Silicon-Germanium Heterojunction Bipolar Transistors

For a listing of related titles, turn to the back of this book.

Silicon-Germanium Heterojunction Bipolar Transistors

John D. Cressler
Guofu Niu

Artech House
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data
Cressler, John D.
Silicon-germanium heterojunction bipolar transistors / John D. Cressler, Guofu Niu.
p. cm.
Includes bibliographical references and index.
ISBN 1-58053-361-2 (alk. paper)
1. Bipolar transistors. 2. Germanium. I. Niu, Guofu. II. Title.

TK7871.96.B55 C74 2002
621.3815'28—dc21 2002038276

British Library Cataloguing in Publication Data
Cressler, John D.
Silicon-germanium heterojunction bipolar transistors
1. Bipolar transistors 2. Silicon 3. Germanium
I. Title II. Niu, Guofu
621.3'81528

ISBN 1-58053-361-2

Cover design by Christina Stone

685 Canton Street
Norwood, MA 02062

International Standard Book Number: 1-58053-361-2
Library of Congress Catalog Card Number: 2002038276

10 9 8 7 6 5 4 3 2 1

For Maria:
My beautiful wife, best friend, and soul mate for 20 years.
For Matthew John, Christina Elizabeth, and Joanna Marie:
God's awesome creations, and our precious gifts.
May your journey of discovery never end.

J.D.C.

For my wife:
Ying Li.
For my parents:
Pinzhang Niu and Xuehua Feng.

G.N.

Honesty of Thought
And Speech and Written Word
Is a Jewel,
And They Who Curb Prejudice
And Seek Honorably
To Know and Speak the Truth
Are the Only Builders of a Better Life.

John Galsworthy

知之为知之，不知为不知，是知也。

——孔子

Contents

Preface

While the idea of combining the semiconductor silicon and the semiconductor germanium for use in transistor engineering is an old one, only in the past decade has this concept been reduced to practical reality. The fruit of that effort is the silicon-germanium heterojunction bipolar transistor (SiGe HBT). The implications of the SiGe success story contained in this book are far-ranging and likely to be quite lasting and influential in determining the future course of the electronics infrastructure fueling the miraculous communications explosion of the twenty-first century.

This book is intended for a number of different audiences and venues. It should prove to be a useful resource as: 1) a hands-on reference for practicing engineers and scientists working on various aspects of SiGe technology, including: characterization, device design, fabrication, modeling, and circuit design; 2) a textbook for graduate or advanced undergraduate students in electrical and computer engineering (ECE), physics, or materials science who are interested in cutting-edge integrated circuit (IC) device and circuit technologies; or 3) a reference for technical managers and even technical support / technical sales personnel in the semiconductor industry. It is assumed that the reader has some modest background in semiconductors and bipolar devices (say, at the advanced undergraduate ECE level), but we have been careful to build "from-the-ground-up" in our treatment.

The spirit and vision for this book from day one was that it be "SiGe HBT from A to Z." That is, the book is intentionally very broad as well as very deep, and proceeds from a basic motivation and history of the subject, to materials, to technology and fabrication issues, to a detailed discourse on a wide range of fundamental aspects of SiGe HBT operation and design, spanning *dc* and *ac* characteristics, including noise and linearity. These fundamental topics are then supplemented by an even closer look at some of the "fine points" which might be confronted by experts in the field, including second-order phenomena, temperature effects, radiation tolerance, and numerical simulation. We conclude with a brief glimpse at likely future directions for SiGe technology. While we recognize that not all readers have need for exposure to all of these subjects, we like the notion of having a complete reference on the subject contained under one cover.

We have written this book in a manner consistent with our own preferences. Hence, it contains detailed, careful expositions of theory, a discussion of key device design trade-offs and constraints, "bottom-line" arguments on how important this or that phenomenon may be in the overall scheme of things, supporting data to bear out the various claims and theoretical arguments presented, and sufficient breadth and depth to be useful to both the novice and the expert. We also prefer a fairly informal writing style to ensure reader friendliness, and believe it is important to grasp the historical background, trends, and evolution of *any* subject. We have gone to considerable length to carefully reexamine and explicitly state assumptions in our theoretical treatments, and we have also intentionally not skipped the intermediate steps in some of the more important derivations – they are not often seen and deserve to be appreciated. We have highlighted what we feel to be the "open issues" associated with SiGe research that are in need of increased attention by the academic and industrial communities. This book contains a fairly substantial body of previously unpublished data and theory, as well as many careful and critical reinterpretations of the various nuances of the theory of SiGe HBTs. We have found again and again that while some particular theoretical discourse may previously reside in the literature (and even be widely cited), the existing presentation is often either confusing, is not correctly applied, does not fit the facts, or in some way is in need of a closer look. We have done that here. As with any in-depth work of this sort, there will be some among you who may disagree with our theory or interpretations. That's what science is all about! Feel free to contact us with any questions (or gripes!).

As any honest professor will readily concede, our graduate students play an absolutely essential role in our research. We professors may supply ideas, give encouragement, and guide interpretations (okay, and chip in some dollars as well!), but in the end, the really hard work belongs to our students. No exception here. Perhaps the greatest pleasure for us as professors is to behold the blooming of our students and the career successes they enjoy once they "leave the nest." We would like to take this opportunity to thank our graduate students, past and present, including *(J.D.C.)* — David Richey, Alvin Joseph, Bill Ansley, Juan Roldán, Stacey Salmon, Lakshmi Vempati, Jeff Babcock, Suraj Mathew, Mike Hamilton, Kartik Jayanaraynan, Greg Bradford, Usha Gogineni, Gaurab Banerjee, Shiming Zhang, Krish Shivaram, Dave Sheridan, Gang Zhang, Ying Li, Zhenrong Jin, QingQing Liang, Ramkumar Krithivasan, Zhiyun Luo, Tianbing Chen, Yuan Lu, and Chendong Zhu; and *(G.N.)* — Jin Tang, Jun Pan, Yan Cui, Yun Shi, Muthu Varadharajaperumal, and Seema Hegde. A special debt is owed to some of our students, since we have borrowed (unpublished) passages from several of their dissertations and theses (thanks especially to Alvin, David, and Stacey).

Much of the work presented in this book would not have occurred if our funding

sponsors had not embraced our vision of what is important in SiGe research. Special thanks are due to *(industry sponsors)* — IBM (Alvin Joseph, Dave Harame, Jim Dunn, Seshu Subbanna, Dave Ahlgren, Greg Freeman, Dean Herman, and Bernie Meyerson), Texas Instruments (John Erdeljac, Lou Hutter, Badih El-Kareh, and Dennis Buss), On Semiconductor (Joe Neel and Julio Costa), Maxim Semiconductor (Stewart Taylor), Analog Devices (John Yasaitis and Brad Scharf), Northrop Grumman (Harvey Nathanson, Bill Hall, and Rowan Messham) Hughes Electronics, now Boeing (Kay Jobe and Dave Sunderland), and the Semiconductor Research Corporation (Justin Harlow, Dale Edwards, and Jim Hutchby); and *(government sponsors)* — NASA Goddard Space Flight Center (Robert Reed, Cheryl Marshall, Paul Marshall, Hak Kim, Ken LaBel), the Defense and Threat Reduction Agency (Lew Cohn), DARPA, the U.S. Army Space and Missile Defense Command (Charlie Harper and Aaron Corder), NAVSEA Crane (Steve Clark and Dave Emily), Mission Research Corporation (Dave Alexander and Mary Dyson), the Office of Naval Research (Al Goodman), the NASA Center for Space Power and Advanced Electronics (Henry Brandhorst), the Naval Research Laboratory (Fritz Kub), the National Science Foundation, EPSCOR, and the Jet Propulsion Laboratory (Jagdish Patel).

It is also true that much of the work presented in this book would have been impossible without the generous support of the SiGe team at IBM, which provided our research group with ready access to their state-of-the-art hardware. In fact, the lion's share of the data presented in this book was measured on IBM devices, and we are especially grateful to Bernie Meyerson, Dean Herman, David Harame, Alvin Joseph, and Seshu Subbana, in particular, for making that happen. Many current members of the IBM SiGe team contributed directly to various aspects of our research, including Greg Freeman, Dave Ahlgren, Rob Groves, Jim Dunn, Chuck Webster, Fernando Guarin, Lou Lanzerotti, Kim Newton, Herschel Ainspan, and Mehmet Soyuer. John Cressler would like to reach back and thank his colleagues at IBM Yorktown who played a pivotal role in his early education in bipolar devices and SiGe HBTs, including Denny Tang, Tak Ning, Hans Stork, Jim Comfort, Emmanuel Crabbé, Achim Burghartz, Jack Sun, Dave Harame, Keith Jenkins, Kent Chuang, Jim Warnock, and Scott Stiffler.

We would like to thank Mike Palmer and Charles Ellis of the AMSTC for assistance in wire-bonding and fabrication, as well as Dick Jaeger and Dave Irwin for their constant support during the arduous writing process. Several individuals scattered across the globe contributed to various ideas or data presented in this book, including: Stewart Taylor, Larry Larson, Joerg Osten, Dieter Knoll, Robert Plana, and Frank Herzel. Our apologies if we have left anyone out. The body of knowledge contained in this work truly represents the efforts of a great many dedicated engineers and scientists.

Whew! This book has been a year-long labor of love, and although "fun" might be too harsh a word, given the countless hours required, SiGe is a subject we care deeply about, love to talk about, and remain fascinated by. There are many interesting puzzles left in SiGe! It has been immensely satisfying to see both the dream of SiGe and this book become a reality. We hope our efforts please you. Enjoy!

John D. Cressler
School of Electrical and Computer Engineering
Georgia Institute of Technology
December 2002

Guofu Niu
Department of Electrical and Computer Engineering
Auburn University
December 2002

Chapter 1

Introduction

Simply stated, silicon-germanium is "an idea whose time has come." [1] While the concept of combining silicon (Si) and germanium (Ge) into an alloy for use in transistor engineering is an old one, only in the past decade has this concept been reduced to practical reality. The implications of this success story are far ranging and likely to be quite lasting and influential in determining the future course of the communications explosion during the twenty-first century. This introductory chapter sets the stage for the detailed look at the silicon-germanium heterojunction bipolar transistor (SiGe HBT) presented in this book. We first examine the compelling features of the semiconductor Si, look at integrated circuit (IC) needs to support the emerging Information Age, and then examine application-induced design constraints. Armed with this background, the notion of using bandgap-engineering in Si to create the SiGe HBT is introduced, and we address why SiGe HBT BiCMOS technology has emerged as an important enabler for twenty-first century communications systems. We conclude with an historical perspective of this fascinating field, some performance trends, and a view of the looming technology battleground between Si, SiGe, and III-V technologies.

1.1 The Magic of Silicon

We live in a silicon world. This statement is literally as well as figuratively true. Silicates, the large family of silicon-oxygen bearing compounds such as feldspar ($NaAlSi_3O_6$), beryl ($Be Al_2(Si_6O_{18})$), and olivine ($(MgFe)_2SiO_4$), make up 74% of the earth's crust. Si is the third most abundant element on planet Earth (15% by weight), after iron (35%) and oxygen (30%). One need go no further than the

[1] "There is one thing stronger than all the armies of the world, and that is an idea whose time has come." Victor Hugo

beach and scoop up some sand to hold Si in your hand. More important, however, Si, with its many compelling characteristics, has almost single-handedly fueled the emergence of the Information Age. Global semiconductor sales, of which Si captured well over 90%, totaled $204,400,000,000 in 2000 [1]. We humans owe a significant debt to this unique element. Indeed, it is the very existence of Si microelectronics that has enabled emergence of the Information Age, which is so profoundly reshaping the way we live and work and play and communicate. Why Si? This profound market dominance of Si rests on a number of surprisingly practical advantages Si has over other competing semiconductors, including the following.

- Si is wonderfully abundant (there are a lot of beaches in the world), and can be easily purified to profoundly low background impurity concentrations (below 10^{10} impurities / cm^3). Given that the atomic density of Si is $5x10^{22}$ atoms / cm^3, this means that in a production-grade Si wafer, the impurities are smaller than 1 part in 10^{12} (0.000001 ppm), making them some of the purest materials on Earth.

- Si crystals can be grown in amazingly large, virtually defect-free single crystals (200 mm diameter wafers are in production today worldwide, and are rapidly evolving to 300 mm). The resultant large Si wafer size translates directly into more ICs per wafer, effectively lowering the cost per IC. Given that a 200 mm Si boule is roughly 6 feet long, Si crystals are literally the largest and most perfect on the face of planet Earth.

- Si has excellent thermal properties, allowing for the efficient removal of dissipated heat. The thermal conductivity of Si at 77 K is actually larger than that of copper.

- Si can be controllably doped with both n-type and p-type impurities to extremely high dynamic range (less than 10^{14} to greater than 10^{21} cm^{-3}), at moderate incorporation temperatures (e.g., < 1000 °C). In addition, the ionization energies of the three principal dopants for Si (boron, phosphorus, and arsenic) are all at shallow levels in the bandgap (< 50 meV), making them essentially 100% ionized (electrically active) at room temperature.

- Si can be very easily grown or deposited in three different material forms: crystalline Si, polycrystalline Si ("poly" Si), or amorphous Si, each of which finds different uses in IC technologies.

- Si can be etched relatively easily, using either "wet" chemistries (e.g., KOH) or with "dry" chemistries (e.g., with reactive ion etching using CF_6).

