INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

**INVESTIGATION OF SiGe HETEROJUNCTION BIPOLAR
TRANSISTORS WITH RESPECT TO
COMPACT MODELLING
FOR INTEGRATED CIRCUIT DESIGN**

Hai Tran

A Thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of
Master of Electrical Engineering

Ottawa-Carleton Institute of Electrical Engineering
Department of Electronics
Carleton University
Ottawa, Canada

September 1997

©Copyright H. Tran 1997

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-27007-6

Canada

The undersigned hereby recommend to
the faculty of Graduate Studies and Research
acceptance of the thesis,

**INVESTIGATION OF SiGe HETEROJUNCTION BIPOLAR
TRANSISTORS WITH RESPECT TO
COMPACT MODELLING
FOR INTEGRATED CIRCUIT DESIGN**

submitted by

Hai Tran

in partial fulfilment of the requirements for the degree of
Master of Electrical Engineering

Thesis Supervisor

Chair, Department of Electronics

Carleton University
September 1997

Abstract

The thesis presents an investigation of SiGe HBTs with respect to compact modeling. The results from the I_C versus V_{CEx} (Output Characteristics) comparison indicate some mismatches in currents, in particular for the semi-physical High Current Model (HICUM) where the predicted collector currents are overestimating measured data for increasing V_{CEx} 's due to the reduced Early Effect at low to moderate current densities and the collector currents are underestimating measured data due to self-heating at high current densities. Consequently, for the defacto standard SPICE Gummel Poon (SGP) model, the reduced Early Effect is easily estimated from the fitting model parameter V_{AF} but the self-heating effect is not modeled. The Generalized Integral Charge-Control Relation (GICCR) representing a closed-form solution for the collector current for all operating regimes, which is valid for SiGe HBTs in contrast to the classical ICCR, is rederived and implemented into HICUM. Various model parameter extraction strategies, in particular for the modified GICCR model parameters and the lumped thermal resistance and capacitance found in the electrical equivalent model for self-heating in HICUM, based on DC and transient measurements are examined. A direct extraction technique and a method to simultaneously extract the emitter series resistance R_E and R_{TH} are developed. The compact model with the modified GICCR is capable of predicting accurately the following DC characteristics, I_C vs. V_{CEx} , I_B vs. V_{CEx} , and forward Gummel plot (I_C , I_B vs. V_{CEx}) for 1D numerical simulation and as well as for measured data. Good agreement was also obtained between simulated and measured data for the two important figures of merit f_T and f_{max} . Finally, the importance of an accurate model for this technology is applied to a simple current source circuit by performing several simulations using HICUM with a realistic SiGe HBT model parameter set. It was concluded that a large deviation in the output current from the reference current was observed with self-heating turned on, whereas a good match in the output current to the reference current with self-heating turned off was observed.

Acknowledgements

I would like to thank my thesis supervisor, Professor David J. Walkey for his encouragement and support throughout my master's program at Carleton University.

I also gratefully acknowledge the assistance and guidance of Dr. M. Schröter of Rockwell International. Special thanks are also due to Mr. Robert Hadaway, director of Technology Access & Applications Department of Nortel for financial support and research facilities, and to Dr. P. Schvan, manager, and members of Nortel scientific staff for their valuable advice and constructive discussions in my thesis work.

I also like to thank Micronet for the financial support and to help make this thesis possible.

I wish to thank my mother, brothers, and sister for their support, understanding, and inspiration. Finally, I gratefully express my sincerest gratitude to Ms. Gwen Fong for her unwavering support and patience.

Contents

	Page
List of Tables	viii
List of Figures	ix
List of Symbols	xiii
CHAPTER 1 - INTRODUCTION	
1.1 Introduction	1
1.2 Thesis Format	2
CHAPTER 2 - TECHNOLOGY	
2.1 Introduction	4
2.2 Device Description of SiGe HBT	4
2.2.1 Fabrication Limitations	6
2.2.2 Types of Doping Profiles	6
2.2.3 IBM SiGe NPN HBTs	7
2.3 Operational Principle of Si and SiGe HBTs	8
2.3.1 Assumptions	8
2.3.2 Qualitative Analysis	9
2.3.3 Quantitative Analysis	10
2.3.3.1 Material and Physical Effects	11
2.3.3.2 Theory of Generalized Integral Charge Control Relation(GICCR)	12
2.3.3.3 Transit Time	18
2.4 Numerical Simulation	21
2.4.1 DC Characteristics	22
2.4.2 AC Characteristics	24
CHAPTER 3 - COMPACT MODEL	
3.1 Introduction	27
3.2 Compact Models-SGPM and HICUM	28
3.2.1 SPICE Gummel-Poon Model	28
3.2.1.1 DC Current	29

3.2.1.2 DC Bulk Resistance	30
3.2.1.3 Junction Capacitances	31
3.2.1.4 Transit Time	32
3.2.2 HICUM	33
3.2.2.1 Quasi-Static Transfer Current	34
3.2.2.2 Transit Time	36
3.3 Determination of Model parameters	37
3.3.1 Extraction of SGPM Parameters	37
3.3.2 Extraction of HICUM Parameters	38
3.4 Device Electrical Characteristics	39
3.4.1 Determination of Transition Frequency and Maximum Oscillation Frequency	39
3.5 Measuring System Setup	42
3.5.1 Calibration and De-embedding of Parasitic/Interconnects	43
3.5.1.1 Calibration	43
3.5.1.2 De-embedding of Parasitics from DUT	44
3.6 Verification of Compact Models	44
3.6.1 Results and Discussion	47
 CHAPTER 4 - A NEW COMPACT MODEL	
4.1 Introduction	49
4.2 Verification of Weighting Factors	49
4.3 Compact Model with Modified GICCR	53
4.3.1 Extraction of Model Parameters	55
4.4 Results and Discussion	56
 CHAPTER 5 - EXTRACTION OF THERMAL PARAMETERS	
5.1 Introduction	61
5.2 MEDICI Simulation	62
5.2.1 Simulation Conditions	64
5.2.1.1 Results and Discussion	64
5.2.1.2 Electrical Characteristics	64
5.2.1.3 Physical Characteristics	67
5.3 A simple Thermal Subcircuit	69
5.3.1 Implementation of Self-Heating Model in a Compact Model	71

5.4 Extraction of Thermal Parameters	72
5.4.1 Extraction of Thermal Resistance	72
5.4.1.1 An Analytical Expression for R_{TH}	73
5.4.1.2 Derivation of A New Analytical for R_{TH}	74
5.4.1.3 Measurement Setup	77
5.4.1.4 Sensitivity Analysis	79
5.4.1.5 Verification of R_{TH}	82
5.4.1.6 Extraction of an “Isothermal” R_E	85
5.4.1.6.1 Evaluate Existing R_E Extraction Methods	86
5.4.1.6.2 Simultaneous Extraction of R_E & R_{TH}	89
5.4.1.6.3 Verification of R_{TH}-R_E Extraction Method	90
5.4.2 Extraction of Thermal Capacitance	96
5.4.2.1 Theory	96
5.4.2.2 Thermal Transient Measurement Setup	99
5.4.2.3 Extraction and Verification of C_{TH}	101
5.5 Scalable Thermal Parameters	103
5.6 Validation of Self-heating Model in Circuits	104
 CHAPTER 6 - CONCLUSION	
6.1 Summary and Discussion	108
6.2 Future Work	113
 APPENDICES	
A. Derivation of GICCR	115
B. Table of Model Parameters	119
C. Derivation of α_β	124
D. Simultaneous Extraction of R_{TH}-R_E	125
 REFERENCES	129

List of Tables

		Page
Table 3.1	S-Parameter Results for the standards	43
Table 3.2	Device Characteristics	45
Table 4.1	Weighting Factor	52
Table 4.2	Extraction Results of SiGe(Ge of 5%) and Si w/o Ge	57
Table 4.3	Comparison of Numerical and Direct Extraction Methods	58
Table 5.1	Simulation Cases	80
Table 5.2	R_E Comparison Using HICUM Data	90
Table 5.3	R_{TH} Extraction Using MEDICI Data	91
Table 5.4	Comparison Using Measurement Data	93
Table 5.5	Simulation Results for τ_{TH}	99

List of Figures

		Page
Figure 2.1	Cross-section of a Self-aligned SiGe Transistor	5
Figure 2.2	Drift SiGe 1D Doping Profile	6
Figure 2.3	Real SiGe 1D Doping Profile	7
Figure 2.4	IBM SiGe 1D Net Doping Profile	7
Figure 2.5	1D Ideal Net Doping Profile for a Diffusion Transistor	9
Figure 2.6	Energy Band Diagram of a Si BJT and SiGe HBT	10
Figure 2.7	Bandgap Energy of a realistic 1D IBM Doping Profile with and without Ge mole fraction	22
Figure 2.8	Gummel Plot	22
Figure 2.9	DC current Gain β vs. I_C	23
Figure 2.10	Output Characteristics (I_C vs. V_{CEx})	24
Figure 2.11	Transition Frequency	24
Figure 2.12a	Electric Field at $I_C=(1,3)\text{mA}$	25
Figure 2.12b	Electron Density at $I_C=(1,3)\text{mA}$	25
Figure 2.13	f_T vs. V_{CEx} at low I_C	26
Figure 3.1	SGPM Equivalent Circuit for vertical NPN or PNP	29
Figure 3.2	Base resistance as a function of collector current	30
Figure 3.3	Junction Capacitance vs. voltage (CV) curve	31
Figure 3.4	Transit Time as a function of collector current density	33
Figure 3.5	HICUM/LEVEL2 equivalent circuit	33
Figure 3.6	Transit Time vs. Collector current I_C	36
Figure 3.7	Gain versus Frequency	39
Figure 3.8	Small Signal Current Gain $ H_{21} $ vs. Frequency	40
Figure 3.9	DC and S-Parameter Measurement System	42
Figure 3.10	High Frequency DUT	43
Figure 3.11	Equivalent Circuit for a Pad/Interconnect	44
Figure 3.12	Gummel Plot for Transistor with $V_{CEx}=(0.8,1.6,2.4)\text{V}$ A_{E0} of $0.5 \times 10 \mu\text{m}^2$	45
Figure 3.13	Gummel Plot for Transistor with $V_{CEx}=(0.8,1.6,2.4)\text{V}$	45
Figure 3.14	Output Characteristics for Transistor	46
Figure 3.15a	$ H_{21} $ vs. Frequency at $V_{CEx}=1.6\text{V}$	46

Figure 3.15b	$UG_{\max}$ vs. frequency at $V_{CEx}=1.6V$	46
Figure 3.16a	f_T versus I_C for HICUM	46
Figure 3.16b	$f_{\max}$ versus I_C for HICUM	46
Figure 3.17a	f_T versus I_C for SGPM	47
Figure 3.17b	$f_{\max}$ versus I_C for SGPM	47
Figure 4.1a	h_{jc} vs. I_C for the four different Ge mole fraction profiles	51
Figure 4.1b	h_{je} vs. I_C for the four different Ge mole fraction profile	51
Figure 4.2	Ratio of collector currents SiGe HBT over Si BJT	52
Figure 4.3a	Comparison of 1D Device Simulation and GICCR for SiGe HBT graded Ge mole fraction of 5%	53
Figure 4.3b	Comparison of 1D Device Simulation and GICCR for SiGe HBT box type Ge mole fraction of 28%	53
Figure 4.4a	Extraction of c'_{10} and Q'_{p0} for SiGe(Ge 5%) and Si w/o Ge	57
Figure 4.4b	Extraction of h'_{jc} for SiGe(Ge 5%) and Si w/o Ge	57
Figure 4.5a	Comparison of 1D Simulation and Compact Model for Si without Ge	58
Figure 4.5b	Comparison of 1D Simulation and Compact Model for SiGe (graded Ge of 5%)	58
figure 4.6	Comparison of 1D Simulation and Compact Model for SiGe(box Ge of 28%) HBT	59
Figure 4.7a	LHS of (4.7) vs. Q_{jC}	59
Figure 4.7b	Output Characteristics for a SiGe HBT	59
Figure 5.1a	A Full cross-section of an advanced Si NPN transistor	62
Figure 5.1b	Zoom in of Figure 5.1a	62
Figure 5.2	Structure with Grid points	63
Figure 5.3	1D Net Doping Concentration	63
Figure 5.4	A Common Emitter Configuration	64
Figure 5.5	A Gummel Plot of a 0.5μm advanced silicon NPN Transistor	65
Figure 5.6	Beta(β) vs. I_C/A_E	66
Figure 5.7	Transition Frequency (f_T) vs. I_C/A_E	66
Figure 5.8	3D plot of Temperature Distribution	67
Figure 5.9a	Temperature Contour along the cross-section of the structure	68

Figure 5.9b	1D Temperature at $y=0.225\mu\text{m}$	68
Figure 5.10	Overview of heat flow through the semiconductor substrate	69
Figure 5.11	An Equivalent Electrical Subcircuit for Self-Heating Self-Heating model	70
Figure 5.12	Variable links between the BJT model and the Self-Heating Model	71
Figure 5.13	Circuit Schematic Measurement Setup	78
Figure 5.14	$\ln(I_B/I_{B0})$ vs. ΔP_T	79
Figure 5.15	Compensate for R_E in $(I_B-I_{B0})/I_{B0}$ vs. ΔP_T	82
Figure 5.16	$T_{j\text{max}}$ vs. Power Density	83
Figure 5.17	LHS(5.26) vs. I_E for an advanced Si NPN BJT	87
Figure 5.18	LHS(5.26) vs. I_E for SiGe	88
Figure 5.19	A Flow Chart of R_{TH} - R_E Extraction Algorithm	89
Figure 5.20	Large-Signal Switching Simulation	91
Figure 5.21	Gummel Plot	92
Figure 5.22	Beta (DC current gain) vs. I_C/A_E	92
Figure 5.23	$(I'_B-I'_{B0})/I'_B$ versus (P_T-P_{T0})	93
Figure 5.24a	Gummel Plots for $V_{CEx}=(0.8,2.4)V$ and a compact model HICUM using R_{TH} and R_E	93
Figure 5.24b	I_C vs. V_{BEx} for $V_{CEx}=(0.8,2.4)V$ and a compact model HICUM using simultaneous method	93
Figure 5.25	I_B vs. V_{CB}	94
Figure 5.26a	Transit time vs. I_C for A_{E0} for $0.5 \times 10 \mu\text{m}$ for HICUM with and without Self-Heating	94
Figure 5.26b	f_T versus I_C without Self-Heating	95
Figure 5.26c	f_T versus I_C with Self-Heating	95
Figure 5.26d	f_{max} versus I_C without Self-Heating	95
Figure 5.26e	f_{max} versus I_C with Self-Heating	95
Figure 5.27	Bias Schematic for Case 1	97
Figure 5.28	Bias Schematic for Case 2	97
Figure 5.29	Bias Schematic for Case 3	97
Figure 5.30	Normalized quantity vs. Time	98
Figure 5.31	Thermal Transient Measurement Setup	100

Figure 5.32	V_{BEx} vs time(s) for $I_B=35(A, V_{CE}=[1,3]V$	101
Figure 5.33	Thermal Transient Response for a high speed Si BJT	101
Figure 5.34	V_{BEx} vs. time(s) for $I_B=24\mu A, V_{CE}=[1,2]V$	102
Figure 5.35	zoom in part Fig. 5.33	102
Figure 5.36	Thermal Transient Response for a SiGe HBT	103
Figure 5.37	Scaling for R_{TH} for advanced Si BJTs	104
Figure 5.38	A Simple Current Source	105
Figure 5.39	Comparison of I_{REF}, I_0 with and without self-heating	105
Figure 5.40	Wilson Current Source	106
Figure 5.41	Relative Error in I_0 with respect to I_{REF}	106

List of Symbols

		Unit
1D	one-dimensional	-
2D	two-dimensional	-
3D	three-dimensional	-
T_{Jmax}	maximum base-collector junction temperature	K
T_0	ambient or room temperature	K
V_T	thermal voltage	V
P_T	total power dissipation	W
V_{gb0}	base bandgap narrowing voltage	V
V_{ge0}	emitter bandgap narrowing voltage	V
A_{E0}	emitter window area	μm^2
A_E	effective emitter area	μm^2
q	electronic charge (=1.602E-19)	C
h	Planck's constant(=6.63E-34 Js)	Js
k	Boltzman's Constant (=1.38E-23)	J/K
μ_{nB}	electron mobility in the base region	cm^2/Vs
μ_{pE}	hole mobility in the emitter region	cm^2/Vs
n_i	intrinsic carrier density	cm^{-3}
$p(x,.)$	total hole concentration in a transistor in 1D	cm^{-3}
E_V	valence energy band	eV
E_C	conduction energy band	eV

E_g	bandgap energy for Si(=1.12eV)	eV
ϕ_p	quasi-Fermi hole potential	V
J_{nx}	electron current density along the x-axis (1D)	A/ μm^2
N_B	net base doping in the neutral base region	cm^{-3}
N_E	net emitter doping in the neutral emitter region	cm^{-3}
n_C	electron minority carrier in the collector region	cm^{-3}
w_E	width of neutral emitter region	μm
w_B	width of neutral base region	μm
w_{Ge}	width of Ge mole fraction	μm
a_G	bandgap voltage per mole fraction (75mV per 10% Ge mole fraction)	V/Ge
$c_{\max}$	maximum Ge mole fraction	p.u.
m_{ac}	electric field due Ge mole fraction	V/cm
X_{jcB}	metallurgical junction of emitter-base	-
X_{jcB}	metallurgical junction of base-collector	-
BV_{CE0}	Breakdown Voltage of collector-emitter with base open circuited	V
SCR	space-charge region	-
h_{je}	model parameter to represent the weighting factor for the hole charge in the base-emitter junction	-
h_{jc}	model parameter to represent the weighting factor for the hole charge in the base-collector junction	-
h_{p0}	model parameter to represent the weighting factor for the hole charge in the neutral base region	-
h_C	model parameter to represent the lumped weighting factors (h_{fC} , h_{fB} , and h_{fE}) for the minority hole charge	-
Q_P	total hole charge in a transistor	C

Q_f	minority hole charge in a transistor	C
Q_{JE}	hole charge in the base-emitter junction	C
Q_{JC}	hole charge in the base-collector junction	C
Q_f	minority hole charge in a transistor	C
v_C	velocity of carriers at base-collector SCR	cm/s
v_s	saturation velocity at base-collector SCR ($1E7$ cm/s)	cm/s
β	DC current gain for common-emitter configuration	-
f_T	transition frequency or cutoff frequency	Hz
f_{max}	maximum oscillation frequency	Hz
τ_f	transit time of transistor	s
τ_{pE}	transit time of neutral emitter region	s
τ_{pC}	transit time of neutral collector region	s
τ_{Bf}	transit time of neutral base region	s
τ_{BE}	transit time of emitter-base space charge region	s
τ_{BC}	transit time of base-collector space charge region	s
$ H_{21} $	small signal current gain for a common-emitter configuration	-
$ UG_{max} $	maximum unilateral power gain for common-emitter configuration	-
k	stability factor	-
I_C	collector current	A
I_B	base current	A
I_E	emitter current	A
I_S	substrate current	A
I_T	transfer current	A

V_{BEX}	terminal base-emitter voltage	V
V_{CEX}	terminal collector-emitter voltage	V
V_{BCX}	terminal base-collector voltage	V
DUT	device-under test	-
HICUM	High Current Model	-
SGPM	SPICE Gummel-Poon Model	-
ICCR	Integral Charge-Control Relation	-
GICCR	Generalized Integral Charge-Control Relation	-
HBT	Heterojunction Bipolar Transistor	-
BJT	Bipolar Junction Transistor	-
SIC	Selectively Implanted Collector	-

Chapter 1 - Introduction

1.1 Introduction

The thriving market in recent years for wireless systems in applications such as communications, computers, and transportation systems that demand high performance and low cost, 800MHz to 80GHz radio frequency (RF) components has prompted renewed demand for more advanced high-speed transistors. Conventional wisdom has it that silicon homojunction bipolar integrated circuit (IC) technology dominates the market with applications operating in several hundred MHz range, while Gallium Arsenide (GaAs) technology dominates applications operating above 1GHz. The arrival of commercial silicon-germanium base (SiGe) heterojunction bipolar transistor (HBT) technology tailored to low voltage, low power consumption, and comparable performance to that of GaAs technology in RF and microwave mixed signal applications provides an alternative solution. SiGe HBTs in ICs can be easily achieved with the cost and integration advantages similar to that of silicon homojunction BJT. Furthermore, the SiGe HBT technology has been successfully demonstrated in numerous RF and mixed signal circuits by circuit designers, in particular at IBM, where they have successfully designed a 12 bit digital-to-analog converter that processes data at 1.0 gigasample per second[1].

With silicon SiGe-based HBT technology becoming available to mainstream circuit designers, modeling of this technology becomes an important consideration. Furthermore, the strong time-to-market pressure emphasis in recent years has made accurate IC modeling or CAD models a major requirement for product success. The need to verify whether the existing compact models, such as the well known SPICE Gummel Poon Model (SGPM) found in HSPICE[2] and other circuit simulators, and the more semi-physical compact models such as High Current Model (HICUM)[3] and Most Exquisite Transistor Model (MEXTRAM) found in some companies and universities, can be successfully applied to the new technology is an important issue.

The motivation for the thesis work is to use the defacto standard SGPM and HICUM, a semi-physical compact model, to model silicon SiGe-based NPN HBTs. Theory based on

the Generalized Integral Charge Control Relation (GICCR) is rederived to illustrate the physical operation of the device, and extensions to the existing compact models must be implemented to correctly account for the observed effects. Verification of the theory is accomplished by using one and two dimensional (1D and 2D, respectively) numerical simulations based on a realistic doping profile. In addition, measurements of several double emitter-poly self-aligned SiGe HBTs of different sizes are performed to verify the theory through comparisons of simulated results from compact models to measured results. Finally, the importance of accurate modelling of this technology is demonstrated with a simple current source circuit.

1.2 Thesis Format

The thesis begins with a comparison between silicon and SiGe bipolar transistors in terms of electrical performance by using the ICCR theory discussed in Chapter 2. Moreover, the strength and limitations of Si and SiGe transistors are verified based on various figures of merit and electrical characteristics obtained from one-dimensional numerically simulated results.

In Chapter 3, a brief discussion of various compact models from a physical basis perspective is examined followed by a detailed examination of the theoretical background of the existing SPICE Gummel Poon Model (SGPM) and HICUM. Some model parameter extraction methods are briefly discussed for the two compact models. Measurement procedures that include de-embedding of parasitics in S-parameter measurements are discussed. Finally, several comparisons using various figures of merit between simulation and measurement for the SiGe technology are performed for each of the compact model.

In Chapter 4, the modified GICCR discussed in Chapter 2 is implemented in HICUM to account for the small output conductance observed in this technology. Direct extraction methods are developed to determine the new model parameters and these methods are verified by directly calculating for these new parameters using 1D numerical data. Finally, the methods are applied to measured data and the results are confirmed by comparing simulation to measurement with the new compact model.

Chapter 5 investigates the influence of self-heating in advanced Si BJTs and in particular, SiGe-based HBTs. A brief discussion on the derivation of a simple thermal network from a one-dimensional heat-flow equation is examined. A detailed discussion on the extraction procedures of thermal parameters is examined followed by verification of extracted parameters by comparing simulation to measurement for both Si and SiGe-based NPN bipolar transistor technologies. The scalable thermal parameters as well as the validation of self-heating model in a simple current source circuit are briefly discussed.

Finally, Chapter 6 summaries the thesis work. In addition, some recommendations regarding future work are discussed.

Chapter 2 - Technology

2.1 Introduction

In this chapter, some qualitative and quantitative analyses are provided to give some insights into the operating behavior and properties of SiGe heterojunction and Si homojunction bipolar technologies. The strengths and limitations of both technologies are verified by using DC characteristics, such as the base and collector currents (Gummel plot), DC current gain (β) and figures of merit, such as transition frequency or cutoff frequency (f_T) obtained from one-dimensional (1D) numerical simulation for a realistic doping profile. The understanding of the results from the analyses is important in the modelling of SiGe HBT technology. Before these are discussed, a brief overview of the SiGe HBT device structure and the types of doping profiles used is examined.

2.2 Device Description of SiGe HBT

2.2.1 Fabrication Limitations

Bandgap engineering is the process of adjusting the material's energy bandgap to the designer's needs, largely to achieve higher speed and higher DC current gain which will be discussed later. Combining either germanium or carbon elements to the main silicon material can effectively change the bandgap energy of the host material. With silicon and germanium elements being chemically compatible, they can be intermixed to form a stable alloy $\text{Si}_{c_{\max}}\text{Ge}_{1-c_{\max}}$, where $c_{\max}$ represents the maximum Ge mole fraction, usually written as SiGe for simplicity. Some experimental SiGe transistors are usually formed in a mesa-structure where the process is similar to creating GaAs MESFETS. An alternative method to fabricating a SiGe transistor is to employ a technique developed by IBM that would enable a graded Ge mole fraction to be diffused into the base, that is, basically an addition to the existing Si technology base[1]. This technology allows full compatibility with the existing CMOS process technology. Fig 2.1 shows a cross-section of the polysilicon emitter self-aligned SiGe HBT developed by IBM.

FIGURE 2.1 Cross-section of a Self-aligned SiGe Transistor (after Cressler J. D. [4])

An inherent problem associated with the fabrication of the SiGe transistor is the stability of the SiGe alloy. The growth of epitaxial films from SiGe alloy is restrained by the inherent lattice constant mismatch between silicon and germanium, where the lattice constant of the germanium is about 4 percent larger than silicon. The process of growing a perfect SiGe film crystal on an epitaxial base by adopting the smaller lattice constant, also known as strained-layer epitaxial, can induce compressive strain in the overlying film. Strained-layer epitaxy is only thermodynamically stable depending on the film's effective strain (percent of Ge mole fraction) and thickness. If the strained-layer epitaxy becomes unstable, crystal defects or amalgam can occur which can create recombination centers resulting in SRH(Shockley Read Hall) recombination effect. and therefore, the transistor performance suffers.

Different types of Germanium mole fraction profile, such as trapezoidal, uniform or box, and triangular shapes have been explored by Richey[5]. It has been determined that the trapezoidal Ge profile gives the best performance since it is a compromise between the triangular profile, where it has the largest enhancements in cutoff frequency due to the steepest possible Ge mole fraction grading across the base that induces a built-in electrostatic field to aid the flow of minority carriers, and the box profile which yields the largest improvement in DC current gain due to a lower bandgap in the base.

2.2.2 Types of Doping Profile

Two common types of doping profiles used to realize SiGe HBTs are the “real” and “drift” SiGe transistors. The “drift” SiGe HBT is a transistor that does not have an abrupt type of doping profile between the metallurgical junctions, but rather an exponential profile which is similar to the one found in an advanced self-aligned Si bipolar transistor is used. Fig. 2.2 shows a typical drift SiGe HBT containing a one-dimensional (1D) net doping profile as a function of depth, x-direction and a triangular Ge mole fraction profile in the base region.

FIGURE 2.2 Drift SiGe 1D Doping Profile (after Schroter et al. [6])

The “real” SiGe transistors have the following distinctive characteristics. Usually, a box profile with Ge mole fraction of 25 to 35% [7] is employed in the base that enable the base to be highly doped without any degradation caused by the bandgap narrowing effect. The high doped base lower the base sheet resistance that can result in a significant increase in maximum oscillation frequency, an important figure of merit in RF circuit and a lower noise figure. In addition, the doping in the emitter near the base-emitter space charge region is considerably lower than the base doping. The reason is to increase the base-emitter breakdown voltage since breakdown voltage is inversely proportional to doping. Fig. 2.3 shows a typical “real” SiGe HBT net doping profile.

