SPICE," contained in R. W. Dutton, "Techniques and Applications of Computer-Aided Circuit Simulation for Integrated Circuit and System Design, Part II: CAD Applications," Stanford Electronics Laboratories Technical Report, SEL-74-017, Stanford University, California, May 1974.

(57) S. M. Sze, "Physics of Semiconductor Devices," pp. 20-21 (for semiconductors), pp. 397-399 (for Schottky barriers), J. Wiley, 1969.

(58) B. Kulke and S. L. Miller, "Accurate Measurement of Emitter and Collector Series Resistances in Transistors," Proc. IRE (Lett.), Vol. 45, p. 90, January 1957.

(59) J. Lindmayer, "Power Gain of Transistors at High Frequencies," Solid-State Electronics, Vol. 5, pp. 171-175, January 1962.

(60) P. Spiegel, "Transistor Base Resistance and its Effect on High Speed Switching," Solid State Design, pp. 15-18, December 1965.

(61) J. A. Strickland, "Time-Domain Reflectometry Measurements," Tektronix Measurement Concepts Publication, 1970.

(62) R. C. Jaeger and A. J. Brodersen, "Low-frequency Noise Sources in Bipolar Junction Transistors," IEEE Trans. Electron Devices, Vol. ED-17, pp. 128-134, February 1970.

(63) S. T. Hsu, "Noise in High-Gain Transistors and its Application to the Measurement of Certain Transistor Parameters," IEEE Trans. Electron Devices, Vol. ED-18, pp. 425-431, July 1971.

(64) D. D. Cohen and R. A. Zakarevicius, "Operational Amplifier Integrators for the Measurement of the Delay Times for Microwave Transistors," IEEE J. Solid-State Circuits, Vol. SC-10, pp. 19-27, February 1975.

(65) A. B. Phillips, "Transistor Engineering," p. 303, McGraw-Hill, 1962.

(66) C. L. Searle, A. R. Boothroyd, E. J. Angelo, Jr., P. E. Gray, and D. O. Pederson, "Elementary Circuit Properties of Transistors," SEEC, Vol. 3, p. 105, J. Wiley, 1964.

(67) I. E. Getreu, "Low-voltage, Micropower Integrated Circuit Amplifiers," University of California, Berkeley, Ph.D. Dissertation, Appendix C, pp. C23-C25, March 1972.

(68) F. S. Jenkins and S. P. Fan, "TIME - A Nonlinear DC and Time-Domain Circuit Simulation Program," IEEE J. Solid-State Circuits, Vol. SC-6, pp. 182-188, August 1971.

(69) L. W. Nagel and R. A. Rohrer, "Computer Analysis of Nonlinear Circuits, Excluding Radiation (CANCER)," IEEE J. Solid-State Circuits, Vol. SC-6, pp. 166-182, August 1971.

(70) G. C. M. Meijer and H. J. A. DeRonde, "Measurement of the Base Resistance of Bipolar Transistors," Electronics Letters, Vol. 11, pp. 249-250, 12 June 1975.

(71) J. Choma, "A Curve-Fitted Model for Bipolar Transistor f_T Roll-Off at High Injection Levels," IEEE J. Solid-State Circuits, Vol. SC-11, pp 346-348, April 1976.

(72) J. M. Pettit and M. M. McWhorter, "Electronic Amplifier Circuits," p. 307, McGraw-Hill, 1961.

(73) J. W. Slotboom and H. C. deGraaff, "Measurements of Bandgap Narrowing in Si Bipolar Transistors," Solid-State Electronics, Vol. 19, pp. 857-862, October 1976.

(74) J. S. Brugler, "Silicon Transistor Biasing for Linear Collector Current Temperature Dependence," IEEE J. Solid-State Circuits (Corresp.), Vol. SC-2, pp. 57-58, June 1967.