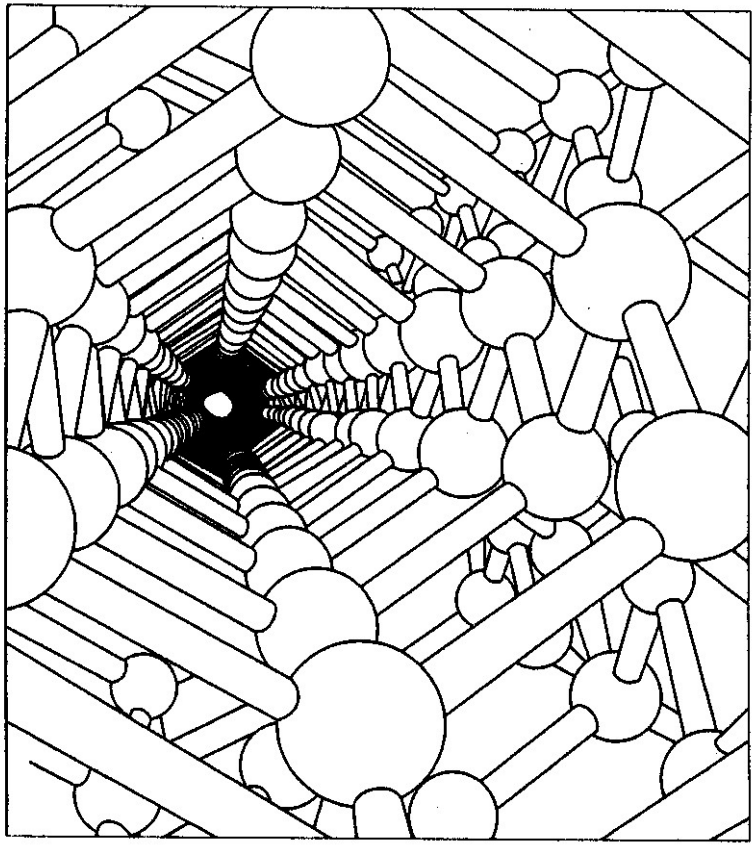

Figure 1.1 End-on view of the Si lattice along the <110> axis (after [2]).

- Like most of the technologically important semiconductors, Si crystallizes in the diamond lattice structure (Figure 1.1 and Figure 2.1). The crystal structure of Si is a direct consequence of its electron orbital configuration ($1s^2 2s^2 2p^6 3s^2 3p^2$), and is thus the underlying reason why Si has so many desirable mechanical and thermal properties. (It's a shame that crystallized Si, despite the fact that it shares an identical lattice structure with that of crystallized C (our beautiful diamonds), ends up with an opaque, fairly uninspiring, greyish-silver appearance.)

- The energy bandgap of Si is of moderate magnitude (1.12 eV at 300 K). If the bandgap were too small (< 0.5 eV), the intrinsic carrier density would be too large at 300 K, making parasitic off-state leakage currents too large. If,

instead, the bandgap were too large (> 2.0 eV), then typically it becomes difficult to etch and diffuse dopant impurities (bandgap is a reflection of atomic bonding strength).

- Si is nontoxic and highly stable, making it in many ways the ultimate green material (although its common dopant sources of di-borane, phosphine, and arsine fall decidedly into the "nasty" category).

- Si has excellent mechanical properties, facilitating ease of mechanical handling during the fabrication process. For a 200-mm diameter crystal, for instance, this allows the Si wafers to be cut to roughly 600-μm thickness, maximizing the number of wafers per Si boule. This mechanical stability also minimizes wafer warpage with fabrication, and in addition allows processing to occur under very large thermal gradients without serious consequences (e.g., under rapid-thermal annealing conditions, ramping from 25°C to 1,000°C in 10 seconds).

- It is remarkably easy to form very low resistance ohmic contacts to Si, using a wide variety of metals and doping conditions. Specific contact resistances below 10–20 $\Omega\mu m^2$ can be achieved, for instance, with a heavily doped polysilicon emitter contact, minimizing parasitic device resistances.

- The damage and resultant interface states associated with cleaving or truncating a Si crystal to produce a crystalline surface are not excessively numerous and, importantly, can be easily passivated to manageable levels (e.g., with hydrogen). In device terms, this results in a low surface recombination velocity for Si, and a reduction in parasitic leakage currents and noise associated with surface leakage phenomena.

- The diffusion coefficients of the common Si dopants are "reasonable," meaning that these dopants can be ion-implanted, and then effectively moved to active substitutional sites using comparatively small thermal cycles (temperature and time). This modest annealing cycle also very efficiently restores the crystalline integrity of the Si lattice. This fact is crucial for allowing the formation of shallow junctions, and the maintenance of the thin doping profiles needed for making high-speed transistors.

- Perhaps most importantly, an extremely high-quality dielectric can be trivially grown on Si, simply by flowing oxygen across the wafer surface at an elevated temperature (or even sitting it on the shelf for a few short minutes). This dielectric, silicon-dioxide (SiO_2, "quartz" to geologists) is one of nature's most perfect insulators (it possesses a breakdown strength greater than

10 MV/cm) and can be used for electrical isolation, surface passivation, a planarization layer, an etch stop, or as an active layer (e.g., gate oxide) in the device. SiO_2 also acts as a wonderful diffusion and ion-implantation barrier to dopants, and thus functions as an ideal masking material for layer-by-layer stenciling of the features of our integrated circuits.

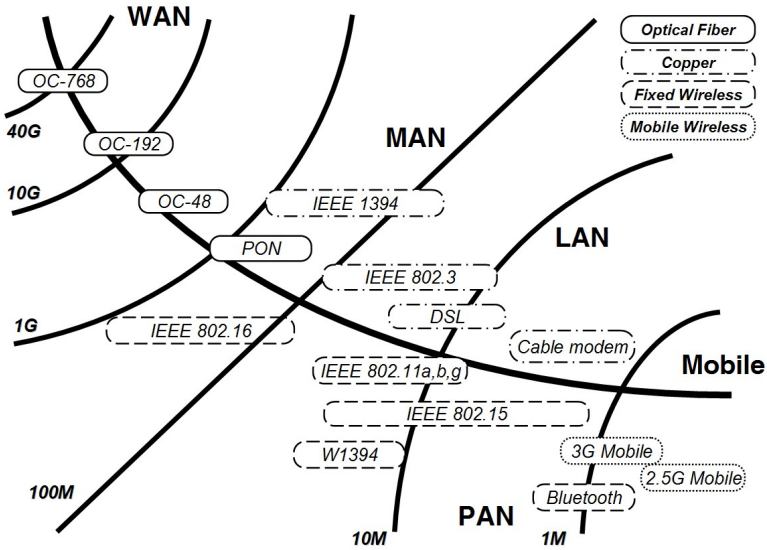

Figure 1.2 The global communications landscape in 2002, broken down by the various communications standards, and spanning the range of: wireless to wireline; fixed to mobile; copper to fiber; low data rate to broadband; and local area to wide area networks. WAN is *wide area network*, MAN is *metropolitan area network*, the so-called "last mile" access network, LAN is *local area network*, and PAN is *personal area network*, the emerging in-home network. (Used with the permission of Kyutae Lim, Georgia Tech.)

Simply put, it is a remarkable fact that nature blessed us with a single material embodying the features one might naively wish for when building low-cost transistors and ICs. From a semiconductor manufacturing standpoint, Si is literally a dream come true. Why is Si the driver of the Information Age? There is literally no other semiconductor that so nicely "fits the bill" as a material from which to construct the roughly $2x10^{20}$ transistors that currently reside today on planet Earth. Interestingly, the wonderful selling points of Si as a fundamental enabler of the

Information Age have little to do with the device or circuit designer's desires, or needs, and in fact are largely driven by manufacturing, yield, and ultimately cost issues. That is, mundane, but nonetheless compelling, economic issues command the driver's seat. They still do, and clearly will far into the future.

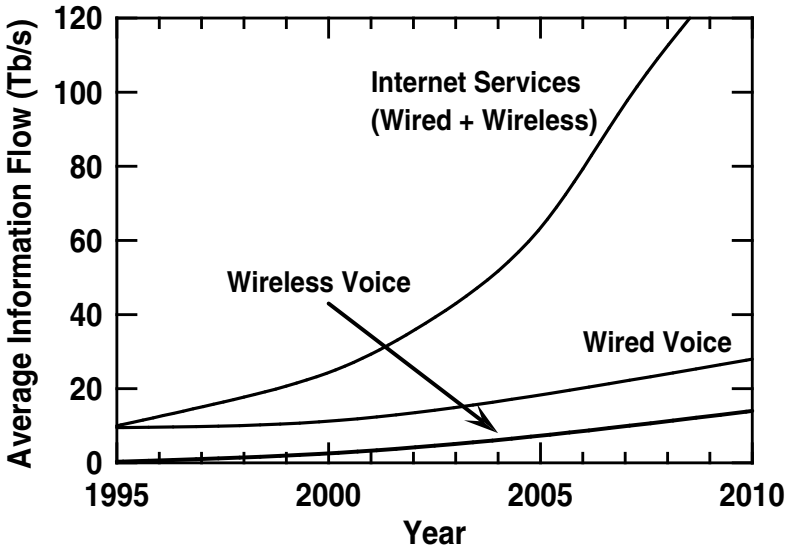

Figure 1.3 Projected global growth in information flow for wired voice, wireless voice, and Internet services (after [3]).

1.2 IC Needs for the Twenty-First Century

Despite the relative infancy of the Information Age, the requirement for integrated circuits and systems is undergoing explosive growth in the global marketplace, a growth that is unlikely to abate in the foreseeable future (Figure 1.2). Indeed, there are few people today who would project any kind of saturation, at least until the physical limits of our conventional semiconductor devices are reached in the 2010–2015 time frame. Even as those horizons inexorably come into view, the frantic search for faster and more complex circuits will only shift directions; it will not cease. Clear evidence for these trends can be found in the growth in average global information flow for wired voice, wireless voice, and Internet applications (Figure 1.3). As can be seen in the evolutionary path of Internet-based services, the Information Age is rapidly evolving into what might be appropriately termed the *Internet Age*, since the Internet appears to be the predominant enabling

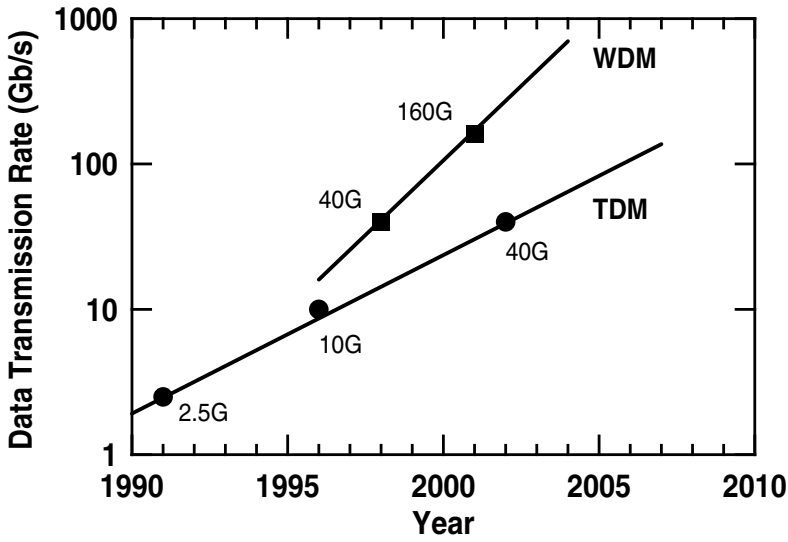

Figure 1.4 Trends in data transmission rates for optical fiber backbone networks. TDM is *time division multiplexing*, and WDM is *wavelength division multiplexing* (after [3]).

medium. The wireline data transmission rates along the global fiber backbone network required to support this projected growth in Internet services are increasing exponentially, fueling what can be termed the *Communications Revolution* (Figure 1.4).

Because of this relentless pace in global information generation, manipulation, storage, and transmission, an insatiable appetite for exponentially greater system-level computational complexity and performance has resulted, translating at the IC level into a demand for increasingly faster logic, increasingly higher memory density, and increasingly higher carrier frequencies for communications channels, as embodied in the well-known Moore's Law growth patterns in the various IC metrics. All at a lower price! Faster, denser, cheaper, the motto of the IC marketer in the twenty-first century. Often a disturbing oxymoron to us IC designers.

1.3 Application-Induced Design Constraints

Where does this evolutionary juggernaut of faster, denser, cheaper ICs leave us poor device and circuit designers? Ask anyone working in the IC trenches, and they will tell you that as IC operational throughput rises, life as a device and circuit designer gets exponentially more difficult! To appreciate why this is so, one simply

needs to consider the design constraints imposed by the various types of IC venues that are required to support emerging Information Age applications. By way of illustration, consider simultaneously a classical digital IC (e.g., a microprocessor), a classical analog IC (e.g., a data converter), and a classical RF or microwave IC (e.g., a low noise amplifier). If we deconvolve the various constraints a device and circuit designer necessarily confronts when designing, modeling, laying out, fabricating, testing, packaging, and selling such ICs, some fundamental observations can be made. (Cost is clearly a primary constraint for all application sectors.) These application-induced IC design constraints include:

- Digital circuits (e.g., a microprocessor):
  - switching speed;
  - power consumption;
- Analog circuits (e.g., a data converter):
  - frequency response;
  - output conductance;
  - current gain;
  - 1/f noise;
  - power consumption;
  - temperature coefficient;
  - device-to-device matching;
  - resistor tolerance;
- RF and microwave circuits (e.g., a low-noise amplifier):
  - broadband noise;
  - 1/f noise;
  - linearity;
  - power gain;
  - power consumption;
  - Q of inductors and capacitors;
  - impedance matching;
  - transmission lines;
  - modulation scheme (e.g., GSM versus CDMA, etc.).