FIGURE 2.3 Real SiGe 1D Doping Profile (after Schroter et al. [6])

2.2.3 IBM SiGe NPN HBTs

An IBM SiGe NPN HBT employs a polysilicon emitter self-aligned technology with a single base, emitter, and collector contact. The cross-section of the device is similar to that shown in Fig. 2.1. Silicides are used to connect from the single base contact to the base on the other side of the emitter contact. In addition, deep trench structures are used to isolate adjacent transistors. The transistor has a drift like net doping profile, a trapezoidal Ge profile with about 5 percent mole fraction graded over the intrinsic base region or base region directly underneath the emitter window, and a selectively ion-implanted collector (SIC) that is, the doping is implanted in the collector of the intrinsic region of the transistor, usually, under the emitter window, and the result is an exponential profile in the collector. A similar one-dimensional net doping profile for the technology can be found in [1]. The net doping profile is captured by fitting analytical expressions which will be used in 1D numerical simulation and the result is shown in Fig. 2.4.

FIGURE 2.4 IBM SiGe 1D Net Doping Profile

The profile is very similar to the one in Fig.2.2 except that a 5 percent Ge mole fraction and a selectively implanted collector are used instead. A SIC instead of a constant collector doping can achieve some improvements in performance such as a reduced transit time and an increased transition frequency, an increased BV_{CE0} (collector-emitter break-down voltage with base open circuit), and a reduced collector resistance. In a constant collector doping profile, a low collector resistance achieved by high doping is required to reduce the RC time constant which ultimately reduces the overall transit time. But the high doping can effectively lower BV_{CE0} which can affect the circuit performance. Thus, some compromises have to be made with respect to the collector doping.

2.3 Operational Principle of Si and SiGe HBTs

2.3.1 Assumptions

The following assumptions are required in the qualitative discussion and in the derivation of analytical expressions.

- i. - A 1D NPN transistor with an ideal abrupt doping profile and a triangular Ge mole fraction profile shown in Fig. 2.10 is used. The dark solid line and dark dashed-line represent electron and hole minority carriers injecting from the emitter to collector via the base and base into emitter, respectively. The reason for using a 1D transistor is to focus some of the important effects observed with SiGe alloy in the analysis without having to take into account the series resistances, current spreading, and other effects that occur in a realistic two-dimensional (2D) or three-dimensional (3D) transistor because these effects are in essence simply superpositions to the 1D transistor.

FIGURE 2.5 1D Ideal Net Doping Profile for a Diffusion Transistor

- ii. - An average value is assumed for the electron mobility μ_n in all regions of the transistor. The added SiGe alloy in the neutral base can change the electron mobility but the change would be small since the amount of Ge mole fraction is small within 5 to 10 percent, so in practice the assumption is still valid.
- iii. - The time derivative of the continuity equations is zero.
- iv. - The influence of the recombination term in the continuity equation is negligible within the neutral base and collector region, that is $I_C = I_p + I_n \approx I_n$, $I_p = 0$ for negligible recombination current; I_p and I_n are transport equations for holes and electrons, respectively.

2.3.2 Qualitative Analysis

The operation of NPN Si BJT and SiGe HBT is identical in nature. Some differences such as enhanced built-in electrostatic field and a higher electron mobility in the SiGe transistor can improve the performance over the Si transistor. This can be explained by using an ideal energy band diagram shown in Fig. 2.6 based on assumption (i). The deviation of the energy conduction band shown as a dashed line in the base region is a result of a graded Ge in the base region. Under DC forward-biased mode that is, base-emitter forward biased and base-collector reverse biased, electrons are injected from the emitter into

base travelling toward the collector via the conduction energy(E_C) band, whereas holes are injected from the base toward the emitter via the valance energy(E_V) band. For an assumed ideal 1D abrupt doping profile Si transistor, electron injection is caused by diffusion that obeys Fick's Law, whereas for the SiGe transistor, injection of electron is caused by a combined diffusion and drift where drift is associated with the built-in electric field created by graded Ge mole fraction in the base region that induces more minority electron carriers to travel toward the collector.

FIGURE 2.6 Energy Band Diagram of a Si BJT and SiGe HBT(dashed line in the base) with solid line for $V_{BE}=0$ and $V_{BC}=0$ and dash-dotted line for $V_{BE} > 0$ and $V_{BC} < 0$.

The result is a higher collector current, and subsequently, a higher DC current gain since the injection of holes from the base into the emitter remains the same for both Si and SiGe transistors. Furthermore, with the graded Ge mole fraction extended to the base-collector junction, a lower output conductance is expected. In terms of AC response, the built-in field in the neutral base of SiGe transistor reduces the base transit time which in turns decreases the overall transit time since the base transit time is a major contribution to the overall transit time, and in contrast without the built-in field in the base of the Si transistor, the base transit time is expected to be larger than the SiGe transistor.

2.3.3 Quantitative Analysis

Four main advantages that SiGe HBTs have over Si BJTs are increase in collector current density, DC current gain, output resistance or lower output conductance, that is higher Early voltage, and a reduction in the base transit time(only for a graded Ge profile) result-

ing in higher a transition frequency. The first three advantages can be illustrated from the analytical collector current expression derived from the Generalized Integral Charge Control Relation (GICCR)[6], and the last advantage can be best illustrated from the base transit time analytical expression derived from the classical approach, but first a brief overview on the influence on SiGe alloy on the basic physical effects is discussed.

2.3.3.1 Material and Physical Effects

The bandgap energy reduction or offset is 75 meV per 10% mole fraction as indicated in [1]. The change in bandgap voltage is linearly dependent on the spatial variation of the Ge mole fraction which is given as[8]

$$\Delta V_{Ge}(C(x)) = a_G C(x) = a_G \cdot c \cdot x \quad (2.1)$$

where $C(x) = c \cdot x$ is the mole fraction as a function of x or depth; c is mole fraction per micron or $c = c_{\max}/w_{Ge}$ with $c_{\max}$ equal to the maximum mole fraction and w_{Ge} equal to the width of the Ge profile; and a_G is the bandgap voltage per mole fraction which is equal to 0.075 V per 10% Ge mole fraction. Taking into account the Ge mole fraction, the bandgap energy is given as

$$E_{GSiGe} = E_{GSi} - \frac{\Delta V_{Ge}}{q} \quad (2.2)$$

The expression shows that the bandgap energy reduces with increasing Ge mole fraction. The built-in electric field given in (2.3) is obtained by taking the derivative of $q \cdot E_{GSiGe}$ with respect to x .

$$\frac{d}{dx}(qE_{GSiGe}) = -m_{ac} \quad (2.3)$$

where $m_{ac} = a_G \cdot c$. This negative electric field in the base region is used to aid the flow of electron carriers toward the collector. For an uniform Ge profile the built-in field is zero according to (2.3). The reduction in bangap energy has the largest effect on the intrinsic carrier concentrations n_i in the base region which is expressed as

$$n_i^2 = n_{i0}^2 \exp\left(\frac{\Delta V_G}{V_T}\right) \quad (2.4)$$

where n_{i0}^2 is the normal intrinsic carrier concentration and ΔV_G is the sum of the bandgap narrowing voltage ΔV_{gb} and Ge mole fraction voltage ΔV_{Ge} in the base region which is given as

$$\Delta V_G = \Delta V_{gb} + \Delta V_{Ge}(x). \quad (2.5)$$

Substituting (2.5) into (2.4), the general intrinsic carrier density equation becomes

$$n_{iB}^2 = n_{i0}^2 \exp\left(\frac{\Delta V_{gb} + \Delta V_{Ge}(x)}{V_T}\right) = n_{iB0}^2 \exp(\Delta V_{Ge}(x)/V_T) \quad (2.6a)$$

or

$$n_{iB}^2 = n_{iB0}^2 \exp((m_{ac}x)/V_T) \quad (2.6b)$$

with

$$n_{iB0}^2 = n_{i0}^2 \exp\left(\frac{\Delta V_{gb}}{V_T}\right) \quad (2.6c)$$

where n_{iB0}^2 is the instrinsic carrier density with bandgap narrowing effect. The instrinsic carrier density described in (2.6b) is higher with the Ge mole fraction than without the Ge mole fraction, and this result translates in a larger collector current which will be shown later. Another advantage of the graded Ge mole fraction in the neutral base is the increased in electron mobility discussed in[9].

2.3.3.2 Theory of Generalized Integral Charge Control Relation (GICCR)

A closed-form solution of collector current densities for all operating regimes under DC bias conditions for high-speed bipolar transistors was difficult to achieve when using the classical theory, also known as Regional Approach[3], which can be used to obtain closed-form analytical solutions for only low and high injection operating conditions. However, a closed-form solution for all operating regimes was possible by utilizing the Integral Charge Control Relation (ICCR) which was first proposed by Gummel in 1970[10]. The ICCR was used as a basis in the development of compact models for bipolar transistors. The theory was subsequently extended to account for silicon-based heterostructure bipolar transistors by generalizing ICCR which became known as GICCR

in Schroter et al.[6]. The derivation of the transfer current using GICCR for HBTs resulted in weighting functions due to added SiGe alloys which only affected the minority hole charge for high current density operation. However, the influence of the added alloy for instance, SiGe alloy in the base region of the transistor, also had an impact on the output conductance or forward Early effect under low to medium current density operations which was not discussed in the derivation. The transfer current is now revisited to take into account the Early effect by using the original GICCR as a basis, and note that, the final expression can also be applied to a compact model which will be discussed in Chapter 4. Furthermore, the final expression can also be used to illustrate some of the advantages found in SiGe over Si transistors.

The derivation of the transfer current equation begins with the general form for the collector or transfer current of a 1D NPN transistor which is given as

$$I_T = \frac{qA_{E0}V_T(\exp(V_{BEi}/V_T) - \exp(V_{BCi}/V_T))}{\int h(x)p(x)dx} \quad (2.7a)$$

where $p(x)$ is the total hole concentration, $h(x)$ represents the weighting function, A_{E0} is the emitter window area, V_T is the thermal voltage, and V_{BEi} and V_{BCi} are the internal base-emitter voltage and base-collector voltage, respectively. The weighting function is defined as

$$h(x) = \frac{\exp\left(\frac{V_{BE} - \phi_p}{V_T}\right)}{\mu_n n_i^2 \frac{|J_{nx}|A_{E0}}{I_T}} \quad (2.7b)$$

where ϕ_p is the quasi-Fermi hole potential and J_{nx} is the electron current density along the x-axis. The weighting function can be simplified by assigning the emitter contact as the potential reference, and in this case, it is at zero volt. This forces ϕ_p to be equal to the terminal voltage base-emitter V_{BE} , and thus, the exponential in the denominator of (2.7b) is equal to one. Also, the term $|J_{nx}|A_{E0}/I_T$ is reduced to one since it is a 1D transistor where the current is only flowing in x direction which is exactly equal to the transfer current. The resulting weighting function becomes

$$h(x) = \frac{1}{\mu_n n_i^2} \quad (2.8)$$

The integration range of the denominator of (2.7a) is the entire transistor region emitter, base, and collector, rather than restricting the integration to the neutral base region as discussed in Kroemer[11]. The reason being under medium and high injection regime of operations, minority hole carriers in the emitter and collector regions become more dominant and errors in transfer current can result.

The results of the integration solved for the entire transistor region discussed in Appendix A for SiGe and Si technologies are as follows. The modified GICCR transfer current expression for SiGe HBTs becomes

$$I_{TSiGe} = \frac{(qA_{E0})^2 \overline{\mu_{nB} n_{iB0}^2} V_T \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right)}{Q_P^*} \left[\exp\left(\frac{V_{BE}}{V_T}\right) - \exp\left(\frac{V_{BC}}{V_T}\right) \right] \quad (2.9a)$$

with Q_P^* . the total hole charge, is a function of DC bias

$$Q_P^* = Q_{p0}^* + Q_{jE}^* + Q_{jC}^* + Q_f^* \quad (2.9b)$$

and $Q_{p0}^* = h_{p0} Q_{p0}$, $Q_{jE}^* = h_{je} Q_{jE}$, $Q_{jC}^* = h_{jc} Q_{jC}$. The corresponding weighting factors are

$$h_{je} = \frac{1}{m_{ac} w_{jE} / V_T} \left(\exp\left(\frac{m_{ac} w_{jE}}{V_T}\right) - 1 \right) \quad (2.9c)$$

and

$$h_{p0} = h_{fB} = \frac{1}{m_{ac} w_B / V_T} \left(1 - \exp\left(\frac{-m_{ac} w_B}{V_T}\right) \right) \quad (2.9d)$$

and

$$h_{jc} = \exp\left(\frac{-m_{ac} w_B}{V_T}\right) \cdot \left(1 + \frac{m_{ac} w_{jC}}{2V_T} \right) \quad (2.9e)$$

Q_f^* is the minority hole charge which is given as

$$Q_f^* = h_{fE}^* Q_{fE} + h_{fB} Q_{fB} + h_{fC}^* Q_{fC} \quad (2.9f)$$

where Q_{fE} , Q_{fB} , and Q_{fC} are the minority hole charge in the emitter, base, and collector, respectively, and the corresponding weighting factors are as follows:

$$h_{fE}^* = \frac{\overline{\mu_{nB} n_{iB0}^2}}{\mu_{nE} n_{iE}^2} \cdot \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right) \quad (2.9g)$$

and

$$h_{fC}^* = \frac{\overline{\mu_{nB} n_{iB0}^2}}{\mu_{nC} n_{iC}^2} \cdot \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right) \quad (2.9h)$$

Q_{p0} , Q_{jE} , Q_{jC} , and Q_f are zero bias hole charge, base-emitter hole charge, base-collector hole charge, and minority hole charge of the entire transistor, respectively. h_{je} , h_{jc} , h_{p0} , h_{fB} , h_{fE}^* , and h_{fC}^* are weighting factors and they are assumed to be independent of bias because according to the analytical expression for each weighting factors, the only parameter that is a function of voltage bias is the width of the space-charge, but they are a weak function of bias such that they can assume constant, and more will be discussed in Chapter 4. In addition, h_{jc} and h_{p0} are usually less than one, whereas h_{je} is approximately equal to one. To quickly illustrate this point, using the following data at zero bias condition (note the subscript “0”) from Fig.2.4 and assuming the Ge profile as a triangular shape, the data are $w_{jE0}=0.001$, $w_{B0}=0.04$, $w_{jC0}=0.004$, $V_T=0.02586V$, $a_c=0.75V/\text{mole}$ fraction, and $c=c_{\max}/w_{GE}$ with $c_{\max} = 0.05$ mole and $w_{GE}=0.045$. The calculated values are $h_{je}=1.014$, $h_{p0}=0.6$, and $h_{jc}=0.35$. More detail analysis will be discussed in Chapter 4.

The original GICCR transfer current equation for Si BJT is determined by setting m_{ac} equal to zero that reduces the values of the weighting factors h_{p0} , h_{fB} , and h_{jc} to one. The result becomes

$$I_{TSi} = \frac{(qA_{E0})^2 \overline{\mu_{nB} n_{iB0}^2} V_T}{Q_P} \left[\exp\left(\frac{V_{BE}}{V_T}\right) - \exp\left(\frac{V_{BC}}{V_T}\right) \right] \quad (2.10a)$$

with

$$Q_P = Q_{P0} + Q_{jE} + Q_{jC} + Q_f \quad (2.10b)$$

and Q_f , the minority hole charge, given as

$$Q_f = h_{fE}Q_{jE} + Q_{fB} + h_{fC}Q_{jC} \quad (2.10c)$$

where

$$h_{fE} = \frac{\mu_{nB}n_{iB0}^2}{\mu_{nE}n_{iE}^2} \quad (2.10d)$$

and

$$h_{fC} = \frac{\mu_{nB}n_{iB0}^2}{\mu_{nC}n_{iC}^2} \quad (2.10e)$$

Again, the hole charges and weighting factors in the above equations retain their usual definitions.

The final transfer current expressions (2.9a) (SiGe HBT) and (2.10a) (Si BJT) illustrate two very important results; the transfer current for SiGe HBT is much larger than the current in Si BJT and the Early Effect is at a minimum or output conductance is smaller compared to Si BJT. The first result (2.11) is obtained by taking the ratio of (2.9a) over (2.10a). The result in (2.11) shows that the ratio is greater than one since Q_P^* is much less than Q_P because the weighting factors in Q_P^* are less than one as mentioned earlier.

$$\frac{I_{TSiGe}}{I_{TSi}} = \frac{Q_P}{Q_P^*} \cdot \exp\left(\frac{m_{ac}X_{jeB}}{V_T}\right) > 1 \quad (2.11)$$

The Early effect describes the variation in the collector current (transfer current) as a direct result of the modulation (changing) of the neutral base region or the variation in the base width w_B [12]. The change in w_B can occur by varying the width of the space-charge region on either side of the neutral base region. To distinguish which side of the space-charge region that causes w_B to modulate, model parameters V_{AR} (Reverse Early Voltage) and V_{AF} (Forward Early Voltage) in SPICE Gummel Poon (SGP) model, a compact model

for bipolar transistor, are used to describe the modulation of w_B from the base-emitter and base-collector space-charge region, respectively. Furthermore, the variations of base-emitter space-charge region (BE-SCR) and base-collector space-charge region (BC-SCR) are directly related to Q_{jE} (Q_{jE}^*) and Q_{jC} (Q_{jC}^*). Hence, the reverse Early voltage for both SiGe and Si transistors is the same since $Q_{jE} = Q_{jE}^*$ for $h_{je}=1$. However, the forward Early voltage is different in the two technologies since $Q_{jC} > Q_{jC}^*$ for $h_{jC} < 1$. The forward Early effect can be observed indirectly by looking at the output conductance, that is the slope of the output characteristics (I_C versus V_{CE}) which is obtained by varying the terminal collector emitter voltage (V_{CEx}) with a fixed terminal base-emitter voltage (V_{BEx}). The output conductance is dependent on the base-collector charge Q_{jC} and the amount of variation is achieved by varying V_{CE} with V_{BE} fixed since the variation of Q_{jC} can effectively change the transfer current via the total hole charge. In other words, the change in the collector current can be kept small by keeping the change in $Q_p(Q_p^*)$ small, that is $Q_{jC}(Q_{jC}^*) \ll Q_{p0}(Q_{p0}^*) + Q_{jE}(Q_{jE}^*) + Q_f(Q_f^*)$. Thus, the forward Early effect for SiGe HBTs is small compare to Si BJT since $Q_{jC} > Q_{jC}^*$ for $h_{jC} < 1$.

Another important advantage is that the DC current gain β_{SiGe} obtained from the ratio of collector current over the base current of the SiGe HBT is expected to be larger than the DC current gain β_{Si} of the Si BJT. The base currents for both SiGe and Si transistor are expected to be the same since the Ge profile starts in the base-emitter space-charge region and extended into the base region, and thus, it has little influence on the injection of hole carriers into the neutral emitter region. The expression for base current is determined by using ICCR, giving [12]

$$I_B = \frac{qA_{E0}\overline{\mu_{pE}n_{iE}^2}V_T\left(\exp\left(\frac{V_{BE}}{V_T}\right) - 1\right)}{\int N_E(x)dx} \quad (2.12)$$

Since I_B remains the same for both SiGe and Si transistors, β_{SiGe} can be expressed as a function of β_{Si} by taking the ratio of the collector current of SiGe by assuming an average ΔV_{Ge} and the collector current of Si transistor. The result is given as

$$\beta_{SiGe} = \beta_{Si} \cdot \exp\left(\frac{\overline{\Delta V_{Ge}}}{V_T}\right) \quad (2.13a)$$

and

$$\beta_{Si} = \frac{\overline{\mu_{nB}} N_E w_E}{\overline{\mu_{pE}} N_B w_B} \cdot \exp\left(\frac{\Delta V_{gB} - \Delta V_{gE}}{V_T}\right) \quad (2.13b)$$

where ΔV_{gB} and ΔV_{gE} are bandgap narrowing voltages for base and emitter regions, respectively, and $w_E(=x_{jeE})$ and $w_B(=x_{jCB}-x_{jeB})$ are widths of neutral emitter and base, respectively. Equation 2.13a showed β_{SiGe} is greater than β_{Si} by a factor of $\exp(\overline{\Delta V_{Ge}}/V_T)$.

2.3.3.3 Transit Time

As mentioned earlier, the graded Ge mole fraction is expect to decrease the transit time or increase the transit frequency due to the built-in electric field as a result of the graded Ge in the base. The transit time of the transistor is given as[12]

$$\tau_f = \tau_{pE} + \tau_{pC} + \tau_{BE} + \tau_{BC} + \tau_{Bf} \quad (2.14)$$

where τ_{pE} , τ_{pC} , τ_{BE} , τ_{BC} , and τ_{Bf} are the transit time of neutral emitter, neutral collector, base-emitter SCR, base-collector SCR, and neutral base, respectively. Usually, τ_{Bf} is the main contribution to the total transit time at low to medium current densities, but at high current densities τ_{Bf} and τ_{pC} are the two major contributions to the overall transit time. For ease of analysis, the derivation of τ_{Bf} is done at low current density. Utilizing the same assumptions used in the derivation of the modified GICCR derivation with an ideal doping concentration, the minority electron density expression in the neutral base region can be readily obtained by integrating the electron transport equation $I_n(=I_C>0$ due to assumption (ii)) over the neutral base region which are bounded by $n_e(x_{jeB})$ and $n_C(x_{jCB})$ and the neutral base width w_B . The result is given as

$$-I_n = -I_C = qA_{E0}n\overline{\mu_{nB}}E + qA_{E0}\overline{\mu_{nB}}V_T \frac{dn}{dx} \quad (2.15)$$

Using (2.15) and the boundary $n_c(x_{jcB})$, the minority carrier expression for electron becomes

$$n(x) = \left(\frac{I_C}{qA_{E0}\mu_{nB}V_T\zeta} \right) (\exp[\zeta(x_{jcB} - x)] - 1) + n_C \exp(\zeta(x_{jcB} - x)) \text{ .for } x_{jeB} < x < x_{jcB} \quad (2.16a)$$

with

$$\zeta = \frac{E}{V_T} = -\frac{m_{ac}}{V_T} = \frac{-a_{GCmax}}{w_B V_T} \quad (2.16b)$$

and

$$n_C(x_{jcB}) = \frac{I_C}{qA_{E0}v_C} \quad (2.16c)$$

where v_C is the velocity at the collector and I_C is the collector current. Moreover, v_C is directly dependent on the base-collector voltage via the electric field in the base-collector space-charge region. This implies that at high $V_{CEi} > 0.8V$, the electric field is above $1E4V/cm$ such that v_C is equal to the saturation velocity $v_S = 1E7$ cm/s. However, at low V_{CEi} less than $0.8V$, the electric field is less than $1E4V/cm$ which means v_C is equal to $\mu_{nC} \cdot E_{field}$ which is less than v_S . The base transit time in the neutral base region is defined as

$$\tau_{Bf} = \frac{dQ_{nB}}{dI_C} \quad (2.17)$$

where Q_{nB} is the minority carrier in the neutral base region as a function of collector current which can be determined from

$$Q_{nB} = qA_{E0} \int_{x_{jeB}}^{x_{jcB}} n(x) dx \quad (2.18)$$

Case 1: $\zeta=0$ Si BJT

By setting $\zeta=0$, simplifying equation (2.16a), substituting the result back into equation (2.18) and solving for Q_{nB} , the base transit time expression determined by differentiating Q_{nB} with respect to I_C becomes

$$\tau_{BfSi} = \frac{w_B^2}{2\mu_{nB}V_T} + \frac{w_B}{v_C} \quad (2.19)$$

The first term of the expression is the standard base transit time which is dependent on the base width, whereas the second term of the expression also known as carrier jam transit time, is a function of base-collector voltage. The carrier jam transit time represents the influence of the finite velocity in the base-collector SCR which leads to a carrier jam in the base that limits the transport velocity through the neutral base region. As mentioned earlier, the velocity is a function of the electric field in the base-collector SCR region which is dependent on the base-collector voltage. Thus, at low V_{CEi} v_C is less than v_S , and as a result the base transit time would be large.

Case 2: $\zeta \neq 0$ SiGe HBT

The expression of the base transit time obtained by solving (2.17) via (2.16) and (2.18) is given as

$$\tau_{BfSiGe} = \frac{w_B^2}{2\mu_{nB}V_T} \cdot \delta_1 + \frac{w_B}{v_C} \cdot \delta_2 \quad (2.20)$$

where w_B is width of neutral base region and δ_1 and δ_2 are scaling factors that are dependent on the Ge mole fraction which are given as

$$\delta_1 = \frac{2[\exp(-a_{GC_{max}}w_B/(w_{Ge}V_T)) + a_{GC_{max}}w_B/(w_{Ge}V_T) - 1]}{(a_{GC_{max}}w_B/(w_{Ge}V_T))^2} \quad (2.20a)$$

and

$$\delta_2 = \frac{[1 - \exp(-a_{GC_{max}}w_B/(w_{Ge}V_T))]}{(a_{GC_{max}}w_B/(w_{Ge}V_T))} \quad (2.20b)$$

w_{Ge} is the width of the graded Ge in the base region. The scaling factors are usually less than one. To illustrate this, using the following data from Fig. 2.4, $a_C=0.65V/\text{mole frac-}$

tion, $c_{\max} = 0.05$ peak mole fraction, $w_{\text{Ge}} = 0.045$ micron, $w_{\text{B}}=0.04$ micron, and $V_{\text{T}}=0.02586\text{V}$, the calculated results for δ_1 and δ_2 are 0.86 and 0.64, respectively. Therefore with δ_1 and δ_2 less than one, τ_{BfSiGe} of the SiGe HBT is less than τ_{BfSi} of the Si BJT. Moreover, the base transit time decreases as the Ge mole fraction increases. Finally, the addition of the Ge mole fraction in the base also reduces the influence of base-collector voltage on the carrier jam transit time. This result has an important implication with respect to compact modeling which will be discussed later.

2.4 Numerical Simulation

To demonstrate some of the underlining advantages in performances of SiGe over Si technologies discussed earlier, several numerical simulations were performed by using DEVICE[6], a numerical mixed-mode device/circuit simulator of MOS(Metal Oxide Semiconductor), bipolar, and heterostructure devices. DEVICE can be used for solution of the basic semiconductor equations such as continuity equation, carrier(electrons and holes) equations, and Poisson's equation in either one (1D), two (2D), or three (3D) spatial dimensions. The modelling of SiGe bipolar transistors in DEVICE was achieved by allowing the user to specify a germanium mole fraction either in an exponential or linear profile and parameters describing the physical relationship between the Ge mole fraction and silicon. In DEVICE, a linear equation was used to describe the dependence of the bandgap voltage on mole fraction. Moreover, the electron and hole mobilities in dependence on mole fraction are modeled on the basis of data given in [9,13] on mobility of electrons in germanium-silicon alloy.

To provide a realistic comparison a 1D IBM doping profile from Fig. 2.4 was used for a 1D simulation. The reason for using a 1D profile, as mentioned earlier, was to get an insight into the operation of the transistor without the influence of series resistances or current spreading that would normally occur in 2D and 3D simulations.

To illustrate the influence of the Ge content has on the electrical performance, SiGe and Si technologies of identical 1D profiles are used. For SiGe technology, two transistors with Ge contents of 5 and 10 percent are used. The DC characteristics such as Gummel

plot (I_C , I_B vs. V_{BEx}), DC current gain (β) vs. I_C , and I_C vs. V_{CE} and AC characteristics such as the figure of merit versus I_C and f_T versus V_{CE} are used in the comparison. These characteristics are used to illustrate some of the theory discussed earlier.