A cursory glance at these three disparate application arenas paints a very clear picture. The performance requirements at the device and circuit level vary radically depending on the intended application. For instance, the key driving force in low-noise amplifier (LNA) design might be transistor noise figure, but a logic designer on a microprocessor design team most likely could care less about noise figure. This design constraint disparity translates to the system level as well. If we consider a generic radio frequency (RF) transceiver, which might, for instance, make up a cell phone, we see that multiple device technologies are required, ranging from: an RF power amplifier capable of large voltage swings, an RF LNA with very low noise capability, RF mixers and oscillators, memory, passives for matching and filtering, data converters, and digital complementary-metal-oxide-semiconductor (CMOS) for baseband processing (Figure 1.5). In today's cell phones, these individual functional blocks are typically packaged as separate ICs using distinct IC technologies in order to achieve acceptable system performance at the lowest possible cost (e.g., GaAs metal-semiconductor field effect transistor (MESFET) or HBT technology for the low-noise amplifier (LNA) and power amplifier (PA), Si BJT technology for the mixer and oscillator and converters, and Si CMOS technology for baseband processing and digital signal processing (DSP)).

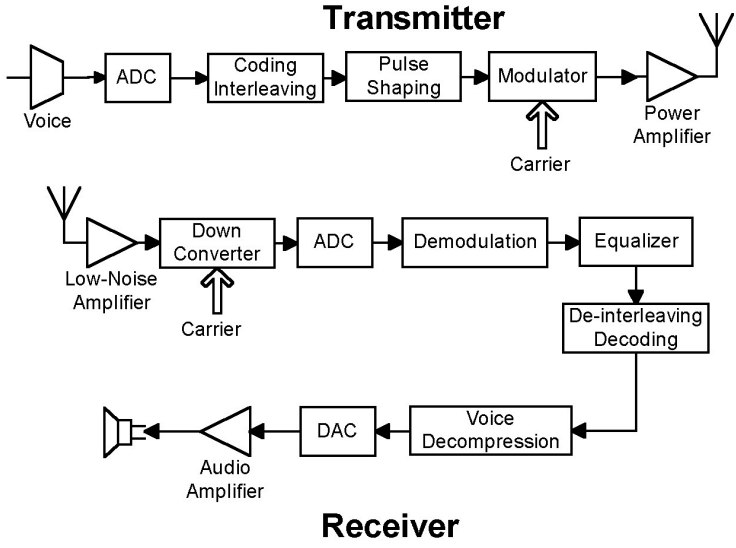

Figure 1.5 A generic RF transceiver architecture.

Given the over-arching theme of cost constraints at the IC level, however, we are led to a logical conclusion. It would be nice if a *single* IC device technol-

ogy was capable of simultaneously supporting all of the types of self-conflicting circuit design needs: digital, analog, and RF. That is, a "one-technology-fits-all" approach would seem to offer compelling advantages from a cost standpoint, potentially enabling "system-on-a-chip" (SoC) integration (Figure 1.6). While the extent to which SoC will dominate the global communications market over the long haul remains a contentious issue, clearly the trend in most foreseeable communications applications favors an increased level of functional integration in order to achieve reduced form factor, lower chip count, longer battery life, reduced packaging complexity, and ultimately lower total system cost.

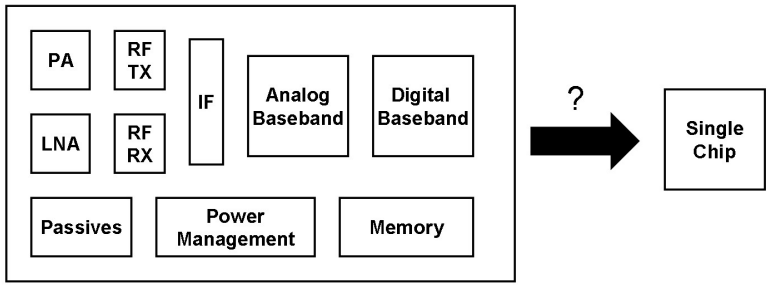

Figure 1.6 Block diagram of a generic cell phone, suggesting a path to single chip integration.

The system-level SoC dream can quickly translate, however, into a device designer's nightmare. Any practicing device engineer will tell you that a single transistor technology simultaneously capable of delivering low-power, high-linearity, low-noise, and high-speed operation for RF, analog, memory, and digital circuits all at a low cost just doesn't exist. Or does it? If we scan the entire field of available IC technologies, we are led inexorably to a logical conclusion. As SoC IC designers we would ideally like to combine the superior RF and analog performance properties of III-V technologies with Si CMOS for digital and memory functions, all married together with the economy of scale and low cost associated with Si IC manufacturing. A Si-compatible, III-V device technology? You bet!

1.4 The Dream: Bandgap Engineering in Silicon

As wonderful as Si is from a fabrication viewpoint, from a device designer's perspective, Si is hardly the ideal semiconductor. The carrier mobility for both electrons and holes in Si is comparatively small, and the maximum velocity that these

carriers can attain under high electric fields is limited to about $1x10^7$ cm/sec under normal conditions. Since the speed of a device ultimately depends on how fast the carriers can be transported through the device under sustainable operating voltages, Si can thus be regarded as a somewhat "slow" semiconductor. In addition, because Si is an indirect gap semiconductor, light emission is painfully inefficient, making active optical devices such as diode lasers impractical. Many of the III-V compound semiconductors (e.g., GaAs or InP), on the other hand, enjoy far higher mobilities and saturation velocities, and because of their direct gap nature, generally make efficient optical devices. In addition, III-V devices, by virtue of the way they are grown, can be compositionally altered for a specific need or application (e.g., to tune the light output of a diode laser to a specific wavelength). This atomic-level custom tailoring of a semiconductor is called *bandgap engineering*, and yields a large performance advantage for III-V technologies over Si [4]. Unfortunately, these benefits commonly associated with III-V semiconductors pale in comparison to the practical deficiencies associated with making highly integrated, low-cost ICs from these materials. There is no robust thermally grown oxide for GaAs or InP, for instance, and wafers are smaller with much higher defect densities, more prone to breakage, poorer heat conductors, etc. These deficiencies translate into generally lower levels of integration, more difficult fabrication, lower yield, and ultimately higher cost. In truth, of course, III-V materials such as GaAs and InP fill important niche markets today (e.g., GaAs MESFETs for cell phones, AlGaAs or InP-based lasers), but III-V semiconductor technologies will never become mainstream if Si-based technologies can do the job.

While Si ICs are well suited to high-transistor-count, high-volume microprocessors and memory applications, RF and microwave circuit applications, which by definition operate at significantly higher frequencies, generally place much more restrictive performance demands on the transistor building blocks. In this regime, the poorer intrinsic speed of Si devices becomes problematic. That is, even if Si ICs are cheap, they must deliver the required device and circuit performance to produce a competitive system at a given frequency. If not, the higher-priced but faster III-V technologies will dominate (as they indeed have until very recently in the RF and microwave markets).

The fundamental question then becomes simple and eminently practical: is it possible to improve the performance of Si transistors enough to be competitive with III-V devices for RF and microwave applications, while preserving the enormous yield, cost, and manufacturing advantages associated with conventional Si fabrication? The answer is clearly yes, and this book addresses the many nuances associated with using strained SiGe alloys to practice bandgap engineering in the

Si material system, a process culminating in the SiGe HBT. [2]

While the basic idea of using SiGe alloys to bandgap-engineer Si devices dates to the 1950s (Shockley considered it early in the transistor game), the synthesis of defect-free SiGe films proved surprisingly difficult, and device-quality SiGe films were not successfully produced until the mid-1980s. This difficulty has a very obvious physical underpinning. While Si and Ge can be combined to produce a chemically stable alloy, their lattice constants differ by roughly 4.2% and thus SiGe alloys grown on Si substrates are compressively strained. This process is referred to as *pseudomorphic* growth of strained SiGe on Si, with the SiGe film adopting the underlying Si lattice constant. These SiGe strained layers are subject to a fundamental stability criterion limiting their thickness for a given Ge concentration [5, 6]. Deposited SiGe films that lie below the stability curve are thermodynamically stable, and can be processed using conventional furnace or rapid-thermal annealing, or ion-implantation without generating defects. Deposited SiGe films that lie above the stability curve, however, are "metastable" and will relax to their natural lattice constant (> Si) if exposed to temperatures above the original growth temperature, generating device-killing defects in the process. For a manufacturable SiGe technology, it is obviously key that the SiGe films remain stable after processing. Stability of SiGe strained layers will be discussed at length in Chapter 2.

1.5 The SiGe HBT

Introducing Ge into Si has a number of consequences. First and most important, because Ge has a larger lattice constant than Si, the energy bandgap of Ge is smaller than that of Si (0.66 eV vs 1.12 eV), and thus SiGe will have a bandgap smaller than that of Si, making it a suitable candidate for bandgap engineering in Si. The compressive strain associated with SiGe alloys produces an additional bandgap shrinkage, and the net result is a bandgap reduction of approximately 75 meV for each 10% of Ge introduced. This Ge-induced "band offset" occurs predominantly in the valence band, making it conducive for use in tailoring *npn* bipolar transistors. In addition, the compressive strain lifts the conduction and valence band degeneracies at the band extremes, effectively reducing the density-of-states and improving

[2] It is technically correct to refer to silicon-germanium alloys according to their chemical composition, $Si_{1-x}Ge_x$, where x is the Ge mole fraction. Following standard usage, such alloys are usually referred to as "SiGe" alloys. Note, however, that it is common in the material science community to also refer to such materials as "Ge:Si" alloys. In this book we will follow standard usage and denote these materials as SiGe alloys. Believe it or not, this field also has its own set of slang pronunciations. The colloquial usage of \'sig-ee\ to refer to "SiGe" (begun at IBM in the late 1990s) has come into vogue recently, although we remain purists in this regard, sticking with the more traditional "silicon-germanium."

the carrier mobilities with respect to pure Si (the latter due to a reduction in carrier scattering). Because a practical SiGe film must be very thin if it is to remain stable and hence defect free, it is a natural candidate for use in the base region of a bipolar transistor (which by definition must be thin to achieve high-frequency operation). The resultant device contains an n-Si / p-SiGe emitter-base heterojunction and a p-SiGe / n-Si base-collector heterojunction, and thus this device is properly called an "SiGe double-heterojunction bipolar transistor," although for clarity we will continue the standard usage of "SiGe heterojunction bipolar transistor" (SiGe HBT).[3] The SiGe HBT represents the first practical bandgap-engineered transistor in the Si material system.

Perhaps most importantly, SiGe HBTs can be quite easily teamed with best-of-breed Si CMOS to form a monolithic SiGe HBT BiCMOS technology. While this might seem at first glance to be a mundane advantage, it is in fact a fundamental enabler for SiGe's long-term success, provided SiGe HBTs can be realized without an excessive cost penalty compared to standard Si ICs. The integration of SiGe HBTs with Si CMOS is also the fundamental departure point between SiGe technology and III-V technologies. If SiGe technology is to be successful in the long haul, it must bring to the table the RF and analog performance advantages of the SiGe HBT, and the low-power logic, integration level, and memory density of Si CMOS, into a single cost-effective IC that enables SoC integration (i.e., SiGe HBT BiCMOS). This merger appears to be the path favored by most companies today. Typically, SiGe HBTs (often with multiple breakdown voltages) exist as an "adder" to a basic CMOS IC building-block core, to be swapped in or out as the application demands, without excessive cost burden. Typical state-of-the-art SiGe HBT BiCMOS technologies generally have a roughly 20% adder in mask count compared to "vanilla" digital CMOS, and are viewed by many as an acceptable compromise between performance benefit and cost, depending on the application. In truth, SiGe HBT BiCMOS technologies are the future of the SiGe HBT, since it enables system-on-a-chip solutions across a very broad market base for both wired and wireless applications, all at an acceptable cost. This is clearly the evolutionary path being traveled today by almost all companies with commercially viable SiGe

[3] A common misconception persists in the literature that the SiGe HBT is not a "true" HBT, but rather some sort of "mutant" bipolar junction transistor (BJT). While it is true that the fundamental doping profile design of most SiGe HBTs in production today does not follow the lines of their III-V HBT brethren, the SiGe HBT is still an HBT. Traditional III-V HBTs exploit a wide bandgap emitter to reduce the back-injected base current (i.e., improve the emitter injection efficiency), thereby allowing an acceptable current gain while using a lightly doped emitter and a very heavily doped base. SiGe HBTs, on the other hand, typically employ a graded-Ge-base design with a heavily doped emitter and moderately doped base, similar to what might be found in a conventional Si BJT. Nevertheless, SiGe HBTs do in fact still contain dual SiGe/Si heterojunctions and thus should be properly referred to as HBTs.

technologies.