2.4.1 DC Characteristics

Several DC simulation were performed to illustrate some of the advantages seen in a SiGe transistor over that of a Si transistor. Fig. 2.7 shows a plot of an energy band diagram for a Si transistor and a SiGe transistor with a 5% mole fraction. The figure shows a small slope in the conduction energy band in the base for a SiGe transistor with 5% mole fraction, whereas a flat conduction energy band in the same is observed for a Si transistor.

FIGURE 2.7 Bandgap Energy of a realistic 1D drift Doping Profile with and without Ge mole fraction

The figure shows a graded conduction energy in base region for the case with Ge as suggested earlier. Fig. 2.8 shows a Gummel plot for a Si transistor and SiGe transistors with the two different Ge mole fractions 5% and 10%.

FIGURE 2.8 Gummel Plot at $V_{CEx}=1.6V$

The collector currents for 5%Ge and 10%Ge transistors are about 1.25 and 1.4 times larger, respectively, compared to the collector current of the Si transistor. The amount of

collector currents for the SiGe transistor is not as large because of the small amount of the Ge mole fraction used. The base currents are approximately equal for the three transistors as expected since the main Ge content occurs in the base region which has little influence on the injection of minority hole carriers in the emitter region - a major contribution of the total base current. Note that, the collector currents for the three transistors at high collector current densities are approximately equal. This illustrates that the influence of Ge at high current densities is negligible because in this regime of operation, the influence of Kirk Effect, that is the extension of neutral base width, dominates. Fig. 2.9 shows the corresponding DC current gain Beta or (β) versus collector current. Again, the beta for the SiGe transistors with the two different Ge percents are approximately 1.25 and 1.4 times larger than the beta of the Si transistor, respectively, which confirms the theory.

FIGURE 2.9 DC current Gain Beta(β) vs. I_C at $V_{CE} = 1.6V$

The verification of Early Effect can be seen in the output characteristics. The output characteristics (I_C vs. V_{CE}) for Si and SiGe(Ge=5%) transistors are shown in Fig. 2.10. The slope of the solid-line curves is larger compared to the slope of the dot-dash line curves. This confirms the theory where the Early effect is more dominant for the Si transistor compared to the SiGe transistor.

FIGURE 2.10 Output Characteristics for $V_{BE} = (0.78, 0.82, 0.84, 0.85, 0.86, 0.87)$

2.4.2 AC Characteristics

A figure of merit often used to verify the transit time τ_f or speed of the transistor is the transition frequency f_T , that is the frequency at unity small signal current gain of a common-emitter configuration. Several small signal numerical simulations were performed for the three 1D transistors, Si, SiGe(5% Ge), and SiGe(10% Ge). Fig. 2.11 shows the transition frequency as a function of the collector current for the three transistors.

FIGURE 2.11 Transition Frequency for $V_{CEx}=1.6\text{V}$

The peak f_T values for 5% and 10% Ge are 50GHz and 55GHz, respectively, whereas the peak f_T value of Si transistor is 45GHz. These results confirm the theoretical analysis discussed earlier, where the transit time for SiGe is expected to be smaller than the transit time of Si. Moreover, the transit time is expected to decrease with increasing Ge mole fractions. At high current densities, the abrupt drop in f_T 's for SiGe transistors is due to the minority carrier built up in the base-collector junction. The reason for the excess charge in that region is because the built-in electric field created by the graded Ge profile in the base-collector junction acts as a retarding electric field that prevents minority electron carriers from crossing into the collector region. Figs. 2.12a and 2.12b confirm the observation by plotting the electric field and electron density at the base-collector junction for two collector currents 1mA and 3mA, respectively, for the corresponding three transistors. Using the same plotting symbols for the three transistors, the accelerating electric field components occurring at a depth of 600Å for a 1mA collector current are observed. These can also be seen more clearly in the zoom in part of the plot. The largest accelerating field observed in a SiGe(Ge of 10%) transistor results in the smallest minority electron

density at the base-collector region and this is confirmed in Fig. 2.12b, whereas the smallest accelerating field and the largest minority carrier density for a Si transistor are observed. However, at 3mA collector current (the beginning of Kirk Effect - base extension into the collector), the conduction energy band offset created by the Ge mole fraction at the base-collector junction generates a retarding electric field shown in the zoom part of Fig.2.12a. The retarding field prevents the flow of minority carriers from crossing into the collector region, whereas the Si transistor does not exhibit any retarding electric field which is as expected. The blocking of minority carriers due to retarding field is confirmed in the zoom part of the curves of Fig. 2.12b, where the largest built-up of minority carriers occurred in that region is observed for the SiGe(Ge=10%) transistor, and the smallest built-up of minority carriers is observed for the Si transistor.

FIGURE 2.12a Electric Field at $I_C=(1,3)\text{mA}$, (Solid line) Si, (dotted-dash line) SiGe(Ge=5%), and (dash-line) SiGe(Ge=10%)

FIGURE 2.12b Electron Density at $I_C=(1,3)\text{mA}$, (Solid line) Si, (dotted-dash line) SiGe(Ge=5%), and (dash-line) SiGe(Ge=10%)

To verify the effect of the Ge mole fraction on the carrier jam component of the transit time at low collector current densities, Fig 2.13 shows a plot of f_T versus V_{CEx} for fixed V_{BEx} 's of 0.78, 0.82, and 0.84V.

FIGURE 2.13 f_T vs. V_{CEx} at low I_C for fixed $V_{BEx}=(0.78,0.82,0.84)$ V

The transit frequency curves for a SiGe(5%Ge) transistor are much flatter compared to the curves for a Si transistor. The flatter curves are attributed to the smaller carrier jam transit time in the base-collector due to the scaling factor δ_2 , discussed earlier, that minimizes the influence of V_{CEx} on the overall transit time variation, whereas the overall transit time of the Si transistor has a slightly stronger dependence on V_{CEx} variations.

Chapter 3 - Compact Models

3.1 Introduction

A compact model consists of lumped elements, such as resistors, diodes, capacitors, and dependent current sources, connected together with minimum number of nodes to form an equivalent circuit to represent a semiconductor transistor. The complexity of a compact model can be as simple as a table fitting method that has little physical basis to a semi-physical model. The term physical refers to the relation of the model equations to the underlying physics describing the operating effect in the device.

The first bipolar transistor compact model was the Ebers-Moll[14]. Since a number of physical effects occur in integrated circuits which are not included in the model, it was replaced by Gummel-Poon Model (GPM)[15], first introduced in the early 1970s which was based on the Integral Charge-Control Relation (ICCR). Several variants of the GPM were derived and one of the well known variants found in most circuit simulators today is the SPICE Gummel-Poon (SGP) model. Over a twenty year period, the model has remained relatively unchanged while the technology has changed dramatically over that time period. As a result, the accuracy of the model has declined to the point where it is no longer sufficiently accurate in chip designs. This has led to the development of more physical-based compact models, such as High Current Model (HICUM)[3], Most Exquisite Transistor Model (MEXTRAM)[16], Charged-Based Model(CBM)[17], and some proprietary models used in industries by extending the existing SGP Model to account for some of the deficiencies. An example of the extension is the Bell SGPM or BSGPM[2] that accounts for the quasi-saturation effects in the epitaxial collector region.

This chapter gives a brief description of the main theory and the shortcomings of both the defacto standard SGP model and the Nortel proprietary version of HICUM. A brief overview on the procedures to determine model parameters is discussed, followed by a discussion on a measurement setup that can be used to perform DC and AC (Scattering or S-Parameter) measurements, and as well as the issue of de-embedding in S-Parameter measurements. Finally, some comparisons are shown between simulated results obtained

from both SGPM and HICUM to measured data for various device characteristics, such as I_C and I_B versus V_{BE} (Gummel Plot) and I_C versus V_{CE} (Output Characteristics) under quasi-static conditions and two figures of merit cutoff frequency(f_T) which is the unity gain frequency of a common emitter configuration and maximum oscillating frequency(f_{max}) which is the unity available power gain also known as the Mason's gain under small-signal conditions.

3.2 Compact Models - SGPM and HICUM

The focus of this section is to provide some of the modeling principles for the quasi-static (DC) operating conditions that include collector and base currents and series resistances and small-signal conditions that include junction capacitances and transit time, associated with each compact model.

3.2.1 SPICE Gummel-Poon Model

Modified versions of SGP model can be found in many commercial simulators such as the well known HSPICETM (Metasoftware), S-SPICE (Silvaco), PSPICE(Microsim), MSPICE (Mentor Graphics), Spectre, and ELDO(ANACAD). Despite these variants in SGP which can turn certain equations on to account for a particular effect, the framework still remains the same. A level 2 equivalent circuit for an integrated vertical NPN or PNP BJT Spice Gummel Poon Model shown in Fig.3.1 is for a four-terminal device showing emitter (E), base (B), collector (C), and substrate (S) terminals with the currents flowing into these terminals as I_B , I_C , I_E , and I_S . The diagram in the dashed-box represents the standard SGPM equivalent circuit level1 which is usually found in most commercial simulators. The additional elements, such as capacitors and resistors added to the existing equivalent circuit to take into account parasitics associated with the high-speed bipolar transistors, elevate the standard equivalent circuit into a level 2 which is only available in HSPICE. The result of the additional elements constitutes a more accurate compact model. The definitions for the model parameters are discussed in Table B1 of Appendix B.

FIGURE 3.1 SGPM Equivalent Circuit for vertical NPN or PNP (after HSPICE[15])

3.2.1.1 DC Currents

The forward and reverse DC operation of a bipolar transistor are described using the following Gummel-Poon large-signal expressions. The collector current expression is given as [2]

$$I_C = I_T - \frac{I_{SC}}{B_R} \left(\exp \left(\frac{V_{BCi}}{n_R V_T} \right) - 1 \right) - I_{SC} \left(\exp \left(\frac{V_{BCi}}{n_C V_T} \right) - 1 \right) \quad (3.1a)$$

where I_T is defined as the transfer current; I_{SC} is the base-collector saturation; B_R is inverse DC current gain; V_{BCi} is the internal base-collector voltage; n_R is the inverse-biased non-ideality factor; n_C is the base-collector non-ideality factor; and V_T is the thermal voltage. I_T defined as the difference between the forward collector current $I_{C\text{forward}}$ and the reverse collector current $I_{C\text{reverse}}$ is given with base charge Q_B normalized:

$$I_T = \frac{I_S}{Q_B} \cdot \left(\exp \frac{V_{BEi}}{n_F V_T} - \exp \frac{V_{BCi}}{n_C V_T} \right) \quad (3.1b)$$

where I_S is the forward saturation current; V_{BEi} is the internal base-emitter voltage; n_F is the forward non-ideality factor; and Q_B is the hole charge in the base which is defined in (3.1c).

$$Q_B = \frac{\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{I_S \exp\left(\frac{V_{BEi}}{n_F V_T}\right)}{I_{KF}} + \frac{I_S \exp\left(\frac{V_{BCi}}{n_R V_T}\right)}{I_{KR}}}}{\left(1 - \frac{V_{BEi}}{V_{AR}}\right)\left(1 - \frac{V_{BCi}}{V_{AF}}\right)} \quad (3.1c)$$

where I_{KF} and I_{KR} are forward and reverse knee current, respectively and V_{AF} and V_{AR} are forward and reverse Early Voltage, respectively. The expression for base current is given as [2]

$$I_B = \frac{I_S}{B_F} \exp\left(\frac{V_{BEi}}{n_F V_T}\right) + I_{SE} \exp\left(\frac{V_{BEi}}{n_E V_T}\right) + \frac{I_{SC}}{B_R} \left(\exp\left(\frac{V_{BCi}}{n_C V_T}\right) - 1\right) + I_{SC} \left(\exp\left(\frac{V_{BCi}}{n_C V_T}\right) - 1\right) \quad (3.2)$$

where B_F is the forward DC current gain; I_{SE} is the base saturation current; and N_E is the base non-ideality factor.

3.2.1.2 DC Bulk Resistances

The modeling of parasitics to account for some of the 2D effects in a bipolar transistor is characterized by the resistances in the bulk region and contacts. The emitter and collector resistances are modeled with constant resistors R_E and R_C , respectively. The modeling of base resistance is somewhat complicated because it must combine the resistance from the neutral base region which is a function of the collector current and the constant extrinsic base resistance (R_{BM}). The base resistance has a characteristic described in Fig.3.2.

FIGURE 3.2 Base resistance as a function of collector current

The resistance in the neutral base region due to the horizontal flow under the emitter is determined from the base doping level and thickness of the neutral base region, together with the horizontal dimensions of the active base region. In other words, the internal base

resistance depends implicitly on the internal base sheet resistance r_{Sbi} which is a strong function of DC operating point. The modeling of the base resistance as a function of base current I_B that also accounts for the emitter current crowding effect is given by [16]

$$R_{BB} = R_{BM} + 3(R_B - R_{BM}) \cdot \frac{\tan z - z}{z \cdot (\tan z)^2} \quad (3.3a)$$

and

$$z = \frac{\pi^2}{24} \cdot \sqrt{\frac{I_{RB}}{I_B}} \cdot \left(\sqrt{1 + \frac{144 I_B}{\pi^2 I_{RB}}} - 1 \right) \quad (3.3b)$$

where R_B is the zero-biased base resistance; R_{BM} is the extrinsic base resistance; and I_{RB} is the collector current that obtained from the average of R_B and R_{BM} .

3.2.1.3 Junction Capacitances

Accurate modeling of depletion (junction) capacitance is important in small signal operations. The classical junction capacitance characteristic is described in Fig. 3.3.

FIGURE 3.3 Junction Capacitance versus voltage (CV) curve

A general junction capacitance expression to describe the above curve is given as[2]

$$C_{j(E, C, S)} = \frac{C_{j0}}{\left(1 - \frac{V_{(E, C, S)}}{V_{D(E, C, S)}}\right)^{m_j}} \quad (3.4)$$

where the subscripts E, C, and S represent base-emitter junction, base-collector junction, and collector-substrate junction, respectively; C_{j0} is the capacitance at zero bias; V_D is the built-in voltage; m_j is the exponent factor which has a value between 1/2 for an abrupt type

doping profile and 1/3 for a linear type doping profile; and V is the applied junction voltage. The junction capacitance equation is the same for all compact models which also include HICUM.

To avoid division by zero in (3.4) when V equal to V_D , a model parameter FC is used to ensure that C_j is continuous by linearizing the formula which is achieved by setting the option DCAP=1 in HSPICE. The extrinsic base-collector junction capacitance C_{jCX} specified in the equivalent circuit is determined from the model parameter XCJC that has a value between zero and one, and it is used to separate the base-collector capacitance into intrinsic and extrinsic components.

3.2.1.4 Transit Time

The empirical expression to describe the transit time of a transistor is given as[2]

$$T_F = T_{F0} \left(1 + X_{TF} \cdot \frac{I_S \exp\left(\frac{V_{BEi}}{n_F V_T}\right)}{I_{TF} + I_S \exp\left(\frac{V_{BEi}}{n_F V_T}\right)} \cdot \exp\left(\frac{V_{BC}}{1.44 \cdot V_{TF}}\right) \right) \quad (3.5)$$

where T_{F0} is the only model parameter that has a physical meaning and is related to τ_{F0} obtained from the transit frequency measurement at low current densities; I_{TF} is related to the zero bias hole charge and T_{F0} ; and the model parameters X_{TF} and V_{TF} have no physical meaning, but rather they are fitting parameters to describe the empirical transit time formula. Fig 3.4 shows a transit time characteristic as a function of collector current density for different terminal collector emitter voltages (V_{CEX}).

FIGURE 3.4 Transit Time as a function of collector current density

An empirical relationship between the transit time parameters and the corresponding curves is shown in the above figure. T_{F0} is the transit time at low I_C . V_{TF} is used to describe the separation between each T_F curve. I_{TF} can be approximated from the product of low current density transit time T_{F0} and zero bias hole charge in the base region Q_{P0} which describes approximately the onset of T_F curve. Finally, X_{TF} describes the curvature of T_F curve.

3.2.2 HICUM

A level2 equivalent circuit for HICUM is presented in Fig.3.5. The circuit contains an additional node B^* compared to the SGPM equivalent circuit. The definitions for the model parameters are discussed in Table B2 of Appendix B.

FIGURE 3.5 HICUM/LEVEL2 equivalent circuit (after Schröter [18], [19])

3.2.2.1 Quasi-Static Transfer Current

The quasi-static transfer current that takes into account the forward and reverse DC operations is given as [18]

$$i_T = \frac{c_{10}}{Q_P} \left(\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right) \right) \quad (3.6a)$$

where c_{10} and Q_P are the current constant containing physical quantities and the transistor area and hole charge, respectively. The expression for the current constant[18] in (3.6b) is also a model parameter.

$$c_{10} = (qA_E)^2 V_T \mu_{nB} \overline{n_i^2} \quad (3.6b)$$

The hole charge expression is given as [18]

$$Q_P = Q_{P0} + Q_{jEi} + Q_{jCi} + Q'_f \quad (3.6c)$$

where Q_{P0} is a model parameter for the hole charge within the effective internal transistor at thermal equilibrium; Q_{jEi} and Q_{jCi} are the depletion charges stored within the base-emitter BE and base-collector BC junction of the effective internal transistor; Q'_f defined in (3.7) is the (weighted) minority charge stored in the transistor which is obtained from the integration of the transit time (τ_f) as a function of i_T .

$$Q'_f = \int_0^{iT} \tau_f di \quad (3.7)$$

The modeling of quasi-static base current, i_B , defined in (3.8a) is the sum of internal injection base current i_{BEi} [19] given in (3.8b) and periphery injection base current i_{BEp} [19] given in (3.8c).

$$i_B = i_{BEi} + i_{BEp} \quad (3.8a)$$

with

$$i_{BEi} = I_{BEiS} \left(\exp \left(\frac{V_{BEi}}{m_{BEi} V_T} \right) - 1 \right) + I_{REiS} \left(\exp \left(\frac{V_{BEi}}{m_{REi} V_T} \right) - 1 \right) \quad (3.8b)$$

where the internal saturation currents I_{BEiS} and I_{REiS} and non-ideality coefficients m_{BEi} and m_{REi} are model parameters,

and

$$i_{BEP} = I_{BEPs} \left(\exp \left(\frac{V_{B^*E}}{m_{BEP} V_T} \right) - 1 \right) + I_{REP s} \left(\exp \left(\frac{V_{B^*E}}{m_{REP} V_T} \right) - 1 \right) \quad (3.8c)$$

where the internal saturation currents I_{BEPs} and $I_{REP s}$ and non-ideality coefficients m_{BEP} and m_{REP} are model parameters.

The bulk resistances for the transistor in the collector and emitter region assume constant values similar to that of SGPM. However, the modeling of base resistance is somewhat different than the one in SGPM. In HICUM, the base resistance is the sum of current depended internal base resistance, r_{Bi} , and a constant extrinsic base resistance, r_{Bx} , value. The expression for r_{Bi} is described in (3.9) which is the ratio of zero bias hole charge to the hole charge in the base region which is given as [18]

$$r_{Bi} = r_{Bi0} \cdot \frac{Q_{p0} + \Delta Q_{r0}}{Q_p + \Delta Q_{r0}} \cdot \Psi(\eta) \quad (3.9)$$

where r_{Bi0} is defined as a zero-biased internal base resistance and is a model parameter; ΔQ_{r0} ($=f_{DQR0}Q_{p0}$ where f_{DQR0} a model parameter) is a charge correction factor; and $\Psi(\eta)$ ($=\ln((1+\eta)/\eta)$) describes the emitter current crowding effect with $\eta(=f_{ge0}r_i i_{BEi}/V_T)$ a current crowding factor where f_{ge0} and r_i are model parameters.

The modeling of base-collector, base-emitter, and collector-substrate junction capacitances is similar to the equations described earlier in SGPM, whereas the modeling of the transit time is different which will be discussed in the next section.

3.2.2.2 Transit Time

Accurate modeling of the transit time is an important criteria in this model. The transit time [19]

$$\tau_f(V_{CEx}, i_T) = \tau_{f0} + \tau_{f1} + \tau_{f2} \quad (3.10)$$

which is determined from measured f_T versus I_C at forward-biased operation, contains three transit time components, τ_{f0} , τ_{f1} , and τ_{f2} which are used to accurately describe the low, medium, and high current density regimes, respectively. Fig.3.6. shows the transit time curve along with its components.

FIGURE 3.6 Transit Time vs. Collector current I_C that includes the various components used in the Transit time model equations.(after Schroter [19])

The currents in the above figure I_{CL} and I_{CK} are used to describe the onset of medium current density from low current density operating regime and the onset of high current density from medium density operating regime, respectively. In addition, the currents also describe the onset of different transit time components.

Unlike the empirical expression for the transit time in SGPM, the description in HICUM is more physical. For the low current density transit time component τ_{f0} [19] described in (3.11), model parameters V_1 and V_0 take into account the finite carrier velocity and Early-effect in the base-collector space-charge region, respectively.

$$\tau_{f0} = \tau_0 \left(1 - \sqrt{\frac{v_C}{V_0}} + \frac{V_1}{v_C} \right) \quad (3.11)$$

where v_C defined as $V_{CEx} - V_{CEs}$ is the auxiliary voltage. Equation (3.12) describes the medium current density transit time component τ_{f1} [19] for the condition $I_{CL} < i_T < I_{CK}$.

$$\tau_{f1} = \frac{\tau_1}{2} \left(\left(\frac{i_T - I_{CL}}{I_{CK} - I_{CL}} \right)^{\alpha_0} + \left(\frac{i_T - I_{CL}}{I_{CK} - I_{CL}} \right)^{\alpha_1} \right) \quad (3.12)$$

where α_0 and α_1 are model parameters and τ_1 [19] is the time constant in (3.13) with model parameters τ_{10} and I_{10} .

$$\tau_1 = \tau_{10} (1 + I_{CK}/I_{10}) \quad (3.13)$$

Equation (3.14) represents the high current density transit time component τ_{f2} [19] with model parameters τ_{20} and s_{20} for the condition $i_T > I_{CK}$.

$$\tau_{f2} = \tau_1 + \tau_{20} (1 - (I_{CK}/i_T)^2 + s_{20}(i_T - I_{CK})) \quad (3.14)$$

3.3 Determination of Model Parameters

Several methods to determine model parameters for both SGPM and HICUM are briefly discussed in the next couple of sections. A program tool to extract model parameters is IC-CAP[20], a measurement system by HP EEsof which is used to perform the necessary measurements. Extraction routines written in IC-CAP Rocky Mountain Basic are utilized to manipulate the measured data, correlate with the correct model equations, then extract the required model parameters.

3.3.1 Extraction of SGPM Parameters

There are several methods to determine SGP model parameters. A method recommended by Getreu[21] involves the extraction of parameters by fitting them to certain measurements, followed by a global optimization of parameters at the end of the extraction process. However, the method yields less than optimal agreement when comparing SPICE simulations to measurements because the assumptions made about the meaning of the parameters are different than those in the model equations.

An alternative approach suggested by Van Halen et al.[22] is to use the model equations to define and establish methods to directly extract parameters that do not require optimization. This involves linearizing the model equation by performing the measurement at certain operating regime, which makes possible the dropping or neglecting of certain terms in the equation and manipulate the equations into a $y=mx+b$ form, from which parameters can be extracted by applying the least square fit with slope($=m$) and intercept($=b$) equal to the parameter values. The advantage of the method is that monitoring of process variations, geometric scaling and statistical modeling are possible due to the mutually consistent method used. This method was refined and implemented by Kendall[23] for the forward DC operation of the transistor in TECAP, an earlier version of IC-CAP developed by HP EEsof. Detail implementations of the direct extraction procedures for the forward SGPM parameters in IC-CAP can be found in Tran[24].

3.3.2 Extraction of HICUM Parameters

Due to the semi-physical nature of HICUM, scaling capabilities are possible. This requires technology specific data such as sheet resistances for base, emitter, and collector regions and junction capacitances in per area and per periphery to be measured from test structures. For instance, the decomposition of junction capacitance into area or bottom (C_j/Area) and periphery components (C_j/Peri) is achieved by measuring two test structures, a large square structure (length=width) and a long rectangular structure (length $\gg$ width). The junction capacitance measured from each structure consists of bottom and periphery component. Consequently, a system of two linear equations can be solved easily for the two unknown. Furthermore, specific HICUM parameters such as c_{10} , Q_{p0} , and the transit time parameters have to be measured and determined from the various transistor sizes. The approach to determining these specific HICUM parameters can be the same method discussed in extracting the SGP model parameters earlier that is, employing the direct extraction technique. These specific technology and HICUM parameter data are entered into TRADICA[25], a program for dimensioning and calculation of integrated bipolar transistors, which is used to generate a HICUM parameter set for any given transistor size. Moreover, TRADICA can also be used to generate SGP model parameter for a

full level2 SGP equivalent circuit by using the technology specific data and some specific SGP parameters. Some of the underlining theory in measuring and determining the specific technology and HICUM parameter data are discussed in [3,19].

3.4 Device Electrical Characteristics

Model verification is often accomplished by comparing simulation results to measured data. Typical electrical characteristics often used are the usual DC figures such as Gummel Plot, current gain, and output characteristics (I_c versus V_{CE} for various base currents) and AC figures of merit such as f_T and f_{max} . Often, these figures of merit are calculated as opposed to measuring them directly. The reasons will be discussed in the next section.

3.4.1 Determination of Transition Frequency and Maximum Oscillation Frequency

The transition or cutoff frequency (f_T) is defined as the frequency in which the small signal current gain $|H_{21}|$ (H-Parameter) in a common emitter configuration is at unity, whereas the maximum oscillating frequency (f_{max}) is defined as the frequency in which the maximum unilateral power gain (UG_{max}) is at unity. Fig. 3.7 show a plot of $|H_{21}|$ and $|UG_{max}|$ as a function of frequency, respectively.

FIGURE 3.7 Gain versus Frequency

Ideally, both f_T and f_{max} can be determined directly from the following Gain versus Frequency as shown in Fig. 3.7. However, direct measurements of both figures of merit may give misleading results. One of the reasons is due to the parasitic capacitances associated with the cables/interconnects from the network analyzer that causes the slope of H_{21} to be less than the -20dB/dec roll off shown as a solid line in Fig. 3.8. As a result, f_T is

grossly overestimated. Another reason is due to the continual advancement in high-speed bipolar transistors that can have maximum transit frequency beyond the operating frequency range of a typical network analyzer which has a frequency range of 45MHz to 26.1GHz. A similar problem also occurs when determining $f_{\max}$. Therefore, an alternative solution is to calculate for f_T and $f_{\max}$ using measured data.

FIGURE 3.8 Small Signal Current Gain $|H_{21}|$ vs. Frequency

The methods to calculate f_T and $f_{\max}$ are as follows.

i. f_T —The expression for a single-pole low-pass behavior shown in Fig. 3.8, is given as

$$H_{21} = \frac{\beta_0}{1 + j\frac{f}{f_C}} \quad (3.15)$$

where f is the frequency, f_C is the -3dB corner-frequency, and β_0 is the small signal current gain at low frequency. f_T obtained from the gain-bandwidth product, $f_C \cdot \beta_0$, can be calculated directly if the corner frequency is known. This method is somewhat tedious, time consuming since a frequency sweep is required to search for the corner frequency at each DC operating point, and it can be difficult to locate the corner frequency for some advanced bipolar transistors due to the lower frequency limit of the network analyzer. An alternative method is to calculate f_T by re-arranging the magnitude of (3.15) along with the gain-bandwidth product expression with the resulting expression for f_T given as [3]

$$f_T = \frac{\beta_M \cdot f_M}{\sqrt{1 - (\beta_M/\beta_0)^2}} \quad (3.16)$$

where f_M has to be greater than f_C but should be on a -20dB/dec slope; β_M is the small-signal current gain at f_M ; and β_0 is the low frequency current gain which is determined from the dI_C/dI_B relation. This method to determine f_T is much faster compared to the previous method since it only requires a single frequency measurement.