1.6 A Brief History of SiGe Technology

The concept of the HBT is an old one, dating to the fundamental BJT patent issued to Shockley in 1951 [7]. Given that the first bipolar transistor was built from Ge, it seems quite likely that Shockley even envisioned the combination of Si and Ge to form a SiGe HBT (he was a bright guy!). The basic formulation and operational theory of the HBT was pioneered by Kroemer, and was in place by 1957 [8, 9]. [4] Reducing the SiGe HBT to practical reality, however, took 30 years due to material growth limitations. Once device-quality SiGe films were achieved in the mid-1980s, progress was quite rapid from that point forward. An interesting historical discussion of early SiGe HBT development is contained in [10]. Table 1.1 summarizes the key steps in the evolution of SiGe HBT technology.

The first functional SiGe HBT was demonstrated in December of 1987 [16], [5] but worldwide attention became squarely focused on SiGe technology in June of 1990 with the demonstration of a non-self-aligned SiGe HBT grown by ultra-high vacuum/chemical vapor deposition (UHV/CVD), with a peak cutoff frequency of 75 GHz [18, 19]. At the time, this SiGe result was roughly twice the performance of state-of-the-art Si BJTs (Figure 1.7), and clearly demonstrated the future performance potential of the technology. Eyebrows were lifted, and work to develop SiGe as a practical circuit technology began in earnest in a large number of laboratories around the world.

In December of 1990, the first emitter-coupled-logic (ECL) ring oscillators using self-aligned, fully integrated SiGe HBTs were produced [20]. The first SiGe BiCMOS technology was reported in December of 1992 [22], and the first LSI SiGe HBT circuit (a 1.2 GSample/s 12-bit digital-to-analog converter) was demonstrated in December of 1993 [23]. The first SiGe HBTs with frequency response greater than 100 GHz were described in December of 1993 [24, 25], and the first

[4] Kroemer was awarded the Nobel Prize in 2000 for his work in bandgap engineering.

[5] It is an interesting and often overlooked historical point that at least three independent groups were simultaneously racing to demonstrate the first functional SiGe HBT, all using the molecular beam epitaxy (MBE) growth technique: an IBM team [33], a Bell Labs team [34], and a Linköping University team [35]. The IBM team is fairly credited with the victory, since it presented (and published) its results in early December 1987 at the IEDM (it would have been submitted to the conference for review in the summer of 1987) [16]. Even for the published journal articles, the IBM team was the first to submit their paper for review (on November 17, 1987), followed by the Bell Labs team (on November 23, 1987), and the Linköping University team (on February 22, 1988). All three papers appeared in print in the spring of 1988. The first SiGe HBT demonstrated using (the more manufacturable) CVD growth technique followed shortly thereafter [17].

Table 1.1 Key Steps in the Evolution of SiGe HBT Technology

Historical Event	Year	Reference
Fundamental HBT patent	1951	[7]
Drift-base HBT concept	1954	[8]
Basic HBT theory	1957	[9, 12, 13]
First growth of SiGe strained layers	1975	[11]
First growth of SiGe epitaxy by MBE	1985	[14]
First growth of SiGe epitaxy by UHV/CVD	1986	[15]
First SiGe HBT	1987	[16]
First ideal SiGe HBT grown by CVD	1989	[17]
First high-performance SiGe HBT	1990	[18, 19]
First self-aligned SiGe HBT	1990	[20]
First SiGe HBT ECL ring oscillator	1990	[20]
First *pnp* SiGe HBT	1990	[21]
First SiGe HBT BiCMOS technology	1992	[22]
First LSI SiGe HBT Integrated Circuit	1993	[23]
First SiGe HBT with peak f_T above 100 GHz	1993	[24, 25]
First SiGe HBT technology in 200-mm manufacturing	1994	[26]
First SiGe HBT technology optimized for 77 K	1994	[27]
First SiGeC HBT	1996	[28]
First high power SiGe HBTs	1996	[29, 30]
First sub-10 psec SiGe HBT ECL circuits	1997	[31]
First SiGe HBT with peak f_T above 200 GHz	2001	[32]

SiGe HBT technology entered commercial production on 200-mm wafers in December of 1994 [26]. The 200-GHz peak f_T barrier was broken in November of 2001 for a non-self-aligned device [32], and for a self-aligned device in February of 2002 [36]. SiGe HBT technologies with f_T above 300 GHz are clearly a realistic goal at this point, making SiGe HBTs quite competitive in performance with competing III-V HBT technologies.

To date, the IC with the highest SiGe HBT device count on a single chip is a 69 × 69 cross-point switch containing greater than 100,000 0.5-μm SiGe HBTs [37]. The highest demonstrated level of SiGe HBT BiCMOS integration to date is a 10.8 × 10.8 mm^2 mixed-signal, single-chip OC-192 10 Gb/s SONET/SDH mapper with integrated serializer/deserializer, clock and data recovery circuits, and synthesis unit, containing 6,000 0.5-μm SiGe HBTs and 1,200,000 CMOS transistors

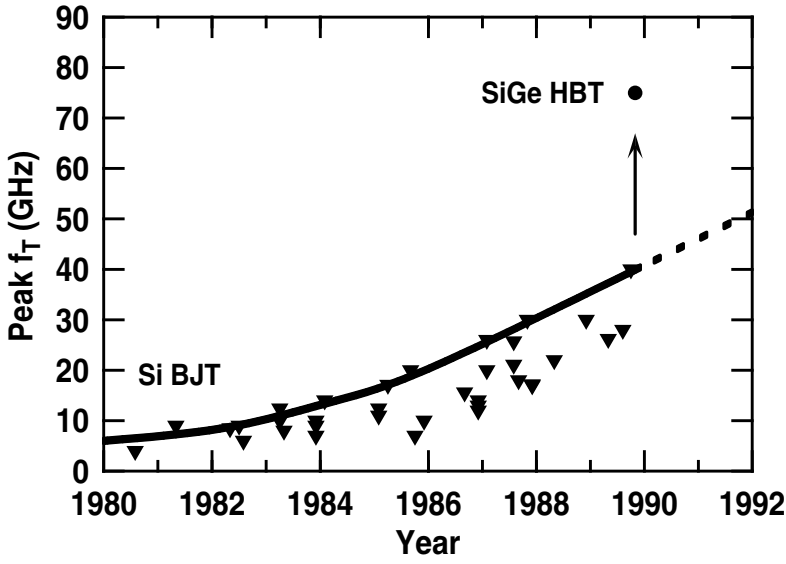

Figure 1.7 Historical trends in published peak cutoff frequency values for various Si BJT technologies compared with the first high-performance SiGe HBT result.

[36, 38].

Not surprisingly, research and development activity in SiGe devices, circuits, and technologies in both industry and at universities worldwide has grown rapidly since the first demonstration of a functional SiGe HBT in 1987. This global interest is nicely reflected in the number of SiGe HBT technical publications in IEEE journals and conferences from 1987 until present, as shown in Figure 1.8.

During the evolutionary path of SiGe HBTs, a large number of SiGe HBT device technologies have been demonstrated at laboratories throughout the world, using a variety of different SiGe epitaxial growth techniques. Commercial SiGe HBT technologies now exist in companies around the world, including: IBM [36], [6] Hitachi [39], Conexant (Jazz) [40], Infineon [41], NEC [42], IHP [43], IMEC [44], TI [45], Philips [46], Lucent [47], ST Microelectronics [48], TEMIC [49], and CNET [50].[7] In recent years, these various SiGe HBT technologies have been leveraged to demonstrate a large number of impressive digital, analog, RF, and microwave circuit results for wireless and wireline communications applications [51]–[100].

[6]For fascinating historical insight into the development of SiGe technology at IBM, see [10].

[7]Only the most recently published version of the SiGe technology from each respective company is given. For the interested reader, each paper contains relevant references to earlier versions of that respective company's SiGe technologies.

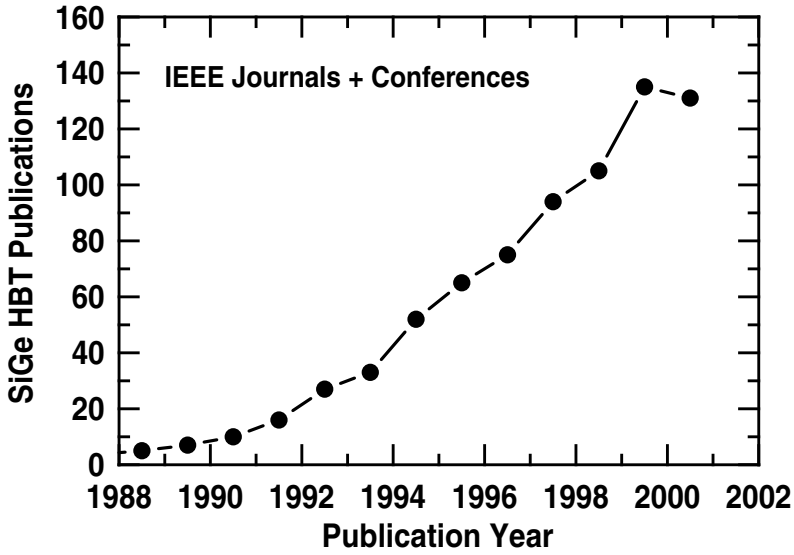

Figure 1.8 Historical trends in the yearly number of SiGe HBT papers published in IEEE journals and conferences (source: IEEE Xplore).

A large number of commercial products using SiGe HBTs are currently on the market, and a foundry service through MOSIS for SiGe HBT BiCMOS technology is available [101], all healthy signs for a new device technology. A variety of review papers on SiGe materials, devices, circuits, and technologies can be found in the literature [102]–[119], and four books (excluding the one you are reading) dealing in one way or another with SiGe materials and devices have been published [120]–[123].

1.7 SiGe HBT Performance Trends

While performance trend charts should always be taken with a grain of salt as to their predictive power, it is nonetheless instructive to examine how SiGe HBT performance has progressed from 1987 until present. For reasons that may or may not meet with approval from all quarters, we have chosen to limit these SiGe HBT trend data in the following manner:

- Consider only results published in the peer-reviewed technical literature.
- Consider either SiGe HBTs or SiGeC HBTs.[8]

[8]A SiGe HBT that has carbon-doping (e.g., less than 0.20% C) in the base to suppress boron

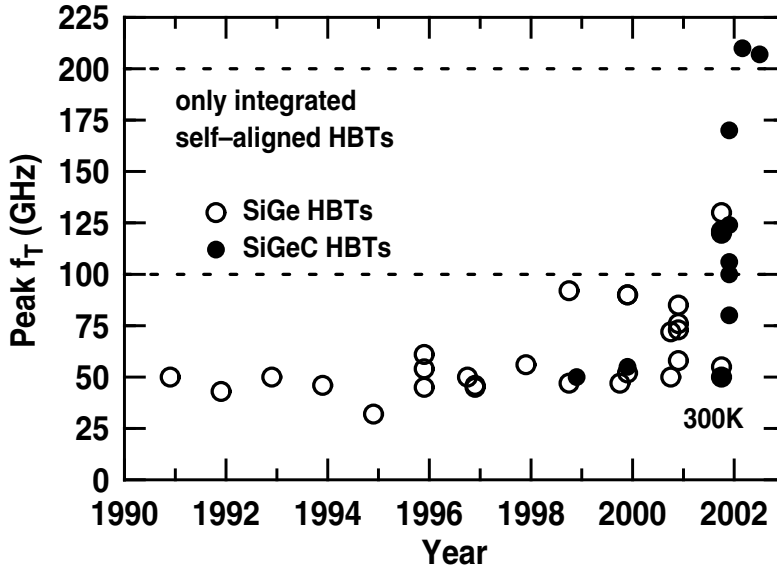

Figure 1.9 Historical trends in peak cutoff frequency for integrated, self-aligned SiGe HBT and SiGeC HBT technologies.

- Consider only self-aligned, fully integrated, Si-processing-compatible SiGe HBT technologies. This eliminates, for instance, non-self-aligned device results that were primarily intended as profile demonstrations. It also eliminates III-V-like mesa-isolated technologies, which cannot be easily integrated with high-transistor-count IC processes, although such device technologies clearly have merit for certain microwave and millimeter wave applications.

- Consider either SiGe HBT or SiGe HBT BiCMOS technologies.

- Consider only room-temperature (300 K) results.

This definition captures greater than 95% of published SiGe HBT results, and limits the trend data to SiGe device technologies that are at least potentially manufacturable and hence in principle commercially viable.

We note that while peak cutoff frequency (f_T) is reasonably straightforward to measure using standard S-parameter techniques (assuming proper calibration and

out-diffusion is properly referred to as a SiGe:C HBT, or simply SiGeC HBT (pronounced "silicon germanium carbon," *not* "silicon germanium carbide"). This class of devices should be viewed as optimized SiGe HBTs, and is distinct from HBTs fabricated using SiGeC alloys with a much higher C content (e.g., 2–3% C) needed to lattice-match SiGeC alloys to Si.