An analytical expression for the transition frequency (f_T) relating to the device parameters and the transistor transit time (τ_f) is given as

$$\tau_f = \frac{1}{2\pi f_T} - \frac{C_{jE} + C_{jC}}{g_m} - (R_C + R_E)C_{BC} \quad (3.17)$$

where C_{jE} and C_{jC} are base-emitter and base-collector junction capacitances. g_m is the transconductance defined as I_C/V_T . R_C and R_E are collector and emitter series resistances, and C_{BC} is the base-collector capacitance. The measured transit time can be used to determine the transit model parameters from the compact model. Note the term, $(R_C + R_E)C_{BC}$, in (3.17) is the additional time constant that exists in the common-emitter configuration has to be subtracted from the transit time because it is already taken into account in the equivalent circuit.

ii. f_{max} - The expression for maximum oscillating frequency [26] is given in (3.17).

$$f_{max} = \sqrt{UG_{max}} \cdot f \quad (3.18a)$$

where f is the frequency and UG_{max} is equal to unilateral power gain which is expressed as

$$UG_{max} = \frac{|S_{21} - S_{12}|^2}{2k \cdot |S_{21}S_{12}| - 2\Re\{S_{21}S_{12}^*\}} \quad (3.18b)$$

and k defined as the stability factor which has to be greater than one is given as

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |S_{11}S_{22} - S_{12}S_{21}|^2}{2|S_{12}||S_{21}|} \quad (3.18c)$$

where S_{ij} is S-Parameter for $i, j = 1$ and 2 .

An analytical expression that relates f_{max} to the device parameters[27] is given as

$$f_{max} = \sqrt{\frac{f_T}{8\pi C_{BC} r_{bb}}} \quad (3.19)$$

where f_T is the transition frequency and C_{BC} and r_{bb} are the base-collector capacitance and base resistance, respectively. The simple expression provides some insights into the parameters that have the most effect on maximum oscillating frequency. This is important in transistor and circuit design.

3.5 Measuring System Setup

The measurement setup shown in Fig. 3.9 can be used to perform either DC or small-signal or S(Scattering)-Parameters measurements. The system consists of IC-CAP software used to control the measuring instruments via HPIB (Hewlett Packard Interface Board) data bus, a high-frequency probe station, and instruments such as a Semiconductor Parameter Analyzer HP4145B which is used to perform DC measurements and a network analyzer HP8510C used to perform S-Parameter measurements for a frequency range of 45MHz to 26.5GHz.

FIGURE 3.9 DC and S-Parameter Measurement System

Fig. 3.10 shows a typical a DUT (Device-Under-Test) with high frequency pads in a common emitter configuration. The substrate and emitter contacts are connected together to ground and signals are applied to base and collector.

FIGURE 3.10 High Frequency DUT pads for Two-Port S-Parameter and DC Measurements

3.5.1 Calibration and De-embedding of Parasitics/Interconnects

3.5.1.1 Calibration

The network analyzer has to be calibrated before an S-parameter measurement is performed. The purpose of calibration is to correct for the parasitic losses as a result of impedance mismatches, losses to the cables and connectors, and the overall system losses. The calibration is corrected for these errors by comparing S-parameters measured on standard structures to their known values. Typical standards normally used are SHORT, THRU, LOAD, and OPEN. The first three standards can be found in a ISS (Impedance Standard Substrate) that comes with the coplanar probe. Table 3.1 contains S-Parameter results for the standards LOAD, OPEN, SHORT, and THRU which is used to verify the calibration to see if it has been performed correctly once it has been completed and the correction has been turned on.

TABLE 3.1. S-Parameter Results for the standards

Standard	S-Parameter Results
LOAD & THRU	S_{11} and S_{22} should be located at the $1+j0$ (50Ω or at the center) of the Smith Chart. In terms of return loss in dB $ S_{11} $ and $ S_{22} $ should be less than -40dB over the specified frequency range.
OPEN	S_{11} and S_{22} should be located at the $\infty+j0$ (infinite resistance or at the outer edge on the right hand side) of the Smith Chart. In terms of return loss in dB $ S_{11} $ and $ S_{22} $ should be ± 0.1 dB over the specified frequency range.
SHORT	S_{11} and S_{22} should be located at the $0+j0$ (zero resistance or at the outer edge on the left hand side) of the Smith Chart. In terms of return loss in dB $ S_{11} $ and $ S_{22} $ should be ± 0.1 dB over the specified frequency range.

3.5.1.2 De-embedding of Parasitics from DUT

In addition to removing losses in the measuring system, the effect of parasitics associated with high frequency pads and interconnects on a transistor must also be removed. The reason being that the continuous reduction in dimensions both laterally and vertically of the transistor to achieve the required speed, while the dimensions of probe pads and interconnects remain relatively unchanged, can affect the measured results and thereby distort the transistor performance if proper de-embedding is not done. Fig. 3.11 show the equivalent circuit of the high frequency pad with interconnect.

FIGURE 3.11 Equivalent Circuit for a Pad/Interconnect

A method propose by Cho et. Al[28] to de-embed parasitics from the DUT is to take the measured Y-parameters of the DUT and subtract the Y-parameters of the on-wafer blank pads. and the result is converted to Z-parameter which is used to subtract the on-wafer short. The purpose of subtracting the blank pad is to remove the capacitance associated with the pad and the substrate, and the subtraction of the short is to remove the resistance, inductance, and capacitance from the metal line of the interconnect. The equation for the de-embedding is given as

$$Y_{dut} = \left(\frac{1}{Y_{dut_pad} - Y_{pad}} - \frac{1}{Y_{short} - Y_{pad}} \right)^{-1} \quad (3.20)$$

where Y_{dut_pad} , Y_{pad} , and Y_{SHT} are measured Y-parameters from DUT, blank pads, and short, respectively.

3.6 Verification of Compact Models

The verification of compact models is accomplished by comparing simulation to measurement for various electrical characteristics. For the simulation, SGP and HICUM model

parameter sets of a SiGe NPN transistor for an emitter window area of $0.5 \times 10 \mu\text{m}^2$ are determined by using the extraction methods and measurement setup discussed in the previous sections. Various DC and small-signal device characteristics with the corresponding figures for each model are organized in Table 3.2.

TABLE 3.2. Device Characteristics

Compact Models	Figures				
	Gummel plot	Output Characteristics	Gain $ H_{21} $ $ UG_{\max} $	Cutoff Frequency	Maximum Oscillating Frequency
HICUM	Fig. 3.12	Fig. 3.14a	Fig. 15	Fig. 3.16a	Fig. 3.16b
SGPM	Fig. 3.13	Fig. 3.14b	-	Fig. 3.17a	Fig. 3.17b

Figs. 3.12 and 3.13 shows the Gummel Plot of measurements and simulated results from HICUM and SGPM, respectively.

FIGURE 3.12 Gummel Plot for Transistor with $V_{CEX}=(0.8,1.6,2.4)V$ for A_{E0} of $0.5 \times 10 \mu\text{m}^2$.

FIGURE 3.13 Gummel Plot for Transistor with $V_{CEX}=(0.8,1.6,2.4)V$ for A_{E0} of $0.5 \times 10 \mu\text{m}^2$.

The output characteristics (I_C vs. V_{CEX} for fixed I_B) for HICUM and SGPM for the same transistor are shown in Figs. 3.14a and 3.14b, respectively.

FIGURE 3.14 Output Characteristics for Transistor with (a) $I_B=(4,9,18,27,36)\mu\text{A}$ and (b) $I_B=(4,9,18,27,36,45)\mu\text{A}$.

The plots of $|H_{21}|$ and maximum unilateral power gain $UG_{\max}$ versus frequency for both measured data and HICUM simulated results are shown in Figs. 3.15a and 3.15b, respectively.

FIGURE 3.15 (a) $|H_{21}|$ vs. frequency (b) $UG_{\max}$ vs. frequency at $V_{CEx}=1.6\text{V}$
Figs. 3.16a and 3.16b show the f_T and $f_{\max}$ curves, respectively, that compare measured results and simulated results from HICUM.

FIGURE 3.16 (a) f_T and (b) $f_{\max}$ versus I_C for $V_{CEx}=(0.8,1.6,2.4)\text{V}$

Figs. 3.17a and 3.17b show the f_T and f_{max} curves, respectively, that compare measured results and simulated results from SGPM.

FIGURE 3.17 (a) f_T and (b) f_{max} versus I_C for $V_{CE} = (0.8, 1.6, 2.4)V$

3.6.1 Results and Discussion

In the Gummel plots presented in Figs. 3.12 and 3.13, the simulated curves appear to deviate from measured curves at low to high current densities. In the low current density region, the simulated curves for HICUM appear to spread more than the measured curves at three different VCEs. The reason for spread in the simulated curves is due to the incorrect modeling of the Early effect in the existing HICUM model. As discussed in Chapter 2, the result from GICCR indicates that a small Early effect is expected to occur for a graded Ge profile in the base region, and in this case the transistor has a similar Ge profile. The small Early effect which is often characterized with a small output conductance is best illustrated in the output characteristics presented in Figs. 3.14a and 3.14b. The measured curves are much flatter compared to the HICUM simulated curves. However, when comparing SGP simulated curves from HSPICE to measured curves, good agreement is obtained at low current densities, but at high current densities the curves between simulated and measured do not agree and the reason will be explained below. The reason for the simulated curves not bending but appear flat is a result of a large V_{AF} value, a fitting SGP model parameter, that behaves similar to the weighting factor, h_{jc} , in GICCR that reduces the base-collector charge, and thus, it minimizes the Early effect. However, the simulated curves from SGP do not agree well with measurement in the saturation region

where the simulated curves seem to rise faster compare to measurement. This discrepancy is due to the incorrect modeling of the collector resistance in SGP where it is treated as a constant value, but the resistance is a function of current[2].

Returning to the Gummel plot, the simulated curves for both models show large deviation from the measured curves in the high current region. This deviation is attributed to the self-heating effect that the transistor is experiencing under high bias conditions because under this operating regime, large currents flowing through the intrinsic region of the transistor cause a large power dissipation that heats up the transistor. The increase in temperature excites more electron carriers up the conduction band and similarly more hole carriers down the valence band, and as a result more current will flow. The self-heating effect is also reflected in the DC current gain curves, output characteristics, and figures of merit at medium to high current densities.

The next two chapters will discuss methods to resolve the inaccuracies in the existing models. In Chapter 4, the correction to the Early Effect in HICUM is realized by replacing the original GICCR with the modified GICCR discussed in Chapter 2, whereas it is not necessary for SGPM because V_{AF} can be used to compensate to the effect since it is only a fitting parameter. A detailed examination of self-heating effects in relation to the modeling issues which also include the extraction of thermal parameters are discussed in Chapter 5. This involves implementing a self-heating subcircuit by hard coding it into the circuit simulator which is only possible in HICUM but not in SGP model since the source code is not available. However, there is some work underway to make the SGP source code available to device engineer. The reason being that the present SGP model is inaccurate in several key areas that includes: a) quasi-saturation effects, b) avalanche effect, c) self-heating effect. An improved bipolar transistor model that semiconductor industries are currently pursuing is to revise the existing SGP model which is known as VBIC95[29], an acronym for Vertical Bipolar Inter-Company Model for 1995.

Chapter 4 - A New Compact Model

4.1 Introduction

This chapter discusses the implementation of the modified GICCR for the transfer current, derived in Chapter 2, into a compact model to correct for the Early Effects observed in the SiGe HBT technology. In section 4.2, the validity of the modified GICCR is verified by comparing the results from the analytical expression to the simulated collector current obtained from device simulation. Section 4.3 discusses the relationship between the modified GICCR and the original GICCR in HICUM. Since, the modified GICCR contains parameters which cannot be measured or de-convolved easily in contrast to the original GICCR, some simplification to the modified expression in the operating range of interest is needed to reduce the complexity of the parameters to only a few simple measurable parameters. An approach to extract these new parameters is also discussed. In section 4.4, the validity of the modified GICCR and the extracted parameters are verified by comparing the model to device simulations for several doping profiles, followed by verification using measured data.

4.2 Verification of Weighting Factors

The modification of the GICCR focuses on the low current operation where it has significant influence on the Early Effect. Some of the weighting factors that are dominant in this regime of operation are h_{p0} (hole charge at zero bias Q_{p0}), h_{je} (hole charge at base-emitter junction Q_{jE}), and h_{jc} (hole charge at base-collector junction Q_{jC}). For high current operation, the GICCR modification is still consistent with the results discussed in Schroter[19] and thus, this regime will not be discussed further.

The validity of the modified GICCR is verified for several SiGe HBTs with different doping profiles and Ge mole fractions and profiles. As a reference, device simulation is used with the preference for the one-dimensional case which normally approximates quite well the situation in the present SiGe HBT technology. This can be easily extended to the two-dimensional case. But first, the one-dimensional case is compared by using the two

doping profiles and Ge mole fraction discussed in Chapter 2, the drift HBT profile with a graded Ge mole fraction profile(cf. Fig. 2.4) and the real HBT profile with a box Ge mole fraction profile(cf. Fig. 2.3). To verify the collector current obtained from the device simulation and compare to the collector current from the calculation of the modified GICCR expression, the weighting factors must be determined. A method that produces accurate results involves calculating for each quantity numerically, but this approach is cumbersome and difficult to realize in practice. A more practical method is to determine these factors by direct extraction which will be discussed in the next section. The numerical calculation for the weighting factors is achieved by using equations (B.6b) of Appendix B rather than using the simplified expression for each factors described in equations (2.9c-e) because this approach can also apply to any Ge profile, in particular the box type profile.

The expression for each weighting factor along with proper biasing scheme is as follows:

i. h_{p0} :

$$h_{p0} = \frac{Q_{p0}^*|_{SiGe}}{Q_{p0}|_{Si}} = \frac{\int_B \frac{P_0(x)}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{SiGe}}{\int_B \frac{P_0(x)}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{Si}} \quad (4.1)$$

where $P_0(x)$, $\mu_{nB}(x)$, and $n_{iB}^2(x)$ are the hole charge, electron mobility, and intrinsic carrier density at zero bias, respectively.

i. h_{je} :

$$h_{je} = \frac{Q_{jE}^*|_{SiGe}}{Q_{jE}|_{Si}} = \frac{\int_B \frac{(p(x) - P_0(x))}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{SiGe}}{\int_B \frac{(p(x) - P_0(x))}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{Si}} \quad (4.2)$$

where $P_0(x)$, $\mu_{nB}(x)$, and $n_{iB}^2(x)$ are the zero bias hole charge, electronic mobility, and intrinsic carrier density, respectively and $p(x)$ is the bias dependent hole charge which is

obtained by requiring $V_{BE} > 0$ and $V_{BC} = 0$ with I_C at low bias, usually, in the range of $1\mu A$ to $20\mu A$.

iii. h_{jc} :

$$h_{jc} = \frac{Q_{jc}^*|_{SiGe}}{Q_{jc}|_{Si}} = \frac{\int_B \frac{(p(x)|_{V_{BC1}} - p(x)|_{V_{BC2}})}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{SiGe}}{\int_B \frac{(p(x)|_{V_{BC1}} - p(x)|_{V_{BC2}})}{\mu_{nB}(x)n_{iB}^2(x)} dx|_{Si}} \quad (4.3)$$

where $P_0(x)$, $\mu_{nB}(x)$, and $n_{iB}^2(x)$ are the zero bias hole charge, electronic mobility, and intrinsic carrier density, respectively and $p(x)$ is the bias dependent hole charge which is obtained by requiring $V_{BC1} = 0$, $V_{BC2} < 0$, constant V_{BE} , and I_C again at low bias, usually in the range of $1\mu A$ to $20\mu A$.

To determine the weighting factors, the following doping profiles are used: a SiGe HBT with two different graded Ge mole fractions of 5% and 10% and a SiGe HBT profile with two box type Ge mole fractions of 8% and 28%. The purpose of having two different graded Ge mole fractions is to illustrate their relationships to the weighting factors. In addition, the influence of the graded and box Ge mole fraction profiles on the weighted factor is also investigated to demonstrate the applicability of the modified GICCR to any of the Ge profile. The results for h_{jc} and h_{je} using the equations along with the proper bias schemes described earlier for the 1D profiles are plotted as a function of collector current are shown in Figs. 4.1a and 4.1a, respectively.

FIGURE 4.1. (a) h_{jc} and (b) h_{je} vs. I_c for the four different Ge mole fraction profiles.

The curves for h_{jc} and h_{je} with different Ge mole fractions appear constant over the collector current range of interest which confirms the initial assumption discussed in Chapter 2 where the simplified expressions(2.9c, 2.9d, and 2.9e) for these weighting factors have some dependence on the space-charge widths. For h_{je} , the space-charge width applies to the base side of the base-emitter junction which changes very little over the V_{BE} bias range of interest and the result is a small variation in h_{je} for different V_{BE} 's. Similarly, the space-charge width for h_{jc} applies to the base side of the base-collector junction which again has a small variation over the V_{BC} bias range, and the result is a small variation in h_{jc} for different V_{BC} 's. Average values of h_{je} and h_{jc} and as well as the calculated value of h_{p0} are summarized in Table 4.1.

TABLE 4.1. Weighting Factors for SiGe HBTs

Parameter	graded Ge (5%)	graded Ge (10%)	box (Ge 8%)	box Ge (28%)
h_{p0}	0.74	0.56	0.25	0.004
h_{je}	0.97	0.97	0.25	0.004
h_{jc}	0.47	0.21	0.25	0.004

The results from Table 4.1 show that the values of weighting factors are the same in the box Ge SiGe transistor, whereas they are different in the graded Ge SiGe transistor. The reason is that the box Ge profile has a constant mole fraction appearing across the base region, and as a result the influence of the Ge mole fraction on the mobility and intrinsic carrier densities is the same across the region. This means that the mobility and intrinsic carrier densities would increase by $1/0.004$ or 250 times for 28% Ge mole fraction from the case without the Ge mole fraction and this constant increase is reflected in the collector current at low to moderate injection which is shown in Fig. 4.2.

FIGURE 4.2. Ratio of collector currents SiGe HBT(box Ge profile of 28%) over Si BJT

The plot contains the ratio of collector currents which is determined by taking the collector current of a Si transistor with the same net doping profile as the box SiGe transistor divided by the collector current of the box SiGe transistor with the Ge mole fraction of 28% and plotted with the collector current of Si transistor. In the plot, the curve remains constant at low injection, then increases at moderate to high injection. For SiGe HBT with the Ge mole fraction graded across the base region, the values of each weighting factor are expected to be different that is, the values are expected to increase to a finite value of one as the Ge mole fraction decreases to zero. For instance, in the case of 5% graded Ge mole fraction, the values are 0.97 for h_{je} where the mole fraction is the smallest and 0.47 for h_{jc} where the mole fraction is the largest.

With the values of the weighting factors known, the collector currents from the modified GICCR and device simulation are compared. Figs. 4.3a and 4.4b show the collector current versus V_{CEx} for the graded Ge (Ge mole fraction of 5%) and box Ge HBTs (Ge mole fraction of 28%), respectively. The modified GICCR predicts accurately for the collector currents at low to moderate injection.

FIGURE 4.3. Comparison of 1D Device Simulation and GICCR for SiGe HBT (a) graded Ge mole fraction of 5% for the bias range of $V_{BE}=(0.7,0.72,0.74,0.76,0.78,0.80,0.82)$ and (b) box type Ge mole fraction of 28% for the bias range of $V_{BE}=(0.6,0.62,0.64,0.66,0.68,0.7,0.72)$

4.3 Compact Model with Modified GICCR

The correction for the Early effect is accomplished by employing the modified GICCR equation for the transfer current derived in Chapter 2. The equation is similar to the trans-

fer current (3.6a) in HICUM discussed in Chapter 3. By expressing the modified GICCR in a form similar to (2.10a), the new equation becomes,

$$I_T = \frac{c_{10}^*}{Q_P^*} \left[\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right) \right] \quad (4.4a)$$

where c_{10}^* and Q_P^* are the new current constants containing physical quantities and the transistor area and hole charge, respectively. The expression for the new current constant in (4.4b) is a model parameter which is given as

$$c_{10}^* = (qA_E)^2 \overline{\mu_{nB} n_{iB0}^2} V_T \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right) \quad (4.4b)$$

A_{E0} is changed to A_E known as effective emitter area which is discussed in [30]. The new total hole charge of the transistor which is a function of DC bias becomes

$$Q_P^* = Q_{p0}^* + Q_{jE}^* + Q_{jC}^* + Q_f^* \quad (4.4c)$$

$Q_{p0}^* = h_{p0} Q_{p0}$, $Q_{jE}^* = Q_{jE}$, $Q_{jC}^* = h_{jC} Q_{jC}$, and Q_f^* is the minority hole charge which is given as

$$Q_f^* = h_{fE}^* Q_{fE} + h_{fB} Q_{fB} + h_{fC}^* Q_{fC} \quad (4.4d)$$

Equation (4.4a) is similar in form to (3.6a) except for the addition of several model parameters h_{p0} , h_{je} , h_{jc} , h_{fe} , h_{fb} , and h_{fc} . Some simplification of the modified GICCR is required to ease the task of determining these parameters - in some cases they are extremely difficult or impossible to obtain from direct measurements. The first step is to simplify the weighting factors h_{fE} , h_{fB} , and h_{fC} associated with the minority carrier hole charges in each region of the transistor. This involves lumping them into a single model parameter h_c [19] which is in agreement to the single lumped value for Q_f representing the minority hole charges of the transistor. The next step is to reduce model parameters h_{p0} , h_{je} , and h_{jc} which affect the collector current at low to moderate injection. This involves dividing (4.4a) by h_{je} . The modified GICCR expression becomes

$$I_T = \frac{c'_{10} \left[\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right) \right]}{Q'_{p0} + Q_{jE} + h'_{jc} Q_{jC} + h'_c Q_f} \quad (4.5)$$

where $c'_{10} = c_{10}/h_{je}$, $Q'_{p0} = Q_{p0}h_{p0}/h_{je}$, $h'_{jc} = h_{jc}/h_{je}$, and $h'_c = h_c/h_{je}$. This simplifies the task of determining these parameters while maintaining the original definition of GICCR. This is particular useful for a box Ge profile where h'_{jc} and h_{p0}/h_{je} are reduced to one since the weighting factors have the same value as discussed in the last section. The model parameters for the new GICCR transfer current are now c'_{10} , Q'_{p0} , h'_{jc} , and h'_c .

4.3.1 Extraction of Model Parameters

The approach to determine these weighting factors discussed in the last section only applies to numerical data. A more practical approach to determine these new parameters is by using a direct extraction technique similar to the one discussed in Chapter 3. By measuring the transistor at different regime of operation, the modified GICCR equation can be simplified to a linear expression. The slope and intercept of the resulting measured data can then be applied to obtain the required model parameters. A systematic method and the sequence in which these parameters c'_{10} , Q'_{p0} , h'_{jc} , and h'_c are determined are discussed below.

(i). Extraction of c'_{10} and Q'_{p0}

This involves biasing the transistor at low injection¹ with the zero potential across the base-collector junction ($V_{BCx}=0$) and the variation of the potential at the base-emitter junction (V_{BEx}) typically in the range of 0.7 to 0.8V. This biasing scheme accomplishes two goals (1) the base-collector charge (Q_{jC}) is zero due to zero voltage at base-collector junction and (2) the low injection reduces the influence of minority hole charge, that is $Q_f \ll (Q'_{p0} + Q_{jE})$, and as a result Q_f can be neglected. With this operating condition, the modified GICCR expression in (4.5) can be simplified and re-arranged into a linear equation as

1. At low injection, the influence of series resistance on the terminal voltage is negligible and as a result $V_{BEx} = V_{BEi}$ and $V_{BCx} = V_{BCi}$.

$$Q_{jE} = -Q'_{p0} + \frac{c'_{10}}{I_T} \left(\exp\left(\frac{V_{BEi}}{V_T}\right) - 1 \right). \quad (4.6)$$

The slope and intercept of the least square fit of the plot of Q_{jE} (calculated from C_{jE} vs. V_{BEx}) versus $(\exp(V_{BEi}/V_T)-1)/I_T$ (calculated from measured data) represent c'_{10} and Q'_{p0} , respectively.

(ii). Extraction of h'_{jc}

For the extraction of h'_{jc} , the transistor is biased at low injection¹ with the base-emitter potential (V_{BEx}) equal to a fixed value between 0.7 to 0.8V and the collector-emitter potential is varied from V_{BEi} to $0.8 \cdot BV_{CE0}$ or an appropriate range to avoid the influence of avalanche effect on the extraction result. Again, at low bias operating condition Q_f can be neglected. By re-arranging (4.5), the result is a linear equation which is given by

$$\frac{c'_{10}}{I_T} \left(\exp\left(\frac{V_{BEi}}{V_T}\right) - \exp\left(\frac{V_{BCi}}{V_T}\right) \right) = Q_{jE} + Q'_{p0} + h'_{jc} Q_{jC} \quad (4.7)$$

Using the results from (i) the left-hand side of (4.7) is calculated and the result is plotted as a function of Q_{jC} (calculated from C_{jC} vs. V_{BC}). The slope from the least square fit of the curve is h'_{jc} .

(iii). Extraction of h'_C

A direct extraction technique is not necessary since it can be obtained by fitting the collector current at high injection where the minority hole charge dominates. The value is usually greater than one.

4.4 Results and Discussion

The verification of the compact model with the GICCR is performed in several steps. The first step is to verify the extraction method and to ensure that it is applicable to both SiGe HBTs and Si BJTs. In the second step, the compact model with the GICCR is verified for two SiGe HBTs in 1D simulated results. Finally, the extraction method and the compact model is applied to measured data.

The extraction method discussed in the last section is verified by applying it to the 1D simulated results of SiGe HBT with graded Ge mole fraction of 5% and to ensure that method is also applicable to Si BJT, the method is applied to the 1D simulated results for the same net doping profile as in SiGe HBT except the Ge profile is removed. Figs. 4.4a and 4.4b contain the required plots in the extraction of model parameters for the collector current and the results are summarized in Table 4.2

FIGURE 4.4. (a) Extraction of c'_{10} and Q'_{p0} (b) h'_{jc} for SiGe(Ge 5%) and Si w/o Ge.

TABLE 4.2. Extraction Results of SiGe(Ge of 5%) and Si w/o Ge

Parameter	drift SiGe (Ge=5%)	Si w/o Ge
Q'_{p0}	1.019E-14	1.375E-14
c'_{10}	2.8E-32	2.7E-32
h'_{jc}	0.44	1.0

The value of h'_{jc} for Si transistor is one which is as expected since all the weighting factors are expected to reduce to one in the absence of Ge mole fraction. The value of h'_{jc} for SiGe is 0.44 which is almost the same value as h_{jc} in Table 4.1 for the same transistor. Using the results from Table 4.2, the determination of each weighting factor is possible by using the relation described in (4.5). For instance, since $c'_{10} = c_{10}/h_{je}$ and with the values of c_{10} ($=2.7E-32$ for Si) and c'_{10} ($=2.8E-32$ for SiGe), the calculated value of h_{je} becomes 0.96. With h_{je} , Q'_{p0} ($=1.019E-14$), and Q_{p0} ($=1.375E-14$) known, the value of h_{p0} determined from the relation $Q'_{p0} = Q_{p0}h_{p0}/h_{je}$ becomes 0.71. Finally, the value of h_{jc}

determined from the relation $h'_{jc}=h_{jc}/h_{je}$ is 0.458. The results of the extraction method and the numerical method for the same SiGe HBT are summarized in Table 4.3.