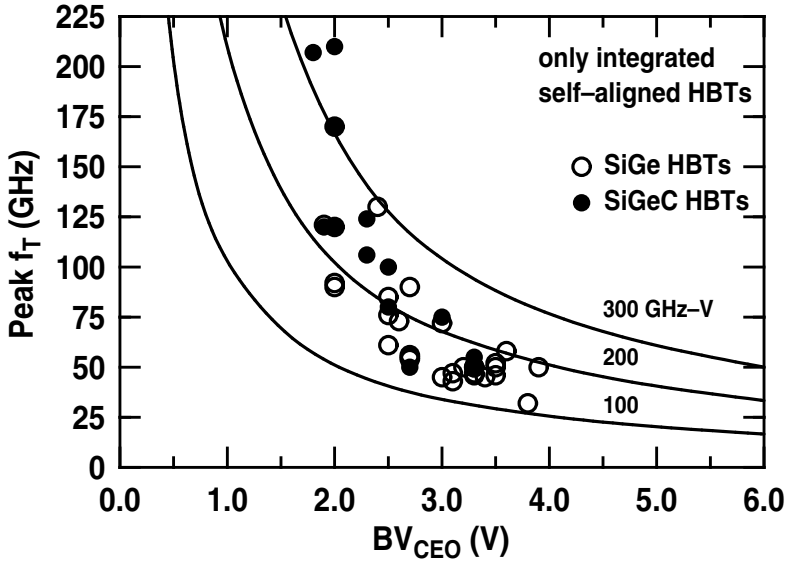

Figure 1.10 Historical trends in peak cutoff frequency as a function of collector-to-emitter breakdown voltage for integrated, self-aligned SiGe HBT and SiGeC HBT technologies.

parasitic de-embedding is performed), the same cannot be said for f_{max}. It has become common practice in the literature to cite f_{max} numbers using unilateral gain (U) extrapolations, which often gives more optimistic numbers than those determined from extrapolations of the maximum available gain (MAG). The f_{max} data presented does not distinguish between the two techniques, and thus adds a level of uncertainty to the f_{max} data presented.

Figure 1.9 shows the historical trends in peak f_T from the first self-aligned device demonstration in December of 1990 [20] until present. It is interesting to note that until about 1998, peak f_T remained in the 50–75 GHz range, suggesting that most research groups were on a profile design and fabrication learning curve, or else attempting to migrate their technologies from research-level demonstrations into commercial IC technologies, and thus worrying less about transistor performance than manufacturability, qualification, and yield. It is interesting to note that both the 100-GHz and 200-GHz f_T barriers were broken within 6 months of each other, 100 GHz being reached in September 2001 by four separate groups [48, 124, 125, 126], and 200 GHz being reached in February of 2002 [36]! [9] It is

[9]Note that non-self-aligned SiGe HBTs with $f_T > 100$ GHz were demonstrated as early as 1993 [24, 25], and a 210-GHz non-self-aligned SiGe HBT was demonstrated in November of 2001 [32].

also clear from Figure 1.9 that after 2001, many groups began migrating towards C-doping of their SiGe HBTs to reduce boron out-diffusion in the base profile, and thereby improve f_T.

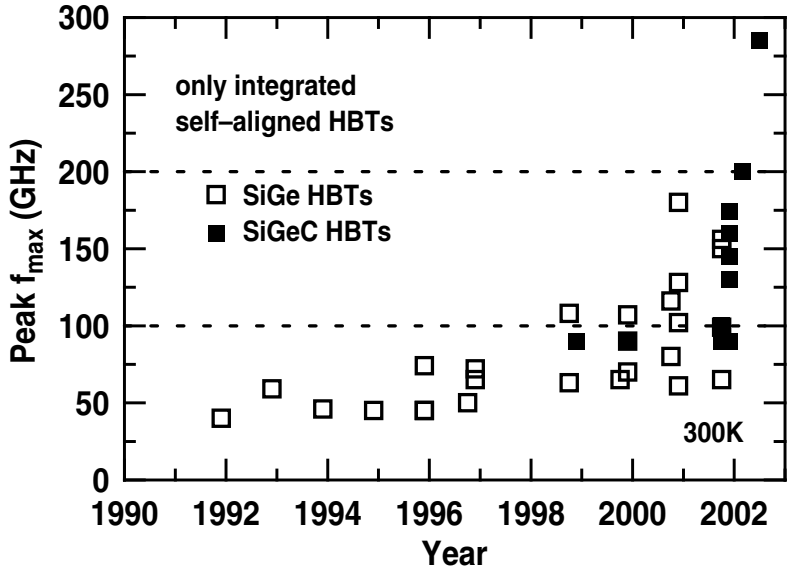

Figure 1.11 Historical trends in peak maximum oscillation frequency for integrated, self-aligned SiGe HBT and SiGeC HBT technologies.

It has been appreciated since 1965 that a fundamental reciprocal relationship exists between transistor peak f_T and BV_{CEO} [127], and the SiGe HBT data qualitatively bear out this trend (Figure 1.10). Most published SiGe HBT results are centered upon an $f_T \times BV_{CEO}$ product of about 200 GHz-V, slightly higher than original "Johnson limit" for Si of 170 GHz-V. More recent results suggest that higher values of the $f_T \times BV_{CEO}$ product are attainable as SiGe device technologies evolve, and have been clustered in the 250–300-GHz-V range over the 1999–2001 time frame. The present record for the $f_T \times BV_{CEO}$ product for a SiGe HBT is 420 GHz-V [36], substantially higher than one might naively expect given past trend data. [10]

Whether this is indirect evidence of an alternative transport mechanism (i.e., ballistic transport) at this level of vertical scaling remains to be seen. It is also worth noting that the current density at which peak f_T is reached has also been steadily rising over time, from about 1.5 mA/μm^2 in 1990 (50 GHz at BV_{CEO} =

[10]Note added in press: the demonstration of a SiGe HBT with 350-GHz peak f_T (BV_{CEO} = 1.4 V), to be presented at the IEDM in December of 2002, increases this number to 490 GHz-V [128]!

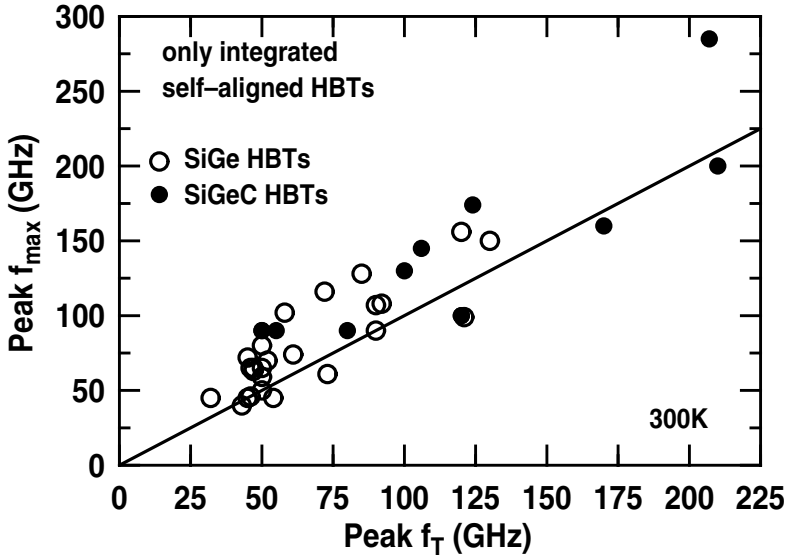

Figure 1.12 Historical trends in peak maximum oscillation frequency as a function of peak cutoff frequency for integrated, self-aligned SiGe HBT and SiGeC HBT technologies.

3.2 V) to about 10.0 mA/μm^2 in 2002 (210 GHz at BV_{CEO} = 2.0 V). This J_C rise over time clearly presents a host of challenges in terms of device reliability and technology metalization needs, not to mention system-level, voltage-compression design constraints induced by the ever-shrinking BV_{CEO}.

While peak f_T is very useful as a technology figure-of-merit, f_{max} is a more relevant circuit-level performance metric. Figure 1.11 shows that the SiGe HBT peak f_{max} data has risen over time is a similar manner to that of peak f_T, as might be naively expected. In 2002, best-of-breed SiGe HBTs have attained peak f_{max} in the 200-GHz range (the present record being 285 GHz [129]), quite impressive even by III-V HBT standards. Figure 1.12 shows that for most SiGe technologies, peak f_{max} is generally comparable to or even exceeds peak f_T.[11] Achieving comparable f_T and f_{max} is highly desirable for many types of high-speed circuit applications, and this trend for SiGe is different from that seen in most traditional III-V HBT designs, where f_{max} greatly exceeds f_T. This ability of SiGe HBTs to simultaneously maintain both high f_T and high f_{max} is a direct result of the inherently low-parasitic nature of highly-scaled, self-aligned Si device structures, and is a decided advantage from an application standpoint.

[11] Please see the above cautionary note concerning the interpretation of f_{max} data extrapolations.

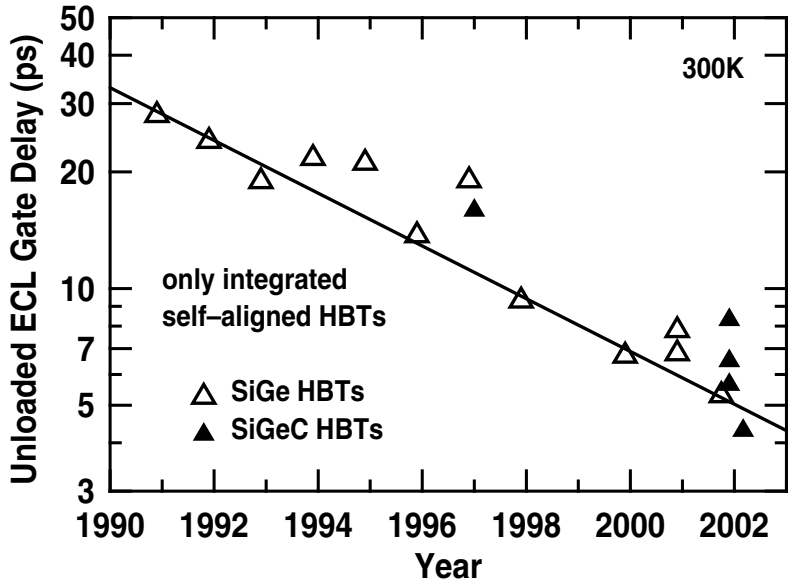

Figure 1.13 Historical trends in unloaded ECL gate delay for integrated, self-aligned SiGe HBT and SiGeC HBT technologies.

Unloaded ECL ring oscillator gate delay has historically been used as a "one-step-better" technology performance metric, since it is easy to implement and is the simplest "real" circuit demonstration vehicle, with a delay that depends strongly on both f_T as well as the resistive and capacitive device parasitics. Figure 1.13 shows the SiGe HBT ECL gate delay trend data. The fact that this data is roughly following a linear decrease on log-linear scales indicates that this performance data is following a classical Moore's Law exponential growth pattern. The long-standing 10 psec ECL delay barrier was broken in December of 1997 [31], and has since marched steadily downward to the present record of 4.3 psec [36].

1.8 The IC Technology Battleground: Si Versus SiGe Versus III-V

And the winner is? From the very beginning it has been, and remains to this day, a highly contentious issue as to whether SiGe technology will be able to successfully position itself to dominate existing and future IC market sectors across a broad array of application fronts. Even broad-brushed comparisons of the relative merits of the competing IC device technologies can be perilous, given that there is no such

Figure 1.14 Headline News: "SiGe 'Pac-Man' gobbles up GaAs competition!" (Used with the permission of Michael W. Davidson, Florida State University.) This SiGe Pac-Man was found on a SiGe RFIC designed by TEMIC Semiconductors. Pac-Man was originally designed by Toru Iwantani and programmed by Hideyuki Mokajima and his associates. The name Pac-Man is derived from the Japanese slang "Paku-paku," which means "to eat." Originally, the Japanese named the game "Puckman," but it was changed to "Pac-Man" upon launching in the United States. Pac-Man is the best-selling video game in history.

thing as a true "apples-to-apples" comparison, and one will inevitably be accused of a personal bias of this or that sort, or be charged with comparing one inferior example of a given technology with a superior example of a competing technology, thus artificially skewing the result. In addition, the potential circuit applications of any given technology are so diverse, some favoring one performance metric, others favoring another performance metric, that sweeping generalizations are simply impossible, and the reader should be wary when they are attempted. Given this disclaimer, however, it is nonetheless instructive in this context to make some general comparisons of the performance metrics that might be encountered, for instance, while designing an radio-frequency integrated circuit (RFIC) using each of the various device topologies (Table 1.2).

From an RF viewpoint, state of the art SiGe HBTs offer frequency response, noise figure, and linearity comparable to current-generation III-V devices, and better than both Si BJTs and Si CMOS (even highly scaled CMOS). SiGe HBTs offer better low-frequency (1/f) and phase noise than all of the competition, with the

Table 1.2 Relative Performance Comparisons of Various Device Technologies for RFICs (Excellent: ++; Very good: +; Good: 0; Fair: −; Poor: −−)

Performance Metric	SiGe HBT	Si BJT	Si CMOS	III-V MESFET	III-V HBT	III-V HEMT
Frequency response	+	0	0	+	++	++
1/f and phase noise	++	+	−	−−	0	−−
Broadband Noise	+	0	0	+	+	++
Linearity	+	+	+	++	+	++
Output conductance	++	+	−	−	++	−
Tranconductance/area	++	++	−−	−	++	−
Power dissipation	++	+	−	−	+	0
CMOS integration	++	++	N/A	−−	−−	−−
IC cost	0	0	+	−	−	−−

possible exception of Si BJTs. Being a bipolar transistor, the transconductance per unit area of a SiGe HBT is much higher than for either Si or III-V FETs, and for profile designs with a graded Ge base, the output conductance of a SiGe HBT is also superior to either Si or III-V FETs. SiGe HBTs also have the beneficial feature that their broadband noise is minimized at very low current densities (typically 10× lower than peak f_T), in direct contrast to FETs (Si or III-V), making them very attractive from a power dissipation point of view for portable applications.[12] III-V devices, especially high-electron-mobility transistors (HEMTs), will continue to provide the very best noise performance, albeit at a higher cost, and given their larger bandgaps and hence breakdown voltages, III-V devices will make the best power devices. The long-term advantage of SiGe HBTs over the competition is a strong function of system-level integration and cost. That is, the ability of SiGe HBTs to integrate easily with conventional CMOS distinguishes them fundamentally from all III-V technologies. SiGe technology is essentially equivalent to Si technology in that sense, and enjoys all of the advantages associated with the economy of scale of Si IC manufacturing, including yield and die cost.