TABLE 4.3. Comparison of Numerical and Direct Extraction Methods

Parameter	Numerical method	Extraction method	Relative Error (%)
h_{p0}	0.74	0.71	4.1
h_{je}	0.97	0.96	1.0
h_{jc}	0.47	0.458	2.6

The small relative error for each weighting factor indicates that the direct extraction method is very accurate. To further illustrate the accuracy, the simulation results from the compact model with the extraction parameter set are compared to 1D numerical simulation. Figs. 4.5a and 4.5b show the output characteristics (I_C vs. V_{CEx}) for Si BJT and SiGe HBT (graded Ge of 5%), respectively. In Fig. 4.5a, the new compact model with the parameter set shows good agreements with the numerical simulation. In Fig. 4.5b, the modified GICCR curve, determined from the direct calculation of the expression discussed earlier, is re-plotted again and two additional curves from the compact model with the modified GICCR and the compact model with the original GICCR. The original GICCR shows significant deviations from device simulations in the active region. However, an excellent agreement is achieved from the modified GICCR in the same region.

FIGURE 4.5. Comparison of 1D Simulation and Compact Model for (a) Si without Ge for the bias range of $V_{CEx}=(0.7,0.72,0.74,0.76,0.78,0.80,0.82,0.83)$ and (b) SiGe(graded Ge of 5%) HBT for the bias range of $V_{BEx}=(0.7,0.72,0.74,0.76,0.78,0.80,0.82)$.

Fig. 4.6 shows a comparison between the compact model and the 1D simulation for SiGe HBT with a box Ge profile of 28% mole fraction. The modified GICCR curve is again re-plotted to illustrate the accuracy of the calculated GICCR using numerical calculated parameters and the compact model. A small deviation is observed between the results from the compact model to that of 1D simulated results, but in general the model prediction is reasonably accurate.

FIGURE 4.6. Comparison of 1D Simulation and Compact Model for (SiGe(box Ge of 28%))HBT for the bias range of $V_{BE} = (0.6, 0.62, 0.64, 0.66, 0.68, 0.7, 0.72)$.

Finally, the extraction method and the compact model are applied to measured data. The transistor used in this case is a SiGe HBT of an emitter area $0.5 \times 10 \mu\text{m}^2$. The plot for the extraction of h'_{jc} is shown in Fig. 4.7a and the extracted value is 0.38. Fig. 4.7b shows the output characteristics comparing measured data to compact model with the original and the modified GICCR.

FIGURE 4.7. (a) LHS of (4.7) vs. Q_{jc} and (b) Output Characteristics for a SiGe HBT of emitter area $0.5 \times 10 \mu\text{m}^2$

The compact model provides an excellent agreement to measured data with h'_{jc} , in the active region except for a small aberration in the saturation region where the model is not

able to accurately capture this operating regime well, whereas large deviations are observed for the case without h'_{jc} . Furthermore, the curves from the compact model with the modified GICCR show some mismatched to the measured curves at quasi-saturation. This is due to the known inherent limitations of HICUM.

Chapter 5 - Extraction of Thermal Parameters

5.1 Introduction

The impact of self-heating is becoming an important issue in high-speed integrated bipolar junction transistors, especially, in SiGe-based integrated heterojunction bipolar transistors (HBTs). Self-heating occurs when the transistor is operating in a forward biased mode of operation, the transistor's internal power dissipation causes the intrinsic junction temperature to increase above the ambient temperature of the chip, resulting in a significant influence in the transistor performances. This can be a problem in compact models because if proper de-embedding of heating effects is not considered, it can cause significant deviations of extracted model parameters from their isothermal values resulting in transistors not accurately modelled. This can result in transistors not fully utilized for maximum performances in circuits. Thus, it is critical that self-heating should be included in a compact model, and as well as proper de-embedding of model parameters from self-heating effects is important in order to obtain good agreement between simulation results to measured data. HICUM contains a thermal network subcircuit[27] to model self-heating. However, some work is underway to include self-heating, using a similar thermal network[27], in VBIC95, an extension of the existing de-facto standard SPICE Gummel Poon Model (SGPM).

In this chapter, the physical reasons for the self-heating in both high-speed Si NPN bipolar transistor technology and SiGe-based HBT technology and how it can influence the electrical characteristics of the transistor are examined using MEDICI[31], a 2D commercial numerical simulator. A brief discussion on the derivation of a simple thermal network from a one-dimensional heat-flow equation is examined. A detail discussion on the extraction procedures of thermal parameters that include some theory, followed by verifications of extracted parameters by comparing simulation to measurement for both Si and SiGe-based NPN bipolar transistor technologies are examined. The extraction procedures will show that the parameters can be easily obtained without requiring complex measurement setups that are often costly and time consuming. A simple

approach is presented. A brief discussion on the scalable thermal parameters and as well as the validation of the self-heating model a simple current source circuit are examined

5.2 MEDICI Simulation

The physical reasons for self-heating in a NPN bipolar transistor and how it can influence the electrical characteristics of the transistor can best be explained using MEDICI, a 2D numerical simulator, which can include solutions of the lattice heat equation in the analysis. Furthermore, it can be used to obtain some insights into the temperature distribution within the transistor.

Fig.5.1a shows a full cross-section of a self-aligned double-poly Si NPN transistor structure with contacts base, emitter, and collector are located along the top of the structure and the substrate contact at about six micron deep of the structure. The shaded regions shown in the figure represent the oxide isolation.

FIGURE 5.1 (a) A Full cross-section of an advanced Si NPN transistor, (b) Zoom along the vertical axis of (a)

The region (surrounded by the dashed box) underneath the emitter window area represents the intrinsic region of the transistor, also known as the active device, where most of the transistor action takes place. The region outside the dashed box represents the extrinsic region of the transistor. Fig.5.1b shows a zoom-in cross-section of the intrinsic transistor region. The six micron thickness for the substrate is adequate to get an idea of heat flows

in the substrate without increasing simulation time. The structure is spatially discretized by using a rectangular grid. Fig. 5.2 shows a plot of the structure with the grid points and most of the grid points are concentrating at around the intrinsic transistor region and few grid points outside that region. There are a total of 1040 grid points. This structure along with the mesh refinement were developed and tested by Nortel.

FIGURE 5.2 Structure with Grid points

Fig.5.3 shows a 1D net doping concentration of the intrinsic transistor taken along the AA' cross section of Fig. 5.1.

FIGURE 5.3 1D Net Doping Concentration

The base region has a decreasing doping profile that is important in terms of transit time because major contribution of the transit time is the base transit time. By introducing a decreasing doping profile in the base an electrostatic field, also referred to as an accelerating electric field, is produced in the base region that can aid the flow of minority carriers through the base region. The result is a smaller base transit time. Similarly, the exponential

doping also known as selectively implanted collector(SIC) and the reasons are already discussed in Chapter 2.

5.2.1 Simulation Conditions

For the simulation, the heat flow equation, Poisson's equation, and current-continuity equations are solved simultaneously. A Homogeneous Neuman boundary condition applies to the boundaries of the structure including the contacts except for the substrate contact where it is used as a heat sink at a room temperature of 300K. In other words, the Neuman boundary is based on the assumption that the temperature outside the boundary of the device including the contacts except for the substrate is an adiabatic condition, that is no exchange of heat or the heat flux represented by dT/dx where T and x represent the temperature and displacement, respectively, is zero. Finally, the impact ionization or avalanche/breakdown effect is also included in the simulation.

A transistor of emitter window width, b_{E0} , $0.5\mu\text{m}$ is used in a 2D simulation with the transistor operating in a common emitter configuration as shown in Fig. 5.4.

FIGURE 5.4 A Common Emitter Configuration

5.2.1.1 Results and Discussion

The next couple of sections will discuss some of the electrical characteristics and internal quantities observed from the simulations.

5.2.1.2 Electrical Characteristics

A Gummel plot (forward bias I_C/A_E and I_B/A_E as function of V_{BEX}) for the cases with and without self-heating from a MEDICI simulation for a V_{CEX} of 1.5V is shown in

Fig.5.5. In the plot, the dashed and solid lines represent the cases with and without self-heating, respectively. The onset of self-heating effect does not occur until the collector current density is about $0.05\text{mA}/\mu\text{m}^2$. The currents with the self-heating case become larger than the case without self-heating for increasing V_{BEX} . The reason for the larger currents is mainly due to the temperature dependent intrinsic carrier density component of the saturation current or ICCR constant[18], depending on the different bipolar models. The rise in junction temperature due to an increase in internal power dissipation can thermally excite more intrinsic carriers into the conduction band that in turns increase the saturation current or ICCR constant. The result is an increase current. This effect, as well as how the current can be used as a thermometer to determine an approximate rise in junction temperature, are discussed later.

FIGURE 5.5 A Gummel Plot of a $0.5\mu\text{m}$ advanced silicon NPN Transistor

Fig. 5.6 shows the DC forward current gain (Beta or β), where β is equal to the ratio of I_C over I_B , for the same V_{CEX} value. Again, the dash-line curve represents the case with self-heating and the solid-line curve represents the case without self-heating. The DC gain for the case with self-heating is observed to be larger than without self-heating and this occurs at high collector current densities. The reason for larger current gain is again due to the larger intrinsic carrier densities as a result of an increase in junction temperature.

FIGURE 5.6 Beta vs. I_C/A_E

Finally, to observe the effect of self-heating on AC small signal, a figure of merit such as the transition frequency (f_T) or unity gain frequency of a common emitter configuration is used. Fig.5.7 shows a plot of the transition frequency as a function of collector current densities.

FIGURE 5.7 Transition Frequency (f_T) vs. I_C/A_E

The plot shows little or no difference between the two curves except for a small deviation at peak f_T . Since most of the temperature increase is due to the DC bias, f_T changes according to a steady state temperature increase during each DC bias point, but the change is very small because the junction temperature change is very small and the various components of f_T can partially compensate each other.

5.2.1.3 Physical Characteristics

To get an idea of what factor(s) contributes to the self-heating effect of the transistor, this section will look at some internal quantities for several DC operating points. Fig. 5.8 show a 3D surface temperature plot for an operating point of V_{CEX} and I_C/A_E of 1.5V and $0.2\text{mA}/\mu\text{m}^2$, respectively. The plot shows the highest temperature occurring in the base-collector region of the intrinsic transistor and it is relatively constant through the emitter contact. The small peak temperature at the base-collector space charge region(BC SCR) of the intrinsic transistor is about 301.5K or an increase of 1.5°C from a room temperature of 27°C . The temperature is decreasing nonlinearly from the peak value of the intrinsic region to the surrounding region, and it is also nonlinear from the intrinsic region down to the substrate contact.

FIGURE 5.8 3D plot of Temperature Distribution

A better view of the temperature distribution can be observed in the temperature contour along the cross-section of the structure shown in Fig.5.9a. The figure shows the maximum temperature occurs roughly in the base-collector junction, directly underneath the emitter window. Fig.5.9b shows a 1D plot of temperature as a function of transistor depth at y equal 2.25 micron which is shown as the dotted line through the intrinsic region of the transistor in Fig. 5.9a.

FIGURE 5.9 (a) Temperature Contour along the cross-section of the structure at $V_{BE} = 0.91V$ ($\Delta T = 0.5^\circ C$) (b) 1D Temperature at $y = 0.225\mu m$ for circle($V_{BE} = 0.94V$), square($V_{BE} = 0.93V$), triangle($V_{BE} = 0.91V$), +($V_{BE} = 0.89V$), and x($V_{BE} = 0.87V$)

The plot shows the temperature slightly nonlinear from the peak to the substrate region and the temperature from the emitter poly to roughly where the base-collector junction regions is relatively constant. The reason for the peak temperature occurring near the base-collector space charge region is a result of a large power density drop in that region causing the lattice to heat up. Power density is defined as the electric field times the collector current densities. In a space charge region where the field is fairly high for substantial reverse bias, minority carriers can be accelerated from the base toward the collector region by a high electric field to a point where carriers can acquire sufficient energy to collide with the lattice atoms that could knock loose electrons from the lattice atom. The collision of carriers on lattice atoms can cause the lattice to heat. Since the collector current is associated with minority carriers, then for a larger collector current hence a larger power density, more minority carriers can travel through the space charge region that results in more chance for carriers and lattice collision to occur. The result is a higher junction temperature. In a SiGe transistor technology, the collector current is about 1.5 to 2 times higher than the normal Si transistor technology for the same bias, this translates to about 1.5 to 2 times higher in junction temperature. An additional note, the peak temperatures shown in Fig. 5. 9b seem to shift away from the dash line, where the dash line represents the peak temperature for V_{BE} of 0.87V, for increasing V_{BE} values.

The reason is due to the Kirk effect that causes the neutral base to extend to epi region that forces the space charge region to epi and buried layer region.

5.3 A Simple Thermal Subcircuit

This section will discuss briefly on the derivation of the thermal network for a compact model. The previous sections conclude that most of heat generated occurs at the base-collector junction in the intrinsic region. Furthermore, the entire intrinsic transistor can be regarded as a heat source sitting on a large substrate and heat tends to flow from the hot region toward the cool region[28]. There are three mechanisms in which heat can flow and they are radiation, convection, and conduction. In terms of an integrated transistor, heat flow is by conduction. The temperature generated from the device will conduct toward the substrate because with the substrate several hundred micron deep, the temperature at the backside of the wafer is held at room temperature, 300K. There is a small heat flow through the surface of the transistor. The amount of heat flowing through the surface is small because usually, the surface is covered with a thick oxide resulting in a low thermal conductivity. whereas the substrate is a plain silicon which has a higher thermal conductivity. and as a result, most of the heat flow would be through the substrate as shown in Fig. 5.10.

FIGURE 5.10 Overview of heat flow through the semiconductor substrate

The temperature distribution of the device in a substrate can be explained with the general three-dimensional Heat Flow Equation,

$$\nabla K(T) \nabla T(x, y, z) - \rho c \left(\frac{\partial T}{\partial t} \right) = -P(x, y, z) \quad (5.1)$$

where K is the thermal conductivity, T the absolute temperature, ρ the density, c the specific heat, and t the time. To apply the heat flow concept to a compact model, the 3D equation has to be simplified to a first-order ordinary differential equation which results in (5.2) incorporating lumped-thermal elements such as thermal resistance R_{TH} and thermal capacitance C_{TH} to represent heating effects of a integrated transistor. Equation 5.2 represents a first order self-heating model that states the junction temperature $T_j(t)$ in unit Kelvins as

$$T_j(t) = T_c + R_{TH}P_T - C_{TH}R_{TH}\frac{dT_j}{dt} \quad (5.2)$$

where T_c (K) represents the chip temperature (or temperature at backside of a wafer) usually equal to room temperature, R_{TH} (K/W) and C_{TH} (Ws/K) or (J/K) are the thermal resistance and capacitance model parameters, respectively, and finally, P_T (W) equals to the total power dissipation. The equation suggests that for any situation in which power, P_T , is being dissipated, with a consequent rise in temperature with time, the larger the thermal capacitance, the lower will be the temperature rise at any given time. The equation can be translated to a very close analogy of an electrical RC network in which voltage replaces by temperature, current replaces by power, and electrical resistance and capacitance replace by thermal resistance and capacitance, respectively. Fig. 5.11 shows an electrical analog representation of a Self-Heating subcircuit[29] with P_T as a current source and T_j as voltage node.

FIGURE 5.11 An Equivalent Electrical Subcircuit for Self-Heating

The equivalent electrical subcircuit for self-heating is a first order approximation. However, higher order or more RC sections can be employed to improve the accuracy of

the self-heating model as discussed in Roulston[27], but with more RC sections, additional nodes are added to the overall equivalent transistor circuit that will increase the matrix size resulting in a substantial increase in computation time. This can make compact modeling impractical from a computational time point of view to justify for a possible small increase in accuracy. Thus, it is desirable to incorporate a self-heating circuit into an equivalent transistor circuit with minimum nodes.

5.3.1 Implementation of Self-Heating Model in a Compact Model

The procedure in which a circuit simulator solves the self-heating subcircuit in conjunction with the bjt equivalent circuit is as follows. The self-heating subcircuit shown in Fig. 5.11 is used to take into account the increase of the transistor's junction temperature T_j as a result of self-heating. This is accomplished by taking the total power dissipation determined from the bjt equivalent circuit and the self-heating model takes the P_T value and solves together with the model parameters, R_{TH} and C_{TH} for DC and transient operation to obtain the junction temperature at node T_j . The new local junction temperature is passed back to the BJT model routine that contains the bjt equivalent circuit to modify temperature dependent parameters. Fig.5.12 shows the exchange of variables between two models.

FIGURE 5.12 Variable links between the BJT model and the Self-Heating Model

The exchange of T_j and P_T variables between the two models is iterated until the solution converges.

The next section will discuss procedures to extract thermal parameters. The self-heating circuit in Fig. 5.12 is already implemented in HICUM, and the compact model, HICUM, will be used in all the analyses unless otherwise stated.

5.4 Extraction of Thermal Parameters

The extraction of thermal parameters begins with the determination of thermal resistance followed by the thermal capacitance. The reason for the order of extraction is because the thermal resistance can be easily obtained under steady state(DC) condition, whereas the extraction of C_{TH} is found from a transient measurement which also includes the effect of R_{TH} . Presently, there are many well known methods[32-38] to determine R_{TH} and C_{TH} but most of these methods require complicated and expensive equipments and are often time consuming. The next several sections will discuss simpler methods to determine thermal model parameters.

5.4.1 Extraction of Thermal Resistance

The determination of thermal resistance from an integrated transistor under a DC operation is found by reducing (5.2) to the following

$$R_{TH} = \frac{T_j - T_c}{P_T} \quad (5.3)$$

where T_j (K) is the steady state junction temperature, T_c is the room temperature 300K, and P_T is the total power dissipation.

One of the simplest method to determine thermal resistance is by evaluating (5.3) directly. T_c and P_T are two known variables since T_c represents the ambient or room temperature and P_T represents the total power dissipation which can be specified as an input power to the transistor. Then, the only unknown is T_j . Several techniques[39] have been shown to be useful in measuring junction temperature and they are optical infra-red (IR) thermometry, liquid crystal technique, gate diode voltage, and gate stripe resistor. Most of these methods require the wafer to be specially prepared that can be time consuming or special instruments that may not be available in most laboratories. An alternative method is to use an analytical expression for thermal resistance to be discussed in the next section that solves for the thermal resistance by using the data acquired from a simple DC IV measurement.

5.4.1.1 An Analytical Expression for R_{TH}

An analytical R_{TH} expression shown in (5.4) is derived by Reisch[34],

$$R_{TH} = \frac{T^2 k}{E_g - q \cdot V_{be}} \cdot \frac{\partial}{\partial P_T} \left(\frac{\Delta I_b}{I_b} \right) \quad (5.4)$$

where T = temperature (K)

V_{be} = base-emitter voltage (V)

I_b = base current (A)

E_g = bandgap energy at 300K for Si (1.12V)

k = Boltzman's Constant 1.38E-23 J

q = electron charge of 1.602E-19 C

R_{TH} is determined by solving the above equation. All the variables in the equation are known except for $d(\Delta I_b/I_b)/dP_T$ or the slope of $\Delta I_b/I_b$ versus P_T (dissipated power) curve and T which is assumed to be equal to the ambient temperature (300K). The slope can be found by monitoring the I_b versus V_{cb} curve for a constant V_{be} value (refer to Section 5.4.1.3 for more detail in measurement setup). To verify the accuracy of the analytical expression, HICUM, a compact model, is used to generate the I_b versus V_{cb} curve. For the simulation, a known R_{TH} value is entered into the compact model. The result from extraction method shows that R_{TH} is about 30 percent lower than the actual R_{TH} value used in the compact model. The large error may be attributed to the assumptions made in the derivation of the analytical expression. Unfortunately, the derivation of the analytical expression was not available in Reisch[34] publication. By using (5.4) as basis, a new analytical expression for R_{TH} can be developed.

The next section discusses the theory involved in deriving an analytical expression for the thermal resistance relating to the electrical characteristics of the transistor, and to employ a direct extraction technique by using slope and intercept acquired from the least square fit of measured data to determine the thermal resistance directly.

5.4.1.2 Derivation of A New Analytical Expression for R_{TH}

As mentioned earlier, the main temperature contribution to self-heating effects in a transistor is due to the intrinsic carrier concentration, $n_i(T)$. Since the saturation current[17] in both base and collector currents includes $n_i(T)$ in the expression, then the currents can be used as a thermometer. However, the base current should be used rather than the collector current because it is only a parasitic current, whereas the collector current is influenced by other physical effects such as Early effect and Kirk effect which make it difficult to separate these effects from self-heating.

The dependence of base current on base emitter voltage can be modeled as[17],

$$I_B = I_{BESi} \cdot \left(\exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) - 1 \right) + I_{RESi} \cdot \left(\exp\left(\frac{V_{BEi}}{m_{rEi} V_T}\right) - 1 \right) \quad (5.5)$$

where $I_{(B,R)ESi}$ =saturation current (injection,recombination) (A)

$m_{(B,R)Ei}$ =Ideality factor

V_{BEi} =Internal Base-Emitter voltage (V)

Equation (5.5) models two currents, recombination and injection. Usually, the recombination current is associated with hole carriers being trapped by impurities residing in the base-emitter space charge region (BE SCR) and the mechanism is referred to as Shockley-Read-Hall (SRH) recombination[12]. In addition to hole carriers being trapped in BE SCR, hole carriers can also recombine with electron carriers in the neutral base region but for modern high-speed bipolar, there is no recombination in the neutral base region because the diffusion length, that is related to life time of electron carriers in this case, is much greater than the base width. Hence, recombination current is largely due to Auger Recombination in BE SCR. On the other hand, the injection current is due to hole carriers injected from the base into the emitter region. From the Gummel plot shown in Fig. 5.5 of section 5.1.1.2, it shows that the effect of self-heating is not significant until I_C/A_E is greater than $0.05\text{mA}/\mu\text{m}^2$, and in this regime of operation the main contribution of base current is due the injection current component. Thus, the recombination current component is assumed to be much less than the injection current component, i.e.,

$$I_{RESi} \cdot \left(\exp\left(\frac{V_{BEi}}{m_{rEi} V_T}\right) - 1 \right) \ll I_{BESi} \cdot \left(\exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) - 1 \right) \quad (5.6)$$

and the exponential term in the injection component is much greater than one for the operating range of interest ($V_{BE} > 0.7V$) i.e.,

$$\exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) - 1 \approx \exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) \quad (5.7)$$

With the above approximations, the base current reduces to:

$$I_B \approx I_{BESi} \cdot \exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) \quad (5.8)$$

The saturation current of the injection component can be modeled as[19]

$$I_{BESi} = \frac{q A_E V_T \mu_{pE} n_i^2}{N_E W_E} \quad (5.9)$$

where N_E and W_E are emitter doping concentration and width, and the temperature dependent variables are μ_{pE} , V_T , and n_i^2 . From the three temperature dependent variables only n_i^2 is strongly influenced by temperature because of its exponential dependence, whereas the product, $\mu_{pE} * V_T$, represents the diffusion coefficient which is relatively constant over the entire temperature range due to the high emitter doping concentrations[27]. The temperature dependence of intrinsic carrier concentration is expressed as[27]

$$n_i(T) = \left[4 \left(\frac{\pi k}{h^2} \right)^3 (m_e m_h)^{3/2} T^3 \right]^{1/2} \cdot \exp\left(\frac{-E_g + \Delta E_{ge0}}{2kT}\right) \quad (5.10)$$

where

$m_{e,h}$ = Effective electron/hole masses (9.11E-31 kg)

T = Absolute Temperature (K)

h = Planck's constant (6.63E-34 Js)

k = Boltzmann's constant (1.38E-23 J/K)

E_g = Bandgap-Energy (1.12eV for silicon)

ΔE_{ge0} = Bandgap Narrowing Energy in the emitter (include bandgap reduction due to SiGe HBT) (eV)

Equation 5.10 assumes that the bandgap-energy (E_g) is relatively constant in a regime where the extraction of R_{TH} takes place since for a small rise in junction temperature in a regime where the extraction of R_{TH} takes place. To justify the assumption, the empirical relation of band-energy[27] with respect to absolute temperature is given by

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \quad (5.11)$$

where α and β are material contents and $E_g(0)$ is the band-energy at absolute temperature. In silicon, $E_g(0)=1.17\text{eV}$, $\alpha=4.73\text{E-}4/\text{K}$ and $\beta=636$, the resultant E_g at 300K is 1.1245eV but for a slight increase in temperature, for example 310K, E_g becomes 1.122eV. Thus, the increase is approximately 0.22% which is small.

The substitution of (5.9) along with (5.10) into (5.8) yields

$$I_B = \frac{q A_E V_T \mu_{pE}}{N_E W_E} \cdot 4 \left(\frac{\pi k}{h^2} \right)^3 (m_e m_h)^{3/2} T^3 \cdot \exp\left(\frac{-E_g + \Delta E_{ge0}}{kT}\right) \cdot \exp\left(\frac{V_{BEi}}{m_{BEi} V_T}\right) \quad (5.12)$$

By taking the ratio of $I_B(T)$ over $I_B(T_0)$ where T_0 equal to room temperature, the ratio becomes

$$\frac{I_B(T)}{I_B(T_0)} = \left(\frac{T}{T_0} \right)^3 \cdot \exp\left[\left(\frac{1}{T_0} - \frac{1}{T}\right)\left(\frac{V_g}{k/q} - \frac{V_{BEi}}{m_{BEi} k/q}\right) - \alpha_\beta \Delta T\right] \quad (5.13)$$

where $\Delta E_{ge0}/k(1/T-1/T_0)=\alpha_\beta \Delta T$ where $\Delta T=T-T_0$ (refer to Appendix C) and α_β as a model parameter. Equation 5.14 expresses junction temperature in terms of P_T by rearranging (5.3)

$$T = R_{TH} P_T + T_0 \quad (5.14)$$

where $P_T = I_C V_{CEX} + I_B V_{BEX}$ with I_C and I_B as collector and base currents, respectively, and V_{CEX} and V_{BEX} as terminal collector-emitter and base-emitter voltages, respectively.

By taking the natural logarithm of (5.13) and then differentiating the equation with respect to P_T , the result becomes

$$\frac{d}{dP_T} \left(\ln \frac{I_B(T)}{I_B(T_0)} \right) = \frac{3dT}{T^2 dP_T} + \frac{1}{T^2 k/q} \left(V_g - \frac{V_{BEi}}{m_{BEi}} \right) \frac{dT}{dP_T} - \alpha_\beta \frac{dT}{dP_T} \quad (5.15)$$

Equation 5.14 is differentiated w.r.t. P_T that gives $R_{TH} = dT/dP_T$ and by substituting the result into (5.15) and solves for R_{TH} , it yields the following expression,

$$R_{TH} = \frac{T^2 k/q}{3T k/q + V_g - V_{BEi}/m_{BEi} - \alpha_\beta T^2 k/q} \cdot \frac{d}{dP_T} \left(\ln \frac{I_B(T)}{I_B(T_0)} \right) \quad (5.16)$$

The R_{TH} equation is very similar to the expression derived by Reisch shown in (5.4) and for a valid comparison, the $\ln(I_B(T)/I_B(T_0))$ term in the derivative of (5.16) is simplified by applying the Maclaurin series. For a first order approximation, the term becomes

$$\ln \left(\frac{I_B(T)}{I_B(T_0)} \right) = \frac{I_B(T) - I_B(T_0)}{I_B(T_0)} = \frac{\Delta I_B(T_0)}{I_B(T_0)} \quad (5.17)$$

By substituting (5.17) into (5.16) and times top and bottom of (5.16) by q , the equation simplifies to

$$R_{TH} = \frac{T^2 k}{E_g - V_{BEi}/(q m_{BEi}) + 3kT - \alpha_\beta T^2 k} \cdot \frac{d}{dP_T} \left(\frac{\Delta I_B(T)}{I_B(T_0)} \right) \quad (5.18)$$

Comparing (5.18) to (5.4), two new terms $3kT$ and $\alpha_\beta T^2 k$ are added and also an ideality factor (m_{BEi}) is added to the denominator of V_{BEi} term. The difference in the two analytical expressions for R_{TH} can introduce some errors in the extraction which will be discussed in the verification of R_{TH} section. The next section will discuss the measurement setup.

5.4.1.3 Measurement Setup

The extraction of R_{TH} is straight-forward because from R_{TH} expression, all the variables are known except for the following: m_{BEi} , α_β and the derivative of $\ln(I_B/I_{B0})$ w.r.t.

P_T . The model parameters m_{BEi} and α_β can be easily determined using the methods discussed in Chapter 3. For the derivative of $\ln(I_B/I_{B0})$ w.r.t. P_T , it can be found from the slope of the least squares fit of $\ln(I_B/I_{B0})$ versus P_T plot. To get the plot, the transistor is biased with fixed base-emitter voltage while a sweeping voltage is applied at the collector. The schematic for the circuit to perform the necessary measurement is shown in Fig. 5.13.

FIGURE 5.13 Circuit Schematic Measurement Setup

To achieve significant self-heating without operating the transistor into the high injection region, a fixed V_{BEEX} should be in the range of 0.8 to 0.9V and V_{CEX} is swept from V_{BEEX} up to BV_{CE0} . The reason for not wanting to operate the transistor at high injection is because of the series resistance, it can influence the base current since base current is used to determine the slope or derivative of $\ln(I_B/I_{B0})$ w.r.t. P_T that is used in the R_{TH} calculation. In addition, by sweeping the collector of a transistor with a voltage source as recommended, the influence of avalanche effect can also affect the slope as well. These effects are analyzed in some detail in the sensitivity section. The procedure to extract R_{TH} is as follows:

- i. use extracted α_β , m_{BEi} and R_E (to be discussed in the sensitivity analysis section)
- ii. use $V_g=1.14V$ to $1.17V$ (Si & SiGe, respectively), $k=1.38E-23$ J/K and $q=1.602E-19$ C
- iii. $V_{BEi} = V_{BEEX} - R_E I_C$

5.4.1.4 Sensitivity Analysis

Since the extraction of R_{TH} must take place in the medium to high collector current density regime, the influence of series resistance more specifically emitter resistance and avalanche effect may affect the accuracy of R_{TH} . To get an idea of the influence of series resistance and avalanche have on R_{TH} , a sensitivity analysis is carried out by systematically turning off each effect. It is much easier to use a compact model such as HICUM to do the analysis rather than using a numerical simulator such as MEDICI because in a compact model, avalanche and emitter resistance are modeled with parameter values and these values can be easily turned off, whereas a numerical simulator can turn avalanche off but for emitter resistance it is difficult to turn off for a 2D simulation. Moreover, the simulation will take considerably longer to do in MEDICI than in a compact model. For ease of analysis, a compact model is used.

Using the measurement setup discuss in section 5.4.1.3, HICUM is used to generate the data to make possible the plot of $\ln(I_B/I_{B0})$ versus P_T shown in Fig. 5.14. The simulation is done for an advanced double-poly Si NPN transistor (emitter window area(A_{E0}) of $0.5 \times 20 \mu m^2$) model parameter set.

FIGURE 5.14 $\ln(I_B/I_{B0})$ vs. δP_T for an advanced double-poly Si NPN Transistor size(0.5x20) model parameter set.

The curves in the plot correspond to various simulation cases given in Table 5.1.

TABLE 5.1. Simulation Cases

Case	Parameter Condition			Curves in plot
	R_{TH}	Avalanche	R_E	
1	ON	ON	ON	curve 1
2	OFF	ON	ON	curve 2
3	OFF	OFF	ON	curve 3
4	OFF	OFF	OFF	curve 4

For case 1 with the normal parameter set including avalanche effect and emitter resistance R_E and a R_{TH} value of 548K/W, curve 1 shows the simulated result. The curve exhibits a positive slope up to about 1mW, and then, a negative slope. The positive slope or the rising part of the curve is due to self-heating and the negative slope or the falling part of the curve is due to the avalanche effect. To justify these observed effects, for case 2, R_{TH} is set to zero and the simulation result is shown as curve 2 in the figure. The curve exhibits a slow decreasing slope until 1mW where it starts to fall much faster. This indicates that without self-heating, the curve doesn't have a positive slope in the range of 0mW to 1mW in contrast to the curve in case 1. For case 3 with R_{TH} still equal to zero and avalanche effect turns off, curve 3 does not have the fast falling effect beyond 1mW as is the case in curve 2. This indicates the influence of avalanche effect is more dominant beyond 1mW. Furthermore, even with the self-heating and avalanche effects removed from curve 3, the curve still shows a small negative slope and this is a result of the series resistances, R_E and R_B but R_E has a larger influence than R_B because I_E which is the sum of I_C and I_B is about beta(DC current gain approximately 80-100) times larger than I_B which means the voltage drop across R_E would be much larger than the voltage drop across R_B . Furthermore, R_E is usually extracted under high injection region where self-heating is a factor which means the value has an error associated the self-heating effect. Thus, a method is needed to decoupled from the self-heating effect which will be discussed next. For case 4, the influence of series resistance is illustrated by setting R_E equal to zero. Curve 4 shown in the plot is relatively linear and constant as expected because with avalanche effect and self-heating turn off the base current should be constant for any changes in V_{CEX} 's.

To account for R_E , the base and collector currents have to be modified. By writing Kirchoff voltage equation along the base-emitter terminal of the transistor in Fig. 5.4, the equation is

$$V_{BEX} = V_{BEi} + I_B R_B + (I_C + I_B) R_E \quad (5.19)$$

With I_C usually much greater than I_B , then the term $(I_C + I_B) R_E$ is approximately equal to $I_C R_E$, and since the voltage drop across R_E is much larger than the voltage drop across R_B , thus, (5.19) is simplified to

$$\Delta V_{BE} = V_{BEX} - V_{BEi} = I_C R_E \quad (5.20)$$

ΔV_{BE} is a result of a voltage drop across R_E . Thus, to account for the drop, the new expressions for base and collector currents are as follow:

$$I_B' = I_B \cdot \exp\left(\frac{\Delta V_{BE}}{m_{BEi} V_T}\right) \quad (5.21)$$

and

$$I_C' = I_C \cdot \exp\left(\frac{\Delta V_{BE}}{V_T}\right) \quad (5.22)$$

However, I_C' can assume approximately equal to I_C because by increasing V_{BEi} by a ΔV_{BE} amount causes the charge Q_p to increase which tends to cancel out the effect of the $\exp(\Delta V_{BE}/V_T)$ term. The corresponding total power dissipation becomes

$$P_T = I_B' V_{BEX} + I_C V_{CEX} \quad (5.23)$$

To verify the validity of the compensating factor for R_E , the data from case 1 is used to evaluate for I_B' and the result is plotted as curve 2 shown in Fig. 5.15. To illustrate the effect that R_E has on the curve, the simulated result from case 1 is replotted and is shown as curve 1 in Fig. 5.15. The slope of the rising part of curve 2 is much steeper than curve 1 as a result of compensating for R_E . To verify whether the compensating factor for R_E in curve 2 is correct, another simulation is done which uses the same model parameter set but

neglecting avalanche effect and R_E , and the result is plotted as curve 3 shown in Fig. 5.15. The plot shows that curve 3 is on top of curve 2 up until where curve 2 started to fall off as a result of avalanche effect. By removing avalanche effect in curve 3, it is a straight line which shows the slope is relatively constant over the entire range. With curves 2 and 3 lying on top of each other up until avalanche effect, the slope for both curves should be the same. This suggests that the least square fit should be applied at the low ΔP_T region, preferably, within the dotted box as shown in Fig. 5.15.

FIGURE 5.15 Compensate for R_E in $(I_B - I_{B0})/I_{B0}$ vs. ΔP_T

5.4.1.5 Verification of R_{TH}

The verification of R_{TH} is done in two steps. The first step is to verify the accuracy of equations between Reisch's equation (5.4) and the extension of Reisch's equation (5.18). To avoid the influence of avalanche effect and R_E for now, the slope is obtained from curve 3 of Fig.15 which is $23.96W^{-1}$. Using the following data, $V_{BEX}=0.85V$, $m_{BEi}=1.0144$, $V_g=1.14V$, $T_o=300K$, and $\alpha_\beta = 6E-3K^{-1}$, (5.4) and (5.18) are evaluated and the resulting R_{TH} values are $557.7K/W$ and $640.5K/W$, respectively. With the actual R_{TH} value used in curve 3 is $548K/W$, the relative errors for (5.4) and (5.18) are 16.8% and 1.77%, respectively. The large error in Reisch's equation were attributed to some approximations made in the derivation that are now included in the extension of Reisch's equation, resulting in a small relative error. Furthermore, if the slope is obtained from curve1 in Fig. 5.15, then the relativity error for R_{TH} using (5.4) is about 20% higher than the actual value

in the model parameter set. The first step in the verification proves that the extension of Reisch's equation is much more accurate compare to the original Reisch's equation, and secondly, the compensation for R_E drop produces a more accurate R_{TH} value than when it is not considered.

The next step in the verification process is to determine the accuracy of the analytical expression for R_{TH} . This can be accomplished by using MEDICI because it can provide the junction temperature at each operating point which can be used to determine an R_{TH} value by using the R_{TH} expression in (5.24) which can be easily obtained by re-arranging (5.14).

$$R_{TH} = \frac{T_{jmax} - T_o}{P_T} \quad (5.24)$$

where T_{jmax} =maximum junction temperature is from the base-collector junction

A MEDICI simulation is done using the advanced Si NPN transistor structure. Fig. 5.16 shows the plot of T_{jmax} versus P_T/A_E at $V_{CEX}=1.5V$.

FIGURE 5.16 T_{Jmax} vs. Power Density

By performing a least square fit on the plot, R_{TH}/A_E is obtained from the slope which is $3765 \text{ K}\mu\text{m}^2/\text{W}$. For a $0.5 \times 20 \mu\text{m}^2$ device, the calculated R_{TH} value is 376.5 K/W . The

extracted value using (5.18) with an extracted isothermal R_E value(will discuss the reason later), the result is 301K/W. The relative error in this case is about 20%.

The large error is attributed by several factors. The first factor is due to the simple R_{TH} expression in (5.24) because instead of using the maximum junction temperature, an average temperature should been used based on a method discussed in Liou *et al.*[40]. This can bring the difference closer to the extracted value.

Another factor is due to the R_E assumption made in the derivation of the analytical expression for R_{TH} . The exclusion of base resistance by assuming the reduction of V_{BEX} is mostly due to the voltage drop across R_E and very little on R_B and, thus can be neglected. However, if R_B is large, then there could be an error. In this case, R_{Bx} could be as high as 20 Ω . This could make the extracted value closer to the value using the simple R_{TH} expression.

Finally, the assumption stated earlier, in Section 5.2.1 regarding the temperature boundary surrounding region of the transistor i.e., the emitter, base, and collector contact, assumes an adiabatic condition, except for the substrate where it is setup as a heat sink taking a room temperature value. However, realistically there is some temperature difference between temperature at the emitter contact and maximum temperature at the base-collector junction. The reason for choosing the emitter contact is because it is right above the base-collector junction where the maximum heat source is located. To get an idea how much impact the assumption has on the extraction of R_{TH} , a couple of simulation is done. The first simulation, which has already been done, assumes an adiabatic condition for all the three contacts and the substrate is used as a heat sink. The extracted R_{TH} value is 301K/W and the value from (5.24) is 376.5K/W. The other simulation involves applying a temperature boundary condition at the emitter contact by setting a temperature at the emitter, say about 25% lower than the maximum junction temperature because there have been no study on the actual temperature at the surface of the device, but common sense says that the temperature should be somewhere 10-25% below the maximum temperature. By setting a temperature boundary condition at the emitter contact, the base-collector

junction temperature is observed to be lower than the value from adiabatic assumption about the contacts. The extracted R_{TH} value is found to be 260K/W which is about 15% lower than without setting the emitter temperature boundary, and the corresponding R_{TH} value from the simple expression in (5.24) is about 315K/W which is about 20%. Thus, the assumption about not putting a temperature boundary at the emitter contact does not affect the extraction method.

In general, the analytical expression for R_{TH} gives a reasonable lumped thermal resistance value by considering the fact that it is obtained from a simple IV curve of a DC measurement. In addition, to get a reasonable accurate R_{TH} , the extension of Reisch's method along with the compensation for R_E drop should be used.

5.4.1.6 Extraction of an “Isothermal” R_E

To prevent self-heating effects from being accounted twice, the extraction of R_{TH} requires the previous extracted model parameters, in particular, R_E , α_β , and m_{BEi} in an isothermal condition that is, determination of model parameters under the condition that the transistor does not heat up. Usually, m_{BEi} is independent of temperature, whereas α_β is a relative temperature coefficient. However, for R_E , it contains some self-heating effects because usually, the extraction occurs at high injection regime, where the transistor has already heating up significantly. This can affect the accuracy of R_{TH} because the effect of heating is accounted twice.

To avoid this problem, an isothermal R_E must be obtained. A possible solution is discussed in[33]. The transistor is biased with short quick pulses usually, in tens of nanosecond pulse steps such that it is less than the thermal time constant because it is usually in the microseconds range(to be discussed in Section 5.4.2). This indicates that the effect of self-heating has been reduced significantly, making possible for the extraction of an isothermal R_E . However, this approach often requires complex and expensive equipments. An alternative solution is to solve simultaneously for R_{TH} and R_E without requiring com-

plex and expensive equipment. Before this approach can be examined in some details, an investigation into several methods to extract the model parameter R_E is briefly discussed.

5.4.1.6.1 Evaluate Existing R_E Extraction Methods

A study was carried by Fong *et al.*[41] which discussed two methods to determine R_E . The first method is a DC measurement which uses the base-collector space-charge region, also known as Early Effect, to bias the transistor to obtain an IV curve with the slope of the curve equals to R_E . This method only works well under an isothermal condition because when including self-heating effects, the method produces a negative resistance value. The second method is an AC frequency measurement that determines R_E from $|Z_{12}|$ at frequencies less than 300MHz or before $\omega_{\beta 0}$ also known as the corner frequency at -3dB in a common emitter configuration. The accuracy of the extracted value is within +/- 20% which is quite large.

A better strategy is to measure base current in a regime where self-heating occurs. correct the measurement to obtain an isothermal result, then extract R_E from this data. Defining I_{B0} to be (3) with V_{BEi} replaced by $V_{BEx}(=V_{BEi}+I_B R_B+I_E R_E)$. and taking the ratio of this quantity to I_B directly from (5.8), and the result is

$$\frac{I_{B0}}{I_B} = \exp\left(\frac{I_B r_B + I_E r_E}{m_{BEi} V_T}\right) \quad (5.25)$$

where $I_E \approx I_C$. The result of taking the logarithm of both sides of (10), is the relationship.

$$m_{BE} V_T \ln(I_{B0}/I_B) - I_B R_B = I_E R_E. \quad (5.26)$$

This equation can be viewed as a linear relationship between the left hand side, which is calculated using measured quantities, and I_E , with slope R_E . I_B and I_E are isothermal currents. The above equation assumes that R_B is approximately equal to R_{Bx} .

To verify the method, again a circuit simulator containing HICUM is used. A simulation is done using an advanced Si transistor of size $0.5 \times 20 \mu\text{m}^2$ model parameter set with

R_E equals to 3.6Ω Fig. 5.17 shows the plot of LHS (5.26), obtained from calculation using measured data, versus I_E for the cases with and without self-heating.

FIGURE 5.17 LHS (5.26) vs. I_E for an advanced Si NPN BJT of $0.5 \times 20 \mu\text{m}^2$

The results show the extracted value for the cases with and without self-heating are 3.015Ω and 3.47Ω respectively, and thus, the corresponding relative errors are 16.3% and 4%, respectively. For the case without considering self-heating, the method gives good agreement between actual and extracted R_E . The 4% error is due to assumption about R_B being approximately equal R_{Bx} but in this case the error is small enough such that the approximation is adequate. Furthermore, the small error indicates that the method produces much better accuracy compare to other methods discussed earlier. For the case with self-heating, the error in R_E is quite large and this method worsens if there is more self-heating. To show this, a similar simulation is also done for a SiGe transistor with emitter area of $0.5 \times 10 \mu\text{m}^2$. Fig. 5.18 shows a plot for the cases with and without self-heating. The extracted R_E values for cases with and without self-heating are 4.79Ω and 6.29Ω respectively. The actual value used in the simulation is 6.5Ω . The relative errors for cases with and without self-heating are 26.3% and 3.2%, respectively. Again, the method produces good accuracy for the case without self-heating, whereas for the case with self-heating, the error becomes worse. The reasons for the larger self-heating in SiGe compare to the Si BJT are probably due to the following: higher current densities in SiGe that results in more power dissipation, and thus more self-heating; and for SiGe, deep trench-isolations are used to isolate between transistors but it can also at the same time prevent heat flow from spreading outward, causing the temperature to be higher.

FIGURE 5.18 LHS(5.26) vs. I_E for SiGe $0.5 \times 10 \mu\text{m}^2$

The above analysis shows without self-heating the method produces good accuracy but becomes inaccurate when include self-heating. To resolve this problem, the self-heating effects that are inherent in the base current I_B and collector current I_C must be removed since both currents are used in the extraction of R_E . Analytical expressions for isothermal base and collector currents can be found without using iterations are as follow. The isothermal base current, $I_B(T_0)$ can be obtained by re-arranging (5.13), and the isothermal $I_C(T_0)$ can be obtained by

$$I_C(T_0) = \beta(T_0) \cdot I_B(T_0) \quad (5.27)$$

with

$$\beta(T_0) = \beta(T) / (1 + \alpha_\beta \cdot \Delta T) \quad (5.28)$$

The temperature T can be calculated by using (5.14). Equation 5.29 is replaced with isothermal currents and the new equation becomes

$$m_{BEI} V_T(T_0) \ln\left(\frac{I_{Bo}}{I_B(T_0)}\right) - I_B(T_0) R_B = I_E(T_0) R_E \quad (5.29)$$

The dependence of R_{TH} , when calculating for the temperature which in turn is used to calculate for the isothermal currents, implies that the extraction of R_E and R_{TH} must be done simultaneously because the extraction of R_{TH} also requires the value of R_E .

5.4.1.6.2 Simultaneous Extraction of R_E & R_{TH}

A flow chart of the simultaneous extraction procedure for R_E and R_{TH} is given in Fig. 5.19.

FIGURE 5.19 Flow Chart of R_{TH} - R_E Extraction Algorithm

For R_E extraction, a DC measurement is performed to obtain I_C and I_B as a function of V_{BEx} for a fixed V_{CEx} in the collector current range $I_C = 0.01$ to $1 \text{ mA}/\mu\text{m}^2$; and for R_{TH} extraction, a DC measurement is done to obtain I_C and I_B with V_{CEx} varying in the range (V_{BEx}, BV_{CE0}) for a fixed V_{BEx} such that an average I_C of $0.1 \text{ mA}/\mu\text{m}^2$ is achieved. The determination of R_{TH} and R_E uses the methods described in Sections 5.4.1.2 and 5.4.1.6.1, respectively. R_E is determined first followed by R_{TH} which uses the previous extracted R_E value. Once R_{TH} is known, it is used to determine the junction temperature, which is used to calculate the isothermal currents. These currents are then used by the R_E extraction routine. The process starts over again by extracting for R_E using the latest corrected isother-

mal currents. The iteration is repeated until the tolerance is satisfied i.e., if the relative error between successive values of R_E and R_{TH} is less than $1E-4$, the process is terminated. Note that for the first iteration, R_E is determined by assuming the currents are isothermal which would provide a good initial guess for the R_{TH} extraction. The convergence rate for the method is about 5 to 10 iterations. A program is written in Matlab (refer Appendix D for the program listing).

5.4.1.6.3 Verification of R_{TH} - R_E Extraction Method

i. Simulated Data:

The R_E extraction method was verified using HICUM with a parameter set determined from measurements to generate the necessary input data. The simulations were done with R_{TH} turned on for two technologies, a silicon double-poly bipolar process and an SiGe HBT process. Table 5.2 contains the extracted results for the two cases with and without using isothermal current equations.

TABLE 5.2. R_E Comparison Using HICUM Data

	I_B and I_C	Actual (Ω)	Extracted (Ω)	Error (%)
Si	Isothermal	3.6	3.47	3.6
	Non-Isothermal	3.6	3.02	16.1
SiGe	Isothermal	6.5	6.29	3.2
	Non-Isothermal	6.5	4.79	26.3

For the case with the isothermal current equation, the same R_{TH} value used in the simulation is also used in the calculation. The method works well when using the isothermal current equations, whereas large errors would occur if they are not used. This large error is further verified with a large-signal switching simulation shown in Fig.5.20 for a simple inverter with a SiGe transistor using results from Table 5.2 and a C_{TH} value of $4E-9$ Ws/K

FIGURE 5.20 Large-Signal Switching Simulation of a simple inverter using data from Table 5.2 for a SiGe HBT.

To demonstrate the improved accuracy of (5.18) over (5.4), HICUM was used to generate data for a thermal resistance extraction with a parameter set determined from measurements which included a thermal resistance of 548K/W. Applying the method described in Section 5.4.1.6.2 with (5.18), the extracted R_{TH} was 557K/W(1.8% error), while the same method using (5.4) gave a value of 640K/W(16.8% error). This demonstrates the added accuracy which arises from using (5.18).

The simultaneous extraction technique was verified by generating current characteristics using MEDICI, which includes self-heating through solution of the lattice heating equation. Realistic doping profiles for an advanced double-poly Si NPN bipolar process were used in the simulation. The values of R_E and R_{TH} were extracted directly from the structure and internal temperature change, respectively, so they represent the “actual” values. The value of thermal resistance calculated without the simultaneous method has an error of 8.3%, which is reduced to 1.9% when the simultaneous method is used.

TABLE 5.3. R_{TH} Extraction Using MEDICI Data

Case	R_E (Ω)	R_{TH} (K/W)	Error(%)	
			R_E	R_{TH}
Actual	4.65	375	-	-
Eqn.(5.4)	-	344	-	8.3
Simultaneous	4.7	368	1.1	1.9

Fig. 5.21 and Fig. 5.22 shown below are the Gummel plots and Beta vs. I_C plot with each plot showing three curves for the following conditions; simulation with self-heating, simulation without self-heating, and calculated to compensate for self-heating.

FIGURE 5.21 Gummel Plot

FIGURE 5.22 Beta (DC current gain) vs. I_C/A_E

The above figures show that the curves from calculation and the curves from simulation without self-heating agree relatively well.

ii. Measured Data

The method outlined in Section 5.4.1.6.2 is used to determine R_E and R_{TH} from the actual measurement data and in this case the transistor is a SiGe transistor with emitter area of $0.5 \times 10 \mu\text{m}^2$. Fig. 5.23 shows the dI_B vs. dP_T curve obtained from measurement with solid line representing the least square fit of the positive slope of the measured data.

FIGURE 5.23 $(I_B - I_{B0})/I_B$ versus $(P_T - P_{T0})$ for $V_{BEx}=0.85V$ and $V_{CEx}=[0.85, 4.5]$

Table 5.4 contains R_E and R_{TH} values extracted from measurements for the Si double-poly self-aligned BJT with an emitter area of $0.5 \times 20 \mu m^2$ and SiGe HBT with an emitter area of $0.5 \times 10 \mu m^2$ (in this Table, the “independent” method uses (5.4) for R_{TH} and has no simultaneous extraction procedure).

TABLE 5.4. Comparison Using Measurement Data

Technology	Independent		Simultaneous	
	R_E (Ω)	R_{TH} (K/W)	R_E (Ω)	R_{TH} (K/W)
Si double-poly	3.56	375	5.0	470
SiGe HBT	5.5	1405	7.6	1335

Fig. 5.24 compares the HICUM Gummel plot prediction for a SiGe HBT using the R_{TH} and R_E values from Table 5.4 for the independent and simultaneous methods.

FIGURE 5.24 (a) Gummel Plots for $V_{CEx} = (0.8, 2.4)V$ and a compact model HICUM using R_{TH} and R_E obtained from (—) independent and (—) simultaneous methods. (b) I_C vs. V_{BEx} for $V_{CEx} = (0.8, 2.4)V$ with simultaneous method. A_E is the emitter area.

The improved prediction with the simultaneous method is evident at high current densities, where self-heating and the voltage drop associated with the emitter series resistance are most important. Fig. 5.25 shows the plot of I_B versus V_{CB} for the simulated data with the extracted R_{TH} value obtained from the simultaneous method and measured data.

FIGURE 5.25 I_B vs. V_{CB}

Using the results from the simultaneous method, the small signal simulations were performed. The transit time and figures of merit, f_T and f_{max} are shown in Figs. 5.26 with and without self-heating.

FIGURE 5.26 (a) Transit Time vs. I_C for A_{E0} of $0.5 \times 10 \mu m$ for HICUM with and without self-heating at $V_{CE} = (0.8, 1.6, 2.4)V$.

With self-heating, the transit time from the model is better matched to measured data compared to the case without self-heating and the reason is because the collector current is now modeled more accurately with the self-heating model. This is reflected in the f_T and f_{max} plots shown in Figs. 5.26c and 5.26e. The f_T and f_{max} plots from Figs. 3.17a and 3.17b are replotted which are shown in Figs. 5.26b and 5.26d, respectively for comparison purposes.

FIGURE 5.26 (b) f_T versus I_C for HICUM without self-heating and (c) f_T versus I_C for HICUM with self-heating

FIGURE 5.26 (d) f_{max} versus I_C for HICUM without self-heating and (e) f_{max} versus I_C for HICUM with self-heating.

Furthermore, the f_{max} predicted curve at V_{CE} of 2.4V shown in Fig. 5.26d matched exactly the measured curve for the same V_{CE} , whereas the predicted curve for the same V_{CE} in Fig. 5.26e does not agree exactly the measured curve because the extraction of R_{TH} and R_E were not performed on the same transistor and as a result, some inaccuracies are expected due to process variations.

5.4.2 Extraction of Thermal Capacitance

5.4.2.1 Theory

The general thermal equation in (5.2) is solved with respect to time, and the solution yields,

$$T_j(t) = T_c + R_{TH}P_T(1 - \exp(-t/\tau_{TH})) \quad (5.30)$$

where τ_{TH} = thermal time constant = $R_{TH}C_{TH}$

τ_{TH} represents the time in which the temperature of device is 63.2% of T_j . The determination of thermal capacitance, C_{TH} , is found by taking the thermal time constant, τ_{TH} , divided by the previous extracted thermal resistance, R_{TH} . However, it is difficult to measure the junction temperature of a transistor directly without using special equipment, but rather it can be determined indirectly by monitoring the change in the electrical characteristics under transient condition.

A direct relation between the electrical quantity such as the terminal base-emitter voltage (V_{BEX}) and base current to junction temperature can be proven empirically by using a compact model with a self-heating model. Although the self-heating model in HICUM may not accurately predict all the physical effects in reality, it can still be used to give a first order approximation. In addition, it allows easy access to internal quantities such as the junction temperature. In order to measure the thermal time response, this involves applying the transistor with a short power pulse which heats up the transistor. To remove the influence of the transient response of the transistor, the rise time or the fall time of the pulse should be much longer than the transit time of the transistor while the pulse width should be long enough such that the thermal effect can be detected. The electrical response time can be measured by monitoring either voltage or current at the base-emitter junction. Three cases to generating a power pulse and measuring the electrical response are discussed. For each case, a collector current pulse is generated by applying a voltage pulse at the collector terminal of the transistor with the proper DC operating. The product of the current and voltage pulse is a power pulse.

i. Case 1- This method involves operating the transistor in a common emitter configuration shown in Fig. 5.27. To turn the transistor on, a fixed voltage is applied at the base-emitter terminal and a voltage pulse is applied at the collector. The time response is obtained by monitoring the base current.