Which device technology is likely to walk away from the IC technology battleground? The SiGe advocates obviously embrace the notion depicted in Figure 1.14, in which SiGe swallows the GaAs competition. Wishful thinking? The III-V ad-

[12]That the SiGe HBT exhibits far lower power dissipation than CMOS at fixed RF noise figure is wonderfully ironic, given that the large power dissipation associated with ECL is ultimately what doomed Si BJT technology to CMOS domination in the digital world. Sweet revenge!

vocates clearly see no need for a SiGe upstart! As a interested spectator on the device technology battleground, however, it has been mildly amusing to witness the III-V camp at technical conferences slowly but surely change from a "SiGe is no threat at all, please go away" mentality in the early to mid-1990s, to a grudging but gradual acceptance of SiGe as a serious contender in the late 1990s, to a recent near-paranoia of being supplanted and marginalized by SiGe technology. SiGe technology is indeed evolving rapidly, and given its happy marriage to conventional high-volume, low-cost Si fabrication (and CMOS), it does embody the best of both the III-V and Si worlds, a decided advantage. The CMOS advocates, of course, confidently and zealously maintain that it is simply inevitable that scaled CMOS will "conquer the world," leading to little or no need for either SiGe or III-V devices. Perhaps. While the CMOS tsunami did in fact effectively gobble up Si BJT-based ECL in the high-end server market in the early 1990s, and now dominates the digital microprocessor world, the wireless and high-data-rate wireline domains are another matter entirely, and place far more stringent demands on the devices than simple digital logic.

To this question of long-term market dominance, there is simply no easy answer: only time will tell. In the end, the outcome is likely to be the obvious one: SiGe, III-V, and CMOS will all be around 10 years from now, and each will continue to hold important market share for the foreseeable future, in sectors that value their respective strengths. Clearly CMOS will continue to dominate the digital world, and will grow in importance in the low-end wireless sector. SiGe will make steady inroads into a broad array of both wireless and wireline markets, particularly as frequency bands and data rates continue to rise. III-V technologies will continue to dominate the small but important microwave market and the RF power amplifier market. Given the fact that the global electronics market is already enormous ($1,128,000,000,000 in 2000 [1]), and growing rapidly with no real end in sight, holding even a small niche market is likely to be sufficient to provide long-term sustenance for a variety of device technologies.

The worldwide interest in SiGe as a commercial IC technology is growing rapidly. Very rapidly. For those with lingering doubts as to the beautiful efficacy of the SiGe solution for a wide variety of twenty-first IC needs, it should be noted that there are virtually no companies in the world with a vested interest in communications ICs that do not at present have SiGe technology either in production or under development, or at least in use via foundry services. That message in itself is instructive, and should be considered carefully by SiGe pundits. With this requisite background, we now dive into the deep, rich, and fascinating subject of SiGe HBTs. Enjoy!

References

[1] Source: Cahners Instat Group (www.instat.com), 2001.

[2] L. Pauling and R. Hayward, *The Architecture of Molecules*, W.H. Freeman and Company, New York, 1964.

[3] M. Nakamura, "Challenges in semiconductor technology for multi-megabit network services," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 16-18, 1998.

[4] F. Capasso, "Band-gap engineering: from physics and materials to new semiconductor devices," *Science*, vol. 235, pp. 172-176, 1987.

[5] J.W. Matthews and A.E. Blakeslee, "Defects in epitaxial multilayers– I: misfit dislocations in layers," *J. Cryst. Growth*, vol. 27, pp. 118-125, 1974.

[6] J.W. Matthews and A.E. Blakeslee, "Defects in epitaxial multilayers– II: dislocation pile-ups, threading dislocations, slip lines and cracks," *J. Cryst. Growth*, vol. 32, pp. 265-273, 1975.

[7] W. Shockley, U.S. Patent 2,569,347, issued 1951.

[8] H. Kroemer, "Zur theorie des diffusions und des drifttransistors, part III," *Arch. Elektr. Ubertragung*, vol. 8, pp. 499-504, 1954.

[9] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proc. IRE*, vol. 45, pp. 1535-1537, 1957.

[10] D.L. Harame and B.S. Meyerson, "The early history of IBM's SiGe mixed signal technology," *IEEE Trans. Elect. Dev.*, vol. 48, pp. 2555-2567, 2001.

[11] E. Kasper, H.J. Herzog, and H. Kibbel, "A one-dimensional SiGe superlattice frown by UHV epitaxy," *J. Appl. Phys.*, vol. 8, pp. 1541-1548, 1975.

[12] H. Kroemer, "Heterostructure bipolar transistors and integrated circuits," *Proc. IEEE*, vol. 70, pp. 13-25, 1982.

[13] H. Kroemer, "Heterostructure bipolar transistors: what should we build?," *J. Vacuum Sci. Tech.: B1*, vol. 2, pp. 112-130, 1983.

[14] R. People, "Indirect bandgap of coherently strained $Si_{1-x}Ge_x$ bulk alloys on <001> silicon substrates," *Physical Review B*, vol. 32, pp. 1405-1408, 1985.

[15] B.S. Meyerson, "Low-temperature silicon epitaxy by ultrahigh vacuum / chemical vapor deposition," *Appl. Phys. Lett.*, vol. 48, pp. 797-799, 1986.

[16] S.S. Iyer et al., "Silicon-germanium base heterojunction bipolar transistors by molecular beam epitaxy," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 874-876, 1987.

[17] C.A. King et al., "Si/Si_{1-x}/Ge_x heterojunction bipolar transistors produced by limited reaction processing," *IEEE Elect. Dev. Lett.*, vol. 10, pp. 52-54, 1989.

[18] G.L. Patton et al., "63-75 GHz f_T SiGe-base heterojunction-bipolar technology," *Tech. Dig. IEEE Symp. VLSI Tech.*, pp. 49-50, 1990.

[19] G.L. Patton et al., "75 GHz f_T SiGe base heterojunction bipolar transistors," *IEEE Elect. Dev. Lett.*, vol. 11, pp. 171-173, 1990.

[20] J.H. Comfort et al., "Profile leverage in a self-aligned epitaxial Si or SiGe-base bipolar technology," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 21-24, 1990.

[21] D.L. Harame et al., "SiGe-base PNP transistors fabrication with n-type UHV/CVD LTE in a "NO DT" process," *Tech. Dig. IEEE Symp. VLSI Tech.*, pp. 47-48, 1990.

[22] D.L. Harame et al., "A high-performance epitaxial SiGe-base ECL BiCMOS technology," *Tech. Dig IEEE Int. Elect. Dev. Meeting*, pp. 19-22, 1992.

[23] D.L. Harame et al., "Optimization of SiGe HBT technology for high speed analog and mixed-signal applications," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 71-74, 1993.

[24] E. Kasper, A. Gruhle, and H. Kibbel, "High speed SiGe-HBT with very low base sheet resistivity," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 79-81, 1993.

[25] E.F. Crabbé et al., "Vertical profile optimization of very high frequency epitaxial Si- and SiGe-base bipolar transistors," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 83-86, 1993.

[26] D.L. Harame et al., "A 200 mm SiGe-HBT technology for wireless and mixed-signal applications," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 437-440, 1994.

[27] J.D. Cressler et al., "An epitaxial emitter cap SiGe-base bipolar technology for liquid nitrogen temperature operation," *IEEE Elect. Dev. Lett.*, vol. 15, pp. 472-474, 1994.

[28] L. Lanzerotti et al., "Si/$Si_{1-x-y}Ge_xC_y$/Si heterojunction bipolar transistors," *IEEE Elect. Dev. Lett.*, vol. 17, pp. 334-337, 1996.

[29] A. Schüppen et al., "1-W SiGe power HBTs for mobile communications," *IEEE Micro. Guided Wave Lett.*, vol. 6, pp. 341-343, 1996.

[30] P.A. Potyraj et al., "A 230-Watt S-band SiGe heterojunction junction bipolar transistor," *IEEE Trans. Micro. Theory Tech.*, vol. 44, pp. 2392-2397, 1996.

[31] K. Washio et al., "A selective-epitaxial SiGe HBT with SMI electrodes featuring 9.3-ps ECL-Gate Delay," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 795-798, 1997.

[32] S.J. Jeng et al., "A 210-GHz f_T SiGe HBT with non-self-aligned structure," *IEEE Elect. Dev. Lett.*, vol. 22, pp. 542-544, 2001.

[33] G.L. Patton et al., "Silicon-germanium-base heterojunction bipolar transistors by molecular beam epitaxy," *IEEE Elect. Dev. Lett.*, vol. 9, pp. 165-167, 1988.

[34] H. Temkin et al., "Ge_xSi_{1-x} strained-layer heterostructure bipolar transistors," *Appl. Phys. Lett.*, vol. 52, pp. 1089-1091, 1988.

[35] D.-X. Xu et al., "$n-Si/p-Si_{1-x}Ge_x/n-Si$ double-heterojunction bipolar transistors," *Appl. Phys. Lett.*, vol. 52, pp. 2239-2241, 1988.

[36] A.J. Joseph et al., "0.13 μm 210 GHz f_T SiGe HBTs - expanding the horizons of SiGe BiCMOS," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 180-182, 2002.

[37] AMCC S2090 preliminary data (www.amcc.com), 2002.

[38] AMCC S4803 preliminary data (www.amcc.com), 2002.

[39] K. Oda et al., "Self-aligned selective-epitaxial-growth $Si_{1-x-y}Ge_xC_y$ HBT technology featuring 170-GHz f_{max}," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 332-335, 2001.

[40] M. Racanelli et al., "Ultra high speed SiGe npn for advanced BiCMOS Technology," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 336-339, 2001.

[41] J. Böck et al., "High-speed SiGe:C bipolar technology," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 344-347, 2001.

[42] T. Hashimoto et al., "A 73 GHz f_T 0.18μm RF-SiGe BiCMOS technology considering thermal budget trade-off and with reduced boron-spike effect on HBT characteristics," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 149-152, 2000.

[43] B. Heinemann et al., "Cost-effective high-performance high-voltage SiGe:C HBTs with 100 GHz f_T and $BV_{CEO} \times f_T$ products exceeding 220 VGHz," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 348-352, 2001.

[44] S. Decoutere et al., "A 0.35 μm SiGe BiCMOS process featuring a 80 GHz f_{max} HBT and integrated high-Q RF passive components," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 106-109, 2000.

[45] F.S. Johnson et al., "A highly manufacturable 0.25 μm RF technology utilizing a unique SiGe integration," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 56-59, 2001.

[46] P. Deixler et al., "Explorations for high performance SiGe-heterojunction bipolar transistor integration," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 30-33, 2001.

[47] M. Carroll et al., "COM2 SiGe Modular BiCMOS technology for digital, mixed-signal, and RF applications, *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 145-148, 2000.

[48] H. Baudry et al., "High performance 0.25 μm SiGe and SiGe:C HBTs using non-selective epitaxy," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 52-55, 2001.

[49] A. Schüppen et al., "SiGe-technology and components for mobile communication systems," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 130-133, 1996.

[50] A. Chantre et al., "A high performance low complexity SiGe HBT for BiCMOS integration," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 93-96, 1998.

[51] M. Soda et al., "Si-analog ICs for 20 Gb/s optical receiver," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 170-171, 1994.

[52] T. Hashimoto et al., "SiGe bipolar ICs for a 20Gb/s optical transmitter," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 167-170, 1994.

[53] F. Sato et al., "Sub-20ps ECL circuits with high-performance super self-aligned selectively grown SiGe base (SSSB) bipolar transistors," *IEEE Trans. Elect. Dev.*, vol. 42, pp. 483-488, 1995.

[54] M. Case et al., "A 23 GHz static 1/128 frequency divider implemented in a manufacturable Si/SiGe HBT process," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 121-123, 1995.

[55] W. Gao et al., "A 5-GHz SiGe HBT return-to-zero comparator," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 166-169, 1995.

[56] J. Glenn et al., "12-GHz Gilbert mixers using a manufacturable Si/SiGe epitaxial-base bipolar technology," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 186-189, 1995.

[57] H. Schumacher et al., "A 3 V supply voltage, DC-18 GHz SiGe HBT wideband amplifier," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 190-193, 1995.

[58] A. Gruhle et al., "Monolithic 26 GHz and 40 GHZ VCO's with SiGe heterojunction bipolar transistors," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 725-728, 1995.

[59] F. Sato et al., "The optical terminal IC: a 2.4 Gb/s receiver and a 1:16 demultiplexer in one chip," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 162-165, 1995.