FIGURE 5.27 Bias Schematic for Case 1

ii. Case 2- In this method, the biasing schematic is shown in Fig. 5.28 where the same transistor configuration as in Case 1 is used. Instead of applying a fixed voltage at the base-emitter terminal, a constant current source is employed at the base of the transistor. The thermal time response is obtained by monitoring the base emitter voltage. This method is very similar to the one used in the extraction of R_{TH} except a steady state increasing voltage is used rather than a voltage pulse.

FIGURE 5.28 Bias Schematic for Case 2.

iii. Case 3- The bias schema in Fig. 5.29 shows the transistor operating in a common-base configuration. A constant emitter current source is applied at the emitter terminal to provide the dc bias. The voltage at the base-emitter terminal is monitored for the time response. This method has been widely used [42 and 43].

FIGURE 5.29 Bias Schematic for Case 3.

For the simulation, a realistic model parameter of a SiGe HBT with an R_{TH} of 1330 K/W and an approximated C_{TH} value of 3.2E-9 Ws/K is used. The results of the simulations for each case are normalized by using

$$T_{norm} = \frac{T_j - T_i}{T_f - T_i} \quad (5.31)$$

where T_f and T_i are the steady state value and the initial value, respectively and T_j is the variable. In each case, the junction temperature as well as the corresponding electrical quantity are normalized and the results are plotted as a function of time shown in Fig. 5.30a for cases 1 and 2 and Fig. 5.30b for case 3. The solid curve known as “Actual” in each figure is the normalized junction temperature that is calculated using (5.30) with known C_{TH} and R_{TH} values that were used in the simulation for each case. The expected τ_{TH} is 4.3 μ s.

FIGURE 5.30 Normalized quantity vs. Time

The results from the above figures show the normalized base-emitter voltage and the junction temperature curves for case 1 and 2 agree well with the “Actual” curve, whereas a large deviation is observed between the normalized base current and junction temperature for case 2 and the “Actual” curve. The thermal constant for each case is obtained by taking 63.2% of the steady state value. The results for each case are summarized in Table 5.5.

TABLE 5.5. Simulation Results for τ_{TH}

Method	τ_{TH}	Relative Error (%)
Case 1	5.4	25.6%
Case 2	4.1	4.7%
Case 3	4.25	1.2%

The relative error for case 2 and 3 is quite small, whereas case 1 has about 26% error. The reason for the large error seen in case 1 is due to a non-constant power pulse caused by the Early Effect. The reason for the larger Early effect seen in case 1 is due to the constant base-emitter voltage that causes the collector current to increase, in contrast the other two cases employ a constant current source which makes the dc operating more stable, creating a constant collector current, and hence constant power pulse. In practice the method in case 1 is not used because it is difficult to measure currents directly using an oscilloscope, whereas the methods in case 2 and 3 are much easier because both methods involve measuring the terminal voltage. However, the characterization of model parameters often requires the transistor operating in a common-emitter configuration because it is easier to extract and validate the model in this configuration, in contrast to a common-base configuration, it is rarely used. Thus, to achieve consistent results from a modeling point of view, the method in case 2 is preferred over the method in case 3 even though case 3 is a bit more accurate in extracting the thermal time constant.

Several conclusions can be drawn. First, case 2 and 3 confirm the assumption about being able to monitor the thermal time response by looking at the electrical response. The method in case 2 is preferred because the transistor configuration (common-emitter) is consistent with the transistor configuration used in the full parameter extraction.

5.4.2.2 Thermal Transient Measurement Setup

The measurement setup and bias conditions for the transient measurement are as follows:

i. Measurement setup- The bias schema is to similar the circuit in case 2, except for additional power splitters so that measurement of terminal voltage is possible using an oscilloscope. The new setup is shown in Fig. 5.31.

FIGURE 5.31 Thermal Transient Measurement Setup

ii. Bias Conditions:

1. pulse of V_{CEX} - lower voltage limit of 1V to upper limit of BV_{CE0}
 - pulse width = 40 μ s or up to where V_{BEx} reaches its steady state value
 - rise time ≤ 20 ns or should be larger than τ_f but faster than the thermal time constant
2. Operating Point - $I_B/A_E = 4 \mu A/\mu m^2$

The change in V_{BEx} and the input pulse are monitored by Port1 and Port2 of the oscilloscope, respectively. To minimize noise during measurement, the averaging option from the oscilloscope should be used because usually the change in V_{BEx} is small, less than 3mV for a large high-speed Si bjt with emitter window area of $0.5 \times 10 \mu m^2$, and the change is much smaller for small transistors but the change in SiGe-based HBTs is much larger, usually around 7 to 10mV. Furthermore, with noise level usually between 1 to 2mV, it makes the signal more difficult to resolve, especially with the oscilloscope having a 1mV/division minimum resolution. The extraction of C_{TH} require the measurement of a thermal time constant, τ_{TH} , which is obtained from the signal at port 1 of the oscilloscope.

5.4.2.3 Extraction and Verification of C_{TH}

Figure 5.32 shows the transient response of a high speed Si BJT size $0.5 \times 20 \mu\text{m}^2$. The bottom curve in the figure is the input pulse and the top curve represents the base-emitter voltage V_{BE} . The V_{BE} curve initially drops quickly in the first $5\mu\text{s}$, and then it slowly decreases until it levels off about $15\mu\text{s}$ later. The two slopes from the curve indicate two thermal time constants instead of one. For the first order approximation, the thermal time constant is found by taking 63% of V_{BE} from where the curve drops rapidly which is approximately $1\mu\text{s}$. With the previous extracted R_{TH} value of 520k/W , then the calculated C_{TH} value is $2.69\text{E-}9 \text{ Ws/K}$.

FIGURE 5.32 V_{BE} vs. time(s) for $I_B=35\mu\text{A}$, $V_{CE}=[1,3]\text{V}$

To verify the extracted value, a simulation is performed using the model parameter set for the transistor along with the thermal parameters. The result of the simulation along with the measured data is plotted in Fig. 5.33 which agrees reasonable well despite the fact that the thermal network model is a first order approximation.

FIGURE 5.33 Thermal Transient response for a high speed Si BJT

Note that the simulated curve has a small peak when the input is in transition from low to high, whereas the measurement didn't show the peak. The peak is due to the Early effect which means the output characteristic between simulation and measurement may not agree well. Despite the peak appearing the measurement, the determination of τ_{TH} should not be influenced by it.

Figure 5.34 shows the plot of V_{BEx} versus time with the bottom curve represent the input pulse and the top curve represent the measured V_{BEx} for a $0.5 \times 10 \mu m^2$ SiGe NPN HBT.

FIGURE 5.34 V_{BEx} vs. time(s) for $I_B=24\mu A$, $V_{CE}=[1,2]V$

Fig. 5.35 shows the same curves except they are zoomed in. Again, the curve drops rapidly for the first $5\mu s$, and then it slowly decreases until it reaches steady state $20\mu s$ later. The thermal time constant for this device obtained by taking 63% is about $4\mu s$ where the curve is the steepest. With the extracted R_{TH} value of $1244K/W$, the calculated C_{TH} value is equal to $3.2E-9 Ws/K$.

FIGURE 5.35 zoom in part of Fig. 5.33

To verify the extracted value, a simulation is performed using the model parameter set for the transistor along with the thermal parameters. The result of the simulation along with the measured data is plotted in Fig. 5.36 which shows reasonable agreement.

FIGURE 5.36 Thermal Transient Response for a SiGe HBT

Additional measurements have been performed for various transistor sizes for the two technologies. It has been determined that the thermal time constant is the same for several different transistor sizes which is about $1.4\mu\text{s}$ for the high speed Si BJT technology, whereas a value of $4.5\mu\text{s}$ has been found for the SiGe HBT technology. The reason for a higher time thermal constant from the SiGe HBT technology is probably due to a lower thermal conductivity through the trench isolation compared to a normal local field isolation which has a higher thermal conductivity.

5.5 Scalable Thermal Parameters

It has been determined that the thermal resistance scales according to the inverse square root of the effective emitter area, whereas the thermal capacitance is scaled according to the square root of the effective emitter area[39]. The relations are given below:

$$R_{TH} = \frac{R_{STH}}{\sqrt{A_E}} \quad (5.32)$$

and

$$C_{TH} = C_{STH}\sqrt{A_E} \quad (5.33)$$

where R_{STH} =specific thermal resistance

C_{STH} =specific thermal capacitance

A_E =effective emitter area

To illustrate the relationship in (5.32), R_{TH} is determined from various Si transistor sizes and the results are shown in Fig. 5.37 as a function of effective emitter area (A_E). In the figure the symbol represents the extracted R_{TH} values and the solid line representing the fit line. The resulting specific thermal resistance value is found to be $1900\mu\text{mK/W}$.

FIGURE 5.37 Scaling of R_{TH} for advanced Si BJTs

The thermal time constant can be calculated by taking the product of (5.32) and (5.33) and the result shows that the time constant is independent of emitter area which agrees with the measured results obtained earlier where τ_{TH} is found to be constant for various emitter area. Due to a limited number of available transistor sizes, the specific thermal resistance cannot be determined for the SiGe HBT technology.

5.6 Validation of Self-heating Model in Circuits

This section discusses the validation of a compact model. HICUM, with the self-heating subcircuit to illustrate the importance of an accurate model in circuit design. A simple current source circuit is used to show the influence of self-heating on the circuit performance. This is achieved by performing several simulations using a realistic model parameter set extracted from a SiGe transistor.

Current source are widely used in analog integrated circuits both as biasing elements, as load devices for amplifier stages and because these active devices occupy less die area compared to resistor which results in less cost. Often the performance of circuits require the bias current to be insensitivity to power supply variation and in particular temperature variation. A simple current source shown in Fig. 5.38 is used in the demonstration. The

circuit consists of two transistors ($0.5 \times 10 \mu\text{m}^2$), a resistor to establish a reference current I_{REF} and a sweeping voltage supply V_o at the output. The output voltage supply is swept from 1V to 3V with the output current I_o monitored. The reason for sweeping V_o is to increase the power dissipation in the transistors to illustrate the effect of self-heating.

FIGURE 5.38 A Simple Current Source

The operation of the circuit is as follows. Transistor Q1 is operating as a diode due to the base-collector connection that forces the voltage to be zero. With base of Q1 and Q2 having the same base-emitter voltage, their collector currents are equal, that is I_o equal to I_{REF} where I_{REF} is determined from the ratio of $\Delta V (= V_{\text{REF}} - 0.7_{\text{diode drop of Q1}})$ divided by R_{REF} . The results from the simulation for two cases with and without self-heating are shown in Fig. 5.39.

FIGURE 5.39 Comparison of I_{REF} , I_o with and without self-heating

Ideally, the output current I_o should agree with the reference current (I_{REF}) for the condition without self-heating but the simulation shows I_o deviates from I_{REF} as much as 5%. The reason is due to the finite output resistance associated with the transistor that causes the current to increase with increasing collector voltage. The output current with self-heating shows large deviations, as much as 20% in error, from the reference current for increasing V_o . Even though there is about 5% difference between I_o and I_{REF} for the

case without self-heating, the net error in I_o between the cases with and without self-heating is still large and the difference can affect the performance of the main circuit because the operation of the circuit depends on the proper DC bias provided by the current source. This effect should be considered during the design stage to ensure that the final circuit functions according to the expected design. A possible solution is to use a Wilson current source[44] shown in Fig. 5.40.

FIGURE 5.40 Wilson Current Source

A couple of simulations are performed for the cases with and without self-heating and the relative error in I_o with respect to the reference current for each case along with results from the simple current source is shown in Fig. 5.41.

FIGURE 5.41 Relative Error in I_o with respect to I_{REF} (Note-CS represents current source)

The relative errors of I_o from the Wilson current source for both cases with and without self-heating are 6% and 2%, respectively, and they remain fairly constant over the operating range. In contrast to the simple current source, the relative errors in I_o are larger for V_o greater than 1.5V and become small for V_o less than 1.5V. This illustrates an improvement

in using the Wilson current source to reduce the influence of self-heating on the output current. Most circuit simulators do not have the self-heating model, however the amount of self-heating effect can still be approximated by calculating the junction temperature using known thermal resistance value and an approximate power dissipation and entering this new temperature into the simulator. This process can be repeated until a solution is found.

Chapter 6 - Conclusion

6.1 Summary and Discussion

This work has presented the results from an investigation of SiGe NPN HBTs with respect to compact modeling.

In Chapter 2, some analyses into the operating behaviour and properties of SiGe HBT and Si BJT technologies were performed. The Generalized Integral Charge-Control Relation (GICCR) which represents a closed-form solution of the collector current for all operating regime of transistor operation is derived which is valid for SiGe HBTs. The modified GICCR and the classical theory were used to illustrate the superior performance of SiGe HBT over that of Si BJT in terms of speed, currents, and DC current gain. The modified GICCR also confirmed the smaller Early Effect phenomena found in the graded Ge profile for SiGe HBT compare to that of Si BJT. The results from several 1D numerical simulations with a realistic net doping profile with the graded Ge mole fraction profile indicated the higher cutoff frequency, DC current gain, and higher output resistance or reduced Early Effect observed in the SiGe HBT over that of Si BJT for the same net doping profile.

Chapter 3 provided an overview of the differences and similarities of the modeling principles in the commercial SGP model and HICUM. Methods to characterize the model parameters for both compact models were subsequently discussed. A common technique often employed involved fitting parameters to certain measured transistor's characteristics, followed by a global optimization. The disadvantage of this method of fitting and optimizing parameter values to a certain characteristic was that there is no guarantee that the values obtained are in anyway unique. Consequently, the results from the procedure produced less than optimal agreement when comparing simulation to measurement. Furthermore, optimization had the tendency to make the parameter values to become non-physical to fit a certain curve, and consequently, correlation between model parameter values to specific technology information such as sheet resistances and capacitances were difficult to achieve. A preferred procedure was to obtain the model parameters by performing the measurement at certain operating regime which made possible the dropping of certain

terms in the model equation to make it linear. The model parameters were obtained from the slope and intercept obtained from the least square fit of the linearized model equation. The advantage of this technique was that monitoring of process variations, geometric scaling, and statistical modeling due to the mutually consistent method used.

Often accurate modeling results depended heavily on proper measurement procedures. The measurement setup and as well as the proper technique for de-embedding of parasitic in S-parameter measurements were explained. Difficulties arose with some measurements due to the instability of the transistor. It was found that the stability of the transistor was very sensitive to the DC biasing method used. A solution was to use a Bias-Tee to remove oscillation. In addition, a practical approach to determine two important figures of merit f_T and f_{max} that can be applied consistently to all technology was presented.

Finally, several comparisons using various DC characteristics and figures of merit between simulation and measurement for the SiGe technology were performed for the two compact models, SGPM and HICUM. The results illustrated some discrepancies in both models. Some of the observed physical effects from the measured transistor's electrical characteristics found evidences of a reduced Early Effect caused by the graded Ge mole fraction profile, high f_T and f_{max} , and high collector currents which resulted in a large β value and as well as a pronounced self-heating effect. It appeared that HICUM had difficulty in capturing the DC characteristic such as the output characteristic curves (I_C vs. V_{CEX}) where the predicted collector currents were over estimating measured data for increasing V_{CEX} 's due to the reduced Early Effect at low to moderate operating regime and the currents were under-estimating measured data due to self-heating at high operating regime. Consequently, for the SGP model, the reduced Early Effect was easily predicted from the fitting model parameter V_{AF} but the self-heating effect was not modeled. The main effort of the work was to analyze in detail the physical reasons for the reduced Early Effects and as well as the self-heating in the transistor and solutions were proposed to account for these phenomenons and implemented them into HICUM.

In Chapter 4, the modified GICCR model discussed in Chapter 2 was implemented into HICUM to account for the reduced Early Effect observed for this technology. The differences between modified GICCR and original GICCR were the weighting factors that described the effect of the germanium on the mobility and the electron carrier density and their subsequent impact on the collector currents. Validation of the modified GICCR was performed by calculating the equation directly with the weighting parameters determined numerically and compare the results to that of the 1D simulated data. Reasonable accuracies were achieved for the graded and box Ge profiles of SiGe HBTs. The drawback of modified GICCR was that it contained parameters that cannot be measured easily in contrast to that of original GICCR. Subsequent simplification of the modified GICCR was needed to reduce the complexity of the parameters to only a few measurable parameters. This involved lumping the weighting factors h_{fB} , h_{fC} , and h_{fE} for the minority hole charge into a single parameter h_C , then divide the modified GICCR equation by h_{je} . This approach was made possible the determination of weighting factors from measurement by employing the direct extraction technique. This involved linearizing the model equation by performing the measurement at certain operating regime which made possible the dropping of certain terms in the expression to make it linear. The model parameters were obtained from the slope and intercept of the least square fit applied to the linearized model equation. The modified GICCR was implemented into a compact model, HICUM. An excellent agreement was achieved when comparing the output characteristic curves from the compact model with the modified GICCR to the curves from 1D numerical simulated data and as well as to the measured data.

Chapter 5 began with a detail explanation of the influence that self-heating had on the electrical characteristics of the transistor. This was achieved by performing several numerical simulations using MEDICI on a realistic doping profile to gain some insights into the physical mechanisms by observing the temperature contour and the heating flow. From these results, it was concluded that most of the heat were generated within the intrinsic base-collector junction due to the large power being dissipated in that region. Consequently, the impact of self-heating had a significant influence on the collector and base currents as observed from several simulations, whereas it had a negligible effect on cutoff

frequency. It was further re-enforced from the I-V measurements of SiGe NPN HBTs. A simple thermal subcircuit, consisting of a thermal resistance (R_{TH}) and a thermal capacitance (C_{TH}) currently found in HICUM was explained.

The task was to determine these thermal parameters without requiring complex measurement setups or equipment, but rather these parameters can be easily obtained from the electrical characteristics. The thermal parameter R_{TH} was determined first under a steady condition to decouple the influence of C_{TH} , followed by the extraction of C_{TH} under transient condition. A modified analytical R_{TH} expression based on Reisch's method was derived. The extraction of R_{TH} involved evaluating for the modified analytical expression with the proper I-V measurement which was performed in the medium to high current operating range where self-heating is significant. Since in this operating regime the emitter series resistance due to the emitter contact and bulk of the emitter material can affect the extracted R_{TH} value. Thus, a method for simultaneous extraction of the thermal resistance (R_{TH}) and emitter series resistance (R_E) was presented which allowed the accurate determination of these values from data measured in an operating region where both series voltage drop and thermal effect are important. The technique combined the R_E extraction method based on DC measurement and the R_{TH} extraction method. The simultaneous technique was verified against data generated from HICUM as well as 2D numerical simulation with lattice heating. Furthermore, using measured data from SiGe HBT, a comparison between the independent extraction method(Reisch's method) and the simultaneous method led to a much more accurate prediction of actual characteristics.

The thermal parameter C_{TH} was calculated from the ratio of the thermal time constant τ_{TH} over the previous extracted R_{TH} value. This required measuring the transistor for the junction temperature change in response to a power pulse in which the thermal storage time or the thermal time constant represented the time that it would take the junction temperature to reach 63.2% of the steady state value. The difficulty in this approach was that direct measurement of junction temperature required complex equipment. An approach was to use the electrical transient response of either the terminal base-emitter voltage

(V_{BEx}) or the base current (I_B) to a power pulse applying at the collector of the transistor to represent the transient response of junction temperature. To illustrate the correlation, three different methods for measuring the electrical response to a power pulse and compare to the thermal time response were analyzed using HICUM with a realistic model parameter set. The reason for using a compact model was that it can easily monitored the incremental change in junction temperature. It was concluded that the monitoring of base-emitter voltage with the transistor operating in a common-emitter configuration provided a reasonable agreement within 5% of relative error of the normalized base-emitter voltage to that of the normalized thermal time response. Moreover, this configuration was preferred since it conformed to the standard configuration used in the modeling and characterization of model parameters.

A brief discussion on the scalable thermal model parameters was presented. It had been shown from the publication that the thermal resistance has an inverse relationship to the square root of emitter area, whereas the thermal capacitance is proportional to the square root of emitter area. They were verified for both Si BJT and SiGe HBT technologies and the results confirmed the relationship.

Finally, a simple current source circuit was used to demonstrate the consequence of self-heating on circuit performance. Several simulations using HICUM with the thermal model for a realistic SiGe HBT model parameter set were performed and the results illustrated that with the thermal model, the output current deviated significantly from the reference current, whereas the case without the thermal model, the output current matched the reference current. It was determined that a Wilson current source can be used to reduce the influence of self-heating on the output current. The results from simulation with the thermal model indicated a small relative error between the output current and the reference current.

6.2 Future Work

The recommendations regarding future are presented. One particular area that requires further investigation is the verification of the new compact model HICUM containing the modified GICCR to the measured data of box Ge SiGe HBT. At the moment, the new compact model shows good first order prediction when comparing results from the compact model to the 1D numerical simulation. However, it is also equally important that the new compact model should be verified with the actual measured data because often the numerical simulation does not capture some the physical effects that occur in the actual device operation, and as a result some modifications to the new compact model may be required to reflect some of the observed effects.

Since most of the discussions relating to the thermal model in this work pertain to the extraction procedure to determine the thermal model parameters, the influence of self-heating on multiple emitter stripes transistor and on avalanche breakdown voltage needs to be investigated. For the multiple emitter stripes transistors, they are important, especially, in power amplifier circuits where self-heating can have an adverse affect on circuit performance. As a result, accurate modeling of the effect for multiple emitter stripes transistors is an important consideration. Since no actual measurements for the multiple emitter stripe transistors were performed, it is uncertain whether the extraction techniques to determine the thermal parameters produce the same accuracy as the case for a single emitter stripe. It is also unknown whether the thermal resistance scales linearly to the number of emitter stripes which is important for scalable model. This can easily be verified by extracting the thermal resistances from several transistors of various number of emitter stripes and plot them as a function of emitter stripes. With self-heating effect becomes dominant at medium to high collector current densities, the influence of self-heating on avalanche breakdown, in particular the breakdown voltage of collector-emitter with base open (BV_{CE0}) should be investigated since it can affect the performance of a circuit if this effect is not modeled properly. The analysis can be performed by using one and two dimensional numerical simulation with lattice heating and avalanche.

The scaling of thermal model parameters according to device dimension is an important issue since it is often desirable to be able to predict the impact of self-heating for any transistor size. In the current work, the thermal resistance seems to have scaled according to the one-over square of the emitter window area, whereas the thermal capacitance scales according to square root of the emitter window area. A good first order prediction is found based on the square root of emitter area relationship, but it was verified only on a limited number of transistor sizes. More verification is required through measurements of a wider selection of transistor sizes.

Finally, an important area of compact modeling is its practical application in circuit design. In order to gain any confidence in a compact model, verification and validation of a compact model often involves comparing simulation to measurement of several benchmark circuits such as CML chain, power amplifiers, LNA, VCO, and mixer. Furthermore, the modeling of self-heating effect on circuit performance, in particular power amplifiers for an array of transistors arranged within the minimum design rule of each other should be investigated. The thermal cross-coupling between transistors at the moment is not modeled since the thermal model in a compact model only focuses on a single transistor and not an array of transistors.

Appendix A - Derivation of GICCR

To apply the ICCR equation to a compact model, the denominator of (2.9a) has to be simplified by using a regional approach, that is separating the integration into the (E)mitter, (B)ase, and (C)ollector regions of the transistor and the equation with the simplified weighting function in (2.10). The result is

$$\int h(x)p(x)dx = \int \frac{p(x)}{\mu_n(x)n_i^2}dx = \int_E \frac{p(x)}{\mu_{nE}n_{iE}^2}dx + \int_B \frac{p(x)}{\mu_{nB}n_{iB}^2} + \int_C \frac{p(x)}{\mu_{nC}n_{iC}^2}dx \quad (A.1)$$

To obtain a closed-form expression for the transfer current, the integration has to be evaluated at each region of the transistor. Also, the integral of hole concentration is equal to hole charge as described below

$$\int p(x)dx = \frac{1}{qA_{E0}} \cdot Q \quad (A.2)$$

where

$$Q = qA_{E0} \int p(x)dx \quad (A.3)$$

The integrand defining the weighting function at each transistor region can assume an average value by using the mean-value theorem for integrals that means the weighting function can be taken out of the integral and assume an average value. The average weighting functions for the emitter and collector regions are given as $(\mu_{nE}n_{iE}^2)^{-1}$ and $(\mu_{nC}n_{iC}^2)^{-1}$, respectively. However, the weighting function in the base has inherent Ge mole fraction dependence which has to be decoupled before obtaining the average weighting value.

The integration of the first term of (A.1) over the neutral emitter region of $[x_0=0, x_{jeE}]$ and the used of average weighting function gives:

$$\int_E \frac{p(x)}{\mu_{nE}n_{iE}^2}dx = \int_{x_0}^{x_{jeE}} \frac{p(x)}{\mu_{nE}n_{iE}^2}dx = \frac{1}{\mu_{nE}n_{iE}^2} \int_{x_0}^{x_{jeE}} p(x)dx = \frac{1/(qA_E)}{\mu_{nE}n_{iE}^2} \cdot Q_{fE} \quad (A.4)$$

where Q_{fe} is the minority hole charge in the neutral emitter region. The integration of the second term of (A.1) over the base region the range of $[X_{je}, X_{jc}]$ which includes the space-charge regions of both sides and the substitution of intrinsic carrier as a function of SiGe alloy in the neutral base from (2.6) gives

$$\int_B \frac{p(x)}{\mu_{nB} n_{iB0}^2} dx = \int_{X_{je}}^{X_{jc}} \frac{p(x)}{\mu_{nB} n_{iB0}^2 \exp\left(\frac{m_{ac} x}{V_T}\right)} dx = \frac{1}{\mu_{nB} n_{iB0}^2} \int_{X_{je}}^{X_{jc}} \frac{p(x)}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx \quad (A.5)$$

where

$$p(x) = P_0 + \Delta p(x) \quad (A.5a)$$

$P_0 = N_B$ at zero bias, that is $V_{BE} = V_{BC} = 0$, is the hole concentration in the neutral base which is equal to N_B . Δp is the excess hole concentration, usually for $V_{BE} > 0$ and $|V_{BC}| > 0$, which consists of the following components: $p_{jeionized}$ and $p_{jcionized}$ are the ionized charge in the base-emitter and base-collector space charge regions, respectively. By substituting the various hole concentrations that are spatially independent in the base region and knowing that the Ge mole fraction profile starts from the metallurgical junction of base-emitter (SCR) to the metallurgical junction of base-collector (SCR), the integration part of (2.16) becomes

$$\int_{X_{je}}^{X_{jc}} \frac{p(x)}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx = \int_{X_{je}}^{X_{jc}} \frac{P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx + \int_{X_{je}}^{X_{jc}} \frac{p(x) - P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx \quad (A.6a)$$

or

$$\int_{X_{je}}^{X_{jc}} \frac{p(x)}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx = \int_{X_{je}}^{X_{jcB}} \frac{p(x) - P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx + \int_{X_{jcB}}^{X_{jc}} \frac{P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx + \int_{X_{je}}^{X_{jc}} \frac{p(x) - P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx + \int_{X_{jcB}}^{X_{jc}} \frac{p(x) - P_0}{\exp\left(\frac{m_{ac} x}{V_T}\right)} dx \quad (A.6b)$$

The results from the integration of each part of (A.6) and substituting back into (A.5) are as follows:

$$\int_{B\mu_{nB}n_{iB0}}^{\frac{p(x)}{2}} dx = \frac{\exp(-m_{ac}X_{jeB}/V_T)}{qA_E\mu_{nB}n_{iB0}} [h_{je}Q_{jE} + h_{p0}Q_0 + h_{jc}Q_{jC} + h_{fB}Q_{fB}] \quad (A.7)$$

where Q_{jE} , Q_0 , Q_{jC} , and Q_{fB} are base-emitter junction charge, zero bias hole charge, base-collector junction charge, and minority hole charge in the base region, respectively. The weighting factors h_{je} , h_{p0} , h_{jc} , and h_{fB} are expressed as

$$h_{je} = \frac{1}{m_{ac}w_{jE}/V_T} \left(\exp\left(\frac{m_{ac}w_{jE}}{V_T}\right) - 1 \right) \quad (A.7a)$$

$$h_{p0} = h_{fB} = \frac{1}{m_{ac}w_B/V_T} \left(1 - \exp\left(\frac{-m_{ac}w_B}{V_T}\right) \right) \quad (A.7b)$$

$$h_{jc} = \frac{\exp\left(\frac{-m_{ac}w_B}{V_T}\right)}{m_{ac}w_{jC}/V_T} \left(\exp\left(\frac{-m_{ac}w_{jC}}{V_T}\right) - 1 \right) = \exp\left(\frac{-m_{ac}w_B}{V_T}\right) \cdot \left(1 + \frac{m_{ac}w_{jC}}{2V_T} \right) \quad (A.7c)$$

where $w_{jE}=X_{jeB}-X_{je}$, $w_B=X_{jcB}-X_{jeB}$, and $w_{jC}=X_{jc}-X_{jcB}$ are widths of the following regions, base-emitter space-charge region in the base region, the neutral base region, and the base-collector region in the base region. Note that the widths are a function of their respective bias. This suggests that the weighting factors are a weak function of bias.