[60] L. Larson et al., "Si/SiGe HBT technology for low-cost monolithic microwave integrated circuits," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 80-81, 1996.

[61] F. Sato et al., "A 2.4Gb/s receiver and a 1:16 demultiplexer in one chip using a super self-aligned selectively grown SiGe base (SSSB) bipolar transistor," *IEEE J. Solid-State Circ.*, vol. 31, 1451-1457, 1996.

[62] M. Soyuer et al., "An 11 GHz 3V SiGe voltage-controlled oscillator with integrated resonator," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 169-172, 1996.

[63] R. Götzfried et al., "Zero power consumption Si/SiGe HBT SPDT T/R antenna switch," *Tech. Dig. IEEE MTT-S Int. Micro. Symp.*, pp. 651-653, 1996.

[64] J. Long et al., "RF analog and digital circuits in SiGe technology," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 82-83, 1996.

[65] W. Gao, W. Snelgrove, and S. Kovacic, "A 5-GHz SiGe HBT return-to-zero comparator for RF A/D conversion," *IEEE J. Solid-State Circ.*, vol. 31, pp. 1502-1506, 1996.

[66] M. Soda et al., "A 1Gb/s 8-channel array OEIC with SiGe photodetector," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 120-121, 1997.

[67] M. Wurzer et al., "42 GHz static frequency divider in Si/SiGe bipolar technology," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 122-123, 1997.

[68] P. Xiao et al., "A 4b 8GSample/s A/D converter in SiGe bipolar technology," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 124-125, 1997.

[69] M. Soyuer et al., "A 5.8 GHz 1 V low-noise amplifier in SiGe bipolar technology," *Proc. IEEE MTT RFIC Symp.*, pp. 19-22, 1997.

[70] H. Ainspan et al., "A 6.25 GHz low DC power low-noise amplifier," *Proc. IEEE Cust. Int. Circ. Conf.*, pp. 177-180, 1997.

[71] S. Shiori et al., "A 10Gb/s SiGe bipolar framer/demultiplexer for SDH system," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 202-203, 1998.

[72] T. Masuda et al., "40Gb/s analog IC chipset for optical receiver using SiGe HBTs, *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 314-315, 1998.

[73] J. Sevenhans et al., "Silicon germanium and silicon bipolar RF circuits for 2.7 V single chip radio transceiver integration," *Proc. IEEE Cust. Int. Circ. Conf.*, pp. 409-412, 1998.

[74] J.-O. Plouchart et al., "A 5.2 GHz 3.3 V I/Q SiGe RF transceiver," *Proc. IEEE Cust. Int. Circ. Conf.*, pp. 217-220, 1999.

[75] Y.M. Greshishchev and P. Schvan, "A 60 dB gain 55 dB dynamic range 10 Gb/s broadband SiGe HBT limiting amplifier," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 382-383, 1999.

[76] M. Bopp et al., "A DECT transceiver chip set using SiGe technology," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 68-69, 1999.

[77] T. Morikawa et al., "A SiGe single chip 3.3 V receiver IC for 10Gb/s optical communication systems," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 380-381, 1999.

[78] D. Goren et al., "2.5 GHz pin electronics SiGe driver for IC test equipment," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 31-33, 1999.

[79] P.D. Tseng et al., "A monolithic SiGe power amplifier for dual-mode (CDMA/AMPS) cellular handset applications," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 153-156, 1999.

[80] G. Grau et al., "A current-folded up-conversion mixer and VCO with center-tapped inductor in a SiGe-HBT technology for 5 GHz wireless LAN applications," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 161-164, 1999.

[81] H. Ainspan and M. Soyuer, "A fully-integrated 5-GHz frequency synthesizer in SiGe BiCMOS," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 165-168, 1999.

[82] Y.M. Greshishchev and P. Schvan, "SiGe clock and data recovery IC with linear type PLL for 10 Gb/s SONET application," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 169-172, 1999.

[83] Y.M. Greshishchev et al., "A fully integrated SiGe receiver IC for 10Gb/s data rate," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 52-53, 2000.

[84] M. Meghelli et al., "SiGe BiCMOS 3.3 V clock and data recovery circuits for 10 Gb/s serial transmission systems," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 56-57, 2000.

[85] T. Masuda et al., "45 GHz transimpedance 32 dB limiting amplifier and 40 Gb/s 1:4 high-sensitivity demultiplexer with decision circuit using SiGe HBTs for 40 Gb/s optical receiver," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 60-61, 2000.

[86] H. Pretl et al., "A SiGe-bipolar down-conversion mixer for a UTMS zero-IF receiver," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 40-43, 2000.

[87] M. Roβberg et al., "11 GHz SiGe circuits for ultra wideband radar," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 70-73, 2000.

[88] K. Ettinger et al., "An integrated 20 GHz SiGe bipolar differential oscillator," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 161-164, 2000.

[89] H.-I. Cong et al., "A 10 Gb/s 16:1 multiplexer and 10 GHz clock synthesizer in 0.25 μm SiGe BiCMOS," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 80-81, 2001.

[90] S. Ueno et al., "A single-chip 10 Gb/s transceiver LSI using SiGe SOI/BiCMOS," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 82-83, 2001.

[91] M. Reinhold et al., "A fully-integrated 40Gb/s clock and data recovery / 1:4 DEMUX IC in SiGe technology," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 84-86, 2001.

[92] D. Belot et al., "A DCS1800/GSM900 RF to digital fully integrated receiver in SiGe 0.35 *mu*m BiCMOS," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 86-89, 2001.

[93] G. Babcock, "SiGe-HBT active receiver mixers for base-station applications," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 90-93, 2001.

[94] D. Johnson and S. Raman, "A packaged SiGe x2 sub-harmonic mixer for U-NII band applications," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 159-162, 2001.

[95] S. Hackl et al., "A 45 GHz SiGe active frequency multiplier," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 82-83, 2002.

[96] T. Masuda et al., "Single-chip 5.8 GHz ETC transceiver IC with PLL and demodulation circuits using SiGe HBT/CMOS," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 96-97, 2002.

[97] T. Masuda et al., "Single-chip 5.8 GHz ETC transceiver IC with PLL and demodulation circuits using SiGe HBT/CMOS," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 96-97, 2002.

[98] V. Aparin et al., "A highly-integrated tri-band/quad-mode SiGe BiCMOS TF-to-baseband receiver for wireless CDMA/WCDMA/AMPS applications with GPS capability," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 234-235, 2002.

[99] A. Bellaouar et al., "A highly-integrated SiGe BiCMOS WCDMA transmitter IC," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 238-239, 2002.

[100] M. Meghelli, A. Rylyakov, and L. Shan, "50 Gb/s SiGe BiCMOS 4:1 multiplexer and 1:4 demultiplexer for serial communication systems," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 260-261, 2002.

[101] MOSIS IC Foundry Service (www.mosis.org).

[102] R. People, "Physics and applications of Ge_xSi_{1-x}/Si strained layer heterostructures," *IEEE J. Quan. Elect.*, vol. 22, p. 1696-1710, 1986.

[103] S.S. Iyer et al., "Heterojunction bipolar transistors using Si-Ge alloys," *IEEE Trans. Elect. Dev.*, vol. 36, pp. 2043-2064, 1989.

[104] G.L. Patton et al., "SiGe-base heterojunction bipolar transistors: physics and design issues," *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, pp. 13-16, 1990.

[105] B. Meyerson, "UHV/CVD growth of Si and SiGe alloys: chemistry, physics, and device applications," *Proc. IEEE*, vol. 80, p. 1592-1608, 1992.

[106] J.D. Cressler et al., "Silicon-germanium heterojunction bipolar technology: the next leap in silicon?," *Tech. Dig. IEEE Int. Solid-State Circ. Conf.*, pp. 24-27, 1994.

[107] C. Kermarrec et al., "SiGe HBTs reach the microwave and millimeter-wave frontier," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 155-162, 1994.

[108] J.D. Cressler, "Re-engineering silicon: Si-Ge heterojunction bipolar technology," *IEEE Spectrum*, pp. 49-55, 1995.

[109] D.L. Harame et al., "Si/SiGe epitaxial-base transistors: part I - materials, physics, and circuits," *IEEE Trans. Elect. Dev.*, vol. 40, pp. 455-468, 1995.

[110] D.L. Harame et al., "Si/SiGe epitaxial-base transistors: part II - process integration and analog applications," *IEEE Trans. Elect. Dev.*, vol. 40, pp. 469-482, 1995.

[111] D.L. Harame, "High-performance BiCMOS process integration: trends, issues, and future directions," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meet.*, pp. 36-43, 1997.

[112] J.D. Cressler, "SiGe HBT technology: a new contender for Si-based RF and microwave circuit applications," *IEEE Trans. Micro. Theory Tech.*, vol. 46, pp. 572-589, 1998.

[113] L.E. Larson, "High-speed Si/SiGe technology for next-generation wireless system applications," *J. Vacuum Sci. Tech. B*, vol. 16, pp. 1541-1548, 1998.

[114] L.E. Larson, "Integrated circuit technology options for RFIC's - present status and future directions," *IEEE J. Solid-State Circ.*, vol. 33, pp. 387-399, 1998.

[115] H.J. Osten et al., "Carbon-doped SiGe heterojunction bipolar transistors for high-frequency applications," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 109-116, 1999.

[116] B.S. Meyerson, "Silicon:germanium-based mixed-signal technology for optimization of wired and wireless telecommunications," *IBM J. Res. Dev.*, vol. 44, pp. 391-407, 2000.

[117] A. Gruhle, "Prospects for 200 GHz on silicon with SiGe heterojunction bipolar transistors," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 19-25, 2001.

[118] D.L. Harame et al., "Current status and future trends of SiGe BiCMOS technology," *IEEE Trans. Elect. Dev.*, vol. 48, pp. 2575-2594, 2001.

[119] D.L. Harame and J.D Cressler, "The SiGe HBT: A Primer," *IEEE Trans. Elect. Dev.*, in press, 2002.

[120] E. Kaspar, editor, *Properties of Strained and Relaxed Silicon Germanium*, EMIS Datareviews Series No. 12, INSPEC, London, 1995.

[121] J.S. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors*, John Wiley and Sons, Inc., New York, 1999.

[122] C.K. Maiti and G.A. Armstrong, *Applications of Silicon-Germanium Heterostructure Devices*, Institute of Physics Publishing, London, 2001.

[123] C.K. Maiti, N.B. Chakrabarti, and S.K. Ray, *Strained Silicon Heterostructures: Materials and Devices*, The Institution of Electrical Engineers, London, 2001.

[124] E. Ohue et al., "5.3-ps ECL and 71-GHz static frequency divider in self-aligned SEG SiGe HBT," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 26-29, 2001.

[125] A. Joseph et al., "A 0.18 μm BiCMOS technology featuring 120/100 GHz (f_T/f_{max}) HBT and ASIC-compatible CMOS using copper interconnect," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 143-146, 2001.

[126] K. Schuegraf et al., "0.18 μm SiGe BiCMOS technology for wireless and 40 Gb/s communication products," *Proc. IEEE Bipolar/BiCMOS Circ. Tech. Meeting*, pp. 147-150, 2001.

[127] E.O. Johnson, "Physical limitations on frequency and power parameters of transistors," *RCA Review*, pp. 163-177, 1965.

[128] J.-S. Rieh et al., "SiGe HBTs with Cut-off Frequency of 350 GHz," to appear in the *Tech. Dig. IEEE Int. Elect. Dev. Meeting*, December 2002.

[129] B. Jagannathan et al., "Self-aligned SiGe NPN transistors with 285 GHz f_{max} and 207 GHz f_T in a manufacturable technology," *IEEE Elect. Dev. Lett.*, vol. 23, pp. 258-260, 2002.

Chapter 2

SiGe Strained-Layer Epitaxy

It is ironic, in the present context, that the first functional transistors were in fact fabricated from Ge. Little time elapsed until the recognition that Si would prove to be a much better commercial platform for the emerging transistor field than Ge, and except for a few niche applications, Ge dropped out of vogue and was soon forgotten as a viable device material. Interestingly, however, it was appreciated very early in the game that the appropriate combination of Si and Ge, being chemically compatible semiconductors with differing bandgaps, would present interesting device engineering opportunities. Unfortunately, it took nearly 30 years to reduce that idea to the practical reality of device-quality SiGe strained-layer epitaxy. In this chapter we examine the creation of strained-layer epitaxy from Si and Ge, and explore the stability constraints that the SiGe world is governed by. We then address the resultant band structure and transport parameters of SiGe alloys, followed by a brief discourse on remaining open issues that merit further attention.

2.1 SiGe Alloys

Si and Ge are both Group IV elemental semiconductors, and crystallize in the diamond lattice structure, as depicted in Figure 2.1. For a comprehensive table of the bulk structural, mechanical, and electrical properties of both Si and Ge, refer to the Appendix. Si and Ge are completely miscible over their entire compositional range, giving rise to chemically stable SiGe alloys that preserve their parent diamond crystal structure and that have a linearly interpolated lattice constant given to first order by Vegard's rule,

$$a(Si_{1-x}Ge_x) = a_{Si} + x\,(a_{Ge} - a_{Si}), \tag{2.1}$$

where a is the lattice constant, and x is the Ge fraction. Diffraction measurements

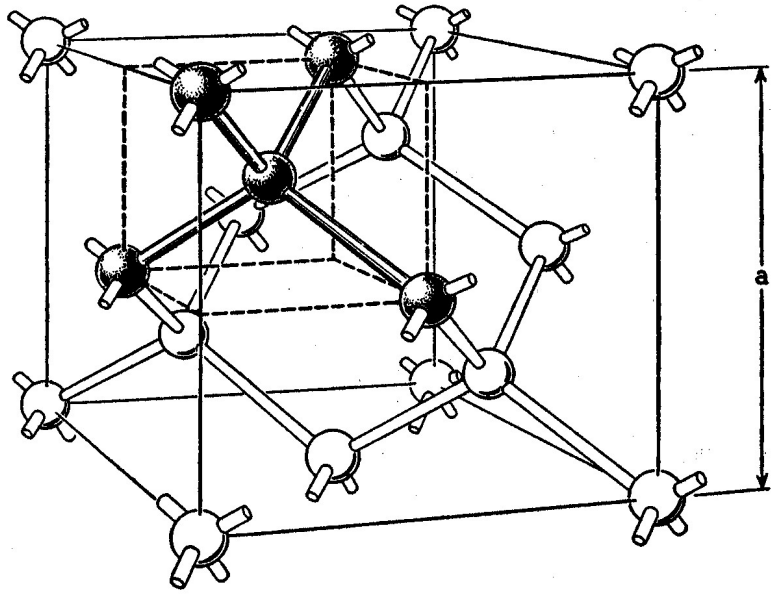

Figure 2.1 Unit cell of the diamond lattice (after [1]).

of actual SiGe films show minor departures of this linear dependence and can be fit by a parabolic relationship of the form

$$a(Si_{1-x}Ge_x) = 0.002733\,x^2 + 0.01992\,x + 0.5431\ (nm), \tag{2.2}$$

as depicted in Figure 2.2 [2].

2.1.1 Pseudomorphic Growth and Film Relaxation

The lattice mismatch between pure Si (a = 5.431 Å) and pure Ge (a = 5.658 Å) is 4.17% at 300 K, and increases only slightly with increasing temperature. When SiGe epitaxy [1] is grown (actually it is more properly said to be deposited) onto a thick Si substrate host, this inherent lattice mismatch between the SiGe film and the underlying Si substrate can be accommodated in only one of two ways.

First, the lattice of the deposited SiGe alloy distorts in such a way that it adopts the underlying Si lattice constant, resulting in perfect crystallinity across the growth interface. In essence, the SiGe film is forced to adopt its host's smaller

[1]The word "epitaxy" is derived from the Greek word *epi*, meaning "upon" or "over."

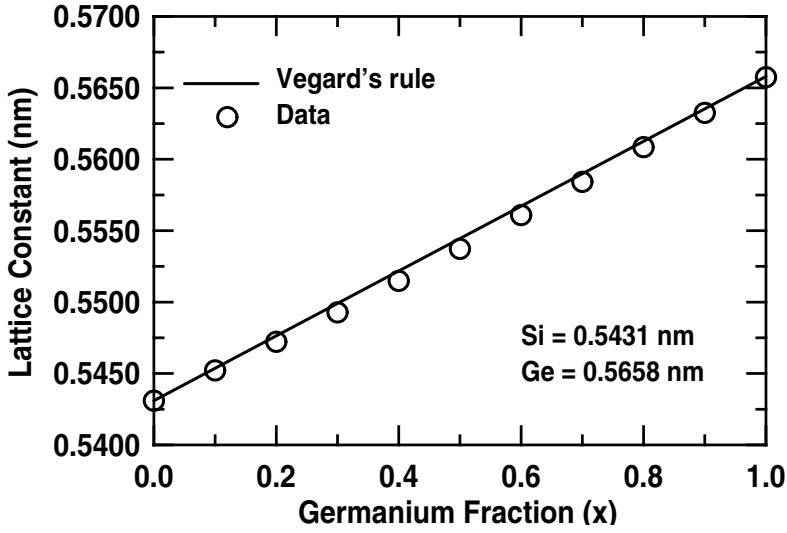

Figure 2.2 Theoretical and experimental lattice constant of a $Si_{1-x}Ge_x$ alloy as a function of Ge fraction.

lattice constant. This scenario is known as "pseudomorphic" [2] growth, and is the desired result for most device applications. Under processing conditions that favor pseudomorphic growth, the SiGe film is forced into biaxial (in-plane) compression. In this case, the SiGe lattice constant in the growth plane is determined by the Si substrate, and the result is a tetragonal distortion (extension) of the normally cubic SiGe crystal in the orthogonal direction, in accordance with the Poisson ratio. The SiGe alloy is now under strain, and "SiGe strained-layer epitaxy" results. Because of the additional strain energy contained in the SiGe film during pseudomorphic growth, it embodies a higher energy state than for an unstrained film, and hence nature does not favor this growth condition except under a very narrow range of conditions, as discussed below.

Second and alternatively, the SiGe film can "relax" during growth to the natural lattice constant determined by its Si and Ge fraction, as given by (2.2). The SiGe film relaxes via misfit dislocation formation, resulting in a break in crystallinity across the growth interface, and a defected film unsuitable for high-yielding device applications. Relaxation during SiGe growth occurs when the pent-up strain energy is sufficiently large that misfit dislocations nucleate and then glide (move).

[2]The word "pseudo" is derived from the Greek word *pseudēs*, meaning "false," and the word "morphic" is derived from the Greek word *morphē*, meaning "form." Hence, pseudomorphic literally means *false-form*.

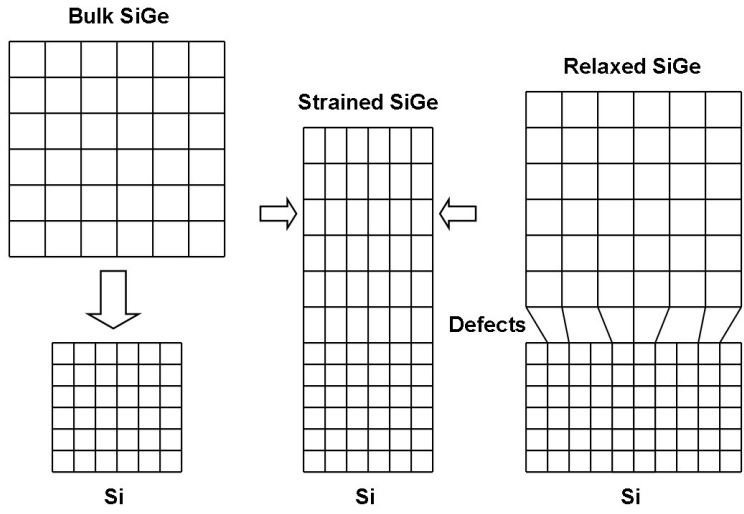

Figure 2.3 Schematic 2-D representation of both strained and relaxed SiGe on a Si substrate.

In essence, when the strain energy in the film exceeds the activation energy required for misfit formation and movement, the film will relax, releasing the stored strain energy. Not surprisingly, this relaxation mechanism is complex, and varying degrees of residual (post-growth) strain can reside in relaxed SiGe films. The misfit dislocations formed during the relaxation process may be either confined to the original growth interface plane, or "thread" their way up through the overlaying SiGe epitaxy, or both, and in either case represent a bad situation from a device design perspective, since such defects can act as generation/recombination (G/R) trapping centers, and high-diffusivity pipes for dopants, which are well-known yield "killers" in bipolar technologies. [3] These two growth scenarios are depicted schematically in Figure 2.3 and Figure 2.4.

Practically speaking, if we imagine an unrestricted SiGe growth process for some arbitrary Ge fraction, it would proceed as follows. Since the Si substrate is very thick (about 600 μm for a 200-mm wafer), and very stiff, it remains essentially unchanged during the epitaxial growth process. Assuming a pristine initial growth interface, the growth of the SiGe film will begin pseudomorphically, adopting the

[3] It is worth pointing out that SiGe-based FET device technologies are inherently less sensitive to such strain relaxation induced defects, simply because the FET is a majority carrier device. Minority carrier devices such as the pn junction and bipolar transistor will always be less tolerant of growth induced defects.

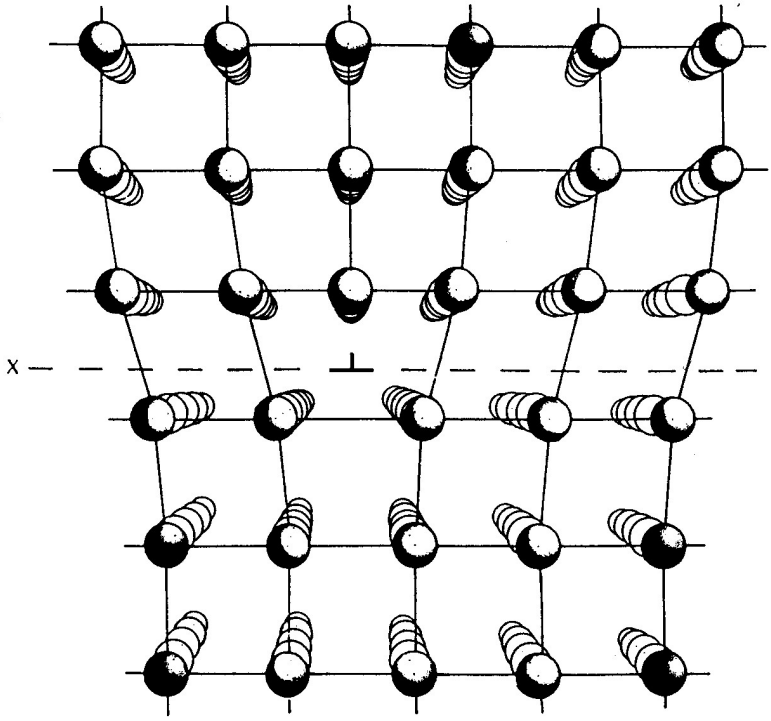

Figure 2.4 Schematic representation of misfit dislocation formed at the Si/SiGe growth interface.

underlying Si lattice constant, but when a given "critical thickness" is reached, the strain energy becomes too large to maintain local equilibrium and the SiGe film will relax to its natural lattice constant, with the excess strain energy being released via misfit formation. In practice, it is also common for the film to remain pseudomorphic until the end of the growth cycle, even though it may have exceeded the critical thickness. Such a SiGe layer is said to be "metastable." Metastable films will relax during subsequent thermal processing steps that add energy to the system, and thus are not suitable for use in Si-fabrication-compatible SiGe technologies. During the relaxation process, whenever it occurs, chaos results, as can be clearly seen in the plan-view TEM micrograph of a relaxed and heavily defected SiGe film shown in Figure 2.5.

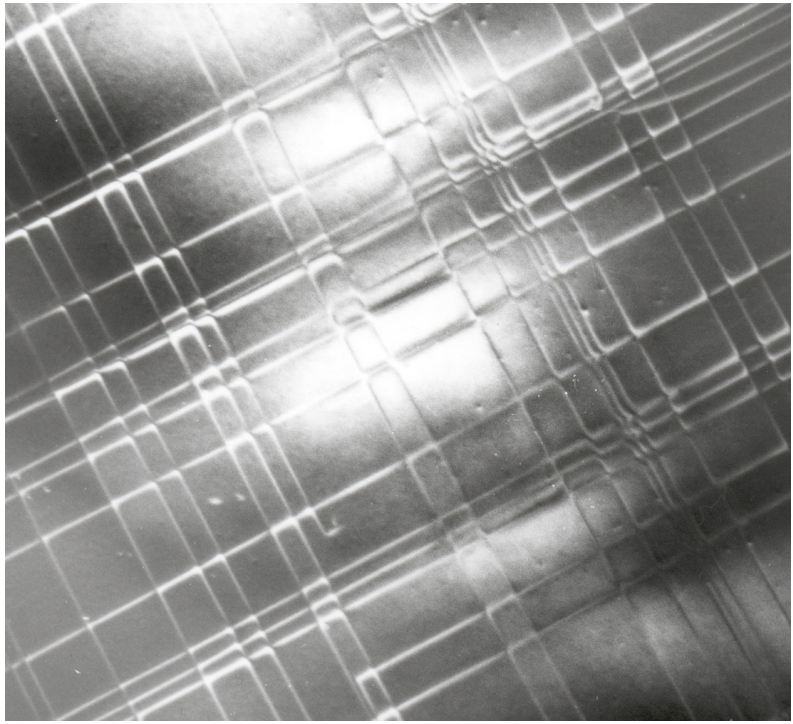

Figure 2.5 Plan-view TEM (top down image) of an unstable SiGe film that has been annealed and undergone relaxation. The visible linear structures are misfit dislocations.

2.1.2 Putting Strained SiGe into SiGe HBTs

Regardless of the growth technique used, or the structure and self-alignment schemes employed in the transistor, strained SiGe films found in today's commercially viable SiGe HBTs all have a similar form. As depicted in Figure 2.6, the deposited SiGe film actually consists of a three-layer composite structure:

- A thin, undoped Si buffer layer;
- The actual boron-doped SiGe active layer;
- A thin, undoped Si cap layer.

The Si buffer layer is used to start the growth process off on the right foot, and serves two purposes. First, the Si buffer layer helps ensure that a pristine SiGe epitaxial growth interface is preserved between the original Si substrate, which was