For the collector neutral region, the integration range is $x=[x_{jc}, x_C]$. The third term of (A.1) is solved with the indicated integration range and an average weighting function, the result is

$$\int_{C\mu_{nC}n_{iC}}^{\frac{p(x)}{2}} dx = \int_{x_{jc}C\mu_{nC}n_{iC}}^{xC} \frac{p(x)}{2} dx = \frac{1}{\mu_{nC}n_{iC}^2} \int_{x_{jc}C}^{xC} p(x) dx = \frac{1/(qA_{E0})}{\mu_{nC}n_{iC}^2} Q_{fC} \quad (A.8)$$

where Q_{fC} represents the minority hole charge in the neutral collector region. Substitution of (A.4, A.7, and A.8) back into (A.1) then substituting back into (2.9), the GICCR transfer current for SiGe HBTs becomes

$$I_{TSiGe} = \frac{(qA_{E0})^2 \overline{\mu_{nB} n_{iB0}^2} V_T \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right)}{Q_P^*} \left[\exp\left(\frac{V_{BE}}{V_T}\right) - \exp\left(\frac{V_{BC}}{V_T}\right) \right] \quad (A.9a)$$

where Q_P^* is the hole charge of the transistor which is a function of d.c bias

$$Q_P^* = h_{p0} Q_{P0} + h_{je} Q_{jE} + h_{jc} Q_{jC} + Q_f^* \quad (A.9b)$$

and Q_f^* is the minority hole charge which is given as

$$Q_f^* = h_{fE} Q_{fE} + h_{fB} Q_{fB} + h_{fC} Q_{fC} \quad (A.9c)$$

$$h_{fE} = \frac{\overline{\mu_{nB} n_{iB0}^2}}{\mu_{nE} n_{iE}^2} \cdot \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right) \quad (A.9d)$$

and

$$h_{fC} = \frac{\overline{\mu_{nB} n_{iB0}^2}}{\mu_{nC} n_{iC}^2} \cdot \exp\left(\frac{m_{ac} X_{jeB}}{V_T}\right) \quad (A.9e)$$

Now, the ICCR transfer current for Si BJT achieved by setting m_{ac} equal to zero that results in the weighting factors h_{je} , h_{p0} , h_{fB} , and h_{jc} equal to one becomes

$$I_{TSi} = \frac{(qA_{E0})^2 \overline{\mu_{nB} n_{iB0}^2} V_T}{Q_P} \left[\exp\left(\frac{V_{BE}}{V_T}\right) - \exp\left(\frac{V_{BC}}{V_T}\right) \right] \quad (A.10)$$

where

$$Q_P = Q_{P0} + Q_{jE} + Q_{jC} + Q_f \quad (A.10a)$$

Appendix B - Table of Model Parameters

This appendix contains model parameters along with some description for both compact models SGPM and HICUM. Table B1 contains parameter list for SGPM/Level2 which are used in HSPICE.

TABLE B1. Spice Gummel Poon Model Parameters (SGPM)

Parameter	Description	Default	Unit
Collector Current			
IS	Saturation current	1E-16	A
nF	Forward current non-ideality factor	1.0	-
VAF	Forward Early voltage	∞	V
VAR	Reverse Early voltage	∞	V
IKF	Forward Knee current	1.1	mA
ISC	Base-Collector leakage saturation current	0	A
IKR	Reverse Knee current	∞	A
NC	Base-Collector non-ideality factor	2	-
NR	Inverse non-ideality factor	2	-
Base Current			
BF	Ideal maximum forward current gain	100	-
ISE	Base saturation current	0	A
NE	Base non-ideality factor	1.0	-
BR	Ideal maximum reverse current gain	1	-
NR	Reverse current non-ideality factor	1	-
Series Resistance			
RC	Collector resistance	0	Ω
RE	Emitter resistance	0	Ω
RB	Zero-bias base resistance	0	Ω
RBM	Minimum(Extrinsic) base resistance	RB	Ω
IRB	Current where base resistance falls halfway to min. value	∞	A
C-E depletion capacitance			
CJE	Zero-bias depletion capacitance	3E-15	F
MJE	Exponent coefficient	0.5	-
VJE	Built-in voltage	0.9	V
FC	Factor for linearization of CJE($V_{BE} > FC * V_{JE}$)	0.8	-
C-B depletion capacitance			
CJC	Zero-bias depletion capacitance	0.46E-15	F
MJC	Exponent coefficient	0.3	-
VJC	Built-in voltage	0.7	V

TABLE B1. Spice Gummel Poon Model Parameters (SGPM)

Parameter	Description	Default	Unit
XCJC	Fraction of base-collector depletion capacitance connected to internal base node	1	-
	C-S depletion capacitance		
CJS	Zero-bias depletion capacitance	0	F
MJS	Exponent coefficient	0.3	-
VJS	Built-in voltage	0.75	V
	Transit Time		
TR	Ideal reverse transit time	4E-9	s
TF	Low current forward transit time	7E-12	s
XTF	Coefficient for bias dependence of TF	0	-
VTF	For modeling the voltage dependence	∞	V
ITF	For Modeling the current dependence	0	A
PTF	Excess phase of transconductance at fT	20	degree
	Temperature Parameter		
XTB	Forward and reverse current gain temperature coefficient	0	-
XTI	Saturation current temperature exponent	3	-
EG	Energy gap for temperature effect on IS	1.11	eV
T	Nominal temperature for simulation at which all model parameters are determined from measured	27	°C
	Noise Parameter		
KF	Flicker-noise coefficient	0	-
AF	Flicker-noise exponent	1	-

Table B2 contains the parameter list for HICUM/Level2

TABLE B2. HICUM Parameters

Parameter	Description	Default	Unit
Transfer Current			
cI0	Constant for GICCR	10^{-28}	A ² s
Qp0	Zero-bias hole charge	25	As
Ich	High-current correction	∞	A
hc	Weighting factor for collector minority charge in HBTs	1.0	-
hjc	Weighting factor for base-collector charge in HBTs	1.0	-
altT	Factor for additional delay time of iT	0.45	-
B-E depletion capacitance			
CjEi0	Zero-bias capacitance	15	F
VDEi	Built-in voltage	0.9	V
zEi	Exponential coefficient	0.5	-
alijEi	Maximum value divided by CjEi0	2.4	-
B-C depletion capacitance			
CjCi0	Zero-bias capacitance	2	F
VDCi	Built-in voltage	0.7	V
zCi	Exponential coefficient	0.4	-
VPTCi	Punch-through voltage	0	V
Forward Transit Time			
T0	Time constant	8E-12	s
V0	Voltage(Early-effect) low current densities	10^5	V
V1	Voltage(carrier jam)	0.05	V
alK	Factor for I_{CL}	0.2	-
T10	Time constant for medium current densities	10E-12	s
I10	Current (T1 adjustment)	15	A
al0	Exponent factor	2.0	-
T20	Time constant for high current densities	30E-12	s
s20	Slope caused by emitter charge	0	s/A
alQf	factor for additional delay time of Qf	0.22	-
Critical Current			
rCi0	Low-field resistance of epitaxial collector	150	Ω
Vlim	Limitation voltage	0.4	V
VPT	Punch-through voltage	3	V
VCEs	Internal C-E saturation voltage	0.1	V
Inverse Transit Time			
TR	Time constant	0	s
Base Current Components			

TABLE B2. HICUM Parameters

Parameter	Description	Default	Unit
IBeiS	B-E saturation current	10^{-19}	A
mBEi	B-E non-ideality factor	1.002	A
IREiS	B-E recombination saturation current	0	A
mREiS	B-E recombination non-ideality factor	2.0	-
IBCiS	B-C saturation current	0	A
mBCi	B-C non-ideality factor	1.5	-
fAVL	Coefficient for Avalanche effect	0	V^{-1}
qAVL	Exponent factor for Avalanche effect	0	As
rBi0	Internal base resistance	0	Ω
fDQr0	Correction factor for modulation	0.2	-
fge0	Geometry factor	0.6557	-
fQi	Ratio of internal to total minority charge	0.75	-
fCrBi	Ratio of high frequency shunt to total internal capacitance	0.2	-
	Peripheral Elements		
CjEp0	Zero-bias capacitance	0	F
VDEp	Built-in voltage	0.9	V
zEp	Depletion coefficient	0.5	-
aljEp	Maximum value divided by CjEi0	2.0	-
	Base Current Components		
IBEps	Saturation current	10^{-19}	A
mBEp	Non-ideality factor	1.002	-
IREps	Recombination saturation current	0	A
mREp	Recombination non-ideality factor	2.0	-
	External Elements		
CjCx0	Zero-bias depletion capacitance	0	F
VDCx	Built-in Voltage	0.7	V
zCx	Exponent coefficient	0.4	-
VPTCx	Punch-through voltage	0	V
CCox	Collector oxide capacitance	0	F
fBC	Partitioning factor for CjCx and CCox	0.7	-
	B-C Base Current Components		
IBCxS	Saturation current	0	A
mBCx	Non-ideality factor	1.5	-
	Other External Elements		
CEox	Emitter oxide(overlap) capacitance	0	F

TABLE B2. HICUM Parameters

Parameter	Description	Default	Unit
rBx	External base series resistance	0	Ω
rE	Emitter series resistance	0	Ω
rCx	External collector series resistance	0	Ω
	C-S Depletion Capacitance and Su resistance		
CjS0	Zero-bias depletion capacitance	0	F
VDS	Built-in voltage	0.6	V
zS	Exponent coefficient	0.5	-
VPTS	Punch-through voltage	0	V
rSu	Substrate series resistance	0	Ω
	Noise Parameters		
kF	Flicker noise factor	0	-
aF	Flicker noise exponent factor	2.0	-
krBi	Noise factor for internal base resistance	1.0	-
	Temperature Dependence		
VGb	Bandgap-voltage	1.17	V
alB	Relative temperature coefficient of current gain	5E-3	K^{-1}
zeta	Temperature coefficient(mobility) for epi-collector	1.6	-
alvs	Relative temperature coefficient of saturation drift velocity	10^{-3}	K^{-1}
alces	Relative temperature coefficient of VCEs	0.4E-2	K^{-1}
zetarBi	Temperature coefficient (mobility) internal base resistance	1.0	-
zetarBx	Temperature coefficient(mobility) external base resistance	0.3	-
zetarCx	Temperature coefficient(mobility) external coll. resistance	0.3	-
zetarE	Temperature coefficient of emitter resistance	0	-
	Self-Heating		
RTH	Thermal Resistance	0	K/W
CTH	Thermal Capacitance	0	Ws/K
TNOM	Temperature for which parameters are valid	300	K
DT	Temperature change for particular transistor	0	K

Appendix C - Derivation of α_β

The derivation for α_β is as follow. The temperature dependence d.c. current gain[8] is described by

$$\beta(T) = \frac{\mu_{nB}(T)}{\mu_{pE}(T)} \exp\left(\frac{\Delta V_{gb0} - \Delta V_{ge0}}{V_T}\right) \cdot \frac{\int_0^{W_B} N_B(x) dx}{\int_0^{W_E} N_E(x) dx} \quad (C.1)$$

By taking the ratio of $\beta(T)$ over $\beta(T_0)$ and assume that the temperature dependence mobilities for both hole and electron are approximately constant, the result becomes

$$\beta(T) = \beta(T_0) \cdot \exp\left(\frac{\Delta V_{gb0} - \Delta V_{ge0}}{V_T} \left(1 - \frac{T}{T_0}\right)\right) \quad (C.2)$$

The above equation shows that the current gain is mainly determined from the difference between emitter and base bandgap narrowing voltages and temperature. In practice, the bandgap narrowing voltage difference can be measured indirectly using the following:

$$x_\beta = \frac{1}{\beta(T)} \cdot \frac{d}{dT} \beta(T) = -\frac{\Delta V_{gb0} - \Delta V_{ge0}}{V_T} \cdot \frac{1}{T} \quad (C.3)$$

where α_β is a model parameter which represents the relative temperature coefficient. By substituting (C.3) into (C.2), the β becomes

$$\beta(T) = \beta(T_0) \exp[\alpha_\beta(T - T_0)] = \beta(T_0) \exp(\alpha_\beta \Delta T) \quad (C.4)$$

where $\Delta T = T - T_0$. Usually, ΔV_{ge0} is much greater than ΔV_{gb0} because the bandgap narrowing voltage is proportional to doping, and since the doping of emitter is much greater than the base doping, this means more bandgap narrowing voltage in the emitter and less in the base. Equation (C.3) can be simplified to

$$\alpha_\beta \approx \frac{\Delta V_{ge0}}{V_T} \cdot \frac{1}{T} \quad (C.5)$$

Appendix D - Simultaneous Extraction of R_{TH} - R_E Routine

A Matlab program listing for the simultaneous extraction of r_E and R_{TH} is shown.

```
% A matlab program to determine simultaneously model parameters,  $r_E$  and  $R_{TH}$ .
% Last change: 27.11.96
% By: Hai Tran
%
% Input requirements:
%   Measurement Data - Gummel plot (IC, IB vs. VBEX for a constant VCEX)
%                     - IB vs. VCEX for a CONSTANT VBEX
%   Model Parameters - alphaBeta(HICUM),  $r_{BX}$ , and  $m_{BEi}$ 
%
% Note-the model parameters,  $r_E$  &  $R_{TH}$ , can be applied to both HICUM and SGPM
%-----
%Clear variables
clear

%Load data from file
load ibm_05x10_rth
load ibm_05x10_gum

%Assign data to local variables
dwh=ibm_05x10_gum;
drth=ibm_05x10_rth; vbei=drth(1,7);

%Model parameters (note IBSEi and  $m_{BEi}$  are automatically determined from IB)
alphab=-8.0e-3;   (SiGe)
vg=1.14;         (SiGe)
%alphab=6.3e-3;   (silicon bjt)
%vg=1.17;         (silicon bjt)
rBx=7.92;

% constants
to=300; kb=1.38e-23;
vto=to*kb/1.602e-19;

%multiplying factor
l=1; mre=length(dwh);
for i=1:mre icwh(i,1)=dwh(i,5)*l;
    ibwh(i,1)=dwh(i,6)*l;
    beta(i,1)=icwh(i,1)/ibwh(i,1);
end

mrth=length(drth);
```



```

for i=1:mrth
    icrth(i,1)=drth(i,5)*1;
    ibrth(i,1)=drth(i,6)*1;
end

%extract IBESi & 1/(mBEi*vt) at low IB
for i=1:mre
    libwh(i,1)=log(ibwh(i,1));
end

lsqfit=polyfit(dwh(1:20,4),libwh(1:20,1),1);
slopewh=lsqfit(1);
mbei=1/(vto*slopewh);
iiswh=exp(lsqfit(2));

%determine Ibonh & Ibowh with no Re & rb influence
for i=1:mre
    ibowh(i,1)=iiswh*exp(dwh(i,4)*slopewh);
end

%plot gummel plot
figure(1) semilogy(dwh(:,4),ibwh,'go',dwh(:,4).ibowh)

%% loop to extract RE & RTH simultaneously
tol=1e-5; err=1;
nloop=1;
while (err > tol)
    if nloop == 1
        ibo=ibwh;
        ico=icwh;
        for i=1:mre
            betafact(i,1)=1;
        end
    end
    end

%check for positive lhs
kl=1;
i=1;
while i <= mre
    if mre ~= i
        error=(ibowh(mre+1-i) - ibo(mre-i+1))/ibowh(mre+1-i);
        if error < 0.01
            kl=i;
            i=mre;
        end
    end
end
end

```

```

    i=i+1;
end

%% extract RE first
kstart=0;
j=1;
for i=k1:mre
    % different method
    lhs(j,1)=(1/slopewh)*log(ibowh(i,1)/ibo(i,1))-rBx*ibo(i,1)
    rhs(j,1)=ico(i,1)+ibo(i,1);
    if i > k1 & kstart == 0
        if lhs(j,1) > lhs(j-1,1)
            kstart=j;
        end
    end
    j=j+1;
end

kend=kstart+20;
lsqfit=polyfit(rhs(:,1),lhs(:,1),1);
ext_re=lsqfit(1);
new_ext_re=ext_re
ydiff=polyval(lsqfit,rhs);

figure(2)
plot(rhs,lhs,'go',rhs,ydiff,'r-')

if nloop == 1
    err=1;
    old_extre=ext_re;
else
    err = abs(old_extre-ext_re)/old_extre;
    old_extre=ext_re;
end

%%% Extract Rth
%calculate total power dissipation
sum_ic=0;
for i=1:mrth
    ibrth1(i,1)=ibrth(i,1)*exp(icrth(i,1)*ext_re/vto);
    p1(i,1)=icrth(i,1)*drth(i,4)+ibrth1(i,1)*drth(i,7);
    sum_ic=sum_ic+icrth(i,1);
end

for i=1:mrth
    dib(i,1)=(ibrth1(i,1)-ibrth1(1,1))/ibrth1(1,1);

```

```

    dp(i,1)=p1(i,1)-p1(1,1);
end

lsqfit=polyfit(dp(1:7,1),dib(1:7,1),1);
slope=lsqfit(1);
y2=polyval(lsqfit,dp);

figure(3)
plot(dp,dib,'go',dp,y2)

newtemp=to*vto/(3*vto/mbei+vg/mbei-vbei/mbei-alphab*to*vto);
ext_rth=slope(1)*newtemp

%find DT
for i=1:mre
    dt=ext_rth*(ico(i,1)*dwh(i,2)+ibo(i,1)*dwh(i,1));
    t=300+dt;
    vt=t*1.38E-23/1.602e-19;
    denom=(t/to)^3*exp(vg/mbei/vt*(t/to-1)-alphab*dt);
    ibo(i,1)=ibwh(i,1)*exp(dwh(i,4)*(1/vto-1/vt))/denom;

    % betafact(i,1)=(1+dt*alphab);
    betafact(i,1)=exp(alphab*dt);
    betao(i,1)=beta(i,1)/betafact(i,1);
    ico(i,1)=betao(i,1)*ibo(i,1);
end
nloop=nloop+1
end

Final_rE=ext_re
Final_rth=ext_rth

```

References

- [1] Harame, D. L., et al., "Si/SiGe Epitaxial-Base Transistors-Part I: Materials, Physics, and Circuits", IEEE Trans. on Electron Devices, Vol.42, No.3, pp.455-468, 1995.
- [2] HSPICE User's Manual, Meta-Software, Inc., 1300 White Oaks Road, Campbell, CA 95008.
- [3] Schroter, M., "Bipolar Transistor Modeling for Integrated Circuit Design", TMC SCG Northern Telecom, 1995.
- [4] Cressler, J. D., "Re-Engineering Silicon: Si-Ge heterojunction bipolar technology", IEEE Spectrum, pp.49-55, March, 1995.
- [5] Richey, D. M., Cressler, J. D., and Joseph, A. J., "Scaling Issues and Ge Profile Optimization in Advanced UHV/CVD SiGe HBT's", IEEE Transactions on Electron Devices, Vol. 44, No. 3, pp. 431-440, 1997.
- [6] Schroter, M., Friedrich, M., and Rein, H., "A Generalized Integral Charge-Control Relation and its Application to Compact Models for Silicon-Based HBT's", IEEE Trans. Electron Devices, Vol. 40, No. 11, pp. 2036-2046, 1993.
- [7] Kasper, E. and Gruhle, A., "Silicon Germanium Heterojunction Transistor for High Speed Operation", BCTM, pp. 23-30, 1993.
- [8] Schroter, M., "DEVICE-A Program for Numerical Simulation of Three-Dimensional Heterostructure Semiconductor Devices with Inclusion of External Circuit Elements", User's Guide to Version 1.6, Ruhr-University Bochum, 1993.
- [9] Glicksman, M., "Mobility of Electrons in Germanium-Silicon Alloys", Physical Review, Vol. 3, pp. 125-128, 1958.
- [10] Gummel, H.K., "A Charge-Control Relation for Bipolar Transistors", BSTJ Vol. 49, pp.115-120, 1970.
- [11] Kroemer, H., "Two Integral Relations Pertaining to the Electron Transport Through a Bipolar Transistor with a Nonuniform Energy Gap in the Base Region", Solid-State Electronics, Vol. 28, No. 11, pp.1101-1103, 1985.
- [12] Early, J.M., "Effects of Space-Charge Layer Widening in Junction Transistors", Proc. IRE, Vol.40, 1952, pp. 1401-1407.
- [13] Manku T., et al., "Drift Hole Mobility in Strained and Unstrained Doped $\text{Si}_{1-x}\text{-Ge}_x$ Alloys", IEEE Transaction on Electron Devices, Vol. 40., pp.1990-1996 1993.

- [14] Ebers, J. and Moll, J., "Large-Signal Behavior of Junction Transistor", Proc. IRE. Vol. 42, 1954, pp. 1761-1772.
- [15] Gummel, H.K. and Poon, H.C., "An Integral Charge-Control Model For Bipolar Transistors", BSTJ Vol. 49, pp.827-852, 1970.
- [16] Graff H.C. and Klaassen F.M., "Compact Transistor Modelling for Circuit Designs", Springer-Verlag Wien, New York, 1990.
- [17] "MMSPICE: A Semi-Numerical Mixed-Mode Device/Circuit Simulator for Advanced Bipolar Technology CAD", Solid-State Electronics, Vol. 33, pp.1283-1291, 1990.
- [18] Rein H. M. and Schroter, M., "A Compact Physical Large-Signal Model For High-Speed Bipolar Transistors at High Current Densities - Part II: Two-Dimensional Model and Experimental Results", IEEE Transact. Electronics.
- [19] Schroter, M., "HICUM-An accurate physical bipolar compact transistor model", Bell Northern Research Documentation, August 1995.
- [20] IC-CAP, "HP Characterization Measuring System Manual", EESOF, 5601 Lindero Canyon Road, Westlake Village, CA 91362.
- [21] Getreu, I., "Modeling the Bipolar Transistor", Elsevier, New York, 1978.
- [22] Van Halen et al "SGP Parameter Extraction", BCTM, pp. 74-77, 1990.
- [23] Kendall, J., "Direct Extraction of Accurate DC Bipolar Spice Parameters for the Forward Active Region Without using Optimization", Proc. IEEE 1991 Int. Conference on Microelectronic Test Structures, Vol. 4., No.1, pp.197-204, 1991.
- [24] Tran, H., "Direct Extraction of SGPM Parameter for a BJT". Internal Report, Northern Telecom, 1993.
- [25] Schroter, M., "TRADICA - A program for dimensioning and calculation of integrated bipolar transistors", Manual Version 4.1, Northern Telecom, 1994..
- [26] Gonzalez, G., "Microwave Transistor Amplifiers". Prentice-Hall, 1984.
- [27] Roulston, D., "Bipolar Semiconductor Devices", McGraw-Hill, New York, 1990.
- [28] Cho, Hanjin and Burk, D. E., "A Three-Step Method for the De-Embedding of High Frequency S-Parameter Measurements", IEEE Transactions on Electron Devices, Vol. 38, No. 6, pp. 1371-1375, 1991.
- [29] Najm, F., "Simulation & Modeling - VBIC95, An Improved Bipolar Transistor Model". Circuits & Devices, pp. 11-15, 1996.

- [30] Walkey, D.J., et al., "Predictive Modeling of Lateral Scaling in Bipolar Transistors", BCTM., pp. 74-77, 1995.
- [31] MEDICI User's Manual, Technology Modeling Associates, Inc., 300 Hamilton Avenue, Palo Alto, CA 94301.
- [32] Meeks, D., "Fundamentals of Heat Transfer in a Multi-layer System", Microwave Journal, January, pp. 165-172, 1992.
- [33] Dennison, R.T. and Walter, K. M., "Local Thermal Effects in High Performance Bipolar Devices/Circuits", IEEE, pp. 164-167, 1989.
- [34] Reisch, M., "Self-Heating in BJT Circuit Parameter Extraction", Solid-State Electronics, Vol.35, No.5, pp.677-679, 1992.
- [35] Sewell, J., Liou, L. L., Barlage, D., Bozada, C., Dettmer, D., Fitch, R., Jenkins, T., Lee, R., Mack, M., Trombley, G., and Watson, P., "Thermal Characterization of Thermally-Shunted Heterojunction Bipolar Transistors", IEEE Electron Dev Lett., Vol. 17, pp. 19-21 1996.
- [36] Thornton, R.D., Dewitt, D., Chenette, E.R., and Gray, P.E., "Characteristics and Limitations of Transistors", Volume 4, John Wiley & Sons, New York, 1966. {tk7872 t73 t485 c.2}
- [37] Schafer, B. and Dunn, M., "Link Measurements to Nonlinear Bipolar Device Modeling", Microwave & RF, February, pp. 114-130, 1996.
- [38] Oettinger, F. F., Thermal Resistance Measurements, 1989 IEEE IRPS Tutorial
- [39] Smith D.S., Fraser A., and O'Neil, J., "Measurement and Prediction of Operating Temperatures for GaAs ICs", SEMI-THERM 86 Symposium, December 9-11 1986.
- [40] Liou, L., Ebel, J., and Huang, C., "Thermal Effects on the Characteristics of AlGaAs/GaAs Heterojunction Bipolar Transistors Using Two-Dimensional Numerical Simulation", IEEE Trans. on Electron Devices, Vol. 40, No. 1, pp. 35-43, 1990.
- [41] Fong, N. and Schroter, M., Evaluation of Emitter Resistance Measurement Methods", Internal Report, Northern Telecom, 1995.
- [42] Ganci, P. R., Hajjar, J. J., Clark, T., Humphries, P., Lapham, J., and Buss, D., "Self-Heating in High Performance Bipolar Transistors Fabricated on SOI Substrates", IEDM, pp. 417-420, 1992.
- [43] "A Method of the BJT Transient Thermal Impedance Measurement with Double Junction Calibration", IEEE Semi-THERM Symposium, pp. 80 -82, 1995.

- [44] Gray, P. R. and Meyer, R. G., “Analysis and Design of Analog Integrated Circuits”, Third Edition, John Wiley & Sons, Inc., 1993.

IMAGE EVALUATION TEST TARGET (QA-3)

APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved