REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Sept 94	Final 1 Jul 93-30 Jun 94

4. TITLE AND SUBTITLE
IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices & Circuits

5. FUNDING NUMBERS
DAAH04-93-G-0183

6. AUTHOR(S)
L.F. Eastman (principal investigator)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cornell University
Ithaca, NY 14850

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 31238.1-EL-CF

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The conference included the following sessions: Plenary,, heterojunction bipolar transistors, poster session, resonant tunneling and novel devices, optoelectronic devices and circuits, device simulation and optimization and heterostructure field effect transistors.

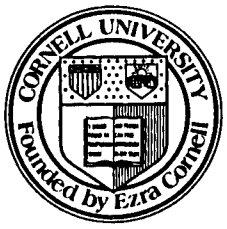
14. SUBJECT TERMS
High Speed Semiconductor Devices, Devices, Semiconductor Devices, Transistors, Optoelectronic Devices, Circuits, Conference

15. NUMBER OF PAGES
556

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

PROCEEDINGS
IEEE/CORNELL CONFERENCE ON ADVANCED CONCEPTS IN HIGH SPEED SEMICONDUCTOR DEVICES AND CIRCUITS



August 2-4, 1993

CORNELL UNIVERSITY
ITHICA, NEW YORK

Sponsored by
IEEE ELECTRON DEVICE SOCIETY
IN COOPERATION WITH
IEEE MICROWAVE THEORY AND TECHNIQUES SOCIETY
AND CORNELL UNIVERSITY

IEEE ISBN Number: 0-7803-0894-8

Proceedings
IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits

P.K. Bhattacharya
GENERAL CHAIRMAN

G. Maracas	M.G. Adlerstein	U. Mishra
PROGRAM	PROCEEDINGS	TREASURER

PROGRAM COMMITTEE

C.O. Bozler, C. Clark, L.F. Eastman
W. Frensley, C. Gaw, J. Hutchby, T. Jackson
A. Khatibzadeh, E. Kohn, L. Larson, L. Lunardi, R. Malik
G. McCoy, R.A. Murphy, D. Pavlidis, M. Reed, R. Trew

August 2-4, 1993
Cornell University
Ithaca, New York

Sponsored by
IEEE Electron Devices Society
and the U.S. Army Research Office
in cooperation with
IEEE Microwave Theory and Techniques Society
and Cornell University

IEEE ISBN Number: 0-7803-0894-8

IEEE Catalog Number 93CH3235-9
ISBN 0-7803-0894-8 (softbound)
ISBN 0-7803-0895-6 (casebound)
ISBN 0-7803-0896-4 (microfiche)
Library of Congress Number 92-75018

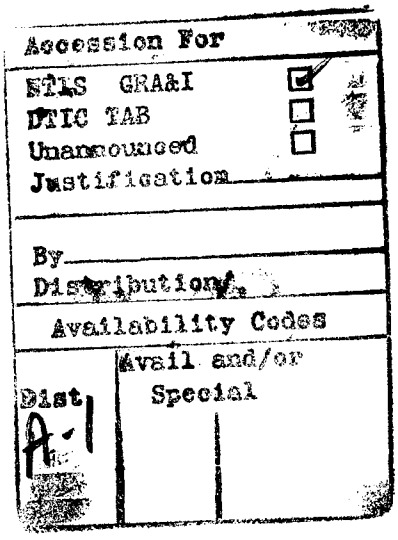


TABLE OF CONTENTS

MONDAY, AUGUST 2, 1993

SESSION I
PLENARY
Chairman: George N. Maracas

MONDAY, AUGUST 2, 1993

SESSION II
HETEROJUNCTION BIPOLAR TRANSISTORS
Chairman: L. Lunardi

MONDAY, AUGUST 2, 1993

SESSION III
POSTER SESSION

TUESDAY, AUGUST 3, 1993

SESSION IV
RESONANT TUNNELING AND NOVEL DEVICES
Chairman: W.R. Frensley

TUESDAY, AUGUST 3, 1993

SESSION V
OPTOELECTRONIC DEVICES AND CIRCUITS
Chairman: C. Gaw

WEDNESDAY, AUGUST 4, 1993

SESSION VI
DEVICE SIMULATION AND OPTIMIZATION
Chairman: E. Kohn

WEDNESDAY, AUGUST 4, 1993

SESSION VII
HETEROSTRUCTURE FIELD EFFECT TRANSISTORS
Chairman: R.J. Trew

X-RAY LITHOGRAPHY FOR HIGH SPEED DEVICES

H.J. Smith, Massachusetts Institute of Technology

PAPER UNAVAILABLE FOR PUBLICATION

Millimeter-Wave Network Analysis using Nonlinear Transmission Lines

M.J.W. Rodwell, R. Yu, M. Reddy, J. Pusl[1], S. Allen, M. Case[2], U. Bhattacharya

Department of Electrical and Computer Engineering,
University of California, Santa Barbara, CA 93106

ABSTRACT

We report systems for network measurements at millimeter-wave frequencies. Active probes are used for on-wafer measurements to approximately 150 GHz. The active probes incorporate monolithic GaAs mm-wave network analysis circuits (comprising a nonlinear transmission line stimulus signal generator and a directional sampling circuit) and low-loss, rugged quartz coplanar-waveguide probe tips. Wideband transmitter and receiver integrated circuits, incorporating nonlinear transmission lines and sampling circuits interfaced to frequency-independent antennas, are used for free-space transmission measurements to approximately 250 GHz.

I. INTRODUCTION

Power-gain cutoff frequencies of heterostructure transistors have exceeded 350 GHz [1], enabling the demonstration of monolithic amplifiers and other receiver components at frequencies as high as 120 GHz. For circuit development at higher frequencies, serving applications in fiber and wireless communications, radar, and imaging, improvements in transistor bandwidth must be accompanied by development of the wideband instrumentation used for device modeling and for circuit evaluation. Monolithic millimeter-wave integrated circuits (MIMICs), in particular, demand accurate transistor and passive element models during the design phase, models which cannot be extrapolated with great confidence from measurements at lower frequencies.

Limited instrumentation is a particularly pressing difficulty in the design of the sub-mm-wave mixers and frequency multipliers used for radio astronomy and atmospheric remote sensing. Device RF models are frequency extrapolated from DC measurements and circuit impedances usually cannot be measured, thus mixers and multipliers are empirically adjusted using waveguide mounts with tuning screws.

Sampling oscilloscopes, network analyzers, counters, and synthesizers use diode sampling bridges for signal measurement or frequency downconversion. Instrument bandwidth is limited by sampling circuit bandwidth, in turn limited by the duration of the

---

[1]Now with Hughes Aircraft Company, Space and Communications Group, S12/W305, P.O. box 92919, Los Angeles, CA 90009.

[2]Now with Hughes Research Laboratories, 3011 Malibu Canyon Road, Malibu, CA 90265

0-7803-0894-8/93 $3.00 © 1993 IEEE

strobe pulse used to switch the diodes. Since 1966, 20-30 ps silicon step-recovery diodes (SRDs) have been used for strobe pulse generation, limiting sampling circuit bandwidth to ≈20–40 GHz.

We have developed nonlinear transmission line (NLTL) pulse generators with over a factor of 10 shorter transition times than such SRDs. The pulse trains generated by NLTLs have significant spectral content to sub-mm-wave frequencies. Using NLTLs as strobe pulse generators, sampling circuits with millimeter-wave bandwidths have been realized. NLTLs and sampling circuits are then used as signal sources and detectors for in systems for on-wafer and free-space mm-wave network analysis.

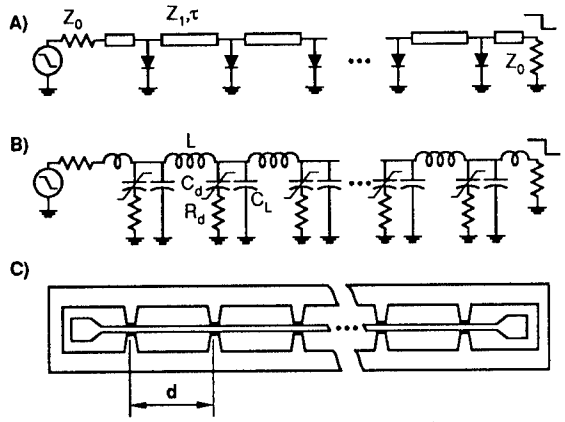


Figure 1: NLTL circuit diagram, a), equivalent circuit, b), and layout, c). C_D is the diode capacitance and R_D its series resistance, $C_L = \tau/Z_1$ is the line capacitance and $L = Z_1\tau$ the line inductance.

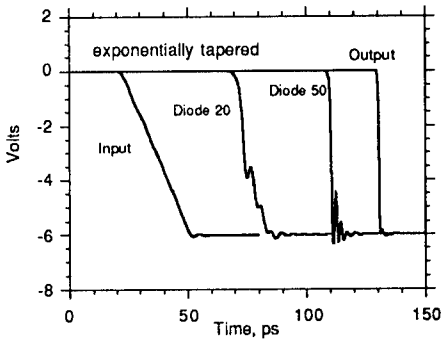


Figure 2: SPICE simulation of NLTL wavefront compression and shock-wave formation.

II. THE SHOCK-WAVE NLTL

The NLTL [2] (fig. 1) is a high-impedance transmission line periodically loaded by reverse-biased Schottky diodes acting as voltage-variable capacitors. The wave propagation velocity varies as the inverse square root of the total (diode plus transmission line) capacitance per unit length and hence increases as the diode reverse bias voltage is increased.

Given a negative-going step function (wavefront) input, the initial portions of the wavefront, near zero volts, will propagate more slowly than the final, more negative, portions of the wavefront. The wavefront transition time (falltime) will progressively decrease with propagation distance. An asymptotic (minimum) compressed falltime is eventually reached (fig. 2) at which the NLTL compression is balanced by various bandwidth limits in the structure. The two dominant bandwidth limits are the varactor diode cutoff frequency $f_D = 1/2\pi R_D C_D$ (defined using the average diode capacitance $\Delta Q/\Delta V$) and the periodic-line (Bragg) cutoff frequency $f_{per} = 1/\pi\sqrt{L(C_D + C_L)}$. If the diode cutoff frequency is dominant, the diode doping is uniform, and the wavefront is 6

Volt amplitude, the minimum compressed falltime is $T_{f,\min} = 1.4/f_D$. Advanced Schottky varactor diodes attain 2-10 THz cutoff frequencies; with further work, NLTL transition times may ultimately reach 0.2–0.3 ps (1–1.5 THz signal bandwidth).

Diode design for the NLTL is a compromise between the objectives of high compression rate (small die area, low attenuation), high diode cutoff frequency (short falltimes), and reverse breakdown voltage. As the diode dimensions are reduced and the active-layer doping increased, the cutoff frequency increases but the reverse breakdown decreases. The larger capacitance variation of hyperabrupt varactors increases the NLTL compression rate (decreasing the required line length and hence both the skin loss and the die area), but hyperabrupt doping decreases the cutoff frequency and the reverse breakdown. The required diode reverse breakdown voltage is determined by the minimum strobe NLTL output voltage required to drive a 2-diode sampling bridge (≈3-4 V) and by the skin-effect losses on the NLTL. As input voltages on current NLTLs are approximately 2:1 larger than the compressed output wavefronts, approximately 8 V reverse breakdown. is required.

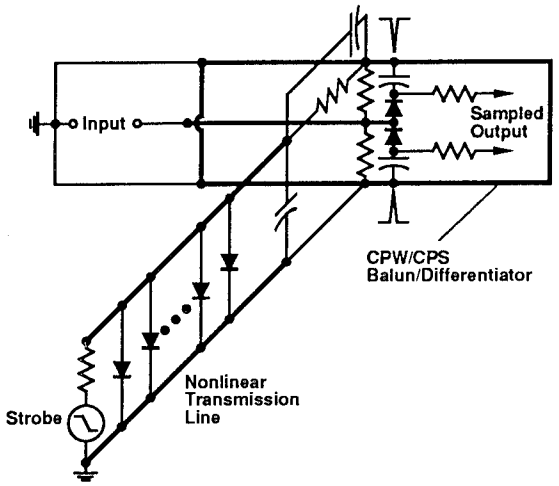


Figure 3: NLTL-gated sampling circuit

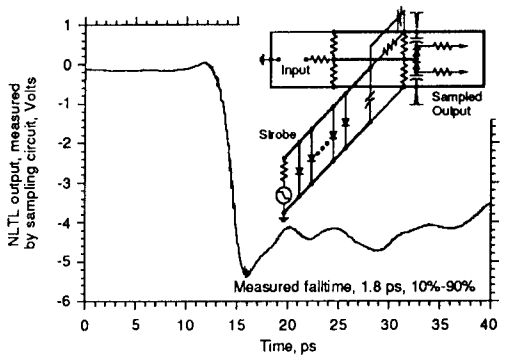


Figure 4: NLTL output measured by an NLTL-gated sampling circuit, both using hyperabrupt diode technology. The measured falltime is 1.8 ps.

III. NLTL-GATED SAMPLING CIRCUITS

A sampling circuit (fig. 3) [2] consists of a strobe pulse generator, a diode/resistor bridge, and a balun/differentiator. An NLTL compresses an input strobe signal, either a step function or a ~10 GHz sinewave, to picosecond falltimes. The sampling diodes are gated by a pair of symmetric positive and negative impulses generated from the strobe NLTL output using a balun / differentiator implemented using the coplanar strip (CPS) mode of the input signal coplanar waveguide (CPW). Coupled through series hold capacitors, the complementary strobe pulses drive the sampling diodes into forward conduction. During this period, the aperture time, the input (RF) signal partially charges the hold capacitors. If the RF frequency is then offset by Δf from a multiple nf_0 of the

strobe frequency f_o, the sampled (IF) signal is mapped out at a repetition frequency Δf. Sampling circuit bandwidth is limited by the sampling diode capacitance and by the aperture time.

To evaluate the NLTL and sampling circuit risetime, the output of an NLTL shock generator is connected to an on-wafer NLTL-gated sampling circuit. The convolved responses of sampling circuit and NLTL shock-wave generator is thus measured. With an NLTL using 1.7 THz exponential hyperabrupt diodes (fig. 4), a 1.8 ps falltime is measured. From this, a 1.3 ps deconvolved NLTL falltime and a 275 GHz sampling circuit bandwidth are determined; NLTLs and sampling circuits with approximately twice this bandwidth will be reported elsewhere [3] by Allen et. al.

III. NLTL SOLITON IMPULSE COMPRESSORS

Although not employed in the current NLTL-based instruments reported here, variants of the NLTL employing soliton propagation effects are more efficient than shock-wave devices in the generation of harmonics of an input signal. Among other applications, these devices can be used as the stimulus signal for network analysis, in place of the shock-wave NLTLs used in the current work.

A solitary wave is a traveling wave having a localized transition (e.g. a pulse) and propagating without distortion in a nonlinear, dispersive medium. Solitons are defined as those solitary waves which preserve their shape and velocity after interactions with other solitons [4]. The soliton is a pulse waveform for which the effects of nonlinearity and dispersion are balanced. If the NLTL Bragg frequency is much smaller than the diode cutoff frequency then propagation is dominated by the interaction between the capacitive nonlinearity and the periodic-network dispersion, and solitons propagate [5]. On NLTLs, soliton duration is inversely proportional to the Bragg frequency and varies approximately as the inverse square root of the peak amplitude. Soliton propagation velocity increases with increasing soliton amplitude. Applied to the NLTL, signals with pulse duration longer than the duration of a soliton correspond to a nonlinear superposition of a set of solitons having differing amplitudes and velocities, and will decompose into this set of two or more solitons during propagation. Figure 5 shows a circuit simulation of a 6 V, 63 ps impulse splitting during propagation into a pair of solitons. Note that the leading output soliton has larger amplitude and shorter duration than the input signal.

A long-duration impulse input to an NLTL will decompose into its characteristic set of solitons. Longer input pulses decompose into progressively larger numbers of solitons, and impulse compression ratios are limited to approximately 2:1 on a homogeneous line. Higher compression ratios can be obtained by using step-tapered lines, consisting of a line with Bragg frequency f_{per1} cascaded with a line with Bragg frequency $f_{per2}=2f_{per1}$. The first line section performs a 2:1 pulse compression, with the second line section performing a further 2:1 compression. While higher compression ratios can be obtained by repeating this scheme in a three-step or four-step fashion, it is

more convenient to approximate such a step-tapered line by tapering the Bragg frequency continuously [6]. Figure 6 shows the output of an exponentially-tapered soliton impulse compressor.

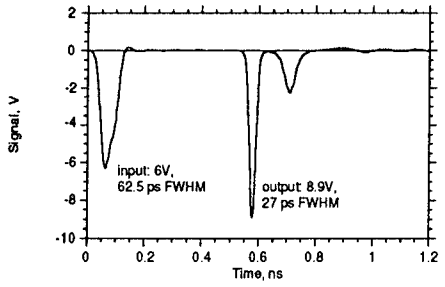


Figure 5: SPICE simulation: splitting of an input impulse into a pair of solitons during propagation on an NLTL.

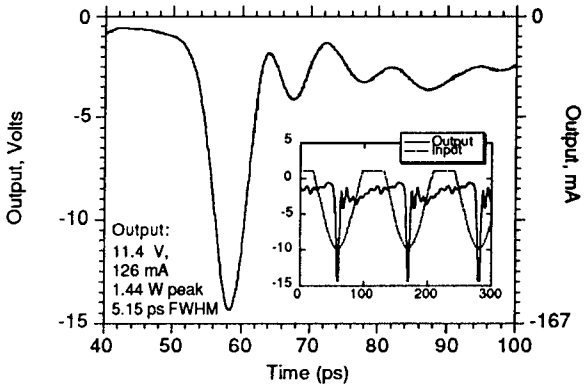


Figure 6: Measured output waveform and calculated input waveform for a tapered soliton impulse compressor

Experimental soliton NLTL output pulse durations are ≈5:1 longer than shock-wave devices, although circuit simulations predict attainable soliton durations only a factor of 1.5:1 larger than shock-wave NLTL falltimes using similar diodes. The discrepancy is not understood, although there is evidence that layout parasitics reduce Bragg frequencies by ≈2:1. If this discrepancy can be solved, soliton compressors will be much more effective than shock-wave NLTLs for high-order mm-wave frequency multiplication, as the conversion loss to the N^{th} harmonic is bounded above by 1/N for the soliton compressor, compared to $1/N^3$ for the shock-wave device.

IV. ON-WAFER MM-WAVE NETWORK ANALYSIS

Using shock-wave NLTLs and NLTL-gated sampling circuits mounted in active wafer probes, on-wafer network analysis can be performed at millimeter-wave frequencies. Figure 7 shows a typical network analyzer. A swept-frequency stimulus signal generator drives the device under test. The incident and reflected waves are separated by a directional coupler, downconverted from microwave to radio frequency by sampling circuits, and subsequently measured by RF vector voltmeters. Bandwidth limits include the sampling circuits and the frequency range of the stimulus signal generator. Available coaxial connectors also limit bandwidth; 110 GHz coaxial connectors were introduced in March 1993, with earlier connectors limited to 65 GHz.

We have developed active probes [7,8] (fig. 8) for on-wafer mm-wave vector network analysis (VNA). These contain an NLTL-based network analyzer IC and a

rugged, wideband quartz probe tip. The IC itself incorporates an NLTL stimulus signal generator and an NLTL-strobe directional sampling circuit which independently measures the forward and reverse waves from the device under test.

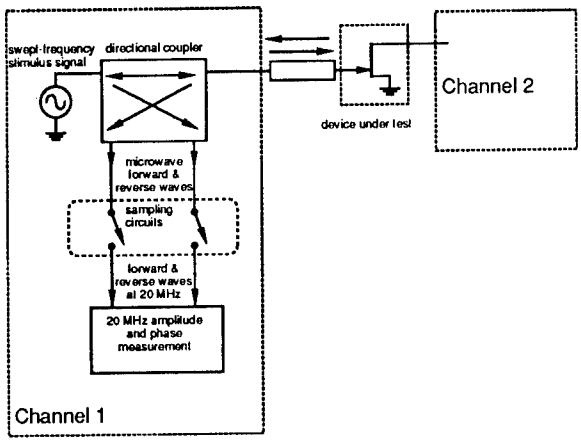


Figure 7: Simplified block diagram of a typical microwave network analyzer. Bandwidth limits are set by the signal source frequency range, by the sampling circuits, and by the coaxial connectors.

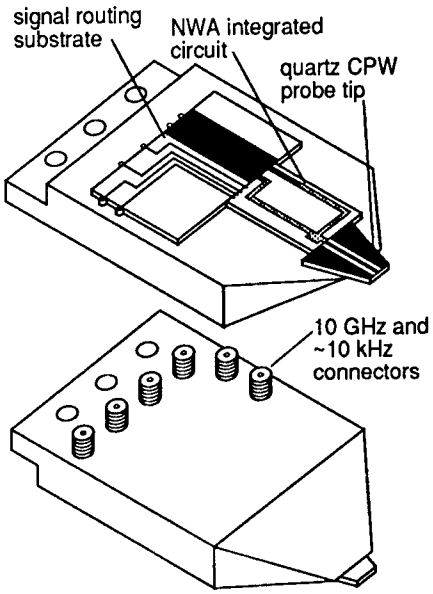


Figure 8: Active Probe for on-wafer mm-wave network analysis

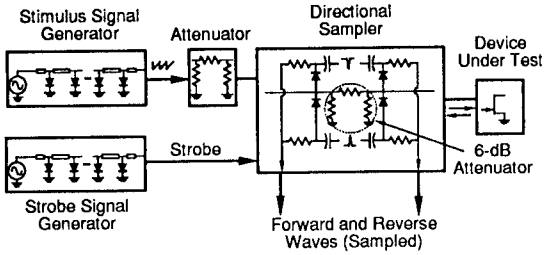


Figure 9: Block Diagram of the network analysis integrated circuit

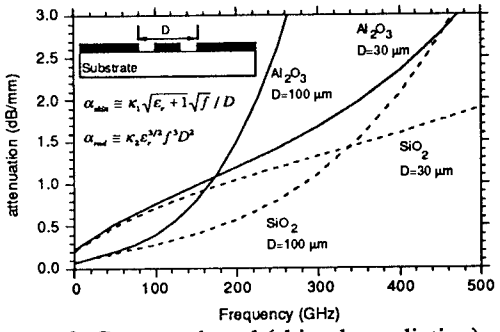


Figure 10: Computed total (skin plus radiation) losses for coplanar waveguide probe tips on quartz and alumina

The network analyzer IC block diagram is shown in figure 9. An NLTL stimulus signal generator converts a 7-14 GHz sinusoidal drive signal into a 5 V sawtooth waveform with ~2 ps transition times and with significant harmonic content to ~250 GHz. The drive signal is attenuated to levels suitable for linear characterization of transistor circuits, and is passed through a directional sampling circuit to the device under test.

The directional sampling circuit is a pair of sampling circuits measuring the port voltages of a 6 dB attenuator placed between the stimulus signal generator and the device under test. A second NLTL on the IC provides the strobe signal for the directional sampler. The attenuator port voltages $V_{out} = V_{incident} + V_{reflected}$ and $V_{in} = 2V_{incident} + V_{reflected}/2$, linear functions of the incident and reflected voltages, are downconverted by the sampling circuits to a 1 kHz intermediate frequency. The IF signals are passed through summing networks to recover $V_{incident}$ and $V_{reflected}$, and are then measured. We emphasize that the directional sampler does not obtain independent measurements of the forward and reverse waves through time-separation of pulsed signals, as in a time-domain reflectometer. Errors associated with time-gating (uncorrected terms in source and load reflections, time truncation of the long-duration impulse responses of narrowband resonators) are therefore avoided.

Tips for the active probe are coplanar waveguide (CPW) on quartz substrates. To leading order, for a given line impedance, skin loss (in dB/mm) varies as $\approx \sqrt{\varepsilon_r + 1}\sqrt{f}/D$ while radiation loss varies as $\approx \varepsilon_r^{3/2} f^3 D^2$ (fig. 10) [9], where ε_r is the relative dielectric constant of the substrate, f is frequency, and D is the ground-ground spacing of the CPW. With appropriate scaling of line dimensions, CPW on a quartz substrate (ε_r=3.8) has lower attenuation than the CPW lines on alumina substrates (ε_r=9.8) used in commercial microwave probes. At 200 GHz, D=100 μm results in a minimum 0.6 dB/mm attenuation on quartz. The probe tips have plated nickel contact bumps, plated gold airbridge ground straps for slot-line mode suppression, and the tips are angle-lapped. The 60 pH wire-bond inductance between the probe tip and the network analyzer IC is acceptable for 200 GHz measurements. Because the short quartz tips are rigid, the probes are mounted on their arms using a complaint joint.

For mm-wave measurements, the active probe stimulus NLTLs are driven at 7-14 GHz, with the strobe NLTL drive frequency offset by 1 kHz to produce an IF at this frequency. To measure device S-parameters at harmonics of the NLTL drive frequency, harmonics of the IF signals are measured by four vector voltmeters (lock-in amplifiers). The system is controlled by a desktop computer which also performs line-reflect-match calibration routines.

Figures 11-14 demonstrate system accuracy and repeatability after calibration over the 10-160 GHz range. Transmission measurements can presently be made to 160 GHz, but poor directivity prohibits measurements of return losses smaller than -20 dB at frequencies below 120 GHz. System performance is currently limited by several factors. Because of a design error, the airbridges used for probe tip slot-line mode suppression are capacitively loading the line, reducing the line impedance to 25Ω and introducing a 160 GHz cutoff; the reflection from the 25Ω loaded tip impedance seriously degrades the uncalibrated directivity. Synthesizer phase noise is also severe, corresponding to a small 1° phase deviation at 10 GHz but a large 20° deviation at 200 GHz. These factors are now being addressed.

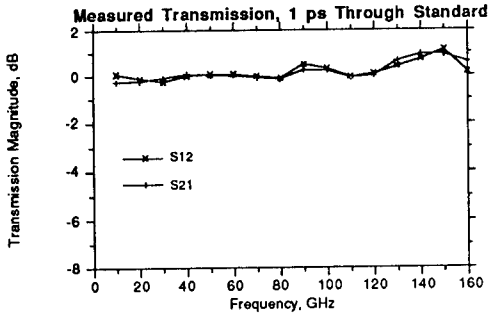


Figure 11: Measured transmission of a 1 ps through line. The deviation from 0 dB shows system accuracy.

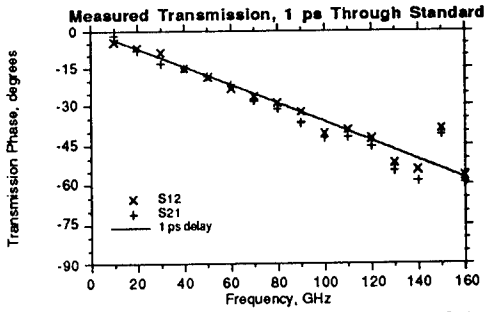


Figure 12: Measured transmission phase of the 1 ps through line. The deviation from 1 ps delay shows system accuracy.

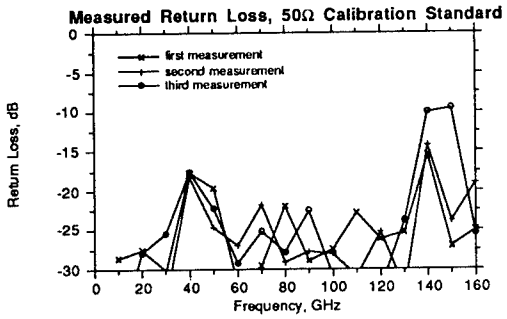


Figure 13: Measured return loss of a 50Ω calibration standard. The deviation from zero reflection (−∞ dB) shows system directivity

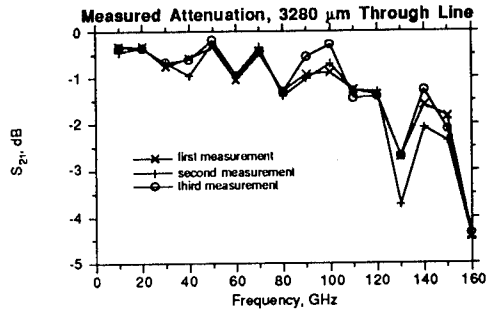


Figure 14: Measured transmission of a long coplanar line. The deviation between measurements shows system repeatability.

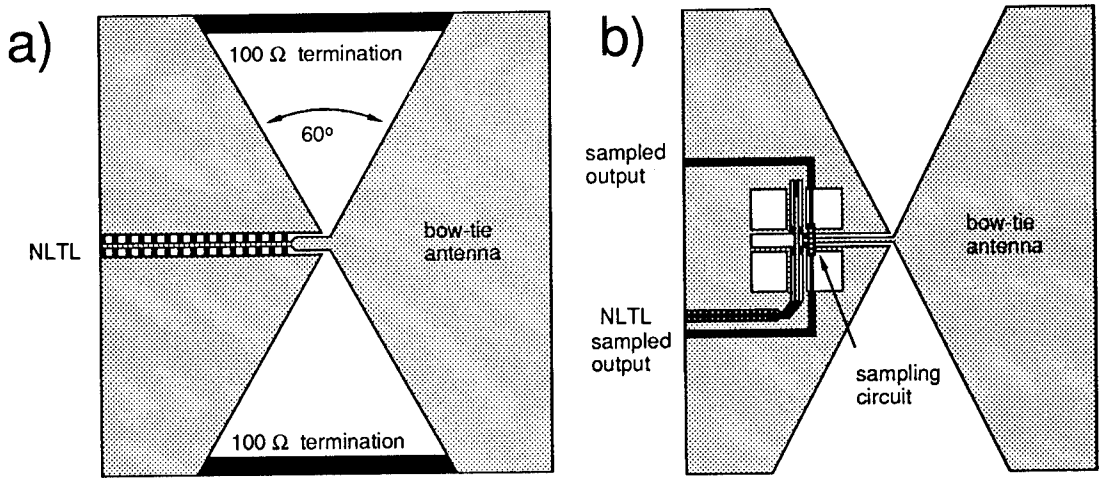


Figure 15: Simplified integrated circuit layouts of a) picosecond transmitter and b) picosecond receiver for free-space mm-wave network analysis

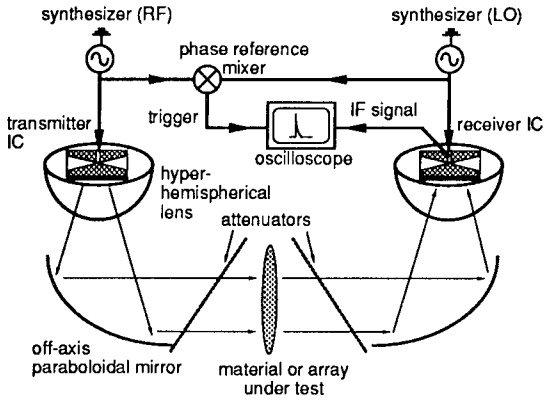


Figure 16: NLTL-based free-space network analyzer

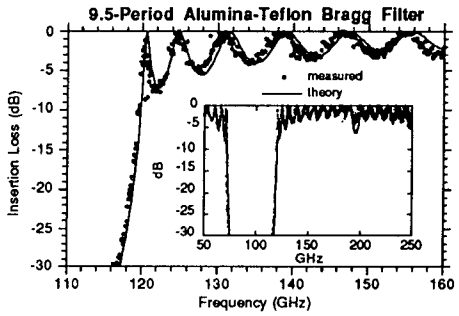


Figure 17: Transmission of a mm-wave Bragg filter

V. FREE-SPACE mm-WAVE NETWORK ANALYSIS

We have also constructed a system for gain-frequency measurements in free-space, for characterization of materials and radiating structures [10]. The system generates and detects picosecond pulses and obtains frequency information through Fourier transformation. Compared to measurement systems using mm-wave harmonic multipliers and harmonic mixers, the system is simple, relatively inexpensive, and covers a 30-250 GHz bandwidth with a single set of components.

For broadband transmission measurements in free space, the NLTL and sampling circuits must be interfaced to broadband antennas. Scale invariant antennas, including the bow-tie antenna, have frequency-independent far-field radiation patterns, and antenna impedance. The transmitter NLTL, fabricated within one electrode plane, drives the 50Ω impedance antenna, through a coplanar waveguide (CPW) balun (fig. 15,a). The receiver is a bow-tie antenna interfaced to an sampling circuit using a similar antenna feed (fig. 15,b). The system is shown in figure 16. Hyperhemispherical substrate lenses partially collimate the antenna radiation, and the radiated beam is collimated with off-axis parabolic mirrors. The receiver uses similar optics. To obtain accurate measurements, cavity resonances between the transmitter and receiver are suppressed by placing 5 dB attenuators at oblique incidence on both sides of the sample under test.

The transmitter NLTL is driven between 7 and 14 GHz, while the sampling circuit is driven at a frequency 100 Hz below the transmitter frequency. The resulting sampled 100 Hz signal is observed directly on an oscilloscope. The received signal at the sampling circuit output is a pulse train with 2.4 ps risetime. Transmission (amplitude and phase) measurements are obtained by taking the ratio of the received Fourier spectrum with the device under test in place with the spectrum of a reference measurement taken with the device under test removed. Fig. 17 shows a measurement of a high-Q Bragg filter.

VI. CONCLUSIONS

Nonlinear transmission lines and NLTL-strobed monolithic sampling circuits are *monolithic solid-state* devices for generation and measurement of electrical signals in the picosecond and millimeter-wave regimes. NLTL technology continues to improve, and 0.3 ps NLTL transition times and 1 THz sampling circuit bandwidths appear to be feasible. Based upon these, systems for sub-mm-wave on-wafer and free-space network measurements will evolve. Robust network instrumentation for device characterization and modeling is imperative for millimeter-wave and sub-mm-wave monolithic circuit design.

ACKNOWLEDGMENTS

The work was supported by the U.S. Air Force Office of Scientific Research (program number F49620-92-0464) and an associated AASERT contract, by a National Science Foundation Presidential Young Investigator Award, and by the DARPA Optoelectronics Technology Center. We acknowledge generous support from Hughes Aircraft Corporation, Wiltron, Shimadzu, Tektronix, and Hewlett-Packard Corporation.

REFERENCES

[1] U. K. Mishra, A.S. Brown, and S.E. Rosenbaum, 1988 International Electron Device Meeting, Dec. 11-14, San Francisco.
[2] Mark Rodwell, Masayuki Kamegawa, Ruai Yu, Michael Case, Eric Carman, and Kirk Giboney, IEEE Transactions on Microwave Theory and Techniques, Vol. 39, No. 7, July 1991, pp. 1194-1204.
[3] S. Allen, M. Reddy, M.J.W. Rodwell. To be presented, 1993 GaAs IC symposium, October, San Jose, CA.
[4] A.C.Scott, F.Y.F. Chu, and D.W.McLaughlin, IEEE Proc, Vol. 61, No. 10, Oct. 1973, pp. 1443-1482.
[5] R. Hirota and K. Suzuki, IEEE Proc., Vol. 61, No. 10, Oct. 1973, pp. 1483-1491.
[6] Michael Case, Eric Carman, Ruai Yu, and M.J.W. Rodwell, Applied Physics Letters, Vol. 60, no. 24, 15 June 1992, pp. 3019-3021.
[7] R.A. Marsland, C.J.Madden, D.W. Van Der Weide, M.S. Shakouri, and D. M. Bloom, 1990 GaAs IC Symposium, New Orleans, Oct. 7-10, pp. 19-22.
[8] Ruai Y. Yu, Joe Pusl, Yoshiyuki Konishi, Michael Case, Masayuki Kamegawa, and Mark Rodwell, IEEE Microwave and Guided Wave Letters, Vol. 2, No. 8, August 1992, pp. 319-321.
[9] D. B. Rutledge, D. P. Neikirk, and D. P. Kasilingam. "Integrated- Circuit Antennas" in Infrared and Millimeter Waves, K. J. Button, Ed., Vol. 10, pp. 1-90, New York: Academic Press, 1984.
[10] Y. Konishi, M. Kamegawa, M. Case, R. Yu, M.J.W. Rodwell, and D.B. Rutledge, Applied Physics Letters, Vol. 61, No 23, 7 Dec. 1992.

MICROCAVITY OPTOELECTRONIC DEVICES
(Invited paper)

L. A. Coldren, D. B. Young, M.G. Peters, F. H. Peters, J. W. Scott, C. C. Barron, B. J. Thibeault, and S. W. Corzine

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106

ABSTRACT

Over the past few years vertical-cavity surface-emitting lasers and modulators have emerged as viable devices with interesting performance characteristics. One of their key features is that they occupy very little substrate area as compared to most optoelectronic devices. As a result, they also require relatively low drive powers. These aspects together with their suitability for wafer-scale fabrication and testing make them appear suitable for low-cost production as well as high performance.

In this paper we shall review recent progress on these devices with emphasis on the vertical-cavity laser. Vertical-cavity lasers with cw powers exceeding 110mW, overall efficiencies exceeding 17%, operating temperatures exceeding 120°C, and output powers insensitive to temperature over ranges exceeding 60°C will be illustrated. In addition, devices have operated cw down to diameters of 2 μm, including a 6 μm device that delivers nearly 2 milliwatts of single-mode output power with greater than 30 dB of spurious mode suppression.

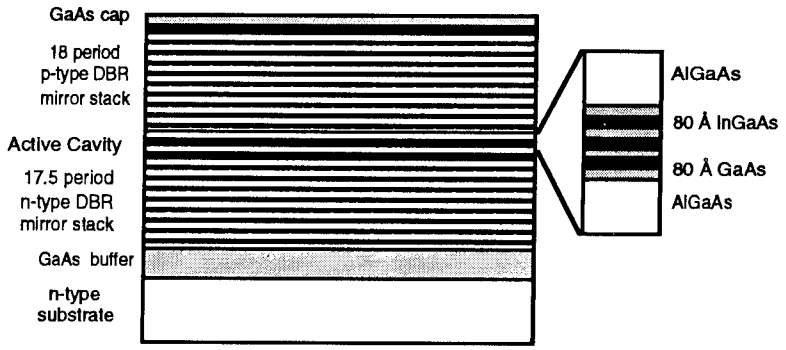
In the vertical-cavity modulator area, reflective asymmetric Fabry-Perot structures have given up to 37 GHz of modulation bandwidth. Insertion losses are about 3 dB, and required voltage swings for 100:1 modulation are ~ ± 2V.

I. INTRODUCTION

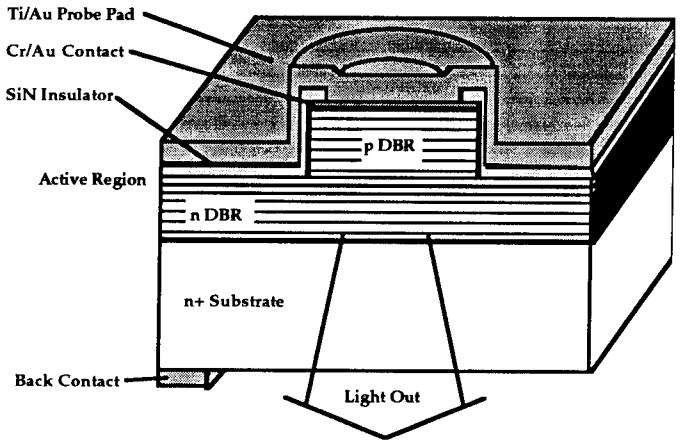
Vertical-cavity surface-emitting lasers (VCSELs) continue to improve and look more practical every year. Within the past year, for example, there have significant advances in three different wavelength regimes: visible (< 0.7 μm)[1], short-wavelength near-IR (0.8 - 1.0 μm)[2-6], and long-wavelength near-IR (1.2 - 1.6 μm)[7-9]. In the visible case, room-temperature pulsed lasing to wavelengths as short as 0.63 μm were demonstrated for the first time. In the short-wavelength near-IR case, overall cw power (wall-plug) efficiencies have approached 20%[6], and power outputs can be stable over temperature ranges exceeding 50°C[3,5] for a constant current drive. In the long-wavelength near-IR case, the first 'cool' room temperature (14°C) cw devices were operated at 1.3 μm[9].

It fact, VCSELs in the 0.8-1.0 μm range should be expected to exceed the performance of their in-plane counterparts in most respects. Reasons include the availability of nearly twice the modal gain for a given cavity fill-factor[10], the availability of low-loss epitaxial mirrors, the relatively small volume of the active region, and the low numerical aperature of the output beam. Low-resistance electrical pumping schemes have also recently been proven[6,11]. Wall-plug efficiencies are already higher in the low-power (≤ 1 mW) range, where many optical interconnect applications exist. Figure 1 illustrates a schematic of the layer stucture used in many of the UCSB devices as well as the through-the-mirror pumped index-guided structure used in most experiments.

0-7803-0894-8/93 $3.00 © 1993 IEEE



(a)



(b)

Figure 1. (a) Typical MBE-grown layer structure and (b) structure of a finished device.

Asymmetric Fabry-Perot reflective modulators have a very similar structure to the VCSEL illustrated above. However, the active region is replaced by an GaAs/AlGaAs multiple-quantum well (MQW) and the top (input) mirror has a considerably lower reflectance (~50%) relative to the rear mirror (which has a reflectance of ~99%). At zero bias the net reflection is dominated by the back mirror, resulting in a value typically greater than 50%. When the device is reverse biased, the resulting quantum-confined Stark shift increases the absorption in the cavity until the round-trip loss equals the difference in reflectance between the top and back mirrors. At this point the cavity is "balanced" at resonance, and the net reflection falls to zero. Thus, the device can function as a variable mirror with high on/off contrast and relatively low drive voltage. Because the undoped MQW region is relatively thick (typically > 0.5 μm), the capacitance can be quite low for devices with effective diameters ~ 15-20 μm, a size which is relatively easy to hit with simple focusing optics. Thus, high-speed operation is facilitated.

II. RECENT RESULTS ON VCSELS

A. AlAs/GaAs Mirrors

As illustrated in Fig. 1, the specific VCSELs to be discussed are bottom-emitting strained-InGaAs quantum-well (QW) index-guided lasers grown using a Varian Gen II MBE system. In initial devices the bottom mirror consisted of binary AlAs/GaAs quarter-wave layers

doped with Si. A similar top AlAs/GaAs mirror was Be doped. In these structures, the doping was $1x10^{18}$ cm^{-3} in the bulk of the layers, but at the interfaces the doping was linearly graded over a 180Å region and doped $5x10^{18}$ cm^{-3}. The active region consisted of 0.12μm $Al_{0.5}Ga_{0.5}As$ confinement layers surrounding a gain region with three 80Å $In_{0.185}Ga_{0.815}As$ QWs with 80Å GaAs barriers. The p-type mirror was capped with a 450Å p^+ GaAs contact layer. In these structures the net device voltage at threshold (~1-2 kA/cm^2) due the series resistance and the accumulated voltage drops at the mirror heterobarriers amounted to 3 to 4 V. In some more recent devices the Al content in alternate mirror layers is dropped to 67% to reduce this series voltage.

Careful Reflection High Energy Electron Diffraction (RHEED) studies were done prior to the growth to determine the growth rate. Photoreflectometry was done on test DBR mirror stacks to determine the relationship between the RHEED growth rate measurements and the optical characteristics of the resulting structure. Use of this relationship allows the subsequent SEL structures to be grown with the DBR mirror stop band and the cavity mode to be within ±0.5% of the design wavelength. Extensive characterization of InGaAs in-plane lasers (IPL) was done to characterize the emission wavelength of the active region in relation to the RHEED growth rates. All aluminum-containing alloys were achieved by use of short period binary superlattices, or digital alloys. The mirrors and cavity were grown at a substrate temperature of 600°C, and the quantum well region was grown at 520°C. Five second smoothing pauses were used on each side of the quantum wells. The entire structure was grown with cracked arsenic, using GaAs and AlAs growth rates of approximately 1 μm/hr. The InGaAs quantum wells were grown at approximately 1.3 μm/hr. In-plane broad-area laser test structures using a 3-QW InGaAs active region grown at this high rate showed threshold current densities as low as 160 A/cm^2, indicating that low growth rate during the InGaAs wells is not required to get good quality material.

To begin the fabrication of discrete VCSELs, Cr/Au/Ni dots are evaporated onto the semiconductor substrate. These dots form the electrical connection to the p-side of the VCSELs and provide added reflectivity. The Ni forms the mask for dry-etching. The material is dry-etched with chlorine reactive ion etching (RIE) down to the active region, yielding straight walled pillars[12]. The devices are isolated and larger contacts are made possible by depositing SiN_x over the entire substrate using plasma enhanced chemical vapor deposition. In order to make contact through the SiN_x, the wafer is masked and holes are etched through the SiN_x using a CF_4 plasma. For small VCSELs where this alignment makes this impossible a self aligned procedure is used. After depositing SiN_x, a planarizing layer of photoresist is applied. Then the photoresist is etched back using an O_2 plasma revealing the tops of the pillars. The SiN_x can then be removed using a CF_4 plasma. After striping the resist, Ti/Au metallization layers are sputtered onto the SiN_x making contact with the VCSELs pillars. Contacts for arrays or single devices are then formed by removing the unwanted metallization. For higher-speed operation, a layer of polyimide is used instead of the resist, and this is left in place following the O_2 etching. Devices have been fabricated with diameters ranging from less than 2 μm to 150 μm. The VCSELs are arranged for individually probing (hundreds per wafer) or as 1X18 and 4X4 packagable arrays.

Most data has been taken by probing arrays of these devices as etched onto the chip being processed. For higher power operation or to provide a portable package, some devices have been flip-chip bonded onto submounts and placed into a package. Figure 2 shows an example using a diamond heat spreader. Such a structure has allowed cw output powers as high as 116 mW from a 70 μm diameter structure[13]. Using 40 μm structures over 30 mW has been coupled into multimode fiber[14]. Figure 3 summarizes these results.

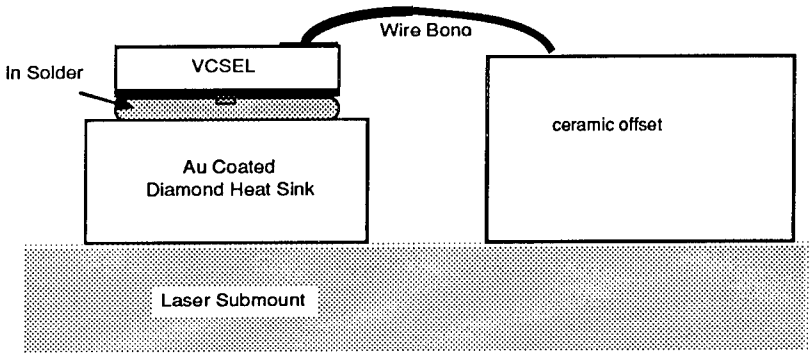


Figure 2. VCSEL heatsink and bonding geometry.

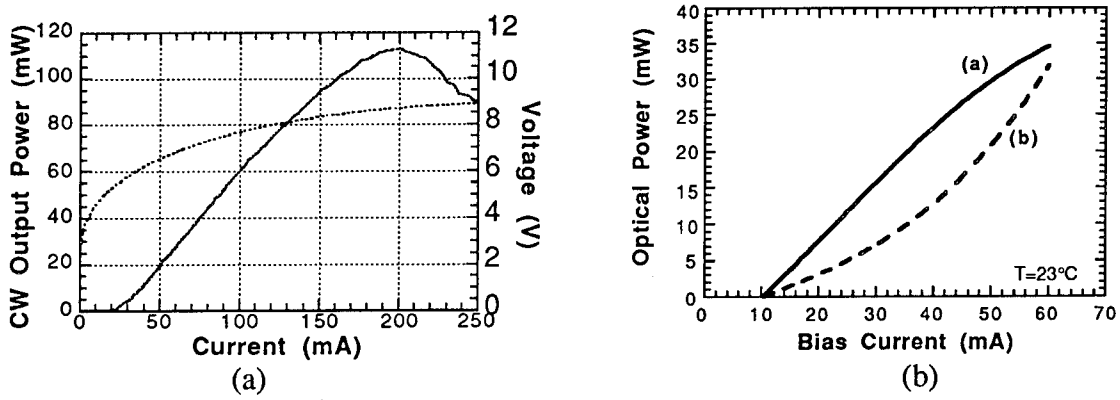


Figure 3. (a) Cw output power and voltage vs. dirve current from a 70 μm diameter VCSEL mounted as in Fig. 2. (b) Solid curve is cw output from a 40 μm diameter VCSEL and dashed curve is the power coupled to a multimode fiber.

The devices of Fig. 3 benefited from the use of higher AlGaAs barriers surrounding the InGaAs quantum wells as compared to the initial work of Geels, et al[12]. The change to higher barriers derived from the investigation of the temperature characteristics of Geels' devices. Modeling of the behavior indicated that the major deficiency in the devices was the low x-value confinement barriers around the active region[15]. The high gain required for lasing in a VCSEL necessitates high carrier densities in the quantum wells. Heating of the active region, due to the resistance and heterobarriers of the mirrors, led to high leakage currents over the x=0.2 $Al_xGa_{1-x}As$ barriers used in the earlier design. Thus, for the devices of Fig. 3, the x-value of the active region around the InGaAs quantum wells was increased to 0.5. Also, the doping at the high-field interface of the first four mirror periods was not increased as it is in the rest of the mirror structure, so as to reduce the free-carrier loss.

Smaller size VCSELs measuring between 8 to 20μm in diameter were also fabricated from this material. The lasing wavelength was measured to be 997nm. Wallplug efficiency of these devices is still low (6%), due to the resistance inherent in the top, p-type mirror design used in this device. Figure 4 shows the L-I curves for the various device diameters. The 8μm diameter device is comparable in area to the 7μm x 7μm devices reported by Geels, et al. [12,17], and the power output is nearly doubled from those devices. The 11μm diameter device

is roughly the same size as Geels' 10x10 μm device, and the power output is 78% higher. The device operates CW up to 140°C, with a maximum output power at that temperature of 0.2 mW.

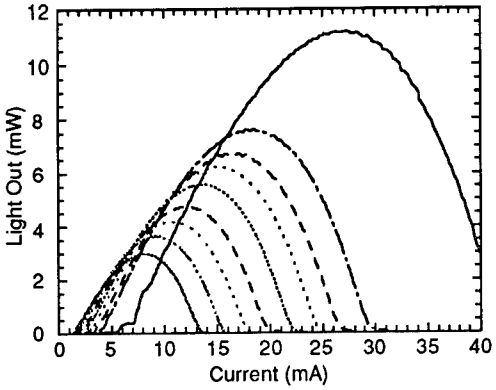


Figure 4. CW L-I curves for index guided cylindrical VCSELs of 8-15 μm (in 1 μm steps), and 20 μm diameter. The lasers exhibit a monotonic increase in power output with diameter. The small ripples are due to residual reflection from the substrate-air interface. This interface has a simple AR coating, but it retains a reflection of ~0.5%.

Another design change that resulted from theoretical investigations [16] was to deliberately offset the mode of the cavity from the gain peak of the quantum well active region. Both the cavity mode and the gain peak shift as the temperature of the active region changes, but while the cavity mode moves at roughly 0.8 Å/°C, the gain peak sweeps at 3.3 Å/°C [17,18]. The difference in the rates at which the cavity mode and gain peak sweep with temperature indicated that it would be possible to create a device which exhibited little threshold current change with temperature by deliberately putting the peak gain of the quantum wells at a higher energy than the cavity mode. Thus, as the temperature in the active region increases, the gain peak should move into alignment with the cavity mode, resulting in efficient operation over a large temperature range. Figure 5 shows the effect of stage temperature changes on drive current required to reach device threshold as well as 0.25, 0.5, 0.75, and 1.0 mW output power. All of the curves reach a minimum at about 45-50°C, evidence of the deliberate offset between the peak-gain wavelength of the quantum wells and the cavity mode wavelength. They are also relatively flat over a broad temperature range. These curves are direct evidence of the infinite T_0 effect expected in these devices.

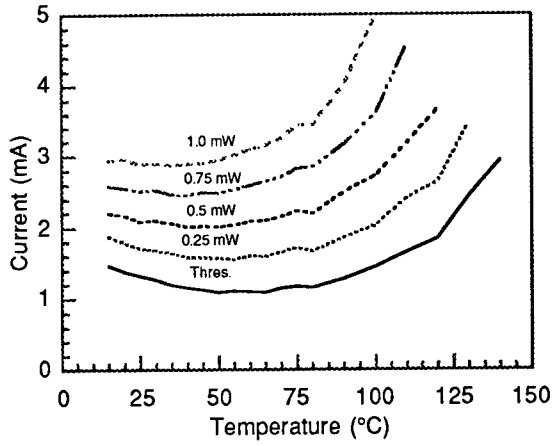


Figure 5. Device drive current as a function of temperature for an 8 μm diameter device. The solid line is the device threshold current, and the dashed lines above it are for power outputs of 0.25, 0.5, 0.75, and 1.0 mW. The curves terminate when the device can no longer reach that power level.

At room temperature, the 8μm device exhibits a CW differential efficiency of 51%. This increases linearly with device area, reaching 66% for the 20μm device due to a combination of effects including increased overlap between the carrier distribution and the optical mode. It

should be noted that the CW efficiency measurements are somewhat exaggerated due to the drop in the threshold current with temperature between 15 and 40°C. Effectively, as the device heats during the CW test, the threshold current drops, causing an increase in the output power observed. This exaggerates the output power at a given current over the equivalent stage temperature pulsed measurement, where the stage temperature would more closely approximate the junction temperature.

B. AlGaAs/GaAs and Parabolically Graded Mirrors

A good practical measure of laser performance is the power-conversion, or wall-plug, efficiency, defined as optical power out divided by electrical power in. High efficiency operation is particularly important for VCSELs to limit the amount of heating caused by wasted power. The performance of the structures discussed above is seriously limited by the voltage drop and the corresponding heating caused by the top p-type AlAs/GaAs DBR mirror. The high resistance of the AlAs/GaAs mirror stack is mainly due to the high energy barriers between the GaAs and AlAs.

Using $Al_{0.67}Ga_{0.33}As$ instead of AlAs[19] lowers the voltage barrier at each interface so that, at equal optical reflectivity, the overall resistance of an $Al_{0.67}Ga_{0.33}As$/GaAs mirror (30 periods) is about half that of an AlAs/GaAs stack (18 periods), even though more mirror periods are necessary to achieve the same reflectance. Also, we have explored the use of grading the composition parabolically at the mirror interfaces[6,20]. If this is combined with a non-uniform doping, the valence band can be made nearly flat through the mirror heterobarriers.

Figure 6 shows a comparison of five different MBE grown wafers that were processed into VCSELs as shown in Fig. 1. Both AlAs/GaAs and $Al_{0.67}Ga_{0.33}As$/GaAs p-type DBRs, and both linear and parabolic interface grades are considered. Also, the different offsets between the mode and the gain peak at room temperature are labeled. The AlGaAs/GaAs samples have slightly lower differential efficiencies (due to extra free-carrier loss in the thicker DBR mirrors) but the reduced heating allows for higher powers before the characteristic rolls over. Note that the samples with a larger gain peak to mode offset generate higher CW output powers due to the better operation at elevated junction temperatures caused by self heating. However, the gain offset does give a slight threshold penalty.

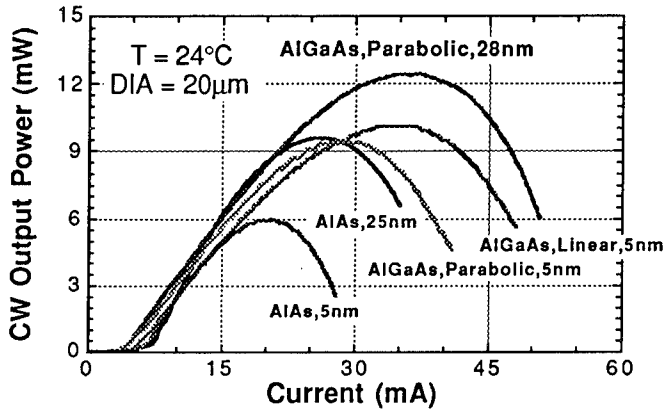


Figure 6. Comparison of both previously reported and recently modified VCSEL structures. CW output power vs. current characteristic demonstrating higher power levels for gain offset, parabolic grading, and $Al_{0.67}Ga_{0.33}As$/GaAs DBR mirrors, with the best results obtained by using a combination of all three.

The real advantage of implementing $Al_{0.67}Ga_{0.33}As$/GaAs p-type DBR mirrors can be seen from Fig. 7. As can be seen in part (b) the threshold voltages of all sizes through 40 μm occur well below 2 V, and the bias voltage remains below 3.5 Volts over the entire ranges of operation. This indicates that much lower electrical powers are required and higher wall-plug efficiencies can be attained. The maximum CW output power for these devices reaches 37.3 mW for a 70 μm device and 29.6 mW for a 40 μm device, which are records for non heat-sunk VCSELs. Figure 8 shows the peak wall-plug efficiencies for the same devices, indicating a record wall-plug efficiency of 17.3% for a 30μm device.

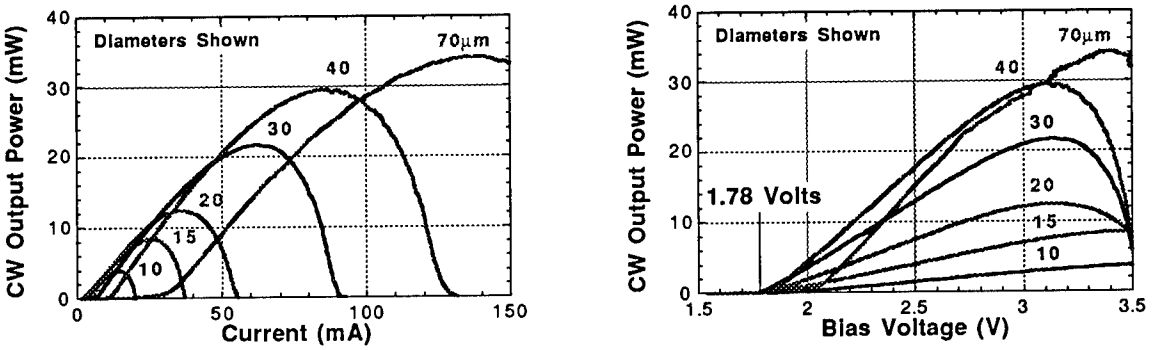


Figure 7. Comparison of characteristics of $Al_{0.67}Ga_{0.33}As$/GaAs parabolically graded VCSELs with various diameters.

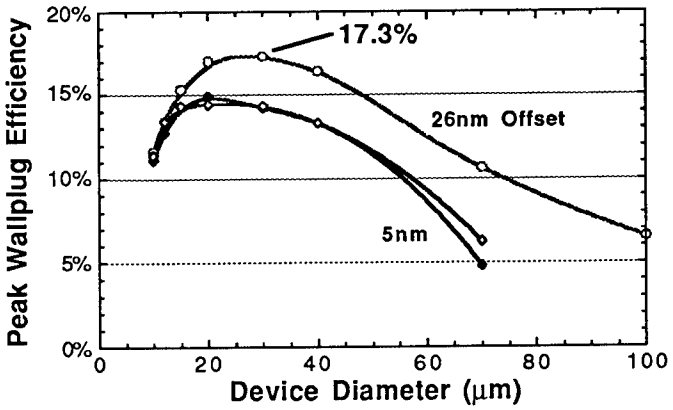


Figure 8. Power-conversion (wall-plug) efficiencies obtained for these devices reach 17.3%. The gain-offset VCSEL exhibits higher wall-plug efficiencies.

Finally, smaller devices were fabricated as outlined above from the same material as for Figs. 7 and 8. As indicated in Fig. 9 the devices operate cw with diameters down to 2 μm. Also, the 6μm diameter VCSEL demonstrated 2 mW single-mode output power with 35dB second-order mode-suppression.

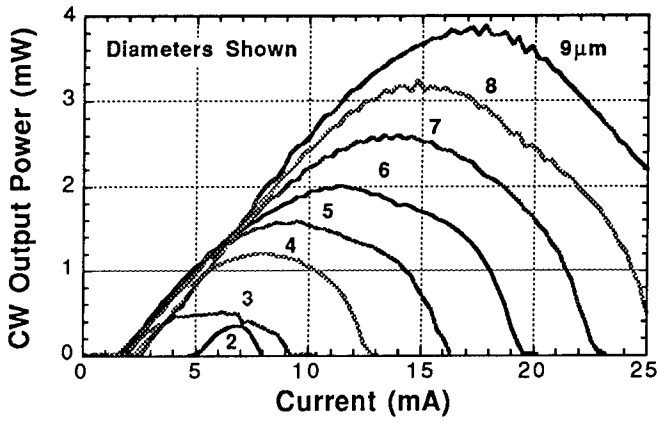


Figure 9. Output power vs. current for smaller devices from the same materials as in Figs. 7 and 8.

III. RECENT RESULTS ON ASYMMETRIC FABRY-PEROT MODULATORS

Asymmetric Fabry-Perot modulators (AFPMs) could serve as the optoelectronic link in high-speed smart-pixel interconnection and switching systems. They are interesting principally because their surface-normal configuration makes large arrays possible and their low voltage swings facilitate integration with electronics. We have designed and fabricated AFPMs with good dc operating characteristics (20 dB contrast, 1.5 dB insertion loss) which also are capable of operating in the mm-wave regime. Previously we had demonstrated large-signal modulation of these devices at 20 GHz and predicted f_{3dB}s approaching 40 GHz[21]. Here we present measurements attesting that these AFPMs do indeed roll off at 37 GHz, a speed far higher than that of any other transverse modulators to date [22,23], and competitive with even the fastest travelling-wave modulators. These high-speed AFPMs should open possibilites of smart interconnection and switching systems with extremely high aggregate bit rates.

ASFPs have much the same structure as the VCSEL depicted in Fig. 1. However, the top (output) mirror typically has a reflectance ~50% rather than ~99%. From our earlier work[21,24] we have concluded that the modulation speed of AFPMs seems to be *RC*-limited (at optical intensities well below the exciton saturation point). At higher optical intensities, the bandwidth is transit-limited because of space-charge effects in the MQW material. To test whether the situation is the same for large-signal modulation, we designed a layer structure with a wide intrinsic region which minimizes the device capacitance while maintaining high-contrast, low-voltage-swing operation [25]. Then 16 x 20 mm diodes with low resistance (≈100 Ω) and capacitance (≈25 fF) were fabricated. At dc, these devices give 20 dB contrast and 1.5 dB (3 dB) insertion loss for a 6 V (4 V) swing about a 12 V dc bias at λ = 864 nm. Large-signal, swept-frequency modulation of these devices to 20 GHz was then demonstrated using a cw Ti:Al_2O_3 laser and a high-speed photoreceiver [25].

More recently, a mode-locked dye laser has been used to illuminate the modulator with a comb of frequencies separated by the repetition rate of the laser (80 MHz) as shown in Fig. 10. The device was then driven with a large mm-wave signal (up to ≈10 dBm, corresponding to ± 2 V at the device) at a selection of frequencies between dc and 40 GHz, each separated by 10 kHz from one of the laser-generated comb frequencies. The resulting IF signal is detected by a low-noise, high-gain detector and fed to a spectrum analyzer. As shown in Fig. 11, the calibrated

response shows the expected 3dB bandwidth of 37 GHz for equivalent cw incident optical powers less than about 100 μW, and a transit-limited 18 GHz for higher optical intensities. This frequency response rivals even the fastest III-V [26] travelling-wave devices and should make possible extraordinarily high-aggregate-bandwidth space-multiplexed photonic switching systems.

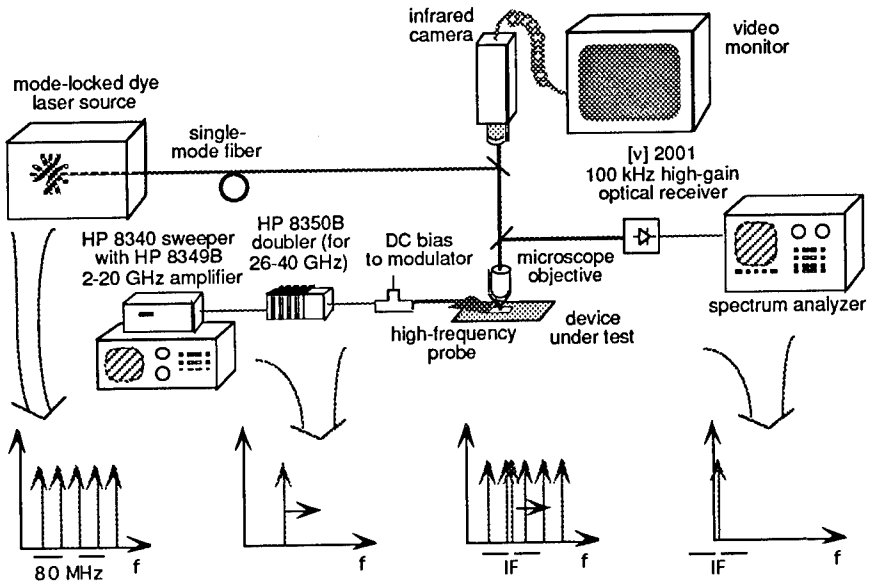


Figure 10. Large-signal 40-GHz measurement apparatus.

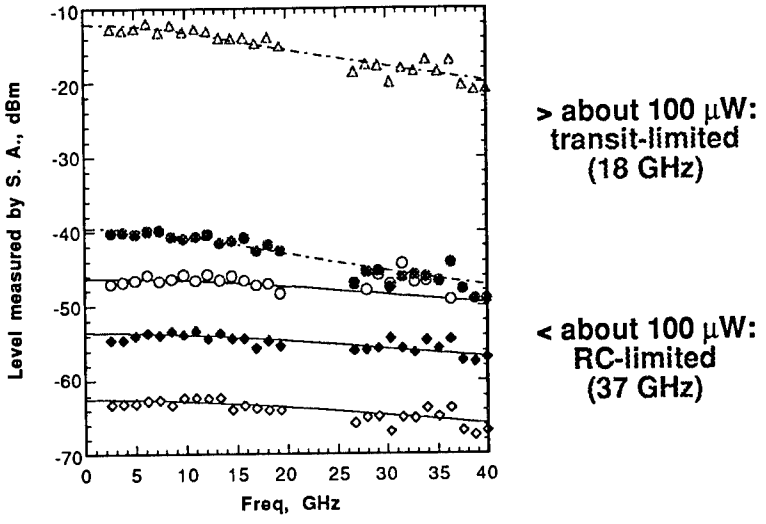


Figure 11. Large-signal measurements of 16 x 20-mm modulators from 2 - 40 GHz with integrated probe pads at various optical power levels (approximately 10 μW, 50 μW, 85 μW, 110 μW, and 3 mW equivalent cw power

incident on the device, respectively). Solid curves show calculated roll-off with f_{3dB} =37 GHz (RC limit); dashed curves are for f_{3dB} =18 GHz (transit limit).

1. J.A. Lott, R.P. Schneider, *CLEO '93*, paper CPD2 (1993).
2. D.B. Young, et al., *IEEE Photon. Tech. Lett.*, **5**, pp. 129-132, (1993).
3. K.L. Lear, S.A. Chalmers, K.P. Kileen, J. Zolper, *CLEO '93*, paper CTuD2 (1993).
4. C.J. Chang-Hasnain, et al., *CLEO '93*, paper CTuM3 (1993).
5. J.M. Catchmark, et al., *CLEO '93*, paper CTuM4 (1993).
6. M.G. Peters, et al., *CLEO '93*, paper CTuM2 (1993).
7. T. Uchida, et al., *Elec. Lett.*, **28**, pp. 550-551 (1992).
8. J.J. Dudley, et al., *Appl. Phys. Lett.* **61**, 3095 (1992).
9. T. Baba, et al., *OSA Topical Meeting on Quantum Optoelectronics,* paper no. PD2, (1993).
10. S.W. Corzine, et al., *IEEE J. Quantum Elec.*, **25**, pp. 1513-1524 (1989).
11. J.L. Jewell, et al., *CLEO '93*, paper CTuM1 (1993).
12. R. S. Geels, et al., *IEEE J. Quantum Electron,* **27**, pp. 1359-1367 (1991).
13. F. H. Peters, et al., *13th International Semiconductor Laser Conference*, paper no. PD-1, pp. 1-2 (1992).
14. M. G. Peters, et al., *Proc. of OFC '93*, San Jose, CA paper no. WH2 (1993).
15. J. W. Scott, et al., *J. of Quantum Electronics*, **29,** (5), 1295-1308, (May, 1993).
16. S. W. Corzine, et al., *LEOS Summer Topical Metting on Integrated Optoelectronics* paper THB5 (1992).
17. R. S. Geels, et al., *IEEE J. Quantum Electronics*, **29**, (1993).
18. D.B. Young, et al., *J. of Quantum Electronics*, Special Issue, (June, 1993).
19. J. D. Walker, D. M. Kuchta, and J. S. Smith, *Appl. Phys. Lett.*, **59**, pp.2079-2081 (1991).
20. M.G. Peters, et al., *Quantum Optoelectronics Topical Meeting.*, paper no.QThA4, (March 17-19, 1993).
21. C. C. Barron *et al, IEEE Photon. Tech. Lett.*, **4**:459-61 (1992).
22. G. D. Boyd *et al, Applied Physics Letters,* **57**:1843-5 (1990).
23. R. J. Simes *et al, IEEE Photon. Tech. Lett.*, **3**:513-5 (1991).
24. C.C. Barron, et al., *51st Device Research Conference,* post-deadline paper no. VIB-9, (June 21-23, 1993).
25. C. C. Barron *et al,* to appear in *Optical and Quantum Electronics* (special issue on e-o mods, 1993), and the refs therein.
26. R. G. Walker, *IEEE J. of Quantum Electronics*, **27**:654-67 (1991).

SILICON GERMANIUM HETEROBIPOLAR TRANSISTOR FOR HIGH SPEED OPERATION

E. Kasper and A. Gruhle

Daimler-Benz AG, Research Center Ulm, Wilhelm-Runge-Str. 11, D-89081 Ulm, Germany

ABSTRACT

The high frequency dilemna (limited high frequency operation only possible with high base sheet resistivities) of silicon bipolar junction transistors can be overcome by the application of heterostructures. In a first part we give an overview about the SiGe base heterobipolar-transistor (SiGe-HBT) concept and about the devices realized so far. In the second part we report about our SiGe-HBT device work based on Si-MBE grown structures. Experimental results obtained include transit frequencies up to 100 GHz, maximum oscillation frequencies up to 65 GHz, low base sheet resistivities (0.7 kΩ/□ to 3 kΩ/□), encouringing noise properties (< 2 dB at 10 GHz) and successful tests in 24 GHz oscillators.

I. INTRODUCTION

The silicon bipolar junction transistor (Si-BJT) is limited in its speed by the base properties (lower limit for the base width w_B, upper limit for the base doping N_B). The principal (punch through, tunneling) and technological limits (doping fluctuations, pinch resistivity) are given in Fig. 1 (calculated for flat doping profiles, but similarily valid for general profiles). Punchthrough of the reverse biased collector-base junction occurs for low base sheet dopings:

$$(N_B w_B)^2 \le \frac{2\hat{\varepsilon}}{e} N_C (V_{BC} + V_d) = 2.25 \cdot 10^{32} / m^4 \qquad (1)$$

$N_B >> N_C$ base and collector doping, w_B base width, V_{BC}, $V_d \cong 1$ V; voltages (base collector and diffusion voltage, respectively), $\hat{\varepsilon}$ (10^{-10} As/Vm) dielectric constant for Si, e electron charge. The numerical values are given for $N_C = 4.5 \cdot 10^{16} cm^3$, $V_{BC} = 3$ V. To avoid complete punchthrough the base sheet resistivity $R_\square$ has to be below (weighted by the mobility μ):

$$(R_\square \mu) \le 2e \hat{\varepsilon} N_C (V_{BC} + V_d) = 400 m^2 / As \qquad (2)$$

The other limit for the base doping is given by tunneling of the emitter-base junction at roughly $N_B = 5 \cdot 10^{18}/cm^3$ ($N_E >> N_B$). The constraints from technological reasons may be even stricter. Consider a thin base with statistically distributed acceptors. Fluctuations and local punchthrough may occur when the number N of acceptors within a cube of length w_B is to small. In Fig. 1 the

0-7803-0894-8/93 $3.00 © 1993 IEEE

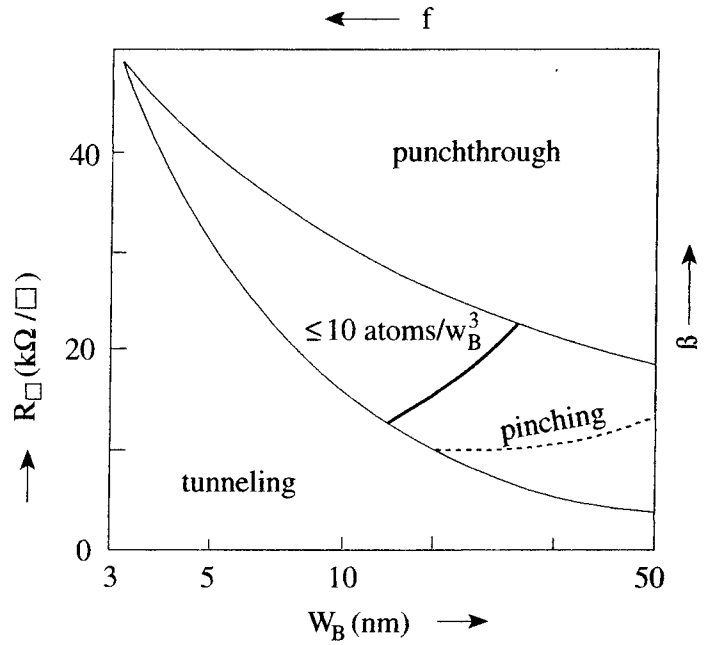


Fig. 1:
Base sheet resistivity $R_\square$ versus base width w_B for Si-BJT with flat profiles.

limit for $N = N_B\, w_B^3 = 10$ acceptors is shown. Outside of the active emitter area where the emitter-base junction is not forward biased the space charge layer from this junction may pinch the connection between base contact and inner base. The pinch resistivity increases to infinity for

$$N_B w_B^2 = \frac{2\hat{\varepsilon}}{e} V_d \cong 1.25 \cdot 10^9 m^{-1} \qquad (3)$$

or

$$R_\square \mu = w_B / (2\hat{\varepsilon} V_d) \cong 5 \cdot 10^9 w_B \left(m^2 / As\right) \qquad (4)$$

The need for a certain current gain β which can be fulfilled only with high emitter dopings $N_E >> N_B$ and the above given base layer limitations set the BJT in a unfavorable position compared to the hetero transistor (HBT) with its much larger freedom for the layer design.

II. SiGe DOUBLE HETEROSTRUCTURE BARRIER

Many of the design limitations of the BJT can be solved by a heterojunction emitter with a larger bandgap than the base material (band gap difference ΔE_g). The current gain β is increased by a factor exp (ΔE_g/kT) allowing a highly doped base together with a low doped emitter. The tunneling limit can therefore be overcome by an emitter doping below $5 \cdot 10^{18}/cm^3$, the acceptor fluctuations and the base pinching are strongly reduced by the high base doping up to $10^{20}/cm^3$. The best choice of a wide gap emitter material on Si would be SiC, but the extremely large lattice mismatch and technological problems will shift a broad application to an undefined future. Within the past few years the SiGe/Si material system has gained considerable progress which made the

realization of a slightly different HBT concept attractive. Instead of a wide band gap emitter a small band gap base (SiGe) is provided which results in heterostructure barriers both at the E-B and the B-C interface (DHBT double heterobarrier bipolar transistor). Very fortunately the band offsets of strained SiGe on Si (Fig. 2) meet the demands of the n-Si/p-SiGe/n-Si transistor nearly

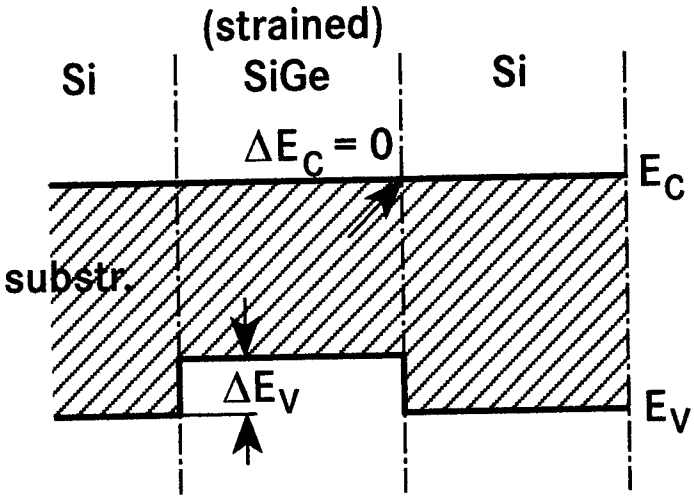


Fig. 2:
Typical band alignment for strained SiGe on Si (unstrained)

ideally. The band gap difference ΔE_g is mainly on the valence band side ($\Delta E_V \cong \Delta E_g$) with near flat band conditions on the conduction band side ($\Delta E_C << \Delta E_g$). The bandgap difference SiGe/Si is naturally dependant on the Ge content x of the $Si_{1-x}Ge_x$ alloy with ΔE_g about 150 meV for 20% Ge-content. Compared to the BJT the collector current I_C of the HBT is increased heavily

$$I_c = I_{co} \exp \frac{V_{BE}}{V_T} - 1 \qquad (5)$$

$$I_{co} = \frac{e\, D_n\, n_i^2(SiGe)}{N_B \cdot w_B} \qquad (6)$$

because the intrinsic carrier density n_i is much higher in SiGe than in Si (roughly by $n_i^2 \sim \exp(\Delta E_g/kT)$). The forward biased B-E junction emitts electrons as excess minority carriers into the base. On the emitter side the electron density is enhanced by exp (V_{BE}/V_T) as shown in Fig. 3. The strain splits the sixfold degenerate conduction band into twofold (electrons flying perpendicular to the interface) and fourfold (electrons moving in plane) states. Also the heavy and light hole states are splitted. This significant change in band structure and the overlap with heavy doping effects prevent at the moment an exact prediction and analysis of the currents although the basic mechanism are now understood. Often data from unstrained material are used which could result in uncertainties of at least a factor two.

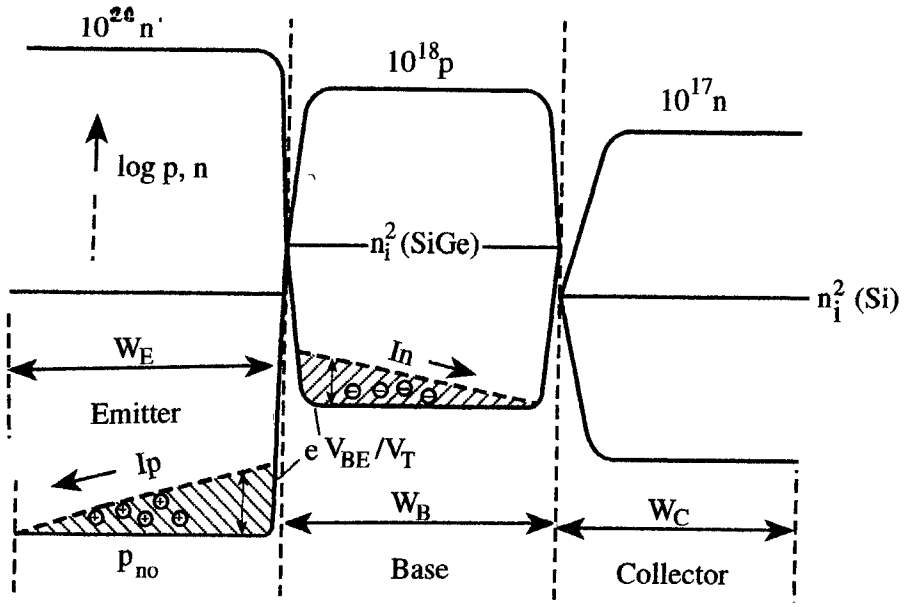


Fig. 3:
Carrier concentrations (majority, minority) within a typical HBT structure for equilibrium ($V_{BE} = 0$) and under current ($V_{BE} > 0$)

The Ge lattice constant is 4.2 % larger than that of Si. As a first approximation one can assume a linear relationship between Ge content x of the alloy and the lattice mismatch η (Vegard's law):

$$\eta \cong 0.042\ x \qquad (7)$$

For exact calculations the slight deviation [1] from Vegard's law has to be considered

$$\eta = 0.042\ x\ [1 - 0.12\ (1-x)] \qquad (8)$$

A lot of the progress now made with SiGe and other strained layer materials is based on the fact that with modern epitaxy techniques (low temperature process) material with good crystal quality can be grown to much higher layer thicknesses [2] than earlier predicted by equilibrium theories [3]. Equilibrium theories predict strained layer growth up to a critical thickness which is about 10 nm for $Si_{0.75}Ge_{0.25}$ on Si. Above the critical thickness misfit dislocations which are undesired from the viewpoint of device physics let the layer strain partly relax. By growth at low temperatures (e.g. 550 °C) the critical thickness can be increased to 100 nm (metastability). It should be mentioned that the Si-cap (emitter) grown on top of the SiGe-layer further stabilize the structure.

III. DEVICE REALIZATION

Within a short time an astonishingly rich variety of technological solutions for the SiGe-HBT was developed. We try it to classify by the following scheme (Tab. 1). One criterion of the cassification is if inversion of the doping levels ($N_B > N_E$) is obtained (true HBT). Especially with strong Ge grading (low Ge content at the B-E junction) the inversion cannot be realized. Other criteria involved include selective growth, growth of the complete structure in one run, mesa type device morphology and postepitaxial processing temperatures.

Table 1: Technical SiGe-HBT solutions. Provisional classification scheme. See text.

Ref.	Doping Inversion	Selective Growth	One Epi Run	Planar (Non Mesa)	Max. Process T
/4/	-	-	-	+	conv.
/5/	+	-	-	+	880 °C, 15 s
/6/	?	-	-	+	800 °C, 10 s
/7/	-	+	-	+	950 °C
/8/	+	-	+	-	900 °C
/9/	+	-	+	(-/+)	800 °C
/10/	+	-	+	-	-

With these different approaches excellent high frequency results were obtained with f_T ranging from 40GHz to 110 GHz, with f_{max} up to 65 GHz. A rather large spread is seen in base sheet resistivities (0.7 kΩ/□ - 17 kΩ/□) with essential improvements compared to BJT only for layer designs with doping level inversion.

IV. MBE GROWTH IN ONE RUN

In the following we report about our own results based on MBE grown material. The complete active structure (collector, base, emitter, emitter contact) was grown in one run. The basic idea behind the "one run concept" is to outbalance the additional efforts for the HBT by a reduction of process steps. The MBE equipment [11] used (Si-MBE B) consists of two UHV-chambers (storage chamber with 10 wafer magazine and the growth chamber) both connected by a transfer mechanism.

Only subsystems proven to be necessary for the process (design principle of highest simplicity) are installed inside the growth chamber. These include the material sources (elemental Si, Ge, B, Sb), the substrate heater, wafer holder with voltage supply (for doping by secondary implantation DSI) and in situ monitoring instruments used for process control. The growth temperature was decreased from 650 °C at the collector side to 325 °C at the emitter contact side. Short anneals were performed before (900 °C) and after growth (600 °C). The collector (130 nm-250 nm) is Sb-doped (typ. $2 \cdot 10^{17}/cm^3$), the following SiGe-layer (20 % Ge - 30 % Ge) is highly doped with B (typ. $6 \cdot 10^{19}/cm^3$) cladded by thin (1-10 nm) intrinsic layers at the interfaces. The 70 nm thick emitter is only doped to $2 \cdot 10^{18}/cm^3$ (Sb) followed by the heavily doped ($2 \cdot 10^{20}/cm^3$, Sb) emitter contact. The chemical profile and the strain status of the multilayer structure are investigated by SIMS and X-ray diffraction (XRD).

V. HIGH FREQUENCY RESULTS

Mesa type test transistors (Fig. 4) for on wafer measurements were fabricated without any additional medium or high temperature processes [12]. Gummel plot measurements in normal and inverse direction were performed and the ideality factors n of the collector currents were very carefully determined. Parasitic barriers at the E-B junction [13] as created by outdiffusion would

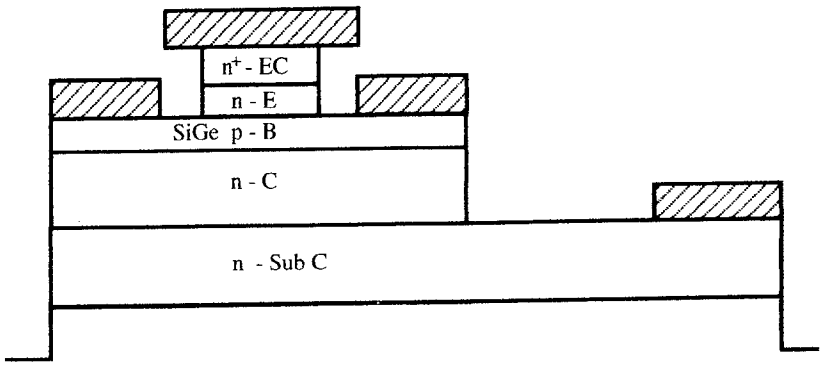


Fig. 4:
Scheme of the mesa type HBT

increase the ideality factor n from 1.0 to up to 1.1. Parasitic barriers at the B-C junction can be seen in inverse operation by an increase of the inverse ideality factor $n_{(i)}$. It can be shown that a simple relationship exists between the Early voltage V_A and the inverse ideality factor

$$V_A = V_T \left(\frac{n_{(i)}}{n_{(i)} - 1} \right) \tag{9}$$

The inverse Early voltage and the normal ideality factor are related in the same way. In our experiments we found out that these electrical measurements are more sensitive to boron outdiffusion (1-2 nm) than SIMS analysis (3-5 nm). Furtheron it seems that the SIMS boron profile is systematically shifted (~ 5 nm) to higher depths (knock on effect of the analysing ions?). Guided by these electrical measurements we have chosen the thicknesses of the intrinsic cladding layers to be 1-2 nm (emitter side) and 10 nm - 15 nm (collector side), respectively. As a next step we have varied the thicknesses of the SiGe layers from 50 nm to 25 nm. The associated decrease of the base transit time τ_B

$$\tau_B = \frac{w_B^{\,2}}{2D} \tag{10}$$

resulted in an increase of the transit frequency f_T (Fig. 5) from 40 GHz to slightly above 100 GHz. Together with an IBM group [6] we demonstrated the operation of a Si-based transistor at the 100 GHz region. In our case this was obtained - due to the superior base sheet resistivities (0.7kΩ/□ - 3 kΩ/□) - with a simple process and relaxed optical lithography (1 μm emitter contact line). Because of the simple process (no outer base, large collector area) the best f_{max} values obtained with 30 nm SiGe layers are significantly below 100 GHz (best f_{max} = 65 GHz). In spite of that this f_{max} value is to our knowledge also a record for SiGe-HBT's.

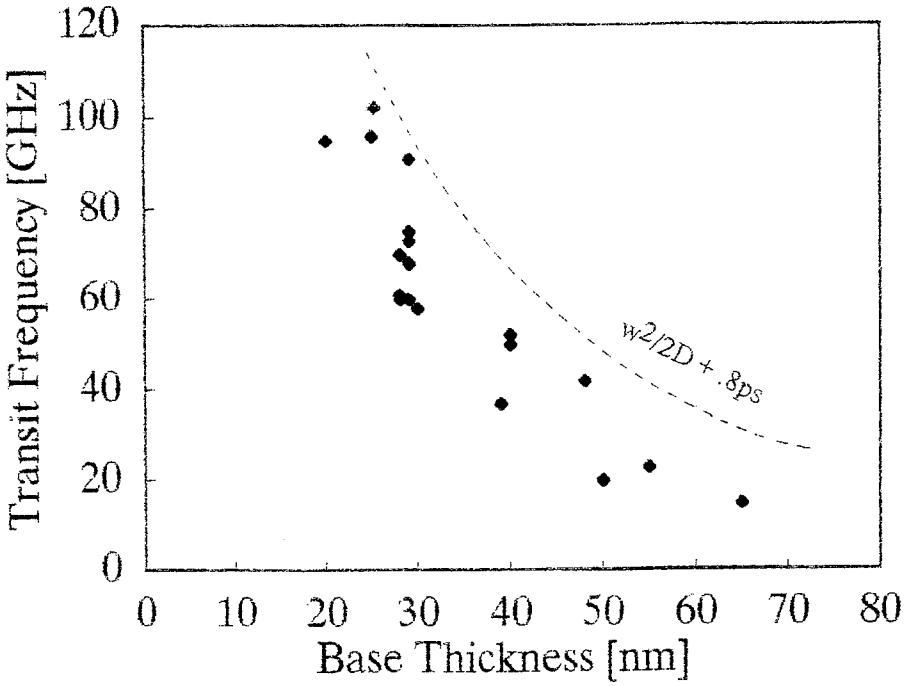


Fig. 5:
Transit frequency f_T versus base thickness (SiGe)

VI. OUTLOOK

Although we believe that by further systematic layer design optimization (base doping, emitter thickness, collector doping) the transit frequency can be increased to 130 GHz - 150 GHz the main efforts in transistor development will shift to increase the f_{max} values to at least up to f_T. The need in process technology is to combine the simple process scheme with standard silicon planar processing [8]. First attempts with our epitaxial material on a production line (Telefunken microelectronic - TEMIC) confirmed our optimistic view with f_T = 40 GHz tran-sistors. Integration started with small scale monolithic circuits (first demonstration of a multi-plexer by our partner Ruhr University Bochum, H.-U. Schreiber) and hybrid integrated microwave oscillators [14] for 10 GHz and 24 GHz. Very encouraging noise measurements [15,16] promise attractive HBT solutions for mobile communication (1-10 GHz). The relaxed production requirements (1 µm lithography) should lead to an attractive economic perspective for SiGe-HBT IC's.

ACKNOWLEDGEMENT

Cooperation with our collegues at Daimler-Benz Center and with the groups of H. Schumacher (Ulm), B. Bosch (Bochum), J. Graffeuil (Toulouse), R. Kaiser (Heilbronn) is gratefully acknowledged. Part of the work was sponsored by the German Ministry of Research and Technology.

REFERENCES

[1] I.P. Dismukes et al., *J. Phys. Chem.* 68, pp. 3021 (1964)
[2] R. People and J.C. Bean, *Appl. Phys. Lett.* 49, pp. 229 (1986)
[3] I.W. Matthews and A.E. Blakeslee, *J. Cryst. Growth* 27, pp. 118 (1974)
[4] G.L. Patton et al., *IEEE-EDL* 11, pp. 171 (1990)
[5] I.N. Burghartz et al., *Symp. VLSI-Techn.*, 5B-1 (1993)
[6] E. Crabbé et al., *Device Research Conf.*, 1993
[7] F. Sato et al., *Techn. Dig. IEDM*, 1993
[8] H.-U. Schreiber, *Electronics Lett.* 28, pp. 485 (1992)
[9] E.I. Prinz et al., *Techn. Dig. IEDM*, 1989
[10] A. Gruhle et al., *Electronics Lett.* 29, pp. 415 (1993)
[11] E. Kasper et al., *Proc. Solid State Devices and Materials, SSDM 93*
[12] A. Gruhle et al., *IEEE-EDL*
[13] A. Gruhle, submitted to *IEEE-EDL*
[14] U. Güttich et al., *Conf. Proc. MIOP*, pp. 146 (1993)
[15] H. Schumacher et al., *Electronics Lett.* 28, pp. 1167 (1992)
[16] R. Plana et al., *ESSDERC* 1993

RECENT ADVANCES IN SILICON CARBIDE HIGH POWER MICROWAVE AND HIGH TEMPERATURE ELECTRONICS

R.C. Clarke, R.H. Hopkins, C.D. Brandt, M.C. Driver, D.L. Barrett, A.A. Burk, G.W. Eldridge, H.M. Hobgood, J.P. McHugh, P.G. McMullin, R.R. Siergiej, S. Sriram,Westinghouse Science Technology Center.

PAPER UNAVAILABLE FOR PUBLICATION

0-7803-0894-8/93 $3.00

Luminescence from Si-Based Materials and Devices

J. C. Campbell, K.-H. Li, and C. Tsai

Microelectronics Research Center
University of Texas
Austin, TX 78712

ABSTRACT

While numerous luminescence effects have been reported for Si-based materials, with few exceptions the efficiencies have been extremely low. The observation of strong, room-temperature, visible emission from porous Si has renewed interest in and stimulated research on light emission from column IV materials. This paper reviews some of the accomplishments that may presage the development of Si-based light emitters.

I. INTRODUCTION

The realization of practical, low-cost Si-based emitters would impact several key technologies including optoelectronic integrated circuits, optical memories and logic, and advanced display systems. Several mechanisms that give rise to luminescence in column IV materials have been documented. Electroluminescence in bulk crystalline Si was observed many years ago in forward [1] and reverse biased p-n junctions. Under forward bias the luminescence arises from band-to-band phonon-assisted recombination. This weak emission (quantum efficiency ~ 10^{-4}%) peaks in the near infrared. In reverse bias, near avalanche breakdown, the luminescence is visible, the efficiency is poor, and its origin is still the subject of controversy. Numerous techniques have been used to improve the luminescence efficiency column IV materials. These approaches can be broadly categorized as (1) elimination of crystallinity, (2) introduction of defects, or (3) the creation of nanostructures. Some of the non crystalline materials such as polysilanes [3], hydrogenated amorphous Si [4,5], and siloxene [6,7] have yielded efficient luminescence but at the price of good electrical transport and thermal stability. The motivation for introducing defects is to confine carriers in coordinate space which creates dispersion in momentum space thereby enhancing no-phonon recombination. Some success has been achieved with isoelectronic isolated impurities and impurity complexes and with rare earth transition metal impurities. A large number of isoelectronic traps have been identified in Si and some, such as carbon, have exhibited efficient luminescence at 77K but in all cases the luminescence is negligible at room temperature [8]. The most successful rare earth ion has been Er [9,10]. An electroluminescence efficiency of 0.05% was achieved at 77K but the relatively low solubility of Er in Si combined with the low radiative transition probability place severe constraints on this approach [11]. Si-based nanostructures have also been the subject of many diverse research efforts. This approach is reminiscent of bandstructure engineering in III-V compounds. Micro-crystalline structures [12-17] and $Si_{1-x}Ge_x$/Si quantum wells [18-23] have shown potential but have not achieved the necessary efficiencies at room temperature. Recently, however, porous Si has exhibited efficient visible, room temperature photoluminescence (PL) [24]. Electroluminescence has also been reported [25-27], but, to date, efficiencies have been well below the PL results and the primary research thrust has been

0-7803-0894-8/93 $3.00

the identification of the physical mechanisms responsible for the luminescence. In this paper we review some of the recent work on Si-based nanostructures.

II. NANOSTRUCTURES

Several column IV nanostructures have exhibited luminescence. In one of the first reports of quantum size effects in a Si structure, Dimaria et al. [15] showed that under certain fabrication conditions off stiochiometric SiO_2 contained small Si islands or microcrystals of average diameter ~ 20Å in a matrix of SiO_2. These microcrystals exhibited characteristics consistent with three dimensional quantum confinement. It should be noted that the authors also postulated that the luminescence might originate at the interface between the Si microcrystals and the SiO_2 host material. More recently, Takagi et. al. [16] showed that the dimensions of the suspended Si islands in SiO_2 can be controlled through the time and temperature of a post oxidation process. Consistent with the quantum confinement model, a reduction in the size of the crystallites caused the photoluminescence to shift to higher energies. Similar behavior for Ge microcrystals in SiO_2 has been reported [14]. The samples were fabricated by co-sputtering Ge and SiO_2. A post annealing step resulted in the formation of Ge microcrystals and the onset of photoluminescence. In each of the cases of microcrystals in SiO_2 the quantum efficiency was ~ 10^{-3} to 10^{-4}%.

Higher photoluminescence efficiencies have been achieved with hydrogen terminated microcrystals. These small Si crystallites have been produced by decomposition of SiH_4 and H_2 with a microwave plasma [16], rf sputtering in a hydrogen gas [12], and eximer laser photolysis of disilane (Si_2H_6) [17]. These diverse fabrication techniques yield structures consisting of nanometer-scale crystalline Si cores with H termination at the surface, possibly in the form $(SiH_2)_n$. Furukawa et al. [12] observed visible photoluminescnece from Si microcrystals that were fabricated by low-temperature sputtering in H. Using Fourier transform infrared spectroscopy to determine the H content and optical transmission to measure the absorption coefficient, they showed that the optical bandgap increased with H content. They concluded that the luminescence originiated in "... small crystalline silicon particles surrounded by hydrogen atoms, which are bonded in the dihydride form." Heath and Jasinski [17] have produced similar hydrogen-terminated Si "balls" using eximer laser photolysis of disilane. It is also worth noting that the wavelength of the PL peak for these particles was not a function of size but rather of crystallinity . It is well known that H a:Si luminescences in the near infrared and the visible. Wollford et al. [3] showed that the optical gap increases with increasing H content and that the higher hydrogen levels also have the fewest dangling bonds. Collectively, these results have raised several issues regarding the role of H in the luminescence process, i.e., does H simply provide an excellent passivation to the microcrystals thus reducing nonradiative surface recombination or does it play a more active role in altering transition levels.

Compared to the structures discussed above, quantum-well devices have the advantages of uniformity, flexibility, and more control over dimensions. It has been suggested that efficient, direct bandgap recombination could be achieved with zone-folded, short-period $Si_{1-x}Ge_x/Si$ superlattices [28]. To date, however, all experimental evidence for direct bandgap behavior has been from material with high dislocation densities which have known luminescence properties. Other effects

including bound and free exciton recombination and no-phonon emission give rise to luminescence in strained $Si_{1-x}Ge_x/Si$ quantum wells. Terashima et al. [29] first reported photoluminescence from exciton recombination in strained $Si_{1-x}Ge_x$ (x = 0.042) films grown by MBE. This was extended to x = 0.2 by Sturm et al. [30] using quantum wells and superlattices grown by rapid thermal chemical vapor deposition. Internal PL quantum efficiencies as high as 31% have been achieved at 4 K from strained layers grown at 400 °C by MBE and subsequently annealed at 600 °C [31]. Recently, Mi et al. [32] have achieved room-temperature electroluminescence at 1.3 μm and 1.5 μm from strained $Si_{1-x}Ge_x$ quantum wells. The internal quantum efficiency at 1.3 μm was estimated to be > $2x10^{-4}$.

III. POROUS SILICON

There is still a great deal of uncertainty regarding the origin of the luminescence from porous Si. The emission was initially modeled as two-dimensional confinement in quantum wire-like structures of highly porous, electrochemically etched Si. There has been substantial support for this interpretation. The quantum size effect accounts for the observed increases in the optical transition energies (~ 0.5 to 1.0 eV) relative to the bandgap of Si and the fact that the peak of the emission spectrum shifts to shorter wavelengths as the feature sizes are reduced by chemical dissolution [33]. These luminescence characteristics require confinement in 3 to 5 nm microstructures. Transmission electron microscopy (TEM) [34,35], Raman spectroscopy [36], and X-ray diffraction measurements [37] have confirmed the presence of Si-crystallites with these dimensions. This model also appears to explain the wavelength dependence of the luminescence lifetime after pulse excitation [38]. Finally, Halimaoui et al. [39] have shown that electroluminescence is obtained during anodic oxidation of porous Si and that the peak wavelength decreases with oxidation time. This was explained in terms of a progressive reduction in the sizes of the silicon crystallites. The efficiency of the electro-oxidized luminescence is estimated to be as high as 3%. [40]

While the quantum confinement model has successfully accounted for many of the luminescence characteristics of porous silicon, there are also difficulties. Brandt et al. [6] have pointed out that the density of silicon crystallites of the correct size to account for visible luminescence is not sufficient to yield efficiencies in excess of 1% and that the size distribution is inconsistent with the PL spectral width (~ 0.3 eV). There is some inconsistency in reports of the low-temperature behavior of the PL spectra, but it appears that the anticipated blue-shift with decreasing temperature is absent [7,41]. It is also somewhat surprising that porous silicon fabricated under a wide range of etching conditions would produce such similar luminescence.

Other hypotheses for the luminescence of porous Si include (1) the formation of wide-bandgap material such as a-SiH_x or siloxene or (2) recombination through boundry layer states. There have been reports of TEM and Raman spectroscopy that reveal amorphous regions in the porous silicon layers [35,42,43]. It has been suggested that the luminescence originates in these amorphous regions [43,44] since it has previously been shown that thin, amorphous polysilane films grown by homogeneous chemical vapor deposition exhibit efficient (>1%), room-temperature photoluminescence [9] and the surface of porous silicon is known to be terminated with hydrogen in the form of SiH and SiH_2 [45]. For this model the specific

mechanism responsible for the luminescence has not been identified but it might involve hydrogen-terminated silicon clusters with polysilane bridges.

A somewhat related model involves the formation of siloxene compounds on the surface of the sponge-like Si skeleton of the porous regions [15,16]. The crystalline structure of siloxene ($Si_6O_3H_6$) consists of six-fold Si rings which are interconnected by oxygen bridges to form a planar arrangement. Siloxene is the only Si-based compound that has a luminescence efficiency (~ 10%) as high as that of porous Si. Intense blue, green, and red luminescence has been demonstrated in siloxene layers grown on Si substrates. The PL, Raman, and IR spectra of siloxene are remarkably similar to that of porous Si. In addition, siloxene exhibits photoenhanced degradation and thermal quenching for T > 300 °C similar to that observed in porous Si [46,47]. Specific details regarding how siloxene derivatives are formed during the anodization process have not been resolved. Of particular concern is the fact that the surface of as-anodized porous Si appears to be free of O. Also, it has not been demonstrated that siloxene can be fabricated electrochemically.

Both the polysilane model and siloxene appear to be inconsistent with thermal annealing studies of the PL intensity. It has been shown that the PL is quenched for annealing temperatures > 300 °C and that the decrease in luminescence is attributable to the desorption of H from the surface [46]. Recently, however, Petrova-Koch et al. [48], have shown that the luminescence can be recovered for annealing temperatures in the range 700 °C to 1000 °C in an O atmosphere. The decrease in the H passivation near 300 °C coincides with an increase in the electron spin resonance signal which then vanishes at the higher temperatures where oxidation occurs. The picture that emerges is that the surface of the as-prepared material is passivated with H which results in a very low surface recombination velocity. As the H is thermally desorbed, the number of dangling

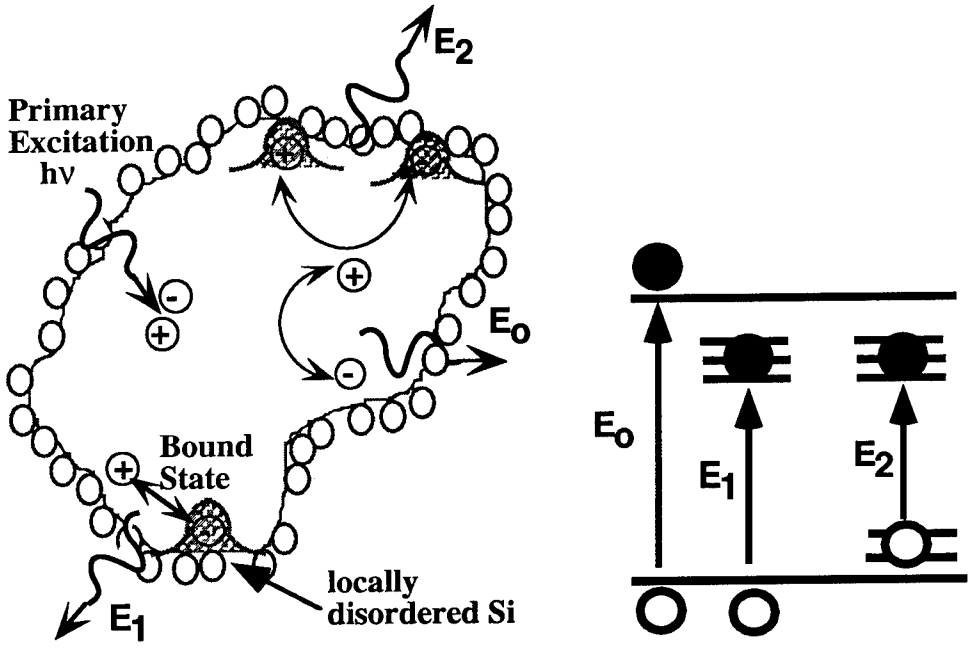


Figure 1. Surface-state model for luminescence in porous Si. Recombination can occur (1) between a photo-excited pair for which the emission energy is E_0, (2) after localization of either the electron or hole (E_1), or (3) between a bound hole and a bound electron.

bonds increases nonradiative recombination via defects and causes the PL intensity to drop. At the higher temperatures O passivation of the Si cores restores the efficiency of the radiative processes. Infrared spectroscopy of the oxidized samples shows a complete absence of H.

Recently, Koch et al. [49] have proposed a model in which absorption occurs in the core of the Si nanometer-crystallites and recombination proceeds through lower energy states at the disordered surface. This model is illustrated in Fig. 1. Owing to the nanostructure features of porous the absorption edge is blue shifted relative to bulk Si. The surface states or boundary traps are postulated to result relaxation and reconfiguration of the outermost atoms of the clusters. These surface experience changes in bond lengths and angles. Recombination can occur between the photogenerated electrons and holes in within the crystalline regions. It has been shown that this is not a strong recombination mechanism however because the transitions retain their indirect characteristics. A more probable recombination channel involves a free electron or hole and a bound carrier at the surface. The third mechanism, between bound electrons and holes, is not favored because of the weak overlap of wavefunctions. This representation is consistent with the size-related blue shifts in the PL spectra and also explains the fact that the PL peak occurs approximately 400 meV below the fundamental absorption edge of porous Si [50]. This model accounts for most of the observed size- and structure-related characteristics and accents the prominant role of surface chemistry in the luminescence process.

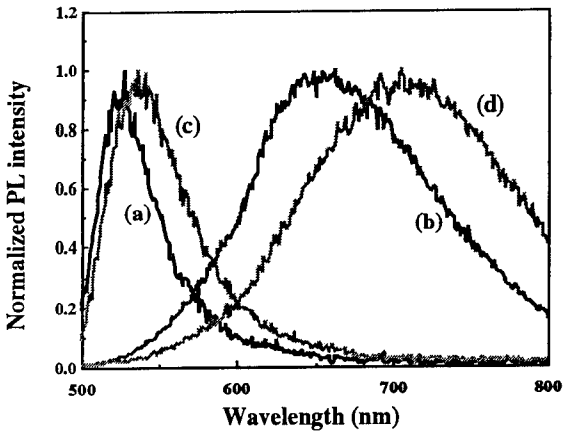


Figure 2 Photoluminescence spectra of p-type Si: (a) as-anodized sample after immersion in H_2O:acetic acid:49% HF=2:1:1, (b) blown dry, (c) immerse in H_2O:acetic acid:49% HF=2:1:1, and (d) blown dry.

The importance of surface chemistry has been underscored by evidence that the luminescence peak can be significantly altered by changing the composition of the electrolyte in which the samples are immersed. Using this approach the emission has been repeatedly cycled (> 100 times) between green and red [51]. It has also been shown that the luminescence changes when the samples are withdrawn from solution. Figure 2(a) shows the green luminescence from an as-anodized sample. The high-energy portion of this spectrum is truncated somewhat owing to the presence of a long-pass filter that was used to block the laser light. After the samples were dipped into DI water and blown dry with nitrogen, the PL changed from green to red (Fig.2(b)). However, returning the samples to a 2:1:1 solution of

water, 49% HF and acetic acid resulted in green luminescence immediately (Fig.2(c)). Finally, when the sample was again blown dry in air the spectrum shifted back to the red (Fig.2(d)). In related work Müller et al. [52] have observed that the PL peak can be tuned and shifted by chemical reactions catalyzed by the presence of light.

The excitement over the efficient luminescnece from porous Si has been primarily due to the expectation that it might be the precursor for the integration of optoelectronic devices with conventional electronics on Si chips. Electroluminescence has been reported using indium tin oxide [25] and Au [26] films on p-type porous Si layers. Both of these devices rely on the injection of high energy carriers into the porous region. Bassous et al. [27] have reported emission from anodized p-n junctions. These devices had turn-on voltages less than 1 V and the optical output varied linearly with input current. Recently, Steiner et al. [53] have achieved an increase in the quantum efficiency (10 to 100x) using a structure consisting of a nanoporous light-emitting n layer sandwiched between a mesoporous p^+-doped cap layer and a macroporous n-type layer. This configuration has efficient carrier injection into the nanoporous recombination and good current-voltage characteristics. The emission is red-orange light with a spectrum very close to the PL spectrum. Optimization of this type of structure may lead to even higher efficiencies.

Photodetectors have also been fabricated from porous Si. A porous Si Schottky barrier photodiode with a quantum efficiency of 97% and a response time of 2 ns has been reported by Zheng et al. [54] and Yu et al. [55] have fabricated a porous Si MSM photoconductor with a responsivity of 0.5 A/W. At present there are two factors that could severely restrict the utility of porous Si for device applications. The first is that aging under ambient conditions changes the luminescence characteristics and the other is that the porous region can be damaged during fabrication. For example, it etches rapidly in photoresist developer. Tsai et al. [56] have shown that rapid-thermal-oxidized porous Si is stable and can withstand conventional processing procedures. They have used the oxidized wafers to fabricate a photodiode with a quantum efficiency of 75% at 740 nm. In addition, these devices appear to have an enhanced UV response relative to bulk Si photodiodes.

IV. CONCLUSIONS

Luminescence from several Si-based structures has been reported. While most of the work to date has concentrated on fundamental materials issues, the potential for device applications is exciting and should stimulate further research.

References

[1] J. R. Haynes and H. B. Briggs, "Radiation produced in germanium and silicon by electron-hole recombination," Phys. Rev. **86**, p. 647 (1952).

[2] R. Newman, "Visible light from a silicon p-n junction," Phys. Rev. **100**, pp. 700-703 (1955).

[3]D. J. Wolford, J. A. Reimer, and B. A. Scott, "Efficient visible photoluminescence in the binary a-Si:Hx alloy system," Appl. Phys. Lett. **42**, pp. 369-371 (1983).
[4] R. A. Street, in *Hydrogenated Amorphous Silicon*, edited by J. I. Pankove, Vol. 21B of Semiconductors and Semimetals (Academic, Orlando, 1984), pp. 194-244.
[5] M. Hopkinson, T. M. Searle, I. G. Austin, and A. J. Rhodes, "A comparison of photo- and electroluminescence in a-Si p-i-n junctions," J. Non-Crystalline Sol. **77&78**, pp. 695-698 (1985).[4] F. Boulitrop, A. Chenevas-Paule, and D. J. Dunstan, "Luminescence and magnetic resonance in post-hydrogenated microcrystalline silicon," Solid State Comm. **48**, pp. 181-184 (1983).
[6] M. S. Brandt, H. D. Fuchs, A. Höpner, M. Rosenbauer, M. Stutzmann, J. Weber, M. Cardona, and H. J. Queisser, "Visible luminescence from silicon: quantum confinement or siloxene?" Materials Research Society Symposium, San Francisco, CA, 1992.
[7] M. Stutzmann, J. Weber, M. S. Brandt, H. D. Fuchs, M. Rosenbauer, P. Deak, A. Höpner, and A. Breitschwerdt, "Visible luminescence from silicon," Adv. Solid State Phys. **32**, to be published.[5] Y. Maeda, N. Tsukamoto, and Y. Yazawa, "Visible photoluminescence of Ge microcrystals embedded in SiO_2 glassy matrices," Appl. Phys. Lett. **59** pp. 3168-3170 (1991).
[8] G. Davies, "The optical properties of luminescence centers in silicon," Phys. Reports **176**, pp. 83-188 (1989).
[9] H. Ennen, J. Schneider, G. Pomrenke, and A. Axmann, "1.54 μm luminescence of erbium implanted III-V semiconductors and silicon," Appl. Phys. Lett. 43, pp. 943-945 (1983).
[10] D. Moutonnet, H. L'Haridon, P. N. Favennec, M. Salivi, M. Gauneau, F. Arnoud, d'Avitaya, and J. Chroboczek, "1.54 μm photoluminescence of erbium-implanted silicon," Mat. Sci. and Eng. **B4**, pp. 75-77 (1989).
[11] Y. H. Xie, E. A.. Fitzgerald, and Y. J. Mii, "Evaluation of erbium-doped silicon for optoelectronic applications," J. Appl. Phys. 70, pp. 3223-3228 (1991).
[12] S. Furukawa and T. Miyasato, "Quantum size effects on the optical band gap of microcrystalline Si:H," Phys. Rev. B **38**, pp. 5726-5729 (1988).
[13] F. Boulitrop, A. Chenevas-Paule, and D. J. Dunstan, "Luminescence and magnetic resonance in post-hydrogenated microcrystalline silicon," Solid State Comm. **48**, pp. 181-184 (1983).
[14] Y. Maeda, N. Tsukamoto, and Y. Yazawa, "Visible photoluminescence of Ge microcrystals embedded in SiO_2 glassy matrices," Appl. Phys. Lett. **59** pp. 3168-3170 (1991).
[15] D. J. DiMaria, J. R. Kirtley, E. J. Pakulis, D. W. Dong, T. S. Kuan, F. L. Pesavento, T. N. Theis, J. A. Cutro, and S. D. Brorson, "Electroluminescence studies in silicon dioxide films containing tiny silicon islands," J. Appl. Phys. **56**, pp. 401-416 (1984).
[16] H. Takagi, H. Ogawa, Y. Yamazaki, A. Ishizaki, and T. Nakagiri, "Quantum size effects on photoluminescence in ultrafine Si particles," Appl. Phys. Lett. **56**, pp. 2379-2380 (1990).
[17] J. R. Heath and J. M. Jasinski, "Structural studies, surface properties, and photoluminescence behavior of silicon nanocrystals," Materials Research Society Symposium Proceedings **256**, pp. 117-122 (1992).

[18] R. Zachai, K. Eberl, G. Abstreiter, E. Kasper, and H. Kibbel, "Photoluminescence in short-period Si/Ge strained-layer superlattices," Phys. Rev. Lett. **65**, pp. 1055-1058 (1990).

[19] U. Schmid, N. E. Christensen, and M. Cardona, "Direct transition energies in strained ten-mono-layer Ge/Si superlattices," Phys. Rev. Lett. **65**, p. 2610 (1990).

[20] K. Terashima, M. Tajima, and T. Tatsumi, "Near-band-gap photoluminescence of Si1-xGex alloys grown on Si (100) by molecular beam epitaxy," Appl. Phys. Lett. **57**, pp. 1925-1927 (1990).

[21] J. C. Sturm, H. Manoharan, L. C. Lenchyshyn, M. L. W. Thewalt, N. L. Rowell, J. P. Noël, and D. C. Houghton, "Well-resolved band-edge photoluminescence of excitons confined in strained $Si_{1-x}Ge_x$ quantum wells," Phys. Rev. Lett. **66**, pp. 1362-1365 (1991).

[22] J. P. Noël, N. L. Rowell, D. C. Houghton, and D. D. Perovic, "Intense photoluminescence between 1.3 and 1.8 μm from strained $Si_{1-x}Ge_x$ alloys," Appl. Phys. Lett. **57**, pp. 1037- 1039 (1990).

[23] Q. Mi, X. Xiao, J. C. Strum, L. C. Lenchyshyn, and M. L. W. Thewalt, "Room-temperature 1.3 μm electroluminescence from strained Si1-xGex/Si quantum wells," Appl. Phys. Lett. **60**, pp. 3177-3179 (1992).

[24] L. T. Canham, "Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers," Appl. Phys. Lett. **57**, pp. 1046- 1048 (1990).

[25] F. Namavar, H. P. Maruska, and N. M. Kalkhoran, "Visible electroluminescence from porous silicon np heterojunction diodes," Appl. Phys. Lett. **60**, pp. 2514-2516 (1992).

[26] N. Koshida and H. Koyama, "Visible electroluminescence from porous silicon," Appl. Phys. Lett. **60**, pp. 347-349 (1992).

[27] E. Bassous, M. Freeman, J.-M. Halbout, S. S. Iyer, V. P. Kesan, P. Munguai, S. F. Pesarcik, and B. L. Williams, "Characterization of microporous silicon fabricated by immersion scanning," Materials Research Society Symposium Proceedings **256**, pp. 23-26 (1992).

[28] M. A. Gell, "Effect of buffer-layer compositon on new optical transitions in Si/Ge short-period superlattices," Phys. Rev. B 38, 7535-7553 (1988).

[29] K. Terashima, M. Tajima, and T. Tatsumi, "Near-band-gap photoluminescence of Si1-xGex alloys grown on Si (100) by molecular beam epitaxy," Appl. Phys. Lett. **57**, pp. 1925-1927 (1990).

[30] J. C. Sturm, H. Manoharan, L. C. Lenchyshyn, M. L. W. Thewalt, N. L. Rowell, J. P. Noël, and D. C. Houghton, "Well-resolved band-edge photoluminescence of excitons confined in strained $Si_{1-x}Ge_x$ quantum wells," Phys. Rev. Lett. **66**, pp. 1362-1365 (1991).

[31] J. P. Noël, N. L. Rowell, D. C. Houghton, and D. D. Perovic, "Intense photoluminescence between 1.3 and 1.8 μm from strained $Si_{1-x}Ge_x$ alloys," Appl. Phys. Lett. **57**, pp. 1037- 1039 (1990).

[32] Q. Mi, X. Xiao, J. C. Strum, L. C. Lenchyshyn, and M. L. W. Thewalt, "Room-temperature 1.3 μm electroluminescence from strained Si1-xGex/Si quantum wells," Appl. Phys. Lett. **60**, pp. 3177-3179 (1992).

[33] L. T. Canham, "Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers," Appl. Phys. Lett. **57**, pp. 1046- 1048 (1990).

[34] A. G. Cullis and L. T. Canham, "Visible light emission due to quantum size effects in highly porous Si," Nature **353**, pp. 335-338 (1991).

[35] M. W. Cole, J. F. Harvey, R. A. Lus, D. W. Echart, and R. Tsu, "Microstructure of visibly luminescent porous silicon," Appl. Phys. Lett. **60**, pp. 2800-2802 (1992).
[36] R. Tsu, H. Shen, and M. Dutta, "Correlation of Raman and photoluminescence spectra of porous silicon," Appl. Phys. Lett. **60**, pp. 112-114 (1992).
[37] V. Lehmann, "X-ray diffraction measurements on as-grown and oxidized porous silicon samples," Proceedings of Light Emitting Si II, Garching, Germany, 1992.
[38] D. L. Naylor, S. B. Lee, J. C. Pincenti, and B. E. Bouma, "Photoluminescence in porous silicon: Evidence for the quantum confinement model," Materials Research Society Symposium Proceedings **256**, pp. 111-116 (1992).
[39] A. Halimaoui, C. Oules, G. Bomchil, A. Bsiesy, F. Gaspard, R. Herino, M. Ligeon, and F. Muller, "Electroluminescence in the visible range during anodic oxidation of porous silicon films," Appl. Phys. Lett. **59**, pp. 304-306 (1991).
[40] J. C. Vial, S. Billiat, A. Bsiesy, G. Fishman, F. Gaspard, R. Herino, M. Ligeon, F. Madeore, I. Mihalcescu, F. Muller, and R. Romestain, "Bright visible light emission from electro-oxidized porous silicon, a quantum effect," 7 th Int. Conf. on Semiconductors, Trieste, Italy (1992).
[41] T. Motohiro, T. Kachi, F. Miura, Y. Takeda, A. Hyodo, and S. Noda, "Excitation spectra of the visible photoluminescence of anodized porous silicon," Jpn. J. Appl. Phys. **31**, pp. L207-L209 (1992).
[42] J. M. Perez, J. Villalobos, P. McNeill, J. Prasad, R. Cheek, J. Kelber, J. P. Estrera, P. D. Stevens, and R. Glosser, "Direct evidence for the amorphous silicon phase in visible photoluminescent porous silicon," Appl. Phys. Lett. **61**, pp. 563-565 (1992).
[43] T. George, M. S. Anderson, W. T. Pike, T. L. Lin, R. W. Fathauer, K. H. Jung, and D. L. Kwong, "Microstructural investigations of light-emitting porous Si layers," Appl. Phys. Lett. **60**, pp. 2359-2361 (1992).
[44] N. Noguchi, I. Suemune, M. Yamanishi, G. C. Hua, and N. Otsuka, "Study of luminescent region in anodized porous silicon by photoluminescnece imaging and their microstructures," Jpn. J. Appl. Phys. **31**, pp. L490-L493 (1992).
[45] P. Gupta, V. L. Colvin, and S. M. George, "Hydrogen desorption kinetics from monohydride and dihydride species of silicon surfaces," Phys. Rev. B **37**, pp. 8234-8243 (1988).
[46] C. Tsai, K.-H. Li, J. Sarathy, S. Shih, J. C. Campbell, B. K. Hance, and J. M. White, "Thermal treatment studies of the photoluminescence intensity of porous silicon," Appl. Phys. Lett. **59**, pp. 2814-2816 (1991).
[47] C. Tasi, K.-H. Li, J. C. Campbell, B. K. Hance, and J. M. White, "Laser-induced degradation of the photoluminescence intensity of porous silicon," J. Electron. Mat. **21**, pp. 589-591 (1991).
[48] V. Petrova-Koch, T. Muschik, A. Kux, B. K. Meyer, F. Koch, and V. Lehman, "Rapid-thermal-oxidized porous Si - the superior photoluminescent Si," Appl. Phys. Lett. **61**, pp. 943-945 (1992).
[49] F. Koch, V. Petrova-Koch, T. Muschik, A. Kux, F. Müller, V. Gavrilenko, and F. Möller, "Thermal oxidation of luminescent porous Si and its implications for the mechanism of light emission," 21st ICPS, Beijing, China (1992); F. Koch, V. Petrova-Koch, T. Muschik, A. Kux, F. Müller, and V. Gavrilenko, Mat. Res. Soc. Symp. Proc. **283**, p. 197 (1993).

[50] H. Koyama, M. Araki, Y. Yamamoto, and N. Koshida, "Visible photoluminescence of porous Si and its related optical properties," Jpn. J. Appl. Phys. **30**, 3606-3609 (1991).
[51] K.-H. Li, C. Tsai, J. Sarathy, and J. C. Campbell, "Chemically-induced shifts in the photoluminescence spectra of porous silicon," Appl. Phys. Lett. **62**, pp. 3192-3194 (1993).
[52] F. Müller, A. Kux, and F. Koch, "Photochemical tuning of the luminescence spectrum of porous silicon," Proceedings of Light Emitting Si II, Garching, Germany, 1992.
[53] P. Steiner, F. Kozlowski, and W. Lang, "Light-emitting porous silicon diode with an increased electroluminescence quantum efficiency," Appl. Phys. Lett. **62**, pp. 2700-2702 (1993).
[54] J. P. Zheng, K. L. Jiao, W. P. Shen, W. A. Anderson, and H. S. Kwok, "Highly sensitive photodetector using porous silicon," Appl. Phys. Lett. **61**, pp. 459-461 (1992).
[55] L. Z. Yu and C. R. Wie, "Fabrication of MSM photoconductor on porous Si using micromachined Si mask," Electron. Lett. **28**, pp. 911-912 (1992).
[56] C. Tsai, K.-H. Li, and J. C. Campbell, "Photodetectors fabricated from rapid-thermal-oxidized porous Si," Appl. Phys. Lett. **62**, pp. 2818-2820 (1993).

POWER HBT'S

Ali Khatibzadeh, Texas Instruments CRL

PAPER UNAVAILABLE FOR PUBLICATION

High Microwave Power Performance of Self-Aligned InGaP/GaAs Heterojunciton Bipolar Transistors

L.W. Yang, S.T. Fu, B.F. Clark, R.S. Brozovich, H.H. Lin, S.M. J. Liu and P.C. Chao
Electronics Laboratory, Martin Marietta (Formerly General Electric Aerospace)
Electronics Park, Syracuse, New York 13221

F. Ren, C.R. Abernathy, S. J. Pearton, J.R. Lothian, P.W. Wisk, T.R. Fullowan,
T.Y. Chiu and S.S. Pei
AT&T Bell Laboratories
Murray Hill, New Jersey 07974

Abstract

The results of experimental study and design criteria of high efficiency power heterojunction bipolar transistors (HBTs) are reported and discussed. Self-aligned Npn HBTs fabricated by dry process technology have shown excellent device characteristics with low collector offset voltage (< 150 mV) and low knee voltage (< 0.7 V), and excellent high-frequency responses well into the millimeter-wave range. The FET-like device topology provides a means to preserve the power gain in device scaling for large emitter area. At X-band (10 GHz), an 8-cell common-base (CB) power transistor with emitter area of 240 μm^2 demonstrated 17.2 -17.5 dB gain. The 12-cell (360 μm^2) common-emitter (CE) HBTs with a breakdown voltage (BV_{ceo}) of 11V delivered 0.53W output power with 57% power added efficiency (P.A.E.) and 11.3 dB power gain at 4 GHz. Peak power of 0.68W, P.A.E. of 52%, and 10.3 dB power gain also was achieved at a collector current density of $5.5x10^4$ A/cm^2.

I. Introduction

The rapid progress of InGaP/GaAs tunneling emitter and heterojunction bipolar transistors (TEBTs and HBTs) has raised much interest in pursuing their microwave applications. The advantage of small conduction band discontinuity (130 ~150 meV) and excellent interface quality of the InGaP/GaAs HBT [1] provides, respectively, much smaller offset voltage (< 200 mV) and high current gain at low current density compared to the conventional AlGaAs/GaAs HBTs [2,3]. Furthermore, the thermal conductivity (W/deg.cm) of InGaP [4] between 300 °K and 700 °K is greater than that of GaAs and AlGaAs. All these characteristics

0-7803-0894-8/93 $3.00

may improve the device power-added efficiency (P.A.E.) and linearity in microwave power operations.

Up to now, however, most reports in the literature are focused on the DC performance - only a few have small signal RF results. In this paper, we report on the design of power transistors and the microwave power performance, as well as small signal results of MOMBE-grown InGaP/GaAs HBTs fabricated with self-aligned technology. Power transistors consisting of 4, 8, and 12 cells were fabricated using self-aligned technology and lapped down to 2 mils for microwave power evaluation. Based on a finite-element 3D analysis, a novel device topology (FET-like) was proposed to optimize the thermal design and eliminate the distributive effect in transistors with large emitter peripheries.

II. Design Consideration of Power HBTs

Heterojunction bipolar transistors have an advantage over field effect transistors (FETs) in cut-off mode operation for high P.A.E.[5] The P.A.E. of HBT power amplifiers operated in classes AB, B, and C can be further improved to nearly reach theoretical limits by enhancing the collector efficiency and power gain. Both factors depend on the heterostructure material system, epitaxial layer design, device layout, DC and RF characteristics (which are the outcome of device technology), and operating conditions. In general, high collector efficiency can be achieved in HBTs through a low microwave saturation voltage and a high voltage bias in device operations. The microwave saturation voltage is determined by the collector offset voltage and knee voltage of device output characteristics in the operating microwave frequency.[6]

To maintain a high efficiency at moderate bias voltages (if system applications limit the operation at high voltage), optimization of device design is necessary to minimize the microwave saturation voltage(MWV_{ces}) and to maximize the power gain. Achieving the minimum MWV_{ces} at chosen bias voltage (V_{cc}), the total output power is then dominated by the peak current at the fundamental frequency of operation, and the power gain is determined primarily by the cut-off frequency, f_T; maximum frequency of oscillation, f_{MAX}; and emitter lead inductance, L_e.[as explained in the Ref. 6] Furthermore, both theoretical analysis [5] and experimental evidence in our laboratories show that low saturation resistance (which leads to a low knee voltage) and sufficient current gain at low collector current densities are essential for

achieving high power gain.[7] High power gain at low bias level is key to high efficiency performance. [8]

In the design of InGaP/GaAs HBTs, a device layout based on self-aligned device technology as well as a heavily carbon-doped GaAs base layer was selected to accomplish high f_{MAX}. To minimize the thermal effects on device performance, a large spacing (66.5 μm) between the emitter fingers and a thin substrate (2 mils) were suggested from the results of thermal modeling study using 3D finite element analysis. Within the power transistor, each unit cell —composed of an emitter finger of 2x15 μm^2— is interconnected through anisotropically etched and Au-plated backside vias on extrinsic emitter pads (Fig. 1). This via interconnection can reduce grounding inductance and aid heat dissipation. Multiple bonding pads at both input and output ports of 12-cell power transistors were utilized to minimize phase delay. With this particular device layout, power combing was achieved, and a junction temperature below 120 °C was estimated (by simulation) using the microwave power operations described below.

To achieve a high frequency response (high f_T) with high current drive capability as well as a high breakdown (with BV_{ceo} ~ 11-12 V), heavily doped InGaAs emitter cap layers, a thin base layer with a high doping level, and optimized collector thickness and doping were included in the device epitaxial layer design. (Table 1) Small conduction bandgap discontinuity in lattice-matched InGaP/GaAs heterointerface and high carrier concentrations in the base and emitter contact layers provide the means to reduce the collector offset and low saturation resistance, thereby a low knee voltage and minimum MWV_{ces}.

In studying the frequency response and its correlation with total emitter area, small signal RF measurements were carried out on common-base transistors with different numbers of unit cells (CB4 and CB8). Results are summarized in Table 2. At 10 GHz, and at collector current density, $J_c = 3x10^4$ A/cm^2, small signal power gain (17.2 - 17.5dB) of an 8-cell CB HBT with total emitter area of 240 μm^2 are comparable to those of 4-cell CB transistors with emitter area of 120 μm^2 at different collector bias voltages (V_{cb}). This indicates that the FET-like device layout along with a large emitter-spacing fabricated with a via-hole interconnect preserves the device microwave gain in scaling the device for large emitter area (more collector current), thus yielding higher output power.

III. Epitaxial Structures and Device Fabrication

The epitaxial layers were grown on 2-inch semi-insulating GaAs substrates at a constant temperature of 525 °C in the a metal-organic molecular beam epitaxial (MOMBE) system. The layer structures are illustrated in Fig. 1. In order to optimize the microwave power performance at relative low operation voltages, a collector thickness of 5000 Å was selected in the epitaxial design. Triethylgallium (TEG) and trimethylgallium (TMG) were used to grow the n-type and p^+ GaAs layers, respectively. Arsine (AsH_3) was the group V source during GaAs growth. The group III organometallic sources utilized for the growth of InGaP were trimethylindium (TMI) and trisobutylgallium (TIBG)[9], while phosphine (PH_3) was introduced to the chamber via a low-pressure cracker maintained at 1100 °C to generate the group V element.

The HBT devices are fabricated using a base-to-emitter self-aligned technology. Device isolation was achieved from a sequence of ion implant steps using proton and fluorine ions. Emitter ohmic metal was utilized as the etching mask to define the emitter fingers and to form emitter mesas via multiple steps of reactive ion etching (RIE) using CCl_2F_2-, and BCl_3-based discharges for etching GaAs layers, while using CH_4/H_2/Ar-based discharge and selective wet chemical etching for InGaAs and InGaP layers.[10] A thin layer of silicon nitride for sidewall protection was formed prior to the deposition of non-alloyed base contact metal using Ti/Pt/Au, followed by the base mesa formation using RIE. The input/output ports of the power transistors were interconnected by a lift-off metallization using Ti/Pt/Au after device passivation.

After finishing the frontside fabrication, the wafer was lapped and polished to a thickness of 2 mils to improve overall heat dissipation. Three additional mask steps were employed to complete the flipside fabrication; These include substrate street definition and etching; via-hole RIE and metallization using electrodeless and electrode-plating techniques; and die separation. The grounding pads of the power transistors, both for common-emitter and common-base configurations, are interconnected only through flipside via-hole dry etching and plated metal.

IV. Device Characteristics and Microwave Performance

The electrical characteristics of devices were measured using a HP 4145A semiconductor parameter analyzer. A HP 8510B automatic network analyzer and a Wiltron 360B system were used in conjunction with Cascade Microtech high frequency probe heads to measure the transistor frequency response. The line reflection match (LRM) method is employed to calibrate the network analyzer system. Small-signal S-parameters were measured from 0.5 to 40 GHz on the HP system, and 0.5 to 60 GHz on the Wiltron instrument. The current gain ($|H_{21}|^2$), maximum stable gain (MSG), and maximum available gain (MAG) as a function of frequency were derived from the measured s-parameters.

Excellent DC performance of InGaP/GaAs HBTs was demonstrated. The unit cell device with emitter dimensions of 2x15 μm^2 demonstrated small signal current gain (β) of 20 - 35. Good junction ideality factor, n = 1.29, was achieved from the emitter-base pN^+ diodes. Low offset (< 150 mV) and low knee voltage (< 0.7 V) for a power transistor consisting of four unit cells, with total emitter area of 120 μm^2, was demonstrated as shown in Fig. 3. The device demonstrated small signal current gain > 25 at low collector current density of 4-5 $x10^3 A/cm^2$. These results are superior to AlGaAs/GaAs HBTs with similar device structures, and are highly desirable for high P.A.E. in microwave power applications.

The unit cell devices were RF tested (small signal S-parameters) from 0.5 to 60 GHz and exhibited a high current gain cut-off frequency, f_T, around 80 GHz, and maximum frequency of oscillation, f_{MAX}, of 70 GHz, in a common-emitter (CE) configuration (Fig. 2a). The common base (CB) transistor with same emitter dimensions demonstrated maximum stable gain (MSG) of 6 dB at 60 GHz (Fig. 2b), which is comparable to the best of AlGaAs/GaAs HBTs. These high frequency results indicate the great potential of InGaP/GaAs HBTs for millimeter-wave power application. The InGaP/GaAs HBT approaches the maximum f_T when biased at a low V_{ce} of 1 V and at a collector current density of $5.2x10^4$ A/cm^2.

Large-signal microwave power evaluations were done with devices mounted on the test fixtures. The discrete power transistors were first die-attached with indium/lead (InPb) alloy onto carriers and then assembled into the fixture. The chips were connected through bonding wires to the 50-Ω microstripe line of the fixture. The power measurement setup was first calibrated by measuring fixture loss with through-line connection. Impedance matching was

done by an external tuner and in-fixture chip tuning. The extra loss introduced from the matching was not removed from the power measurements, therefore the power results are conservative.

Power transistors were tested at 4 GHz with V_{ce}=5 V. This relative low bias of power transistors (or amplifiers) is highly desirable for wireless communication. A maximum output power of 0.6 W with a gain of 12.7 dB and 50% P.A.E. was measured under CW bias. The device also showed high linearity with gain of 15-15.5 dB up to the input power level of 12 dBm. (Fig. 4) Under the pulse bias mode with a pulse width of 50 μs and V_{ce} = 5 V, a different power transistor with 12 cells (360 μm^2) and a breakdown voltage (BV_{ceo}) of 11 V delivered 0.53W output power with 57% P.A.E. and 11.3 dB gain. Peak power of 0.68W, 52% P.A.E., and 10.3 dB gain were also achieved at a collector current density of $5.5x10^4$ A/cm^2.

To explore the power performance at even lower voltages of operation, a common-base transistor with 12 unit cells was evaluated under CW bias. At V_{cb} = 4.0 V and emitter current (I_e) of 100 mA, it exhibited 24-25 dBm output power with 9.7-12.8 dB gain, and peak P.A.E. of 52% at 4 GHz.

V. Conclusions

In summary, we have reported the results of experimental study in pursuing the microwave power applications of InGaP/GaAs HBTs. The design criteria of power HBTs with high efficiency were discussed. The proposed FET-like device topology provides a means to preserve power gains when scaling up power transistors for large emitter area. The advantage of this new material system for Npn heterojunction bipolar transistors, in conjunction with state-of-the-art self-aligned device technology, leads to excellent device characteristics such as low collector offset voltage (< 150 mV) and low knee voltage (< 0.7 V), —both are essential for high P.A.E. in microwave power operations. The unit cell devices exhibited high frequency response with f_T, f_{max} well into millimeter-wave range. At 4 GHz, high power gain of 12.7 dB and high linearity with P.A.E. ≥ 50% were achieved from CE HBTs at relatively low voltage bias of 5 volts. Transistors with a CB configuration operating at 4 V, demonstrated P.A.E. of 52% and > 9.7 dB gain at a peak output power level of 25 dBm.

These results suggest that InGaP/GaAs HBTs are suitable not only for high-speed digital but also for microwave small-signal and power applications.

Acknowledgment

The technical support of N.J. Nordheden and D.E. Houston at Martin Marietta's Electronics Laboratories is gratefully acknowledged, as well as the continuied support of A.W. Swanson, A.A. Immorlica, J.A. Kinzel, and W.J. Butler. The continued support of D.V. Lang at AT& T is also acknowledged.

REFERENCES

[1] W. Liu and S. K. Fan, "Near-Ideal I-V Characteristics of GaInP/GaAs Heterojunction Bipolar Transistors", EDL-13 (10) pp. 510-512, 1992

[2] C. R. Abernathy, F. Ren, P.W. Wisk, S.J. Pearton, and R. Esagui, "Improved performance of carbon-doped GaAs base heterojunction bipolar transistors through the use of InGaP", Appl. Phys. Lett. 61(9), pp. 1092-1094, 1992

[3] L.W.Yang, P.A. Martin, J.S. Mazurowski, J.M. Ballingall, Y.Yin, D. Yan, F.H. Pollak, W. West and D.B. Davito, "Material and Device Characterization of High Quality InGaP/GaAs HBT Structures Using Photoreflectance, Capacitance-, and Current-Voltage Measurements", Electronic Materials Conference, University of California, Santa Barbara, California, June 23-25, 1993, paper in preparation for publication in Journal of Electronic Materials.

[4] S. Adachi, "physical Properties of III-V Semiconductor Compounds", John Wiley & Sons, 1992

[5] M.G. Adlerstein and M. P. Zaitlin, "Cutoff Operation of Heterojunction Bipolar Transistors", Microwave Journal, Vol.34, No.9, pp.114-125, 1991

[6] G. Gao, H. Morkog, and M. C. F. Chang, "Heterojunction Bipolar Transistor Design for Power Applications", IEEE Trans. Electron Dev. vol. 39, no. 9, pp. 1987-1997, 1992

[7] L.W. Yang, J. D. Kennedy, J.J. Komiak and A. Chin; unpublished results.

[8] C.S. Wu, C.K. Pao, M.Hu, A.T. Igawa, D.C. Wang and C.P. Wen,"High Efficiency X-band Power HBT's",Tech.Digest, pp.254-262, 14th GaAs IC Symposium, Oct.4-7,1992

[9] P. Zwicknagl, U. Schaper, L. Schleicher, H. Siweris, K.H. Bachem, T. Lauterbach and W. Pletschen, " High speed non-self aligned GaInP/GaAs TEBT", Electron. Lett., vol. 28, pp. 327-328, 1992

[10] J. R. Lothian, J. M. Kuo, F. Ren, and S. J. Pearton, " Plasma and wet etching of InGaP", Journal Electronic. Mater., vol.21, pp. 441-445, 1992

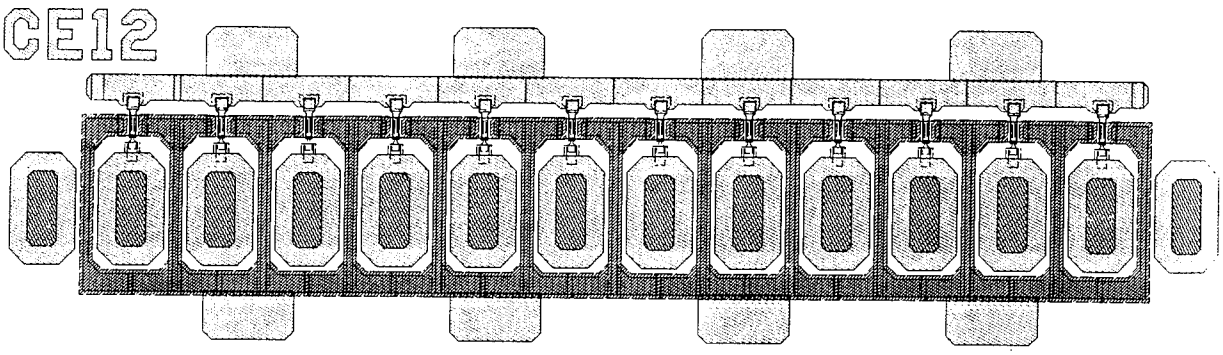


Fig. 1 The layout design of a 12-cell power transistor with "FET-like" topology. All unit cells are interconnected through metallized via holes and backside metal.

Table 1. Layer structures of InGaP/GaAs Power HBT.

Emitter Cap 3	$In_{0.75}\ Ga_{0.25}\ As$	n^+, $3x10^{19}cm^{-3}$, 300Å
Emitter Cap 2	$In_x\ Ga_{1-X}\ As$	n^+, $3x10^{19}cm^{-3}$, 300Å
Emitter Cap 1	GaAs	n^+, $1.5x10^{19}cm^{-3}$, 3500Å
Emitter	InGaP	n , $5x10^{17}cm^{-3}$, 500Å
Base	GaAs	p^+, $7x10^{19}cm^{-3}$, 700Å
Collector	GaAs	n^-, $3x10^{16}cm^{-3}$, 5000Å
Subcollector	GaAs	n^+, $3x10^{18}cm^{-3}$, 8000Å
GaAs Substrate		

Table 2. Comparison of gains (MSG/MAG) measured at 10 GHz and Jc = $3x10^4$ A/cm^2 for common-base HBTs with different emitter areas.

Device Configuration	Emitter Area (μm^2)	Gain (dB) at Vcb = 1.7V	Gain (dB) at Vcb = 3V	Gain (dB) at Vcb = 4V
CB (4 cells)	120	17.5	17.5	17.5
CB (8 cells)	240	17.5	17.5	17.2

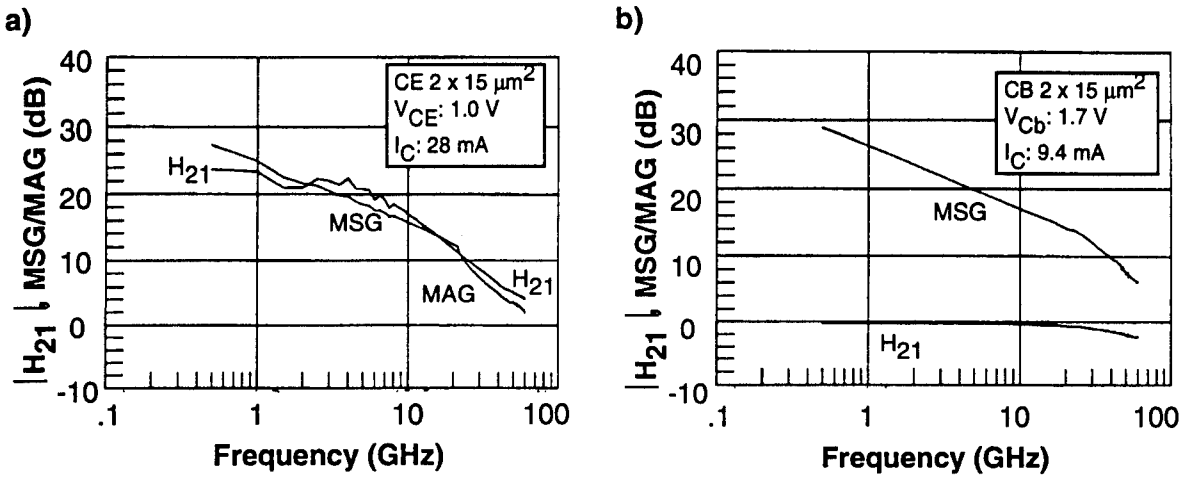


Figure 2. The current gain, $|H_{21}|^2$, maximum stable gain (MSG) and maximum available gain (MAG) extracted from S-parameters measured from 0.5 - 60 GHz for a) common emitter and b) common base configurations.

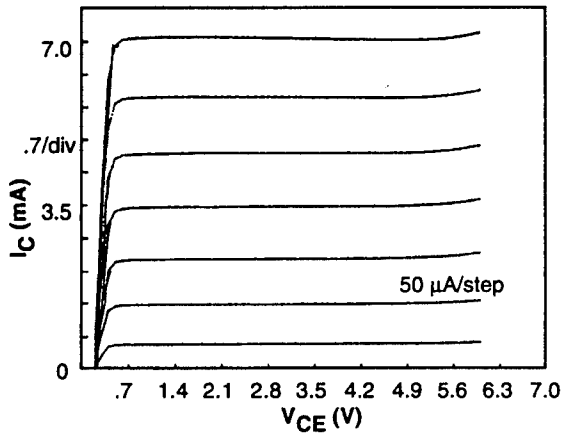


Figure 3. The I-V characteristics of an InGaP/GaAs power HBT with common-emitter configuration (measured from a 2 mil substrate). The transistor consists of four unit cells with a total emitter junction area of 120 μm^2 .

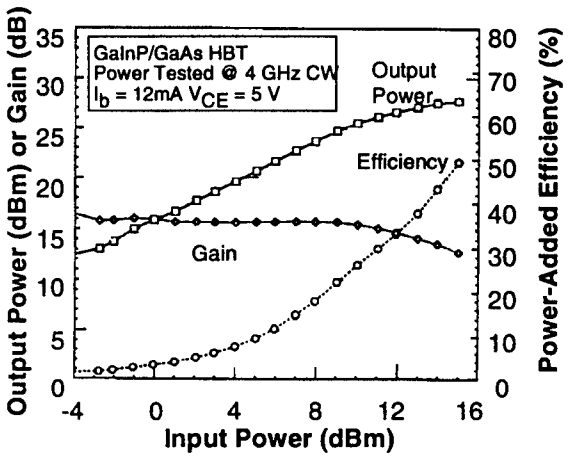


Figure 4. Measured CW power-added efficiency, output power and gain as a function of input power at 4 GHz for InGaP/GaAs (CE) HBT with 360 μm^2 total emitter junction area.

JUNCTION BARRIER EFFECTS ON THE MICROWAVE POWER PERFORMANCE OF DOUBLE HETEROJUNCTION BIPOLAR TRANSISTORS

B. R. A. Sugeng, C. J. Wei and J. C. M. Hwang
Lehigh University, 19 Memorial Drive West, Bethlehem, PA 18015

J. I. Song, W. P. Hong and J. R. Hayes
Bellcore, 331 Newman Springs Road, Red Bank, NJ 07701

ABSTRACT

The barrier effects in InP based double heterojunction bipolar transistors have been investigated both theoretically and experimentally. It was found that the gradual saturation of collector current with increasing collector voltage was caused by the collector-base heterojunction barrier. Whereas the gain compression with increasing base current was caused by electron accumulation at the collector-base heterojunction. These effects have been included in terms of nonlinear currents and capacitances in an equivalent circuit model for the prediction of DHBT large signal performance. By using an InAlAs emitter and an InP collector with an InGaAs spacer to alleviate these undesirable effects, improvement in microwave power performance was verified.

INTRODUCTION

InP based heterojunction bipolar transistor exhibited higher cutoff frequencies than its GaAs based counterpart [1]. However, its power performance was limited by high collector-emitter offset voltage and low breakdown voltage [2]. Recently, these limits were overcome by using a wideband material such as InP in the collector, thereby achieving greater than 1 W/mm output power in such a double heterojunction bipolar transistor (DHBT) [3][4]. But the DHBT is characterized by gradual saturation and gain compression hence low efficiency and linearity. In this paper, the effects introduced by the collector-base heterojunction are analyzed using a physical model. From the physical insight gained through the analysis, nonlinear elements are incorporated in an equivalent circuit model of unique topology for the prediction of DHBT large signal performance.

EXPERIMENTAL

Figure 1 illustrates the three types of DHBT structures that were analyzed. Type A and Type B contain two InP/InGaAs heterojunctions while Type C contains an InP/InGaAs and an InAlAs/InGaAs heterojunction. Type A structure includes an InGaAs spacer at the collector-base junction to suppress Zn outdiffusion from the base. A n^+-InP layer is added to Type B to reduce the collector-base junction barrier. In Type C the InP emitter is replaced by an InAlAs emitter to lower the emitter resistance. All the structures are grown by metal-organic chemical vapor

0-7803-0894-8/93 $3.00

deposition using Zn and Si as *p*-dopant and *n*-dopant, respectively. The DHBT growth and fabrication techniques have been reported elsewhere [3].

Figure 2 shows typical dc I-V characteristics measured on the DHBTs. It can be seen that, although Type A DHBT exhibits lower than 0.1 V offset voltage and higher than 10 V breakdown voltage, the collector current saturates only after the collector voltage exceeds 4 V. In Type B the saturation voltage is reduced to approximately 0.5 V, but at high base current the collector current continues to increase gradually with increasing collector voltage. In Type C the current gain is compressed at high base currents. The apparent saturation at high base currents is due to self-heating. Without self-heating, the pulsed I-V characteristics exhibit both gradual saturation and gain compression as shown in Fig. 3.

The measured small signal RF characteristics are shown in Fig. 4. S_{22} of Type A DHBT contains a kink in the middle of the frequency band suggesting two different charge accumulation mechanisms at the collector-base junction. Although less obvious, Type B and Type C DHBTs show similar tendency of midband deviation from the constant resistance circle. The unity current gain frequency improves from approximately 30 GHz for Type A, 35 GHz for Type B, to 50 GHz for Type C. The maximum frequency of oscillation is 20 GHz for Type A, 30 GHz for Type B, and 60 GHz for Type C.

PHYSICAL MODEL AND ANALYSIS

To analyze the DHBT I-V characteristics, a two dimensional physical device simulator, *SEMICAD* [5], was modified to include both thermionic emission and tunnelling field emission. Published velocity-field dependence for InP and InGaAs, including negative differential mobility, was used in the simulation. A Si like mobility model was assumed for InAlAs. The dopant outdiffusion length was assumed to be 3 nm.

Figure 5 shows the simulated energy band diagram along with electron density in the three types of DHBTs. It can be seen that, for Type A, a conduction band barrier exists near the collector-base junction at a collector voltage of 2 V and the barrier vanishes only after V_{ce} exceeds 4 V. Thus this barrier appears to be the culprit for the gradual saturation behavior. For Type B, the barrier is located farther from the collector-base junction with the top of the barrier below the conduction band in the base, resulting in improved saturation behavior over that of Type A.

Figure 5 also shows strong electron accumulation at the InGaAs/InP heterojunction near the collector-base junction. The accumulation increases with increasing base-emitter voltage and, at high V_{be}, can exceed the accumulation at the collector-base or emitter-base *p-n* junctions. This retards electron diffusion across the base resulting in gain compression. With both gradual saturation and gain compression accounted for in the model, the simulated I-V characteristics (Fig. 6) are in reasonable agreement with measurement.

EQUIVALENT CIRCUIT MODEL AND ANALYSIS

With the above physical insights, an existing DHBT equivalent circuit model [6] was modified to more accurately reflect the nonlinear effects. As shown in Fig. 7, the modification consists mainly nonlinear currents and capacitances at the base and collector, respectively.

The nonlinear base current, I_{be}, is expressed as:

$$I_{be} = I_s[\exp(\frac{V_{be}-\Delta E_{eb}}{kT}) - 1]$$

where I_s is the saturation current; ΔE_{eb} is the emitter-base barrier; k is the Boltzmann constant; T is absolute temperature.

The nonlinear collector current, I_c, is expressed in terms of a nonlinear base transport factor, α, which contains functions of I_{be} and V_{cb} to account for gain compression and gradual saturation, respectively:

$$\alpha = \alpha_o[1+a_1\exp(a_2 I_{be})]^{-1}[1+a_3\exp(-\Delta E_{cb}/kT)]^{-1}$$

$$\Delta E_{cb} = \Delta E_{cbo}\exp(-a_4 V_{cb})$$

where α_o is the base transport factor at low injection levels; ΔE_{cb} is the collector-base heterojunction barrier; ΔE_{cbo} is ΔE_{cb} at V_{cb}=0; a_1, a_2, a_3 and a_4 are fitting parameters.

Finally, a non-linear capacitance in terms of $Q_{ce} = -t_c I_c$ is included in the collector node to account for electron accumulation at the collector-base junction. Here t_c is collector transit time.

The nonlinear currents and emitter and collector resistances are determined from dc I-V characteristics. The base resistances and all capacitances and inductances are determined from small signal S-parameters. Figures 8 and 9 show simultaneous fit with both dc I-V characteristics and S-parameters.

This nonlinear equivalent circuit model has been incorporated into a harmonic balance simulator, *LIBRA* [7], as a user defined element. The model, in spite of its complexity, is very efficient for large signal analysis. Figure 10 shows the simulated output power and power added efficiency for a Type C DHBT biased at V_{ce}=6V and I_b=0.1mA (Class AB). The simulated maximum power output of 1W/mm and efficiency of 23% agrees with the measured data.

CONCLUSION

Unique DHBT characteristics in terms of gradual saturation and gain compression have been analyzed. The collector-base heterojunction barrier was found to be responsible for gradual saturation while gain compression is due to electron accumulation at the barrier. These effects have been included in terms of nonlinear currents and capacitances in an equivalent circuit model for the prediction of DHBT large signal performance. By using an InAlAs emitter and an InP collector with an InGaAs spacer to alleviate the undesirable power limiting effects, improvement in microwave power performance was verified.

ACKNOWLEDGEMENT

This work was partially supported by the Air Force Wright Laboratory under Contract No. F33615-91-C-1793.

REFERENCES

[1] Y.-K. Chen, R. N. Nottenburg, M. B. Panish, R. A. Hamm and D. A. Humphrey, "Subpicosecond InP/InGaAs heterostructure bipolar transistors," *IEEE Electron Device Lett.*, vol. 10, pp. 267-269, June 1989.

[2] J. F. Jensen et. al., "AlInAs/GaInAs HBT IC technology", *IEEE J. Solid-State Circuits*, vol. 26, pp. 415-421, 1991.

[3] J.-I. Song, W.-P. Hong, R. Bhat, K. B. Chough, J. R. Hayes, B. Sugeng, C. J. Wei and J. C. M. Hwang, "Microwave power InP/InGaAs/InP double-heterojunction bipolar transistors", *Electronic Lett.*, vol. 29, pp. 724-725, Apr. 1993.

[4] W. E. Stanchina, T. Liu, D. B. Rentsch, P. MacDonald, M. Hafizi, W. W. Hooper, Y. K. Allen, T. V. Kargodorian, R. Wong-Quen and F. Williams, "Performance of AlInAs/GaInAs/InP microwave DHBTs," in *Proc. Fifth Int'l Conf. Indium Phosphide Related Materials*, Apr. 1993.

[5] Dawn Technologies, Inc., Sunnyvale, CA.

[6] C. J. Wei, B. R. A. Sugeng, J. C. M. Hwang, W. P. Hong, J. I. Song and J. R. Hayes, "A microwave SPICE model for double heterojunction bipolar transistors," *IEEE Sarnoff Symp. Digest*, Mar. 1993.

[7] EEsof, Inc., Westlake Village, CA.

(a)

Cap	n^{++}	InGaAs	$2x10^{19}/cm^3$	200nm
Emitter	n^{++}	InP	$5x10^{18}/cm^3$	100nm
Emitter	n	InP	$5x10^{17}/cm^3$	250nm
Spacer	i	InGaAs	undoped	15nm
Base	p^{++}	InGaAs	$1x10^{19}/cm^3$	85nm
Spacer	i	InGaAs	undoped	50nm
Collector	n	InP	$5x10^{16}/cm^3$	400nm
Collector	n^{++}	InP	$5x10^{18}/cm^3$	100nm
Sub-Coll	n^{++}	InGaAs	$2x10^{19}/cm^3$	500nm
	S.I.	InP		

(a) Type A

(b)

Cap	n^{++}	InGaAs	$2x10^{19}/cm^3$	200nm
Emitter	n^{++}	InP	$5x10^{18}/cm^3$	100nm
Emitter	n	InP	$5x10^{17}/cm^3$	250nm
Spacer	i	InGaAs	undoped	15nm
Base	p^{++}	InGaAs	$1x10^{19}/cm^3$	85nm
Spacer	i	InGaAs	undoped	10nm
Collector	n^{-}	InP	$1x10^{18}/cm^3$	15nm
Collector	n	InP	$5x10^{16}/cm^3$	400nm
Collector	n^{++}	InP	$5x10^{18}/cm^3$	100nm
Sub-Coll	n^{++}	InGaAs	$2x10^{19}/cm^3$	500nm
	S.I.	InP		

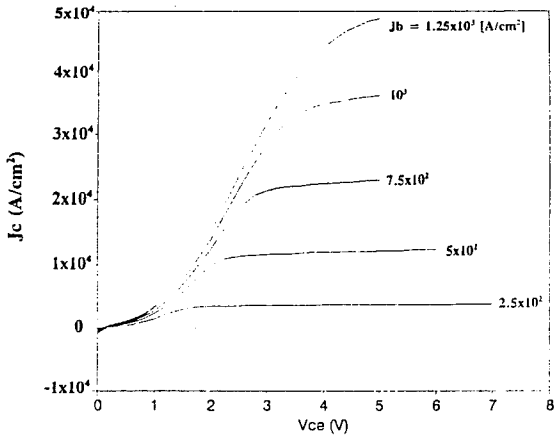
(b) Type B

(c)

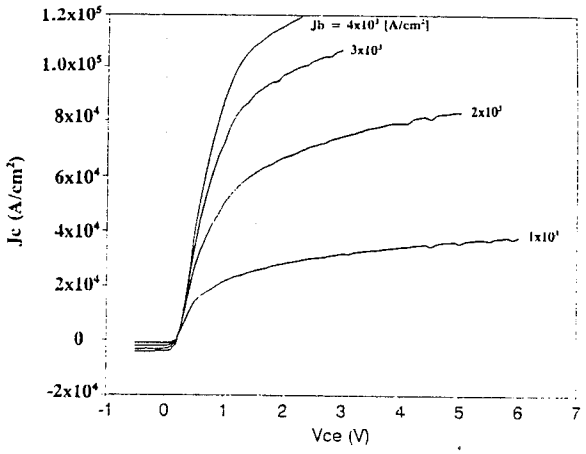
Cap	n^{++}	InGaAs	$2x10^{19}/cm^3$	200nm
Emitter	n^{++}	InAlAs	$2x10^{19}/cm^3$	230nm
Emitter	n	InAlAs	$1x10^{17}/cm^3$	70nm
Emitter	n	InAlAs	$1x10^{18}/cm^3$	5nm
Spacer	i	InGaAs	undoped	15nm
Base	p	InGaAs	$2x10^{19}/cm^3$	60nm
Spacer	i	InGaAs	undoped	20nm
Collector	n	InP	$2x10^{18}/cm^3$	15nm
Collector	n	InP	$5x10^{16}/cm^3$	100nm
Collector	n	InP	$1x10^{16}/cm^3$	300nm
Collector	n	InP	$2x10^{19}/cm^3$	100nm
Sub-Coll	n	InGaAs	$2x10^{19}/cm^3$	500nm
	S.I.	InP		

(c) Type C

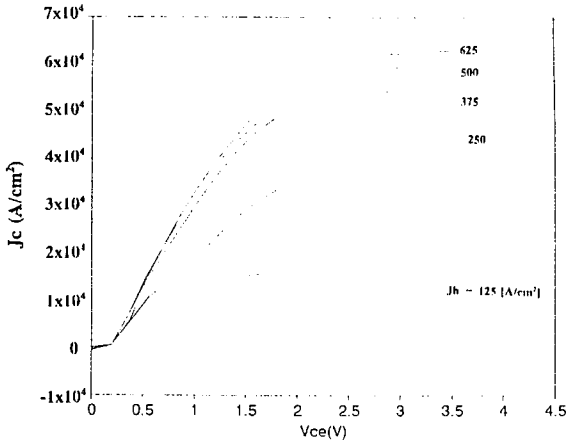
Fig 1. Structure of DHBT



(a) Type A



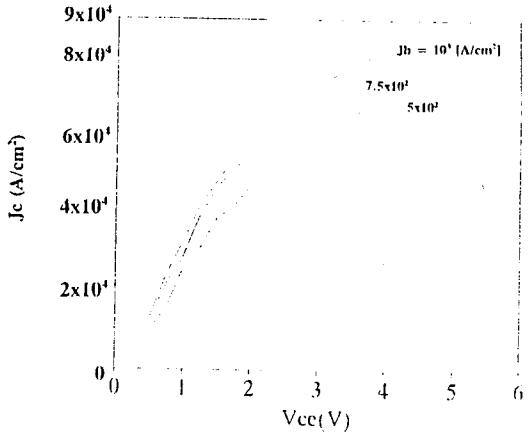
(b) Type B

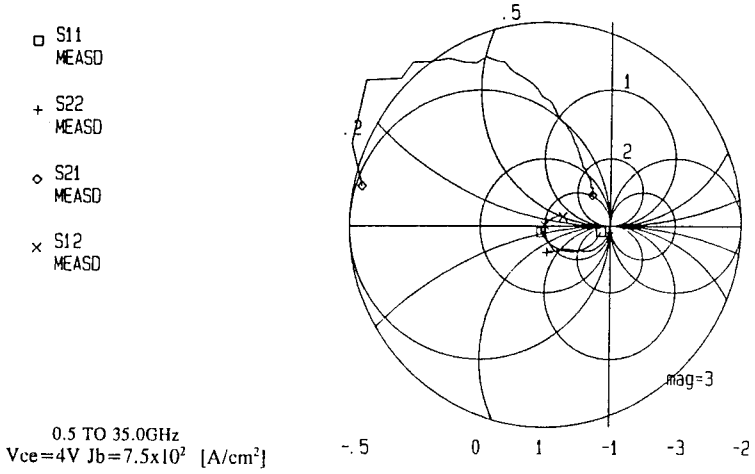


(c) Type C

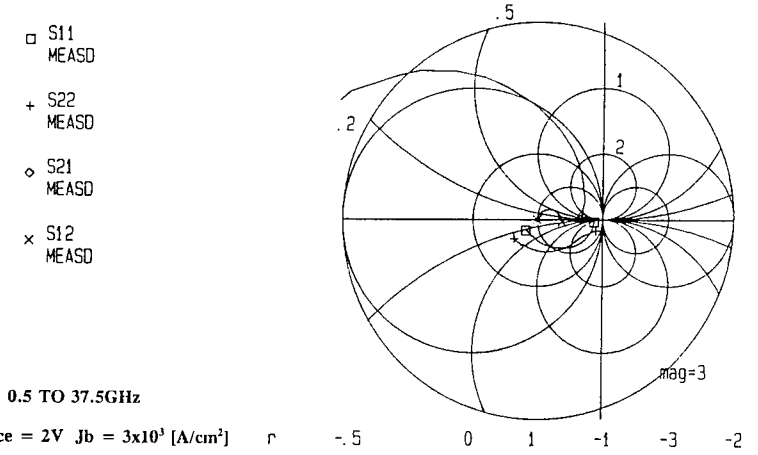
Fig 2. Measured IV characteristics of DHBT

Fig 3. Pulsed IV characteristics of DHBT Type C

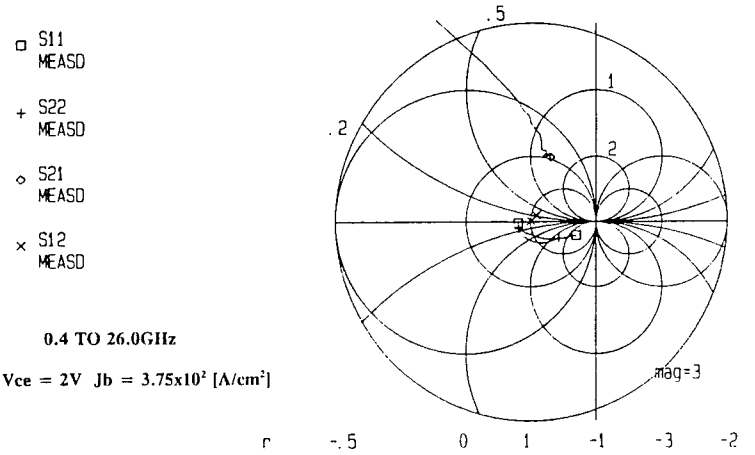




(a) Type A

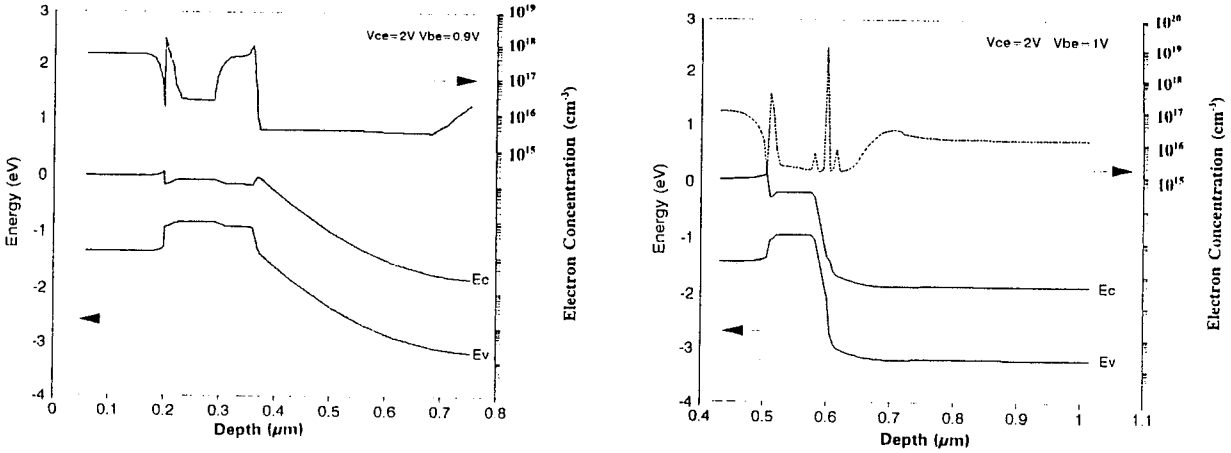


(b) Type B



(c) Type C

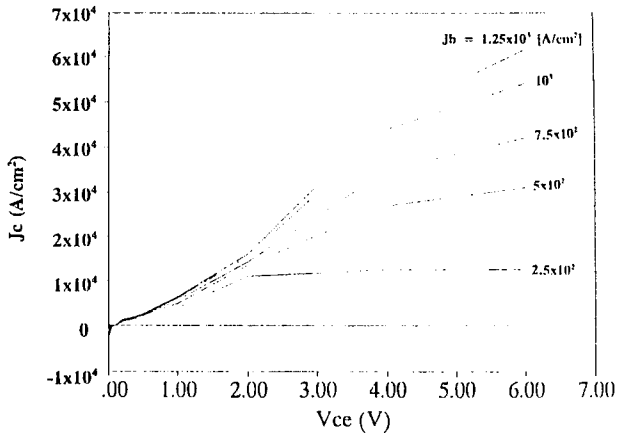
Fig 4. Measured small signal S-parameters of DHBT



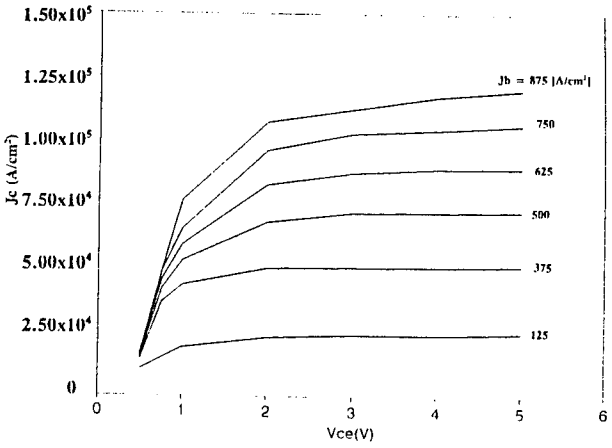
(a) Type A at Vce=2V and Vbe=0.9V

(b) Type C at Vce=2V and Vbe=1V

Fig 5. Energy band diagram (simulated) of DHBT



(a) Type A



(b) Type C

Fig 6. Simulated IV Characteristics of DHBT

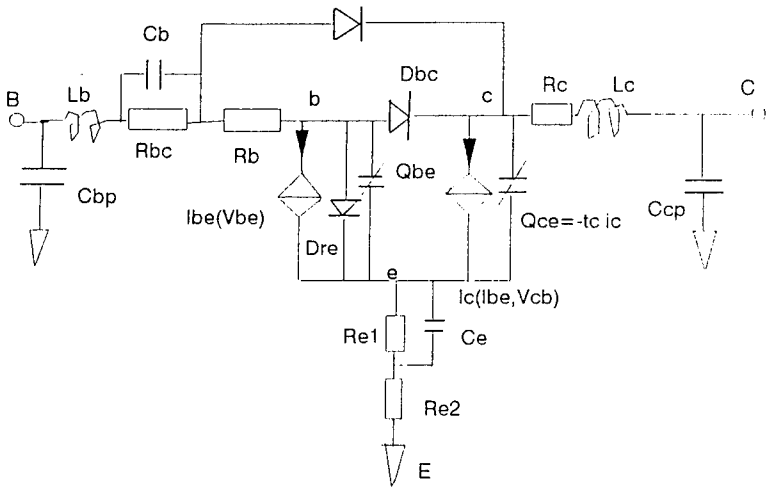


Fig 7. Large-signal Equivalent Circuit.

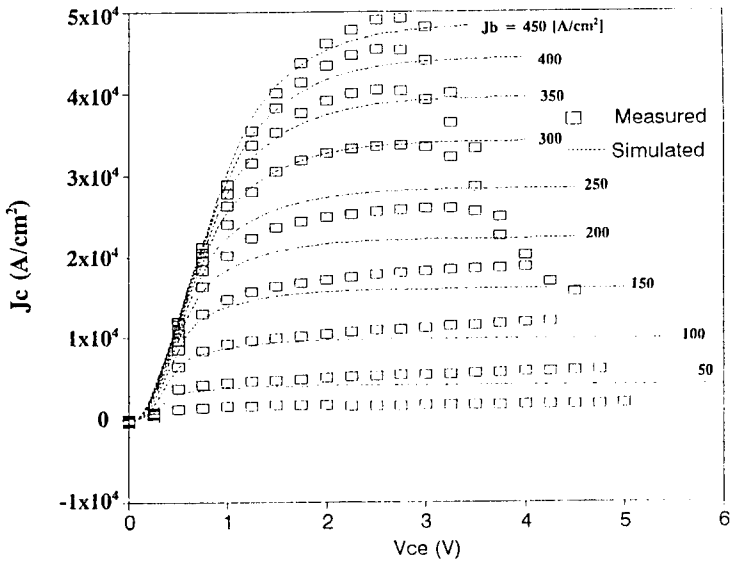


Fig 8. Measured and simulated IV characteristics of DHBT Type C (InAlAs/InGaAs/InP DHBT).

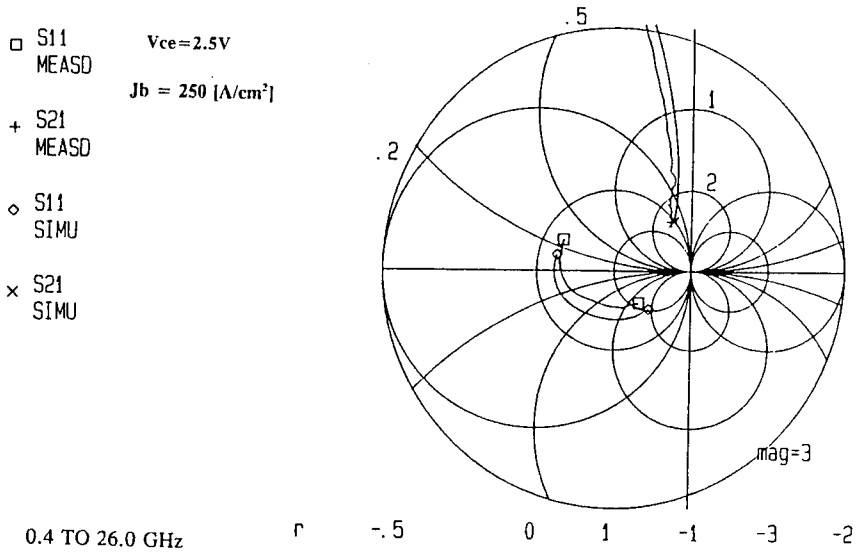


Fig 9. Measured and simulated small-signal S-parameters of DHBT Type C.

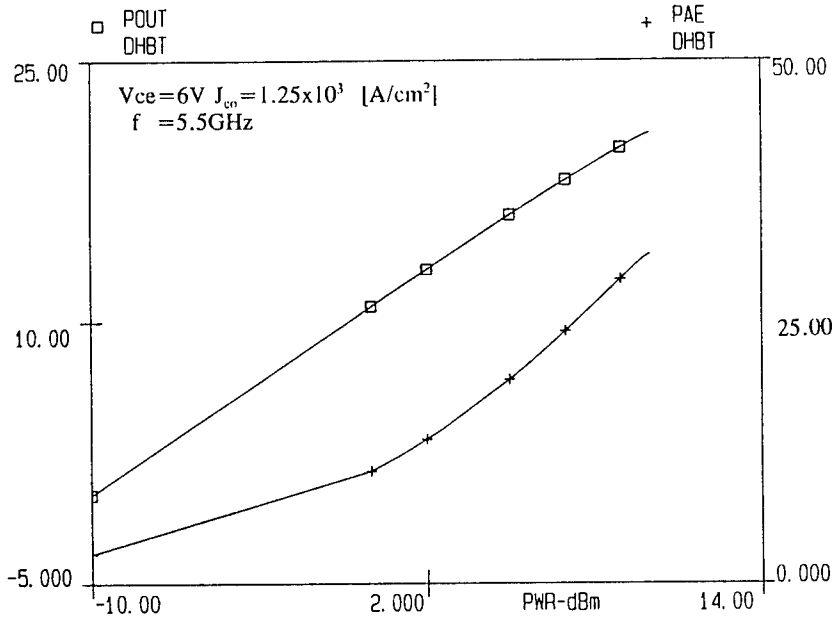


Fig 10. Simulated output power and power added efficiency of DHBT Type C biased at class AB.

InAlAs/InGaAs(P) Double Heterojunction Bipolar Transistors with High Breakdown Voltage Grown by Chemical Beam Epitaxy (CBE) [1]

John C. Cowles, W.L. Chen, G.O. Munns, and G.I. Haddad

Center for High Frequency Microelectronics
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109-2122
(313)-764-3305

ABSTRACT

InAlAs/InGaAs single(S) and double(D) heterojuction bipolar transistors (HBTs) have been grown by Chemical Beam Epitaxy (CBE). The DHBT features an InGaAsP precollector that greatly increases the breakdown voltage from 2V to 8V and decreases the collector emitter offset voltage from 360mV to 210mV in comparison to the SHBT. The common emitter current gain β remains high in both structures. Microwave measurements performed on the DHBT reveal that f_t increases monotonically with J_c upto at least $10^5 A/cm^2$ showing no signs of base pushout. These results show that CBE grown DHBTs are promising devices for high power applications.

1. Introduction

Heterojunction bipolar transistors (HBTs) lattice matched to InP have demonstrated excellent high frequency performance and are promising candidates for integration into optoelectronic lightwave communication systems. InAlAs/InGaAs and InP/InGaAs single(S) HBTs have been reported with f_t in excess of 110 GHz and 165 GHz, respectively [1][2]. However, InGaAs collector devices typically suffer from low collector breakdown voltage due to avalanche multiplication and tunneling across the base-collector junction. Furthermore, due to the drastic assymmetry in the two junctions, a significant offset voltage appears in the output characteristics. For power applications where output voltage swing and current drive capacity are essential, these devices are clearly unsuitable. Double(D) HBTs with InP collectors have shown greatly increased breakdown voltages and good RF performance but have relied on a 2-step growth sequence of MOVPE followed by MBE [3]. More recently, DHBTs grown by gas-source MBE,MOCVD and MOMBE have shown excellent speed and power performance [4][5][6]. We report an InAlAs/InGaAs/InGaAsP DHBT grown entirely by CBE showing improved DC characterisics and RF performance comparable to a SHBT processed in parallel.

[1]This work is supported by: ARO contract DAAL03-92-G-0109

2. Device Growth and Layer Structure

The InAlAs/InGaAs/InGaAsP DHBT and the InAlAs/InGaAs SHBT layer structures were grown in a first generation Varian Intervac CBE system using trimethyl indium (TMI), triethyl gallium (TEG), trimethylamine alane (TMAA), 100% arsine and phosphine. Conventional solid source silicon and beryllium were used as n- and p-type dopants, respectively. The substrate temperature was maintained at 520 C for the entire growth sequence. All layers were grown lattice-matched on Fe-doped semi-insulating InP substrates. The InGaAsP quaternary was calibrated to have a bandgap of 1.0 eV and a lattice mismatch of less than 5×10^{-4} arcsec according to photoluminescence and double crystal X-ray diffraction measurements. Further details on the growth characterization can be found in [7][8].

Figure 1 illustrates the layer structures for the SHBT and the DHBT. Both designs feature an n+/N+ InGaAs/InAlAs isotype heterojunction to reduce the effect of the conduction band spike on the emitter access resistance. A thin InP spacer layer was inserted at the emitter-base interface to prevent any beryllium out-diffusion during growth and to act as an etch-stop layer during the base etch. The base and sub-collector layers were heavily doped to minimize the ohmic contact and bulk access resistances. The pre-collector layer thickness and doping level for both devices were identical. The SHBT precollector was entirely InGaAs while the DHBT had a composite collector featuring a 300$\mathring{A}$ InGaAs spacer followed by a 3200$\mathring{A}$ InGaAsP layer. The spacer served to reduce the blocking effect of the base-collector heterojunction by displacing it slightly away from the electrical junction into the collector. Note that no delta-doping was introduced at the discontinuity as is commonly done to further inhibit the effect of the spike.

3. Device Fabrication

Both HBTs were fabricated using a triple mesa fully self-aligned emitter and base process [9]. First, a NiGeAuTiAuTi emitter ohmic metal was evaporated. The second Ti-layer served as a robust mask for the RIE emitter-base mesa etch which was found to sputter off the Au contact layer. The base etch relied on a $CH_4 : Ar : H_2$ RIE followed by a $H_2O_2 : H_2O : H_3PO_4$ selective wet chemical etch. This minimized emitter metal undercut while avoiding the deleterious effects of exposing the base surface to the RIE plasma [10]. The InP etch-stop layer prevented any overetching into the base from occuring. Next, a PECVD SiO_2 layer was deposited and then anisotropically etched in a CF_4 plasma to form sidewalls on the emitter-base mesa. This prevented the base metal from shorting to the exposed emitter sidewall. The effects of sidewall formation on the surface of the base layer and of the emitter sidewall were found to be negligible. Diode samples fabricated without being exposed to any plasma treatments showed comparable ohmic contact resistances and diode ideality factors.

After sidewall formation, the InP spacer was selectively removed in a $H_3PO_4 : HCl$ dip and the PdGeTiAuTi base metal was then deposited self aligned to the emitter metal/ SiO_2 sidewall. This non-alloyed metal system forms a shallow ohmic contact with excellent contact resistivity when deposited on heavily-doped p-type InGaAs [11]. Furthermore, its anealing temperature is compatible with the n-type ohmics and with device reliability.

Again, the top Ti-layer protected the integrity of the mask during the subsequent self-aligned RIE of the base-collector mesa. In order to remove the RIE damaged region, the last part of base-collector mesa was wet-chemical etched in $H_2O_2 : H_2O : H_3PO_4$ for the InGaAs SHBT and in saturated bromine water solution for the InGaAsP DHBT.

Once the subcollector layer was reached, NiGeAuTiAu collector metal was deposited. All the ohmics were anealed simultaneously at 300 C for 20 sec. in a N_2 ambient. Higher temperature annealing was found to degrade the base-collector junction of the DHBT. Then each device was isolated by etching down to the semi-insulating InP substrate. For interconnect support, a low temperature 6000$\AA$ PECVD SiO_2 layer was deposited and via holes were defined and etched. Finally, a thick interconnect metal was evaporated for microwave testing. The final structure of a typical device is illustrated in Figure 2.

4. DC Characterization

DC measurements were performed on a HP-4145 curve tracer. Common emitter output characteristics and Gummel plots for both structures are shown in Figures 3 and 4. Both devices showed adequate β of 75 and 100 for the SHBT and DHBT, respectively, that were independent of device geometry and emitter perifery. The collector breakdown voltage with open base increased dramatically from 2V in the SHBT to 8V in the DHBT. Correspondingly, the collector-base breakdown voltage improved from 2.5V to 12V. The Gummel plots show relatively uniform β down to $10^{-10}A$ of collector current suggesting that even with $V_{cb} = 0$, the DHBT collector barrier does not impede collector current flow. The measured ideality factors are given in the figures.

A careful inspection of the saturation region in the common emitter characteristics shows a decreased collector offset voltage ΔV_{ce} for the DHBT as shown in Figure 6. This is consistent with a corresponding increase in the base-collector turn-on voltage for the InGaAsP device. Note that previosusly reported DHBTs with a delta-doped spike showed no ΔV_{ce} reduction because tunneling through the spike reduced the junction turn-on voltage. By using InGaAsP instead of InP as the precollector material, no delta doping was required to remove the barrier effect and thus the ΔV_{ce} was reduced. The particular I_b-dependence of ΔV_{ce} has been attributed to the difference in ideality factors of the individual base-emitter and base-collector junctions [12].

5. RF Characterization

Microwave measurements were made on a cascade probe station using an HP-8510 microwave network analyser. Figure 6 shows the bias dependence of f_t for a $5\mu m \times 5\mu m$ emitter area DHBT. The monotonic increase of f_t with J_c indicates no sign of the Kirk effect. The valence band discontinuity at the base-collector junction is expected to contain the holes in the base thus preventing base pushout and subsequent degradation of f_t at high collector currents.

A comparison of the current dependence of f_t between both devices biased at $V_{ce} = 2V$

is given in Figure 7. At low and medium values of J_c the SHBT is clealy superior but at $J_c \approx 4 \times 10^4 A/cm^2$ the f_t of the SHBT saturates. This agrees well with a simple calculation of the critical current density required for base pushout given the precollector doping and the electron saturated velocity.

6. Conclusion

InAlAs/InGaAs/InGaAsP DHBTs have been grown and fabricated using a fully self aligned process. The devices demonstrate greatly improved DC characteristics from the SHBT with increased breakdown voltages and lower ΔV_{ce} while maintaining good current gain. RF results confirm the fact that the base-collector discontinuity prevents the reduction of f_t at high currents. Therefore, these devices appear to be excellent candidates for high power application because of large collector voltage swing and good high frequency operation at high current levels.

REFERENCES

[1] Chen et al, "Subpicosecond InP/InGaAs Heterostructure Bipolar Transistors," IEEE EDL, Vol. 10, No. 6, June 1989 pp.267-269.

[2] Jensen et al, "36 GHZ Static Digital Frequency Dividers in AlInAs-GaInAs HBT Technology," Dig. 49th DRC, June 1991, Via-5.

[3] Stanchina et al, "60 GHz AlInAs/GaInAs/InP DHBTs Grown by MOVPE and MBE," Dig. 49th DRC, June 1991, Via-6.

[4] Stanchina et al, "Performance of AlInAs/GaInAs/InP Microwave DHBTs", 5th Conf. InP and Related Materials, April 1993 MA2.

[5] Hong et al, "Microwave Power InP/InGaAs/InP Double-heterojumction Bipolar Transistors", 5th Conf. InP and Related Materials, April 1993 MA3.

[6] Chau et al, "High-Speed, High Breakdown Voltage InP/InGaAs Double Heterojunction Bipolar Transistors Grown by MOMBE", Dig. 51st DRC, June 1993, Session IV-A. [7] Sherwin et al, "Optmization of InGaAs and InP Grown by CBE," J. Crystal Growth, (111), 1991, pp. 594-598.

[8] Sherwin et al, "Chemical Beam Epitaxy of High Quality InAlAs using TMAA," J. Vac. Sci. Technol. B, 10(2), March/April 1992, pp. 943-945.

[9] J.Cowles et al, "InAlAs/InGaAs Heterojunction Bipolar Transistors Grown by Chemical Beam Epitaxy (CBE)", 5th Conf. InP and Related Materials, April 1993 TuP16.

[10] S. J. Pearton, "Dry Etching Techniques and Chemistries for III-V Semiconductors," Mat. Res. Soc. Proc., Vol. 216, 1991, pp. 277-290.

[11] Chen et al, "Ohmic Contact Study for Quantum Effect Transistors and Heterojunction Bipolar Transistors with InGaAs Contact Layers," J. Vac. Sci. Technol. B, 10(6), Nov/Dec 1992, pp.2354-2360.

[12] Chand et al, "Collector-emitter offset voltage in AlGaAs/GaAs heterojunction bipolar transistors," Appl. Phys. Lett. 47 (3), August 1, 1985 pp. 313-315.

InAlAs/InGaAs SHBT			
n+	InGaAs	2000 Å $1x10^{19}$ cm^{-3}	cap
N+	InAlAs	500 Å $3x10^{18}$ cm^{-3}	cap
N	InAlAs	1500 Å $1x10^{18}$ cm^{-3}	emitter
i	InP	100 Å	spacer
p+	InGaAs	700 Å $1x10^{19}$ cm^{-3}	base
n-	InGaAs	3500 Å $1x10^{16}$ cm^{-3}	precollector
n+	InGaAs	6000 Å $1x10^{19}$ cm^{-3}	subcollector
semi-insulating substrate			

InAlAs/InGaAs(P) DHBT			
n+	InGaAs	2000 Å $1x10^{19}$ cm^{-3}	cap
N+	InAlAs	500 Å $3x10^{18}$ cm^{-3}	cap
N	InAlAs	1000 Å $1x10^{18}$ cm^{-3}	emitter
i	InP	100 Å	spacer
p+	InGaAs	700 Å $1x10^{19}$ cm^{-3}	base
n-	InGaAs	300 Å $5x10^{15}$ cm^{-3}	spacer
n-	InGaAsP	3200 Å $1x10^{16}$ cm^{-3}	precollector
n+	InGaAs	6000 Å $1x10^{19}$ cm^{-3}	subcollector
semi-insulating substrate			

Figure 1. Epitaxial Layer Structures for the SHBT and DHBT.

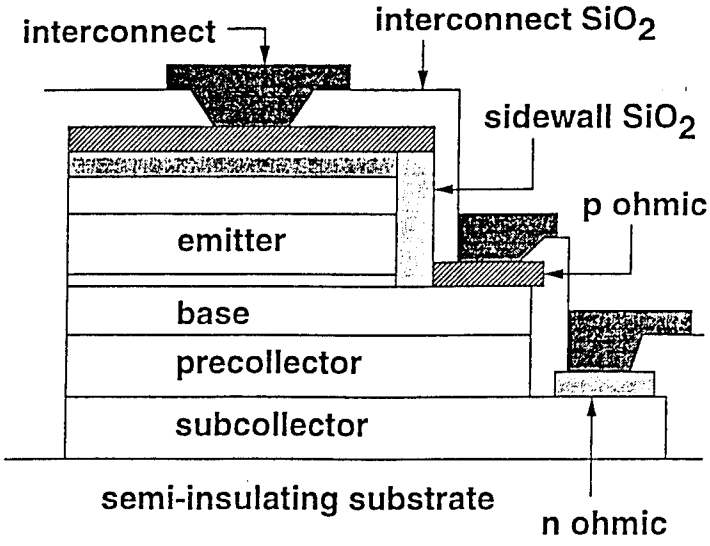


Figure 2. Cross-sectional View of the Device Structure.

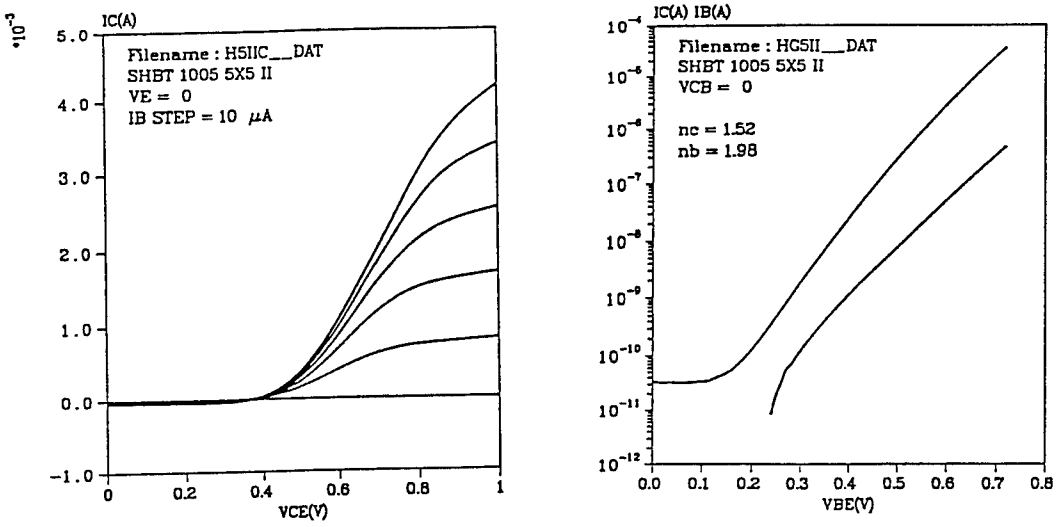


Figure 3. Common Emitter and Gummel Characteristics for the InAlAs/InGaAs SHBT.

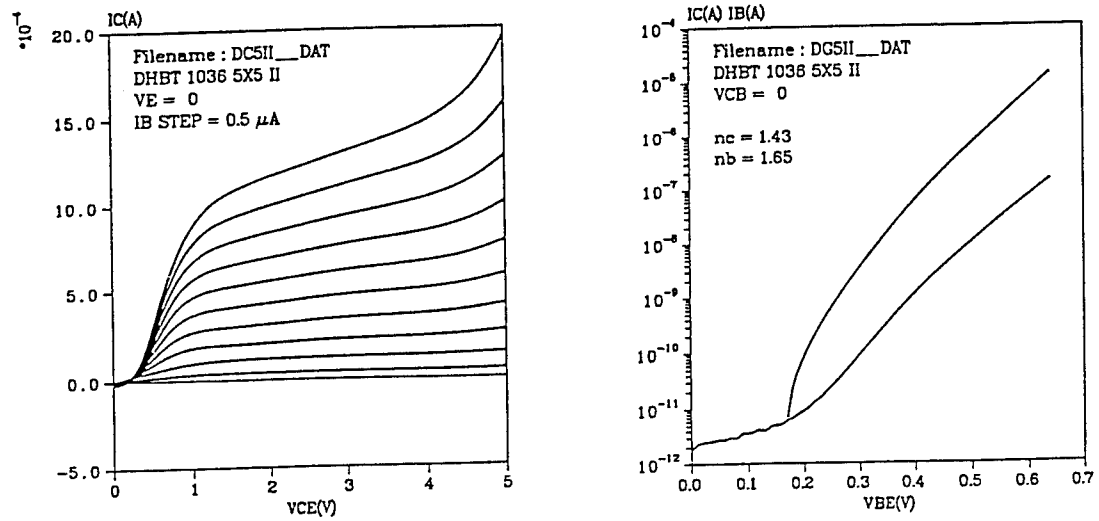


Figure 4. Common Emitter and Gummel Characteristics for the InAlAs/InGaAs/InGaAsP DHBT.

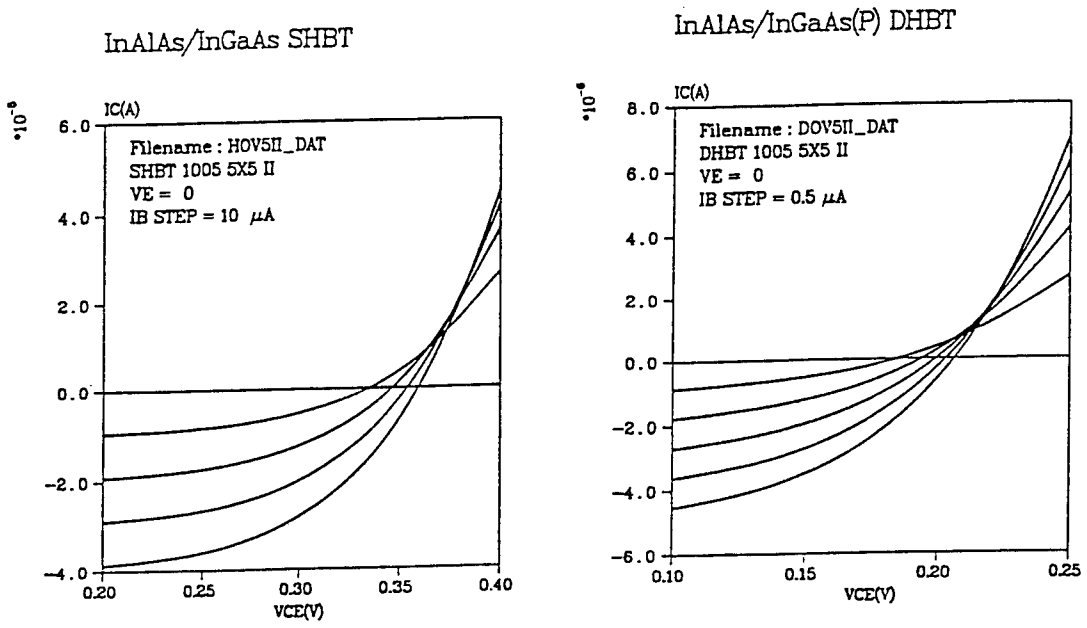


Figure 5. Saturation Region of the SHBT and DHBT Showing ΔVce and its Bias Dependence.

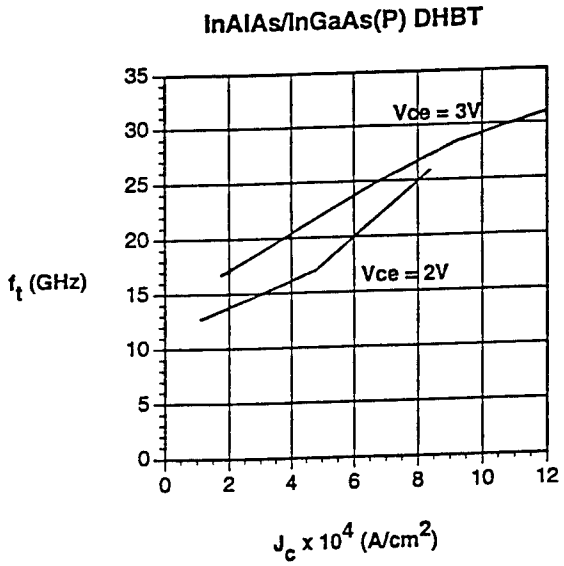


Figure 6. Bias Dependence of f_t for a 5μm x 5μm Emitter DHBT

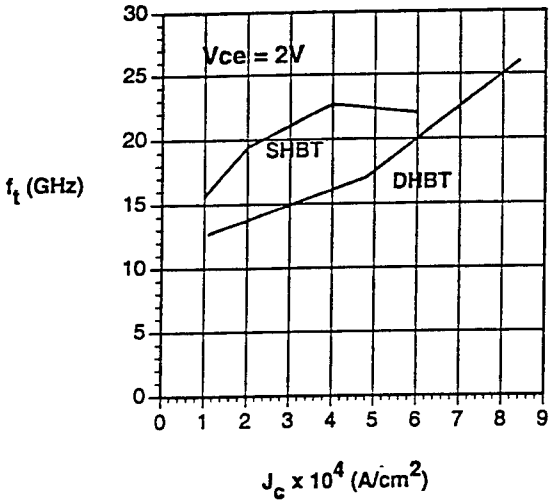


Figure 7. Comparison of f_t vs. J_c at Vce = 2V

GaInP/GaAs Heterostructure Bipolar Transistors with High Gain and High Breakdown Voltages

Z. Abid, S. P. McAlister, and W. R. McKinnon

Institute for Microstructural Sciences,
National Research Council, Ottawa
Canada

Abstract

GaInP/GaAs/GaInP double heterostructure bipolar transistors ,emitter area of 50 x 50 $(\mu m)^2$, with a DC gain of 445 and breakdown voltage (V_{CB0}) of more than 17V have been fabricated. The Gummel plots give an ideality factor of 1.01 for the collector current and 1.1 for the base current, and the devices show gain down to collector currents of 10^{-9} A. A higher gain, more than 800, was recorded for a smaller device, 20 x 50 $(\mu m)^2$ emitter, but the breakdown voltages were lower.

1. Introduction

GaInP/GaAs HBTs offer improved carrier injection into the base compared with the AlGaAs/GaAs devices, due to the higher bandgap of GaInP (1.85eV) [1]. Most of the band offset is taken up in the valence band which leads to better npn devices. GaInP also has a better etching selectivity relative to GaAs and a lower concentration of deep levels compared to AlGaAs material. The breakdown behavior of a GaInP/GaAs HBT can be improved by using a GaInP collector instead of GaAs to create a double heterojunction device (DHBT). The second heterojunction introduces a barrier to the electrons in the collector resulting in a reduced gain [2] and can affect the speed of the device. By shifting this GaInP/GaAs interface away from the base, improvements can be made to the breakdown keeping the high gain [3]. This composite collector design has not been tried in GaInP/GaAs DHBTs. Thus we wished to test this approach to see what improvements could be made to the breakdown behavior in such GaAs based HBTs without significantly compromising the gain. The fabrication and the DC characteristics of devices with different emitter areas will be presented in this paper.

2. Layer structure and fabrication

We have fabricated DHBTs using a conventional wet-chemical and lift-off processes. The wafer was grown by gas-source MBE and, as designed, consists of GaAs cap layer (1500 Å, doped to 5 x 10^{18} cm^{-3} with Si), a GaInP wide bandgap emitter (500Å, doped to 3 x 10^{18} cm^{-3} and 1000Å, doped to 5 x 10^{17} cm^{-3} with Si), an 80Å

0-7803-0894-8/93 $3.00

undoped GaAs spacer layer, then a GaAs base (800Å, doped to 10^{19} cm^{-3} with Be). A 700Å GaAs spacer layer, doped with Si to 10^{16} cm^{-3}, was inserted between the base and the GaInP collector which is doped to about 10^{16} cm^{-3} with Si and has a thickness of 4000Å. The GaAs subcollector is 4000Å thick and is doped with Si to 2 x 10^{18} cm^{-3}. The external GaAs base was not exposed and a thin GaInP layer was left, while defining the emitter mesa, as a passivation layer. Non-alloyed TiPtAu contacts were made on the GaAs layers. The contacts could be optimized to improve the performance of the devices at low applied biases and at high frequency. The base sheet resistance was 300Ω/ sq. The spacing between the metal contact and the mesa edge was 5 μm. Fig. 1 shows the band diagram calculated at equilibrium and at applied biases of V_{BE} =1.0 V and V_{CB} =3V. In these calculations we used the designed doping levels. These diagrams show that the electric field is distributed over both the GaAs and GaInP layers so that the breakdown behavior of the device should improve, but not at the expense of the gain. If the GaInP in the collector is spaced too far from the base then the device will behave as a SHBT and improvements in the breakdown will be lost.

3. Results

DC measurements were performed on the 50 x 50 $(\mu m)^2$ emitter DHBTs. The common emitter characteristics are shown in Fig. 2 for a series of base currents up to 20μA. The offset voltage is 0.41 V and is higher than what one would expect for a DHBT. This is not related to the composite collector design since for unpassivated, DHBTs, made from the same wafer, the offset voltage was ~0.05V. Clearly the variation of the characteristics in the low-bias region is complicated by the use of a passivation ledge as well as the GaAs spacer, and this needs further study. The offset voltage is dependent not only on the band discontinuities of the two heterojunctions but also the recombination loss between the two heterojunctions [4]. The breakdown voltage (V_{CE0}) is 13V while (V_{CB0}) is more than 17V as extracted from the common-base characteristics shown in Fig. 3. The Gummel plot for this device gives an ideality factor of 1.01 for the collector current compared with 1.1 for the base current. For DHBTs made without the passivating ledge the base ideality factor was ~2.1. Fig. 4 shows the Gummel plot of two DHBTs with the same emitter area, 70 x 70 $(\mu m)^2$, but only one of them is passivated. In such unpassivated HBTs the collector and base currents cross in a Gummel plot. This is caused by surface recombination in exposed GaAs. In our present devices the surface recombination component is not dominant in the base current. The passivation ledge thickness and width are important in determining its effectiveness in improving the device characteristics [5]. The ledge thickness in our devices was ~200Å, and its length 5μm. For the ledge to be useful it should be fully depleted and sufficiently long to suppress the base contact recombination current. This can limit the device performance since minority carriers diffusing sideways along the ledge can reach the base contact before recombining. The success of our devices suggest the dimensions we have are not close to limiting values. A DC and small signal gain close to 450 were measured over a wide range of applied bias, V_{CE}=1.4V to 2.4V and at V_{CB}=3.0V (Fig. 5). The base-collector voltage was applied to decrease the blocking effect of the heterojunction at the collector, thus reducing the excess charge storage. The comparatively flat gain curve is

common in GaInP wide bandgap HBTs, but is not easily achieved for AlGaAs/GaAs HBTs.

Similar measurements were performed on the 20 x 50 $(\mu m)^2$ emitter DHBTs. The common-emitter I/V curves are presented in Fig. 6. The DC and small signal gain shows higher gains (as high as ~850) compared with the previous DHBTs, but the breakdown voltages were lower, V_{CE0} is 9.3V while V_{CB0} is about 12.5 V.

In small devices, in particular the smaller of our devices, and at collector currents approaching 100mA the collector currents decreased slightly with common-emitter bias. This is caused by self-heating in the device and the negative temperature dependence of the current gain [6]. This heating is dependent on the current density and the thermal resistance. Both of these are smaller in the larger of our devices

4. Conclusion

We have designed and fabricated high gain GaInP/GaAs DHBTs which exhibit high breakdown voltages. A 20 x 50 $(\mu m)^2$ emitter device showed a DC and small signal gain of more than 800. A constant gain of around 400 over more than five order of magnitude of the collector current was measured for a 50 x 50 $(\mu m)^2$ device. This indicated the importance of the composite collector design and the passivation of the base surface in increasing the gain even at very low collector currents, while keeping the breakdown voltages high.

Aknowledgment :

We would like to thank P. Marshall, E. Guzzo, M. Lumsden for their technical help, and to Dr F. Williamson and Prof. M. I. Nathan from the University of Minnesota for supplying the epitaxial structures.

REFERENCES :

[1] H. Kroemer, J. Vac. Sci. Tech. B., **1**, 126 (1983).
[2] A. W. Hanson, S. A. Stockman, G. E. Stillman, IEEE Elec. Dev. Lett., **14**, 25 (1983).
[3] A. Feygenson, D. Ritter, R. A. Hamm, P. R. Smith, R. K. Montgomery, R. D. Yadvish, H. Temkin and M. B. Panish, Electronics Lett., **28**, 607 (1992).
[4] S.-C. Lee and H.-H. Lin, J. Appl. Phys., **59**, 1688 (1986).
[5] W. Liu and J. S. Harris, Solid-St. Electron., **35**, 891 (1992).
[6] G. B. Gao, Z. F. Fan and H. Morkoc, Appl. Phys. Lett., **61**, 198 (1992).

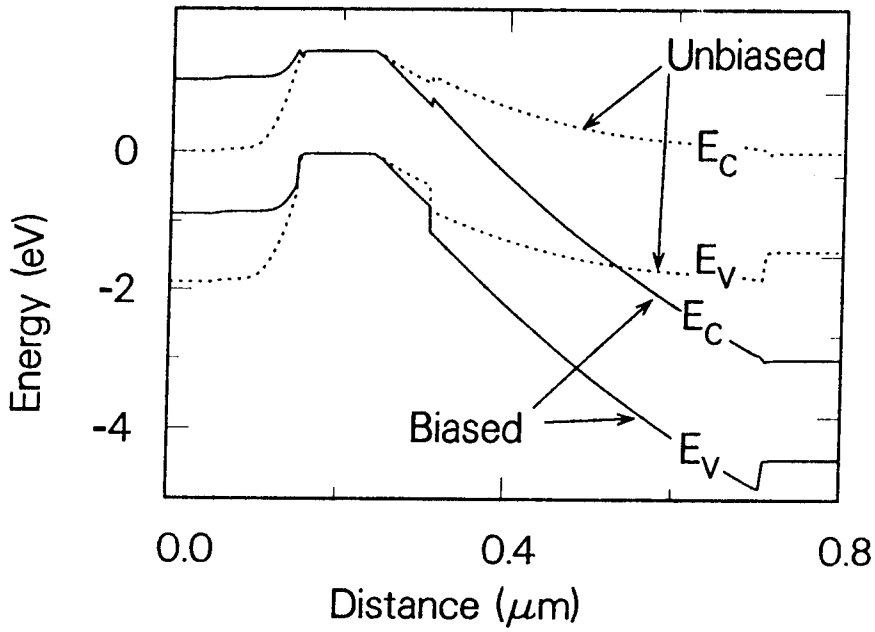


Fig. 1 : Band diagram for the DHBT layer structure. The dashed line is for equilibrium and the solid line is for the device biased at V_{BE}=1.0 , V_{CB}=3.0V

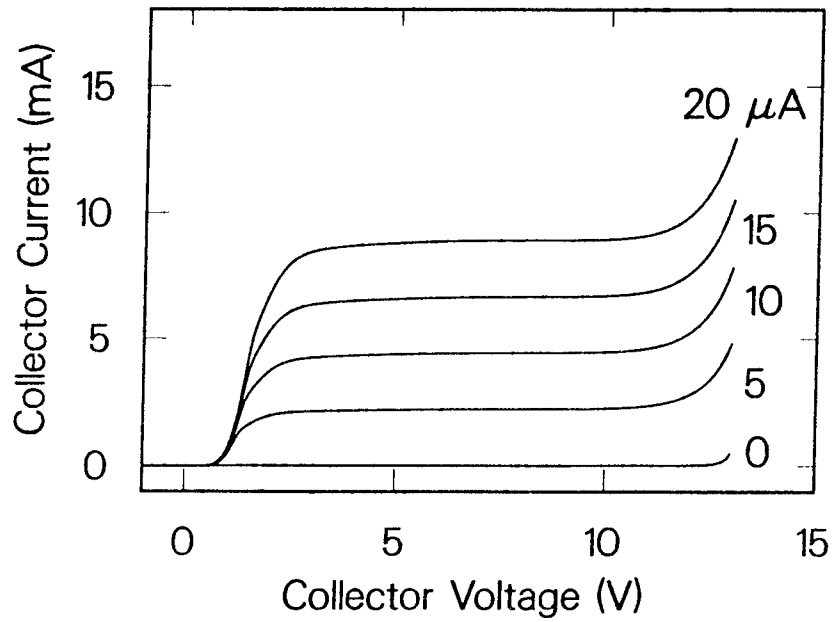


Fig. 2 : Common emitter characteristics for a 50 x 50 $(\mu m)^2$ emitter DHBT for a series of base currents

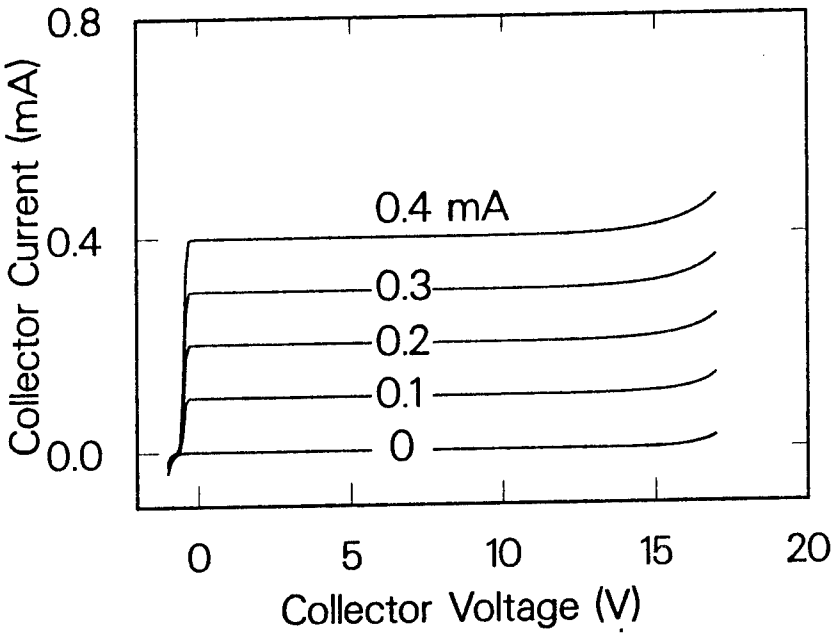


Fig. 3 : The common-base characteristics for the 50x50 $(\mu m)^2$ emitter DHBT for the emitter currents indicated.

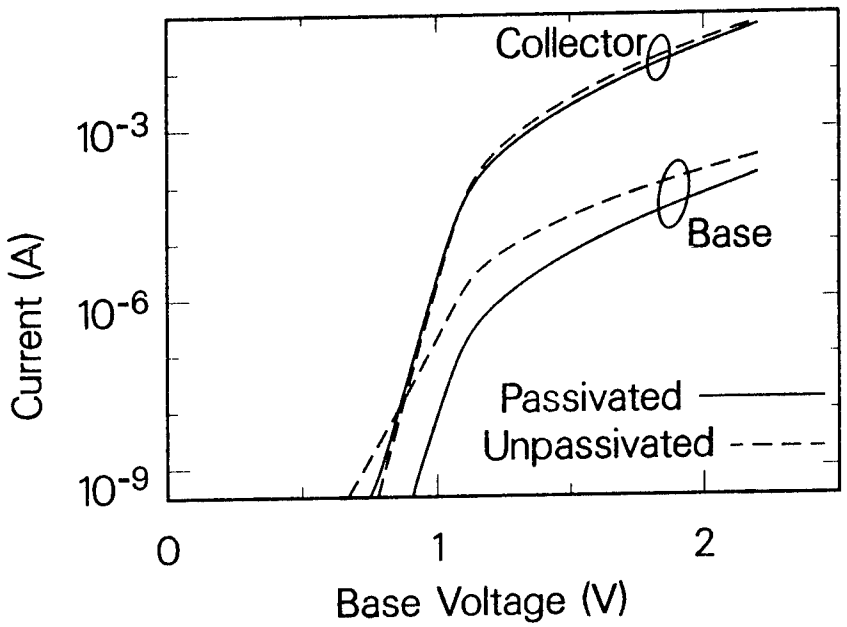


Fig. 4 : Gummel plot for two 70x70 $(\mu m)^2$ emitter DHBTs. The passivated device had V_{CB}=3.0V, and for the unpassivated device V_{CB} was 2.0V.

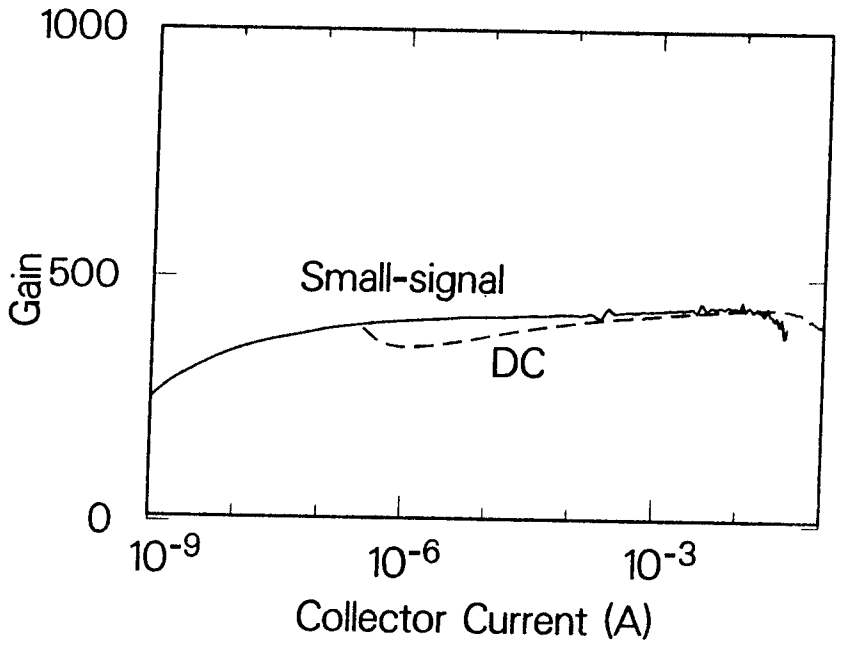


Fig. 5 : The DC and small signal gain at V_{CB}=3.0V, for the 50 x 50 $(\mu m)^2$ emitter device.

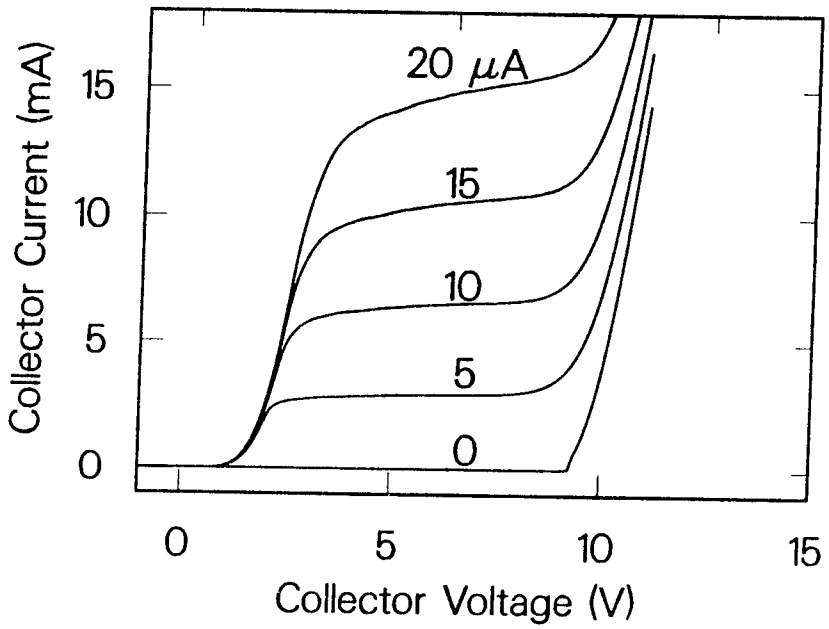


Fig. 6 : Common-emitter I-V curves for a 20x50 $(\mu m)^2$ emitter DHBT for the indicated base currents.

New Analytic Determination of f_T, f_{MAX} and the Frequency Dependence of Current Gain and Power Gain in HBTs

David R. Pehlke and Dimitris Pavlidis

Solid-State Electronics Laboratory
University of Michigan
1301 Beal Avenue
Ann Arbor, Michigan 48109-2122
(313)-763-6132

ABSTRACT

A new analytic formalism describing the frequency dependence of current gain and power gain of HBTs based on the T-Model equivalent circuit is presented. It is found that while extrapolation serves as a conservative benchmark for HBT high-speed device comparison, it often significantly underestimates the actual intercept frequency for f_T, with an experimental example presented here with error of 50%. The frequency dependence of G_{MAX} shows it to be unsuitable for extrapolation to find f_{MAX}, while extrapolation of Unilateral Gain (U) will generally in the absence of resonances yield accurate values for f_{MAX}. The conditions for resonance and resonance frequencies for U are analytically investigated here, and new analytic expressions are developed and experimentally demonstrated for highly accurate determination of the f_T and f_{MAX} intercept frequencies of HBTs.

1. INTRODUCTION

Heterojunction Bipolar Transistors (HBTs) have recently demonstrated significant advances in measured current gain and power gain at high frequencies. An AlGaAs/GaAs HBT with ballistic collector design has exhibited an $|h_{21}|^2$ of approximately 12.5 dB at 25 GHz [1], and unilateral gains have been measured exceeding 17.5 dB at 40GHz [2]. The $|h_{21}|^2$,U, and G_{MAX} unity intercept frequencies f_T and f_{MAX}, respectively for these and other state-of-the-art devices lay well beyond the maximum frequencies of currently available measurement systems. In order to determine f_T and f_{MAX}, extrapolation assuming a single pole roll-off of 20 dB/decade is often employed from the highest frequency for which gain can be measured, and in the two cases cited above, yielded both those f_T and f_{MAX} values above 100 GHz and 300 GHz respectively. In many cases for HBTs; however, $|h_{21}|^2$, G_{MAX} and U do not obey a single pole approximation and extrapolation does not yield an accurate indication of the actual f_T and f_{MAX} intercepts, but merely a conservative benchmark for comparison of high-speed devices. There have been many reports of f_T and f_{MAX} relationships derived for HBTs, but they have relied on hybrid-pi equivalent circuit topologies, with most being single pole extrapolation models [3,4,5], and the most accurate involving transcendental expressions requiring iterative or numerical solution [6,7,8]. In order to better describe HBT gain roll-off and calculation of the actual intercept frequencies f_T and f_{MAX}, this paper presents a new formalism based on the T-model equivalent circuit that allows analytic determination of the frequency dependent current gain and power gain roll-off along with results from new and more

0-7803-0894-8/93 $3.00

complete expressions for f_T and f_{MAX} that take into account coupling between various charging times, coupling between charging and transit times, and the effects of parasitic elements such as inductances and other series access components that are not included in traditional expressions. The analysis is then used to investigate analytic techniques for HBT optimization and conditions for resonance in unilateral gain.

2. DEVICE FABRICATION

The experimental results presented here are from measured S-parameters of self-aligned GaAs/AlGaAs HBTs fabricated at the University of Michigan using an MBE-grown layer structure with 1000 Å base doped $1 \cdot 10^{19}/cm^3$ and 3000Å pre-collector doped $5 \cdot 10^{16}/cm^3$. The base ohmic metal is Pd/Zn/Pd/Au and is self-aligned to the emitter using a dry and wet etching sequence that creates an emitter ohmic overhang profile for lift-off. A combination of trench mesa isolation and hydrogen (H^+) ion implantation are used to isolate parasitic areas from the active device region and reduce the overall device dimensions. Figure 1 shows a three dimensional technology schematic of a typical device through the trench isolation step. The active base-emitter area for the single emitter finger geometry investigated here is 2μm x 10μm, and the devices were tested using a Cascade prober and an hp8510B S-parameter test system from 0.5GHz to 25.5GHz.

3. EQUIVALENT CIRCUIT MODEL

The HBT equivalent circuit model used for this work is shown in Figure 2. It is a conventional T-topology with the addition of a distributive capacitance element (C_{BX}) in the base series access impedance [9]. The base-emitter impedance is represented so that only the diffusion capacitance (C_D) and base-emitter resistance (R_{BE}) constitute the path for current which is multiplied by the gain term $\alpha(\omega)$ in the collector branch. The base-emitter depletion capacitance is modeled as a parasitic leakage element that adds displacement current between base and emitter at high frequencies, but does not contribute particle current in the form of minority carrier flow in the base which might then be seen at the collector[3].

4. HBT CURRENT GAIN ROLL-OFF

The gain roll-off behavior of HBTs is generally thought to be dominated by a single pole, but closer investigation of the $h_{21}(s)$ transfer function yields a slightly more complicated picture of coupled transit times, charging times and higher order combinations of equivalent circuit elements that manifest themselves as additional zeros and poles. For various combinations of these equivalent circuit values, the additional poles and zeros may be low enough in frequency to distort the single pole 20dB/decade slope and affect the current gain unity intercept f_T. In order to calculate the pole and zero frequencies, we start with the current gain expression written in terms of $s = j\omega$:

$$h_{21}(s) = \frac{N(s)}{D(s)} = \frac{\alpha(s)Z_{BC}(s) - Z_{BE}(s) - z_E(s)}{Z_{BE}(s) + z_E(s) + Z_{BC}(s) + z_C(s) - \alpha(s)Z_{BC}(s)} \quad (1)$$

and expand $N(s)$ and $D(s)$ out into polynomial expressions whose roots can then be solved for in terms of ω. If the zeros and higher order poles are much larger in frequency than the first order pole (f_{3dB}), then extrapolation will predict f_T accurately and it may be expressed as :

$$f_{T_{EXTRAP}} = \beta_0 f_{3dB} = \frac{1}{2\pi\tau_{EC}} \quad (2)$$

τ_{EC} is the total forward delay for the device and may be expressed as :

$$\tau_{EC} = \frac{C_{BC}(R_{BE} + R_E + R_C)}{\alpha_0} + R_{BE}\left(\frac{C_{BE}(2-\alpha_0)}{\alpha_0} + \frac{C_D}{\beta_0}\right) + \tau_B\left[1 + \frac{1}{\beta_0\left(\frac{\pi^2}{8}\right)}\right] + \tau_{PCD} \quad (3)$$

where τ_B is the base transit time and τ_{PCD} is the pre-collector transit delay. Note that this is slightly more complete than, but reduces to, the conventionally accepted form of τ_{EC} for the case where $\alpha_0 = 1$:

$$\tau_{EC} = C_{BC}(R_{BE} + R_E + R_C) + R_{BE}C_{BE} + \tau_B + \tau_{PCD} \quad (4)$$

For the case where the first order zero (f_{zero}) is lower in frequency and closer to the first order pole (f_{3dB}), then its presence will make the slope of the $|h_{21}|^2$ curve less negative and extend out the actual unity intercept f_T. An example of this effect of the zeros presence at lower frequency is shown in Figure 3. The f_{3dB} and f_{zero} are calculated analytically from the expanded polynomial expression of equation (1) and correspond well to the 3dB points of both the T-model and measured data.

The actual f_T intercept often does not correspond with $f_{T_{EXTRAP}} = \frac{1}{2\pi\tau_{EC}}$ because of the low frequency positioning of f_{zero}, and so for the calculation of the actual intercept $f_{T_{ACTUAL}}$, we evaluate the condition where $|h_{21}| = 1$ and solve for f_T. Based on the expression (1), we find that $f = f_T$ when :

$$Re[\alpha] - \frac{1}{2} = Re\left[\frac{Z_{BE} + z_E + \alpha z_C}{Z_{BC} + z_C}\right] \quad (5)$$

and expanding this into a polynomial expression in ω, we are able to find a complicated but direct and analytic expression for $f_{T_{ACTUAL}}$.

For the case of our experimental device and directly extracted equivalent circuit [10] of Figure 3, we find that $f_{T_{ACTUAL}} = 60.7GHz$. As a matter of contrast, extrapolation can produce a range of numbers, from the correct extrapolation from f_{3dB} of $f_{T_{EXTRAP}} = 30.6GHz$ all the way up to $36.3GHz$ depending upon which of the measured points are chosen for the extrapolation. It will

always be true that extrapolation will not overestimate $f_{T_{ACTUAL}}$ (as long as the equivalent circuit remains valid up to that frequency) but will generally underestimate $f_{T_{ACTUAL}}$ and in some cases, significantly.

As an investigation of the extrapolation accuracy, our reference result with $C_{BC} = 0.064$pF was modified in order to see the effect a range of C_{BC} values on $f_{T_{EXTRAP}}$ and $f_{T_{ACTUAL}}$ keeping all other element values the same. For the equivalent circuit investigated, a wide range of extrapolation accuracy is seen in Figure 4 as both $f_{T_{ACTUAL}}$ and $f_{T_{EXTRAP}}$ change with C_{BC}, with up to a factor of 2 difference between the two at $C_{BC} = 0.03$pF. At small C_{BC} values, the zero frequency is very large and doesn't significantly impact the single pole approximation up to f_T, so that $f_{T_{EXTRAP}}$ is very close to the actual intercept. As C_{BC} increases, the zero frequency starts to decrease and get closer to $f_{T_{EXTRAP}}$, so that $f_{T_{ACTUAL}}$ is extended out further in frequency and starts to increase to a peak at 80.25 GHz at $C_{BC} = 0.03$pF. As C_{BC} increases further still; however, it starts to dominate the h_{21} expression as part of a single dominant pole, and return h_{21} to a valid single pole approximation, with $f_{T_{ACTUAL}}$ once again approaching $f_{T_{EXTRAP}}$.

In order to best determine the extent to which devices are near optimum, it is important to evaluate the limitations to the devices behavior and figure out which dominates it. For the HBT current gain intercept at f_T, we can determine the transit time limited f_{T_τ} by assuming all the RC charging components are zero and likewise for the RC charging time limited $f_{T_{RC}}$ by assuming all its transit time components are zero. It is found that at small pre-collector thicknesses, $f_{T_{EXTRAP}}$ is dominated by its RC charging terms as C_{BC} becomes large. The opposite is true at large W_{BC} where transit time becomes large and dominates f_T. The use of analytic expressions for these limitations $f_{T_{RC}}$ and f_{T_τ} allows direct optimization to maximize $f_{T_{EXTRAP}}$ and $f_{T_{ACTUAL}}$ by looking at what limits their behavior most severely.

5. POWER GAIN

The evaluation of power gain must involve both the unilateral gain (U) and the maximum available gain (G_{MAX}). Their common intercept at f_{MAX} may be evaluated analytically by expressing unilateral gain for the HBT T-Model as

$$U = \frac{|\alpha Z_{BC}|^2}{4[Re[z_B]Re[Z_{BE} + z_E + Z_{BC} + z_C - \alpha Z_{BC}] + Re[Z_{BE} + z_E]Re[Z_{BC} + z_C]]} \tag{6}$$

One can then expand it into a polynomial in $\omega = 2\pi f$ and find the frequency at which it goes to 1. Figure 5 shows the measured maximum available gain and unilateral gain and their excellent agreement with the corresponding power gain terms of the T-model which intersect almost exactly at the analytically calculated $f_{MAX}^{CALC} = 38.3GHz$. Although the unilateral gain roll-off is fairly well behaved at 20dB/decade, G_{MAX} deviates significantly from a single pole approximation and extrapolation for this parameter is generally in error. Unilateral gain can also deviate significantly

from single pole behavior as transit times can induce resonances and cause severely non-ideal peaking in the gain as the denominator of equation (6) goes to 0 [11]. The condition for resonance to occur from equation (6) is

$$C_{BC}\left(Re[z_B](Re[Z_{BE}+z_E]+R_C)+Re[Z_{BE}+z_E]R_C\right) < \alpha_0 Re[z_B]\left(\frac{8\tau_B}{\pi^2}+\tau_{PCD}\right) \qquad (7)$$

Figure 6 shows the experimental result with $C_{BC} = 0.064$pF and $R_C = 14.9\Omega$ along with its T-model calculated U up to 800GHz and because it fails to satisfy the condition for resonance , does not exhibit such an effect. Once C_{BC} and R_C are decreased to 0.01pF and 2Ω respectively, the left hand side of equation (7) becomes small enough to satisfy the condition and a significant resonance shows up at 260 GHz. These conditions for resonance and the extent to which they affect unilateral gain at higher frequencies are important to understand and predict.

6. CONCLUSION

A new formalism for the description of gain roll-off and calculation of intercept frequencies f_T and f_{MAX} is presented along with results verified by experimental measurements on GaAs/AlGaAs self-aligned HBTs. Extrapolation approaches are often found to be inaccurate, but always in underestimating performance, and still serve as a conservative benchmark for comparison of device gain at higher frequencies. The analytic calculation of pole and zero frequencies is used to explain non-idealities in gain roll-off for $|h_{21}|^2$ and provides the basis for accurate extrapolation criteria in determining f_T. Also, the behavior of resonance in unilateral gain is expressed analytically based on the T-Model equivalent circuit topology, with the resonance condition derived in analytic form.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Army Research Office, URI program [Contract No. DAAL-03-92-G-0109], and Bell Northern Research.

[1] T. Ishibashi and Y. Yamauchi. *A Possible Near-Ballistic Collection in an AlGaAs/GaAs HBT with a Modified Collector Structure* IEEE Transactions on Electron Devices, April 1988, Vol.35, No. 4, pp.401-404

[2] W.J. Ho, N.L. Wang, M.F. Chang, A. Sailer, and J.A. Higgins. *Self-Aligned, Emitter-Edge-Passivated AlGaAs/GaAs Heterojunction Bipolar Transistors with Extrapolated Maximum Oscillation Frequency of 350GHz*, Proceedings of 1992 Device Research Conference, p. IVA-1

[3] R. L. Pritchard. *Electrical Characteristics of Transistors*, McGraw-Hill, New York, 1967, pp. 278-326

[4] W. Liu, D. Costa, and J.S. Harris Jr. *Derivation of the Emitter-Collector Transit Time of Heterojunction Bipolar Transistors* Solid-State Electronics, Vol.35, No. 4, pp. 541-545, 1992

[5] L. H. Camnitz and N. Moll. *An Analysis of the Cutoff Frequency Behavior of Microwave Heterostructure Bipolar Transistors* Compound Semiconductor Transistors : Physics and Technology, IEEE Press, ed. S. Tiwari, pp. 21-45, 1993

[6] M. Das. *High-Frequency Performance Limitations of Millimeter-Wave Heterojunction Bipolar Transistors* IEEE Transactions on Electron Devices, May 1988, Vol.35, No. 5, pp.604-614

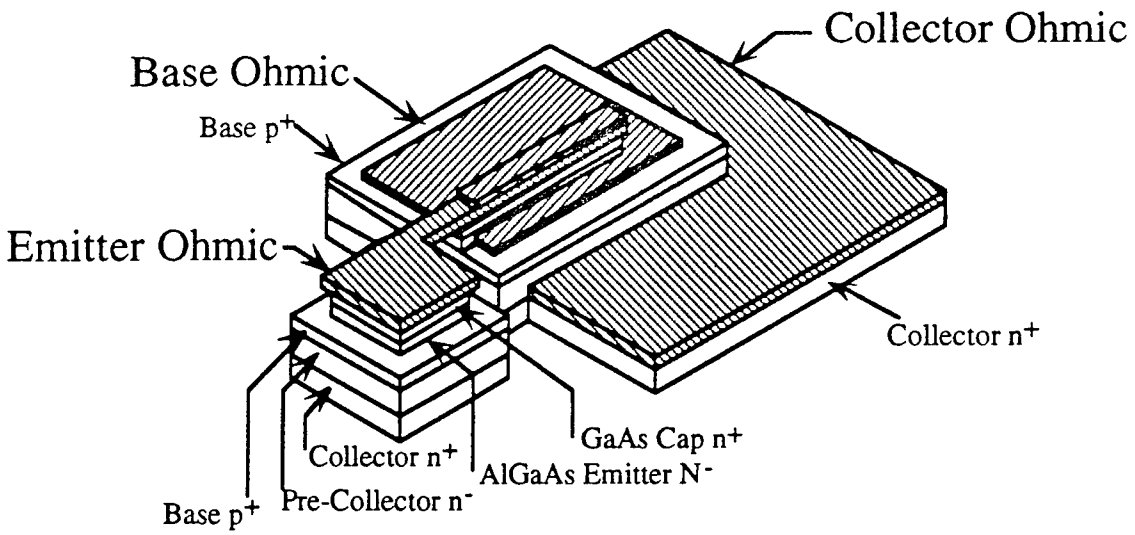
[7] A. Laser and D. Pulfrey. *Reconciliation of Methods for Estimating f_{MAX} for Microwave Heterojunction Transistors* IEEE Transactions on Electron Devices, August 1991, Vol.38, No. 8, pp.1685-1692

[8] S. Tiwari. *Frequency Dependence of the Unilateral Gain in Bipolar Transistors* IEEE Electron Device Letters, December 1989, Vol.10, No. 12, pp.574-576

[9] W. Liu, D. Costa, and J.S. Harris Jr. *A Simplified Model for the Distributed Base Contact Impedance in Heterojunction Bipolar Transistors* Solid-State Electronics, Vol.35, No. 4, pp. 547-552, 1992

[10] D. R. Pehlke and D. Pavlidis. *Evaluation of the Factors Determining HBT High-Frequency Performance by Direct Analysis of S-Parameter Data* IEEE Transactions on Microwave Theory and Techniques, December 1992, Vol.40, No. 12, pp.2367-2373

[11] N. Nagli. *Physical Origin of the Negative Differential Output Resistance of Heterojunction Bipolar Transistors* IEEE Electron Device Letters, 1988, Vol.9, No. 3, p.113



- Base Metal Self-Aligned to Emitter Ohmic
- Combination H^+ Ion Implantation/Trench Isolation
- 2um x 10um Single Emitter Finger Geometry

Figure 1 Technology schematic showing typical device completed through the trench isolation process step.

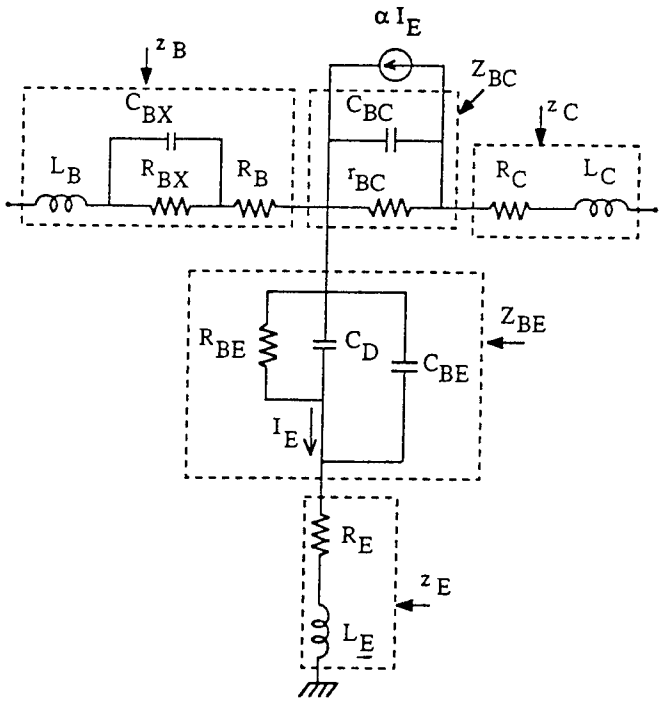


Figure 2 HBT T-Model equivalent circuit.

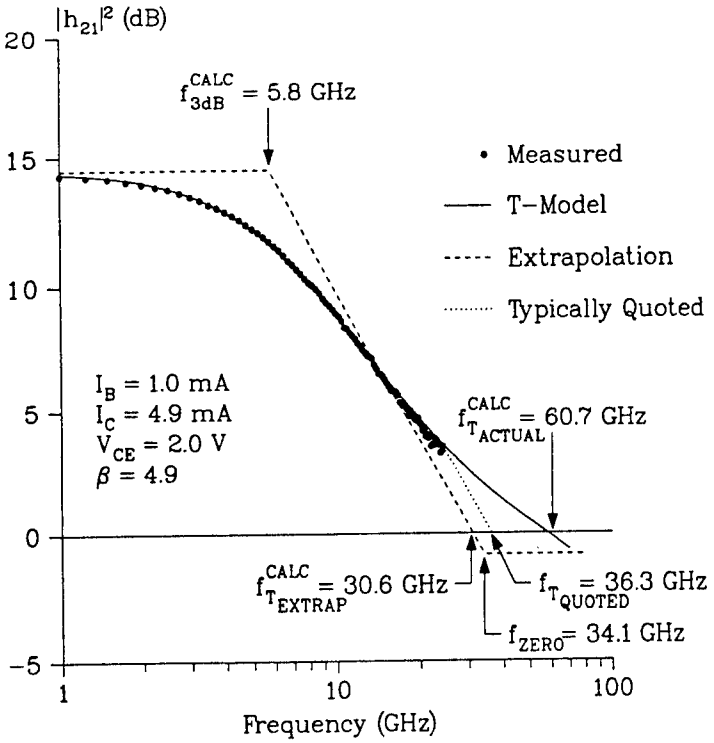


Figure 3 Magnitude of $|h_{21}|^2$ as a function of frequency showing the the effect of a low frequency zero in extending f_T out further in frequency, the actual f_T intercept frequency ($f_{T_{ACTUAL}}$), the extrapolated $f_{T_{EXTRAP}}$ and the typically quoted extrapolation from the last measured frequency $f_{T_{QUOTED}}$. Because extrapolation only underestimates $f_{T_{ACTUAL}}$, $f_{T_{EXTRAP}} < f_{T_{QUOTED}} < f_{T_{ACTUAL}}$.

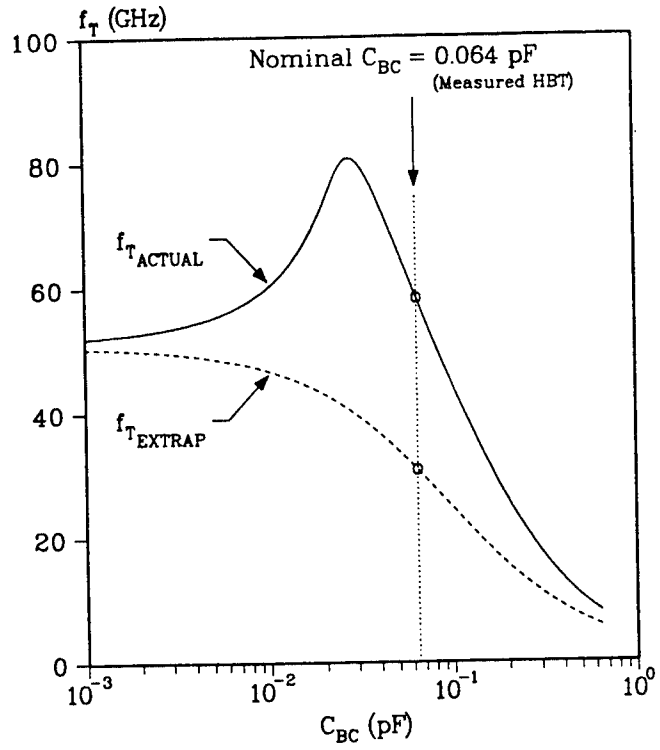


Figure 4 Effect of C_{BC} on both the extrapolated and actual f_T. A factor of two difference between the two is achieved at C_{BC} = 0.03pF

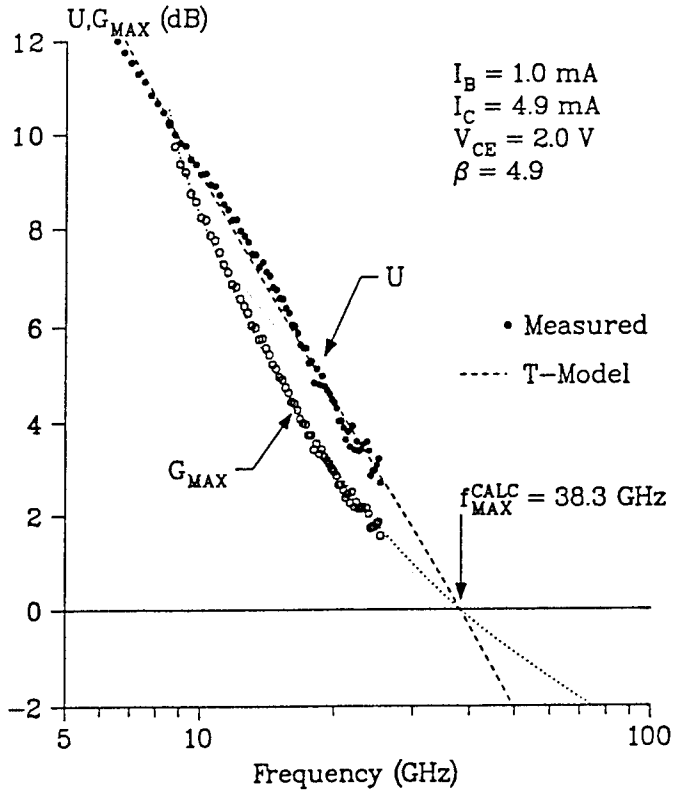


Figure 5 Power gain as a function of frequency showing the unilateral gain (U), and maximum available gain (G_{MAX}) along with the excellent agreement to its T-model equivalent circuit and common unity intercept at its analytically calculated $f_{MAX}^{CALC} = 38.3GHz$

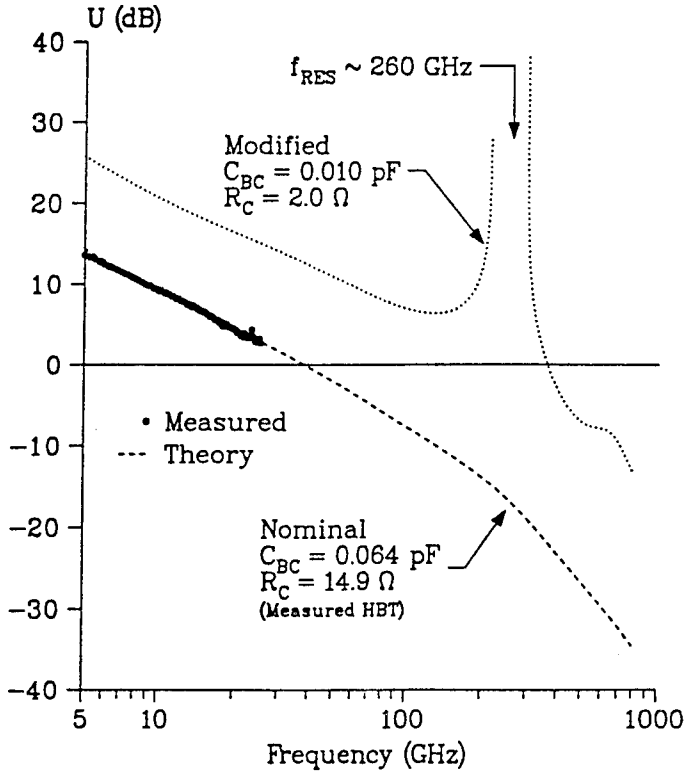


Figure 6 Unilateral gain vs. frequency with no resonance in measured device when theoretically calaulated out to 800GHz. Modifications to reduce C_{BC} and R_C serve to force the resonance condition to be satisfied, and the device resonates at $f_{res} = 260GHz$.

DESIGN AND FABRICATION OF THERMALLY-STABLE AlGaAs/GaAs MICROWAVE POWER HBTs

B. Bayraktaroglu, R. Fitch, J. Barrette, R. Scherer, L. Kehias, and C.I. Huang

Wright Laboratory, Solid State Electronics Directorate
Wright-Patterson AFB, OH 45433

ABSTRACT

Record power density performance of AlGaAs/GaAs microwave power heterojunction bipolar transistors (HBTs) was accomplished through the use of novel design and fabrication techniques. Thermally-stable operation of HBTs up to their electronic limitation (10 mW/μm^2 output power density at 10 GHz with 0.6 W CW output power, 7.1 dB gain and 60% PAE) was attained. The design of the HBT was based on a detailed electro-thermal device analysis which revealed the necessity to provide an effective heat transfer path between heat sources in a multi-emitter power device. Excess heat was transferred out of the device using thermal shunt and thermal lens techniques. The thermal resistance of the device was lowered by a factor of 2.5-3 compared to conventional devices.

I. INTRODUCTION

One of the main applications of AlGaAs/GaAs HBTs is power generation and amplification at microwave frequencies. The vertical structure of an HBT, like its Si bipolar junction transistor (BJT) counterpart, results in highly compact devices. This compactness, when coupled with large emitter areas (2-3 μm emitter finger widths) resulting from low base sheet resistance, produces high power microwave devices. Rapid progress has been made in this area over the last 5 years to make AlGaAs/GaAs HBTs competitive with advanced GaAs FETs also used for this application. High efficiency and high power [1-2] operations have been demonstrated with discrete devices up to Ku-band. MMIC applications have also been demonstrated up to X-band frequencies [3-5].

All power HBTs produced to date have shown thermally-limited operation. That is, thermally induced instabilities in the device set the highest performance limitations. These thermal instabilities may be in the form of high junction temperature rise causing device burn out or non-destructive but performance limiting "current-crush" effects [6]. The electronic limitations of GaAs HBTs are considerably higher than thermal limits as demonstrated by substantial power density improvements achieved under short pulse operation [7].

Electrothermal analysis of conventional HBTs fabricated on GaAs substrates has shown that "current-crush" is primarily responsible for the thermal limitations of multi-emitter finger power devices. This instability is caused by the negative temperature coefficient of the base-emitter voltage, V_{be}, giving rise to uneven distribution of base current among emitter fingers. After a certain threshold power level (determined by parasitic resistances in the device), all the base current flows through a single emitter causing a drastic increase in the temperature of this finger, i.e. hot spots. Ballast resistors are commonly used with bipolar transistors to increase the threshold of instability to higher values [8]. This approach, however, lowers the already limited microwave gain of bipolar transistors limiting their use as efficient amplifiers at higher microwave frequencies.

0-7803-0894-8/93 $3.00

Recently, a thermal shunt technique was developed for AlGaAs/GaAs HBTs to prevent "current-crush" [9]. This technique relies on thermally and electrically connecting all emitter fingers of a power transistor to avoid temperature differential between them. Because the temperature of all fingers are kept the same with this approach, the base current distribution is also uniform thereby eliminating the occurrence of current-crush and the attendant hot spot formation.

This paper discusses design and fabrication issues related to thermally-stable HBTs employing "thermal shunt" and "thermal lens" techniques. Although these techniques were shown to stabilize AlGaAs/GaAs HBTs without degrading their microwave performances, they are equally applicable to all other bipolar and heterojunction transistors including Si BJTs.

II. DESIGN AND FABRICATION

The samples whose epitaxial layer structures are shown in Figure 1 were grown using Metal Organic Chemical Vapor Deposition (MOCVD). Si and C were used as n and p-type dopants. The base layer doping was maintained at 5 x 10^{19} cm^{-3} to achieve low base sheet resistance. The common-emitter current gain value (ß) achieved with this doping level was 30. A rapid drop in ß beyond 5 x 10^{19} cm^{-3} prevented the use of higher doping levels. The collector layer was 1 µm thick. The thickness of this layer strongly influences the current gain cutoff frequency, f_T, and device breakdown voltage. The thickness chosen was a compromise to maintain both values high. The collector doping level was chosen to maximize collector current density while maintaining full collector depletion under high bias conditions. An $In_{0.5}Ga_{0.5}As$ emitter cap layer was employed with all structures to facilitate the use of non-alloyed emitter contacts. The emitter-base junction was not intentionally graded, although a transition region of 10 nm may be present at this interface due to finite response time of gases in a MOCVD system.

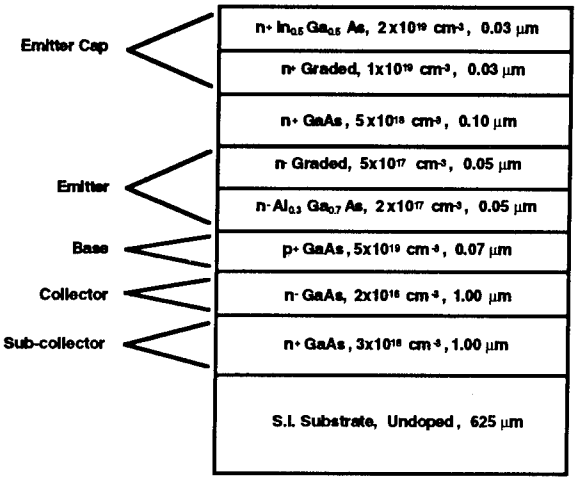


Figure 1. HBT structure: material, doping concentration and thickness.

Several device layout configurations were investigated. Common to all designs was the emitter-base contacts, which were self-aligned to maximize microwave gain performance. The emitter of all devices was made of a parallel connection of multiple emitter discs arranged linearly. As shown in Figure 2a, base contacts completely surround each emitter. The diameter of emitter discs was constant within a given device but was varied between 2 and 3 µm from device-to-device. The spacing between emitter discs was kept the same as the diameter. Collector contacts were spaced 1.5 µm from the edge of the base contact and surrounded it on 3 sides. An oxygen ion implantation was used to convert all areas outside of the active device to semi-insulating. Note that both base and collector contacts extend continuously from the active area to implant isolated areas to facilitate electrical contacts between device terminals and bonding pads. Emitter discs were connected in parallel using a thick air bridge metal as shown in Figure 2. This bridge was extended over to the isolated region and was terminated in large metal pads. As will be discussed below, thick air bridges used as pads on GaAs acted as effective heat spread areas. The ratio of heat spread to emitter areas was typically 20.

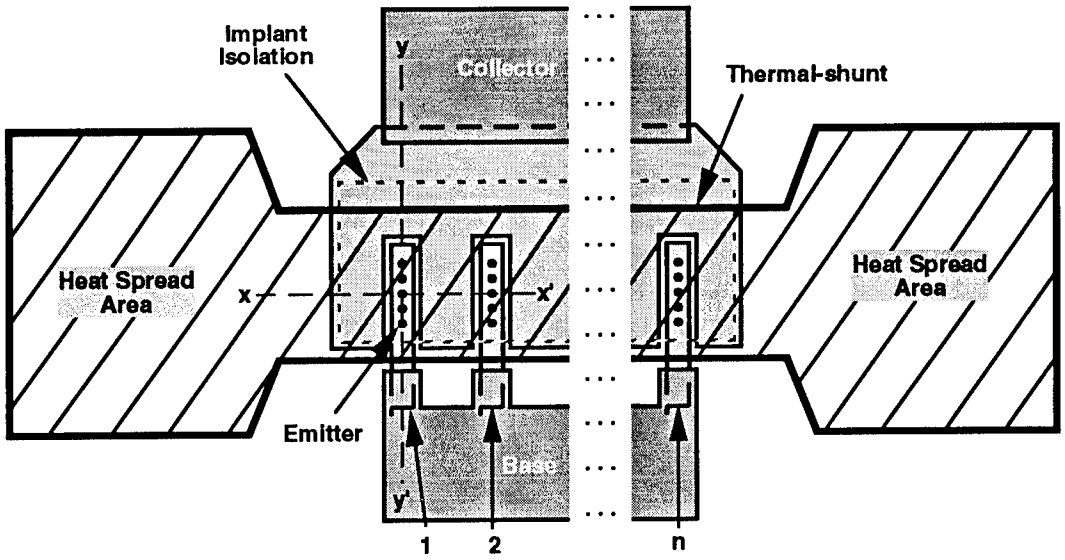


Figure 2a. Layout of the HBT having n base fingers.

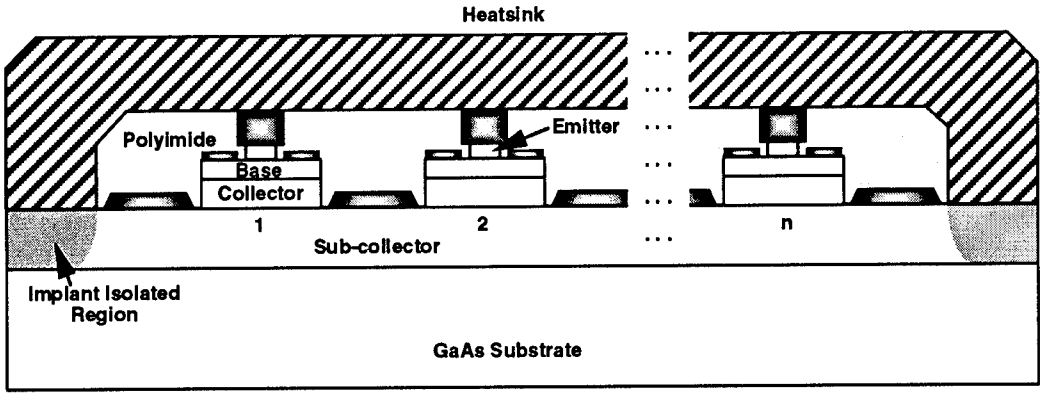


Figure 2b. xx' cross-sectional view of n base firgers of the device.

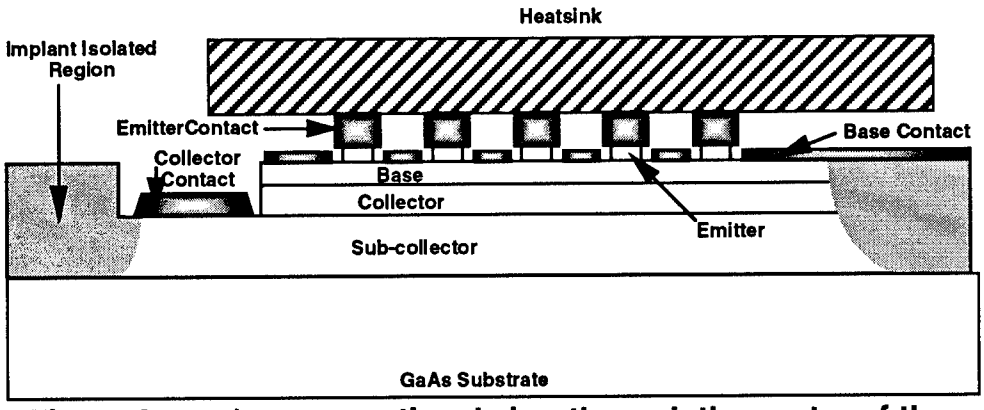


Figure 2c. yy' cross-sectional view through the center of the center of the base-collector mesa and emitters.

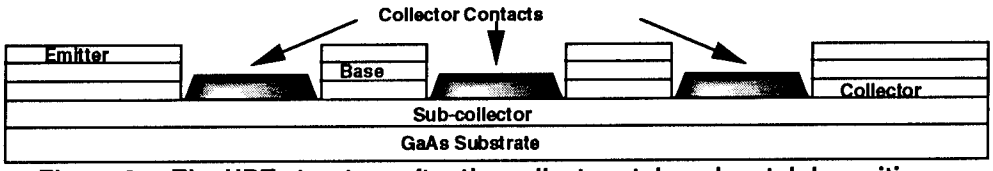


Figure 3a. The HBT structure after the collector etch and metal deposition.

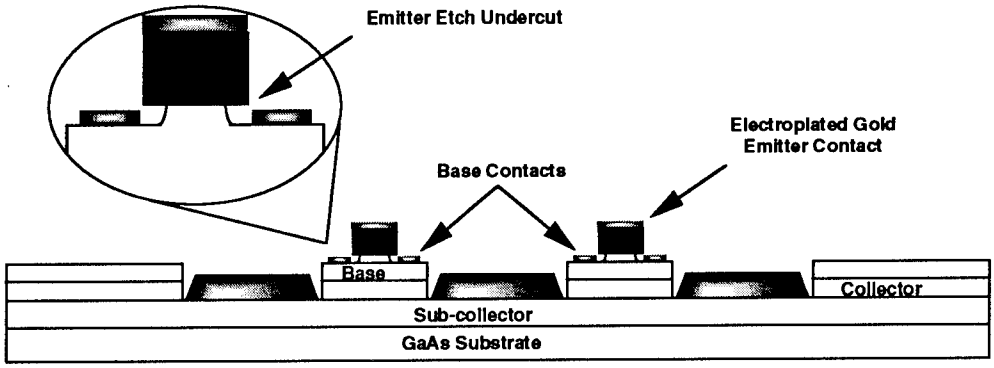


Figure 3b. Emitter and base metal contacts and associated undercut of the emitter layer.

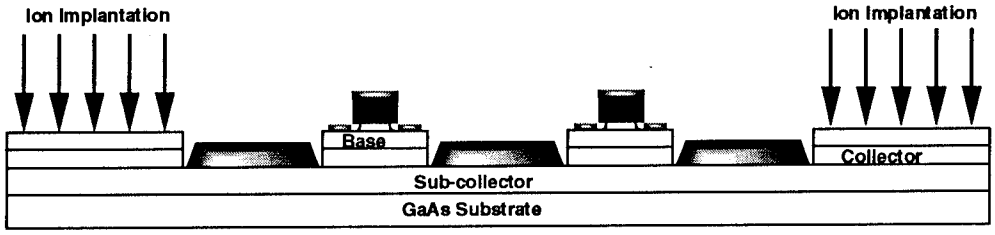


Figure 3c. Ion implantation of inactive areas and base trim etch.

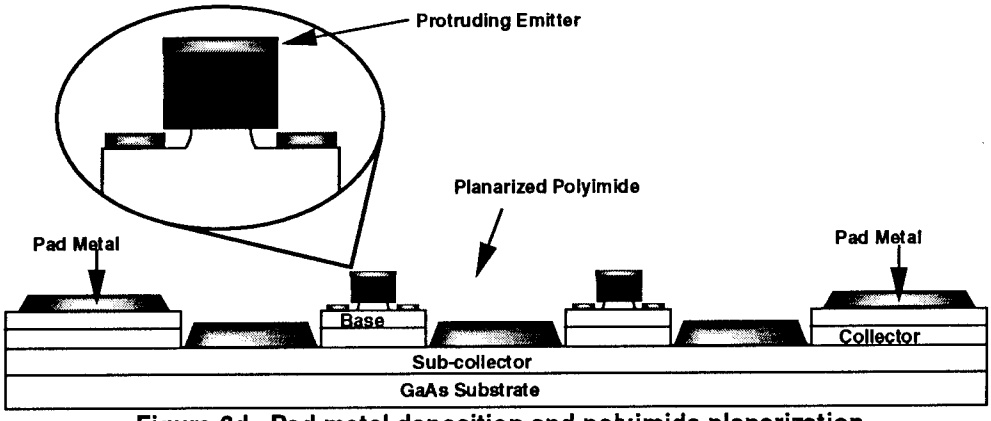


Figure 3d. Pad metal deposition and polyimide planarization.

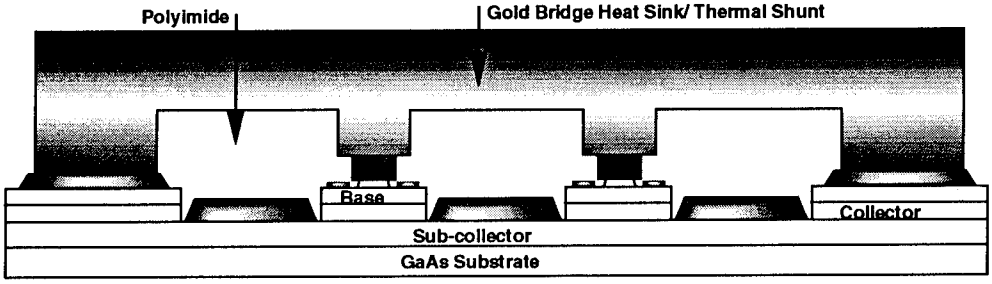


Figure 3e. Gold heat sink securely contacting the emitters and supported by polyimide.

Cross-sectional views of thermally-stabilized HBTs are shown in Figures 2b and 2c.

Device fabrication began with collector lithography and wet chemical etch to the sub-collector layer. Standard ohmic metal Ni/Ge/Au/Ni/Au was deposited and alloyed in a rapid thermal annealer at 420°C for 15 seconds. The resulting structure is shown in Figure 3a. The collector contact is the only alloyed contact in this design. To avoid any possible degradation of the exposed emitter-base junction or emitter and base contacts during high temperature alloying, the collector contact was fabricated first. The wafer was not subjected to temperatures higher than 250°C for the rest of processing.

The gold emitter contacts were electroplated to a height of 1 μm and centered on the base-collector mesas. InGaAs provided an excellent seed layer for electroplating. The emitter material surrounding the emitter contacts was etched using citric acid/hydrogen peroxide based chemical etching. A citric:peroxide ratio was chosen to minimize selectively between InGaAs and GaAs and to optimally undercut the emitter metal [10]. TiPtAu base contacts, self-aligned to the emitter metal, were then deposited as shown in Figure 3b. The active device areas were isolated using oxygen implantation of the material between devices as well as a trim etch to remove the top 1500Å of the implanted surface. A thin layer of Si_3N_4 was deposited across the wafer and then etched using a CF_4/O_2 plasma to leave a thin Si_3N_4 passivation layer on the sidewalls on the emitter/base periphery. Ti/Au pad metal was then deposited on the inactive surface to provide electrically probable contacts to the base and collector metal. The resulting device structure is shown in Figure 3c.

The final processing steps were planarization and heat sink formation. The wafer was planarized by applying a 3 μm layer of DuPont polyimide (PI-2555) to the wafer. PI was etched in the Reactive Ion Etcher (RIE) using an oxygen plasma to expose the top surface of the emitter contacts shown in Figure 3d. An evaporated Ti/Au seed layer provided continuity for electroplating posts and the gold heat sink bridge to the emitter contacts. The posts were 3 μm tall and the heat sink bridge was electroplated to a thickness of 22 μm. The final HBT structure is shown in Figure 3e. An SEM picture of a completed 250 μm^2 emitter area device is shown in Figure 4.

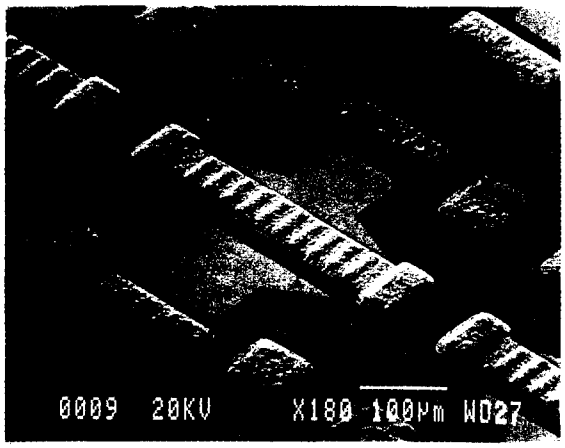


Figure 4. SEM photo of a 250 μm^2 power device.

III. THERMAL ANALYSIS

Thermal limitations of an HBT can be reached due to self-heating under high bias conditions which may: (a) cause long-term reliability problems (mostly at metal contacts), (b) degrade its electronic properties, or (c) cause thermal runaway. The first two limitations require temperature rise of at least 300°C to be significant, whereas the third limitation can occur at much lower temperatures.

The device layout and size are important factors in determining which thermal limitation is the dominant mechanism. The bias of single-emitter, small-area devices can be increased until a critical junction temperature is reached, beyond which the device reliability is unacceptable. Thermal limitations of multi-emitter microwave power devices, however, is determined by the onset of the thermal runaway condition [6]. This is illustrated in Figure 5, where electrical and

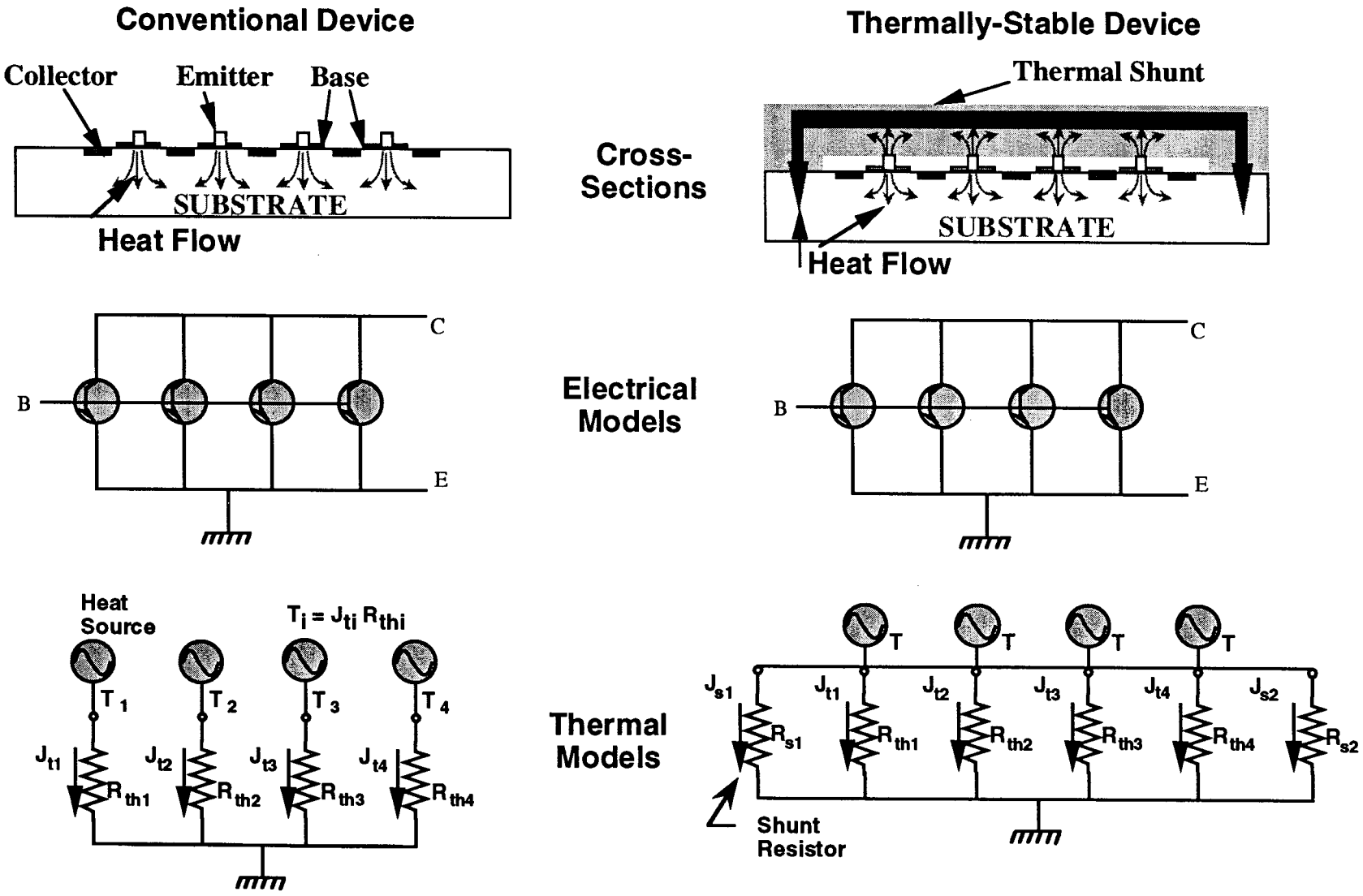


Figure 5. Cross-sections, electrical and thermal models of conventional and thermally-stable HBTs.

thermal models of a conventional multi-emitter HBT are shown.

In the electrical model, each cell corresponds to a single emitter finger, base finger and collector of a multi-finger device. Each set of terminals of the HBT (emitters, bases and collectors) are electrically connected. Due to the negative temperature dependence of $V_{be(on)}$, the electrical and thermal characteristics are codependent.

The thermal model of the conventional device in Figure 5 shows that the excess heat from each part of the device is dissipated through thermal resistors $R_{th1}...R_{th4}$. Since very little cross-coupling exists between these resistors, the temperature of each element is independently determined by the local bias conditions. Because all terminals of the device are electrically connected in parallel, the base current distribution to each element is a function of the local temperature. Negative temperature coefficient of base-emitter turn-on voltage, $V_{be(on)}$, can cause a local increase in base current (and therefore collector current) in the hotter parts of the device. The increased bias current in these regions can, in turn, increase the local temperature even further. This positive feedback mechanism can result in thermal runaway in bipolar transistors. However, because of the negative temperature dependence of current gain, thermal runaway in GaAs HBTs does not always result in device burn out. Instead, a "current-crush" is observed, where the collector current suddenly drops [6]. A device operating in a "current-crush" mode develops local hot spots within the device affecting device reliability. Further, the dc and microwave performance of such a device is severely degraded [6].

As an alternative to ballast resistors, thermal shunt resistors can be used between emitters of a power transistor, as shown in Figure 5. A thermal shunt between emitters is fabricated using thick metal air bridges, such that the thermal resistance of the shunt resistors is very small compared with the heat sink resistance. The electrical model of a thermally-stable transistor is identical to that of conventional devices. Thermal characteristics are, however, significantly different, as shown in Figure 5. Because of the presence of the thermal shunt between emitter elements, the device temperature is the same at all points. This in turn dictates a uniform base current distribution.

Thermal stabilization of the power device removes the "current-crush" related thermal limitation. The CW performance of such a device is similar to that found with single emitter and small area devices operating at the same bias conditions. Low thermal conductivity of GaAs, however, severely limits the heat removal rate from the active device area. The overall thermal resistance of the device now becomes the thermal limitation. To lower the thermal resistance of our devices, we have employed heat spreaders at the ends of thermal shunt bridges (see Figure 2a). In this design, the majority of the excess heat is removed through the emitter metal, thermal shunt bridge, and thermal resistors R_{s1} and R_{s2}. Because the heat sink area is considerably larger than the emitter area, and the heat path from the device to the heat sinks is all metal, this technique produces an apparent increase in heat sink area of the device active area. Such a magnification of the heat sinking area is termed "thermal lens" technique.

IV. RESULTS AND DISCUSSION

Typical Gummel plots and common-emitter I-V characteristics of a 150 μm^2 emitter area device with 3 μm diameter emitter discs are shown in Figures 6 and 7, respectively. The current gain value obtained with small emitter area devices (microwave devices) was almost the same as that obtained with large-area devices (DC devices). This insensitivity of current gain to emitter periphery/area ratio is indicative of reduced recombination current density at the perimeter of the emitter mesa and close to the surface of the base layer. The temperature coefficient of emitter-base (V_{be}) turn-on voltage was measured to be between -1.95 and -2.05 mV/°C, nearly

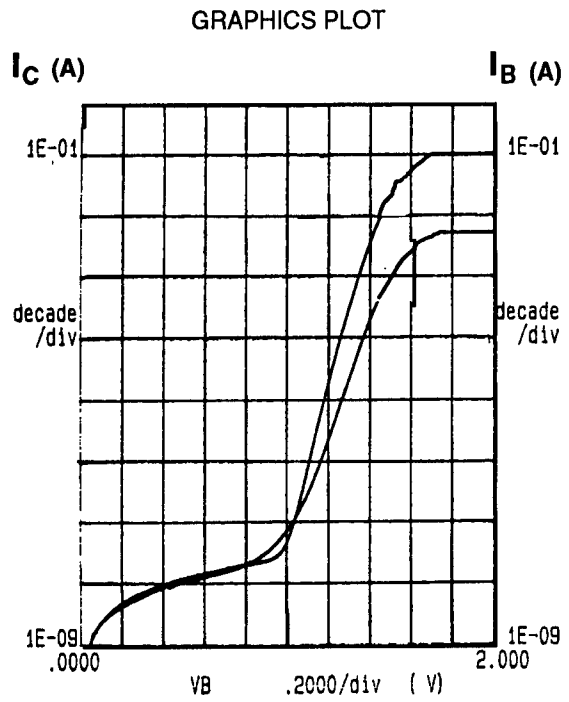


Figure 6. Gummel plot of a 60 μm² device.

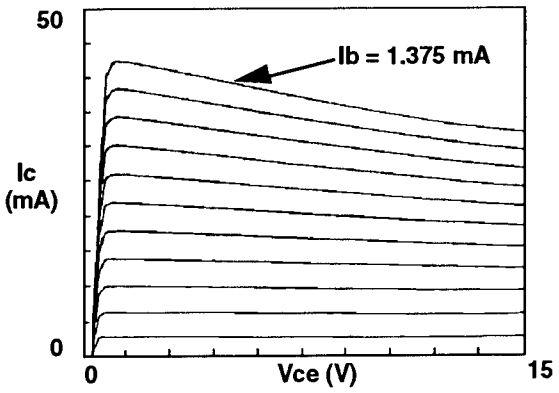


Figure 7. DC characteristics of the thermally-stabilized HBT.

independent of device size and base current. Temperature dependent variation of V_{be} was used to estimate the thermal resistance [11].

A total thermal resistance of a 60 μm² emitter area device with 3 μm emitter discs mounted in test fixtures with silver epoxy was measured as 350°C/W. The thermal resistance of the same device without the thermal bridge was estimated to be 2.5 to 3.0 times higher than this value. Emitter resistances of various size devices were measured using the "open-collector" method on a curve tracer. It was found that the emitter resistance (contact resistance plus bulk emitter resistance) was 2.6 x 10^{-6} Ω-cm² of emitter area and scaled well with emitter area. This value of emitter resistance is too small to function as a ballast resistor for HBTs [8].

Small-signal measurements were made using an HP8510C automatic network analyzer. On-wafer tests of the devices in the frequency range from 500 MHz to 62.5 GHz showed the current gain cutoff frequency, f_T to be 35 GHz. The maximum frequency of oscillation, f_{max} was 50 GHz. These measurements were determined from the computed $|h_{21}|^2$ and unilateral power gain curves. The devices were biased at V_{ce} = 3 V and I_c = 2 x 10^4 A/cm².

Large-signal measurements at 3, 7.5 and 10 GHz were performed on carrier-mounted devices of various sizes. Off-chip coaxial mechanical tuners were used to match the input and output impedances of the devices. The results shown in Table 1 were obtained with devices operating under common-emitter mode with V_{CE} = 13V. It is seen that the power density is the highest for smaller area devices due to more effective heat sinking. The output power density of devices was in the range of 9-10 mW/μm² of emitter area for CW output powers up to 600mW at 10 GHz. Similar power densities were obtained at 3 and 7.5 GHz, but with higher power gain. A maximum power density of 10 mW/μm² was obtained at 10 GHz with 0.6 W CW output power and 60% power-added efficiency (PAE). The same device, when tuned for best efficiency, produced 67.2% PAE (81.4% collector efficiency) with 0.56 W CW output power and 7.6 dB gain. The power density under this condition was 9.3 mW/μm². The collector current density and the estimated junction temperature at this operating condition were 9.5 x 10^4 A/cm² and 49°C, respectively. The current density

was limited by the onset of the Kirk effect. The junction temperature rise was low because of effective heat spreading and high collector efficiency.

Table 1. Microwave performance of discrete HBTs.

Emitter Feature Size (μm)	Total Emitter Area (μm^2)	Freq. (GHz)	Gain (dB)	Power Output (W)	Power Density ($mW/\mu m^2$)	PAE (%)
2	30	3.0	12.2	0.27	9.0	60.6
2	75	7.5	7.6	0.63	8.4	48.8
2	250	7.5	4.3	1.76	7.0	36.0
2	250	7.5	7.4	1.50	6.0	52.0
3	600	7.5	3.4	2.52	4.2	31.8
3	600	7.5	5.4	2.50	4.2	40.1
3	60	10.0	7.6	0.56	9.3	67.2
3	60	10.0	7.1	0.60	10.0	60.0

The highest output power achieved was 2.5 W at 7.5 GHz with a 600 μm^2 device. The power density of this device was thermally limited to 4.2 $mW/\mu m^2$ because of lower collector efficiency, which is mostly a result of the limited tuning range of the microwave test system used. However, the power density is still considerably higher than any previous HBTs with similar output levels.

IV. CONCLUSIONS

Design and fabrication techniques were developed for AlGaAs/GaAs microwave HBTs to eliminate thermal instabilities associated with high power operation. The use of a thermal shunt across emitter contacts of a multi-emitter HBT was shown to eliminate "current-crush" effects. A significant reduction was achieved in the thermal resistance of devices by spreading the heat over a larger area using thermal lens techniques. With these design features applied together, we have demonstrated thermally-stable operation of AlGaAs/GaAs HBTs up to their electronic limitations. A record power density of 10 $mW/\mu m^2$ of emitter area was achieved with 0.6 W CW output power and 60% power-added efficiency at 10 GHz. The thermal stabilization technique developed is applicable to all other bipolar transistors.

ACKNOWLEDGMENTS

The authors would like to thank P. Cassity, C. Isbill and L. Callahan for technical assistance, M. Mack, K. Nakano, and G.L. McCoy for discussions.

REFERENCES

[1] N.L. Wang, N.H. Sheng, M.F. Chang, W.J. Ho, G.J. Sullivan, E.A. Sovero, J.A. Higgins, and P.M. Asbeck, "Ultrahigh power efficiency of common-emitter and common-base HBTs at 10 GHz," *IEEE Trans. Microwave Theory and Tech.*, vol. 38, pp. 1381-1389, 1990.

[2] P. Bartusiak, T. Henderson, T. Kim, A. Khatibzadeh, and B. Bayraktaroglu, "High-efficiency Ku-band HBT amplifier with 1 W CW output power," *Electronics Lett.*, vol. 27, pp. 2189-2190, 1991.

[3] J.A. Higgins, "GaAs heterojunction bipolar transistors: A second generation microwave power amplifier transistor," *Microwave J.*, pp. 176-194, 1991.

[4] M.A. Khatibzadeh, B. Bayraktaroglu, and T. Kim, "12 W monolithic X-band HBT power amplifier," *IEEE MTT-S Microwave and Millimeter Wave Monolithic Circuits Symp. Techn. Dig.*, pp. 47-50, 1992.

[5] P. Bartusiak, T. Henderson, T. Kim, and B. Bayraktaroglu, "High-efficiency Ku-band HBT MMIC power amplifier," *IEEE Electron Dev. Lett.*, vol. 13, pp 584-586, 1992.

[6] L.L. Liou, B. Bayraktaroglu, and C.I. Huang, "Thermal stability analysis of multiple finger microwave AlGaAs/GaAs heterojunction bipolar transistors," *IEEE MTT-S Microwave Symp. Tech. Dig.*, pp. 281-284, 1993.

[7] M. Adlerstein, M. Zaitlin, G. Flynn, W. Hoke, J. Huang, G. Jackson, P. Lemonias, R. Majorone, and E. Tong, "High power density pulsed X-band heterojunction bipolar transistors," *Electron. Lett.*, vol. 27, pp. 148-149, 1991.

[8] G.B. Gao, M.S. Unlu, H. Morkoc, and D.L. Blackburn, "Emitter ballasting resistor design for, and current handling capability of AlGaAs/GaAs power heterojunction bipolar transistors," *IEEE Trans. Electron Dev.*, vol. 38, pp. 185-196, 1991.

[9] B. Bayraktaroglu, J. Barrette, R. Fitch, L. Kehias, C.I. Huang, R. Neidhard, and R. Scherer, "Thermally-stable AlGaAs Microwave Power HBTs," presented at *1993 Device Research Conference*, Santa Barbara, CA.

[10] G.C. DeSalvo, W.F. Tseng, and J. Comas, "Etch rates and selectivities of citric acid/hydrogen peroxide on GaAs, $Al_{0.3}Ga_{0.7}As$, $In_{0.2}Ga_{0.8}As$, $In_{0.53}Ga_{0.47}As$, $In_{0.52}Al_{0.47}As$ and InP," *J. Electrochem. Soc.*, vol. 139, pp. 831-835, 1992.

[11] D.S. Whitefield, C.J. Wei, and J.C.M. Hwang, "Temperature dependent large-signal model of heterojunction bipolar transistors," *IEEE GaAs IC Symp. Tech. Dig.*, pp. 221-224, 1992.

NEW HIGH FREQUENCY AlGaAs/GaAs PNP - HBT

D.B. Slater, Jr., P.M. Enquist, J.A. Hutchby, Research Triangle Institute,
A. Morris and R.J. Trew, North Carolina State

PAPER UNAVAILABLE FOR PUBLICATION

0-7803-0894-8/93 $3.00

FABRICATION AND CHARACTERIZATION OF PLANAR INTEGRATED SCHOTTKY DEVICES FOR VERY HIGH FREQUENCY MIXERS

I. Mehdi, P. Siegel, and M. Mazed
Jet Propulsion Laboratory
California Institute of Technology, Pasadena CA 91109

ABSTRACT

Many millimeter-wave mixers and frequency multipliers today still employ a whisker contacted Schottky diode as the nonlinear device. In order to reduce the risk and assembly cost associated with these critical receiver components for NASA's present and future space missions, the authors have developed a novel fabrication procedure that integrates a planar Schottky diode with the mixer circuitry thus greatly simplifying the assembly and testing of the diode circuits. The process and DC results obtained so far will be discussed along with some preliminary RF results at 200 GHz.

INTRODUCTION

Present generation NASA space-borne instruments employing millimeter-wave radiometers, such as the Microwave Limb Sounder(MLS) operating on the Upper Atmosphere Research Space Satellite(UARS), use whisker contacted Schottky diodes in both the mixer and frequency multiplier circuitry. However, the delicate nature of the whisker contact decreases the reliability of the instrument. Recently, it has been shown that, at least up to 350 GHz planar integrated diodes perform as well as whisker contacted diodes [1,2,3] and it is believed that the next generation of space-borne instruments will utilize this advancement in technology. Though mounting of the planar diode chip in the mixer circuit is not as labor intensive as the contacting of a whisker diode , it is still not a trivial task. The planar diode chips are typically 150x50 microns and require some expertise in order to place them correctly in a given circuit.

In order to alleviate this problem we have suggested integrating the Schottky device with additional mixer filter circuitry which makes it easier to handle as the device is now a part of a much larger structure. Unfortunately, GaAs does not lend itself as readily to typical waveguide mixer circuitry and generally quartz is preferred. In this paper we discuss the procedure that has been developed to combine the planar GaAs Schottky diode with the mixer quartz filter circuitry thus resulting in a

0-7803-0894-8/93 $3.00

mechanically robust, easy to handle structure. Initial DC and RF results will also be discussed.

PROCESS DEVELOPMENT

At higher frequencies it is believed that losses due to the GaAs substrate can degrade performance. The process that was designed to fabricate these Schottky diodes removes all of the GaAs on the chip except under the Schottky and ohmic contacts. It has been demonstrated at the University of Virginia that thinned substrate devices can be transplanted on to quartz using UV cured optical cement [4]. However, that exact process can not be replicated in our case since the heavily doped layer of GaAs extending all the way in the waveguide channel will be highly detrimental to the circuit's RF performance. In our method, once the devices and the associate circuitry is fabricated on the GaAs,the wafer is mounted upside down on the quartz plate with a UV curable optical cement. Once the GaAs substrate has been completely removed the mixer filter structures, which include the planar Schottky diode can be diced and tested. Though we believe that this process is portable to other planar devices thus far we have only investigated it for the fabrication of the 220,440, and 640 GHz subharmonically pumped waveguide mixers. This mixer uses a split block waveguide circuit with a single quartz insert. The circuitry that is integrated with the planar diode consists of two in-line hammerhead filters, signal and LO coupling probes, and a DC return for diode biasing and monitoring. Details of the block and the filter structures are presented in [2]. The salient steps of the fabrication process are detailed below.

1. 200-400 nm of Silicon-dioxide is deposited via an Electron Cyclotron Resonance(ECR) machine, Fig. 1(a).

2. Holes are opened lithographically in the oxide and the epilayer is etched until contact to the n+ layer is made. The same mask is used to etch the oxide and the GaAs and then lift-off the ohmic contact metal, Fig. 1(b).

3. The Schottky metal(Ti/Pt/Au) is evaporated, Fig. 1(c). Optical lithography is used for anode diameters of 1.0 micron or more. At 600 GHz, submicron anodes are required and direct-write E-beam will be used.

4. The filter pattern is deposited as Cr/Au including the anode contact fingers on the planar Schottky diodes which are in an antiparallel configuration, Fig.1(d). The metal is electroplated up with gold to at least three skin depths at the desired operating frequency.

5. The devices are electrically isolated by etching a channel under the anode fingers through to the GaAs buffer layer. The etchant used is sensitive to the crystal direction and undercuts the fingers without etching the GaAs surrounding the Schottky anode, Fig. 1(e). This technology was developed at the University of Virginia and is discussed in [4].

6. The processed GaAs wafer is bonded device side down on the quartz plate which at 200 GHz is 150 micron thick. The wafer is held via a commercially available UV cured optical adhesive , Fig. 1(f).

7. The semiconductor substrate is removed by first lapping it down mechanically to about 50-100 micron and then etching it in a selective etch which stops when the AlGaAs layer is exposed, Fig.1(g).

8. The remaining GaAs on the backside of the device is completely removed, except in the area around the Schottky diode. This is accomplished using reactive ion etching. It is important to use a dry etch at this point in order to avoid severe undercut and a drastic reduction in the yield, Fig. 1(h).

Once the GaAs from the back is completely removed the filter structures can be diced and placed in the waveguide block for testing. Fig. 2 shows the completed device both (a) the metal side, and (b) the quartz side. Fig. 3 shows a SEM of a portion of the filter used at 200 GHz.

RESULTS

The diode structure that has been designed and fabricated has a 100 nm top GaAs epilayer doped at 4.0E17 cm^{-3} followed by a 3.0 micron thick heavily doped (2E18 cm^{-3}) GaAs layer. The buffer layer is undoped GaAs and 700 nm thick. Finally the etch stop layer is AlGaAs(Al=55%) and is 400 nm thick. The substrate is semi-insulating. The diodes produced from this layer structure have typically 1 micron diameter anodes with an ideality factor of 1.4, saturation current of 5.E-14 A, series resistance of 13-16 ohms and a barrier height of approximately 1.05 eV.

With discrete GaAs devices of similar layer structures we have obtained single sideband noise temperatures of less than 2000 K in our waveguide mixer block [2]. However, for the devices already integrated with the filters the best results to date have been a noise temperature of 10,000 K. Moreover, the LO power required to drive the unbiased diode pair is much larger (6-8 times) than the power required for the discrete devices. We believe the poor performance of the integrated devices is

caused by two as yet unresolved fabrication problems. First, the integrated structures have exceptionally high pad-to-pad capacitance (larger than 30 fF). This accounts for the poor LO coupling and much of the high noise temperature. Second, the diode ideality factor and saturation current are higher than expected(the better discrete diodes have an ideality factor of 1.2 and saturation current of less than 1.E-15 A). Efforts are currently underway to correct these shortcomings.

CONCLUSIONS

A fabrication method has been discussed and demonstrated that integrates a planar air-bridge-type Schottky-diode pair with associated quartz mixer circuitry. This procedure will allow diode circuitry to be used at frequencies too high to allow a separate diode chip to be used. The Schottky diodes obtained via this process have so far had poor DC characteristics and large parasitic capacitance which have limited the RF performance. Efforts are underway to resolve these problems so better RF performance can be obtained.

REFERENCES

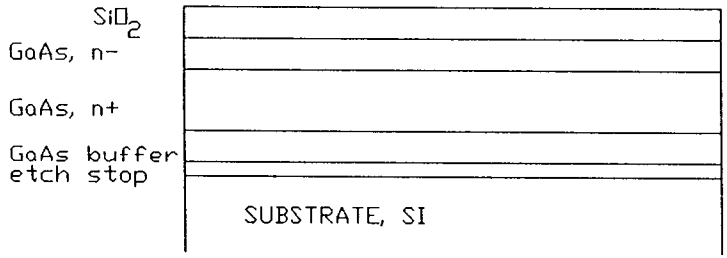
1. D. Garfield, R. Mattauch, and S. Weinreb, "RF Performance of a Novel Planar Millimeter-Wave Diode Incorporating an Etched Surface Channel," IEEE-MTT vol. 39, no. 1, pp.1-5, January 1991.

2. P. Siegel, R. Dengler, I. Mehdi, J. Oswald, W. Bishop, T. Crowe, and R. Mattauch, "Measurements on a 215 GHz Subharmonically Pumped Waveguide Mixer using Planar back-to-back Air Bridge Schottky Diodes," to be published in IEEE-MTT.

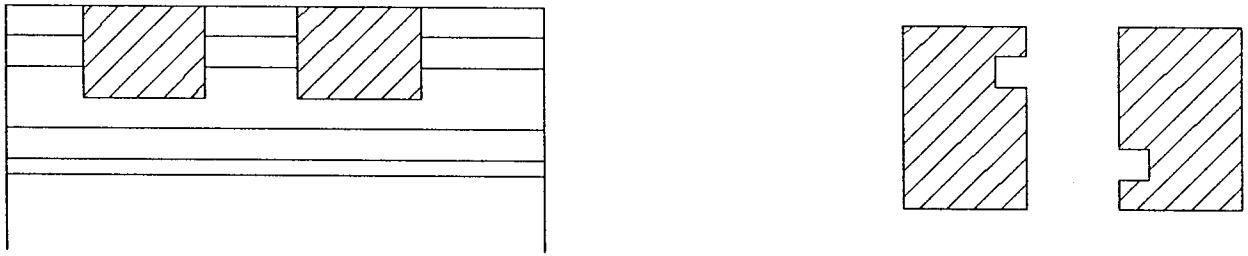
3. T. Newman, W. L. Bishop, K. T. Ng, S. Weinreb, "A Novel Planar Diode Mixer for Submillimeter-Wave Applications," IEEE-MTT vol. 39, no. 12, pp. 1964-1971, December 1991.

4. W. Bishop, K. McKinney, R. Mattauch, T. Crowe, and G. Green, "A Novel Whiskerless Diode for Millimeter and Submillimeter Wave Applications," 1987 IEEE-Int. Microwave Symposium Digest, pp. 607-610, June 1987.

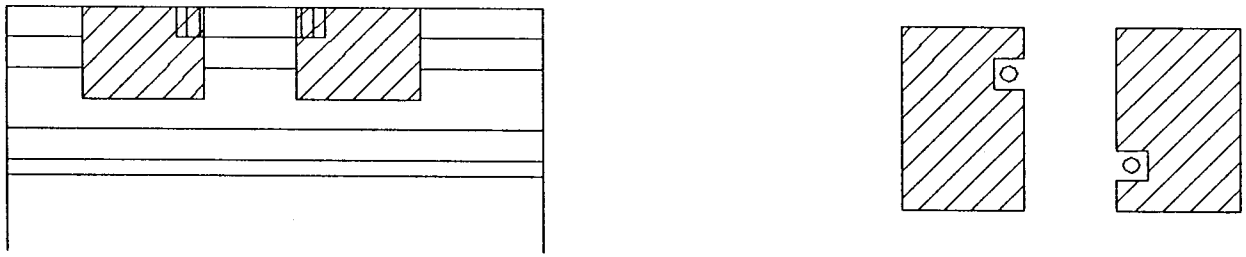
Fig. 1. Process flowchart



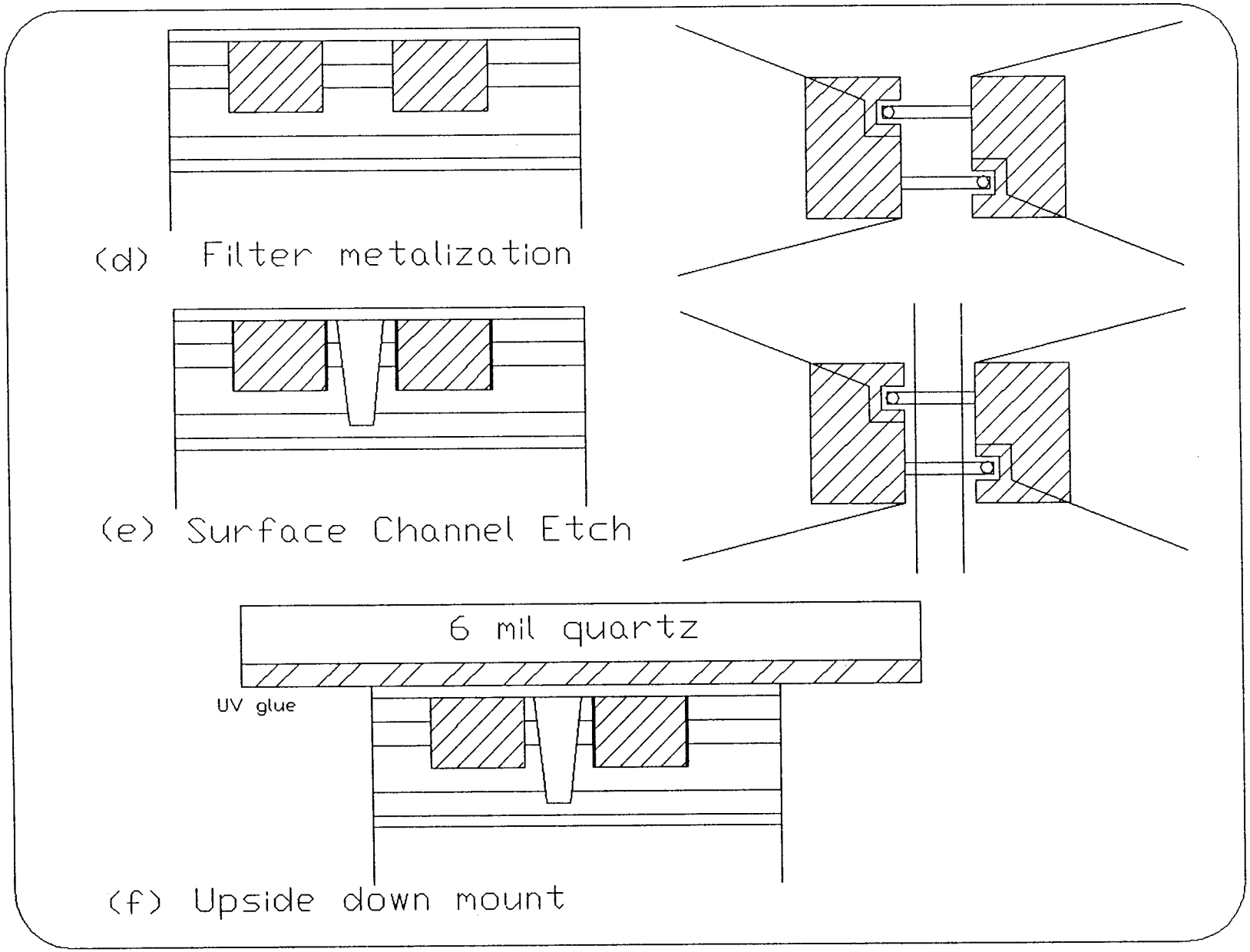
(a) Oxide deposition

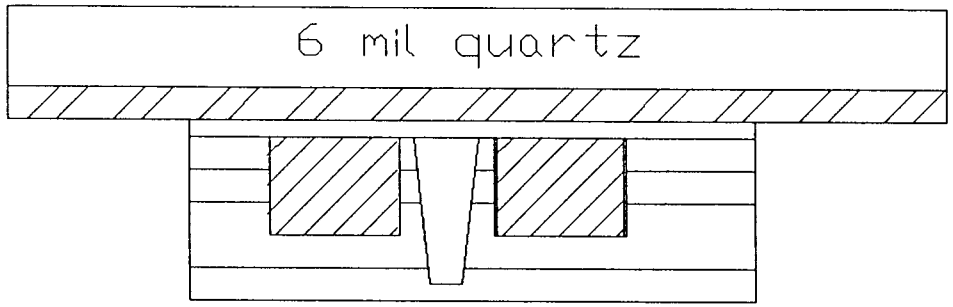


(b) Ohmic metalization

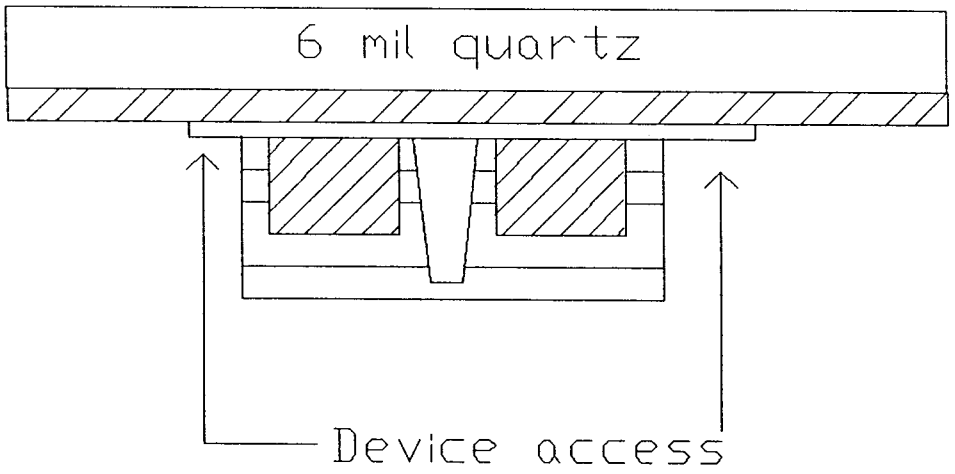


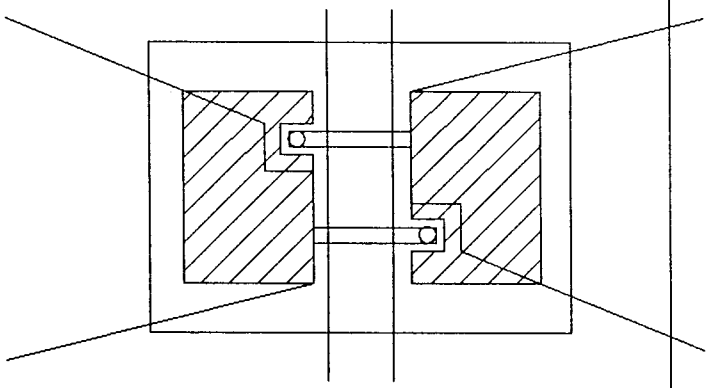
(c) Schottky metalization


(d) Filter metalization
(e) Surface Channel Etch
6 mil quartz
UV glue
(f) Upside down mount

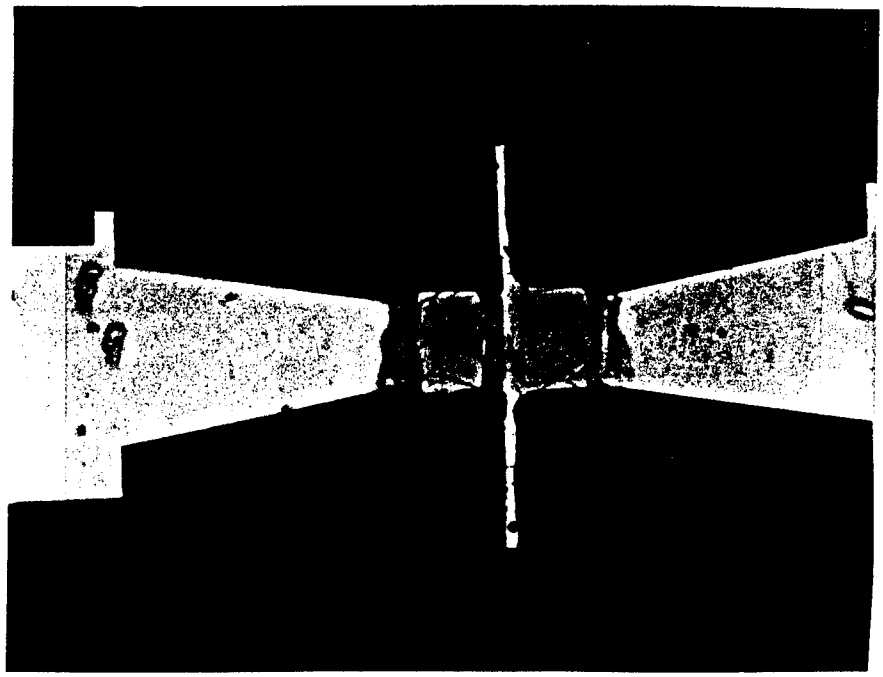


(g) Substrate removal

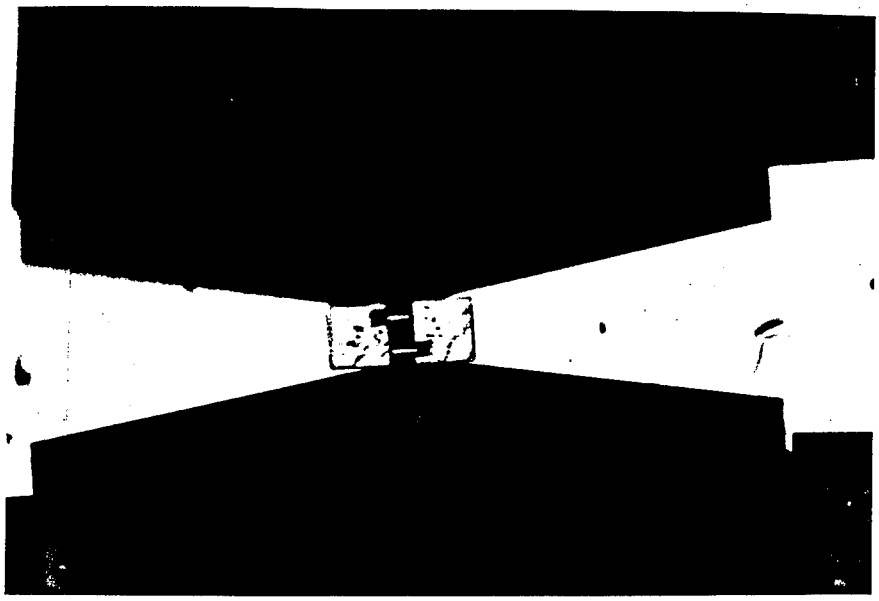




(h) Device isolation and etch



(a)



(b)

Fig 2: A back-to-back diode pair integrated with a filter structure, (a) the front side of the structure showing the GaAs pads over the Schottky diodes, (b) the bottom side (quartz) of the structure showing the fingers contacting the anodes.

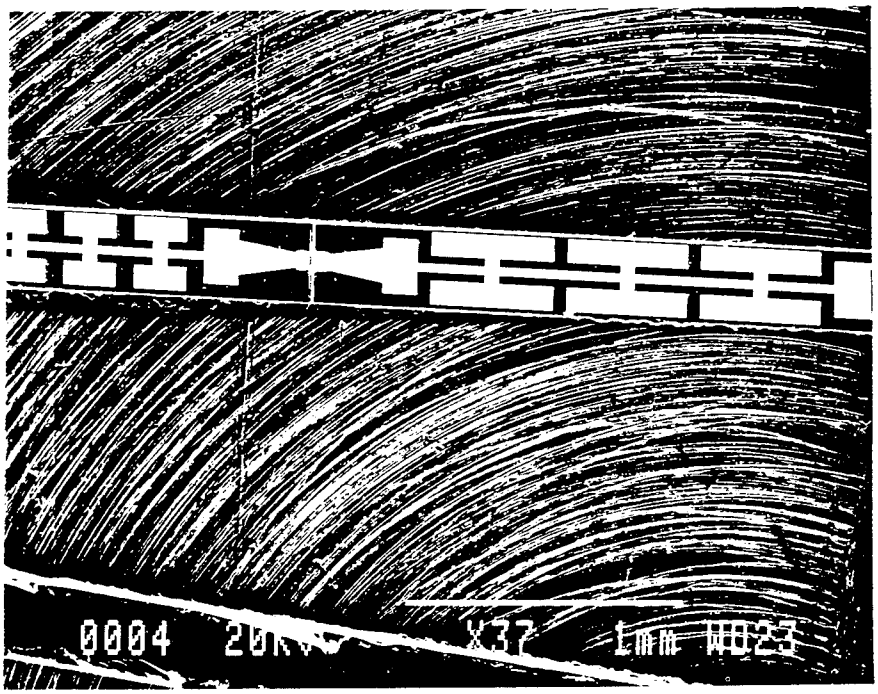


Fig 3. A portion of the mixer hammerhead filter and the planar back-to-back Schottky diode medullization side facing up. The structure is ready to be mounted in a waveguide block.

LT-GaAs-MIS-Diode Characteristics and Equivalent Circuit Model

K.-M. Lipka, B. Splingart, U. Erben, E. Kohn

University of Ulm
Department of Electron Devices and Circuits
Oberer Eselsberg, D-89069 Ulm, Germany
Tel.: +49 731 502-6151, Fax: +49 731 502-6155

Abstract

LT-GaAs MISFETs had been realized indicating a record 2.7W/mm RF power handling capability. To optimize such LT-GaAs power MISFET structures, the MIS system containing a LT-GaAs insulator and an AlAs interfacial diffusion barrier to the channel has been analysed. A noticeable parallel conductance was found in the insulator which is thought to be one of the key parameters to realize high gate to drain breakdown voltages. This conductivity however leads also to a g_m-dispersion in the MHz range. Locus-curves of this system demonstrate a higher resistivity in the AlAs layer than in the LT-GaAs layer, indicating that the simple model of a single lossy capacitance does not describe the MIS diode completely. An extended electronic equivalent circuit for use in the FET model has been established.

1. Introduction

Low Temperature-GaAs layers, grown at extremly low temperatures, containing a large amount of excess As, have been successfully used as gate insulator or surface passivation in GaAs FET structures. Both, high gate-drain and high drain-source breakdown voltages had been shown [1], leading to a record power capability of 1.57W/mm for GaAs based FETs at 1.1GHz [2]. Additionally very high drain to source breakdown voltages had been shown at full current [3].

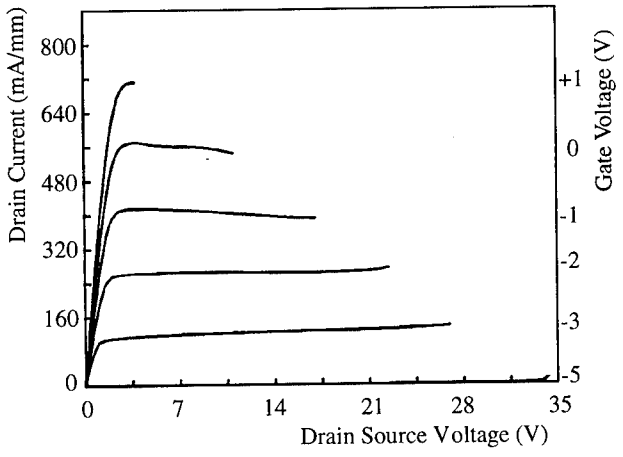


Figure 1: DC-output characteristics of record LT-GaAs power MISFET structure

I_{dmax} =720mA/mm
U_{dsmax}=34V

Our recent results shown in figure 1 indicate 2.7W/mm RF power are feasible. Gate-Drain breakdown voltages above 50V are observed independent of the doping concentration in the channel as shown in figure 2. A detailed overview of the material properties and its applications is found in [4,5]. This indicates that the limitation of the output power capability predicted by the lateral spreading model in MESFETs [6] is not applicable in such devices and therefore further improvements seem to be possible.
Using a LT-GaAs insulating layer in GaAs MISFETs an interfacial diffusion barrier has to be introduced to avoid As-outdiffusion from the LT-GaAs insulator into the channel [7]. The commonly used AlAs diffusion barrier is part of the insulating layer system. Nevertheless, many details concerning the MIS structure, composed of Ti/Pt/Au, LT-GaAs-insulator, AlAs-barrier and the n-doped channel, are still controversy. Therefore important device related parameters have to be investigated in order to obtain information on the mode of operation in FET-structures. This study focuses on the electrical properties of the insulator and an equivalent circuit describing the LT-GaAs/AlAs insulating layer system.

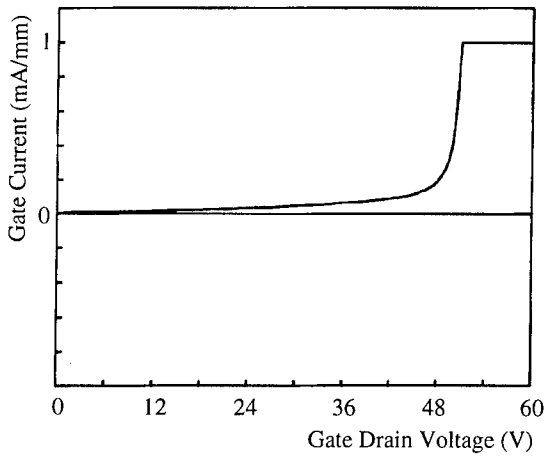


Figure 2: Gate-Drain-breakdown characteristics of MISFET structure
Current limitor is set at 1mA/mm

MISFETs are of special interest for power applications due to suppressed gate leakage current and therefore high gate breakdown voltage independent of the channel doping. However GaAs MISFETs can only operate in the depletion mode due to the insulator interface potential pinning. Therefore, typical problems of these devices are a low transconductance and instabilities due to charging of the interface states, which are insulated from the gate electrode by the dielectric insulator. In the case of the LT-GaAs insulator system a noticeable leakage current is observed, comparable to the Deep Depletion MISFET-case using a silicon rich silicon nitride insulating layer [8]. This controled leakage is also thought to be one of the key reasons for the large gate-drain breakdown voltage shown in figure 2.

2. Device Structure and Fabrication

The investigated MIS structures contain a silicon doped GaAs-channel and the insulator, consisting of a 20nm AlAs diffusion barrier and a 80nm LT-GaAs top layer. The 250nm thick channel was $1x10^{17}cm^{-3}$ silicon doped.

All structures had been grown in a Riber 32P MBE system on (100) semi-insulating LEC GaAs substrates. The 600nm GaAs-buffer layer and channel layers have been grown under standard MBE conditions. After growing the AlAs layer at 680°C the growth was interrupted and the substrat cooled down to 200°C, the growth temperature of the LT-GaAs, determined by thermocouple reading.
Structures had been insulated by wet chemical mesa etching. Ohmic contacts were made by evaporating and alloying Au/Ge/Ni after removing the LT-GaAs and AlAs layers by selective wet chemical etching and lift-off. Finally Ti/Pt/Au Schottky gate contacts have been deposited by e-beam evaporation. The size of the Schottky contact in this FATFET-like structure was 20µmx100µm.

3 Models

At RF the LT-GaAs/AlAs MIS diode, representing a bias dependent network of RC-components, can be characterized by its impedance Z(ω) or admittance Y(ω) in the complex plane, the so-called locus-curves. The MIS structure, being part of a FET, is preferably presented as an impedance Z(ω), since the structure is a layered structure of the LT-GaAs, the diffusion barrier and the GaAs space charge layer (SCL). It seems therefore reasonable to expect a series connection of parallel RC elements. Due to a highly doped channel and forward biased gate further capacitive elements are negligible, however a channel series resistance is included.

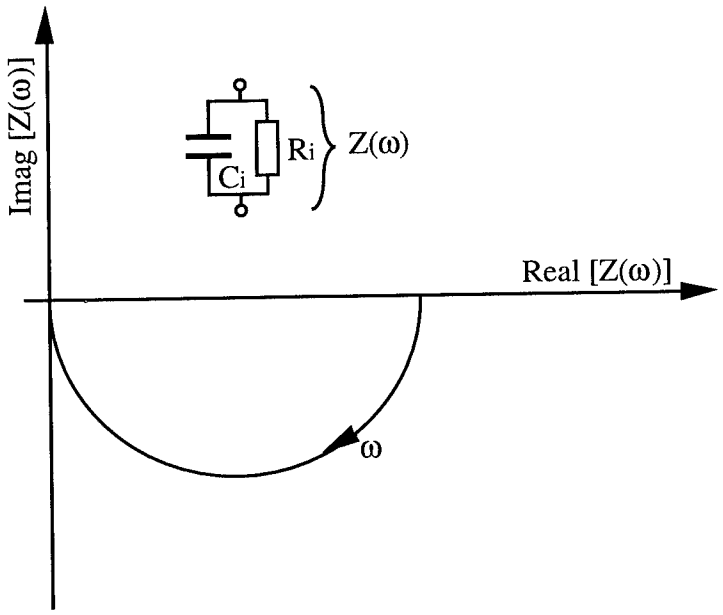


Figure 3: Equivalent circuit of the one layer model including a conduction path, and its locus-curve

Three different cases are in principle imaginable:

(1) One layer model
The resistivity of the LT-GaAs and the AlAs layers are similar. Therefore the entire structure of both layers acts as one composed layer. The equivalent circuit consists of a capacitor C_i in parallel to a resistance R_i representing a leakage current (see figure 3). In case of an ideal insulator the resistance becomes infinite. In principal this circuit results in a locus-curve, calculated by equation (1) describing a semicircle in the complex plane. The radius of this semicircle is $1/2R_i$ and the

effective MIS capacitance, part of the FET input, is calculated by equation (2) resulting in an frequency dependent $g_m(\omega)$ in the FET structure.

$$Z(\omega) = \frac{R_i}{1+\omega^2 R_i^2 C_i^2} - j\frac{\omega R_i^2 C_i}{1+\omega^2 R_i^2 C_i^2} \tag{1}$$

$$C_{eff}(\omega) = C_i + \frac{1}{\omega^2 R_i^2 C_i} \tag{2}$$

$$C_i = \varepsilon_i \frac{A}{d_i} \tag{3}$$

(2) Two layer model

If the resistivity of both layers is essentially different, this will result in an equivalent circuit of two parallel RC elements in series. The locus-curve is composed of two semicircles, shown in figure 4 as a principle diagram. The insulator part of the g_m-controling capacitance in the MISFET is then calculated by equation (4). The minimum value of this capacitance $C_{eff}(\omega)$ is determined for high frequencies by the capacitance C_i of the entire insulator. At low frequencies the resistors bypass the capacitors and only the space charge layer capacitance in the GaAs is seen.

$$C_{eff}(\omega) = C_i + \frac{C_{LT\text{-}GaAs}\left(1+\omega^2 R_{LT\text{-}GaAs}^2 C_{LT\text{-}GaAs}^2\right) + C_{AlAs}\left(1+\omega^2 R_{AlAs}^2 C_{AlAs}^2\right)}{\omega^2\left(C_{LT\text{-}GaAs}+C_{AlAs}\right)^2\left(\omega^2 R_{LT\text{-}GaAs}^2 R_{AlAs}^2 C_{LT\text{-}GaAs} C_{AlAs} + \frac{R_{LT\text{-}GaAs}^2 C_{LT\text{-}GaAs} + R_{AlAs}^2 C_{AlAs}}{C_{LT\text{-}GaAs}+C_{AlAs}}\right)} \tag{4}$$

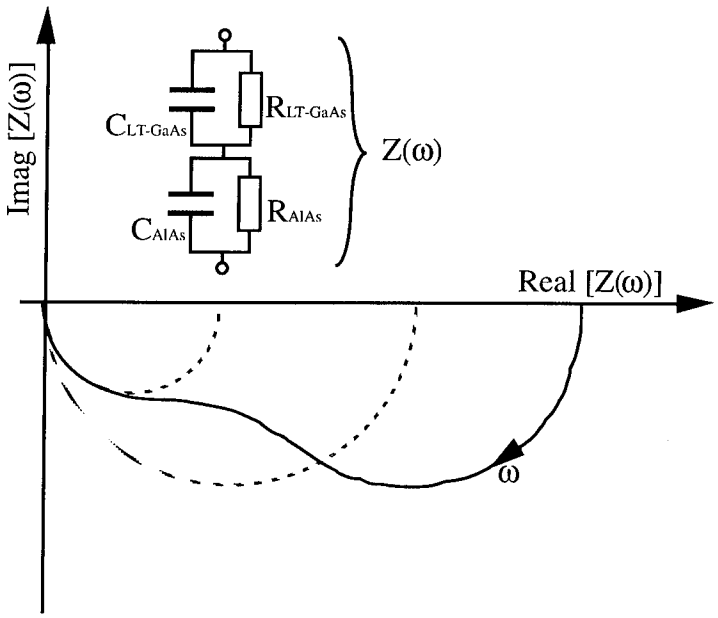


Figure 4: Equivalent circuit and locus-curve of a two layer MIS model

(3) Distributed layer model
The resistivity of the insulating layer changes gradually across the structure, due to a diffusion of excess As. Therefore the discrete RC-elements become dispersed resulting in a stretched semicircle-like locus-curve. To model this case a dispersion function needs to be assest. This has not been attempted in this study. However this effect may attribute to the differences of the two layer model with the measured data.

Measuring the MIS structure at forward bias (as discussed above) the calculation and modelling of the insulator is simplified, due to the absence of the GaAs-SCL capacitor. However a channel series resistance and parasistic elements mainly caused by the coupling between the measuring pads have to be considered at high frequencies.

4. Measurements and Modelling

The locus-curve of the investigated LT-GaAs/AlAs MIS structure is shown in figure 5 measured at frequencies between 45MHz and 10 GHz. The measurement has been performed at 1V forward bias in order to avoid a SCL below the insulator. Equivalent circuit elements have been calculated or measured as listed below:

Element	Determination	Value
$R_{channel}$	measured between two ohmic contacts in close proximity the MIS structure	18.5Ω
$C_{LT\text{-}GaAs}$	calculated using equation (3)	3pF
C_{AlAs}	calculated using equation (3)	12pF
$C_{insulator}$	calculated by equation (3)	2.4pF
$R_{parasitic}$	estimated	1Ω
$R_{LT\text{-}GaAs}+R_{AlAs}+R_{channel}$	extracted from DC measurements	73Ω

In all cases the dielectric constant of the As rich insulator was approximated by the ε_r of GaAs. The parasitic capacitor had been fitted at high frequencies. The rather high value of Cp=165fF is due to the large size of the pads used in the structure. All elements of the one layer MIS model are predetermined and the measured locus-curve can be compared to the calculated one. The result, shown in figure 5, indicates clearly that this model does not describe the LT-GaAs/AlAs MIS system sufficiently, since no acceptable fit is possible despite parasitic corrections.

In case of the double layer MIS structure the resistances of the layers were used as fitting parameters. Since the sum of the resistors $R_{LT\text{-}GaAs}+R_{AlAs}=54.5\Omega$ had been determined by DC measurements, only the ratio of these resistors can be varied. The best fit was obtained by $R_{LT\text{-}GaAs}=19\Omega$ for the LT-GaAs layer and $R_{AlAs}=35.5\Omega$ for the AlAs layer as shown in figure 7. Figures 9 and 10 illustrate finally the accuracy of this fit. The remaining discrepancy between measurement and calculation is believed to be due to the distributed nature of the structure modeled by lumped elements.
The essential higher resistance of the AlAs layer agrees to the fact that the upper limit of the effective MIS capacitance is given by the capacitance of the AlAs layer for low frequencies [9].

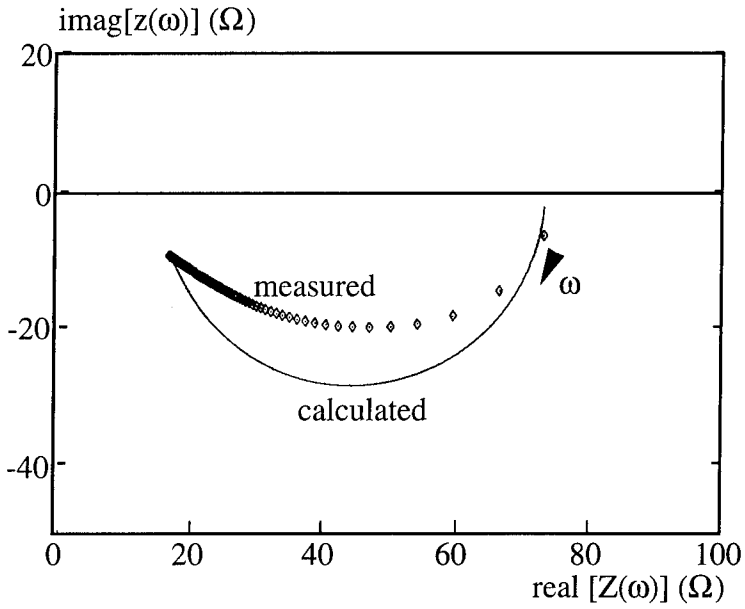


Figure 5: Locus-curve of the MIS structure compared to the one layer model

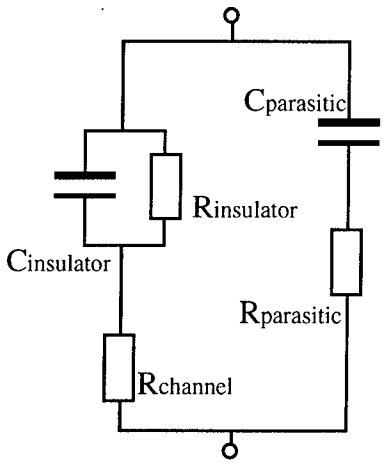


Figure 6: Equivalent circuit of the one layer model

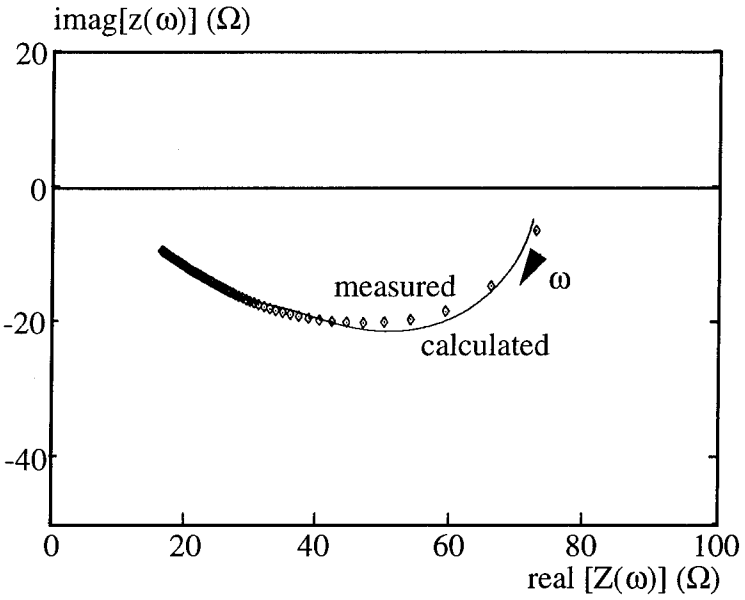


Figure 7: Locus-curve of the MIS structure compared to the two layer model

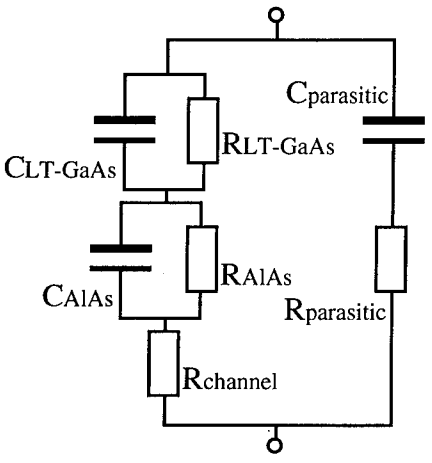


Figure 8: Equivalent circuit of the two layer model

The two layer model for the LT-GaAs/AlAs MIS system can now describe the transition from low frequency to high frequency behavior and can be used to estimate the g_m-dispersion of MISFET structures. In the bias range applied to the FET the input capacitance C_{gs} is then composed of the effective insulator capacitance and the capacitance of the GaAs-SCL. Compared to normal GaAs-MISFET structures no interface state capacitance is expected, because the dc-leakage current and high recombination rate (fs electron and hole capture time constant) in the LT-GaAs system prevent insulated or loosely coupled interface states to appear. Figure 9 and 10 indicate a substantial dispersion between 0.1 and 10GHz. However in the FET the GaAs SCL-capacitance will reduce the effect substantially. For a small dispersion the depth of the GaAs SCL needs to be large compared with the insulator thickness.
However the absolut values of the resistors included in the modeled case depend on the specific LT-GaAs properties, that means MBE growth conditions as well as post growth thermal treatments. The presented data may therefore only apply for the MIS structure in question, but the analysing technique will apply in general.

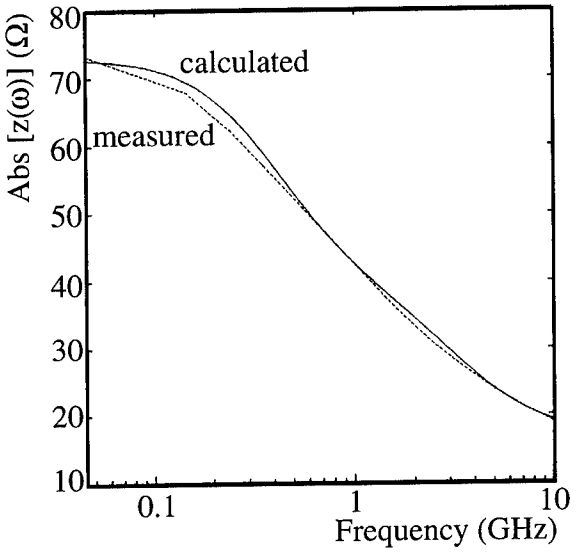


Figure 9: Absolut value of measured impedance compared to calculated values for the two layer model

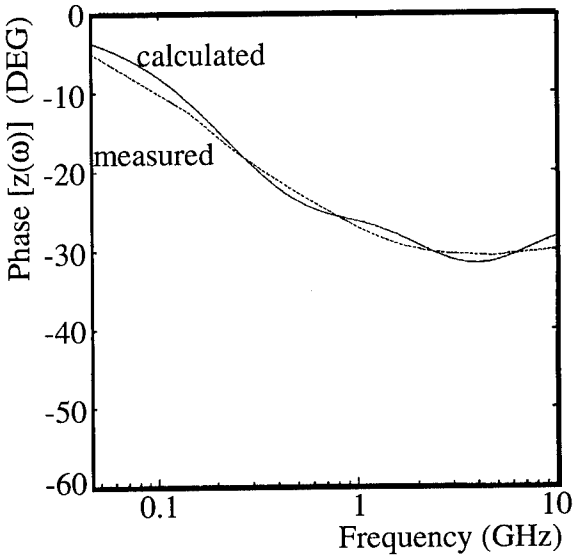


Figure 10: Phase of measured impedance compared to calculated values for the two layer model

5. Discussion and Conclusions

LT-GaAs MISFETs are attractive for power applications, because of their high demonstrated RF-output power of 1.5W/mm and an even higher IV-power product. Devices with $N_D=8x10^{17}cm^{-3}$ of 62.5nm channel thickness have displayed a maximum current densisty of 720mA/mm and a maximum drain source voltage of 34V (see fig.1). This represents a record IV-power product of 23.9W/mm. In class A operation this will translate to $P_{RFmax}=I_{max}(V_{max}-V_{sat})/8=2.7W/mm$. Thus, high DC-breakdown voltages can be obtained and combined with high output currents, indicating a power capability of these devices superior to those predicted by the lateral spreading

model. Additionally the lower interface potential compared to other MISFET structures increases the available current [10]. However, it has been shown that the MIS-gate configuration introduces necessarily a frequency dispersion, which in turn will also translate into a time and temperature instability. To fully explore and utilize the potential of these structures, these instabilities have to be minimized by tailoring the structure. The upper limit of the effective insulator capacitance of the MIS structure by the capacitance of the AlAs layer for low frequencies requires an optimized diffusion barrier. Reducing the thickness of the diffusion barrier is therefore of special interest in order to minimize the dispersion and increase the g_m in MISFETs. First experiments in this direction are encouriging. First structures with a 5nm modified diffusion barrier have been processed resulting in a maximum low frequency g_m of 180mS/mm for 1μm gate length. However, the frequency dispersion in the MIS-system means that a larger RF-drive voltage is needed than indicated by the quasi static output characteristics.

Acknowledgement

The authors would like to thank Y. Druelle and G. Salmer from IEMN Lille (France) for many helpful discussions and making their MBE equipment available for the growth of most of the samples. Thanks also to K. Panzlaff for growing some additional MBE structures. This work was supported by the CEC ESPRIT project 6849 "TAMPFETS".

Literature

[1] L.-W. Yin, Y. Hwang, J. H. Lee, R. M. Kolbas, R. J. Trew, U. K. Mishra
Electron Device Letters, vol 11, No 12, 561 (1990)
[2] C.-L. Chen, F. W. Smith, B. J. Clifton, L. J. Mahoney, M. J. Manfra, A. R. Calawa
Electron Device Letters, vol 12, No 6, 306 (1991)
[3] K. Lipka, B. Splingart, E. Kohn, Electronics Letters, vol 29, No13, 1170 (1993)
[4] F. W. Smith, Mat.Res. Soc. Symp. Proc. Vol 241, 3 (1992)
[5] U. K. Mishra, Applications of GaAs grown at a low temperature by MBE,
Proc. E-MRS Spring Meeting Strasbourg, paper B-15 (1993)
[6] S. H. Wemple, W. C. Niehaus, H. M. Cox, J. V. Dilorenzo, W. O. Schlosser
Transactions on Electron devices, Vol. ED 27 No. 6, 1013 (1980)
[7] L.-W. Yin, University of California, Santa Barbara, ECE Technical Report # 92-23 (1992)
[8] J. M. Dortu, E. Kohn, Inst. Phys. Conf. Ser. No 74: Chapter 7, 563 (1984)
[9] K.-M. Lipka, B. Splingart, U. Erben, E. Kohn, LT-GaAs MISFET structure for power application, Proc. 20th International Symposium on GaAs and Related Compounds,
Freiburg, Germany, TuP4 (1993)
[10] K. M. Lipka, B. Splingart, X. Zhang, M. Poese, K. Panzlaff, E. Kohn, Interfacial Barrier Characteristics for LT-GaAs on low doped GaAs Layers, Proc. E-MRS Spring Meeting,
Strasbourg, France, paper B-10 (1993)

On the suppression of phonon-electron scattering in short periodic AlAs/GaAs multiple quantum well structures.

T.T.LeTran, W.J.Schaff, B.K.Ridley*, Y.P.Chen, A.Clark, S.O'Keefe and L.F.Eastman.

School of Electrical Engineering, Cornell University.
Ithaca, NY-14853-5401
* *Department of Physics, University of Essex.*
Colchester, CO4-3SQ, England

Abstract.

The suppression of longitudinal optical phonon (LOP)-electron scattering was sought in multiple quantum well (MQW) structures. The structures had GaAs well widths = 12, 15 and 20 mono layers (ML) and AlAs barrier widths = 2 and 4 ML. The MQWs were grown in the channel of GaAs/$Al_{0.3}Ga_{0.7}As$ modulation doped field effect transistors (MODFETs) without gates. The Hall mobility and carrier sheet density were measured by the van der Pauw method. The Hall mobility of the MQW samples was found to be less than that of the control samples (without MQW) at room temperature, but was better at temperatures lower than 50 K. The reduction of the room temperature mobility was due to interaction of the well electrons with the interface polaritons from the AlAs barriers. The increase of the low temperature mobility was due to reduced remote ionized impurity scattering of the well electrons. The evidence of performance improvement of MQW devices at room temperature due to suppression of electron-LOP scattering is thus disputed by this study. The results of an experiment made elsewhere, which appeared to show the contrary, can be interpreted using arguments other than the suppression of electron-LOP scattering in MQWs.

0-7803-0894-8/93 $3.00

Introduction

In recent years, the possibility of electron mobility enhancement by suppression of phonon-electron interaction in multi quantum well (MQW) structures has raised considerable interest [1-5]. The suppression is sought by the confinement of longitudinal optical phonons (LOP) inside the quantum wells that have different coupling constant and ion masses than the barriers. The interaction of 2D-LOP / 2D-electrons gets weaker as the well width decreases, in contrast to the 3D-LOP / 2D-electrons system [2]. On the other hand, interface polariton (IP) interaction with free carriers becomes more important as the well width decreases in systems of finite barriers [3-5].

A study of several different GaAs/AlAs multiquantum well samples is reported here. It indicates that there may be an upper limit of electron mobility parallel to the AlAs/GaAs MQW plane due to the interplay between LOP and IP scattering of electrons. This study confirms the importance of IP scattering in very narrow wells (12 to 20 atomic layers). The latter is shown to outweigh the benefit of suppression of LOP scattering.

Samples

Two control samples and MQW structures of different well and barrier widths were examined. Fig. 1 shows the first control sample's (REF) structure. The 200Å n-GaAs cap layer, the 800Å $Al_{0.3}Ga_{0.7}As$ uniform donor layer, the 150Å spacer layers and the GaAs channel were grown by MBE on (100) GaAs at 620°C. The donor concentration was 2×10^{18} cm^{-3}. The second control sample, REF+4, is similar to the REF, except that there is a 4 ML thick AlAs-barrier between the spacer and the channel. All parameters in the other samples were kept the same as the REF's, except that they have AlAs/GaAs multiple quantum well in their channel. Six such samples were studied: MQW(12/2), MQW(15/2), MQW(20/2), MQW(12/4), MQW(15/4), MQW(20/4), where the numbers in brackets represent the well and barrier widths in number of ML,

respectively. Fig. 2 shows the band structure of the MQW(15/4) sample. Note the large density of the low mobility electrons in the cap and donor layers.

The barriers in the MQW samples were kept thin to minimize the possibility of carriers being scattered into the X-valleys of AlAs, while still being large enough to confine LOP to within the wells. For example, the energy gap between the X-level of the AlAs barrier and the first Γ-level of the GaAs well is 0.38 eV in the MQW(15/2) sample. The fraction of the well's electrons being scattered into the barriers is approximately 10^{-7}. In the MQW(15/4), the fraction is just 1%. The upper limit of the well width is roughly 30 ML, above which the confinement of LOP may not be appreciable [2]. If the wells were narrower than 10 ML, the electronic subband levels would rise above the barriers causing the carriers to be deconfined by the MQW, as shown by computer simulations that solve the Schrödinger and Poisson equations self-consistently.

Data and discussions

The samples' Hall mobility and carrier sheet density were measured by the van der Pawv method. The magnetic field was 2 kiloGauss. The samples were exposed only to diffused ambient light shortly before the measurements. To minimize the contribution of the parallel conduction in the donor layer to the total mobility, the samples were etched with the solution H_3PO_4:H_2O_2:H_2O=3:1:50 until their individual Hall mobility reached a maximum value.

The mobilities after etching and the carrier sheet densities are shown in Table 1. The REF's mobility was 7590 $cm^2V^{-1}s^{-1}$ at room temperature, comparable to the value of REF+4, taking into account its lower carrier sheet density. The MQW samples, however, all showed lower room temperature mobilities than the control samples. Among them, the MQW(15/x) samples had the highest mobilities. At T=77K, the mobility of the MQW samples approached that of the control samples, except for the MQW(20/x) samples.

Table 1: The Hall mobility and carrier sheet density of the samples at room and liquid Nitrogen temperature. The control sample REF has the basic MODFET structure; REF+4 has an additional 4 ML thick AlAs between the AlGaAs spacer and the GaAs channel. The samples MQW(x/y) have x ML of GaAs as well and y ML of AlAs as barriers.

	T=300K		T=77K	
	μ [$cm^2V^{-1}s^1$]	N_S [10^{11} cm^{-2}]	μ [10^3 $cm^2V^{-1}s^1$]	N_S [10^{11} cm^{-2}]
REF	7590	5.44	189	5.52
REF+4	8200	4.12	193	4.15
MQW(12/2)	5287	2.94	191	1.80
MQW(12/4)	5189	2.00	180	1.48
MQW(15/2)	6318	3.16	198	2.20
MQW(15/4)	6682	2.49	181	2.66
MQW(20/2)	6024	3.72	124	2.48
MQW(20/4)	5055	3.31	153	1.90

The true mobility of the GaAs channel (and its MQW) is given by the Petriz model [6]:

$$\mu_{channel} = \frac{1}{2}\mu_{AlGaAs}\left(1+\sqrt{1-4\mu N\frac{(\mu_{AlGaAs}-\mu)}{\mu^2_{AlGaAs}\,N_{channel}}}\right)$$

where μ and N are the effective mobility and carrier sheet density at T=300 K, respectively; $N_{channel}$ is the sheet density measured at $T \leq 50K$; $\mu_{channel}$ and $N_{channel}$ are the true mobility and sheet density at T = 300 K, respectively. μ_{AlGaAs} is the mobility of the free electrons inside the donor layer at 300K; its value is conventionally taken to be 500 to 1000 $cm^2V^{-1}s^{-1}$. Thus $\mu_{channel} \approx 9600$ $cm^2V^{-1}s^{-1}$ and 8700 $cm^2V^{-1}s^{-1}$ for the REF and MQW(15/x) samples, respectively.

The lower mobility of the MQW samples is readily understood by the hybrid model of phonons in superlattice structures. This model is presented in a paper submitted elsewhere [7]. In case of the AlAs/GaAs MQW structure, the LOP-like hybrid modes of the AlAs barriers are not coupled to the interface polaritons (IP). The latter thus interact strongly with the electrons in the well region. Although the interaction of electrons and LOP is reduced in the MQW samples, the increasing strength of the electrons-barrier IP scattering enhances the overall scattering rate as the well width decreases, and hence reduces the mobility. The channel mobility calculated by this model was 6120 $cm^2V^{-1}s^{-1}$ for an infinite MQW with 15 ML thick GaAs wells and 2 ML thick AlAs barriers, compared to the actual value of 8700 $cm^2V^{-1}s^{-1}$ of the (short periodic) MQW(15/2) sample.

The dependence of the Hall mobility and carrier sheet density on temperature was also investigated. To minimize the persistent photo conductivity effect, the samples were kept in the dark for at least 10 hours before being measured in a continuous flow cryostat fed by liquid helium. The results for the REF, MQW(15/2) and MQW (15/4) are shown in Fig. 3. One observes that the mobility of the MQW samples raised above that of the REF at $T \leq 50$ K. The MQW(15/2)'s mobility was 1.4×10^6 $cm^2V^{-1}s^{-1}$, twice as high as the REF's (6.9×10^5 $cm^2V^{-1}s^{-1}$) at T=11 K. The

sheet density of the MQW samples were considerably lower than that of the REF because they were etched, as mentioned above.

The improved mobility of the MQW samples indicates that the electron-IP interaction is no longer the dominant scattering mechanism at low temperatures. Our low temperature simulations showed a smaller field across the spacer layer in the MQW samples than in the REF, which reduced the effect of remote ionized impurity scattering. More important, the improved mobility shows that the interface roughness scattering is not a factor in lowering the MQW sample's mobility at room temperature.

Conclusions

We have tested several short periodic MQW structures for the enhancement of their room temperature mobility due to the suppression of electron-LOP scattering. The results were interpreted as negative. The best room temperature mobility was achieved by the MQW with 15 ML thick GaAs wells and 2 or 4 ML thick barriers. It was roughly 10% lower than that of the samples without the MQW. This result was readily explained by the strong interaction of the well electrons with the IP-like hybrid modes from the AlAs barriers. At low temperatures, the MQW samples' mobility was higher than that of the control samples, due to reduced remote ionized impurity scattering.

To the best of our knowledge, the enhancement of room temperature mobility by suppression of electron-LOP scattering has yet to be experimentally proven. The results of Zhu et.al.[8] that appeared to confirm the opposite may actually be explained using some alternative arguments. The energy band structure and the bulk carrier density of the samples are shown in Fig. 4. The MQW's wall and barrier thickness are 8 ML and 4ML, respectively. These narrow wells raise the electron subbands over the barriers, spilling most of the electrons onto the AlAs/GaAs-substrate interface, which is both smooth and far away from the ionized donors. The major contribution to the total mobility is from the 2DEG at this interface, not from the electrons in the MQW structure. In case of the control sample, the electrons stay inside the GaAs well and are subjected to

more interface roughness and ionized impurity scattering, as compared to the majority of the electrons in the MQW sample; hence the mobility is considerably lower. Please note that both of these samples have lower room temperature mobility than bulk GaAs.

Acknowledgments

The authors wish to thank C.Lehman for his assistance in fixing the variable temperature Hall measurement system. The project was funded by the Army Research Office, the GEO-centers and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR) of Québec.

References

1. F.A.Riddoch and B.K.Ridley, J. Phys. C: Solid State Phys. **16**, 6971 (1983).
2. N.Sawaki, J.Phys.C:Solid State Phys.**19**, 4965 (1986).
3. K.W.Kim, A.R.Bhatt, M.A.Stroscio, P.J.Turley and S.W.Teitsworth, Appl.Phys.Lett. **72**, 2282 (1991).
4. B.K.Ridley, Phys.Rev.B **47**, 4592 (1993).
5. M.A.Stroscio, GJIatrate, KWKim, MALittleJohn, HGoronkin, G.N.Maracas, Appl.Phys.Lett. **59**, 1093 (1991).
6. R.L.Petritz, Phys.Rev.**110**, 1254 (1958).
7. T.T.LeTran, W.J.Schaff, B.K.Ridley, Y.P.Chen, A.Clark, S.O'Keefe and L.F.Eastman, submitted to J.Appl.Phys.
8. X.T.Zhu, H.Goronkin,G.N.Maracas, R.Droopad, M.A.Stroscio, Appl.Phys.Lett **60**, 2141 (1992).

Figure captions:

1. The structure of the control sample REF. It consists of a 200Å n-GaAs cap layer, a 800Å $Al_{0.3}Ga_{0.7}As$ uniform donor layer (2 x 10^{18} cm^{-3} n-doped), a 150Å spacing layers and the GaAs channel. The REF+4 has an additional 4 ML of AlAs between the spacer and the channel. The energy band E_c is shown by the solid line; the bulk density of free electrons is shown by the dotted line.

2. The structure of a typical MQW sample, the MQW(15/4). It consists of the REF plus a short periodic MQW structure in its channel. The energy band E_c is shown by the solid line; the bulk density of free electrons is shown by the dotted line.

3. The mobility μ and carrier sheet density N_s of the REF, MQW(15/4) and MQW(15/2) as functions of temperature. μ of REF is higher than the other two at room temperature, but decreases relatively at $T \leq 50$ K. The sheet densities of the MQW samples are lower than that of REF because the latter were etched.

4. A different design of MQW structures, as of Zhu et.al. a) shows the control sample with its GaAs well at the center, where the free electrons are concentrated. The sharp gap on the right is the AlAs/GaAs-substrate interface. b) shows the MQW structure. The MQW with well width = 8 ML and barrier width = 4 ML is grown onto the GaAs. Since the wells are small, the free electrons are spilled into the smooth AlAs/GaAs-substrate interface.

Figures

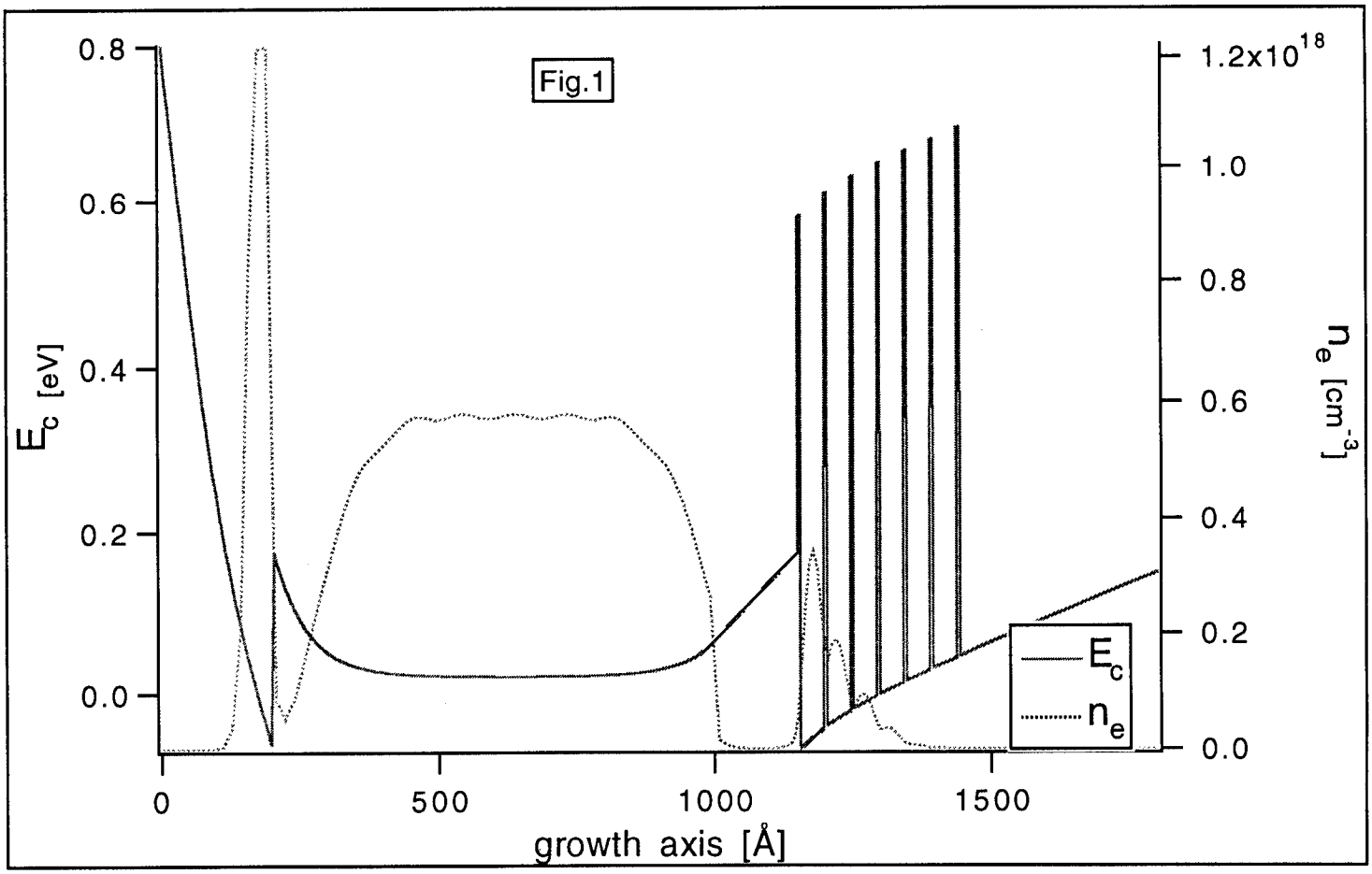


Fig.1
0.8
0.6
0.4
0.2
0.0
E_c [eV]
1.2x10^{18}
1.0
0.8
0.6
0.4
0.2
0.0
n_e [cm^{-3}]
0
500
1000
1500
growth axis [Å]
E_c
n_e

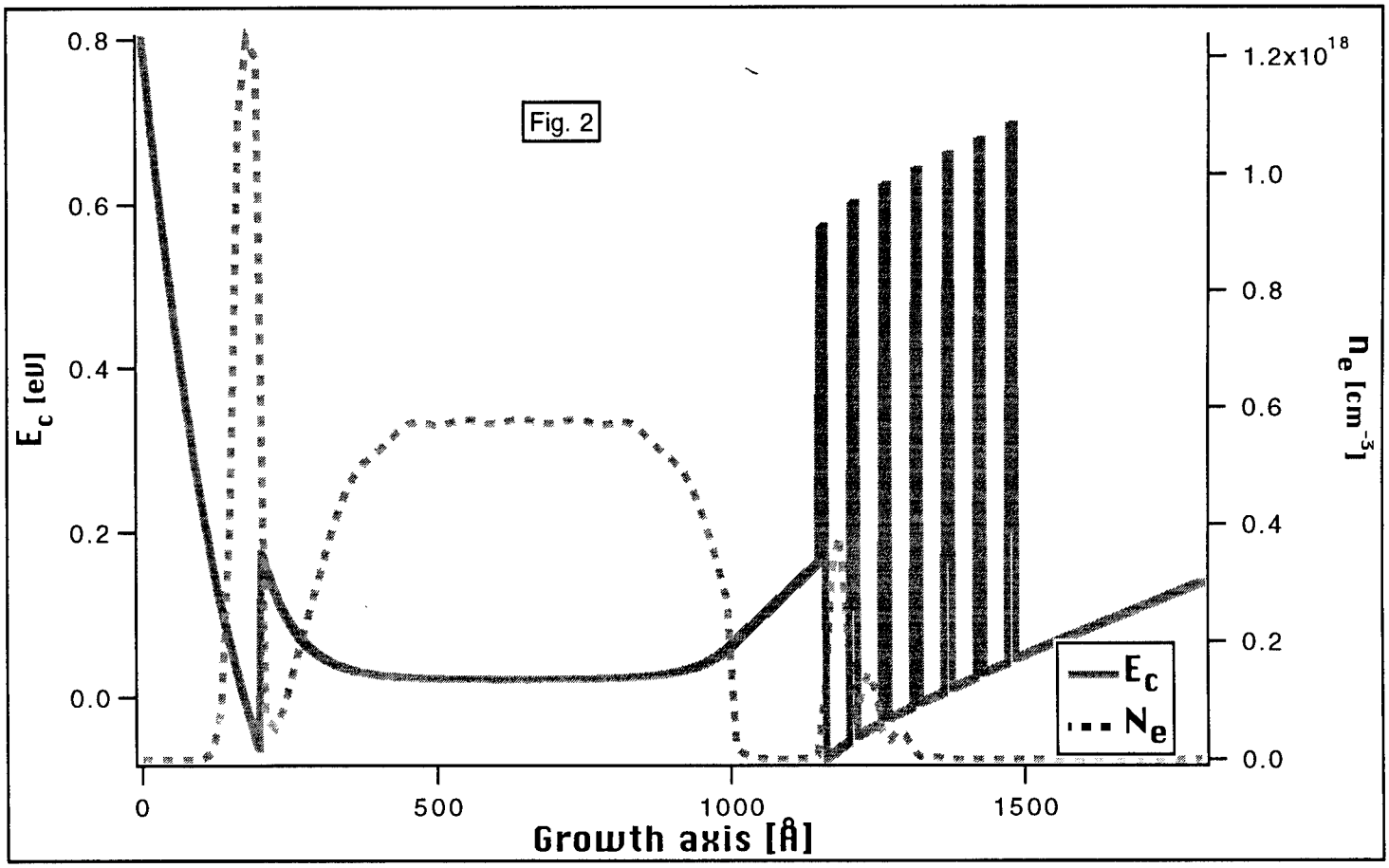


Fig. 2
0.8
0.6
0.4
0.2
0.0
E_c [eV]
1.2x10^{18}
1.0
0.8
0.6
0.4
0.2
0.0
n_e [cm^{-3}]
0
500
1000
1500
Growth axis [Å]
E_c
N_e

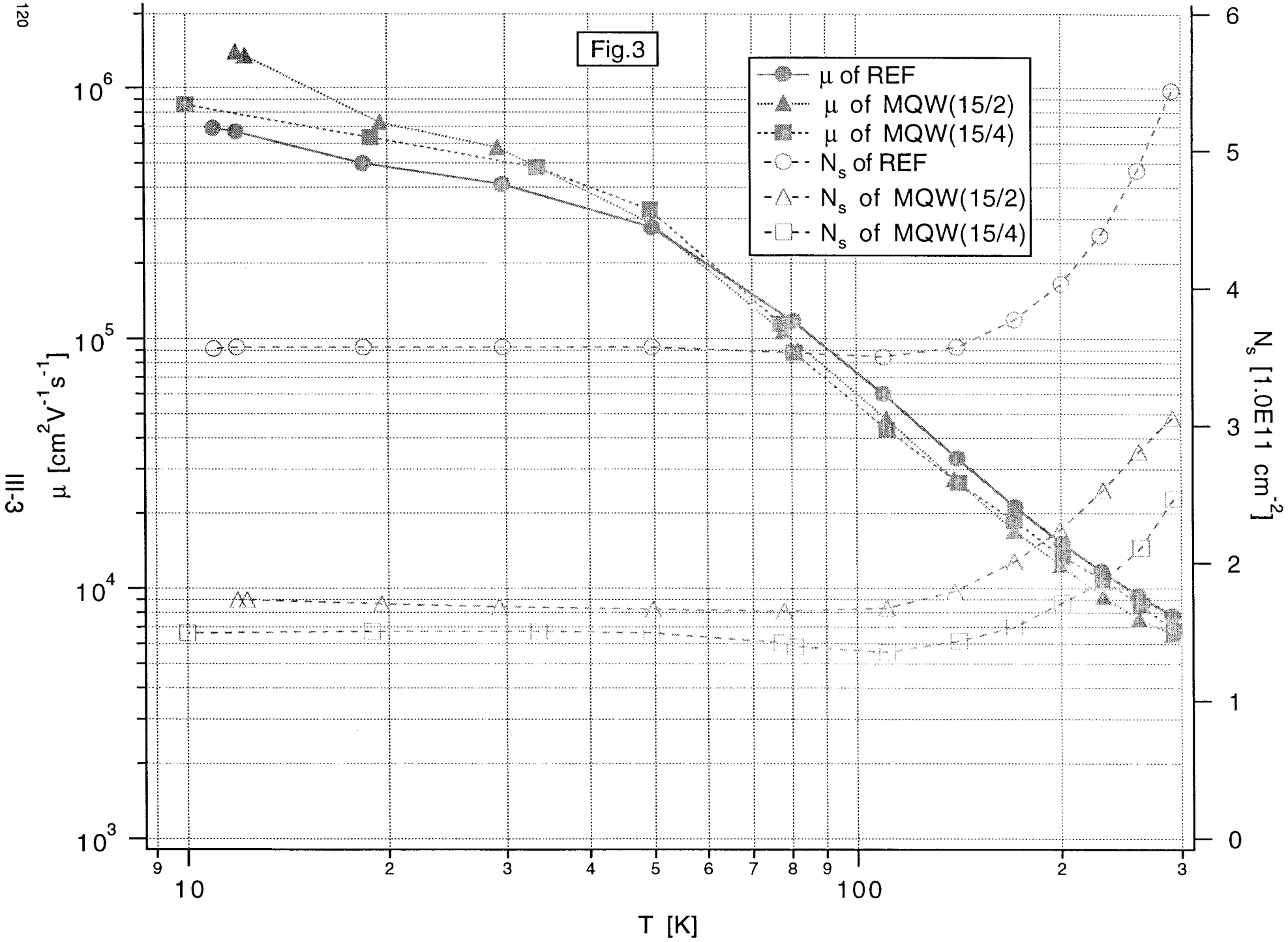


Fig.3
μ of REF
μ of MQW(15/2)
μ of MQW(15/4)
N_s of REF
N_s of MQW(15/2)
N_s of MQW(15/4)
μ [cm²V⁻¹s⁻¹]
N_s [1.0E11 cm⁻²]
T [K]
10³
10⁴
10⁵
10⁶
0
1
2
3
4
5
6
9
10
2
3
4
5
6
7
8
9
100
2
3

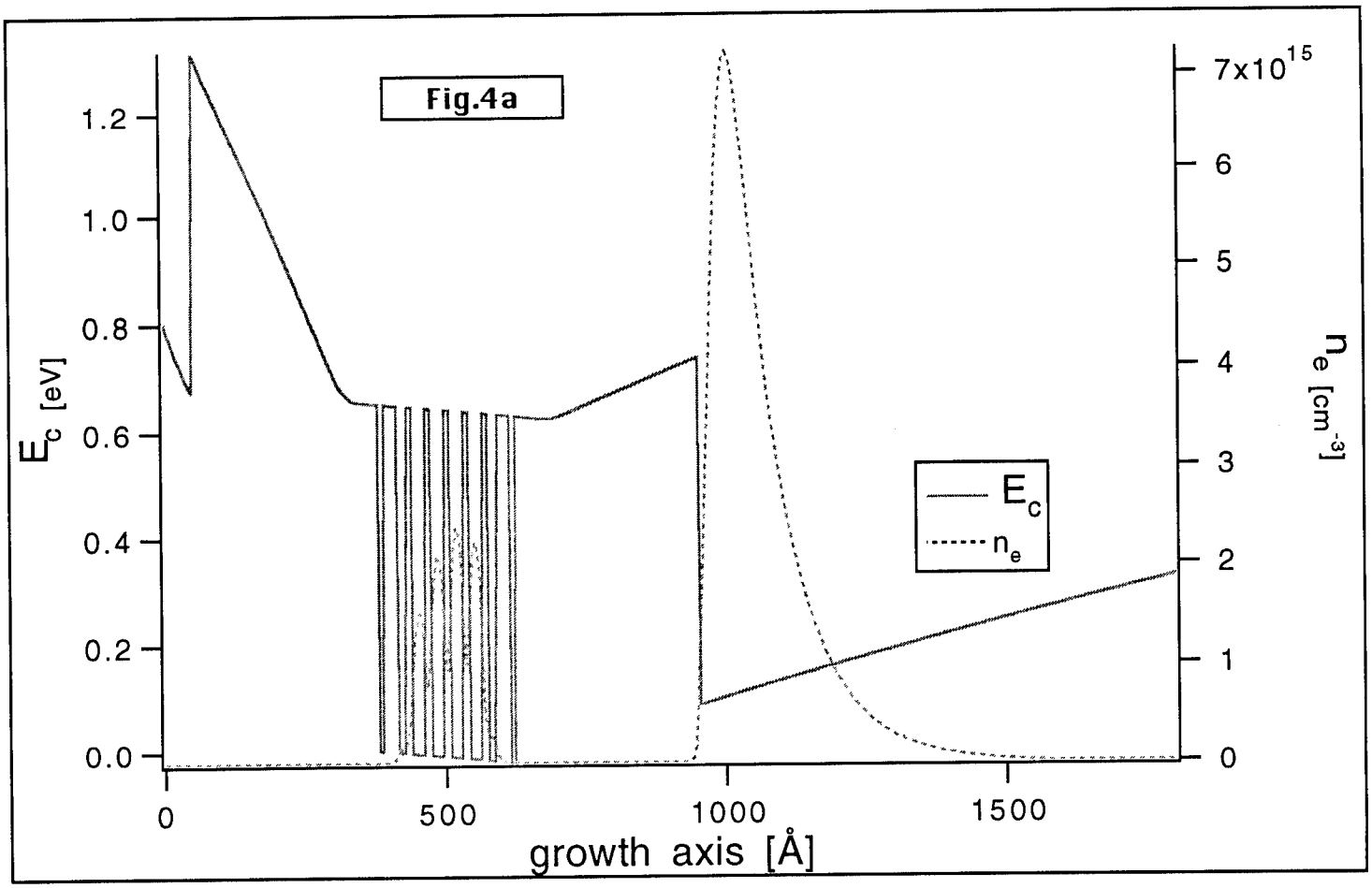


Fig.4a
1.2
1.0
0.8
0.6
0.4
0.2
0.0
E_c [eV]
7x10^{15}
6
5
4
3
2
1
0
n_e [cm^{-3}]
E_c
n_e
0
500
1000
1500
growth axis [Å]

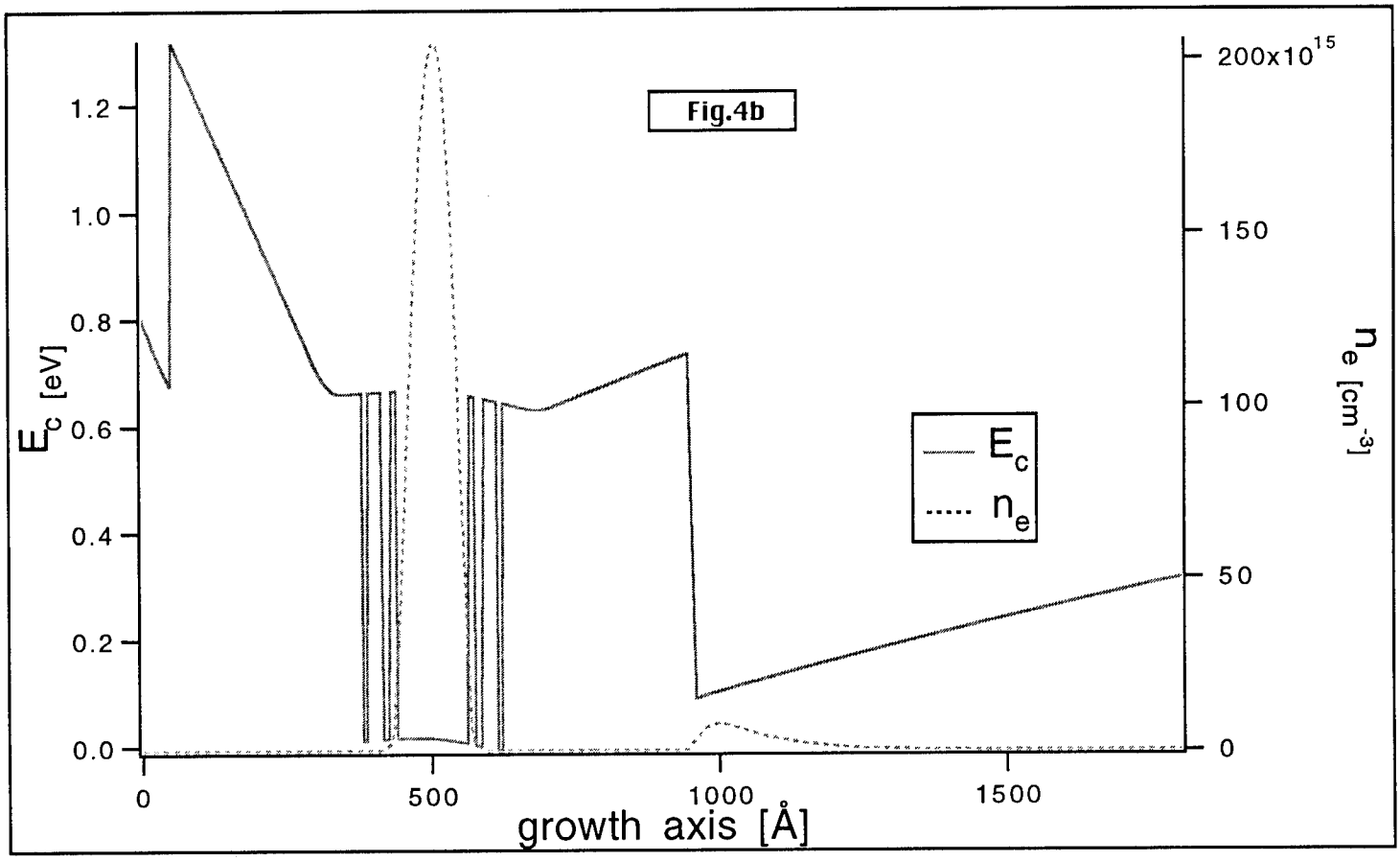


Fig.4b
1.2
1.0
0.8
0.6
0.4
0.2
0.0
E_c [eV]
200x10^{15}
150
100
50
0
n_e [cm^{-3}]
E_c
n_e
0
500
1000
1500
growth axis [Å]

Flow Modulation Growth of III-V Compound Semiconductors Using a Multichamber OMVPE Reactor

B.L. Pitts, D.T. Emerson, M.J. Matragrano†, K.L. Whittingham,
B.P. Butterfield and J.R. Shealy

OMVPE Facility, Cornell University
School of Electrical Engineering, † Dept. of Material Science and Engineering
Ithaca, N.Y. 14853

Abstract

The flow modulation growth of high purity InP and GaAs based III-V compound semiconductors is demonstrated using a multichamber Organometallic Vapor Phase Epitaxy apparatus. Flow modulation is performed by rotating substrates through spatially separated group III and group V rich zones without valve switching. This system has multiwafer capability and excellent compositional and thickness uniformity. In addition, using the appropriate flow conditions, phosphide and arsenide based superlattice structures can also be produced without the use of mechanical valve switching. High purity InP and GaAs have been realized, yielding low temperature (77 K) mobilities exceeding 110,000 and 115,000 cm^2/V s, respectively. Excellent transport and optical properties were also observed for GaInP/GaAs and GaInAs/InP structures. The multichamber reactor is a useful tool for both research and manufacturing of optoelectronic and high speed compound semiconductor devices.

I. INTRODUCTION

The production of high purity III-V compound semiconductors has led to the realization of many optoelectronic and high speed electronic devices. Optical devices spanning from the infrared to the visible and electronic devices with picosecond switching speeds have been fabricated using InP and GaAs based material systems. Over the years, there has been a growing demand for the manufacture of these devices in high volume at a relatively low cost. This, however, places strict demands on crystal growth technology. Abrupt heterojunctions, sharp doping profiles and high purity material are all necessary requirements in obtaining high performance device structures. In addition, a multiwafer process which has excellent wafer to wafer thickness and compositional uniformity is also desired. Advanced growth techniques such as migration enhancement (flow modulation epitaxy) and atomic layer epitaxy have been suggested to possibly meet many of these demands. Commercial vendors have offered a wide variety of crystal growing systems utilizing both Molecular Beam Epitaxy and Organometallic Vapor Phase Epitaxy (OMVPE) techniques to mass produce III-V semiconductor devices.

0-7803-0894-8/93 $3.00

In this paper, we describe a multichamber OMVPE reactor with multiwafer capability, which produces high purity III-V compound semiconductor structures with excellent compositional and thickness uniformity. In this reaction chamber, substrates are rotated through spatially separated group III and group V rich regions without valve switching [1]. These regions are confined by injecting the proper gas flows to reduce zone intermixing. This enables the formation of abrupt heterostructures as well as the use of advanced growth techniques (flow modulation epitaxy) to be performed without requiring mechanical valve switching.

We will demonstrate the versatility of the multichamber reactor by discussing various experiments employing flow modulation epitaxy (FME). AlAs/GaAs superlattice structures were produced by rotating substrates through localized growth zones separated by large hydrogen fluxes. Excellent compositional and thickness uniformity will be presented for GaInP lattice matched to GaAs. High purity GaAs and InP have been produced yielding low temperature (77 K) mobilities exceeding 115,000 and 110,000 cm^2/V s, respectively. Excellent transport properties were also observed for GaInAs/InP and GaInP/GaAs structures.

II. MULTICHAMBER REACTOR

The vertical barrel, multichamber reaction cell is illustrated schematically in Figure 1. The outer cell wall is made of 6-inch diameter high purity quartz and the inner wall is a 2 inch diameter quarz ampoule which serves to localize growth regions. This system is operated at reduced pressure (76 torr) to eliminate gas recirculation due to convection forces. Palladium diffused H_2 is used as the carrier gas. Typically, the gas flow is 30 slm, while the gas velocity is maintained at 30 cm/s. A barrel style susceptor holding up to eighteen 1.5 inch wafers, is inductively heated by RF radiation. The susceptor has three tiers, with each tier having six wafer slots. Two group III rich regions are spatially separated by large hydrogen fluxes. Triethylgallium (TEG), Trimethylindium (TMI) and dopants sources are injected in the left zone, while trimethylaluminum and a second TEG source are injected in the left zone. Hydrides (i.e. arsine and phosphine) are uniformly injected into the growth chamber through a distributor ring located at the uppermost portion of the cell. Flow modulation is carried out by rotating substrates through group III and group V rich regions without valve switching. The flow modulation configurations can be varied by adjusting either reactant fluxes, carrier flows or susceptor rotation speed. The reaction cell is periodically cleaned by heating the susceptor to 900 °C (hot wall) while injecting HCl through the distributor ring. It has been reported that operating with a clean reaction cell eliminates oval surface defects in In-containing compounds [2]. The pumping port is located directly under the right growth zones. Experiments were also carried out using a distributed pumping ring.

The standard vent-run configuration is used to reduce switching transients. Organometallics (OMs) and hydrides are switched at the head the reaction cell to enhance heterostructure abruptness. Gas lines are made of stainless steel tubing which are bakeable to 200 °C. To enhance

transport capabilities, the dopants, hydrides and each individual OM gas loop have independent H_2 carriers. The OM pressures are controlled using a needle valve which enables better control of fluxes. An ultrasonic analyzer is used to monitor and regulate the TMI transport to the chamber. Al-Ga-In gettering melts are used to reduce oxygen and H_2 contamination in the reactant gases [3]. This system is also equipped with a quadrapole mass analyzer for gas sampling prior to every run.

III. FLOW MODULATION GROWTH

Flow modulation Epitaxy (FME) has been reported to be a useful technique for obtaining smooth heterointerfaces, sharp doping profiles and reduced deep level traps [4]-[6]. Typically, FME is performed by alternately supplying group III and group V species to the growth surface which enhances the surface mobility of the adatoms. Unlike atomic layer epitaxy, where the growth rate is independent of the amount of injected source material, FME is group III mass transport limited. Therefore, high growth rates can be realized. Flow Modulation in a conventional reaction chamber requires the use of repetitive valve switching which ultimately reduces the lifetime of the valves. Replacing these valves usually requires exposing gas lines to H_2O and O_2. However in the multichamber configuration, flow modulation is executed by alternately exposing substrates to group III and group V rich environments by substrate rotation. As a result, this eliminates the need for valve switching.

Many flow modulation configurations have been investigated by modifying the flow configuration and pumping scheme. Experiments were performed without susceptor rotation to determine the substrate exposure cycle. In all configurations, a measurable growth rate was observed on the opposite side of the reaction cell corresponding to the left and right zone inlets. Several flow modulation schemes were investigated and compared by investigating the film quality of GaInP closely lattice matched to GaAs. Figure 2a shows the exposure cycle for what was determined to be the highest quality GaInP lattice matched to GaAs. This configuration was created by injecting high carrier flow between the group III zones (10 slm) and 4.5 slm in both the left and right zones. Strong photoluminscence intensity and carrier concentration in the mid 10^{16} cm^{-3} were observed. Other flow configurations produced GaInP which exhibited broad X-ray diffraction features. A distributed circular pumping port was installed to further localized zones. However, X-ray diffraction spectra displayed weak and broadened features from the GaInP films. This was observed for all flow configurations using the distributed circular pumping port.

IV. SPATIALLY SEPARATED GROWTH ZONES

Applying the appropriate flow conditions, localized growth regions can be obtained which enables the formation superlattice structures. Changing group III flux and rotation of the susceptor changes the periodicity of the superlattice structure. This is a flexiable technique to

prepare superlattices without valve switching. As stated earlier, although intermixing occurs in all configurations, device quality superlattices can be produced by this method. Figure 3, shows the a TEM image of a 40 ÅAlAs/GaAs superlattice produced in the multichamber cell. Equal amounts of gas were injected into each inlet to balance the flows, which minimizes zone intermixing.

V. COMPOSITIONAL AND THICKNESS UNIFORMITY

The compositional and thickness uniformity is presented for GaInP lattice matched to GaAs. A compositional uniformity map, generated using X-ray diffraction, is shown in Figure 4. Indium composition varied $\pm 0.5\%$ over a 6 cm^2 area. The wafer to wafer composition varied $\pm 0.7\%$ for wafers placed on the top tier of the susceptor. The compositional uniformity was measured on wafers located on the second tier during the same run. There, the composition varied $\pm 1.5\%$ over a 4 cm^2 area. We also observed that the In composition decreased down stream along the wafer. A decrease of 3% In concentration was observed from the top to the second tier. Thickness uniformity was measured using SEM for the wafer located on the first tier. The thickness varied $\pm 3.3\%$ over a 6 cm^2 area.

VI. GROWTH OF GaAs AND InP BASED MATERIAL

Both InP and GaAs based material were produced using FME. High purity GaAs, AlGaAs, and GaInP were on produced GaAs substrates. A process has been developed for arsenides based compounds which produces high purity material using minimal arsine flows [7]. The V/III ratio dependence on 77 K mobility for GaAs grown at 635 °C is given in Figure 5. Low temperature (77 K) mobilities exceeding 115,000 cm^2/V s and room temperature of 8000 cm^2/V s for V/III ratio less than 5. A 77 K mobility of 93,000 cm^2/V s was measured for films produced with a V/III ratio of 1.8. Strong excitonic features were observed for films produced using a V/III ratio of unity. Using trimethylamine alane, TEG and arsine, excitonic features were observed for $Al_{0.15}Ga_{0.85}As$ for V/III ratios as low as 7.5 [8]. GaInP was grown over a wide range of growth conditions yielding both ordered and disordered layers. Low temperature (1 K) and room temperature PL linewidths of less than 10 and 26 meV, respectively were observed for FME grown GaInP. A room temperature mobility of 3,000 cm^2/V s and 77 K mobility of 14,000 cm^2/V s for GaInP grown at 645 °C and a V/III ratio of 118.

InP and GaInAs lattice matched to InP substrates exhibited excellent photoluminscence and electronic properities. InP grown at 600 °C and at a V/III ratio of 175 yielded a 77 K mobility of 110,000 cm^2/V s. Low temperature PL revealed neglegible acceptor related transitions, indicative of n-type material with low compensation. GaInAs nearly lattice matched to InP also showed excellent low temperature photoluminscence properties having linewidths as low as 1.7 meV. In addition, room temerature mobilities as high 11,000 cm^2/V s were observed for GaInAs nearly lattice matched to InP.

VII. CONCLUSIONS

In summary, the advantages of a multichamber OMVPE system have been demonstrated. This system has multiwafer capability which also has excellent compositional and thickness uniformity for GaInP lattice matched to GaAs. Spatially separated group III rich and group V rich regions enable the formation of superlattice structures and flow modulation growth without valve switching. Flow modulation exposure cycles can be altered by adjusting the carrier flows and source flux. High purity GaAs and InP was produced, both having low temperature mobilities exceeding 110,000 cm^2/V s. The multichamber concept has been extended by introducing a four chamber reaction cell. Presently under contruction, this system can accommodate six 3 inch wafers. Both FME and ALE will be accomplished without valve switching. This system represents the next generation of multichamber designs which are capable of producing III-V semiconductor device structures in high volume.

Acknowledgements

The authors wish to thank B. Sun for providing the compositional uniformity data. This work was supported by the Joint Services Electronics Program under grant No. F49620-90-C-0039, the Strategic Defense Initiative Objective under contract No. N00014-89-J-1311, and the Defense Advanced Research Projects Agency under contract No. MDA97290C0058 Optoelectronics Technology Center.

REFERENCES

[1] J.R. Shealy,"Characterization of AlGaAs/GaAs Heterostructures Grown in A Multichamber OMVPE Apparatus", J. of Crystal Growth, vol. 87, pp.350-356, 1988.

[2] B.L. Pitts, M.J. Matragrano, D.T. Emerson, B. Sun, D.G. Ast and J.R. Shealy,"Coherency Limits of Tetragonal III-V In-Containing Alloys on GaAs and InP Substrates", to be published in Inst. Phys. Conf., 1993.

[3] J.R. Shealy, V.G. Kreismanis, D.K. Wagner and J.M. Woodall,"Improved Photoluminescence of Organometallic Vapor Phase Epitaxial AlGaAs Using a New Gettering Technique on the Arsine Source", Appl. Phys. Lett., vol. 42, pp.83-85, 1983.

[4] N. Kobayashi, T. Makimoto, Y. Yamauchi and Y. Horikoshi,"Flow-Rate Modulation of GaAs and AlGaAs", J. of Appl. Phys., vol. 66, pp.640-651, 1989.

[5] N. Kobayashi and Y. Horikoshi,"Improved Flatness in GaAs/AlGaAs heterointerfaces Grown by Flow-Rate Modulation Epitaxy", Appl. Phys. Lett., vol. 50, pp.909-911, 1987.

[6] N. Kobayashi, T. Makimoto and Y. Horikoshi,"Efficient Si Planar Dopiinng In GaAs by Flow-Rate Modulation Epitaxy", Jpn. J. of Appl. Phys., vol. 25, pp.L746-L748, 1986.

[7] N. Kobayashi and Y. Horikoshi,"Improved Flatness in GaAs/AlGaAs heterointerfaces Grown by Flow-Rate Modulation Epitaxy", Appl. Phys. Lett., vol. 50, pp.909-911, 1987.

[8] B.L. Pitts, D.T. Emerson and J.R. Shealy, "Arsine Flow Requirement for the Flow Modulation Growth of High Purity GaAs using Adduct-Grade Triethylgallium", Appl. Phys. Lett., vol. 61, pp. 2054-2056, 1992.

[9] B.L. Pitts, D.T. Emerson and J.R. Shealy, "Gas Phase Reactions of Trimethylamine Alane in Low Pressure Organometallic Vapor Phase Epitaxy of AlGaAs", Appl. Phys. Lett., vol. 62, pp. 1821-1823, 1993.

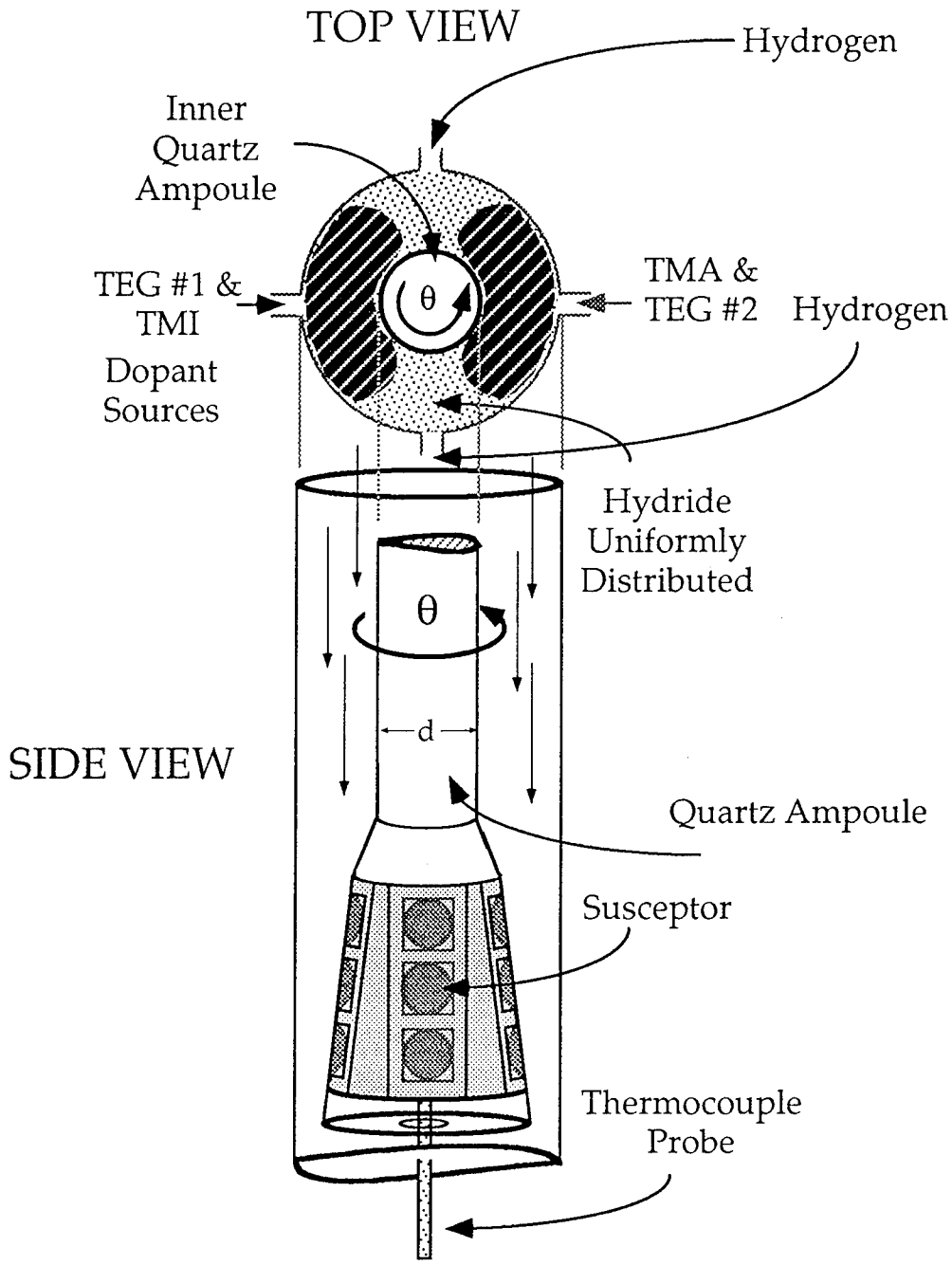


Figure 1. Schematic illustration of implementation of flow modulation epitaxy in the multichamber cell. Dopant sources, TEG and TMI are injected in the left zone, while TMA and TEG are injected into right zones. Two hydrogen carriers serves to isolate the group III rich zones. The hydrides are uniformly around the cell. The inner quartz ampoule (diameter-*d*) serves to separate the reactant fluxes of each deposition zone.

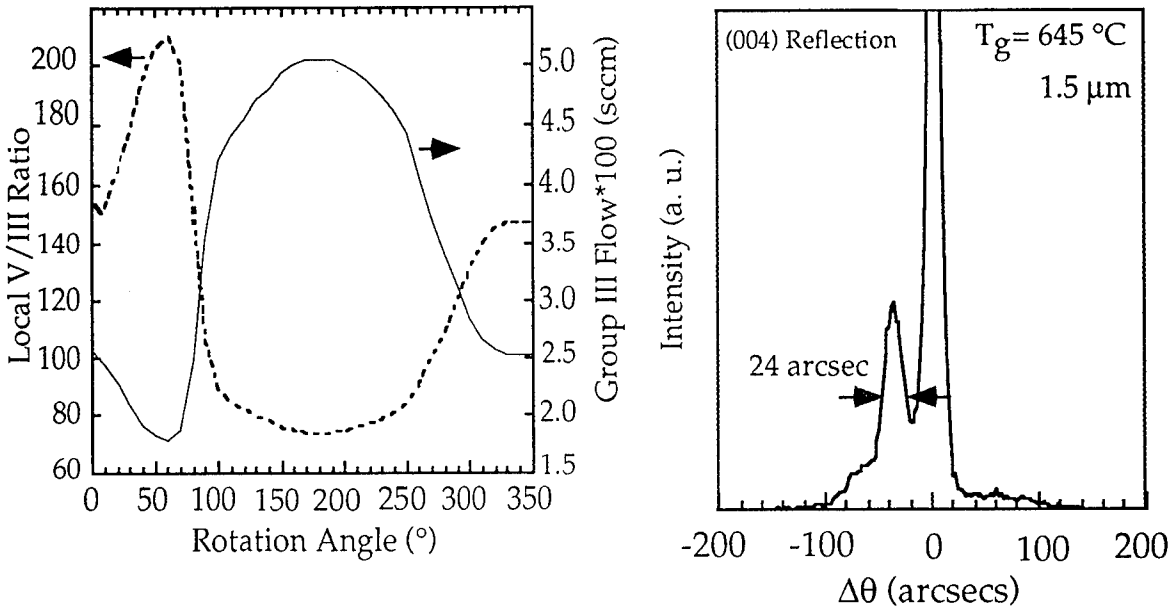


Figure 2. (a) Flow modulation exposure cycle for premixed GaInP lattice matched to GaAs. The exposure cycle was calculated by performing no rotation growth experiments to determine growth rate around the reaction chamber. (b) Corresponding X-ray spectra of GaInP produced using the flow modulation scheme in part (a).

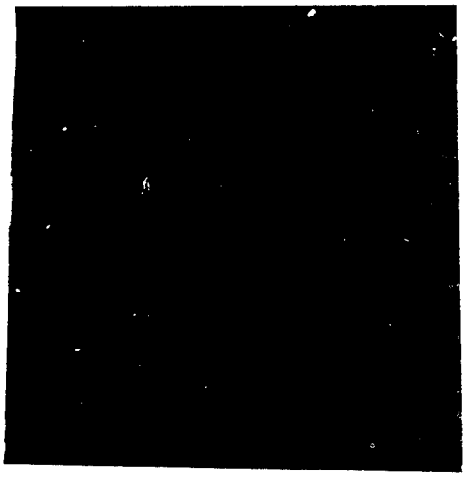


Figure 3. TEM image of an AlAs/GaAs 40 Å superlattice produced in the multichamber reactor. Equal flows were injected in each inlet port to provide optimum spatial confinement.

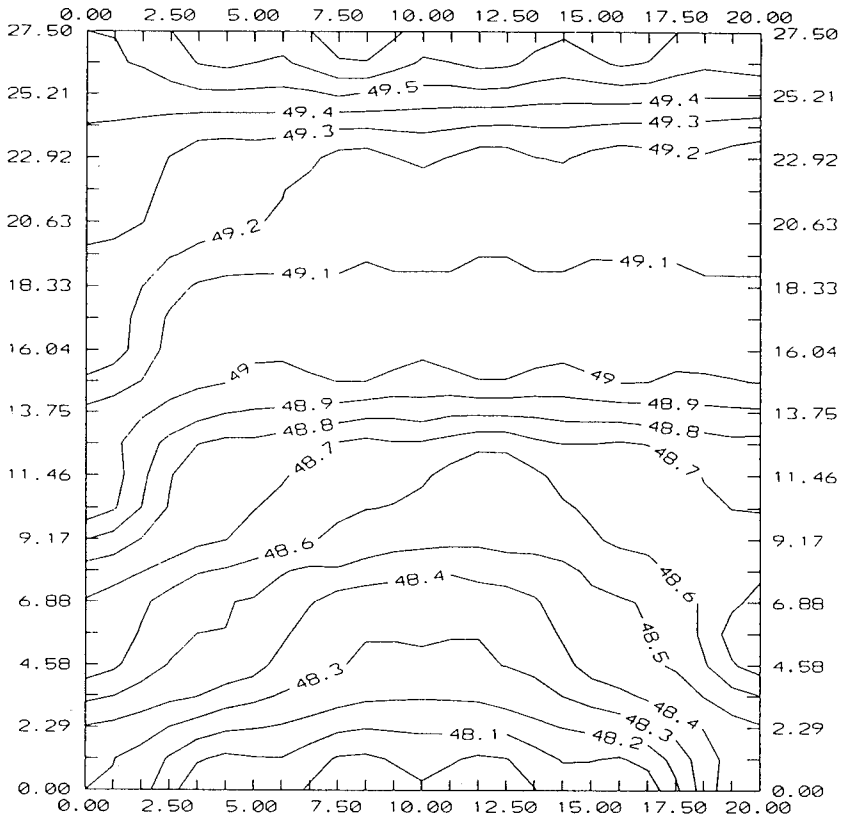


Figure 4. Compositional uniformity map of GaInP lattice matched to GaAs. The composition varies ±0.5% across 6 cm^2 area. The map was created using double crystal X-ray diffraction.

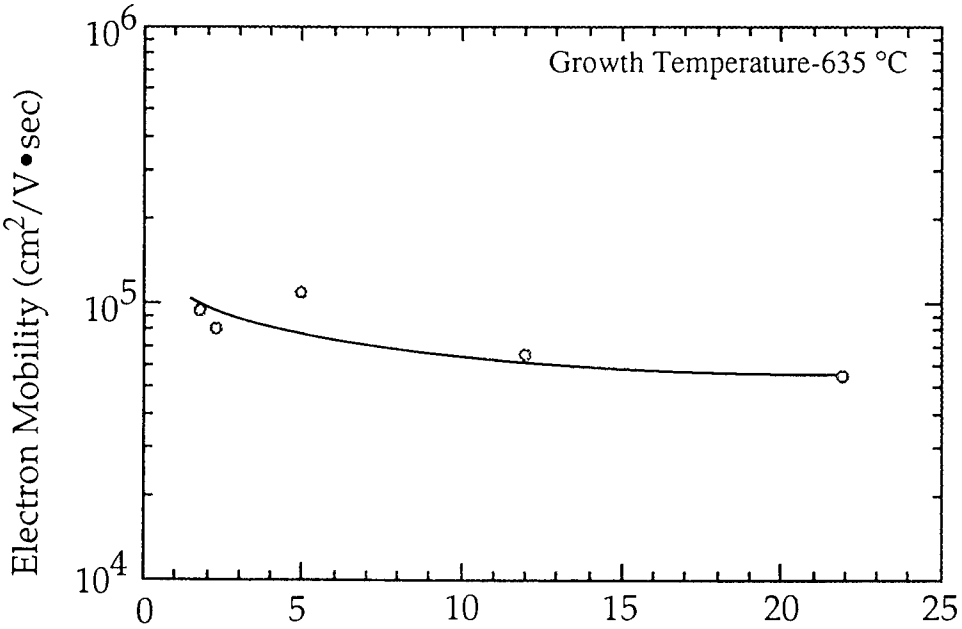


Figure 5. Low temperature (77 K) mobility dependence on V/III ratio for GaAs produced by flow modulation epitaxy.

High Speed Optical Detectors for Monolithic Millimeter Wave Integrated Circuits

K. Litvin, J. Burm, D. Woodard, W. Schaff, and L.F. Eastman
School of Electrical Engineering, Cornell University
Ithaca, New York 14853

ABSTRACT

Metal-semiconductor-metal photo diodes with interdigitated Schottky barrier fingers are being developed for applications in monolithic optical receiver circuits with the purpose of detecting millimeter wave modulation signals being transmitted via an optical carrier. The devices are planar and incorporate submicron finger spacings and a thin absorption region for speed with a buried stack of tuned Bragg reflectors for enhanced sensitivity at the carrier wavelength. These devices are being integrated with short-gate MODFET amplifiers to form the complete monolithic integrated optical receiver circuit.

DEVICE DESCRIPTION

Metal-semiconductor-metal photo diodes with interdigitated Schottky barrier fingers are being fabricated in the AlGaAs material system using electron beam lithography techniques. Both linear and circular detector finger geometries have been examined (Fig. 1). Such photo detectors are capable of detection bandwidths in excess of 100 GHz [1-3]. The devices have a planar construction making them ideal candidates for monolithic integration [4,5]. The detector structures were designed with the aid of a computer program and then grown by molecular beam epitaxy. The photo diodes in this design incorporate sub-micron electrode spacings and a thin absorption layer of GaAs, these features limit the transit time of the optically generated carriers thus providing the mechanism for millimeter frequency operation (Fig. 2).

Device sensitivity to low intensity light is enhanced with the inclusion of a highly reflective (94%) stack of tuned Bragg reflectors buried below the absorption layer. The Bragg reflector stack is composed of nonabsorbing alternating quarter wavelength layers of $Al_{0.9}Ga_{0.1}As$ and $Al_{0.35}Ga_{0.65}As$. The reflective stack causes incident photons, of the appropriate wavelength, to make a second pass through the absorption layer thus creating additional carriers for collection at the electrodes.

0-7803-0894-8/93 $3.00

Additional sensitivity refinements were made by including a top surface optical impedance matching layer of $Al_{0.35}Ga_{0.65}As$. This layer reduces top surface reflectance to nearly negligible values (0.5%). The interfaces between the GaAs absorption layer and the matching layer above it and the bottom of the absorption layer and the top surface of the Bragg reflector stack also form a potential energy well for the photo generated carriers. This is due to the energy band discontinuity of the material layers. The potential energy well serves to confine the photo generated carriers to the thin absorption region where they can be rapidly collected by the electrodes on the top surface. (Also prevents carrier loss by rapid surface recombination (v_S @ $3x10^6$ cm/sec)). The combined effects of the finite line width of the Bragg reflector stack, absorption layer and top surface matching layer introduces some wavelength discrimination capability into the device because there is an increased optical sensitivity at the design wavelength and a decreased sensitivity at other wavelengths away from this line (Fig. 3).

MODEL and MEASUREMENT RESULTS

The topology of a microwave equivalent circuit model for the detector diodes was determined using S-parameter measurements of the devices and of associated short circuited test structures (Figs. 4 & 5). The S-parameter measurements on the short circuit test structures readily allow for the extraction of the finger electrode resistance and inductance. In addition to the dc resistance in the fingers there is an additional frequency dependent skin effect resistance . The skin effect resistance is proportional to the square root of the frequency. The proportionality constant can be determined by comparing the frequency dependence of the measured device input impedance with that of the model. The MSM detector capacitance is extracted from the device measurements and the excess parasitic capacitance is determined by comparing the measured capacitance to that calculated for an MSM structure using the elliptic integral method [6]. The parasitic capacitance is the result of the finite length of the interdigitated fingers and the close proximity of the contact pad of opposite polarity to the open ends of the fingers. The equivalent circuit model is used to determine the passive element frequency response and time domain impulse function (Fig. 6) .

A Monte Carlo simulation of the transport behavior of the optically generated carriers, which takes into account the material dimensions, electrode geometry, bias conditions ,and optical absorption properties, is used to produce modeling information about the time domain response of the carriers [7,8]. The FFT of the Monte Carlo results provides the carrier frequency response function (Fig. 7). This information is combined with the circuit model information (multiplied in the frequency domain or convolved

in the time domain) to give the composite frequency response and impulse function of the MSM detector (Fig. 8).

Static optical tests were done with a white light source and monochromator in order to characterize the spectral response of the diodes. The response signal was normalized to the white light spectrum (Fig. 9). A large size detector was used in the experiment for better sensitivity. Figure 9 shows that the expected peak in the response for a 800 nm wafer design was down-shifted slightly, this probably indicates that the angle of incidence was not quite normal to the sample. Photospectrometer measurements of the wafer reflectance support this conclusion.

Optical measurements using a short pulse (120 fs) mode locked laser operating at 780 nm with a beam diameter of 10 μm indicated a carrier transit time limited bandwidth in excess of 40 GHz (this was the limit of the measurement system). The detector response was observed using a HP2782 spectrum analyzer. The measurement on a detector diode with 23 μm long fingers, 0.5 μm finger width and 0.5 μm gap between the fingers fabricated on a GaAs wafer designed for 750 nm light had the following results: At 7 volts applied bias and 80 mW average incident power, it showed a flat response up to at least 40 GHz. When the average incident light power was doubled to 160 mW the response at 39 GHz was -6 dB down due to excess carrier generation. The dynamic range of the detectors was then measured using a Ti:sapphire laser operating at 770 nm with a variety of neutral density filters . The dynamic range was found to be 33 dB.

Optical measurements employing a semiconductor mode locked laser as a short pulse optical source were also performed and the detector response was observed in the time domain using a sampling oscilloscope. The full width half maximum of the output pulse was 15 ps. This value was again essentially the limit of the measurement system's capability.

DEVICE INTEGRATION

The MSM photodiodes are being integrated with short-gate MODFETs for use as monolithic integrated optical receiver circuits capable of detecting a 44 GHz modulation signal being transmitted via an optical carrier [9]. The transistor layers are grown on top of the detector layers in the MBE system. A slow etch-rate citric acid solution is used for the transistor mesa isolation. The wafer top surface reflectance is monitored with a photospectrometer until the reflectance minima shifts to the design

wavelength. The MODFETs will be used as low noise amplification devices in the optical receiver circuits. The current transistor design has a source-drain spacing of 2 μm with a gate length of 0.25 μm and gate width of 100 μm. We are in the process of developing a transistor with a gamma-gate and self-aligned ohmic metalization in order to reduce the source resistance and therefore improve the noise and frequency characteristics of the devices [10,11]. In test structures we have achieved gates with 0.25 μm footprint and 0.6 μm cross-section along the top. The gate had an overhang ratio of more then 2:1 and a source-drain spacing of 0.6 μm. With self-aligned metalization the source-gate spacing was 0.1 μm and the drain-gate spacing was 0.25 μm (Fig. 10). Coplanar waveguide interconnections are being used in the circuit design.

SUMMARY

Metal-semiconductor-metal photodetector diodes have been designed with the aid of a computer program. The material was grown by MBE and the devices were fabricated with electron beam lithography. The detectors incorporate features for high speed operation, (submicron electrode spacings, thin absorption layer, and a potential well to confine optically generated carriers to the absorption layer), with features to enhance the device sensitivity, (buried Bragg reflector stack, and a top surface optical impedance matching layer). Device measurements indicate top surface reflectance of 0.5% with an operation bandwidth in excess of 40 GHz and a dynamic range of 33 dB. The devices are being integrated with short-gate MODFETS for applications in monolithic millimeter wave optical receiver circuits.

ACKNOWLEDGMENTS

This work is supported by the United States Air Force Rome Laboratory and Hughes Aircraft Company Research Laboratory. The authors would also like to thank Kirk Giboney and John Bowers of U.C. Santa Barbara for valuable assistance with the time domain measurements, Martin Jaspan and Clifford Pollock of Cornell University for their work with the Ti:sapphire laser. The guidance of Richard Tiberio of Cornell University with gate level processing technology is also appreciated

REFERENCES

[1] B. J. Van Zeghbroeck, W. Patrick, J. Halbout, and P. Vettiger, "105-GHz Bandwidth Metal-Semiconductor-Metal Photodiode," IEEE Electron Device Lett. , vol. 9, pp. 527-529, 1988.

[2] S.Y. Chou, Y.Liu, W.Khalil, T.Y. Hsiang, and S. Alexanddrou, "Ultrafast nanoscale metal-semiconductor-metal photodetectors on bulk and low temperature grown GaAs," Appl. Phys . Lett., vol 61, no. 7, pp. 819-821, 17 Aug. 1992.

[3] S.Y. Chou, Y. Liu, and P.B. Fischer,"Tera-hertz GaAs metal-semiconductor-metal photodetectors wiyh 25 nm finger spacing and finger width," Appl. Phys. Lett., vol. 61, pp. 477-479, 1992.

[4] S. J. Wojtczuk, J. M. Ballantyne, S. Wanuga, and Y. K. Chen, "Comparative study of easily integrable photodetectors," IEEE J. Lightwave Technol., vol. LT-5, pp. 1365-1370, 1987.

[5] D. L. Rogers, "Integrated optical receiver using MSM detectors," in Proc. LEOS '90.

[6] Y. C. Lim, and R. A. Moore, "Properties of Alternately Charged Coplanar Parallel Strips by Conformal Mappings," IEEE Transactions on Electron Devices, vol ED-15, pp. 173-180, March 1968.

[7] L. Reggiani (ed), *Hot-Electron Transport in Semiconductors* , Springer-Verlag, 1985

[8] J. B. D. Soole,and H. Schumacher, " Transit Time Limited Frequency Response of InGaAs MSM-Photodetectors," IEEE Tranactions on Electron Devices, vol ED-37, p. 2285, 1990

[9] L. D. Nguyen, P. J. Tasker, D. C. Radulescu, and L. F. Eastman, "Design, Fabrication, and Characterization of Ultra High Speed AlGaAs/InGaAs MODFET's," IEDM Tech. Digest, pp. 176-179, 1988.

[10] L.D. Nguyen, A.S. Brown, M.A. Thompson, L.M. Jelloian, L.E. Larson and M. Matloubian, " 650 A Self-Aligned-Gate Pseudomorphic AlInAs/GaInAs High Electron Mobility Transistors," IEEE Electron Device Letters, vol 13, no. 3, pp 143-145, March 1992.

[11] R.C. Tiberio, J.M. Limber, G.J. Galvin, and E.D. Wolf, "Electron Beam Lithography and Resist Processing for the Fabrication of T-Gate Structures", SPIE Vol. 1089, Electron-Beam, X-ray and Ion Beam Technology: Submicrometer Lithographies VIII, pp. 124-131, 1989.

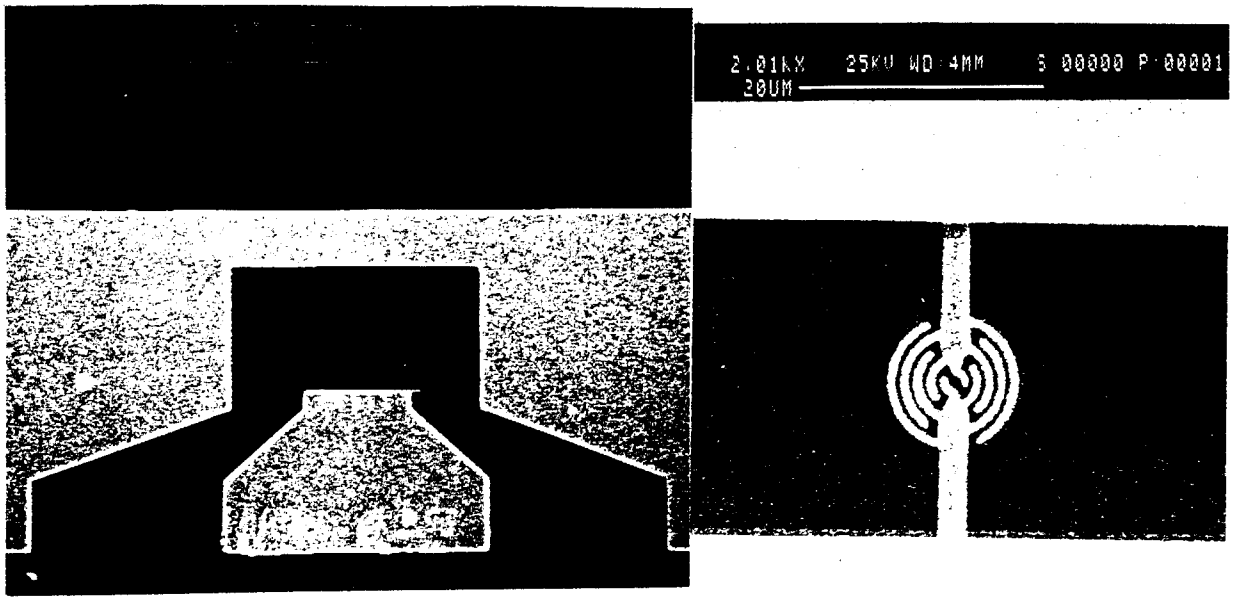


Figure 1. Two MSM Photodiode diode geometries. The linear device has 0.35μm fingers and gaps with a 20μm overlap. The circular device has 0.5μm fingers and gaps and a 10μm diameter.

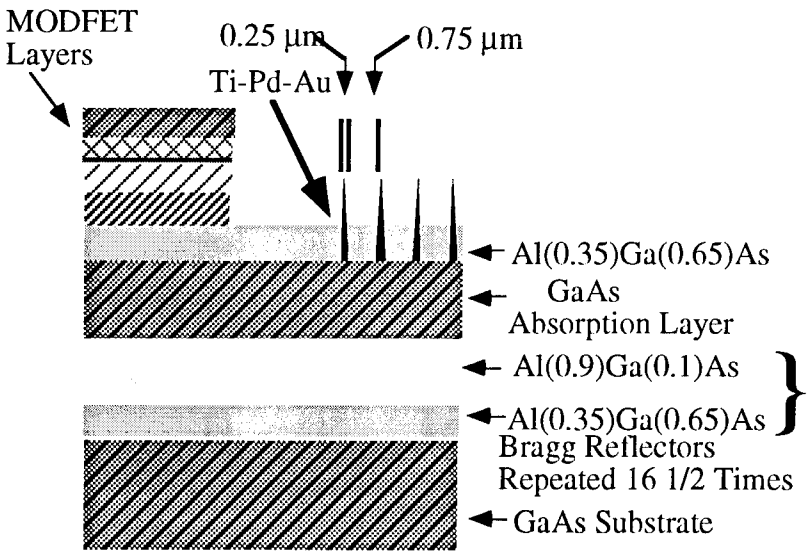


Figure 2. Cross-sectional view of the MSM photodiode material structure showing the top surface impedance matching layer, absorption layer, and Bragg reflector stack.

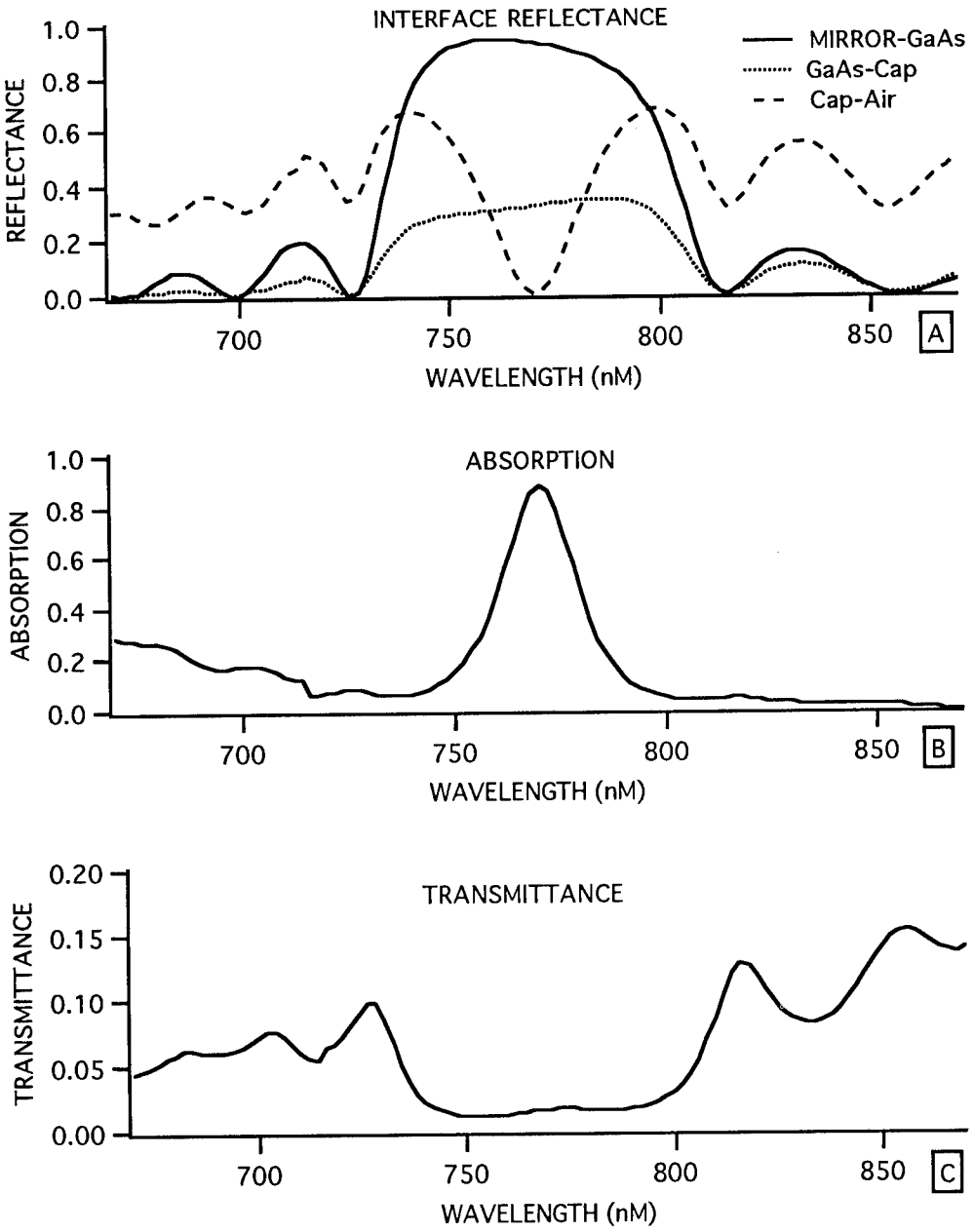


Figure 3. Calculated optical response of the detector material layers for a 770 nm light design: A). reflectance at important interfaces of the structure. B). fraction of incoming light absorbed. C). fraction of incoming light transmitted into the substrate.

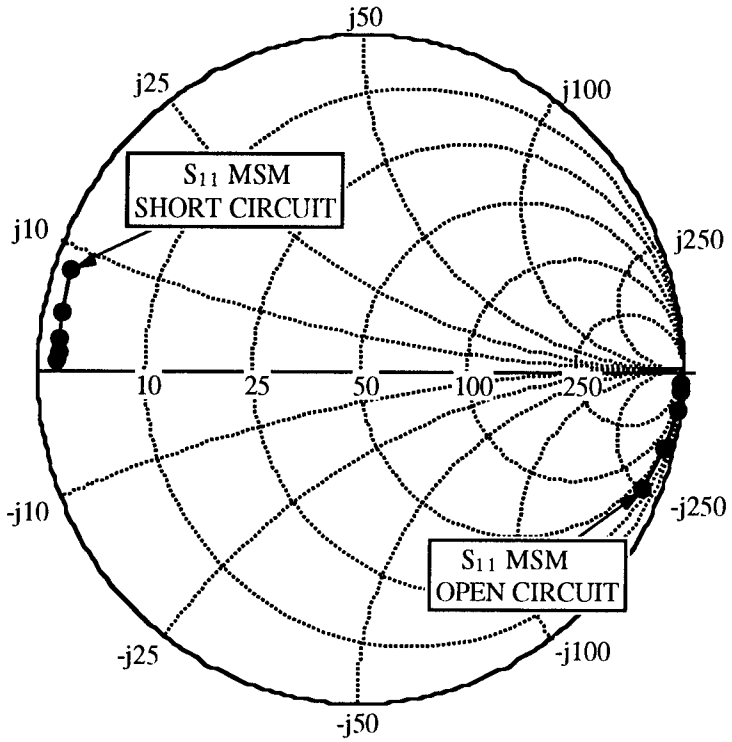


Figure 4. Measured S-parameters, (45 MHz-26.5 GHz), for a MSM detector and its associated short circuit test device. The geometry is linear with 15 fingers, 0.5 μm wide, 0.5 μm gap and 23 μm overlap

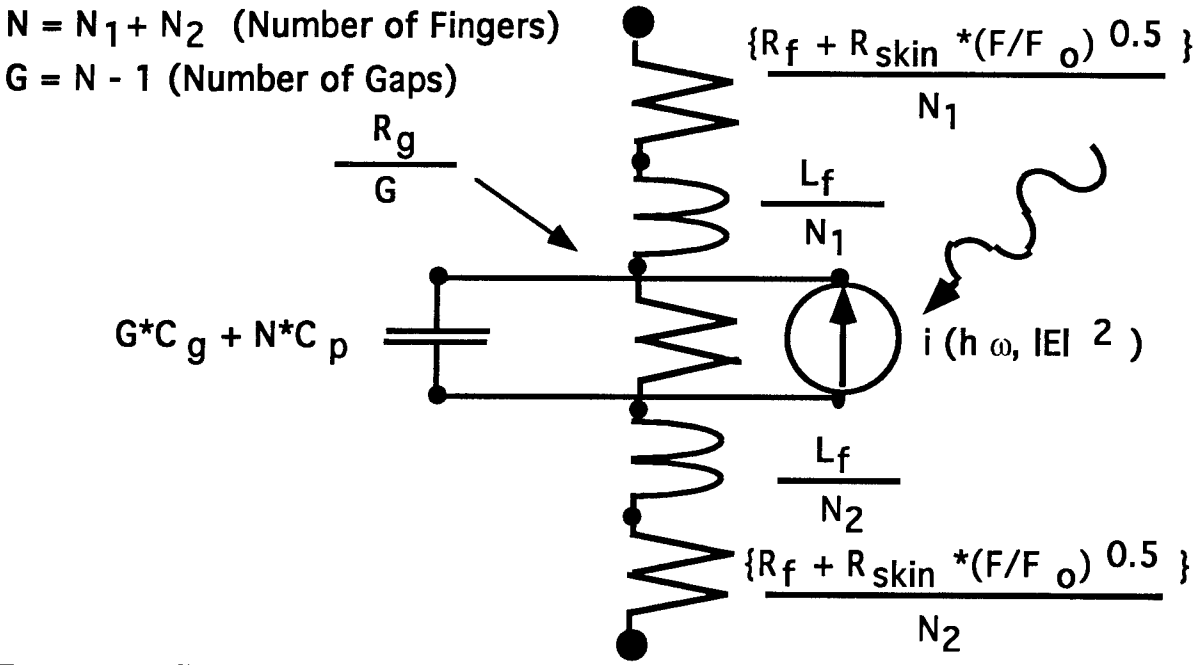


Figure 5. MSM detector diode equivalent circuit model.

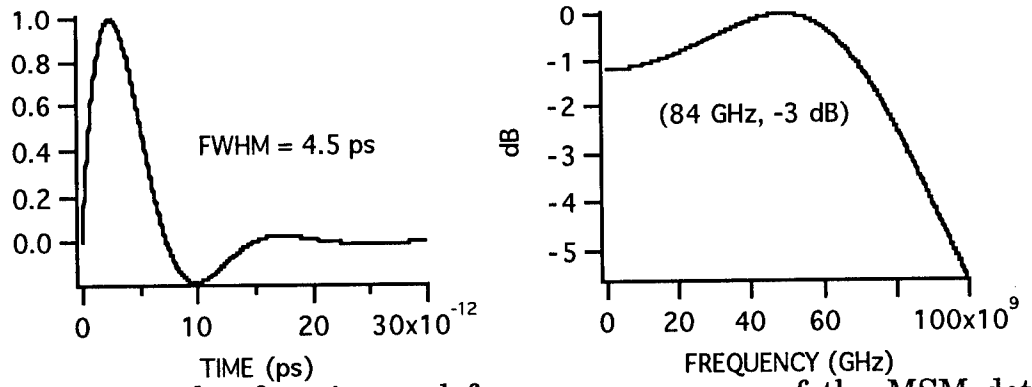


Figure 6. Impulse function and frequency response of the MSM detector passive circuit element model (50 Ω load). For the geometry described above.

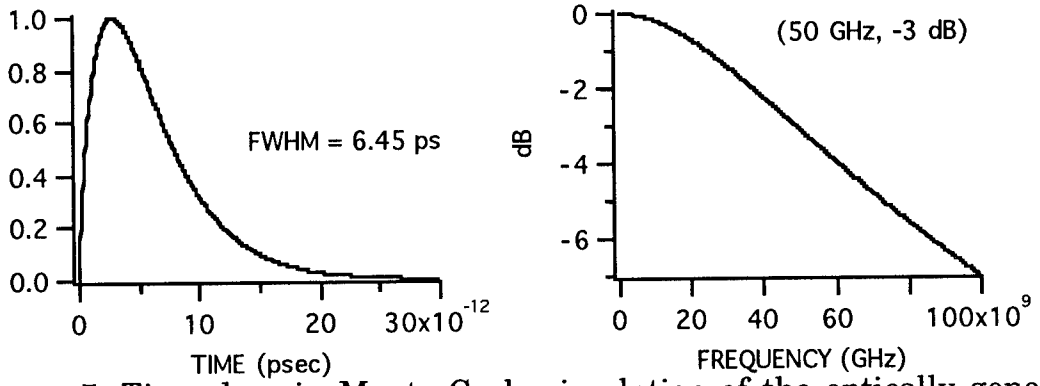


Figure 7. Time domain Monte Carlo simulation of the optically generated carriers in the MSM structure and the FFT of this function.

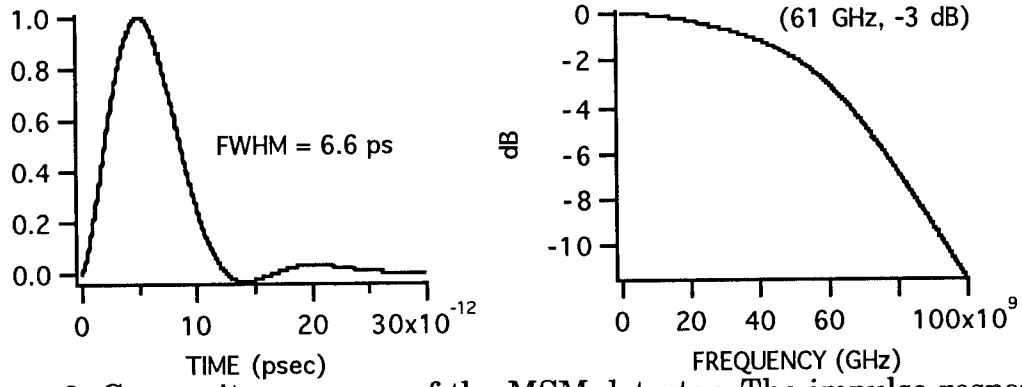


Figure 8. Composite response of the MSM detector. The impulse response is the convolution of the time domain curves of figures 6 and 7. The frequency response is the product (addition of dB) of the frequency responses of figures 6 and 7.

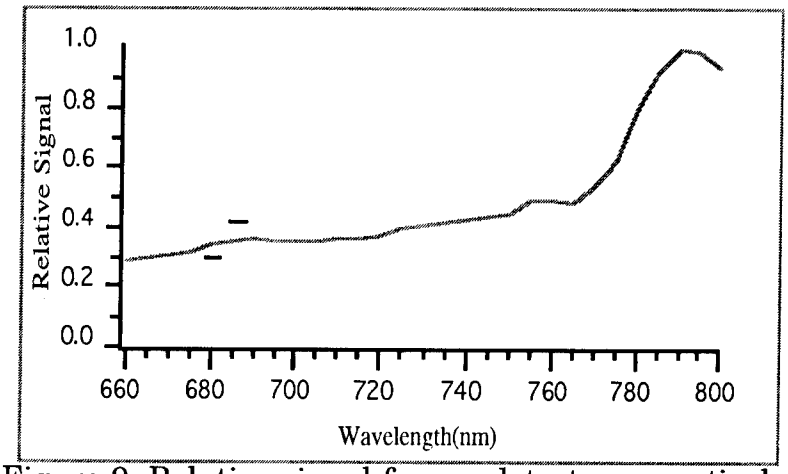


Figure 9. Relative signal from a detector vs. optical wavelength. The detector was fabricated on a wafer with the Bragg reflectors and top surface optical impedance matching optimized for 800 nM light.

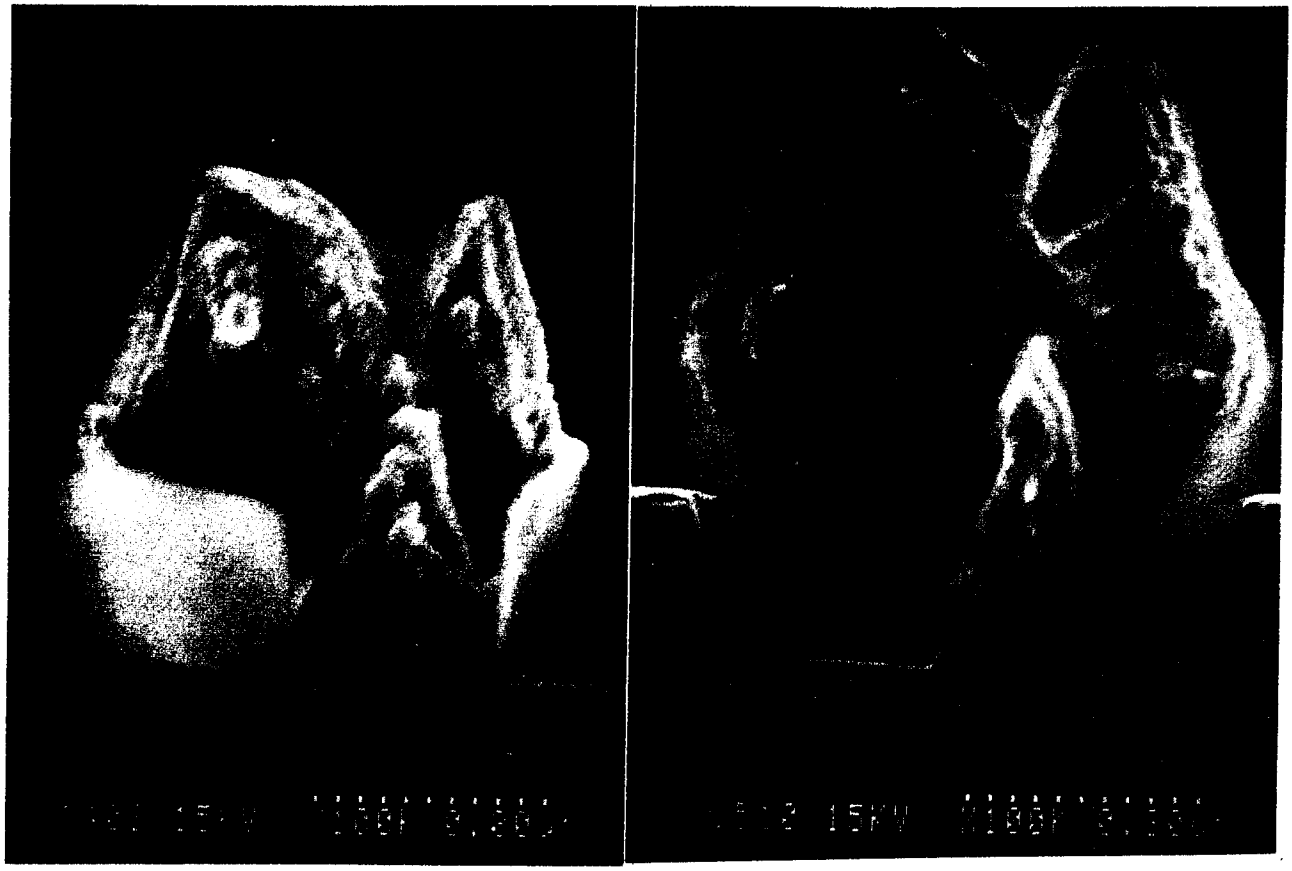


Figure 10. Gamma-gate with 0.25 μm footprint and 2:1 overhang ratio. The photograph also shows the 0.6 μm drain - source spacing after the self-aligned ohmic metalization step.

EXTRACTION OF HIGH-FREQUENCY EQUIVALENT NETWORK PARAMETERS OF HBT's BY LOW-FREQUENCY EXTRAPOLATION OF MICROWAVE S-PARAMETER DATA

Eugene John, Mukunda B. Das, Li-Wu Yang† and S. M. J. Liu†
Department of Electrical Engineering, and
Electronic Materials and Processing Research Laboratory
The Pennsylvania State University
University Park, PA 16802

† Electronics Laboratory
General Electric Company
Syracuse, NY 13221

Abstract

This paper reports a comprehensive low-frequency extrapolation method for the purpose of separating the intrinsic and extrinsic or parasitic equivalent network parameters of high-frequency HBT's. The method involves measurements of device's small-signal parameters over a wide frequency range, and low-frequency extrapolation (LFE) of various transformed two-port parameters using a range of collector bias currents. For the extraction of the equivalent network parameters, the measured s-parameters are first converted into y- and h-parameters and they are used directly or in composite fashions in order to extract the parasitic series resistance elements ($r_{bb'}$, and $r_{ee'}$), the shunt capacitance elements (c_{je}, c_{bc} and c_c), and the intrinsic emitter-to-collector transit time after allowing for the parasitic time constant effects.

Introduction

A precise knowledge of the high-frequency equivalent network parameters of high-speed transistor i.e. heterojunction bipolar transistors, is of considerable interest to circuit designers. To the device engineers, the same information, if interpreted in terms of the device structural dimensions and electronic material parameters, can be even more valuable since they can be used to optimize the device structure for improved performance. Unfortunately, it is difficult to extract the device's intrinsic properties from its externally measured characteristics which are often severely modified, particularly at high frequencies, due to the presence of extrinsic or parasitic elements e.g. shunt capacitances, and series resistance and inductance elements as illustrated in the equivalent network representation of the high frequency HBT (without the inductors) in Figure 1. [1-3]. However, by exploiting the linear dependence of the intrinsic transconductance on the collector bias current of the transistor, and extrapolating the results of transformed s-parameter data presented in suitable fashions, this paper demonstrates that it is possible to consistently separate the intrinsic and extrinsic equivalent network parameters. This approach is different from the approach of de-embedding of parasitic from the overall measured data [4], and that of calculating the T-model parameters from the overall measured microwave s-parameter data over a wide frequency range [5]. In this work the majority of experimental data were obtained from the s-parameter measurements carried out over a wide frequency range i.e. from 50 MHz to 40 GHz. For greater accuracy and cross-checking purposes, data could also be

0-7803-0894-8/93 $3.00

obtained for the transconductance g_m, the output conductance g_o, and the common-emitter current gain β_o by independent measurements at low frequencies (e.g. 1 KHz to 10 KHz).

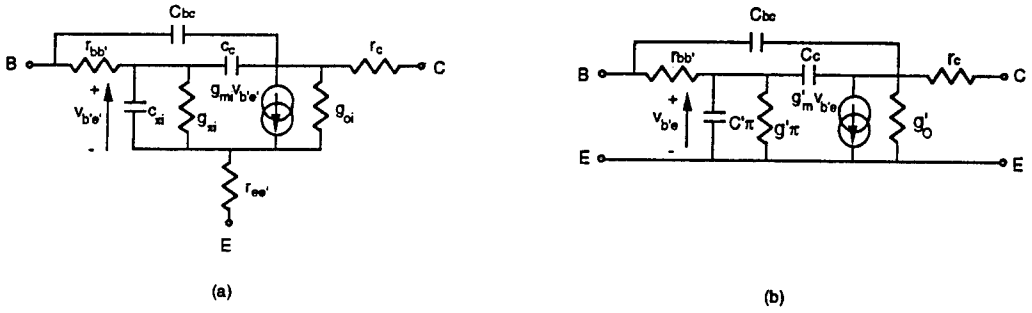


Figure 1 Small-signal high-frequency equivalent network model for the HBT: (a) Physics-based model (b) Modified model when the effects of $r_{ee'}$ is absorbed into other elements.

In the following sections, we first present an outline of the theoretical basis on which the low-frequency extrapolation (LFE) technique has been developed, then present the experimental results that illustrate the methodology of extraction of all the key parameters and finally we discuss the effectiveness of the technique and the usefulness of the extracted intrinsic parameters for understanding the limitations of the material and structure of the HBT.

Theoretical Basis

DETERMINATION OF $r_{ee'}$: By analysing the small-signal equivalent network model as shown in Figure 1, the extrinsic or as measured transconductance (g_m) can be expressed as

$$\frac{1}{g_m} = \frac{1}{g_{mi}} + r_{ee'} + \frac{r_{bb'}}{(1+\beta_o)} \tag{1}$$

where $g_{mi} = qI_c/nkT$, n is the ideality factor and β_o is the small-signal common-emitter current gain at low frequencies. From equation (1) it is obvious that by plotting $\frac{1}{g_m}$ versus I_c^{-1}, it is possible to extract the total resistance $[r_{ee'} + r_{bb'}/(1+\beta_o)]$, and knowing β_o and $r_{bb'}$ one readily obtains $r_{ee'}$ (see Figure 2).

DETERMINATION OF $r_{bb'}$: Again from the equivalent network model, it can be readily seen that the real part of the input impedance, excluding the effect of the extrinsic capacitance c_{bc}, can be expressed as,

$$Re\left(\frac{1}{Y_{11} - j\omega c_{bc}}\right) \approx (r_{bb'} + r_{ee'}) + \frac{r_{\pi'}}{1 + j\omega r_{\pi'}\bar{c}_{\pi'}} \tag{2}$$

where $r_{\pi'} = r_\pi/(1 + g_{mi}r_{ee'})$, and $\bar{c}_{\pi'} = c_{\pi i}/(1 + g_{mi}r_{ee'}) + c_c(1 + g_m r_{cc})$. The second term in equation(2) at high frequencies can be quite small and can be expressed in terms of measurable Y-parameters, namely

$$\frac{r_{\pi'}}{1 + j\omega r_{\pi'}\bar{c}_{\pi'}} = \frac{Re\left(\frac{Y_{21}}{Y_{11}}\right)}{Im\left(\frac{Y_{21}}{Y_{11}}\right)} Im\left(\frac{1}{Y_{11} - j\omega c_{bc}}\right) \tag{3}$$

By plotting the impedances given by equations (2) and (3) as functions of frequency, the difference impedance can be interpreted as $(r_{bb'} + r_{ee'})$ for a wide range of frequencies as can be seen from results presented in Figure 3.

DETERMINATION OF c_{bc} AND $r_{bb'}c_c$: The output capacitance of the HBT when observed with the input short- circuited can be obtained by dividing the imaginary part of Y_{22} by ω. At different operating collector bias currents and the same can be expressed in the following manner [1,6]

$$c_{22} = c_{jc} + c_c r_{bb'} g'_m \tag{4}$$

where $g'_m = g_{mi}/(1 + g_{mi} r_{ee'})$.

It is obvious from equation (4) that when c_{22} is plotted as a function the transconductance g'_m, corresponding to the selected collector bias currents, the capacitance $c_{jc} = (c_{bc} + c_c)$ and the product $r_{bb'}c_c$ can be readily extracted as illustrated in Figure 4.

DETERMINATION OF THE INTRINSIC TRANSIT TIME, τ_{ec}: Perhaps the most difficult parameter to extract, using the measured s-parameter data or the transformed Y-parameter data, is the intrinsic transit time (τ_{eci}) which arises due to the carrier transport in the transistor base region and in the collector depletion region, namely [2]

$$\tau_{eci} = \frac{W_B^2}{2D_n} + \frac{W_B + (X_{dc}/2)}{v_{sat}} \tag{5}$$

where W_B is the basewidth, X_{dc} is the collector depletion-width and D_n is the electron diffusion constant in the base, and v_{sat} is the high field carrier saturation velocity. This parameter can be best obtained by utilizing a wide range of extrapolated common-emitter unity current gain frequency (f_τ) at different collector bias currents. The common practice is to plot the current gain magnitude $|h_{21}|$ versus frequency and exploit its 6dB/octave fall-off behavior at low-frequencies to extrapolate f_τ [5,6] Unfortunately, this approach does not work when the low-frequency value of h_{21} or β_o is low. For β_o values between 10 and 20, it is more appropriate to obtain $\omega_\tau (= 2\pi f_\tau)$ from the ratio of the input capacitance (c_{11}) and the transconductance (g'_m) at low-frequencies. The capacitance c_{11} can be determined from the frequency dependent behavior of $\mathrm{Im}(Y_{11})$ at various operating collector bias currents as illustrated in Figure 5. From a careful analysis of equivalent network model, it can be readily shown that

$$\frac{1}{\omega_\tau} = \frac{c_{11}}{g'_m} = \frac{(c_{je} + c_{jc})}{g_{mi}} + \tau_{eci} + c_{jc}(r_{ee'} + r_c) \tag{6}$$

where $1/g'_m = 1/g_{mi} + r_{ee'}$ and $c_{jc} = c_{bc} + c_c$.

As illustrated later in Figure 6, it can be seen that by plotting $1/\omega_\tau$ versus $1/g_{mi}$ one can separate $(c_{jc} + c_{je})$ and the total time constant, $\tau_{eci} + c_{jc}(r_{ee'} + r_c)$, where $c_{jc} = (c_{bc} + c_c)$. In order to be able to separate $c_{jc}(r_{ee'} + r_c)$ and τ_{eci} and also for cross-checking purposes the parameter τ_{eci} can also be obtained exploiting the frequency dependent behavior of $|Y_{21}|$, and the following expression:

$$\frac{1}{g'_m r_{bb'} \omega_{21}} = \frac{c_c + c_{je}}{g_{mi}} + c_c(r_{ee'} + r_{cc'}) + \tau_{eci} \tag{7}$$

In equation (7) ω_{21} is the 3-dB cut-off frequency of $|Y_{21}|$, and it can be obtained by low-frequency extrapolation of the measured quantity $[|g_m/Y'_{21}|^2 - 1]$ versus frequency as illustrated later in Figure 7, since

$$\left|\frac{g_m}{Y'_{21}}\right|^2 - 1 = \left|\frac{\omega}{\omega_{21}}\right|^2 \tag{8}$$

where $Y'_{21} \approx Y_{21} + j\omega c_{bc}$.

As illustrated in Figure 6, it can be seen that by plotting the composite parameter $(g'_m r_{bb'} \omega_{21})^{-1}$ in equation (7) versus $1/g_{mi}$, it is possible to extract the total time-constant $c_c(r_{ee'} + r_c) + \tau_{eci}$. The difference between this and the previously obtained time constant can be used to determine $c_{bc}(r_{ee'} + r_c)$, and since c_{bc} is already known from $(c_{jc} - c_c)$, the sum of the resistances $(r_{ee'} + r_c)$ follows. Hence, follows the intrinsic transit time τ_{eci}, provided the latter remains constant while the operating collector current is varied without exceeding the value where τ_{eci} tends to attain its peak value.

DETERMINATION OF v_{sat}: The dependence of τ_{eci} on the collector depletion width (X_dc) as given in equation (5) can be exploited in order to obtain the carrier saturation velocity, v_{sat}. For this purpose, the equivalent network parameters need to be extracted for different collector bias voltages by repeating the entire process of parameter extrapolation as outlined above. A knowledge of the capacitance c_c at different collector bias voltages and that of the emitter area is sufficient to obtain X_{dc}. Once v_{sat} becomes known, a knowledge of the basewidth (W_b) enables one to determine D_n or vice versa.

Experimental Results

The LFE method of extracting the equivalent network parameters of the HBT's as outlined above has been implemented on several $n - p - n$ AlGaAs/GaAs devices using the HP 8510 network analyzer for the frequency range of 0.5 to 40 GHz. The on-the-wafer measurements of s-parameters were performed using the co-planar waveguide technique using collector bias current ranging from 2 mA to 30 mA at $V_{CE} = 2Volts$. Only representative data obtained from a single device will be presented here.

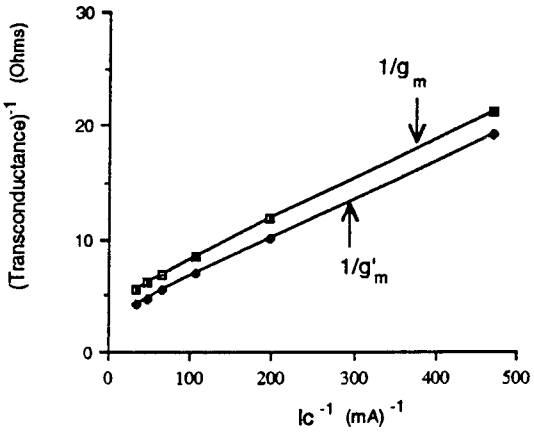


Figure 2 Reciprocal transconductance g_m^{-1} versus the reciprocal collector bias current plot: determination of $r_{ee'}$.

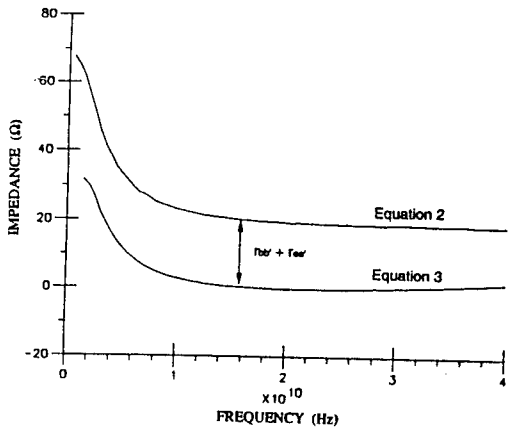


Figure 3 Frequency dependent behaviors of $Re(Y_{11} - j\omega c_{bc})^{-1}$ and $r'_\pi/(1+j\omega c_\pi r_\pi)$: determination of $(r_{ee'} + r_{bb'})$

The details of experimental determination of the emitter series resistance can be best understood from the results presented in Figure 2. Because of a rather low value for β_o in the particular device tested, it was necessary to account for the contribution of the resistance $r_{bb'}/(1 + \beta_o)$, and this was obtained iteratively between the results presented in Figure 2 and Figure 3 with a knowledge of β_o at different collector currents. After subtracting $r_{bb'}/(1 + \beta_o)$ from $1/g_m$ a plot of $1/g'_m$ was obtained which is also shown in Figure 2. From the slope of the latter plot, the ideality factor n is seen to change from 1.2 at lower currents to 1.5

at higher currents. These values are larger than that determined from the collector current versus the base emitter voltage plot (Gummel plot) of the device at lower currents ($n \approx 1.15$). The results also indicate that due to the rise of junction temperature at higher collector currents, the diode differential resistance may be increasing while the emitter contact resistance is decreasing as may be expected.

The experimental data presented in Figure 3 clearly demonstrate that the real part of $r_{\pi'}/(1+j\omega r_\pi \tilde{C}_\pi)$ systematically approaches zero value at high frequencies revealing the frequency independent series resistance $(r_{ee'}+r_{bb'})$. The difference between the two plots appears to remain essentially constant over a wide frequency range indicating the validity of the method of extracting $(r_{ee'}+r_{bb'})$and that of the equivalent circuit model used.

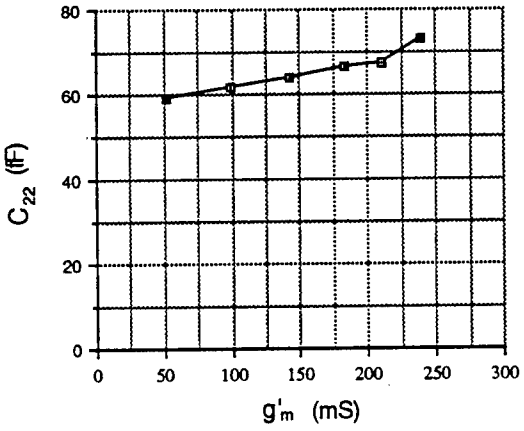


Figure 4 Dependence of the output capacitance (c_{22}) on the transconductance g'_m: determination of c_{jc} and $r_{bb'}c_c$ product.

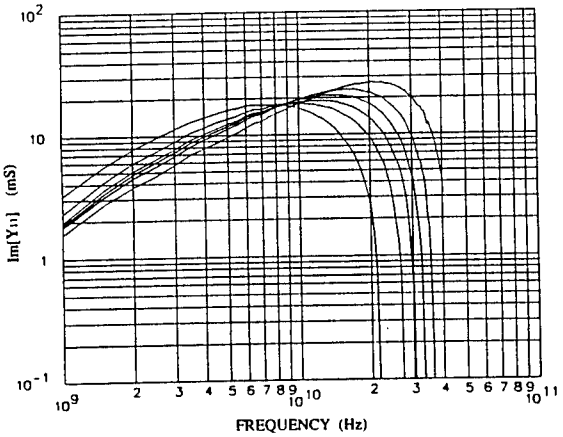


Figure 5 Frequency dependence of $Im(Y_{11})$: determination of the input capacitance by low-frequency extrapolation

The collector current dependence of the output capacitance c_{22}, obtained from the low frequency extrapolation of $Im(Y_{22})$ are presented in Figure 4, versus the transconductance g'_m, does indicate a good degree of linearity for the middle range bias current values. This allowed us to extract the total collector base junction capacitance (c_{jc}) and the $r_{bb'}c_c$ product. This yielded $c_{jc} \approx 59fF$ and $c_c \approx 3fF$ when $r_{bb'} \approx 17\Omega$. We found $r_{bb'}$ varying between 19Ω and 15.5Ω when I_c changed from 2 mA to 30 mA. From a knowledge of the collector-base depletion layer width $(X_{dc} \approx 0.55\mu m)$ at $V_{CB} \approx 0.7V$, we calculate a total collector-base capacitance of 28 fF. This yields an effective value for $c_{bc} \approx 25$ fF. This suggests that there is an excess collector-to-base parasitic capacitance of 34 fF which could have arisen due to the wafer probing of the device where the emitter is coupled to the co-planar wave-guide in a manner which is not physically grounded. This is in contrast to the situation when the device is mounted on a microstrip carrier. Similar additional extrinsic parasitic capacitance could also be presented between the base and the emitter electrodes.

For the determination of ω_τ we were unable to exploit the expected 6dB/octave frequency dependency of $|h_{21}|$, because of the low value of β_o and a large value of c_{jc}. It is not difficult to see that the former lowers the slope of $|h_{21}|$ versus frequency below the 6dB/octave limit at lower frequencies, while the latter lowers the same at higher frequencies. However, we obtained $\omega_\tau = c_{11}/g'_m$ by extracting the total input capacitance from the expected linear dependence of $Im(Y_{11})$ at low frequencies (see Figure 5), and using the values of g'_m already obtained (see Figure 2). The results are graphically presented in Figure 6 indicating some degradation of the intrinsic transit time above the collector current density of 2×10^4 A/cm^2 (where

$N_{DC} \approx 10^{16}\ cm^{-3}$) which corresponds to approximately 8mA for the test device. However, with the limited linearity of the curve we obtain $\tau_{eci} \approx 2.75\ p\ sec$, after allowing for $c_{jc}(r_{ee'} + r_c)$. From the slope of this curve we obtain $c_{je} \approx 154\ fF$ and this most likely consists of an excess parasitic capacitance of some unknown magnitude due to the co-planar waveguide measurements. Unfortunately, c_{je} could not be precisely determined since we do not know the built-in potential involving a graded $Al_xGa_{1-x}As$ emitter.

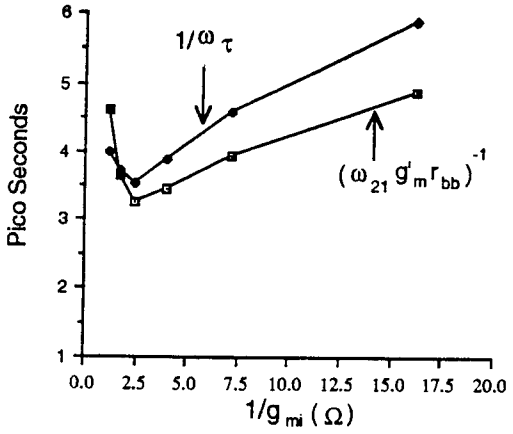


Figure 6 Dependence of $1/\omega_\tau$ and $(g'_m r_{bb'}\omega_{21})^{-1}$ on the reciprocal intrinsic transconductance $(g_{mi})^{-1}$: determination of τ_{eci} and parasitic resistance and capacitance elements.

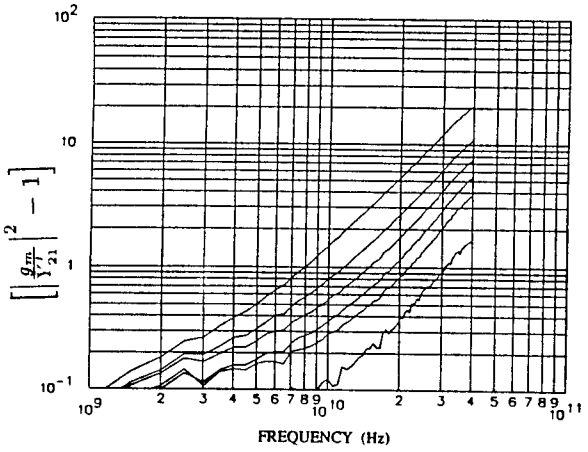


Figure 7 Frequency dependent behavior of $\left[\left|\frac{g_m}{Y'_{21}}\right|^2 - 1\right]$: determination of ω_{21} by low-frequency extrapolation approach ($Y'_{21} \approx Y_{21} + j\omega c_{bc}$).

We present the data showing dependence of $[|g_m/Y'_{21}|^2 - 1]$ in Figure 7 in order to determine ω_{21}, which clearly indicates ω^2 dependence when extrapolated. The composite time constant $(g'_m r_{bb'}\omega_{21})^{-1}$ when plotted versus $1/g_{mi}$ (see Figure 6), we observe behavior similar to that obtained for the $1/\omega_\tau$ versus frequency plot, but with a slightly lower intercept and a lower slope. This is what one would expect by comparing equations (6) and (7), due to the difference between c_{jc} and c_c. With the limited data points we obtain the capacitance difference to be $\approx 69\ fF$ which is slightly higher than the values obtained from data presented in Figure 3. However, the two plots in Figure 6 indicate reasonable agreements in so far as the extraction of τ_{eci} is concerned and they provide a basis for a high degree of confidence in the proposed LFE method of extraction of the HBT equivalent circuit parameters. Finally, we present the extracted equivalent circuit parameter values in Table I.

Table I

Extracted Equivalent Circuit Parameter Values

Parameter	Extracted value
c_{jc}	59fF
c_c	3 fF
c_{bc}	25fF
$r_{bb'}$	15 - 17 Ω
c_{je}	$\approx 154fF$
τ_{eci}	2.75 psec
$r_{ee'}$	3 Ω
r_c	1 Ω
β_o	≤ 10
η	1.2 – 1.5

Conclusion

A comprehensive new low-frequency extrapolation (LFE) technique has been developed and implemented for the extraction of small-signal equivalent network parameters of HBT's. The technique utilizes the microwave s-parameter data over a wide frequency range with varying collector bias currents. For the parameter extraction purposes the s-parameters are transformed into Y-parameters which readily relate to the physics-based equivalent network model of the HBT. New approaches for obtaining the unity current gain frequency (ω_τ), and the 3dB cut-off frequency (ω_{21}) of $|Y_{21}|$ are introduced which provide the basis for extracting the intrinsic transit time parameter (τ_{eci}) with built-in cross checking capability. A composite parameter versus frequency approach has been introduced for the extraction of resistance ($r_{bb'} + r_{ee'}$), and visualization of its frequency independence. It is demonstrated that the LFE technique is quite powerful in extracting all of the equivalent network parameters. It is believed that the proposed LFE method would provide the basis for a better understanding of the behavior of the intrinsic and extrinsic parameters of emerging high speed HBT's with f_τ-values extending beyond 100 GHz.

Acknowledgement

This work has been supported by the National Science Foundation under its Material Synthesis and Processing Program, under grant No. ECS-9202642

References

[1.] R. L. Pritchard, "*Electrical Characteristics of Transistors*", McGraw Hill, pp. 298-413, 1967.

[2.] M. B. Das, " High-Frequency Performance Limitations of Millimeter-Wave Heterojunction Bipolar Transistors", *IEEE Trans. Electron Devices*, vol. 35, pp. 604-614, 1988.

[3.] M. B. Das, "HBT device physics and models" Ch.4, pp., in :*HEMTs and HBTs : Devices, Fabrication, and circuits*, editors: F. Ali and A. Gupta, Artech House, pp. 191-251, 1991.

[4.] D. Costa, W. U. Liu and J. S. Harris Jr., " Direct Extraction of the AlGaAs/GaAs Heterojunction Bipolar Transistor Small-Signal Equivalent Circuit", *IEEE Trans. Electron Devices*, vol. 38, pp. 2019-2024, 1991.

[5.] D. R. Pehlke and D. Pavlidis, " Evaluation of the Factors Determining HBT High-Frequency Performance by Direct Analysis of S-Parameter data", *IEEE Trans. Microwave Theory and Techniques*, vol. 40, pp. 2367-2373, 1992.

[6.] M. B. Das "On the Determination of the Extrinsic Equivalent Parameters of the Diffusion Transistor" *J. Electr. Control*, vol. 8, pp. 351-363, 1960.

HIGH FREQUENCY PERFORMANCE OF GaAs/AlGaAs MULTIPLE QUANTUM WELL (MQW) ASYMMETRIC FABRY PEROT (ASFP) REFLECTION MODULATOR

M. NAWAZ, B. T. OLSEN*, K. McILVANEY*
UNIK Centre for Technology, P.O.Box 70, 2007, Kjeller, Norway
* Norwegian Telecom Research P.O. Box 83, 2007, Kjeller, Norway

we report the high frequency performance of GaAs/AlGaAs multiple quantum well (MQW) based Asymmetric Fabry Perot (ASFP) reflection modulator. The measured frequency response bandwidth of the modulator was ~ 600 MHz at an applied voltage of only 5 V at a wavelength of ~ 854 nm.

The Asymmetric Fabry Perot (ASFP) reflection modulators based on quantum confined stark effect [1,2] (QCSE) with GaAs/AlGaAs multiple quantum well (MQW) system are considered promising candidates in the area of photonic switching, optical information processing and an interface device in an optical interconnect for high speed electronics [3,4]. Much work has been performed to study the optical contrast, reflectivity change, insertion loss, but little or no attention has been drawn for the high speed performance of ASFP modulators. In this letter, we present the results of the frequency response of the ASFP modulator.

The modulator layer structure was grown by molecular beam epitaxy (MBE). It consists of 1 um n doped buffer layer grown on the n doped GaAs substrate, followed by a dielectric mirror design consisting of 20 periods of 2.83 Å n type AlAs and 25.44 Å n type GaAs. Over this are 75 undoped multiple quantum wells (MQW) consisting of 96.11 Å GaAs well between 59.38 Å AlGaAs barriers. On top of this, is 250 Å undoped AlGaAs buffer layer, followed by p doped 4556 Å AlGaAs layer, and finally there are two cap layers of 10 Å and 40 Å p doped GaAs to complete the PIN diode structure. All AlGaAs layers in the structure contain 30% Al content. For n type doping, Si was used with doping concentration ~ 1 x 10^{18} cm^{-3}, while Be was used for p type doping with doping concentration ~ 1 x 10^{18} cm^{-3}. All AlGaAs layers in the structure contain 30% Al content. The buffer layer was used to separate the intrinsic region from any background doping diffusion. The semiconductor - air interface forms the front mirror, whose reflectivity is 0.31, while dielectric stacks grown on the n doped GaAs substrate forms the back mirror, whose reflectivity is 0.96. The structure is like an asymmetric Fabry Perot cavity, because of unequal front and bottom mirror reflectivities. The device was prepared by etching mesas of 220 um circular diameter, connected to a contact pad of 100 um square. The calculated capacitance of the device was 4.8 pF, for 1.16 um thick MQW region with an area of 4.799 x 10^4 um^2 of the device.

The device was connected to the microstrip transmission line through the Au bonded wire (5 mm long). The reverse breakdown voltage of the device was 30V. The white light source was used to study the reflection spectrum of the modulator. The light was coupled through 2 x 2 fiber coupler, and perpendicular incident onto the sample surface. The reverse dc bias (from 0 to -16 V) was applied perpendicular to the quantum well layers. To determine the driving voltage and working wavelength for the speed response measurements, results of the differential contrast ($\frac{\Delta C}{\Delta V}$) defined as the contrast difference between two voltages divided by the corresponding voltage difference are presented in figure 1. The differential contrast peaks increases with the increase of reverse bias $V = (V_1 + V_2)/2$ in the wavelength range from 852 to 862 nm (which is the most sensitive region), reaching a maximum value of ~4.0 at 12 and 13V (i.e, cr(13, 12) in fig. 1) and then decreases at higher voltages (i.e., cr(14,10) in fig. 1) because the reflectivity increases (due to decreased absorption) once reaches a minimum value due to the QCSE[1,2]. It was noted that lower voltage difference (i.e, $\Delta V = V_2 - V_1$) resulted higher sensitivity of the device.

An AlGaAs laser diode was used for the frequency response measurements. The incident light was 11 mW and set at a wavelength of 854 nm. The reflected light was detected by a PIN diode receiver (HP83401A). The bias dependent modulation response at different modulation frequencies at a wavelength of 854 nm is shown in figure 2. The voltage sensitivity of the device was maximum at an applied voltage of 5V at 854 nm. The modulator was

0-7803-0894-8/93 $3.00 © 1993 IEEE

driven by an RF signal from the swept network analyser (HP 8702A). The RF signal was 10 dB and a reverse bias of 5V was superimposed through a bias Tee. For the integration of MQW modulators with VLSI circuits as optical interconnects a significantly lower operating voltage is desirable (i.e, 5V or less). The frequency responses of the receiver, the cables, and the bias T were separately calibrated to enhance the accuracy of measurements. To determine the suitability of the device for the speed response measurements, the return loss of the device was first measured, without illumination of the light and plotted in Fig. 3. The frequency scan was from 300 KHz to 3 GHz. The return loss of the device is insensitive (as expected) of the frequency. However, a small resonance of - 5 dB at 450 MHz, was observed in the return loss due to the mismatched impedance of the device and transmission line. The frequency response of the modulator at an applied voltage of 5V at 854 nm is shown in figure 4. The frequency response of the modulator is almost flat upto 550 MHz and drops rapidly at higher frequencies. The measured 3 dB frequency response bandwidth of the modulator was 600 MHz. If the modulator is modelled as a simple RC series circuit in which the response is proportional to the voltage across the capacitor, then the 3 dB frequency cutoff ($f_c = 1/2\pi RC$= 650 MHz for 50 ohm) of the modulator is in excellent agreement with the measured frequency response. The measured frequency response is limited by the circuit consideration such as the capacitance of the device, which can be made very small by minimizing the device area, increasing the thickness of the intrinsic region and using a relatively thick polymide layer under the contact pad [5]. Using these considerations, the speed of the device can be increased further, which is the object of our future research.

In summary, we have demonstrated the fabrication, differential contrast, and frequency response of the ASFP modulator with GaAs/AlGaAs MQW system. The maximum differential contrast of the modulator was 4.0 at an applied voltage of 12.5V at 862 nm. The frequency response of the modulator was 600 MHz at an applied voltage of only 5V at 854 nm. The results are encouraging for the subsequent development of ASFP modulators for their high speed of response.

REFERENCES

1. R. H. Yan, R. J. Simes and L. A. Coldren, Appl. Phys. Lett. 55 (19), (1989).
2. R. H. Yan, R. J. Simes and L. A. Coldren, IEEE J. Quantum Electron. Vol. 27, No. 7, (1991).
3. T. H. Wood, J. Lightwave Technol. 6, 743, (1988).
4. R. J. Simes, R. H. Yan, R. S. Geels, L. A. Coldren, J. H. English, A. C. Gossard and G. Parry, Electron. Lett. 53, p. 637-639, (1988).
15. O. Mitomi, I. Kotaka, S. Nojima, K. Kawano, Y. Kawamura and H. Asai, Appl. Optics. Vol. 31, No. 12, (1992).

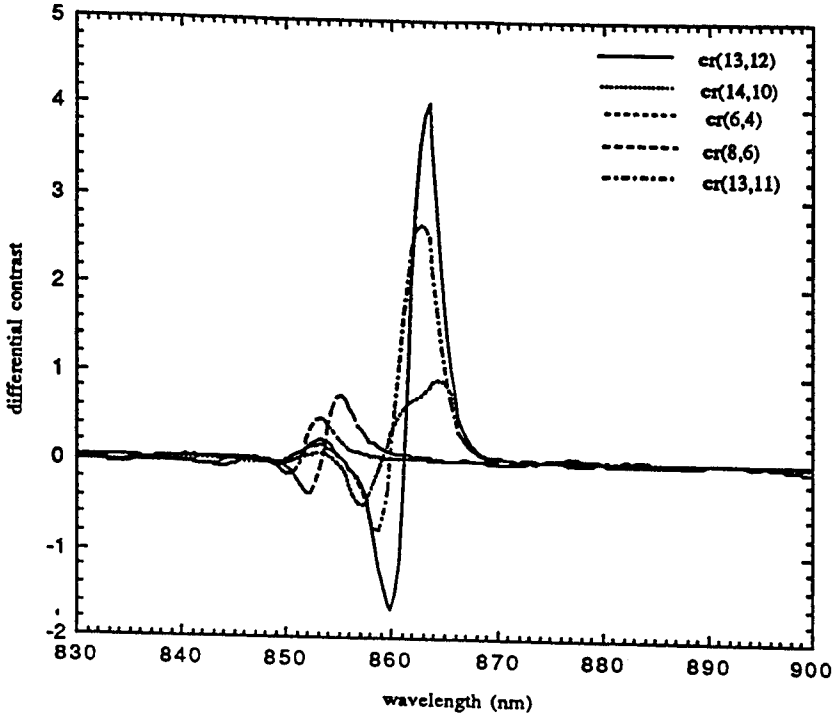


Fig. 1. differential contrast of the modulator as a function of wavelength at various voltages.

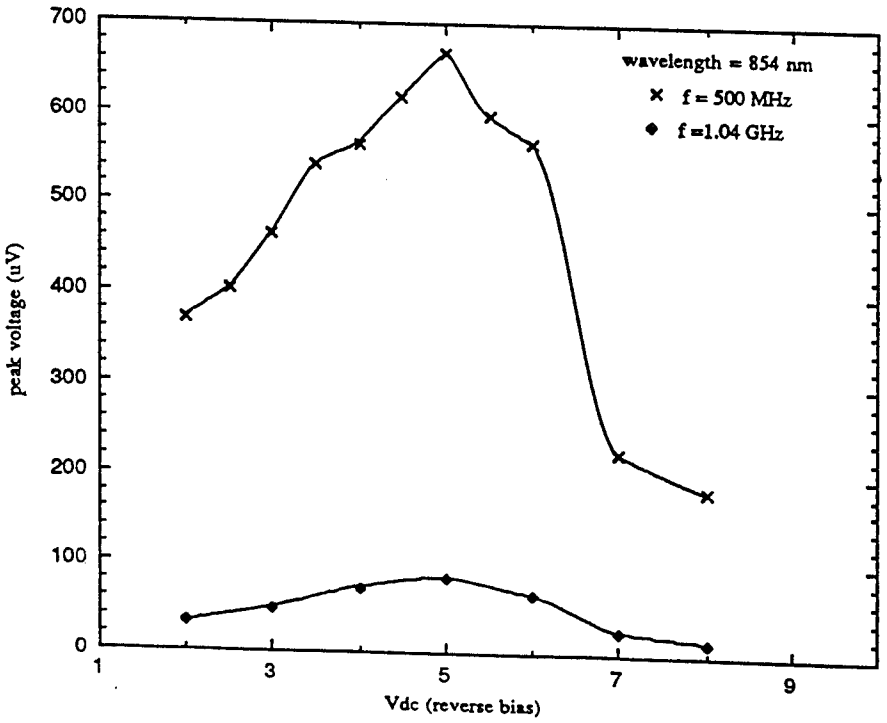


Fig. 2. voltage sensitivity of the device at a wavelength of 854 nm for different modulation frequencies.

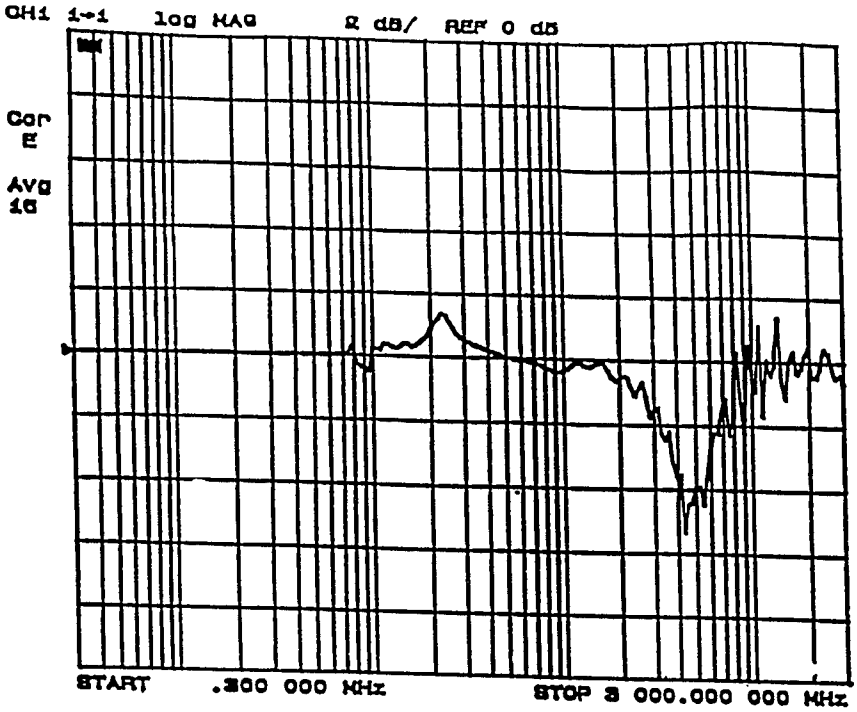


Fig. 3. return loss measurement of the device.

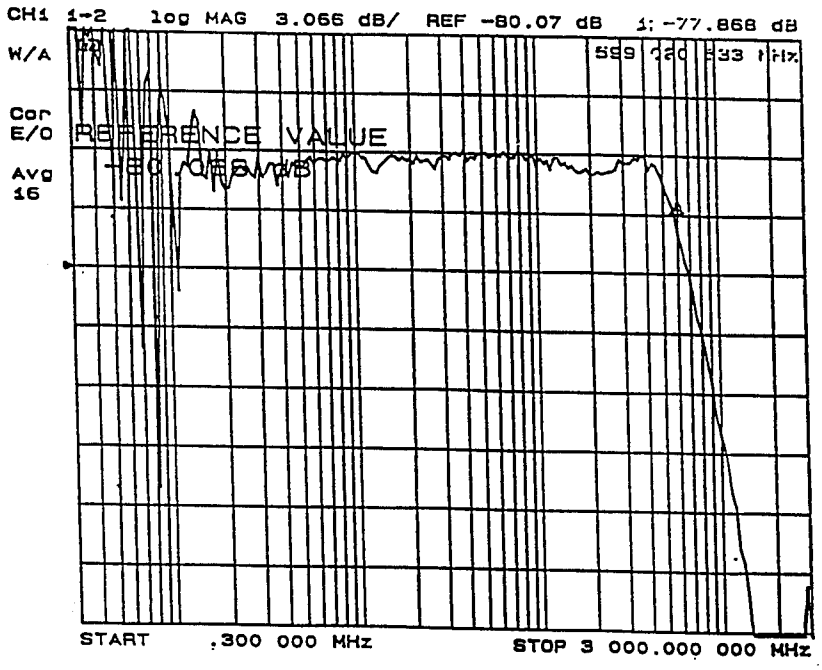


Fig. 4. swept frequency response of the modulator for an applied voltage of 5V at a wavelength of 854 nm.

A Hybrid Quantum-Classical Model for Transport in Tunneling Heterostructures

Chenjing L. Fernando, William R. Frensley
Program in Electrical Engineering
Erik Jonsson School of Engineering and Computer Science
The University of Texas at Dallas
Richardson, TX 75083

ABSTRACT

Several models have been developed to evaluate the current-voltage characteristics of the resonant tunneling diode (RTD) . The current density predicted by both flat-band model and Thomas-Fermi approximation (or zero-current model) fails to reproduce the experimental results. We believe that this disagreement is due to the assumption of perfect electron coherence in the tunneling theory. An adjusted Thomas-Fermi model which includes a series resistance at either of the contact layers has been tried out, but the result is still unsatisfactory. A hybrid quantum-classical approach is suggested to evaluate self-consistently the electron current. The method couples the quantum-tunneling current obtained by solving the Schrödinger equation in the tunneling region with the classical drift-diffusion current in the contact layers. The resulting current continuity equation is solved self-consistently with Poisson's equation. It shows that the hybrid quantum-classical model gives much more realistic $I(V)$ curves than other models.

1 Introduction

The resonant tunneling diode (RTD) is a simple tunneling heterostructure device which has shown promising high frequency performance [1]. The dc $I(V)$ characteristic of the RTD is among its most fundamental properties, and the details of the $I(V)$ curve are quite sensitive to the device design. Two basic methods (the Schrödinger equation and the Wigner function) have been used to evaluate the $I(V)$ curve which has proved to be very difficult to predict theoretically [2]-[7]. The ultimate solution to this problem would be a complete quantum transport theory, but such a theory is not yet available. To assist in the design of tunneling diodes and transistors, it is desirable to develop a simple but efective semiquantum model. Several approaches based on the tunneling theory have been developed to evaluate the electron current of such devices.

The current density according to the tunneling theory is expressed by integrating over the

0-7803-0894-8/93 $3.00

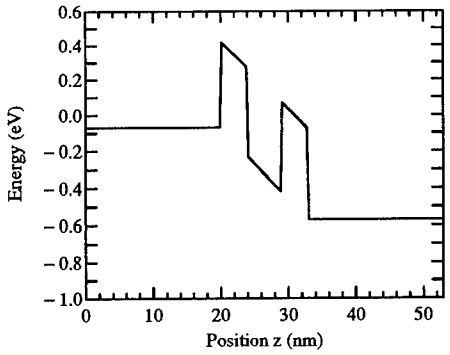


Figure 1: Flat-Band model.

transverse energies as [8]:

$$j = \frac{emkT}{2\pi^2\hbar^3}\int_{E_0}^{\infty}\frac{dE}{2\pi\hbar}Tln\left(\frac{1+exp[(E-E_{fl})/kT]}{1+exp[(E-E_{fr})/kT]}\right) \tag{1}$$

where T is the transmission probability in terms of the probability current as a function of the longitudinal energy E and the bias voltage. E_{fl} and E_{fr} are the Fermi levels at the left and right boundaries, respectively. The transmission probability T depends on the potential distribution across the device. Several models of potential distribution are to be discussed in the following sections.

2 Theoretic Models

The simplest model is the flat-band model. The applied bias voltage is assumed to be dropped linearly across the well and the barriers (quantum region) as shown in Fig. 1.

The second model is the Thomas-Fermi approximation (or zero-current model) which is based on the self-consistent solution of the electrostatic potential and the carrier distribution at the two boundary regions. The carrier densities are evaluated within a local quasi-equilibrium approximation. The equations are Poisson's:

$$-\nabla\cdot[\epsilon(x)\nabla\phi(x)] = \rho(x) \tag{2}$$

and the local approximation for the net charge density ρ :

$$\rho = e(\sum_{v} N_v\mathcal{F}_{1/2}[\beta(E_v - e\phi) - E_f)] - \sum_{c} N_c\mathcal{F}_{1/2}\beta[E_f - (E_c - e\phi)] + \sum_{D} N_D^{+}(\phi) - \sum_{A} N_A^{-}(\phi)) \tag{3}$$

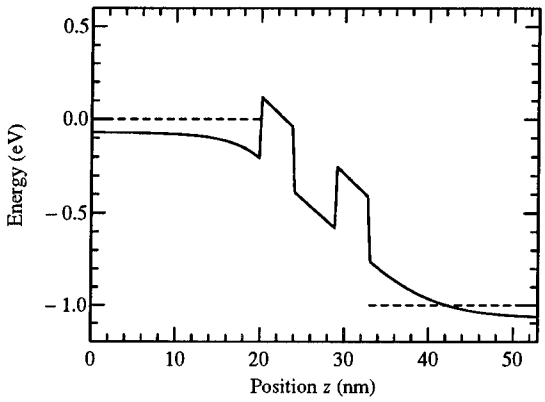


Figure 2: Thomas-Fermi model.

here ϕ is the electrostatic potential, E_f is the local Fermi energy, and $\mathcal{F}$ is the Fermi integral of order 1/2. E_v and E_c are the band-edge energies which contain the heterojunction band discontinuities, and N_v and N_c are the effective densities of states. N_D^+ and N_A^- are the ionized donor and acceptor densities, respectively. Within the quantum region, it remains linear potential distribution as shown in Fig. 2.

Unfortunately, the current-voltage curve predicted by the above tunneling calculations fails to reproduce the experimental results (as we will see some examples in the next section). This disagreement, we believe, is because the perfect electron coherence over the whole range of the device is assumed and all the dissipative processes are neglected in the calculations.

An adjusted Thomas-Fermi model has been carried out to take into account the imperfect electron coherence (or limited range coherence), as shown in Fig. 3. A series resistance r is attached to one of the boundary regions. We assume that the quantum region is $[x_l, x_r]$ which is considered the range of the electron coherence, instead of the whole range of the device as in the original Thomas-Fermi model. The boundary positions x_l and x_r can be adjusted, so can the series resistance r. By adjusting x_l,x_r and r, one can expect better result than the other two models, but the disageement is still significant, especially in the case of thick barriers.

A hybrid quantum-classical model couples the quantum tunneling current obtained by (1) in the quantum region (the range of the quantum region can be adjusted) with the classical drift-diffusion current in the two boundary regions. The resulting discretized current equation can be derived as follows.

For the nondegenerate case, the tunneling current within the quantum region $[x_l, x_r]$ can

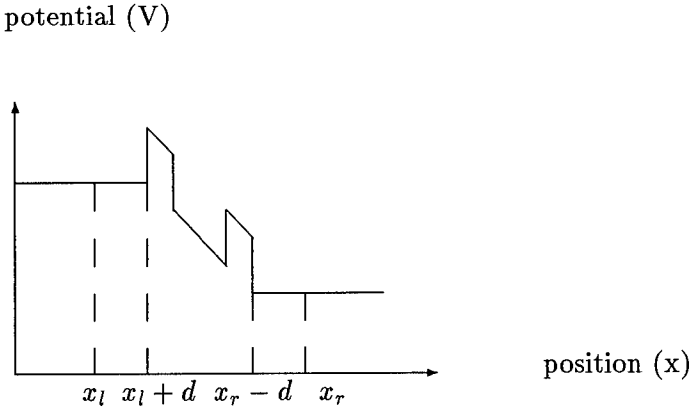


Figure 3: Illustration of the parameters used in the adjusted Thomas-Fermi model and the Hybrid Quantum-Classical model.

be written as (only consider electron current) :

$$j_t \approx a_1 n_l - a_2 n_r \tag{4}$$

where n_l and n_r are the electron densities at x_l and x_r respectively. and

$$a_1 = \frac{emkT}{2\pi^2\hbar^3} \int_{V_0}^{\infty} \frac{dE}{2\pi\hbar} \frac{T}{N_{cl}} e^{\beta(V_l - E)} \tag{5}$$

$$a_2 = \frac{emkT}{2\pi^2\hbar^3} \int_{V_0}^{\infty} \frac{dE}{2\pi\hbar} \frac{T}{N_{cr}} e^{\beta(V_r - E)} \tag{6}$$

For the mesh points at the two boundaries, the drift-diffusion current between x_i and x_{i+1} is (δ_i is the mesh spacing between x_i and x_{i+1}):

$$j_{i+1/2} = (-\alpha_{i+1} n_{i+1} + \beta_i n_i) \tag{7}$$

where

$$\alpha_{i+1} = \begin{cases} c_i \frac{\beta(V_{i+1} - V_i)}{1 - e^{\beta(V_i - V_{i+1})}} & \text{if } V_i \neq V_{i+1} \\ c_i & \text{if } V_i = V_{i+1} \end{cases}$$

$$\beta_i = \begin{cases} c_i \frac{\beta(V_{i+1} - V_i)}{e^{\beta(V_{i+1} - V_i)} - 1} & \text{if } V_i \neq V_{i+1}) \\ c_i & \text{if } V_i = V_{i+1} \end{cases}$$

$$c_i = \frac{1}{\beta} \frac{\mu_i \delta_i + \mu_{i+1} \delta_{i+1}}{\delta_i + \delta_{i+1}}. \tag{8}$$

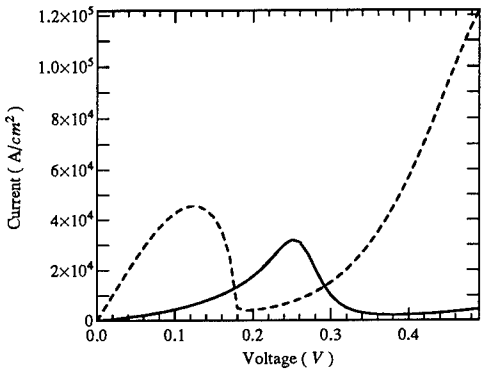


Figure 4: Flat-Band and original Thomas-Fermi models at 300k. Solid line is the Thomas-Fermi model, and dashed line is the flat-band model.

Therefore the matrix form of the continuity equation $j_{i-1/2} - j_{i+1/2} = 0$ can be written as:

$$\begin{bmatrix} 1 & 0 & 0 & & & & & & \\ -\beta_1 & \alpha_1+\beta_1 & -\alpha_2 & & & & & & \\ & \ddots & & & & & & & \\ & & -\beta_{l-1} & \alpha_l - a_1 & a_2 & & & & \\ & & & a_1 & \beta_{r+1} - a_2 & -\alpha_{r+1} & & & \\ & & & & \ddots & & & & \\ & & & & & -\beta_p & \alpha_p+\beta_p & -\alpha_{p+1} & \\ & & & & & 0 & 0 & 1 & \end{bmatrix} \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_l \\ nr \\ \vdots \\ n_p \\ n_{p+1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (9)$$

Then the above discretized current continuity equation(9) is solved self-consistently with Poisson's equation (2).

3 Results and Conclusion

One example is given in Fig. 4 and Fig. 5 calculated by the above models (well width = $37A$, barrier thickness = $43A$, mole factor = 0.244, undoped layer thickness = $150A$, the experimental data is from [9]). The results obtained from the Flat-Band and the original Thomas-Fermi models are illustrated in Fig. 4. Fig. 5 gives the results of the adjusted Thomas-Fermi and Hybrid quantum-classical models. It is obvious that the hybrid quantum-classical model gives the best matching result. The remaining valley current disagreement suggests better predictions will require a more accurate model, which can take into full account the incoherent processes.

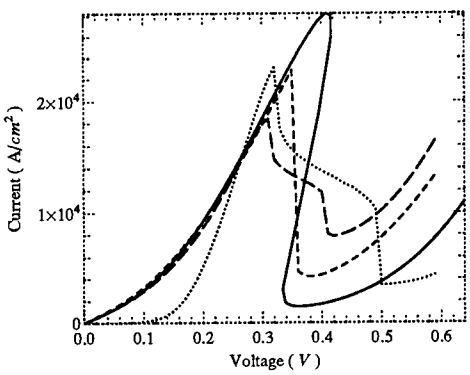


Figure 5: Adjusted Thomas-Fermi Model and Hybrid Quantum-Classical model compared with experimental results. Solid line is the adjusted Thomas-Fermi Model at 300K ($d = 11A$, $r = 4.5 \times 10^{-6}\Omega cm^2$, r is choosen such that the simulation matches the turn-on characteristics obtained from the experimental result), dashed line is the Hybrid Quantum-Classical model ($d = 44A$, $\mu = 10cm^2/(Vsec)$) at 300K, long-dashed line is the experimental result at 300K, and dotted line is the experimental result at 77K.

One also notices that the adjusted Thomas-Fermi model gives better result than the orginal model.

We have observed that the $I(V)$ curve obtained by this hybrid model is sensitive to the values of mobility μ, quantum region range x_l and x_r. With smaller mobility, peak and valley current densities are smaller and the peak and valley voltages are also smaller. Generally, as d increases, the peak and valley voltages move to the left, i.e. smaller values, but the peak and current densities can be either larger or smaller, depending on detailed quantum structures.

In conclusion, the calculation from the hybrid quantum-classical model matches best with experimental result out of the three models based on the tunneling theory. However, there are still some problems remaining (such as the low mobility used in the calculation) which need further research work.

REFERENCES

1. Sollner, T. C. L. G., W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck, 1983, Appl. Phys. Lett. **43**, 588.

2. Frensley, W. R., 1990, Rev. Mod. Phys. **62**, 745.

3. Mains, R. K., and G. I. Haddad, 1988, J. Appl. Phys. **64**, 5041.

4. Tsuchiya, H., M. Ogawa, and T. Miyoshi, 1991, IEEE Trans. Electron Devices, **38**, 1246.

5. Mains, R. K., and G. I. Haddad, 1988, J. Appl. Phys. **64**, 3564.

6. Hu, Y., and S. P. Stapleton, 1993, IEEE J. Quantum Electronics, **29**, 327.

7. Kluksdahl, N. C., A. M. Kriman, D. K. Ferry, and C. Ringhofer, 1989, Phys. Rev. B **39**, 7720.

8. L. Esaki and R. Tsu, 1973, Appl. Phys. Lett. **22**, 562.

9. Reed, M. A., Frensley, W. R., Duncan, W. M., Matyi, R. J., Seabaugh, A. C., and Tsai, H.-L., 1989, Appl. Phys. Lett. **54**, 1256.

A PROPOSED STARK SHIFT ELECTROOPTIC DEVICE OPERATING IN VISIBLE WAVELENGTHS

Andalib A. Chowdhury, M. Mahbub Rashed and C. M. Maziar

Microelectronics Research Center, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712

ABSTRACT

We report calculations of transition energies as a function of applied electric field for different combinations of GaP and AlP thicknesses. Most of the calculated red shifted transition energies fall within the orange/yellow region of the visible spectrum. These results may prove to be useful in designing Stark shift electrooptic devices operating in visible wavelengths.

INTRODUCTION

Electric fields applied perpendicular to quantum wells may be observed to shift the associated electron transition energies and optical absorption edges (the quantum confined Stark effect or QCSE). The QCSE has been studied extensively for Type I, III-V compound semiconductor systems [1], in which electron and hole quantum wells are formed in the same layer. This effect has also been exploited in Type II (electron and hole quantum wells are formed in adjacent layers) strained $Si_{1-x}Ge_x/Si$ heterostructures for electrooptic device applications [2]-[3]. The GaP/AlP material system, which also has Type II band alignment, may be more suitable than the strained $Si_{1-x}Ge_x/Si$ system for this particular application since it is almost lattice matched. In addition to having the characteristic advantage of Type II heterostructures, namely large transition shifts under modest electric fields, it offers a potential application in the visible wavelength region, as is shown in this work. This material system has another interesting feature: it is almost lattice matched with Si [Fig. 1].

Many of the investigations discussed in the literature dealing with the optical properties of GaP/AlP heterostructures were based on the quasi-direct band gap created by Brillouin zone-folding in short period superlattices (SLs) [4]-[7]. This effect has been observed for periodic

0-7803-0894-8/93 $3.00

multilayer structures in which each period is only a few monolayers thick. Such superlattices exhibit a considerable amount of overlap of quantum well wavefunctions so that bands of energy are formed. Interestingly, in this work, we show that the large Stark shift in GaP/AlP multiple quantum wells (MQWs) can be observed for relatively thick layers (>50Å) with negligible overlap of wavefunctions between quantum wells.

In this work, we report calculations of transition energies as a function of GaP and AlP layer thickness, and applied electric field for GaP/AlP MQWs grown on GaP substrates. Most of the calculated red shifted transition energies fall within the orange/yellow region of the visible spectrum. These results should be useful in designing GaP/AlP Stark shift electrooptic devices operating in visible wavelengths. From the calculation of the square of the overlap integral, we also show that the transition probability decreases with the applied electric field. We believe these calculations are the first reported of their kind for non-zone-folded transitions.

MODELS/RESULTS

Figure 2 illustrates the QCSE in the GaP/AlP system, where E_1 and HH_1 are the energies of the minimum bound electron and hole levels in the quantum well, respectively, and L is the width of the quantum well. The band alignment of these structures is Type II [8]. Under an applied electric field, the bands become tilted and the transition energies shift accordingly. For a Type II band alignment under applied electric field ε , the transition energy from the conduction band to the heavy hole valence band is lowered by $\Delta E_1+\Delta HH_1-e\varepsilon L$ near one wall of the well and raised near the opposite wall of the well by $\Delta E_1+\Delta HH_1+e\varepsilon L$. In these structures, the transition is indirect in real-space because the electrons and holes are spatially separated, and indirect also in k-space because the conduction band minima is at the X point while the valence band maxima is at the Γ point [5]. Transitions occur due to the overlap of the wavefunction at the interfaces. The presence of a sharp interface breaks the translational symmetry, allowing different k values to be mixed at the heterointerface. Though these effects have not yet been experimentally verified for GaP/AlP systems, these mechanisms are believed to be responsible for radiative recombination in Type II AlGaAs/GaAs [9] and in $Si_{1-x}Ge_x/Si$ MQW structures [3].

The band offset value was taken from the recently reported capacitance-voltage (C-V) measurements [8]· In Figure 3, the variation of minimum transition energies from the bound electron to the bound heavy hole states are shown as a function of applied electric field for

different well/barrier thicknesses. The bound energy states were calculated using the Envelope Function method [10], neglecting the intervalley interaction. Similarity of conduction and valence band structures of GaP and AlP justifies the use of this simple but useful method. Using this method the following is solved

$$\cos[q(L_A + L_B)] = \cos(k_A L_A)\cos(k_B L_B) - \frac{1}{2}\left(\frac{k_A}{m_A^*}\frac{m_B^*}{k_B} + \frac{k_B}{m_B^*}\frac{m_A^*}{k_A}\right)\sin(k_A L_A)\sin(k_B L_B) \quad (1)$$

with

$$k_i = \sqrt{\frac{2m_i^*}{\hbar^2}(E - V_i)} \qquad i = A, B \quad (2)$$

where L is the layer thickness, m_i^* is the effective mass, V_i is the band edge, and the subscripts A and B denote the different layers of the superlattice. The value E represents the bound state energy of the quantum wells. For these calculations the discrete energy levels are obtained in the limit for "thick" superlattices. Table I lists the fundamental parameter values used in the calculations [11].

Table I. List of fundamental parameter values used in the calculations.

	GaP	AlP
Electron effective mass m_e^*/m_o	6.9	3.67
Heavy hole effective mass m_h^*/m_o	0.419	0.513
Lattice constant a_o (Å)	5.45	5.45
Valence band offset	0.41 eV	

Note the red shift in transition energies under an applied electric field and that most of the red shifted transition energies fall within the orange/yellow region of the visible spectrum. The shift of minimum bound state energies (ΔE_1 and ΔHH_1) under an electric field was calculated using the variational method of Bastard *et al.* [12] In these calculations of the shift in bound state energies, each quantum well of the MQW structure was treated independently of other quantum wells. The quantum wells are sufficiently separated so that negligible wavefunction overlap occurs between neighboring quantum wells. Though the change in

transition energy shift equals $\Delta E_1+\Delta HH_1 \pm e\mathcal{E}L$, the last term (which represents the potential drop between the center of two adjacent layers) prevails and that leads to large transition shifts. Figure 3 demonstrates that under the same electric field, a larger transition energy shift should occur for thicker layers. Since the oscillator strength or the optical transition probability of the band-to-band transition is proportional to the square of the overlap integral between the carriers envelope functions, we have calculated (Fig. 4) the square of the overlap integral between the lowest electrons and heavy holes states as a function of electric field for different GaP/AlP MQWs. These indicate the rapid decrease of the transition probability with electric field, particularly with wider wells and barriers. We see that there is high absorption at zero electric field and low absorption at high electric field. The loss of interband absorption is the result of the decreased overlap of electron and heavy hole wavefunctions at the heterointerface. It may be mentioned here that in Type I band aligned case, a high electric field need to be applied to increase the Stark shift. That may result in the problem of excess leakage current. These results indicate that GaP/AlP MQWs may have potential application in light modulation by electric field. We believe that these simple calculations demonstrate the concept that GaP/AlP MQW structures may be good candidates for Stark shift electrooptic device applications operating in visible wavelengths. However, more rigorous calculations should incorporate the physics of excitons.

CONCLUSION

In summary, the results of transition energy calculations for the lattice-matched GaP/AlP system are reported. Large transition energy shifts, even under modest electric fields, suggest the possibility of designing optimum GaP/AlP MQW structures for present day optoelectronic devices such as electrooptical modulators. We believe that these results should be useful for designing GaP/AlP MQW optoelectronic devices exploiting the QCSE at the visible wavelengths.

ACKNOWLEDGMENT

This work was supported by the Joint Services Electronics Program, F49-620-92-C-0027 and the National Science Foundation Presidential Young Investigator Program, ECS-9057633. We would like to acknowledge R. D. Dupuis and J. Neff for their helpful discussions and comments.

REFERENCES

[1] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood and C. A. Burrus, "Electric field dependence of optical absorption near the band gap of quantum-well structures," Physical Rev. B, vol.-32, pp. 1043-1060 (1985).

[2] J. S. Park, R. P. G. Karunasiri and K. L. Wang, "Observation of large Stark shift in Ge_xSi_{1-x}/Si multiple quantum wells," J. Vac. Sci. Technol. B, vol.-8, 217-220, 1990.

[3] Andalib A. Chowdhury, K. H. Jung, D. L. Kwong, and C. M. Maziar, "Non-zone-folded transition-energy calculations for quantum-confined Stark effect in $Si_{1-x}Ge_x/Si$ quantum wells,"J. Appl. Phys., vol.-70, pp. 3946-3948, 1991.

[4] M. Kumagai, T. Takagahara and E. Hanamura, "Optical properties and indirect-to-direct transition of GaP/AlP (001) superlattices," Phys. Rev. B, vol.-37, pp. 898-915, 1988.

[5] A. Morii, I Ohno, A. Kanda, K. Arai, K. Tokudome, K. Hara, J. Yoshino and H. Kukimoto, "Short-period superlattices of $(GaP)_n(AlP)_n$ grown by metalorganic vapor phase epitaxy," Jpn. J. Appl. Phys., vol.-30, pp. L1244-L1246, 1991.

[6] K. Asami, H. Asahi, T. Watanabe, M. Enokida, S. Gonda and Sg. Fujita, "Optical properties of GaP/AlP short-period superlattices grown by gas source molecular beam epitaxy," Appl. Phys. Lett., vol.-62, pp. 81-83, 1993.

[7] J. Neff, R. V. Chelakara, K. G. Fertitta, Andalib Chowdhury, R. D. Dupuis and C. M. Maziar, "Short-period $(GaP)_n(AlP)_n$ superlattices grown by MOCVD," Elec. Mat. Conf., paper no. E1, Univ. of California, Santa Barbara, June 23-25, 1993.

[8] A. Morii, H. Okagawa, K. Hara, J. Yoshino and H. Kukimoto, "Band discontinuity at $Al_xGa_{1-x}P/GaP$ heterointerfaces studied by capacitance measurements," Jpn. J. Appl. Phys., vol.-31, pp.L 1161-L1163, 1992.

[9] B. A. Wilson, C. E. Bonner, R. C. Spitzer, R. Fischer, P. Dawson, K. J. Moore, C. T. Foxon and G. W. 't Hooft, "Radiative recombination mechanisms in staggered-alignment (GaAs)/(AlAs) heterostructures," Phy. Rev. B, vol.-40, 1825-1835, 1989.

[10] G. Bastard and J. A. Brum, "Electronic states in semiconductor heterostructures," J. Quantum Electronics, vol.-22, pp. 1625-1644, 1986.

[11] Landolt-Bornstein, *Numerical Data and Functional Relationships in Science and Technology* (Springer, New York, 1982), Group III, vol 22a.

[12] G. Bastard, E. E. Mendez, L. L. Chang and L. Esaki, "Variational calculations on a quantum well in an electric field," Phy. Rev. B, vol.-28, pp. 3241-3245, 1983.

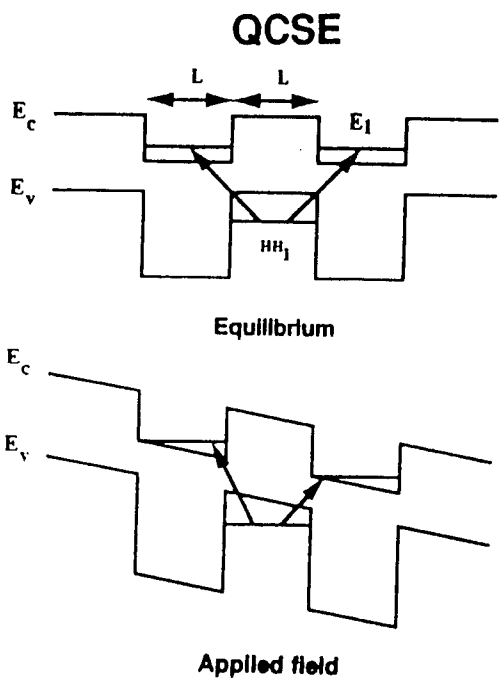


Figure 1: Energy band diagram illustrating the QCSE in equally thick GaP/AlP MQWs.

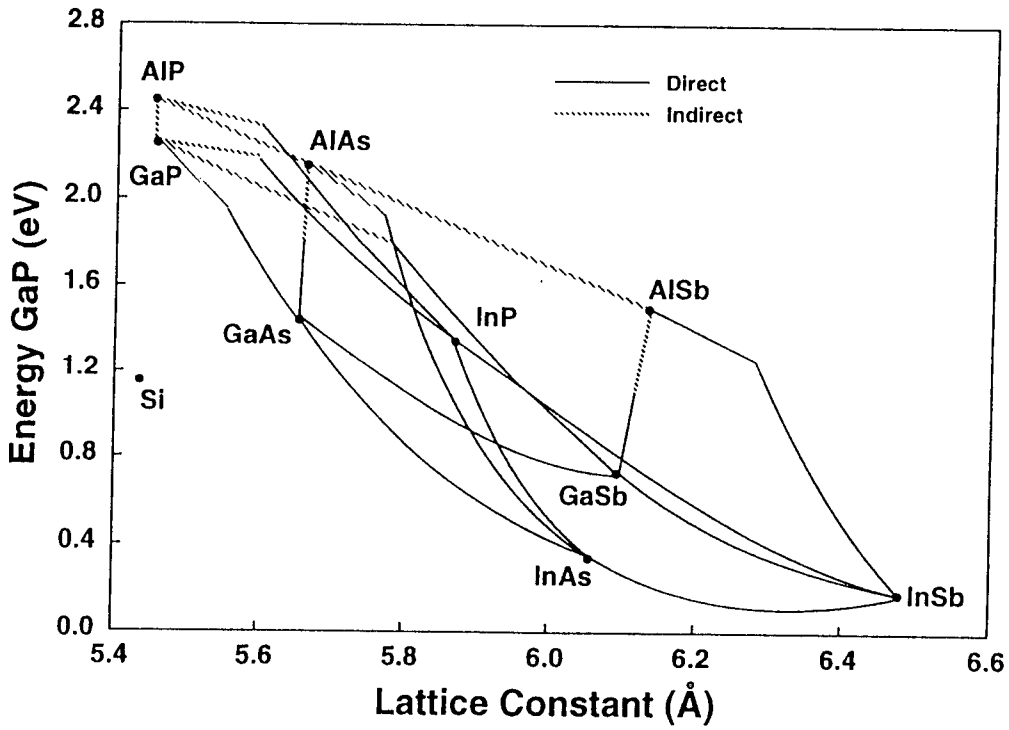


Figure 2: Band gaps and lattice constants of different III-V materials [7]. Lattice constant of silicon is 5.43 Å.

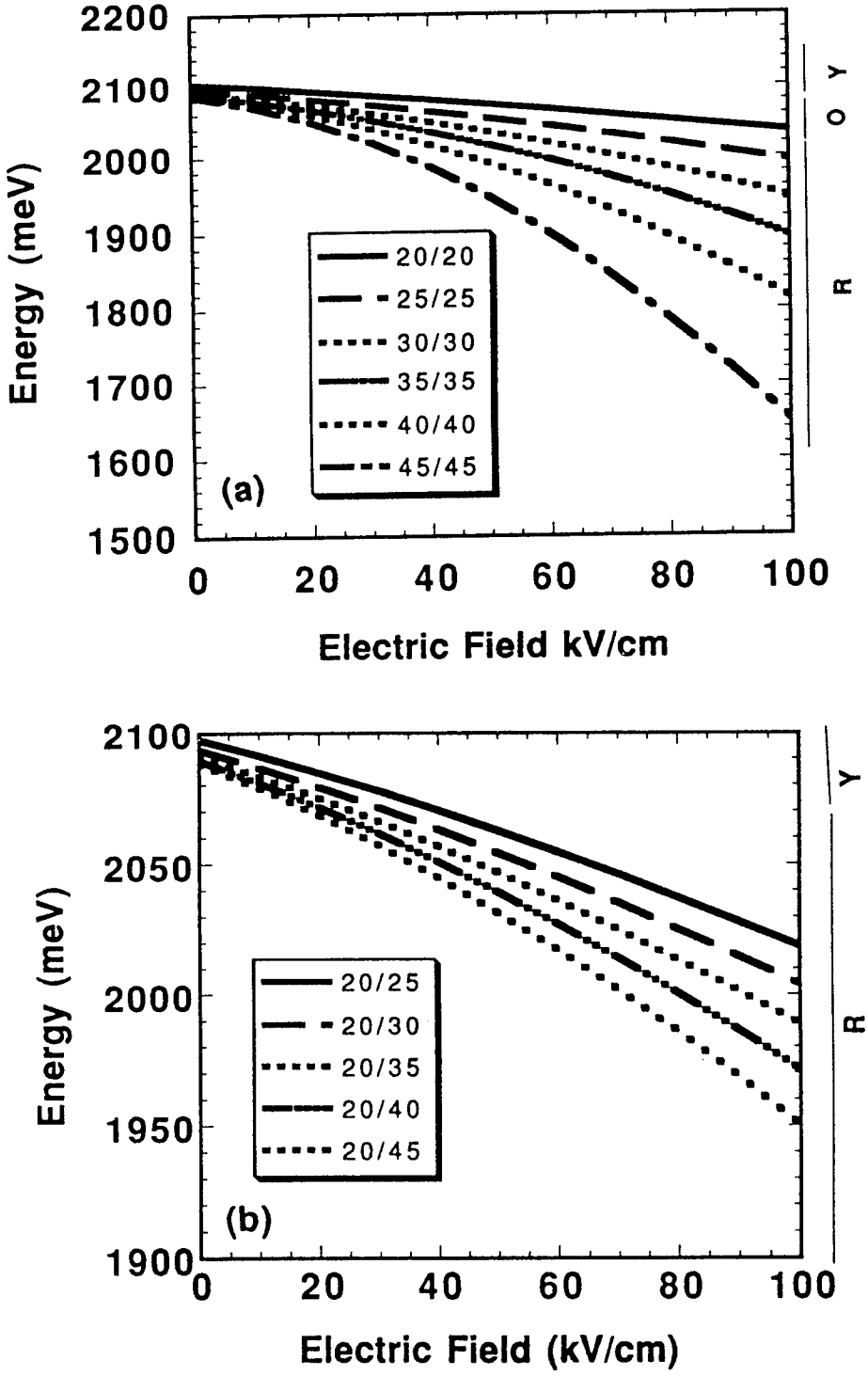


Figure 3: Variation of the transition energies as a function of electric field for (a) $(AlP)_m(GaP)_m$ and (b) $(AlP)_m(GaP)_n$ MQWs. The m's and n's in monolayer are shown in the legends. Y stands for yellow, O for orange and R for red wavelengths.

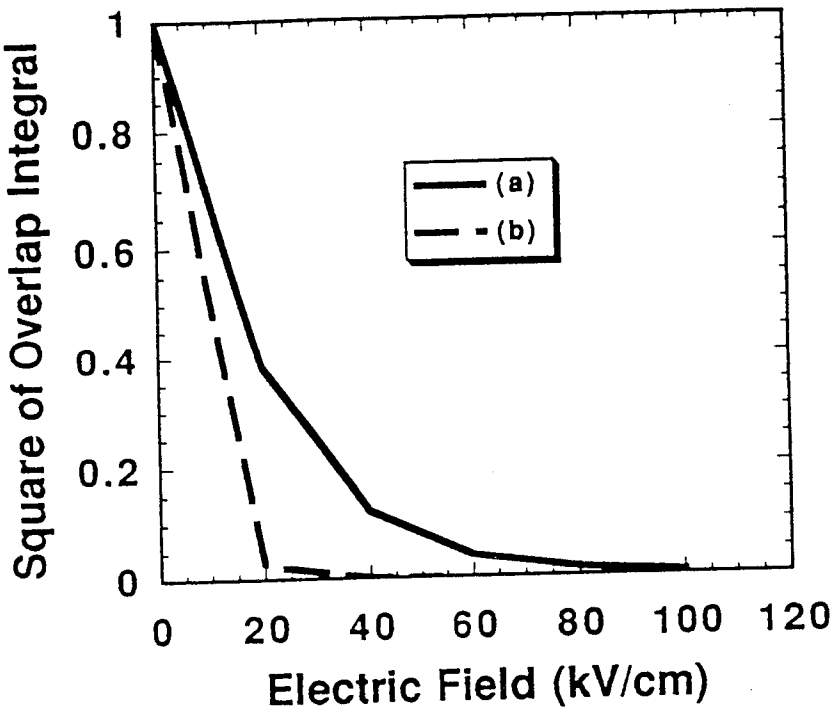


Figure 4: Square of the normalised overlap integral between electrons and heavy holes as a function of electric field for (a) $(AlP)_{20}(GaP)_{20}$ and (b) $(AlP)_{30}(GaP)_{30}$ MQWs. These indicate that the transition probability decreases with electric field.

On the Feasibility of Intersubband Transition Lasers *

A. Afzali-Kushaa, G. I. Haddad, and T. B. Norris

Solid-State Electronic Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, Michigan 48109

ABSTRACT

The feasibility and potential of laser sources based on intersubband transitions in quantum wells and strained layers will be presented. The basic schemes and proposed structures for both electrically and optically pumped devices are discussed. Both conduction band and valence band quantum wells as well as strained layers may be used as the active layer of these lasers. These sources can be either optically or electrically pumped with each having its own advantages. Various material systems which are appropriate for these applications will be described.

I. INTRODUCTION

Recent advances in crystal growth techniques such as MBE, CBE, and MOCVD have enabled us to grow thin layers of semiconductor materials on top of bulk materials. Depending on the thickness and the material properties of the grown layers, one can realize either a bulk like or a quantum well layer. If the thickness of the grown layer is less than the de Broglie wavelength and the energy gap of the layer is different from the surrounding layers, a quantum well can be realized in which quantum size effects become easily observable. When the carriers are confined in the valence band of the layer, the well is referred to as "valence band quantum well" while when the carriers are confined in the conduction band of the layer the well is referred to as "conduction band quantum well." The confinement of carriers in one direction and the lack of confinement in the other two directions in the structure lead to the formation of subbands of bound states. The energy separation of subbands is inversely proportional to the square of the well width. In the strained system, the lattice constant of the grown layer does not match the lattice constant of the substrate which leads to lifting the degeneracy of the of light- (LH) and heavy-hole (HH) subbands at the Brillouin zone center. If the strain is compressive, i.e., the lattice constant of the epilayer is larger than that of the substrate, the induced non-degeneracy gives rise to a smaller energy difference between the HH subband and the conduction band compared to the LH subband to conduction band energy separation. If, however, the strain is tensile, i.e., the lattice constant of the epilayer is smaller than that of the substrate, the non-degeneracy leads to a closer LH subband to the conduction band. The energy difference between the top of the LH and HH subbands is proportional to the strain in the system. Energy separation between subbands in both quantum well and strained layers therefore can be designed properly to be in the THz range (10-20 meV) making these layers a suitable material system for detectors, modulators, and lasers in this range of frequency.

* This work is supported by the NASA Center for Space Terahertz Technology under contract No. NAGW-1334 and the Army Research Office the URI program under contract No. DAAL03-92-G-0109.

II. PROPOSED STRUCTURES

1) Intersubband Conduction Band Quantum Well Lasers

Recently, several research groups [see e.g. 1-14] have reported studies on intersubband transitions in quantum wells with applications such as infrared lasers, light modulators (switches), and detectors. Strong optical absorption [8] and spontaneous emission [9] due to intersubband transitions have been observed experimentally. Although these observations are encouraging, in order to realize intersubband lasers, enough photon gain should be achieved to compensate all the photon losses in the quantum well. The gain in a laser system is directly proportional to the population inversion in the active layer. To achieve population inversion in the quantum well, both optical and electrical pumping schemes have been suggested [1-7]. Figure 1a shows the band diagram of the simplest electrically pumped intersubband laser suggested by Mehdi *et al.* [1] and Loehr *et al* [2]. Carriers are injected from the left contact and by tunneling through the left barrier, reach to subband 2 and radiatively relax to subband 1 and finally, after tunneling through the right barrier, they are collected by the right contact. Another proposed electrically pumped intersubband laser is shown in Figure 1b. To improve population inversion in the quantum well (the active layer of the laser), resonant tunneling filters have been utilized in the structure to selectively inject carriers into subband 2 and remove carriers from subband 1 [4,5].

The major problems associated with the electrically pumped laser schemes is that the confinement factor in the system, which accounts for the reduction in gain that occurs because of the spreading of the optical energy beyond the active layer, is very low. The net gain therefore is too small which makes the laser action very difficult. This factor can be improved by utilizing the optically pumped laser scheme where one period of its active layer is shown in Figure 2. The structure may utilize 10-20 periods of the well shown in the figure to enhance the confinement factor about a factor of 100-400. The proposed scheme uses a heavily n-doped layer to provide carriers in subband 1 where a pumping laser excites carriers from the subband into subband 4. After relaxing to subband 3, these hot electrons radiatively relax to subband 2, and finally return to subband 1 and the carrier cycle completes.

In conduction band quantum wells, the symmetry of the subbands is the same. Therefore, the tunneling and relaxation rates in the well only depend on the relative energy difference of the subbands. Population inversion, which is a crucial requirement for laser action, may be created when the injection and removal rate of the carriers from the subbands are not identical. Therefore, inverting the population of two different subbands in these wells is rather difficult.

2) Intersubband Valence Band Quantum Well Lasers

Instead of using conduction band quantum wells, one can use valence band quantum wells in similar structures. In these structures, the tunneling and relaxation rates of the hole from one subband to the other is not only dependent on the energy difference of the two subbands, but it also depends on the symmetry of the subbands. This additional feature of the subbands in valence band quantum wells, makes them an excellent choice for inter-subband quantum well lasers. The origins of different transition rates of the hole from different subbands are briefly expressed in the following:

a. The tunneling rate is a function of the carrier mass in the growth direction. The different masses of the holes in the growth direction, therefore give rise to different tunneling rates for light and heavy holes.
b. Under the application of a strong magnetic field parallel to the interface, the motion of holes passing through the barrier is altered [14]. This can be considered as a

decrease of the kinetic energy in the tunneling direction or an effective increase of the tunneling barrier. The additional voltage needed, to obtain the same tunneling rate as in the case of zero magnetic field, is inversely proportional to the effective mass of the hole in the growth direction.

c. In strained systems, the barrier seen by the light hole is different from the barrier seen by the heavy hole due to the induced non-degeneracy in the system. Since the tunneling rate decreases exponentially as the barrier increases, the tunneling rate would have different values for the heavy-hole and light-hole.

d. Depending on the symmetry, namely, LH or HH, of the subband, the tunneling rate of the carriers may be different. This provides another feature which can be utilized in designing systems to obtain a better population inversion. The non-radiative relaxation rate between two subbands in the same well also might be reduced if the subbands have different symmetry.

Any of a-d or some combinations thereof may be exploited in designing structures for intersubband lasers where achieving the population inversion between the two subbands involved in the radiation is a vital requirement. The conclusion reached from the above argument suggests that the threshold of the input power (or carrier density) for lasing should be lower in properly designed valence band quantum wells compared to a similar structure but based on conduction band quantum wells. Despite the aforementioned potential of the valence band quantum well, it has not been studied as extensively as conduction band quantum wells. White *et. al.* [15] have observed population inversion with a ratio of 300/1 between LH1 and HH1 in a p-i-n structure. The energy separation between the two subbands is 21 meV which is less than the LO phonon energy. Since the population inversion has been achieved in this simple quantum well, one can consider a simple quantum well, similar to the one shown in Figure 1a, as a potential structure for THz sources.

Figure 3 shows an electrically pumped laser scheme proposed by our group. The structure is the valence band quantum well version of the scheme shown in Figure 1b. In this structure, the transition between subband 3 and subband 1, as well as the transition between subband 2 and subband 0 are undesired and detrimental to the population inversion between subband 2 and subband 1. In the electron version, we do not have much control over reducing these rates except for adjusting the energies of the subbands. In the hole version of the proposed scheme, however, one can make use of the properties mentioned before to reduce these rates more. Figure 3 is an example of a design where due to the different symmetry between subband 3 (HH band symmetry) and subband 1 (LH band symmetry), the transition rate between these two subbands is reduced. The same argument holds for the transition from subband 2 (HH band symmetry) to subband 0 (LH band symmetry). In addition, the lighter effective mass of the carrier in subband 1 provides a higher tunneling rate for holes in this subband. This further improves the population inversion between subband 2 and subband 1.

Optically pumped laser schemes can take advantage of the relaxation properties of the holes in quantum wells. Figure 4 shows a valence band quantum well version of the optically pumped laser structure discussed earlier (see Figure 2). To invert the populations of subband 3 and subband 2, the non-radiative relaxation of holes from subband 4 to 2 as well as from subbands 3 to 1 should be as slow as possible while the non-radiative relaxation of holes from subband 4 to subband 3, as well as from subband 2 to subband 1 should be as fast as possible. A desired symmetry for the subbands of the laser is shown in the figure where unwanted relaxation corresponds to different symmetry transitions while desired relaxation occurs between the subbands with the same symmetry.

In all these proposed structures, in order to have coherent light, we would like to retain carriers in the higher energy subband as long as possible and let photons stimulate their radiation coherently. In addition, to maintain population inversion in the structure, the carriers in the lower

energy subband should be emptied out as soon as possible. Consequently, relaxation times between different subbands in a quantum well, as well as tunneling rates between different subbands in adjacent quantum wells are the key factors in determining the achievable population inversion due to a certain current density or pumping rate in an intersubband quantum well laser. Several studies have previously been devoted to intersubband or interwell scattering [see e.g. 16, 17]. These studies mostly have addressed the case of simple quantum wells and transition from subband 2 to subband 1.

In order to enhance the confinement factor of the laser, the use of optically pumped structures was suggested earlier. Although, a considerable improvement may be obtained, however, the factor is still far below unity in the frequency range of interest.

3) Bulk Strained Lasers

As has been mentioned earlier, the energy separation between LH and HH subbands in a strained layer is proportional to the amount of strain in the layer. By adjusting the strain, one can design the desired energy separation of the subbands. Depending on the strain, hoewever, the thickness of the strained layer may not exceed a certain thickness called critical thickness for the pseudomorphic or coherent layer, i.e., a layer without any dislocations. The critical thickness is *inversely* proportional to the strain on the system. For the energy separations that lead to THz radiation (around 10 to 20 meV), this thickness is more than 100 nm which can be a great advantage for THz lasers based on these systems as will be discussed later. Figure 5 shows the critical thickness for InGaAs alloys as a function of the frequency which can be radiated due to the transition of the hole from one subband to the other.

The far-infrared (FIR) amplification of the electromagnetic waves due to direct transition of the hot holes in *strong* electric and magnetic field (larger than 1 kV/cm and 1 T, respectively, induced by a voltage greater than a few kilovolts and a few Amperes) has been under investigation [see e.g. 18, 19]. The observed stimulated emission covers a broad wavelength range 80-210 mm with a line width of about 20 1/cmand with pulse emitted power up to 10W. The FIR radiation from uniaxially stressed p-type Ge in the absence of a magnetic field and at lower electric fields also has been observed [see e.g. 20, 21]. The uniaxial strain in the Ge layer split the degenerate valence band edge at the Brillouin zone center into two subbands separated by the energy Δ. In the measured structure which is a bar of the semiconductor, at zero electric field essentially all of the holes occupy the lower energy band with small effective mass in the stress direction. When an electric field is applied to the sample, the light holes heated by the field are transferred to the upper band with larger effective mass and higher density of states. The direct optical transitions may take place only due to the holes with energies above the band-split energy Δ.

Instead of using external uniaxial strain, one can use biaxial strained layers which are grown on a lattice mismatched substrate. An electrically pumped p-i-p laser structure based on strained bulk systems proposed by our group is shown in Figure 6. The active layer, which is intrinsic, is under tensile strain. When an electric field is applied to the system, the top of the emitter valence band is aligned to the top of the active layer HH band while the top of the collector valence band is aligned to the top of the active layer LH band. The applied field moves holes from the emitter valence band to the active layer HH band where holes radiatively relax to the active layer LH band. Holes are then collected by the collector layers. If the structure is designed for the light radiation with the photon frequency of 4 THz (~16 meV energy separation), the thickness of the active layer may be as high as 200 nm as can be extracted from Figure 5. Therefore, in these structures the confinement factor of the laser gain can be improved by a factor about 40 compared to that of the quantum well system.

An optically pumped laser which is based on strained bulk layers is shown in Figure 7. The active layer is under tensile strain and is p-doped. A laser pump is used to excite electrons from split-off band to LH band. Hot electrons in this band radiatively relax to HH band which create photons at a THz frequency. The relaxed holes to HH band then relax back to split-off band and complete the circle. The relaxation time between LH band and HH band will be long, if the energy separation between the HH band and the LH band at $k = 0$ (where k is the wave number) is less than LO phonon energy. However, the energy separation between the split-off band with the LH and HH bands are more than LO phonon energy leading to shorter transition rates between these bands.

III. Material System

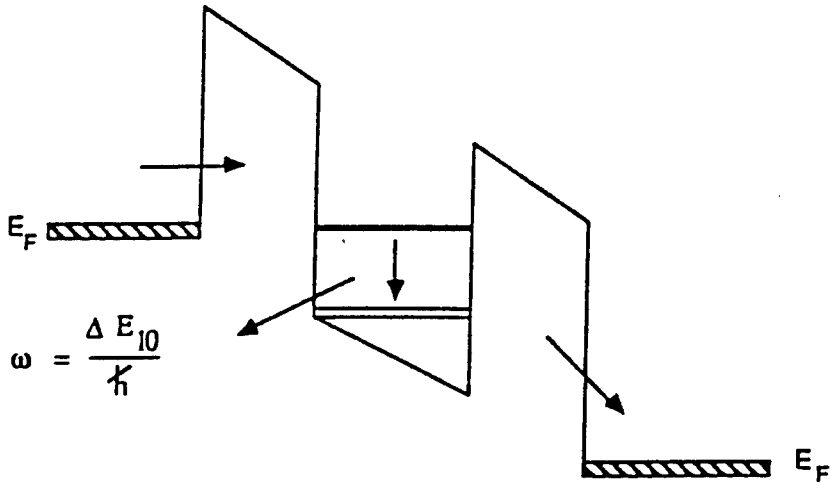
The choice of materials in none of the proposed structures is restricted to direct band gap semiconductors and indirect band gap materials, like Ge and Si, also can be utilized in the laser structure. Therefore, these system have the potential to be integrated eventually with other semiconductor devices in silicon based integrated circuits.

IV. Summary

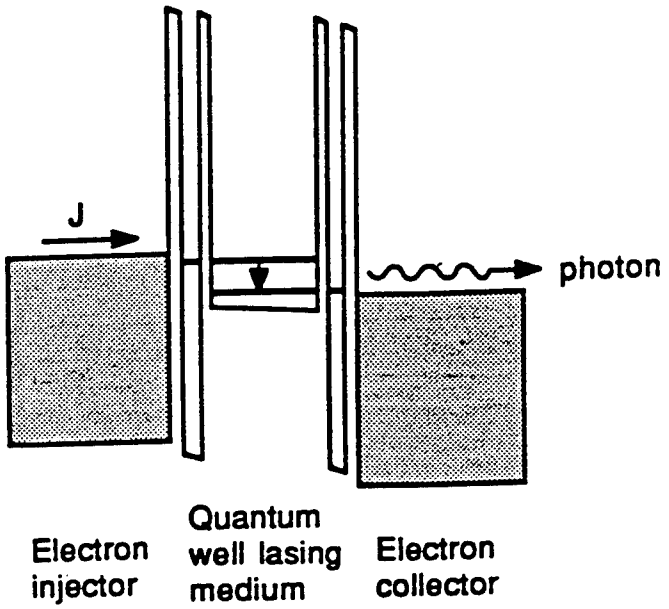
The potential application of the transition between subbands in both conduction band and valence band quantum wells as well as the transition between HH and LH subbands in strained layers for THz sources was discussed. Both electrically and optically pumped lasers based on quantum well and strained systems were proposed. The optically pumped schemes offer a larger confinement factor compared to electrically pumped structures in the same system. At present time, the strained system seems to be the most promising system in both electrically and optically pumped lasers among the proposed structures. The confinement factor in the optically pumped laser based on strained layers can be as high as 1. In addition, the factor is higher in the proposed electrically pumped scheme compared to similar structures based on quantum wells. The choice of materials in these structures is not limited only to the direct band gap semiconductors and indirect band gap semiconductors such as Si and Ge may be used in the active layer of these systems. To asses the performance and characteristics of the proposed structures further study is being carried out both theoretically and experimentally.

REFERENCES

1. I. Mehdi, G. I. Haddad, and R. K. Mains, Superlatt. Microstruc. 5, 443 (1989).
2. J. P. Loehr, J. Singh, R. K. Mains, and G. I. Haddad, Appl. Phys. Lett. 59(17), 2070 (1991).
3. H. C. Liu, J. Appl. Phys., 63(8), 2856, (1988).
4. Borenstain S. I. and Katz J., Applied Physics Letters, 1989, Vol 55, 654-656.
5. Hu Q. and Feng S., Applied Physics Letters, 1991, Vol 59, 2923-2925.
6. P. Yuh and K. L. Wang, Appl. Phys. Lett. 51(18),1404 (1987).
7. K. M. Lau and W. Xu, IEEE J. Quantum Electron. QE-28, 400 (1992).
8. L. C. West and S. J. Eglash, Appl. Phys. Lett. 46, 1156 (1985).
9. H. Helm, E. Colas, P. England, F. DeRosa, and S. J. Allen, Jr., Appl. Phys. Lett. 53, 1774 (1988).
10. D. D. Yang, P. Boucaud, F. H. Julien, L. Chusseau, J. M. Lourtioz, and R. Planel, Electronics Letter, Vol. 26, pp. 1531-1532, 1990.
11. S. Noda, T. Uemura, T. Yamashita, and A. Sasaki, J. Appl. Phys., vol. 68, pp. 6529-6531, 1990.
12. Oberli D. Y., Wake D. R., and Klein M. V., Physical Review Letters, 1987, Vol. 59, 696-699.
13. I. V. Altukhov, M. S. Kagan, K. A. Korol'ov, V. P. Sinis, F. A. Smirnov, Proceedings 1991 International Semiconductor Device Research symposium, 371-374.
14. R. P. G. Karunasiri and K. L. Wang, J. Vac. Sci. Technol. B 9 (4), Jul/Aug 1991, 2064-271.
15. C. R. H. White, H. B. Evans, L. Eaves, P. M. Martin, and Henini, Phys. Rev. B, 1992, pp. 9513-9516.
16. Ferreira R. and Bastard G., Physical Review B, 1989, Vol. 40, PP 1074-1086.
17. B. Deveaud, A. Chomette, F. Clerot, P. Auvray, and A. Regreny, R. Ferreira, and G. Bastard, Physical Review B, 1990, Vol. 42, 7021-7032.
18. Keilmann, V. N. Shastin, and K. Till, Appl. Phys. Lett. 1991, 58, pp. 2205-2207.
19. L. E. Vorobjev, S. N. Danilor, D. V. Dontzky, Yu. V. Kochegarov, V. I. Stadfeev, and D. A. Firsov, Proceedings: 1991 International Semiconductor Device research Symposium, pp. 387-389.
20. I. V. Altukhov, M. S. Kagha, and V.p. Sinis, Optical and Quantum Electronics, 23, No. 2, 1991, pp. S211-S216.
21. I. V. Altukhov, M. S. Kagha, K. A. Korol'ov, V.p. Sinis, and F. A. Smirnov, Proceedings: 1991 International Semiconductor Device research Symposium, pp. 371-374.



(a)



(b)

Figure 1. Electrically pumped laser based on conduction band quantum wells a) single quantum well [1,2] b) coupled quantum well [4,5].

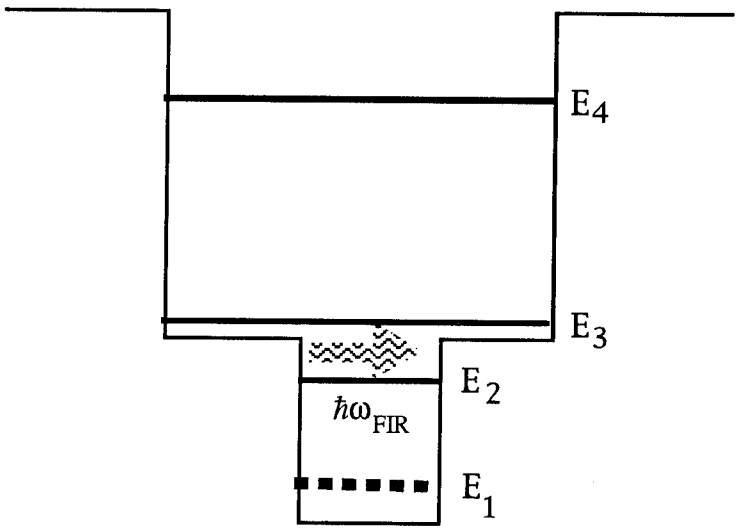


Figure 2. Optically pumped laser structure based on conduction band quantum wells.

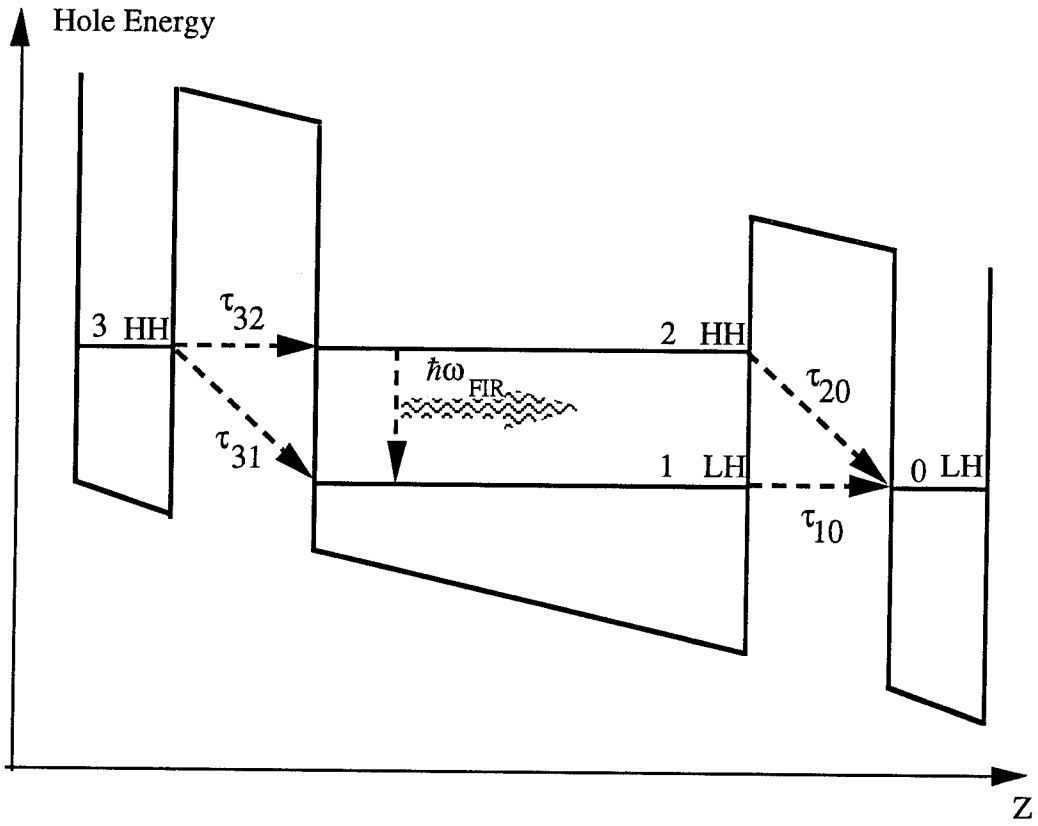


Figure 3. Electrically pumped laser based on valence band quantum wells

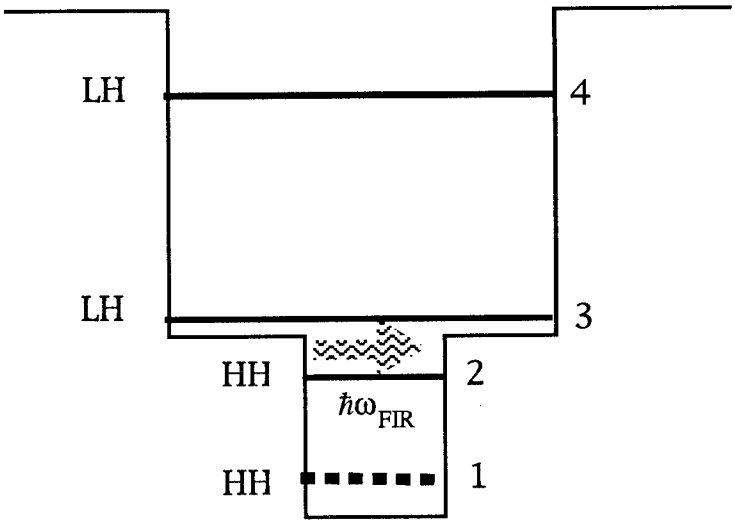


Figure 4. Optically pumped laser structure based on valence band quantum well lasers

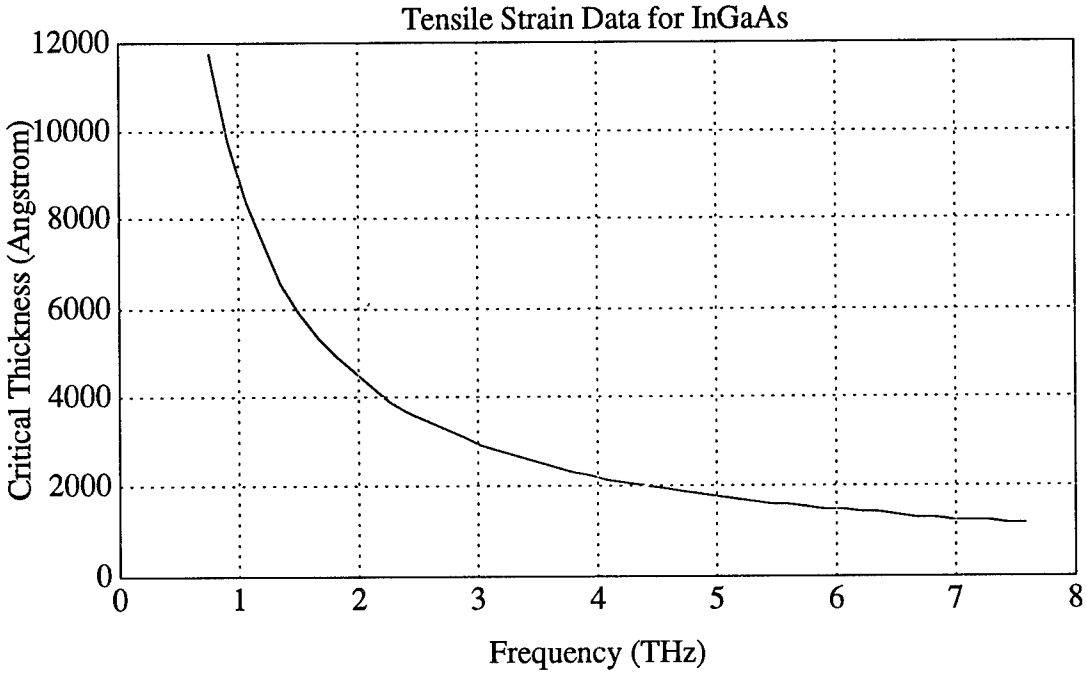
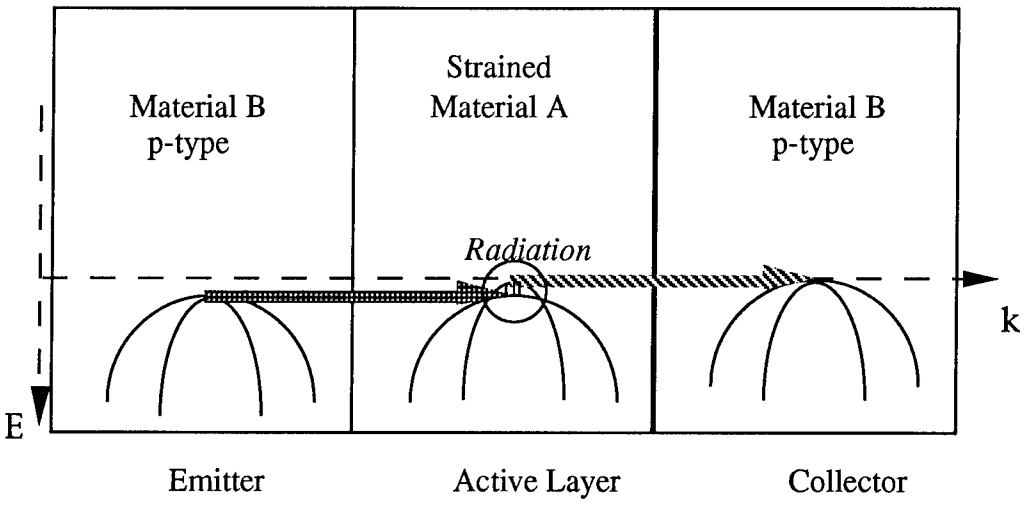


Figure 5. Critical thickness of InGaAs/GaAs systems as a function of frequency which can be radiated due to transition of holes from one subband to the other.



$a_{substrate} = a_B$

$a_{substrate} > a_A$ Tensile Strain

Figure 6. Electrically pumped laser structure based on strained layers

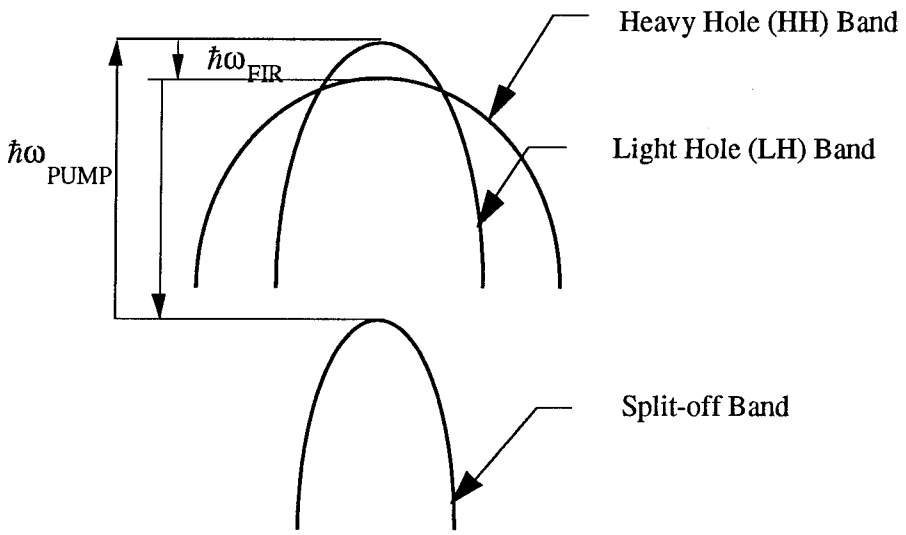


Figure 7. Optically pumped laser structure based on strained layers

OBSERVATION OF NOVEL CONDUCTANCE STRUCTURE IN $GaAs/Ga_xAl_{1-x}As$ RESONANT TUNNELING HETEROSTRUCTURES

M.R.Deshpande[(1)], N.H.Dekker[(1)], J.W.Sleight[(1)], J.L.Huber[(1)], E.S.Hornbeck[(1)], M.A.Reed[(1)], R.J.Matyi[(2)], Y.C.Kao[(3)], C.J.L.Fernando[(4)], W.R.Frensley[(4)]

[(1)]Yale University
[(2)]University of Wisconsin at Madison
[(3)]Texas Instruments
[(4)]University of Texas, Dallas

ABSTRACT

A novel pre-resonant conductance structure is observed in single and double well $GaAs/Ga_xAl_{1-x}As$ resonant tunneling heterostructures. This structure is attributed to single electron tunneling through donor bound states in the quantum well as they cross the Fermi level. Peaks in conductance are observed in devices as large as $(64\mu)^2$. We observe for the first time donor binding energies in quantum wells as large as 35 meV, which is probably due to the formation of a donor complex. An impressive impurity tunneling conductance structure is observed in double quantum well structures with conductance peak heights varying over 3 orders of magnitude. Conductance peaks are observed before and after the main resonant current peak which are attributed to tunneling of electrons through impurities in one quantum well and the quantum state in the other well.

INTRODUCTION

The subject of vertical transport through confined heterostructures is of great interest in view of the continuous trend towards miniaturized devices [1,2]. Quantum well devices are being used in applications such as oscillators, switching devices, quantum well transistors, multistate memory and analog - digital converters [3]. Recently additional structure is observed in the conductance - voltage characteristics of sub-micron quantum well structures which is attributed to tunneling through donor bound states in the quantum well [4,5]. We observe similar structure in large mesa size $\{(2\mu)^2$ to $(64\mu)^2\}$ single and double quantum well heterostructures.

The structures under study have an undoped active region, bounded by 150 A of undoped GaAs spacer layers separating the Si doped n^+ GaAs emitter and collector regions. The single well structures have a 44 A quantum well and $Ga_{0.27}Al_{0.73}As$ symmetric barriers of thicknesses 118 A, 85 A and 65 A for devices 2013, 2014 and 2015 respectively. The double well structure (device 1850) has an active region consisting of two quantum wells of width 50 A and 80 A separated by an AlAs barrier of width 15 A, and bounded by 40 A AlAs barriers on either sides.

0-7803-0894-8/93 $3.00

Results and Discussion

Figure 1 shows the current and conductance vs bias voltage for the single quantum well device 2014. The current peaks are due to resonant tunneling through the quasi bound quantum well state. The conductance plot on the expanded logarithmic scale shows peaks in conductance prior to the resonant current peak. These peaks in conductance are randomly distributed in voltage. For a given epitaxial structure the number of peaks increases with area but the peak height (ΔG) is approximately the same. The peak height (Δ
epitaxial structures and is smaller for devices with thicker barriers.

We propose that these peaks in conductance are due to single electron tunneling through donor bound states in the quantum well [4]. Assuming an unintentional impurity concentration of $10^{14}\,cm^{-3}$ we estimate approximately two donor impurities in the quantum well for the 2 micron device. Calculation for donor atoms in a quantum well predict binding energies of 15 meV for centrally placed donors for these sample structures [6]

When these donor states cross the Fermi level single electrons from the emitter can resonantly tunnel through these states. This results in a sudden increase in current which shows as a step in I(V) and as a peak in the conductance. The current step is determined by -

$$\Delta I = e/\tau$$

e : electronic charge

τ : life time of the donor state

Impurities centrally located in the quantum well have higher binding energies than the impurities at the edge. Impurities are randomly distributed in the well and hence we see conductance peaks randomly distributed in voltage. With an increase in bias voltage the number of electrons that can tunnel through the impurity state increase. But these channels of conduction become saturated with single electron tunneling and hence we do not observe a further increase in current through this channel unless the lifetime of the state changes.

The lifetime of a donor state in the quantum well is larger in a device with a thicker barrier. Experimentally we observe the conductance peak heights (ΔG) in device 2014 (85 A barrier) to be smaller by approximately a factor of 10 than the conductance peak heights in device 2015 (65 A barrier). This is in agreement with the model indicating that the height of the current step (and the corresponding conductance peak) is inversely related to the lifetime of the donor state. We do not observe conductance peaks in device 2013 (118 A barrier) possibly because of current steps being below the measurement threshold.

For donor impurities physically located nearer to one barrier in the quantum well the "leakage " of the donor state wave function through the barriers is more for the bias direction when the near barrier is favorably biased. Thus the donor states have different lifetimes for different bias directions. This possibly explains the difference in conductance peak heights for different bias directions as observed in Fig. 2 for peaks 3 and 4. Notice that the first two peaks are symmetric implying that they are due to centrally located donors.

Calculations for donor atoms in quantum wells indicate donor binding energies of approximately 15 meV for the devices under study. A spectroscopic study with a numerical calculation using Thomas-Fermi approximation indicates that at a bias of approximately 0.15 V the quantum well state is approximately 14 meV above the Fermi level [7]. This suggests that the conductance peaks observed at 0.15 V are due to tunneling through donor states with binding energies of approximately 14 meV [Fig.3]. We also observe peaks at 0.1 V. Energy band calculation at 0.1 V show the quantum well state 35 meV above the Fermi level. This indicates

that we observe tunneling through donor states with binding energies of approximately 35 meV. These high binding energies are possibly due to an impurity complex formed in the quantum well.

The double quantum well device 1850 shows an impressive structure in conductance due to impurity tunneling. Fig 4 shows the current for positive bias voltage on a logarithmic scale. The peak in current at 0.17 V is due to resonant tunneling through the n=1 quantum well state of the far well. The peak at 0.67 V is due to an overlap of resonant tunneling through the n=1 state in the near well and n=2 state in the far well. The new step like structure in current at approximately 0.3V may be due to tunneling through impurity states in both the quantum wells.

Fig 5 shows the current and conductance for negative bias. The main current peak is very sharp, due to perfect overlap of resonant tunneling through the n=1 states of both the wells. The conductance plotted on a logarithmic scale shows a variation in peak heights over three orders of magnitude. Fig 6 shows the energy band diagrams obtained from the spectroscopic study of this device at various bias positions. It indicates that the low bias structure (-0.15 to -0.21) is due to tunneling through donor states in the far well as they cross the Fermi level. At a higher bias (-0.22 to -0.28) when the impurities in the far well align with the quantum state in the near well their lifetime decreases suddenly, resulting in another current step with a higher magnitude.

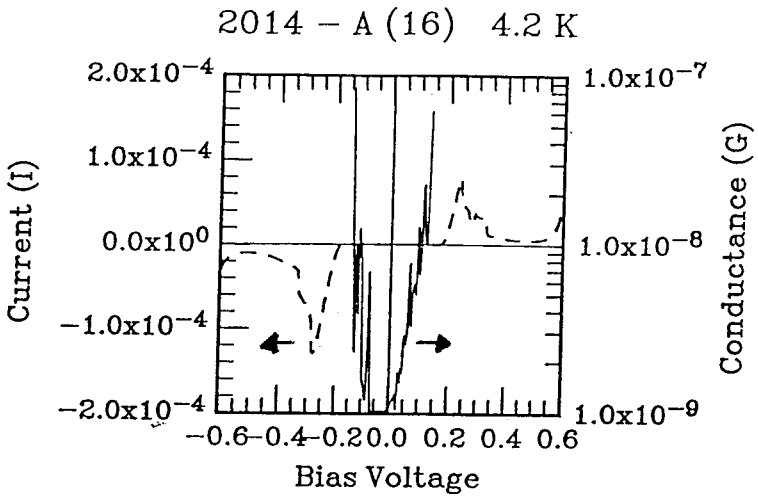
We also observe structure very near the current peak (which is probably due to tunneling through impurities in both wells), and after the main resonance. The post resonance structure may be due to a combination of two mechanisms. The lower bias structure near -0.5 V is possibly tunneling through impurities in the near well and the quantum state in the far well. The higher bias structure near -0.6 V is possibly due to tunneling through impurity states bound to the n=2 quantum state in the near well.

Conclusions

Conductance structure is observed in large area $\{(64\ \mu)^2\}$ resonant tunneling devices attributed to single electron tunneling through donor bound states in the quantum well. We observe for the first time donor binding energies of 35 meV, in addition to the normal binding energies of 14 meV. We observe an impressive conductance structure in the double well devices with peak heights varying over three orders of magnitude. In devices with doped quantum wells tunneling through donor states will be a dominant factor. Impurity tunneling will contribute to noise in resonant tunneling devices.

References

1) M.A.Reed *et al.*, Phys. Rev. Lett. **60**,535 (1988)
2) A.Groshev, Phys. Rev. **B 42**, 5895 (1990)
3) F.Capasso *et al.*, IEEE Tran. Electron Devices **36**, 2065 (1989)
4) M.W. Dellow *et al.*, Phys. Rev. Lett. **68**, 1754 (1992)
5) P.Gueret *et al.*, Phys. Rev. Lett **68**, 1896 (1992)
6) R.L.Green and K.K.Bajaj, Solid State Commn. **45**, 825 (1983)
7) M.A.Reed *et al.*, Appl. Phys. Lett. **54**, 1256 (1989)



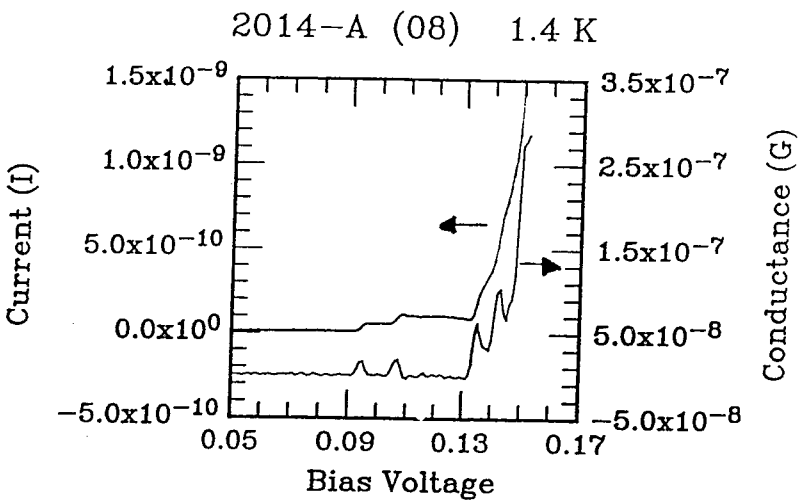


FIG. 1

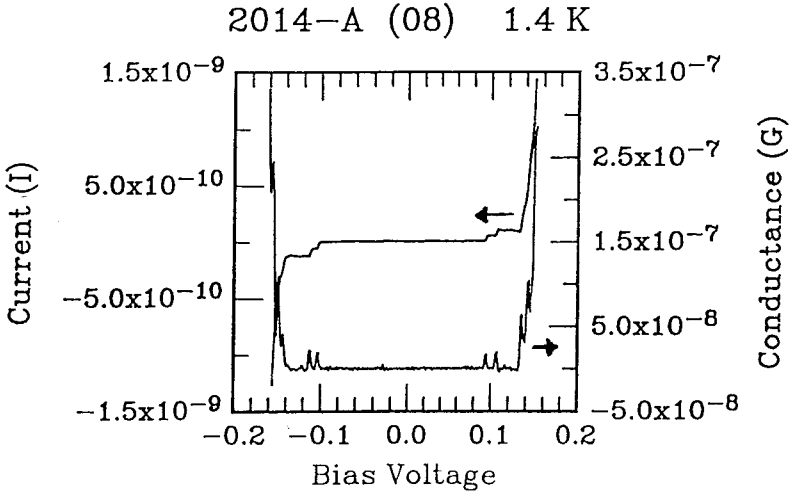


FIG. 2

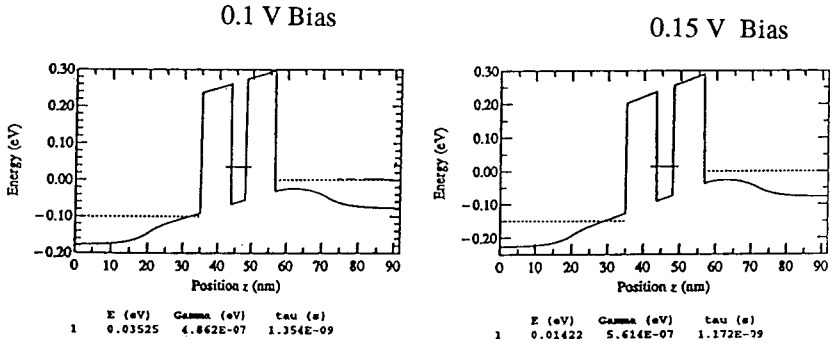


FIG. 3

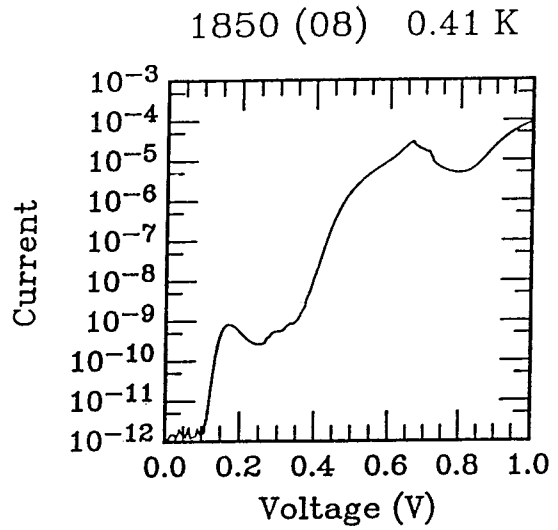


FIG. 4

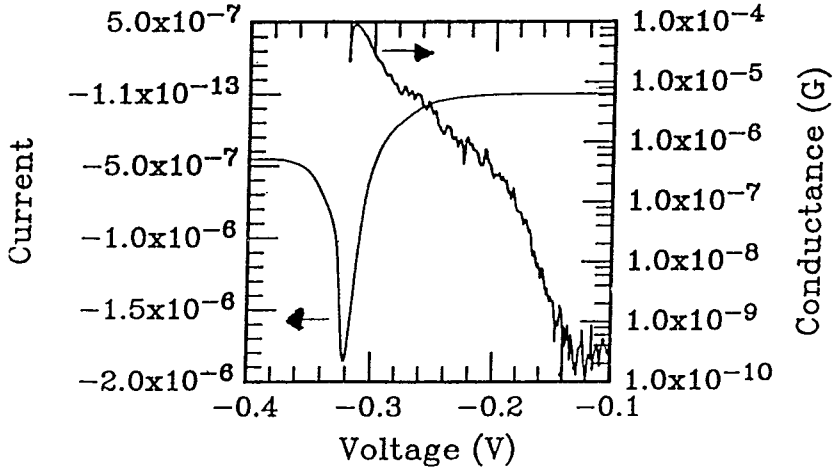


FIG. 5

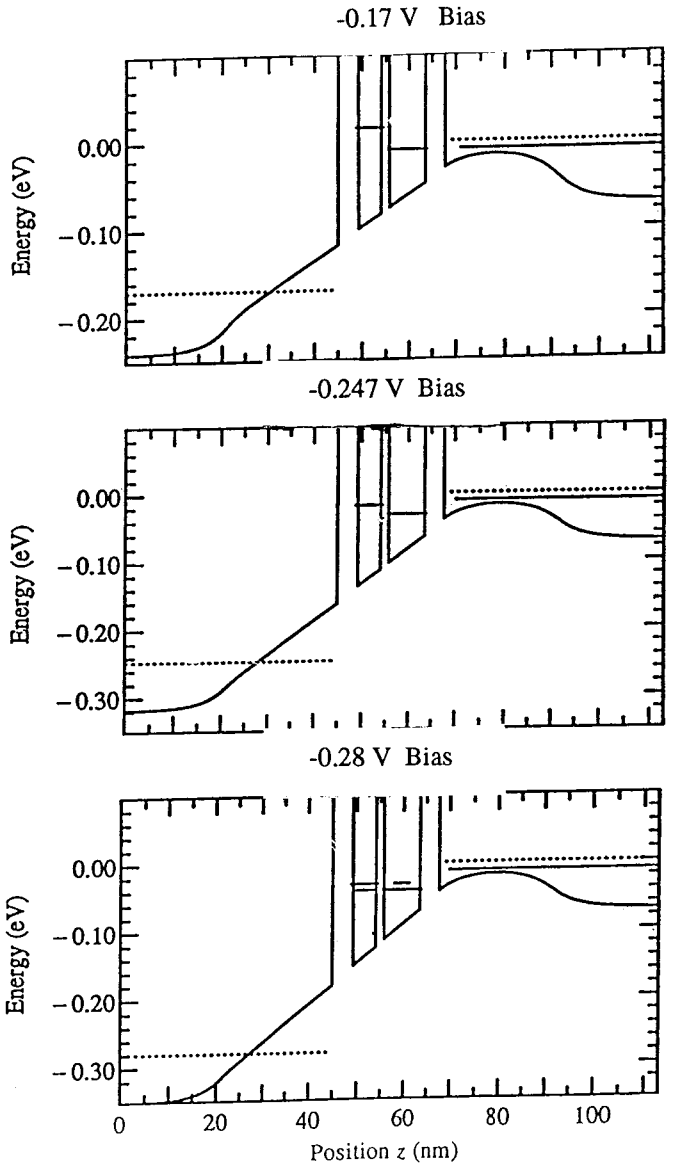


FIG. 6

A Coupled Mode Theory for Electron Wave Directional Couplers

Andrew M. Sarangan, Wei-Ping Huang
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada

Abstract

In this paper, a coupled mode theory for electron wave directional couplers is presented. The theory includes the dephasing effect on electrons due to temporally random phase destroying collisions. Using this scheme, the switching characteristics of the electron wave directional coupler is studied.

The rapid progress in semiconductor fabrication technology over the last few years has resulted in significant improvements in material characteristics and lithography. As a result, it has become possible to fabricate extremely small devices in the nanometer dimensions using very high mobility materials. In this regime of small dimensions and high mobilities, electrons begin to behave as waves, an entirely new physical phenomenon described by quantum physics. If suitably utilized, this feature can result in novel, efficient and fast electronic switches [1]. One such device that has been proposed and studied is the electron wave directional coupler switch [2, 3, 4]. This is the electronic analogy of the optical directional coupler. In this paper, a coupled mode theory for the analysis of this directional coupler device is presented [5]. In particular, the random dephasing effect has been taken into account in this model by including all phase destroying collisions into a single lumped parameter.

Figure 1 shows the top view of the proposed electron wave directional coupler device. It consists of two parallel quantum wires fabricated on a two-dimensional electron gas (2DEG). The 2DEG is created at the GaAs/AlGaAs hetero-junction interface of the semiconductor crystal. The quantum wires are defined by depositing patterned metal gates on the surface of the crystal and by applying a negative voltage to them, so that the electrons under the gates become depleted. This is shown on figure 2. Current is injected

0-7803-0894-8/93 $3.00

from one of the two inputs (Source 1 or Source 2) while the other input is grounded, and the outputs are taken from the drains (Drain 1 and Drain 2). The coupling between the two quantum wires causes an interference between the electron wavefunctions. This makes the electrons to switch back and forth between the two quantum wires in an oscillatory fashion along the length of the device producing a switching action. The strength of the coupling, and consequently the switched current, can be controlled by the potential barrier V_{g3} between the two wires. The conduction band model of this device is shown on figure 3.

The z direction will be taken to be along the length of the device (direction of propagation), and the x direction will be taken to be along the width of the device (direction of confinement). If the ground state wave functions in the individual quantum wires are represented as $\psi_1(x)$ and $\psi_2(x)$, then the wavefunction of the coupled system can be expressed as a linear superposition of these wavefunctions, such as

$$\psi(x) = \sum_j^2 a_j(z)\,\psi_j(x).$$

The effective mass Schrödinger wave equation that governs the motion of electrons in a two-dimensional system is

$$-\frac{\hbar^2}{2}\left[\frac{\partial}{\partial x}\left(\frac{1}{m^*}\frac{\partial}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{1}{m^*}\frac{\partial}{\partial z}\right)\right]\psi = (E - E_c(x))\,\psi.$$

By substituting $\psi(x)$ into this wave equation, the following set of coupled equations can be derived:

$$\mathbf{X}\frac{\partial^2 \mathbf{a}(z)}{\partial z^2} = \mathbf{M}\mathbf{a}(z),$$

where $\mathbf{X}$ is the overlap matrix and $\mathbf{M}$ is the coupling matrix [5]. Further, $\mathbf{M}$ can be shown to be hermitian, which makes the solution of this equation easier to obtain.

There are well established techniques for solving this system of equations [6]. When solved, the amplitudes of the mutually orthogonal states of the coupled system, $W_1(z)$ and $W_2(z)$ can be obtained. The switching action in the device can be explained as due to the interference pattern between these two orthogonal states. However, due to temporally random dephasing collisions with phonons and other electrons, this interference pattern will not be persistent for indefinite z distances. In fact, the interference effect can be shown to decay with a characteristic *mean coherence length* of L_c. This L_c is a quantity that includes all the different phase destroying collisions in the material as a single lumped parameter.

From the quantum mechanical expression for electron current

$$J = \frac{j\hbar e}{2m^*}\left(\psi^* \nabla\psi - \psi \nabla\psi^*\right),$$

the current density in the device can be expressed as

$$J = \sum_{ij}^{2} J_{ij}(z)\,\psi_i(x)\,\psi_j(x)$$

where $J_{ij}(z)$ is the current density matrix element. Further, $J_{ij}(z)$ can be split into

$$J_{ij}(z) = J_{ij}^M + J_{ij}^I(z).$$

J_{ij}^M contains the term $|W_n(z)|^2$ and $J_{ij}^I(z)$ contains $W_n(z)\,W_m^*(z)$. Therefore, J_{ij}^M can be thought of as representing the magnitude of the current density matrix element, and $J_{ij}^I(z)$ as representing its spatially dependent interference term. The dephasing effect can now be introduced into the interference term $J_{ij}^I(z)$ as an exponentially decaying factor, such as

$$J_{ij}^I \sim e^{-z/L_c}.$$

The implication of this statement is that, after a distance equal to many times the mean coherence length L_c, the interference effects in the device will completely cancel out due to these phase destroying collisions. This is apparent from the fact that $J_{ij}^I \to 0$ as z gets large. Therefore, to observe the switching effect, it is desirable to have the device dimensions within the same order of magnitude as the mean coherence length L_c.

As an example, applying this model to the quantum directional coupler switch shown on figure 1 consisting of two 60nm quantum wires separated by a distance of 100nm each with a confining potential of 1mV, the current density distribution in the device will be as shown on figures 4 and 5. The energy of the incoming electrons was assumed to be 1meV. As clearly seen from figure 5, the effect of the shorter dephasing length is a heavily damped interference pattern that degrades the switching performance. Figures 6 and 7 show the total current in each quantum wire as a function of the longitudinal distance z for $V_a = 0$ and $V_a = 0.33$mV, where V_a is a modulation voltage superimposed on the 1mV gate voltage, such that $V_{g3} = 1\text{mV} - V_a$. It can be seen that, switching action can be realized with $V_a = 0.33$mV (or for a barrier voltage of about 0.66mV) for a device length of about 4.8μm. On figure 8, the I-V characteristics of this device is shown. As expected, the shorter dephasing length results in a smaller, or damped, switching effect.

In conclusion, a coupled mode theory for electron wave directional couplers has been developed. This model takes the dephasing effect on electrons into account by lumping all the phase breaking collisions in the material into a single mean coherence length L_c.

References

[1] Cristopher C. Euster, Peter R. Nuytkens, and Jesús A. del Alamo. A novel analog-to-digital conversion architecture using electron waveguides. *IEDM*, pages 495–498, 1992.

[2] Nadir Dagli, Gregory Snider, Jonathan Waldman, and Evelyn Hu. An electron wave directional coupler and its analysis. *J. Appl. Phys.*, 69(2):1047–1051, 1991.

[3] Jesús A. del Alamo and Cristopher C. Euster. Quantum field-effect directional coupler. *Appl. Phys. Lett.*, 56(1):78–80, 1989.

[4] Mason Thomas, Nadir Dagli, Jonathan Waldman, and Arthur Gossard. Zero gap electron waveguide coupler. *Integrated Photonics Research, Proc.*, 10:103–105, 1993.

[5] Andrew M. Sarangan and Weiping Huang. A coupled mode theory for quantum electron waveguides. *IEEE J. Quantum Elec.*, 1993. manuscript submitted.

[6] H.A. Haus, W.P. Huang, S. Kawakami, and N.A. Whitaker. Coupled-mode theory of optical waveguides. *J. Lightwave Tech.*, 5(1):16–23, 1987.

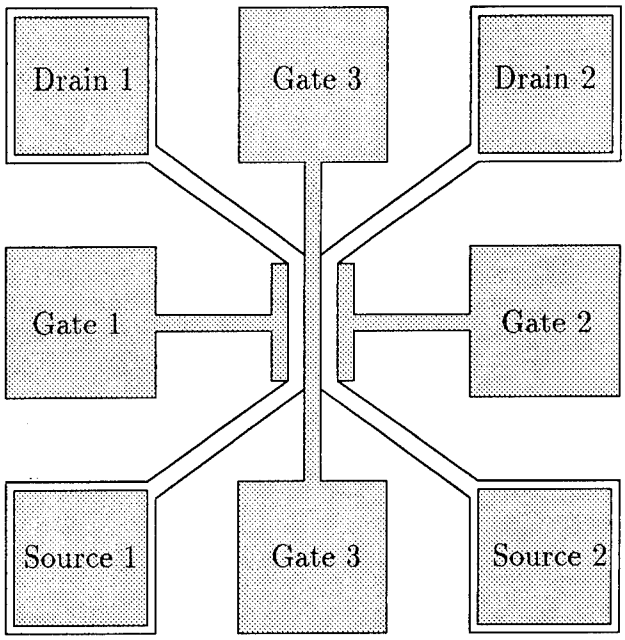


Figure 1: Top view of the quantum directional coupler [2].

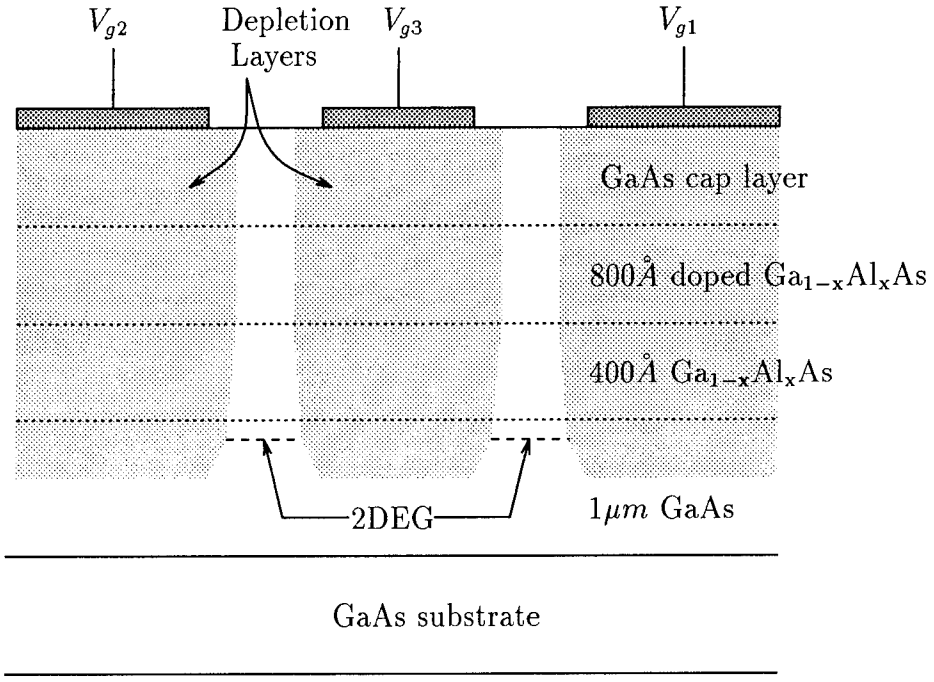


Figure 2: Cross sectional view of the quantum directional coupler [2].

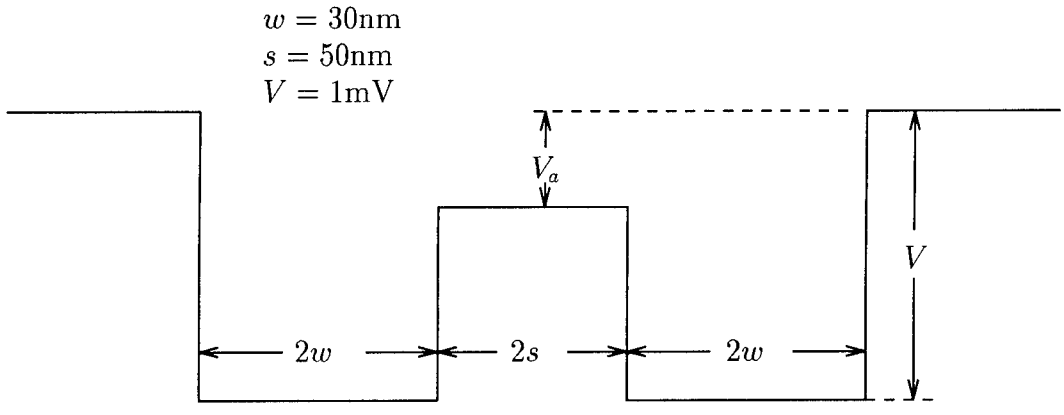


Figure 3: Conduction band model of the proposed quantum directional coupler.

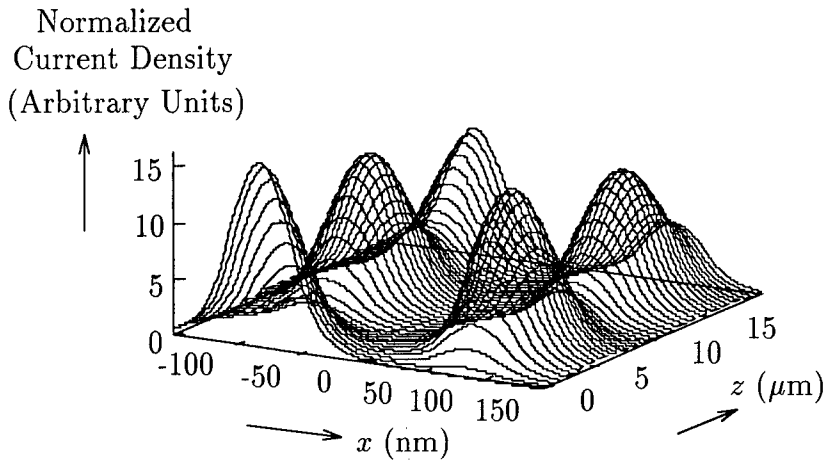


Figure 4: A three dimensional plot of the current density distribution in the lateral quantum directional coupler with $L_c = 20\mu$m and $V_a = 0$.

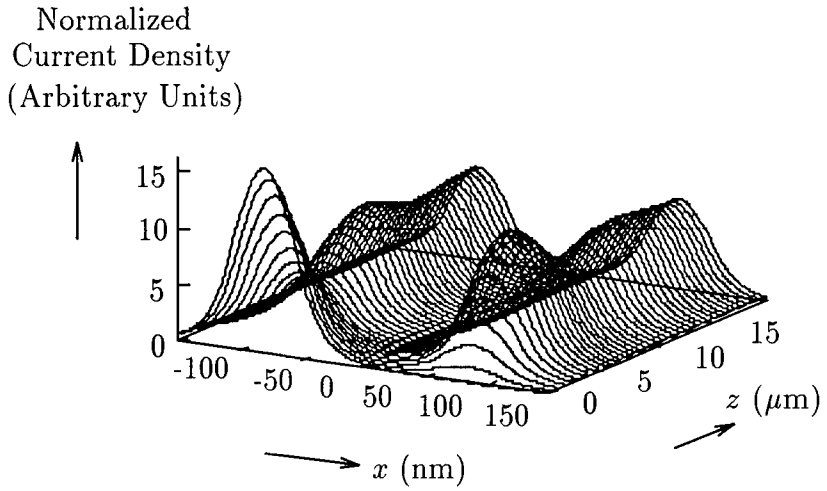


Figure 5: A three dimensional plot of the current density distribution in the lateral quantum directional coupler with $L_c = 5\mu$m and $V_a = 0$.

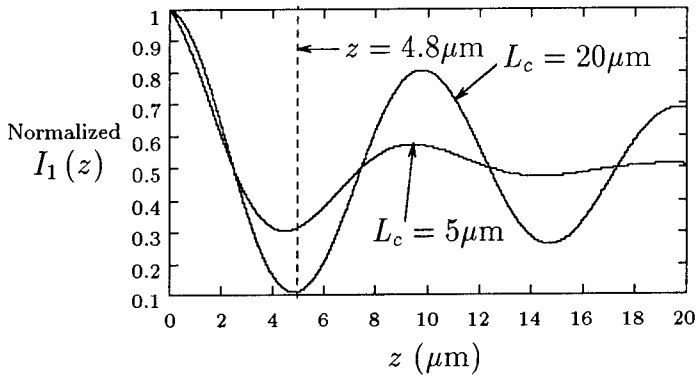


Figure 6: Current in the first wire, $I_1(z)$, versus the distance z along the device for both coherence lengths of $L_c = 20\mu$m and $L_c = 5\mu$m with an applied voltage $V_a = 0$.

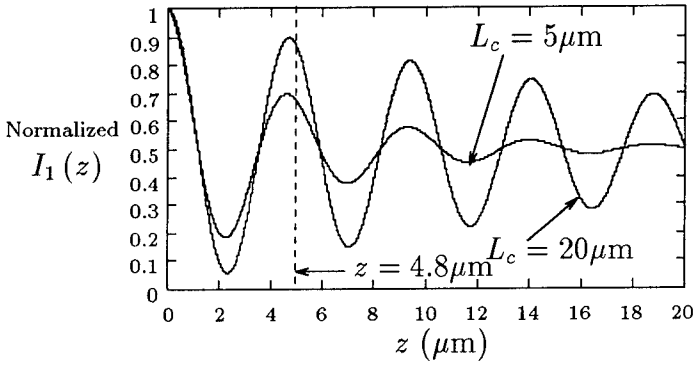


Figure 7: Current in the first wire, $I_1(z)$, versus the distance z along the device for both coherence lengths of $L_c = 20\mu$m and $L_c = 5\mu$m with an applied voltage of $V_a = 0.33$mV.

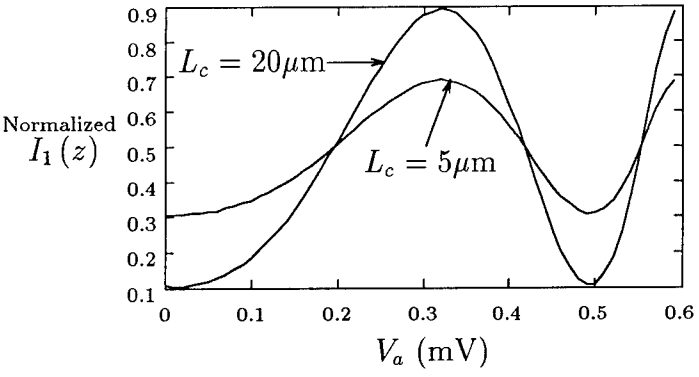


Figure 8: Current in the first wire, $I_1(z)$ at $z = 4.8\mu$m, for $L_c = 20\mu$m and $L_c = 5\mu$m with increasing V_a

Study of Shot Noise in a Double Barrier Resonant Tunneling Structure

M. M. Jahan and A. F. M. Anwar

Electrical and Systems Engineering Department
The University of Connecticut
Storrs, CT 06269-3157

ABSTRACT

Shot noise is calculated in a Double Barrier Resonant Tunneling Structure (DBRTS) by taking the space charge accumulated inside the quantum well into account. The calculation is self-consistent in nature and is obtained by simultaneously solving the Schrödinger and Poisson's equations. The calculation manifests the suppression of shot noise in the positive differential resistance (PDR) region and an enhancement in the negative differential resistance region (NDR) of the DBRTS. The behavior is explained in terms of the fluctuation of the eigen energy of the structure due to the stored charge in the quantum well.

INTRODUCTION

Shot noise [1-3] is observed in almost all electronic devices and is due to the fluctuations in device current. The fluctuations are attributed to the randomness of the carrier flowing through the device. In the resonant tunneling device, the current is a very strong function of eigen energies and thus current gets modulated by any parameter which changes eigen energies. One such parameter is space charge developed inside the quantum well. In devices where eigen energies are dependent on current, significant deviations from full shot noise may occur. As the current affects the eigen energies through the space charge stored in the quantum well [4], a significant deviation from the full shot noise behavior is expected in the quantum well device. The experimental observation of the suppression as well as enhancement of shot noise in a double barrier resonant tunneling structure has already been reported [1]. However, in order to appreciate the phenomena theoretically, one has to perform the calculation of shot noise self-consistently. Brown et. al. [1] has presented a theoretical analysis of the suppression and enhancement of shot noise without taking self-consistency into account. In this paper, the analysis of shot noise is extended to incorporate the effect of space charge accumulated inside the quantum well self-consistently. As electron traverses through a DBRTS, it spends a finite amount of time inside the structure and thus the space charge gets accumulated in the quantum well. The magnitude of space charge contributed by a traversing electron is a function of the eigen energies. The contribution of space charge due to an electron is also affected by

0-7803-0894-8/93 $3.00

the presence of other electrons. This gives rise to two different types of correlations, auto-correlation that depends on the incident energy and cross-correlation that includes the effect of the space charge due to other eigen energies. To make a self-consistent calculation of shot noise, these correlation terms need to be computed by solving the Schrödinger and Poisson's equation simultaneously. In this paper, these correlation terms are computed by calculating space charge for every allowed energy level separately and then evaluating the dependence among different eigen energies that arises through the space charge. Calculated shot noise shows a suppression in the PDR ($\gamma = 0.91$) and enhancement in the NDR ($\gamma = 50$).

THEORY

The self-consistent calculations are performed by using the logarithmic derivative of wavefunction. Once the logarithmic derivative $\Xi(x, E_x)$ [4] is obtained as a function of distance x along the structure and the incident electron energy E_x, the group velocity $v_g(x, E_x)$ [5] is calculated as

$$v_g(x, E_x) = \frac{1}{2} Re[\Xi(x, E_x)]. \tag{1}$$

The determination of the current density at every energy E_x, $J(E_x)$, enables the evaluation of the space charge $n(x, E_x)$

$$n(x, E_x) = \frac{J(E_x)}{v_g(x, E_x)}. \tag{2}$$

Space charge is calculated for each allowed eigen energy and the contribution of the space charge to the potential profile is determined by solving the Poisson's equation. Accordingly, the potential profile is modified and the effect of the modified potential profile on the other energy levels and the level itself is determined by calculating $J(E_x)$'s [4] again and comparing the new current density with the one calculated before.

$$\Delta J^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} [\Delta J(E_x^i, E_x^j)]^2 \tag{3}$$

where, $\Delta J(E_x^i, E_x^J)$ is the difference of current density corresponding to energy level E_x^i due to the space charge accumulated for energy level E_x^j, and N is the total number of allowed eigen energies. ΔJ is used to calculate shot noise in the structure.

Shot noise power spectrum S is obtained as

$$S = \frac{\Delta J^2}{\delta f} \tag{4}$$

where, δf is the measurement bandwidth. Shot noise factor γ is calculated as

$$\gamma = \frac{S}{2qJ}, \tag{5}$$

where, J is the total current density [4], and q is the electronic charge, 1.6×10^{-19} Coulomb.

RESULTS AND DISCUSSIONS

Results for a symmetric DBRTS are presented in this section. The barrier and the quantum well are each 50Åwide. The barrier height is assumed to be $0.275eV$. The effective mass of the electron is $0.067m_0$ and $0.096m_0$ in the quantum well and in the barrier, respectively and m_0 is the electron rest mass, $9.1 \times 10^{-31} Kg$. In the calculation , the fermi level E_F and temperature T are assumed to be $0.03eV$ and 4.2 degree Kelvin, respectively.

In Figs 1 and 2, theoretical current voltage characteristics and shot noise factor as a function of bias voltage are plotted, respectively. As expected, current starts increasing with bias voltage in the positive differential resistance region (PDR) followed by the current peak at $0.12Volt$. A further increase of bias voltage causes a reduction of current (NDR) until it reaches the valley current at $0.18Volt$. Thereafter, current continues to increase. The rise in current beyond the valley is caused by the onset of current flow through the second quasibound level in the quantum well. Fig.2 shows a suppression of shot noise factor in the PDR and an enhancement in the NDR. The suppression and enhancement can be explained in terms of the behavior of quasibound level of the quantum well. As space charge gets accumulated in the quantum well, it will reduce the effective voltage across the well. As a result, the quasibound level in the quantum well moves up along the energy axis. However, it leads to a decrease in the total current in the PDR as the quasibound level is moving away from the fermi level. On the other hand, in the NDR, as the quasibound state moves up along the energy axis due to space charge, it moves towards the fermi level and hence total current will be increased. More current implies more shot noise and greater shot noise factor and vice versa. This explains the suppression and the enhancement of shot noise factor in the PDR and NDR, respectively.

Fig.3 shows the space charge that is stored in the quantum well as a function of bias voltage. It is to be noted that the peak in the space charge distribution corresponds to the current peak. Thus, the greatest contribution to the shift of quasibound level due to space charge occurs when the space charge becomes the maximum ($0.12Volt$). Since, this voltage is situated just at the onset of the negative resistance region, eigen energies are shifted up by a considerable amount and again there are relatively few electrons available to tunnel through causing a reduction of shot noise factor. From Fig.2, it is also noted that the enhancement of shot noise gets peaked not at the voltage corresponding to peak current rather at a higher voltage. It can be explained in the following manner. The current becomes maximum when the first eigen energy is aligned with the fermi level of the emitter. Moreover, the space charge moves the eigen energy up along the energy axis. Thus, effectively, eigen energy will be aligned when device is operated in the negative resistance region. This explains why shot noise becomes the maximum in the negative resistance region not at the onset of negative resistance region.

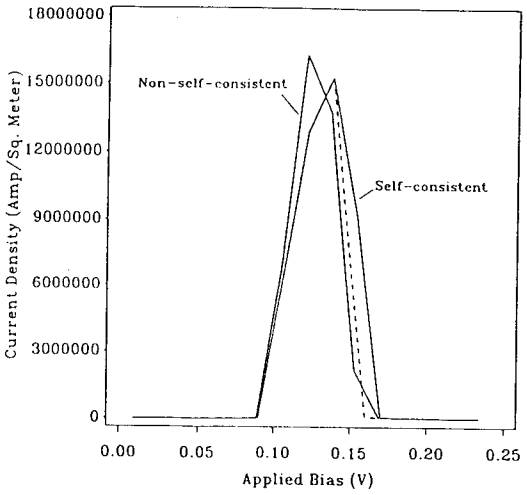


Figure 1: Current-voltage characteristics of a DBRTS.

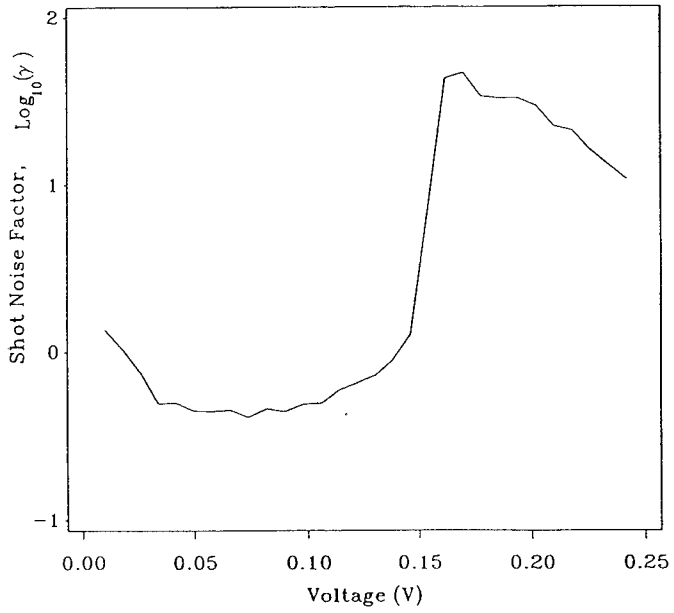
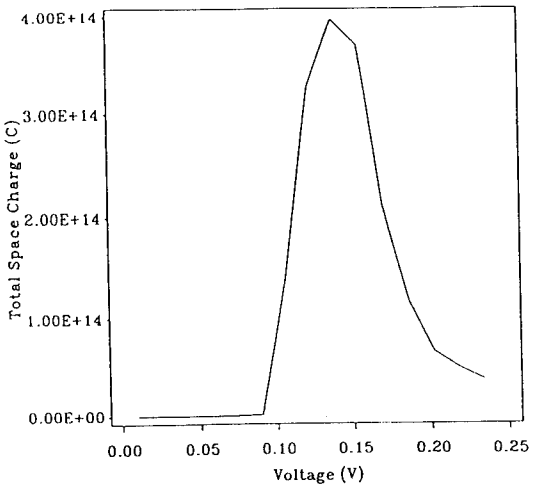


Figure 2: Shot noise factor is plotted as a function of bias voltage.



CONCLUSION

The computation of shot noise shows suppression in the PDR and enhancement in the NDR of a DBRTS. The observation agrees qualitatively with the experimental results [1]. The two opposite behaviors are explained in terms of space charge stored in the quantum well.

REFERENCES

1. E. R. Brown, "Analytic model of shot noise in double barrier resonant tunneling diodes," Dev. Res. Conf., Colorado, 1991.

2. J. Han and F. S. Barnes, " A theory of shot noise in quantum wells and applications in resonant tunneling heterojunction bipolar transistors," IEEE Transaction on Electron Devices, vol. ED-38, no 2. , p.237, 1991.

3. Y. P. Li, A. Zaslavsky, D.C. Tsui, M. Santos, and M. Shayegan," Noise characteristics of double-barrier resonant-tunneling structures below 10 kHz," Phys. Rev. B, vol. 41, p.8388, 1990.

4. M. M. Jahan and A.F.M. Anwar, " Space charge buildup and bistability in double barrier resonant tunneling structures in the presence of a transverse magnetic field," Submitted to Journal of Applied Physics.

5. A. F. M. Anwar, A. N. Khondker and M. R. Khan, " Calculation of traversal time in resonant tunneling devices," J. Appl. Phys. 65(7), p. 2761, 1989.

Structure and Simulation of GaAs TUNNETT and MITATT Devices for Frequencies Above 100 GHz[†]

Chien-Chung Chen, Richard K. Mains, George I. Haddad and Heribert Eisele

Solid-State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, Michigan 48109-2122

Abstract

A numerical simulation program for two-terminal transit-time devices based on the energy-momentum transport model, with valence band to conduction band tunneling phenomena incorporated, has been developed. This program can deliver accurate TUNNETT and MITATT device simulation results in the millimeter and submillimeter range, and therefore provides a useful tool for high frequency device structure design and optimization.

Simulation results for GaAs TUNNETT and MITATT devices for frequencies above 100 GHz will be presented. As simulation results show, the negative resistance of the device decreases rapidly as the operating frequency increases. Under such circumstances, the contact resistance severely degrades the device's RF performance. When a diode's negative resistance becomes lower than the contact resistance, no RF power can be generated. To overcome this difficulty, device structures using no ohmic contacts are investigated. In these devices, Ohmic contacts are replaced by Schottky contacts, and also p-n junctions are replaced by Schottky junctions for single-drift structures. Since the metal-semiconductor contact resistance is eliminated or greatly reduced in such devices, they are very promising as RF power sources at extremely high frequencies.

1 Introduction

Tunnel injection transit-time (TUNNETT) and mixed tunneling and avalanche transit-time (MITATT) diodes are two potentially useful RF power sources at extremely high frequencies[1]. The valence band to conduction band tunneling by electrons in such devices has a profound effect on the device performance. In the past, many numerical simulation programs have been developed to simulate transit-time devices, mostly for simulation of impact ionization avalanche transit-time (IMPATT) diodes[2, 3]. The free charge carrier transport models used in such simulations are the drift-diffusion model and energy-momentum model[4]. Programs based on the drift-diffusion model are more popular because of its simplicity. The energy-momentum model, due to the large number of partial differential equations to solve, is less frequently adopted in simulation.

Simulation of the interband tunneling process is still rare today. Recently Dash and Pati developed a generalized method to include the tunneling mechanism in their simulation[5]. Their simulation was based on the simple drift model (no diffusion currents), and was limited to DC or small-signal analysis. And their simulation lacks verification with measurement data for real devices.

We developed the large-signal energy-momentum program to simulate TUNNETT and MITATT diodes. Tunneling currents are modeled and incorporated in the transport models. A real W-band diode

[†]This work was supported by the Army Research Office under contract No. DAAL03-92-G-0109 and Navy under contract N60530-92-M-E959.

is simulated using both models, and simulation results agree with experimental measurements[6].

From previous simulation work, it is known that the series resistance of two-terminal transit-time devices seriously degrades the RF performance especially at high frequencies[6]. First of all, it consumes part of the RF power generated by the device. Second, it reduces the device's efficiency. Third, it counteracts a device's negative resistance, rendering a more difficult matching situation. When the negative resistance gets small, the diode's cross-sectional area has to be reduced for matching to a specific load, and therefore the generated RF power is reduced. Furthermore, if a device's internal negative resistance is smaller than the series resistance, no RF power can be delivered to the external load. To overcome such difficulties, Schottky contact TUNNETT/MITATT's are currently considered for RF power generation above 100GHz. In such devices, the Ohmic contacts are replaced with Schottky contacts, which, when depleted, impose no potential barrier to electrons or holes except a small probability of quantum reflection as shown in Figure 1. Outflowing carriers near the Schottky contact regions are pushed by the electric field toward the contacts and collected. Little contact resistance is expected in such devices.

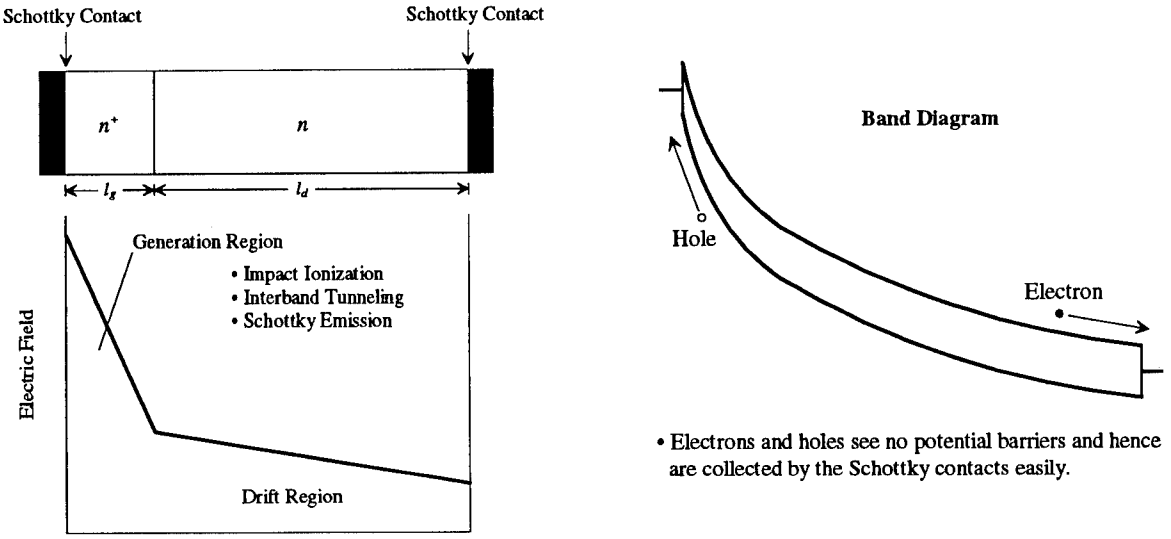


Figure 1: The operating principle of Schottky contact TUNNETT/MITATT devices.

In this work, four GaAs Schottky contact TUNNETT/MITATT devices operating from 170 GHz to 370 GHz are simulated using the energy-momentum transport model. In the simulation, three major carrier generation mechanisms are considered, i.e. *impact ionization, interband tunneling* and *metal-semiconductor tunneling.*

2 Transport Theory

The two most commonly used transport models for semiconductor device simulation are the drift-diffusion (DD) model and the energy-momentum (EM) model. In the DD model, charge carriers respond immediately to the external electric field and reach steady state without any delay in time, while the energy-momentum model requires relaxation times for carriers to adjust their energy and momentum before reaching steady state. Due to this transient nature of the EM model, free carriers exhibit velocity overshoot and undershoot effects in response to the change of electric field. This kind of transient response

can also be observed in Monte Carlo simulations. By using correct relaxation times, the EM model can depict carrier behavior in remarkable agreement with Monte Carlo observations[7].

In the EM simulation program developed by us, holes still follow the drift-diffusion equation, while electrons are described by a variation of the two-valley drift-diffusion equation with the mobilities and diffusion coefficients being functions of electron energies instead of the electric field. Two energy equations govern the electron energies. The following equations are solved by the EM simulation program:

$$J_p = qp\mu_p(E)E - qD_p(E)\frac{\partial p}{\partial x}$$

$$J_{n1} = qn_1\mu_{n_1}(w_1)E + qD_{n_1}(w_1)\frac{\partial n1}{\partial x}$$

$$J_{n2} = qn_2\mu_{n_2}(w_2)E + qD_{n_2}(w_2)\frac{\partial n2}{\partial x}$$

$$\frac{\partial w_1}{\partial t} = -\frac{J_{n1}E}{n_1} - \frac{w_1 - w_{th}}{\tau_{w1}(w1)}$$

$$\frac{\partial w_2}{\partial t} = -\frac{J_{n2}E}{n_2} - \frac{w_2 - w_{th}}{\tau_{w2}(w2)}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + G_p$$

$$\frac{\partial n1}{\partial t} = \frac{1}{q}\frac{\partial J_{n1}}{\partial x} + G_{n1}$$

$$\frac{\partial n2}{\partial t} = \frac{1}{q}\frac{\partial J_{n2}}{\partial x} + G_{n2}$$

where $w_{th} = \frac{3}{2}k_BT$ is the thermal equilibrium energy for electrons, and G_p, G_{n1} and G_{n2} are given by

$$G_p = G_{II} + G_T + G_{th} - R_{th}$$

$$G_{n1} = G_{II} + G_T + G_{th} - R_{th} + \frac{n2}{\tau_{n_{2\to1}}(w2)} - \frac{n1}{\tau_{n_{1\to2}}(w1)}$$

$$G_{n2} = \frac{n1}{\tau_{n_{1\to2}}(w1)} - \frac{n2}{\tau_{n_{2\to1}}(w2)}$$

Close examination of the above equations reveals that the momentum relaxation times for electrons are actually neglected in the program. This is justified by the fact that the momentum relaxation times are generally much smaller than the energy relaxation times.

Metal-semiconductor tunneling at Schottky contacts is modeled as part of the current boundary condition. Every time the current boundary condition is to be applied, the program evaluates terminal inflow currents by using the contact electric fields and an electric field–current table for Schottky contacts.

3 Material Parameters

The material used for devices in this work is GaAs. Since these devices are intended for RF power generation, the material lattice temperature is usually very high. 500°K is assumed throughout this work. The majority of the material parameters used by the simulation programs are generated by a traditional Monte Carlo (MC) program[8]. This MC program provides the values of electron velocities, diffusion

coefficients, relaxation times and electron energies as functions of the electric field up to 3000 kV/cm. The methods used to evaluate the relaxation times are similar to those proposed by Stewart *et. al.*[7], while the electron (longitudinal) diffusion coefficients are calculated in the MC program using the method mentioned by Fawcett[10]. Other electron parameters such as the material intrinsic concentration and electron-hole recombination lifetime and all the hole parameters are set according to empirical expressions or assumed values. Figure 2 shows the various time constants, electron velocities and diffusion coefficients obtained from Monte Carlo for GaAs at 500°K, assuming the impurity doping is $5 \cdot 10^{17}$ cm^{-3}.

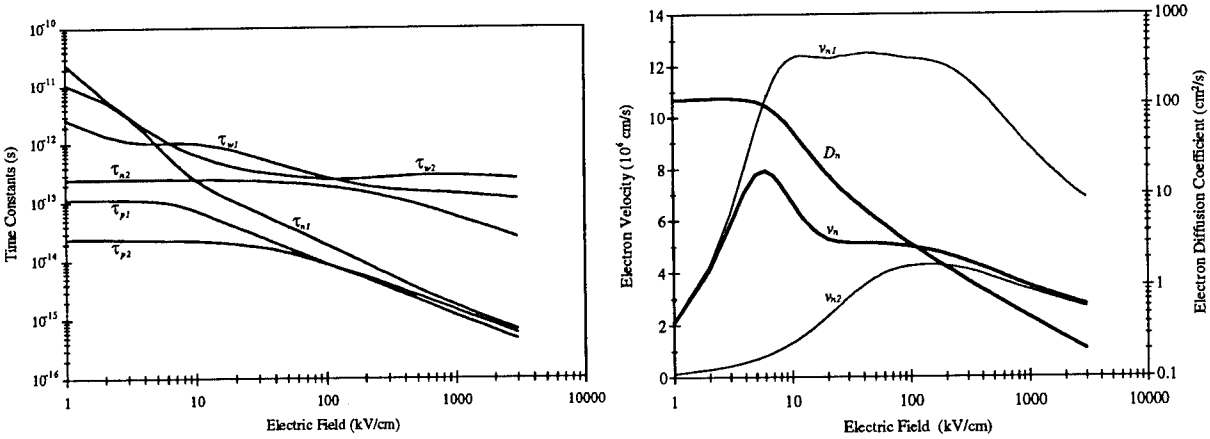


Figure 2: Time constants, electron velocities and diffusion coefficients for GaAs at 500°K used in simulation.

Two important material parameters used by the simulation programs are the impact ionization and tunneling rates. They have very profound effects on the device operation, but still lack adequate determination in the electric field range where the devices are operated (up to 3000 kV/cm). The impact ionization generation rate is expressed as

$$G_{II} = (\alpha_n J_n + \alpha_p J_p)/q$$

with α_n and α_p being given by

$$\alpha_n = A_n \exp\left[-(B_n/E)^2\right]$$

$$\alpha_p = A_p \exp\left[-(B_p/E)^2\right]$$

For GaAs at 500°K, we use

$$A_n = A_p = 2.1205 \times 10^5 \quad \text{cm}^{-1}$$

$$B_n = B_p = 6.71 \times 10^5 \quad \text{V/cm}$$

As to the interband tunneling rate, we adopt the expression which complies with the highly idealized form proposed by Kane[11] as shown below

$$G_T = A_T E^2 \exp(-B_T/E)$$

For GaAs at 500°K, we use the following values for A_T and B_T

$$A_T = 1 \times 10^{20} \quad \text{cm}^{-1}\text{s}^{-1}\text{V}^{-2}$$

$$B_T = 1.2 \times 10^7 \quad \text{V/cm}$$

The dead space for impact ionization and interband tunneling is considered in the simulation program. For impact ionization, the energy relaxation effect automatically includes the dead space in the model. As to interband tunneling, we use the bandgap to determine the dead space as illustrated in Figure 3. The principle is that tunneling only occurs where there are empty states in the conduction band and filled states in the valance band.

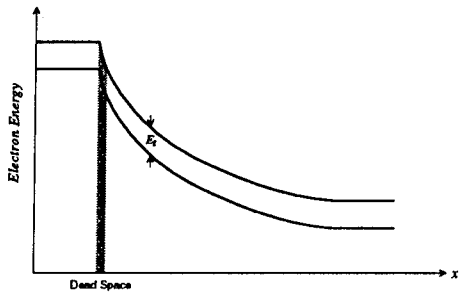


Figure 3: Dead space for interband tunneling.

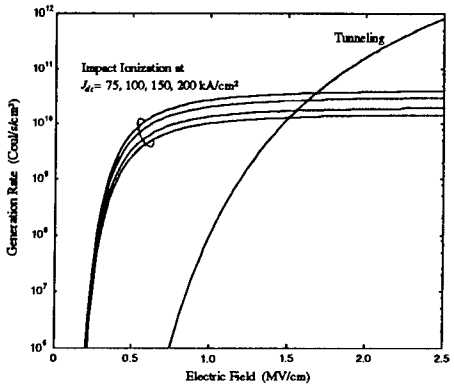


Figure 4: The impact ionization and tunneling generation rates for GaAs at 500°K.

In the EM program, all material parameters for electrons except the tunneling rate are tabulated according to the electron energy. In other words, they are treated as functions of the electron energy instead of the electric field. The tunneling rate is considered as a function of the electric field since it is related more to the electric field that electrons in the valence band experience than to the energy of electrons in the conduction band. Figure 4 shows the impact ionization and tunneling generation rates for 500°K GaAs that are used in the simulation programs. Inside the simulation program, the impact ionization rate is transformed to be a function of electron energy.

4 Device Structure

Four GaAs Schottky contact TUNNETT/MITATT's are investigated in this work. These four diode structures to be simulated have not been optimized for maximum RF power generation.

Diode	l_g (nm)	l_d (nm)	N_{d1} (cm^{-3})	N_{d2} (cm^{-3})	J_{dc} (kA/cm^2)
D1	33.75	166.5	$4 \cdot 10^{18}$	$1.3 \cdot 10^{17}$	75
D2	33.25	115.5	$4 \cdot 10^{18}$	$2 \cdot 10^{17}$	100
D3	33.75	67.5	$4 \cdot 10^{18}$	$4 \cdot 10^{17}$	150
D4	37	37	$4 \cdot 10^{18}$	$7 \cdot 10^{17}$	200

Table 1: Device Structures under simulation.

Table 1 shows the diode structures, where l_g is the length of the heavily doped generation region with doping N_{d1}, l_d is the length of the drift region with doping N_{d2}, and J_{dc} is the DC bias current for each structure. All the devices have the same generation region doping ($N_{d1} = 4 \cdot 10^{18}$ cm^{-3}) since we want the doping to be as high as possible to make l_g short while $N_{d1} = 4 \cdot 10^{18}$ cm^{-3} is close to the

maximum doping achievable by MBE. When the operating frequency gets higher, we apply higher DC bias current in order to increase the negative resistance.

To understand the need of high bias current for high negative resistance, we write down the negative resistance of the diode as

$$-R = \frac{-G}{G^2 + B^2} \tag{1}$$

where B and $-G$ are the diode's susceptance and negative conductance, respectively. At frequencies as high as 300 GHz, B is usually dominated by the diode's capacitance C and is much higher than $-G$. Therefore the diode's negative resistance is approximated by

$$-R \approx \frac{-G}{\omega^2 C^2} \tag{2}$$

In the above expression for negative resistance, the capacitance is essentially independent of the bias current density. To increase $-R$, we may try to increase $-G$. Increasing the bias current density has the effect of increasing the negative conductance since at higher current density the same amount of V_{rf} causes greater AC current.

Also note the doping in the drift region, N_{d2}, must be high enough to accommodate the space charge effect. Therefore, N_{d2} gets higher in higher frequency devices. The generation region length, l_g, is determined by making the electric field at the junction of the generation region and drift region $400 \sim 500$ kV/cm to suppress carrier generation in the drift region.

The labels D1, D2, D3 and D4 refer to diodes with $J_{dc} = 75$, 100, 150 and 200 kA/cm^2, respectively. This labeling will be used consistently during the report.

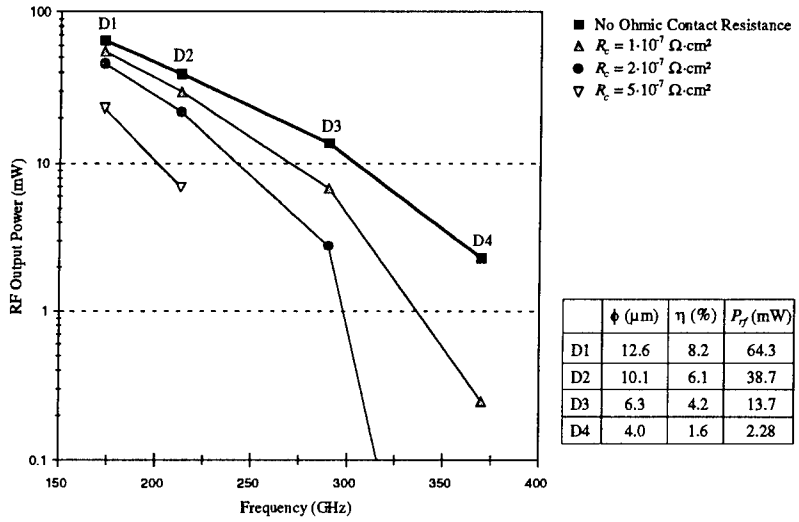
5 Simulation Results and Discussion

The simulation can be divided into two parts: the DC and AC simulations. DC simulation is necessary to provide the initial condition to AC simulation. Table 2 shows the bias conditions calculated from the DC simulation. From it, one can see that impact ionization, instead of interband tunneling, dominates the charge generation process even in high frequency diodes. Two factors contribute to the dominance of impact ionization — one is the high bias current density, which enhances only the impact ionization; the other is the significant generation region expansion phenomenon for impact ionization at high frequencies[6].

Diode	Bias Voltage (V)	Bias Current (kA/cm^2)	Tunneling Current (kA/cm^2)
D1	10.06	75	32.5
D2	8.89	100	38.2
D3	7.52	150	54.9
D4	6.37	200	83.0

Table 2: Device DC bias conditions calculated from the EM DC simulation.

Note that diode D4 has slightly higher tunneling current than the other diodes. This is because D4's drift region is designed to have lower electric field so that electrons in the drift region can have higher drift velocity. The low electric field in the drift region also makes the generation region expansion effect less pronouced in D4, which leads to reduced contribution from impact ionization.



	ϕ (μm)	η (%)	P_{rf} (mW)
D1	12.6	8.2	64.3
D2	10.1	6.1	38.7
D3	6.3	4.2	13.7
D4	4.0	1.6	2.28

Figure 5: Energy-momentum AC simulation results for GaAs Schottky contact devices matched to 1 Ω.

AC simulation results are summarized in Figure 5, in which the external load is assumed to be 1 Ω. The diode diameters are determined by matching the impedance to 1 Ω. V_{rf} is swept in small steps to find the maximum output power. Simulations are also carried out over frequency range to search for the optimum operating frequency. The figure also shows the RF output power if there existed contact resistance as is the case in conventional TUNNETT/MITATT devices which employ Ohmic contacts. From it, one can see that the series contact resistance severely reduces the device performance at high frequencies because of the decreased device negative resistance. It is obvious that Schottky contact TUNNETT/MITATT devices have superior RF performance at high frequencies compared with conventional structures.

The low negative resistance at high frequencies is mainly due to the device capacitance, whose effect on the negative resistance can be seen in Equation 2. Since the capacitance is proportional to the operating frequency that the device is designed for, it turns out that the relationship of negative resistance and frequency is given by

$$-R \propto \frac{1}{f^4} \tag{3}$$

To increase the negative resistance $-R$, the following measures can be considered:

- Use high drift velocity materials such as InP to increase the drift region length, and therefore decrease the device capacitance.
- Design structures with lower electric field in the drift region to have higher electron velocity. This again can lead to longer drift region length.
- Adopt the double-drift instead of single-drift structures to achieve longer device length. However, tunneling current is lower in such devices since the generation region length is also longer. The operating frequency will decrease.
- Cool the device in liquid nitrogen to increase the carrier velocity. Nevertheless, the tunneling current is suppressed at low temperature.

- Bias the device with higher current to increase $-R$ as explained in the preceding discussion. Note that the bias current is limited by the heat dissipation capability.
- Use low bandgap materials such as InGsAs in the generation region to increase the negative conductance. InGaAs/InP heterojuction TUNNETT/MITATT design may be the direction since we get enhanced generation currents (especially the tunneling currents), high electron drift velocity and low diffusion currents in the drift region.

Future simulation work will be to design optimized structures using the developed simulation program, simulate devices with different materials such as InGaAs and InP, and simulate heterojuction structures.

6 Conclusion

Interband tunneling has been included in the energy-momentum transport model program to simulate operation of TUNNETT and MITATT diodes. Also the tunneling currents at Schottky contacts are modeled in the current boundary condition so that Schottky contact TUNNETT/MITATT's can be simulated.

At high frequencies, conventional TUNNETT and MITATT diodes' performance is severely limited by the contact resistance. Schottky contact TUNNETT/MITATT's, which exhibit much lower contact resistance than those having Ohmic contacts, are promising as RF power sources at frequencies well above 100 GHz.

References

[1] G. I. Haddad *et. al.*, "Tunnel transit-time (TUNNETT) devices for terahertz sources," *Microw. Opt. Technol. Lett.*, Vol. 4, No. 1, pp. 23–29, 1991.

[2] P. E. Bauhahn and G. I. Haddad, "IMPATT device simulation and properties," *IEEE Trans. Electron Devices*, vol. 24, No. 6, pp. 634–642, 1977.

[3] R. K. Mains *et. al.*, "Simulation of GaAs IMPATT diodes including energy and velocity transport equations," *IEEE Trans. Electron Devices*, Vol. 30, No. 6, pp. 1327–1338, 1983.

[4] K. Bløtekjær "Transport Equations for Electrons in Two-Valley Semiconductors," *IEEE Trans. Electron Devices*, Vol. 17, pp. 38–47, 1970.

[5] G. N. Dash and S. P. Pati, "A generalized simulation method for MITATT-mode operation and studies on the influence of tunnel current on IMPATT properties," *Semicond. Sci. Technol.*, Vol. 7, pp. 222–230, 1992.

[6] C. C. Chen *et. al.*, "Numerical simulation of TUNNETT and MITATT devices in the millimeter and submillimeter range," *Fourth International Symposium on Space Terahertz Technology*, pp. 362–376, 1993.

[7] R. A. Stewart *et. al.*, "Improved relaxation-time formulation of collision terms for two-band hydrodynamic models," *Solid-State Electronics*, Vol. 32, No. 6, pp. 497–502, 1989.

[8] C. Jacoboni and Paolo Lugli, "The Monte Carlo method for semiconductor device simulation," Springer-Verlag, Vienna, 1989.

[9] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev.*, Vol. 38, No. 14, pp. 9721–9745, 1988.

[10] W. Fawcett, "Non-ohmic transport in semiconductors," *Electrons in Crystalline Solids*, International Atomic Energy Agency, Vienna, pp. 531–618, 1973.

[11] E. O. Kane, "Zener tunneling in semiconductors," *J. Phys. Chem. Solid*, vol. 12, pp. 181–188, 1959.

OBSERVATION OF EXTREMELY LARGE SHEET HOLE DENSITIES IN UNCAPPED UNDOPED AlSb LAYERS

Bijan Tadayon, Carl S. Kyono, David J. Godbey, Saied Tadayon[a], and Jeffrey A. Mittereder[b]
Code 6856, Naval Research Laboratory, Washington, DC 20375
(a) COMSAT Laboratories, Clarksburg, MD 20871
(b) SFA, Landover, MD 20785

ABSTRACT

In this paper, we present a study of the electrical and physical properties of uncapped AlSb, grown by molecular beam epitaxy. Large concentrations of oxygen have been observed in the uncapped undoped AlSb layers, resulting in extremely large sheet hole densities (as high as 3 X 10^{15} cm^{-2} with the corresponding mobilities of about 60 $cm^2V^{-1}s^{-1}$). However, for the AlSb layers capped with 50 Å of GaSb, oxygen does not penetrate beyond the GaSb cap layer, and the underlying AlSb layer is highly resistive.

X-ray photoelectron spectroscopy (XPS) has been used to analyze the surface oxides for the uncapped AlSb layers. Three different oxides have been observed on the surface of the uncapped layers: Al_2O_3, Sb_2O_5, and Sb_2O_3.

I. INTRODUCTION

AlSb/InAs heterojunction system is finding increased use in various electronic structures, such as heterojunction field effect transistors and heterojunction bipolar transistors.[1-7] It is known that AlSb oxidizes very rapidly in air. To prevent oxidation, AlSb epilayers are usually capped with 50 to 100 Å of GaSb.[1-7] The electrical properties of the capped AlSb have

0-7803-0894-8/93 $3.00

been investigated extensively.[1-7] However, there has been no study on the properties of uncapped AlSb. In this paper, we present a study of the electrical and physical characteristics of uncapped AlSb.

II. EXPERIMENTAL

AlSb layers (0.3-1 μm thick) were grown by molecular beam epitaxy on semi-insulating (100) GaAs substrates, with a substrate temperature of 500 °C and a growth rate of 0.3 μm/h. The Sb (Sb_4) to Al beam equivalent pressure ratio was 30. Capped layers were also grown, with a thin (50-100 Å) GaSb layer deposited on top of the AlSb layer. Neither the capped nor the uncapped layers were intentionally doped. Typical background carrier concentration levels obtained in epitaxially grown GaAs were in the range of 10^{14}-10^{15} cm^{-3}, due to unintentional impurities. Similar background doping levels are expected for the AlSb layers.

The capped and uncapped samples were characterized by Hall and transmission line model (TLM) measurements. Conventional photolithographic and wet chemical etching techniques were used to make Hall and TLM structures on the samples. Before and after each wet chemical etch, HCl:H_2O (1:1) was used to remove surface oxide layers. In addition, immediately prior to loading of the samples into an e-beam evaporator, a low-power RF plasma clean in O_2 and a final HCl:H_2O (1:1) oxide strip were performed. Cr/Au (100 Å/2000 Å) contact metallization was then deposited on the samples. Hall and TLM measurements were taken before and after contact anneals.

X-ray photoelectron spectroscopy (XPS) was performed to identify the surface oxides present on the uncapped samples. Auger electron spectroscopy (AES) was used to determine the material composition of the capped and uncapped samples.

III. RESULTS AND DISCUSSION

For Hall and TLM measurements, the as-deposited Cr/Au metal contacts were non-linear and exhibited high contact resistance for both the capped and uncapped samples. After annealing the contacts, the capped layers were still highly resistive. In contrast, after annealing the uncapped layers at 300 °C (for 1 min. in forming gas), the contacts showed linear characteristics.[8] The uncapped layers were conductive and exhibited extremely large p-type conduction, with sheet hole density as high as 3×10^{15} cm^{-2} with the corresponding mobility of 60 $cm^2V^{-1}s^{-1}$.

A piece of a capped layer was etched to remove the GaSb cap and expose the AlSb layer to air. The sample became conductive with extremely large sheet hole density, similar to values obtained for the uncapped layers.

A piece of an uncapped layer was also etched to reduce the thickness of the AlSb layer from 1 μm to 0.5 μm. The AlSb layer was still conductive, with about the same sheet hole density and mobility as the original (thicker) uncapped layer.

AES measurements showed that oxygen in the air penetrated into the uncapped layers to a depth of more than 1400 Å (Figure 1). Near the surface, the oxygen concentration was more than 50 atomic percent. However, for the capped layers, oxygen did not penetrate beyond the (50 Å) GaSb cap layer (Figure 2), and the cap layer prevented the oxidation of the underlying AlSb layer.

Two XPS spectra from the surface of the same uncapped AlSb sample are shown in Figures 3 and 4. Figure 3 shows the aluminum 2p peak. Its position at 75.1 eV is consistent with the presence of Al_2O_3. Figure 4 shows the antimony 3d peaks. The pair of peaks at 531.5 and 540.8 eV indicate the presence of Sb^{5+} (Sb_2O_5). The pair of peaks at 529.1 and 538.5 eV are due to either Sb^{3+} (Sb_2O_3) or AlSb. Since the Al region shows

no contribution from an unoxidized species, it is concluded that these pair of peaks are only due to Sb^{3+} (Sb_2O_3). The relative intensities of the those peaks indicate that for every Sb^{3+} (Sb_2O_3), there are about two Sb^{5+} (Sb_2O_5) on the surface. Overall, three different oxides were observed on the surface of the uncapped layers: Al_2O_3, Sb_2O_5, and Sb_2O_3.

The oxidation process for the uncapped layers is very dependent on environment (for example, on room temperature and humidity). Therefore, the oxidation process may occur at different rates in different laboratories. After a significant amount of oxidation, stress in the layer causes the entire layer to peel off from the substrate upon probing. For our samples, the exposure of the uncapped layers to air for more than six months after growth did not make the layers peel off. This is a much longer period than those reported by other laboratories.[9] Keeping the uncapped layers under nitrogen flow makes them stable for a longer period of time.

IV. SUMMARY

In summary, large concentrations of oxygen have been observed in the uncapped undoped AlSb layers, resulting in extremely large sheet hole densities (as high as 3×10^{15} cm^{-2} and the corresponding mobility of 60 $cm^2V^{-1}s^{-1}$). However, for the capped layers, oxygen does not penetrate beyond the 50 Å GaSb cap layer, and the underlying AlSb layer is highly resistive.

XPS has demonstrated that the entire surface of the uncapped layers is oxidized. Three different oxides have been observed on the surface of the uncapped layers: Al_2O_3, Sb_2O_5, and Sb_2O_3.

It has been found that the oxidation process causes extremely large sheet hole densities in the uncapped AlSb layers. However, more work is needed to understand the mechanism for this phenomenon. Samples are currently being

examined by X-ray and transmission electron microscopy (TEM) techniques.

ACKNOWLEDGEMENT

We would like to thank Dr. Harry Dietrich for his technical advice and encouragement.

REFERENCES

(1) K. Longenbach, R. Beresford, and W. Wang, Trans. Electron Devices **37**, 2265 (1990).
(2) J. Werking, C. Bolognesi, L. Chang, C. Nguyen, E. Hu, and H. Kroemer, IEEE Electron Device Lett. **13**, 164 (1992).
(3) S. Ideshita, A. Furukawa, Y. Mochizuki, and M. Mizuta, Appl. Phys. Lett. **60**, 2549 (1992).
(4) C. Nguyen, B. Brar, H. Kroemer, and J. English, Appl. Phys. Lett. **60**, 1854 (1992).
(5) S. Subbanna, G. Tuttle, and H. Kroemer, J. Electronic Materials **17**, 297 (1988).
(6) A. Furukawa, and M. Mizuta, Electronics Lett. **24**, 1378 (1988).
(7) J. Pekarik, H. Kroemer, and J. English, J. Vac. Sci. Technol. B **10**, 1032 (1992).
(8) Note that annealing at higher temperatures (such as 400 °C, for 1 min.) caused the entire mesa peel off.
(9) Other laboratories have reported periods of time in the range of a few hours to a few days.

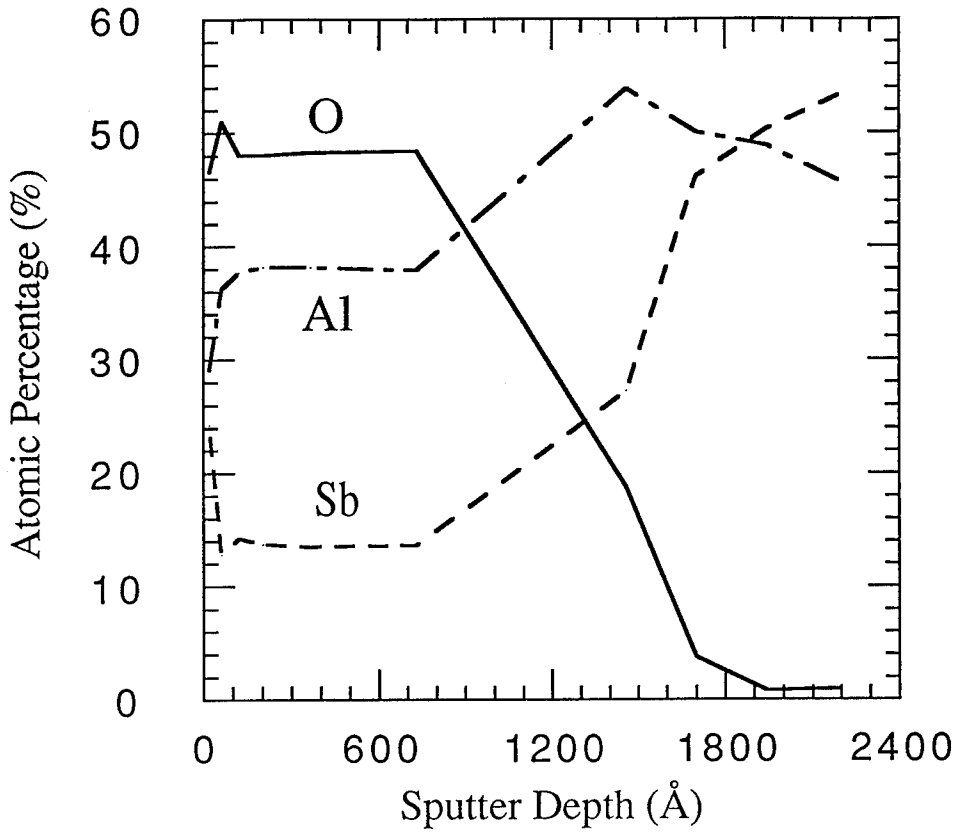


Figure 1. Auger sputter depth profile for an uncapped AlSb layer (1 μm thick).

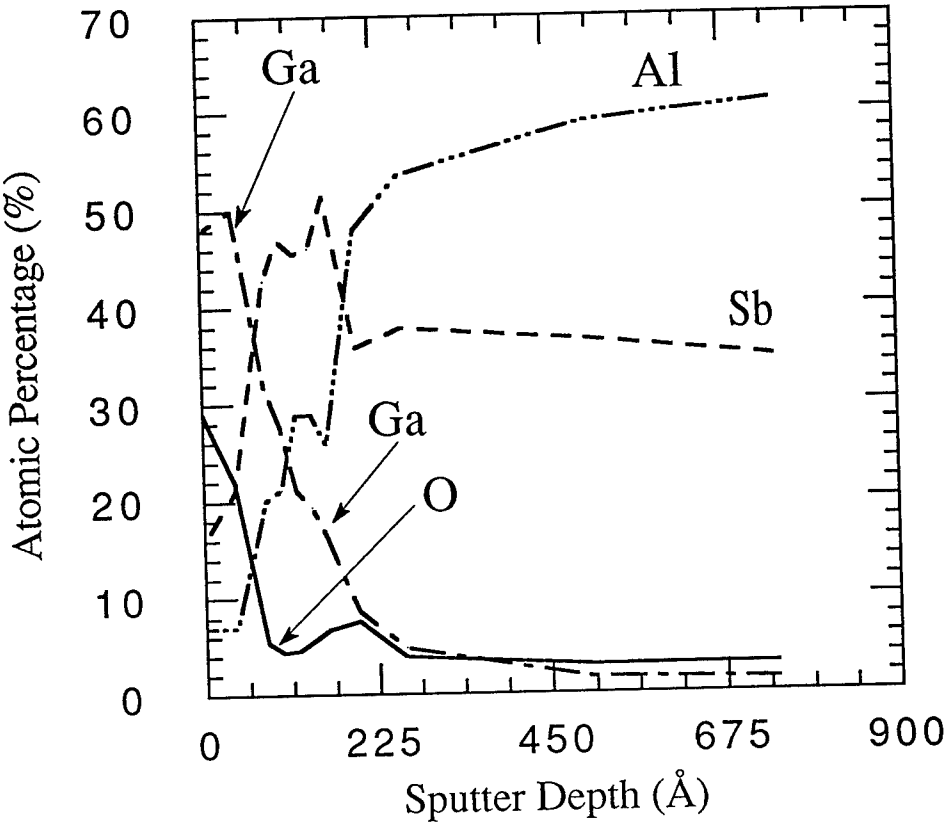


Figure 2. Auger sputter depth profile for a capped AlSb layer (0.3 μm AlSb, capped with 50 Å GaSb).

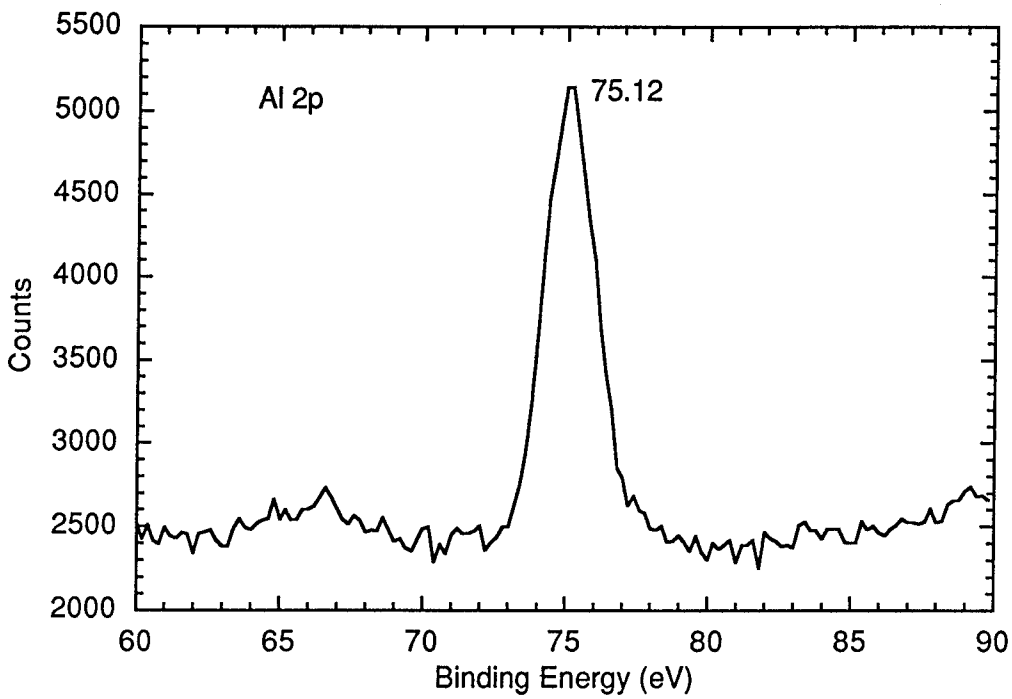


Figure 3. XPS spectrum obtained from the surface of an uncapped AlSb layer, showing the range between 60 and 90 eV.

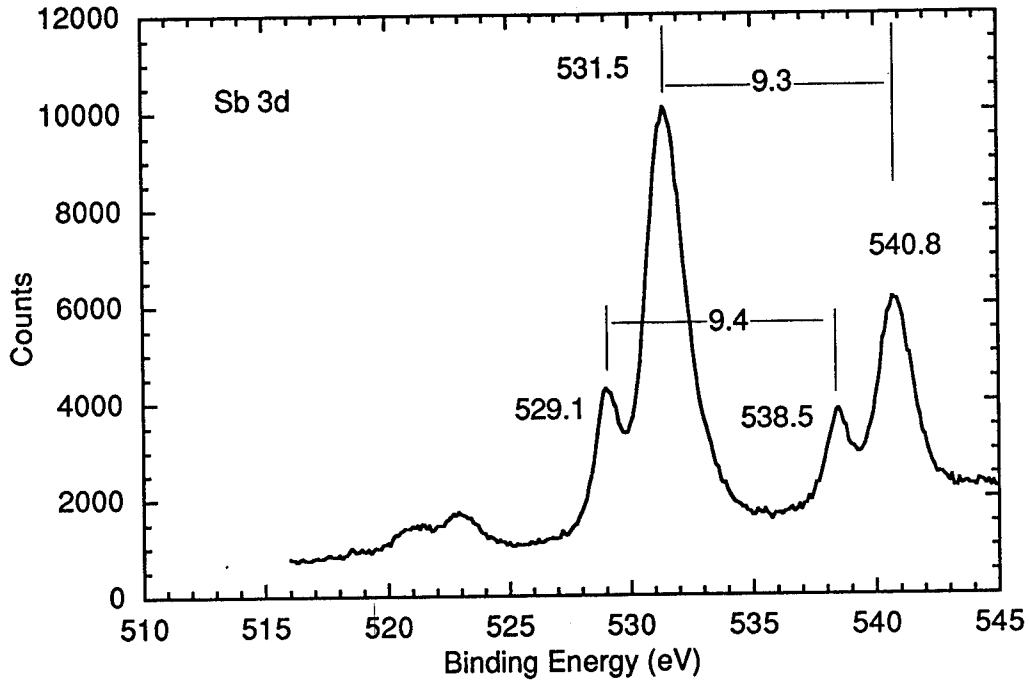


Figure 4. XPS spectrum obtained from the surface of an uncapped AlSb layer (the same sample as that of Figure 3), showing the range between 516 and 545 eV.

STATISTICAL MODELING OF PSEUDOMORPHIC HEMTS FROM AUTOMATED NOISE AND SCATTERING PARAMETER MEASUREMENTS

by

A. Caddemi, G. Martines, M. Sannino

Dipartimento di Ingegneria Elettrica, University of Palermo
Viale delle Scienze- 90128 Palermo - Italy
(Phone +39 91 590404; Fax +39 91 488452)

ABSTRACT

A method for the complete characterization of microwave transistors in terms of noise, gain and scattering parameters using a computer-controlled noise figure measuring set-up only is presented. A modeling procedure exploiting both scattering and noise parameters has been employed to extract the equivalent circuit which gives the best fit of the experimental data.

Results are reported concerning the complete characterization and modeling of a series of ten pseudomorphic HEMTs in the 8-16 GHz range.

1. INTRODUCTION

For the (computer-aided) design of low-noise amplifiers , the designer needs the complete characterization of the transistors in terms of noise, gain and scattering parameters.

The scattering parameters (S-parameters, [S]) are measured by an Automatic Network Analyzer (ANA) , whereas the noise parameters (N-parameters, {N}) and the gain parameters (G-parameters, {G}) can not be measured directly by any instruments.

The N-parameters describe the dependence of the transistor noise figure F on the input termination (source) reflection coefficient Γ_s through the relationship

$$(1) \qquad F(\Gamma_s) = F_o + 4N_n \frac{|\Gamma_s - \Gamma_{on}|^2}{(1 - |\Gamma_s|^2)(1 - |\Gamma_{on}|^2)}$$

where Γ_{on} is the optimum value of Γ_s corresponding to the optimum noise figure F_o, and N_n is a terminal invariant parameter which represents how the noise figure departs from the minimum when Γ_s differs from Γ_{on}; N_n is related to the more known noise resistance R_n by $N_n = R_n G_{on}$, where G_{on} represents Γ_{on} in terms of conductance .

Therefore, the determination of the N-parameters requires measurements of $F(\Gamma_s)$ for some values of Γ_s (more than four, for accuracy) and the data processing of the experimental data through (1), by means of an error minimizing procedure (such as the least-squares method) .

The same procedure is required to determine the four G-parameters by gain measurements by using the relationship similar to (1)

$$(2) \qquad \frac{1}{G_a(\Gamma_s)} = \frac{1}{G_{ao}} + 4N_g \frac{|\Gamma_s - \Gamma_{og}|^2}{(1 - |\Gamma_s|^2)(1 - |\Gamma_{og}|^2)}$$

where Ga is the available power gain, which depends on Γ_s , and G_{ao} (maximum available power gain), $|\Gamma_{og}|$, $\angle\Gamma_{og}$ and N_g are the four G-parameters of the device. Alternatively, in order to avoid gain measurements, the G-parameters can be computed by [S].

The noise figure measurements are the most difficult to be carry out. The conventional measuring set-up is assembled with a 50-ohm noise source which injects noise in the device under test (DUT) through a tuner, used

This work was supported by National Research Council (C.N.R.) and Italian Space Agency (A.S.I.).

as an admittance transformer network to realize the values of Γ_s, a receiver and a noise figure meter. The noise figure $F(\Gamma_s)$ is extracted from the noise figure F_m as indicated by the meter, according to the following relationship

$$F_m = \alpha_{\Gamma s} \left(F(\Gamma_s) + \frac{F_r(\Gamma_{out}) - 1}{Ga(\Gamma s)} \right) \quad (3)$$

where $\alpha_{\Gamma s}$ represents the loss of all the passive stages preceding the DUT, which depends on the configuration of the tuner (i.e. on Γ_s), and $F_r(\Gamma_{out})$ is the noise figure of the "receiver", i.e. the cascade of all the active stages following the DUT, which depends on the DUT output reflection coefficient Γ_{out}. For the evaluation of the loss $\alpha_{\Gamma s}$, the computation through the S-parameters has been suggested.

Examples of graphical representations of the device noise figure $F(\Gamma_s)$ (see (1)) , available power gain $Ga(\Gamma s)$ (see(2)) and loss $\alpha_{\Gamma s}$ of a slide-screw tuner are shown in Fig. 1 on the Smith chart.

This conventional methodology requires different measuring procedures to be carried-out in separate times by means of different noise figure, gain and S-parameters measuring set-ups and manual action to disassemble and re-assemble the instruments. Consequently, it suffers from low accuracy, low repeatability and large time consumption.

To fully evaluate the last problem it has to be accounted for that the determination of the four noise- and the four gain-parameters of the DUT requires measurements of noise figure and available gain for several values of Γ_s . Furthermore, to correctly carry-out the measurements it is fundamental to realize values of Γ_s on the Smith chart close to loci selected on the basis of well assessed criteria [1,2,3]. These loci are usually different even for each device of the same series because they depend on the optimum value Γ_{on}. This requires that F_o is approximately determined by acting on the input tuner to obtain the minimum F_{mo} of the meter reading F_m.

A further consideration regards the direct determination of the optimum input termination condition from the noise point of view. This procedure is often preferred by the experimenters for characterizing the devices in place of the time consuming procedure for determining all the four {N}. This condition is searched by adjusting the (manual or computer-driven) input tuner to obtain the minimum F_{mo} of the meter reading F_m; then F_o is computed through (3) and Γ_{om} is measured. This procedure corresponds to determine the minimum of the composition of the graphical representations of F, G_a and α of Fig.1, which obviously does not coincide with the optimum of F. From this viewpoint too, the noise parameters test-set commercially available are not useful.

All the above disadvantages are avoided by the measuring methodology here presented, which has been expressly devised to allow the complete characterization of microwave transistors by means of an automated noise figure measuring set-up. All the elements of (3) (F_m,G_a and also $\alpha_{\Gamma s}$ [7]) are simultaneously measured for each values of Γ_s. Then, the DUT noise figure F is computed in real time by (3), and {N} and {G} are derived from (1) and (2). All the scattering parameters which are needed for amplifier design are also derived by computation from the G-parameters, i.e. without further experimental steps. In addition, since the DUT is thus driven at noise level, possible non-linear effects of the transistors (MESFET, HEMT) due to insufficiently low signals, such as those furnished by a network analyzer, are avoided.

State-of-the-art modeling procedures for microwave and millimeter-wave low-noise transistors are based on the extraction of a noiseless equivalent circuit from measured scattering parameters. The noise performance is then predicted by associating noise models whose fitting coefficients, in the simplest case, can be determined from a few measurements of F_{min} vs. frequency. By this approach, the solution is not unique since different noise models associated to a small-signal model can reproduce identical values of the minimum noise figure F_{min} with remarkable differences among the remaining parameters.

In the present work, a novel modeling procedure is presented which allows the extraction of a consistent noisy model from the simultaneous fitting of the measured N- and S-parameters among which a strong correlation exists.

2. THE AUTOMATIC NOISE, GAIN AND SCATTERING PARAMETERS TEST-SET

In order to present the measuring system, we refer to the simplified scheme of Fig. 2 . The noise is injected in the system by a variable admittance noise source. A step-motor-driven microprocessor-controlled slide-double-screw-tuner is employed as input mismatching-network. The reflection coefficient Γ_{eq} of the variable admittance noise source so obtained is measured by the reflectometer (an ANA used in the reflection mode only)

connected by a switch. The corresponding value Γ_s of the input termination of the DUT is then derived by computation. To this end, the cascade of all the passive stages (bias tee, 1/2 transistor test fixture,connectors,...) which separates the transistor input from the Γ_{eq} measuring plane is previously characterized vs. frequency in terms of scattering parameters.

The DUT output is terminated on the 50 ohm input admittance of the receiver through the output bias tee and a switch; in other words, it is noteworthy that an output matching tuner is not used. In conventional measuring systems, this second tuner is used to match the DUT output impedance Γ_{out} to the receiver input. In this way, the receiver noise figure F_r is constant against Γ_s variations.

Unfortunately, the use of this output tuner causes two very serious drawbacks:

a) the output matched condition causes very frequently the rise of oscillations;

b) automating the noise measuring system becomes impossible. After adjusting the input tuner, it would be necessary to adjust the output tuner as well with the aid of a "matching detector". Consequently, the automatic searching of the optimum input matching for noise by adjusting the input noise would becomes unfeasible.

In order to avoid these disadvantages the DUT output is connected directly to the low impedence Γ_r of the receiver (50 ohm nominal); the noise figure $F_r(\Gamma_{out})$ of the receiver is then computed on the basis of the measured values of $F_r(0)$ and Γ_{out} [5,6] .

A programmable step attenuator is inserted as first stage of the receiver. By acting on the attenuator for each Γ_s, different values of the receiver noise figure F_r are realized and the DUT available power gain is computed [5-8].

Once α and $(F_r - 1)/G_a$ are determined, the DUT noise figure F is derived by (3).

Controlled by a (unpublished) software, the automatic set-up initially performs the calibrations of the instruments and some guiding measurements. On the basis of the results so obtained, the software selects for each frequency the values of Γ_s and other parameters (e.g. the values of the attenuation to introduce to the receiver input for the determination of the available gain) in order to realize the best measuring conditions for accuracy . Furthermore, it drives all the measuring steps vs. frequency and DUT bias condition fach Γ_s, evaluates the accuracy, processes the data and furnishes the {F}, $\{G_a\}$ and [S] parameter sets [11-12]. This is accomplished without the action of an (unskilled) operator.

3. MEASUREMENTS

The automatic measuring set-up has been recently used to characterize the performance of ten pseudomorpfic HEMTs (CFB001-3, by Celeritek) over the 8-16 GHz range at the bias conditions suggested in the data sheets for the best noise behavior to the aim of extracting an accurate noisy model which also accounts for the parasitics due to the transistor package. The measured N- and S-parameters of this transistor series are reported in Fig. 3. The method can be also applied for the automatic direct (and correct) measurements of the optimum noise parameters F_o and Γ_{on} by adjusting the tuner, because the system furnishes directly the minimum noise figure F_o computed from (3) while measuring α, F_r and G_a.

In conclusion the measuring system here described can be seen as an instrument (at present not commercially available) for the direct measurement of the DUT noise figure by removing on-line from the meter reading the noise effects of all the other stages of the set-up .

4. MODELING

We considered the measured spreading of the parameter values as the modeling optimization goals for the device series. To include noise in the modeling procedure, we have adopted the representation based on effective temperatures of the chip resistive elements R_{ch} and R_{ds} since the above coefficients are not influenced by the impedance-transforming properties of the device package [9]. In order to determine the global model topology which also accounts for the package effects, we start from the well-assessed chip equivalent circuit and sequentially add elements to the input section accoding to previous analysis results which have shown a definite correlation between the trends of S_{11},Γ_{on} , R_n and different circuit configurations. Subsequently, we complete the feedback paths by introducing (if required for the good match of S_{12} and S_{21}) additional coupling capacitors between the gate and drain nodes, and refine the output section for matching S_{22} and, partially, S_{21}.

Once this basic topology has been obtained, we perform an analysis of the interdependence between the calculated model parameters and the (model) element values.

The results are reported in form of a matrix which, after proper manipulation,allows the decomposition of

the overall optimization problem into separate suboptimization cycles.

By this procedure we obtain a fine tuning of the element values since each of them is optimized in order to minimize the error function relevant to a scattering or noise parameter which is strongly affected by that element (or, more frequently,a group there of) [10,11].

For the above HEMT series, the model performance (bold line) is compared with the measured parameters in Fig.3. The corresponding noisy model is reported in Fig.4 and the relevant element values are listed in Tab. 1.

It has to be noted that the circuit model represents the performance of packaged device as a whole, thus taking into account the package parasitic effects and the chip-to-packge interactions.

Therefore the model is not predictive from the device manufacturing-process viewpoint, but it has all the qualifications required for CAD applications.

REFERENCES

(IEEE publications only)

[1] - G. Caruso, M. Sannino,"Computer-aided determination of microwave two-port noise parameters," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-26, pp. 639-642, Sept. 1978.

[2] - M. Sannino, "On the determination of device noise and gain parameters" , *Proc. IEEE* (Letters), vol. 67, pp. 1364-1366, Sept. 1979.

[3] - G. Martines, M. Sannino, "A method for measurement of losses in the noise-matching microwave network while measuring transistor noise parameters", *IEEE Trans. on Microwave Theory Tech.*,vol. MTT-35, pp. 71-75,Jan. 1987.

[4] - M. Sannino, "Simultaneous determination of device noise and gain parameters through noise measurements only" , *Proc. IEEE* (Letters), vol. 68, pp. 1343-1345, Oct. 1980.

[5] - M. Sannino, "Comments on 'Simultaneous determination of device noise and gain parameters through noise measurements only', reply of the author" , *Proc. IEEE* (Letters), vol. 70, pp. 100-101, Jan. 1982.

[6] - G. Martines, M. Sannino, "Determination of microwave transistor noise and gain parameters through noise-figure measurements only", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-30, pp. 1255-1259, Aug. 1982.

[7] - E. Calandra, G. Martines, M. Sannino, "Characterization of GaAs FET's in terms of noise, gain, and scattering parameters through a noise parameter test set", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, pp. 231-237, March 1984 .

[8] - G. Martines, M. Sannino,"Simultaneous determination of transistor noise, gain and scattering parameters for amplifier design through noise figure measurements only", *IEEE Trans. Instrum. Meas.*, vol. IM-34, pp. 89-91, March 1985.

[9] M. W. Pospieszalski, "Modeling of noise parameters of MESFETs and MODFETs and their frequency and temperature dependence", *IEEE Trans. Microwave Theory Tech.* ,vol.37, pp. 1340-1349,Sept. 1989.

[10] - A.Caddemi, G. Martines, M. Sannino, "HEMT for Low Noise Microwaves: CAD Oriented Modeling", *IEEE Trans. Microwave Theory Tech.,* Special Issue on "Process-Oriented Microwave CAD and Modeling", Vol. 40,pp. 1441-1445, July 1992.

[11] A. Caddemi, M. Sannino, "Modeling of low-noise microwave HEMTs for CAD-oriented applications", *Int. J. of Microwave Millimiter-Wave Computer-Aided Engineering,* Special Issue on Modeling of MESFETs and HEMTs for microwave CAD, pp. 29-36, Jan. 1993.

Ref. [2,3,5,] are also available in "Low-noise microwave transistors and amplifiers," edited by H. Fukui, IEEE Press, 1981.

Ref. [2,7,12] are also available in "MODFETs: Applications and Circuits," edited by H. Daembkes, IEEE Press, 1990.

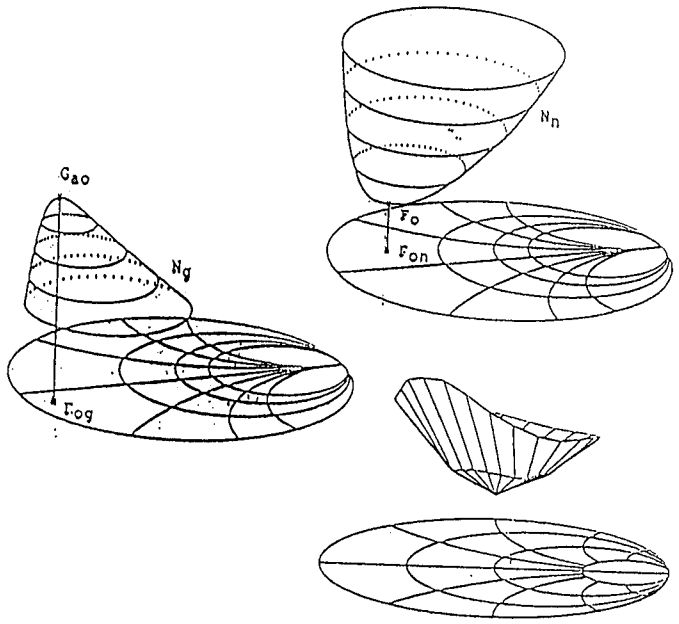


Fig. 1 - Example of graphical representation of $F(\Gamma_S)$, $G_a(\Gamma_S)$ and $\alpha_{\Gamma S}$ on the Smith chart. The graphical representation of (1) is the composition of them. From this it is apparent that the optimum value Γ_{mo} of the DUT input admittance corresponding to the minimum $F_{m\,o}$ of the noise figure given by (3), which can be found by adjusting the tuner, may differ considerably from the optimum Γ_o corresponding to DUT F_o.

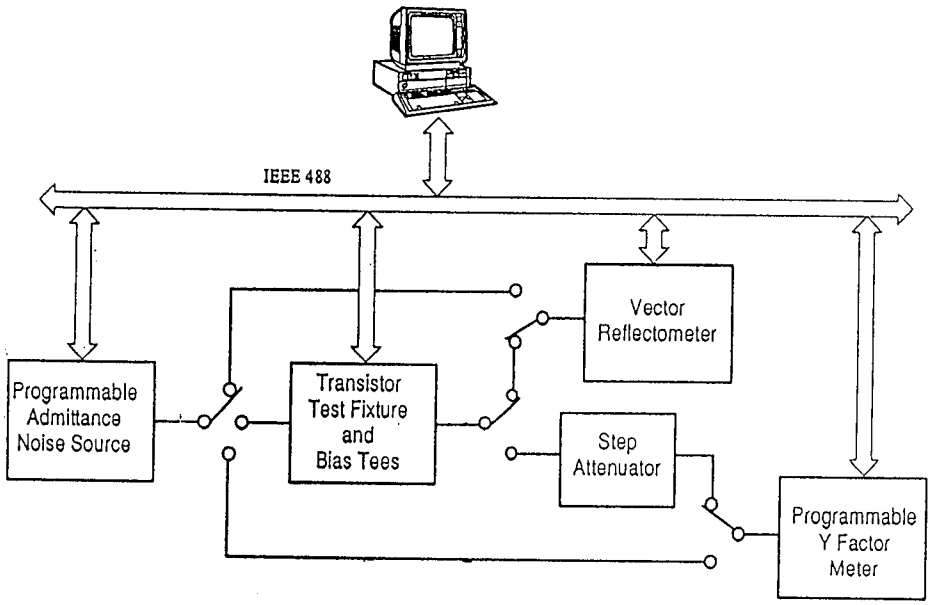


Fig. 2 - Simplified scheme of the automatic noise figure test-set.

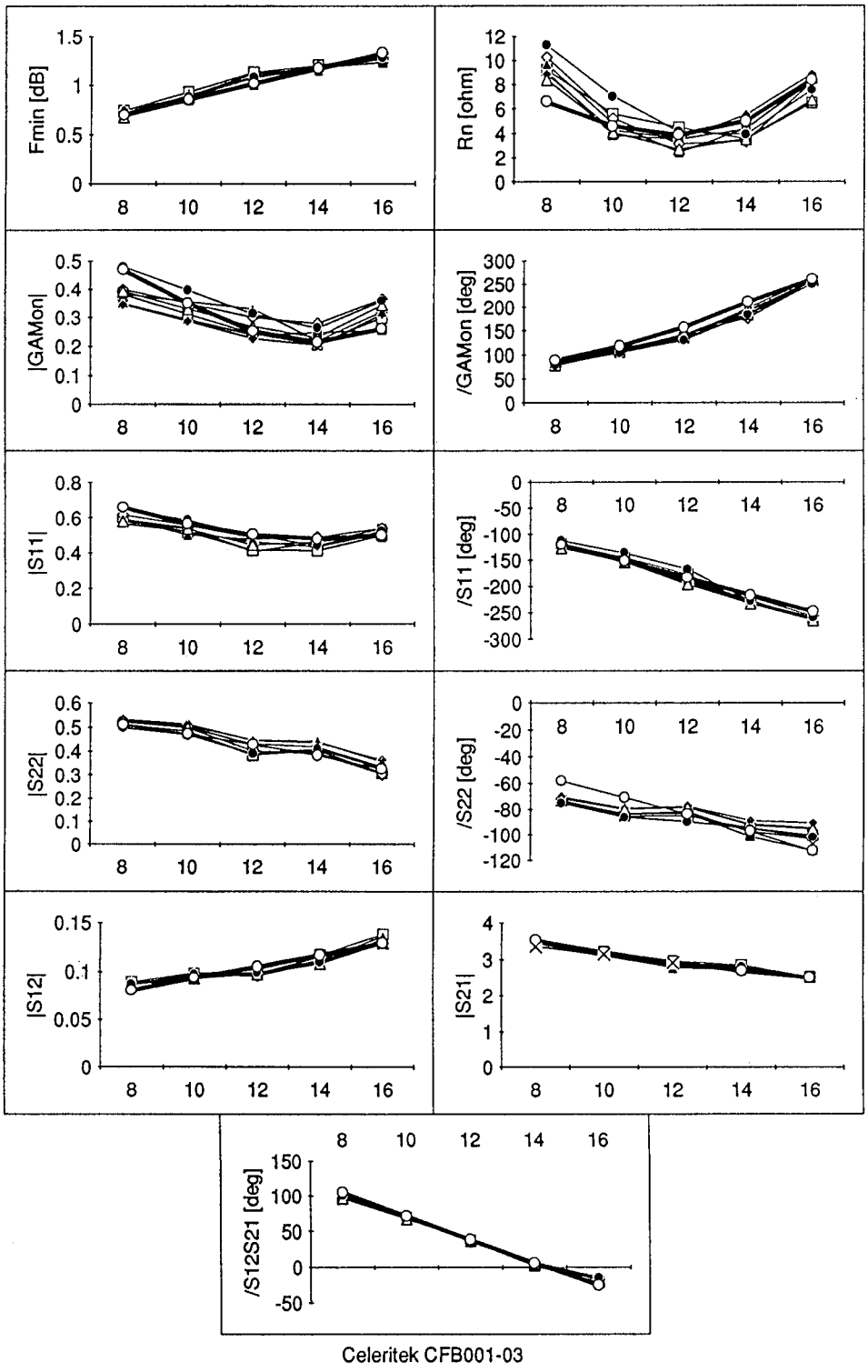


Fig. 3 - Measured and modeled (bold line) N- and S-parameters of the CFB001-03 series (by Celeritek) over the 8-16 GHz frequenzy range.

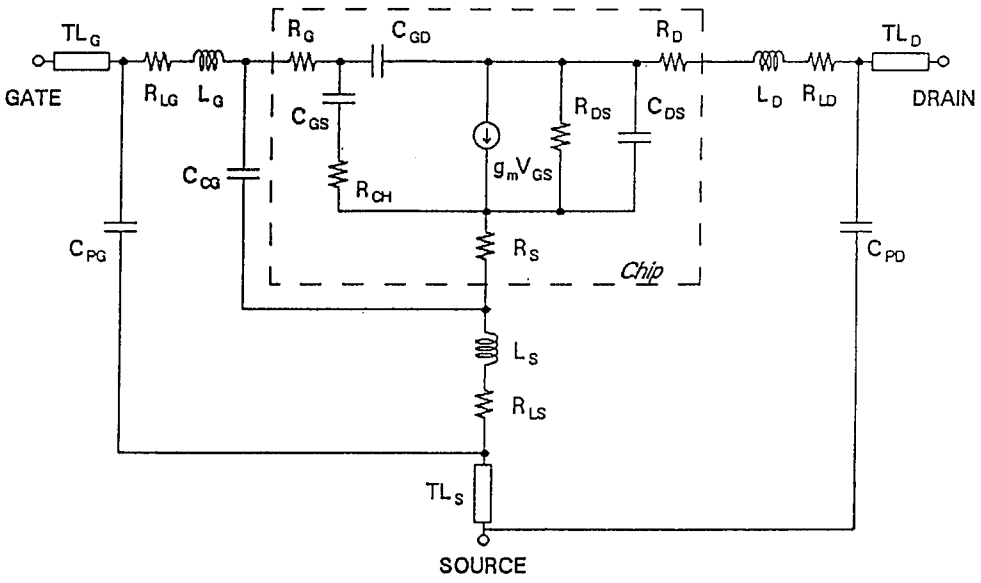


Fig. 4 - Equivalent circuit model of the CFB001-03 series.

R_{LG}	= 0.3 ohm	R_D	= 1.5 ohm
L_G	= 0.47 nH	L_D	= 0.49 nH
R_G	= 2.2 ohm	R_{LD}	= 1 ohm
C_{CG}	= 0.03 pF	R_S	= 1.1 ohm
C_{PC}	= 0.06 pF	L_S	= 0.033 nH
C_{GS}	= 0.23 pF	R_{LS}	= 1.2 ohm
R_{CH}	= 1.4 ohm	R_{DS}	= 180 ohm
C_{GD}	= 0.023 pF	g_m	= 66 mS
C_{DS}	= 0.07 pF	τ	= 2.1 ps
TL_G	L= 840	TL_D	L= 280
TL_S	L= 500		

E = line physical lenght in microns @ 8GHz;
Z=50 ohm
All resistors are warmed @ 290 ^{o}K;
R_{DS} @ 3950 K

Tab. 1
Values of the circuit elements of the model shown Fig. 4.

QUALITATIVELY MODELING HETEROJUNCTION BIPOLAR TRANSISTORS FOR OPTIMIZATION: A NEURAL NETWORK APPROACH

Mankuan Vai, Zhimin Xu, and Sheila Prasad

Department of Electrical and Computer Engineering
Northeastern University
Boston, MA 02115

ABSTRACT

A neural network approach is developed to qualitatively model the relationship between fabrication process parameters and the characteristics of a heterojunction bipolar transistor (HBT). An equivalent circuit model is used as an intermediate representation format for this objective. The goal of this research project is to develop a method that can predict and explain changes in the behavior of a device without the need for precise problem formulations and computationally intensive methods. The primary use of such a neural network model is in a reverse modeling process which performs device optimization.

I. INTRODUCTION

This paper describes a research project that investigates the development of a parallel distributed analog computing network (i.e., a neural network) to qualitatively model the relationship between fabrication process parameters and device characteristics. The objective is to provide a model for a reverse modeling process which performs device optimization. The initial result of applying this qualitative model to the modeling of an heterojunction bipolar transistor (HBT) is presented. Although HBTs are targeted in this project, the method described here is general and can be readily adapted for other devices.

Since Kroemer [1] predicted the great potential of HBTs for high frequency/speed circuits, there has been a considerable research effort directed toward understanding the physics of the device and studying its applications. The characteristics of HBTs can be evaluated by scattering parameter (S-parameter) measurements. But, such data generally do not explain the causes of the observed characteristics. Furthermore, the relationship between the fabrication process and the device characteristics is generally very complicated.

Equivalent circuit models serve as an excellent intermediate representation format for device optimization. Ladbrooke has advocated the idea of *reverse modeling* in discussing

0-7803-0894-8/93 $3.00 © 1993 IEEE

the relationship between device structure/material parameters and its equivalent circuit model [2].

Reverse modeling is similar to a standard modeling procedure except that the goal in this case is to minimize the difference between the modeled and desired performance, and hence the name *reverse modeling*. The primary difficulty in reverse modeling lies in the fact that the optimized model has to be physically consistent so that the desired processing parameters can be developed. A change in any single processing variable typically results in correlated changes in many, if not all, equivalent circuit elements.

For example, if the given goal is the improvement of the power gain at a certain operating frequency, the reverse modeling process must be able to determine the required changes to the model. The constraints resulting from both the equivalent circuit model (e.g., KVL, KCL) and the fabrication process (e.g., doping concentration) have to be satisfied in this search operation. The fact that all these constraints have to be incorporated makes the development of a computer-aided design tool very complicated and the product difficult to use.

On the other hand, the manufacturing process for a semiconductor device as complicated as the HBT often involves considerable experimental data. Engineers often make decisions based on their intuitive knowledge of the device structure and fabrication process. A rule-based system can be used to describe the condition-action relationship between parameters. The problem can then be solved by inference and induction. However, a typical production system has the inherent limitations of sequentiality, directionality, exact matching, and determinism.

We have investigated the use of a parallel distributed analog computing network, commonly called a neural network, to model the relationship between processing parameters, equivalent circuit elements, and characteristics of HBTs. The advantages of this approach include not only the fact that multiple constraints can be easily and globally enforced, but also that many competing hypotheses can be simultaneously explored. In addition, the model can be developed without the need for precise problem formulation and computationally intensive methods.

II. QUALITATIVE MODELING

We have observed that the above-mentioned problems and constraints can be addressed by a qualitative model implemented with a parallel distributed analog computing network. Parallel distributed analog computing networks are commonly referred to as artificial neural networks in the literature [3]. Although artificial neural networks are

famous for their capability of *learning* the solutions to the problems that they are designed to solve, they also provide a framework for constructing special parallel machines to solve specific problems [4].

In contrast to conventional parallel processors, the function of each neuron is extremely simple, and the overall behavior of a neural network is determined predominately by a set of interconnections. Conceptually, neural networks are formed by interconnecting many simple processors (i.e., neurons) each connected to many others. Every connection entering a processor has an adaptive coefficient called a weight assigned to it. This weight is used to amplify, attenuate, and change the sign of the signal in the incoming connection. The processors used in a neural network are nonlinear and are typically analog. An input vector is entered into the network. Each processor operates on the outputs of other processors according to its transfer function and delivers a single output to other processors. Often, the transfer function sums the incoming signals to determine the value of the next output value. The result is an output vector representing some characteristics associated with the input. The major attraction of these neural networks is their high level of parallelism. Another commonly claimed feature of artificial neural networks is that they are inherently fault tolerant.

The qualitative modeling approach developed here is based on the paradigm of Hopfield type neural networks [5]. Consider a Hopfield type neural network of N nodes. If the activation of a node is updated according to the equation:

$$V_i(t+1) = sgn(\sum_{j=1}^{N} T_{ij} V_j(t) + I_i), \tag{1}$$

where $V_i(t) \in \{0,1\}$ is the state of node i at moment t, T_{ij} is the weight associated with the link between nodes i and j, I_i is the internal threshold parameter of node i, and

$$sgn(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}, \tag{2}$$

it can be shown that an energy function defined as

$$Energy = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij} V_i V_j - \sum_{i=1}^{N} I_i V_i + K, \tag{3}$$

where K is a constant, is minimized.

The significance of a Hopfield type network is in its capability of performing associative inference. There is no specific distinction between input and output vectors and a network perturbed by changing one or more neuron state will evolve into one of its consistent states which are the minima of its energy equation (3).

This Hopfield type neural network is applied to implement a qualitative model. The behavioral relationship between three variables (X, Y, Z) in an arithmetic equation

$$X = Y \bullet Z \text{ (operator } \bullet \in \{+, -, \times, \div\}) \quad (4)$$

is mapped into a neural network building block as follows. An arithmetic equation $X = Y + Z$ can be qualitatively represented by the consistent states listed in Table 1. Each row of Table 1 describes the change of a variable caused by varying the other two variables. For example, the first row of the table indicates that if two of the variables are remaining unaltered, then the third one is also unchanged. This table is mapped into a Hopfield type neural network with 6 nodes. The mapping involves the determination of all the link-weights (T_{ij}'s) and thresholds (I_i's) and is done by formulating a linear programming problem with the 64 constraints in the forms of equalities and inequalities according to the energy equation (3). There are 13 consistent states and 51 inconsistent states. All consistent states have an energy of 0 while all inconsistent states have energies larger than zero. The simplex method [6] is used to solve for the unknown T_{ij}'s and I_i's. The resulting network is shown in Figure 1.

Table 1 The neural network states consistent with the arithmetic equation $X = Y + Z$ (I: increasing, D: decreasing, and S: staying the same).

$X (V_1V_2)$	$Y (V_3V_4)$	$Z (V_5V_6)$
00 (S)	00 (S)	00 (S)
00 (S)	10 (I)	01 (D)
00 (S)	01 (D)	10 (I)
10 (I)	10 (I)	00 (S)
10 (I)	10 (I)	01 (D)
10 (I)	10 (I)	10 (I)
10 (I)	00 (S)	10 (I)
10 (I)	01 (D)	10 (I)
01 (D)	01 (D)	00 (S)
01 (D)	01 (D)	10 (I)
01 (D)	00 (S)	01 (D)
01 (D)	10 (I)	01 (D)
01 (D)	01 (D)	01 (D)

The neural network shown in Figure 1, which will be referred to as a fundamental neural network, is very powerful since it can be used to model any basic arithmetic relations. It can be easily shown that the same network can also be used to qualitatively

model arithmetic operations $X = Y \times Z$, $Y = X - Z$, $Y = X \div Z$. If we set X = voltage, Y = current, and Z = resistance, this neural network becomes a qualitative model of Ohm's law.

The modeling approach described above can be iteratively and hierarchically applied to build a system model. This process is explained here through the development of a model to represent the relationship between two serially connected resistors. Referring to Figure 2, there are five arithmetic constraints in this model ($E_1 = I * R_1$, $E_2 = I * R_2$, $E = E_1 + E_2$, $R = R_1 + R_2$, and $E = I * R$) and thus 5 fundamental neural networks are required.

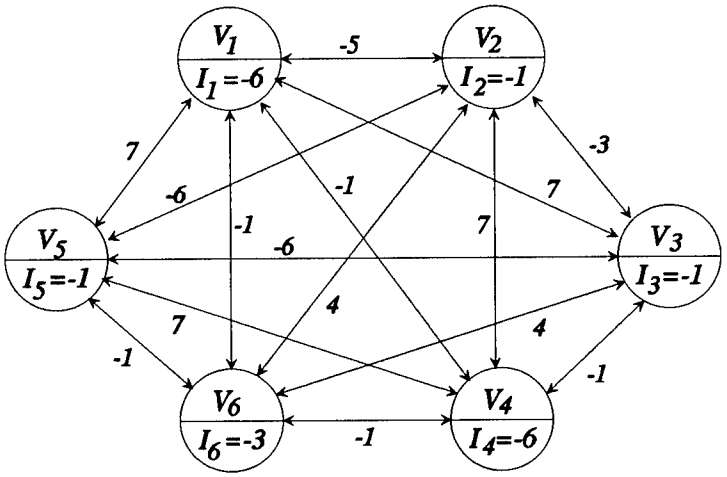


Figure 1 The qualitative model of $X = Y + Z$.

Common feature nodes in individual fundamental neural networks are merged to yield the model shown in Figure 3 (some links were left out for clarity) as follows. Nodes with identical labels are replaced by a single node with a threshold (I) equal to the sum of the original node thresholds. Similarly, identical links are merged into a single link with link-weight (T) equal to the sum of the original link-weights. In Figure 3, the pair of nodes representing a feature is grouped and labeled accordingly. In addition, nodes were grouped into overlapping fundamental networks to show the 5 modeled constraints.

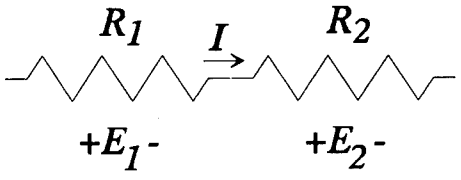


Figure 2 Two serially connected resistors for the illustration of model building.

The obvious use of this model is to answer questions such as: what is the qualitative effect of changing one element on other circuit values? However, our goal is to use the qualitative model of an HBT to guide the reverse modeling process since a solution found by this model must satisfy all the imposed constraints. A global view is thus always maintained. The setup for fulfilling this objective is shown below in Figure 4.

According to Figure 4, an automatic mapping process has been implemented to map an equivalent circuit into a corresponding neural network. When used in a modeling process, the difference between the modeled and desired characteristics is translated into the demanded qualitative changes of certain values (e.g., voltage or current) in the equivalent circuit. These qualitative changes are then treated as the input vector of the neural network model to determine the associative changes to other variables so that a consistent state is maintained.

The neural network qualitative model of the HBT is implemented in a modeling process which extracts equivalent circuit parameters from S-parameter measurements to evaluate the effectiveness of this approach. The detailed information about this collector-up HBT being modeled is reported in [6]. The modeling process implements the simulated annealing (SA) technique described in [7] to avoid the problem of local minima. The role of the neural network HBT model is to qualitatively guide the changes of equivalent circuit parameters to minimize the difference between the measured and simulated S-parameters.

The same SA modeling process, implemented to generate intermediate solutions by a random search, is applied independently to the same HBT for the purpose of comparison. The chart in Figure 5 shows the convergence rates of both processes. The modeling process that is guided by a neural network qualitative model converges to its final solution in less than half of the number of iterations required by the method using a random search. An advantage of the neural network guided approach not shown in Figure 5 is that only a single solution is generated and evaluated at each iteration. Comparing with the random search approach that processes a large number of solutions at each iteration to satisfy the slow cooling procedure of an SA optimization, the saving in the modeling effort is significant.

III. SUMMARY

In summary, a novel approach that uses a neural network to qualitatively model an HBT is developed. The initial result of applying this neural network HBT model to guide the fitting of an equivalent circuit model to measured S-parameters is presented.

Encouraging results in the convergence rate of the modeling process have been observed. The ultimate potential of this qualitative modeling technique is to enforce physical constraints and encode experimental and intuitive knowledge in the modeling and optimization (reverse modeling) of semiconductor devices.

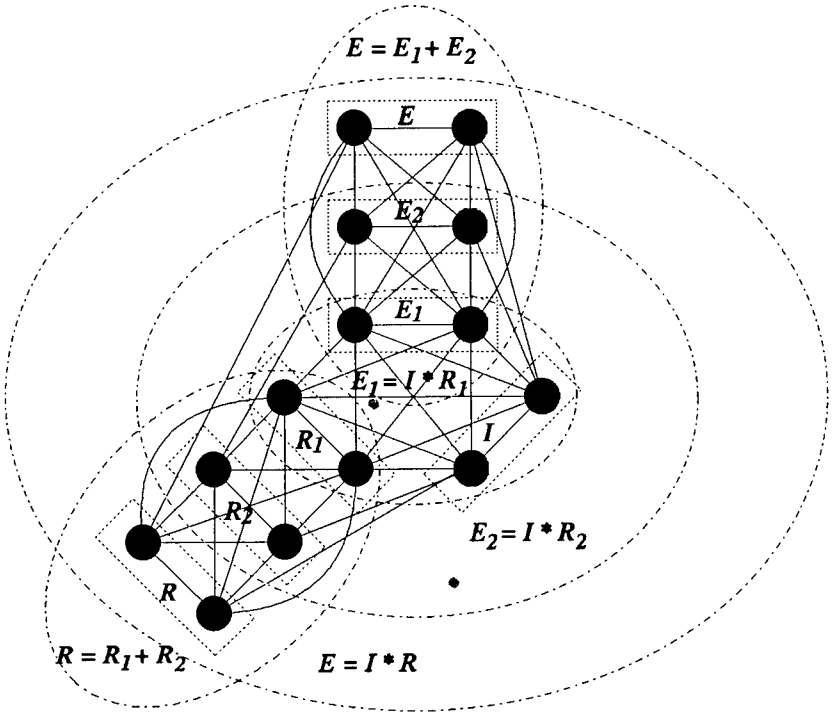


Figure 3 The qualitative model of two serially connected resistors (with some links left out).

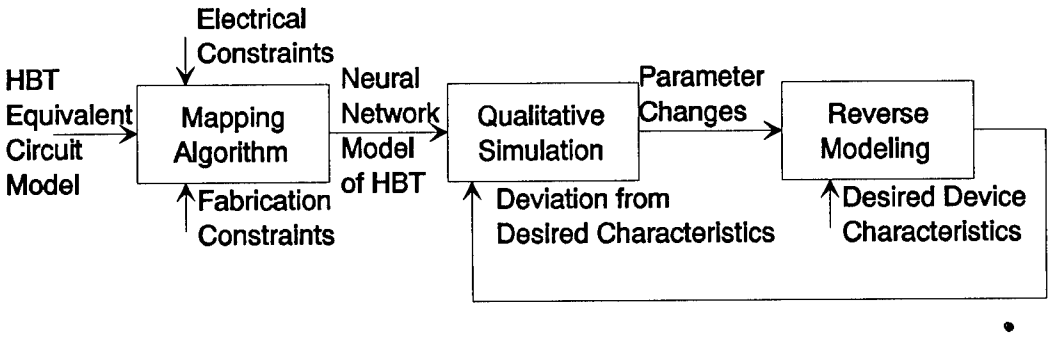


Figure 4 A setup for applying the neural network model of an HBT to reverse modeling.

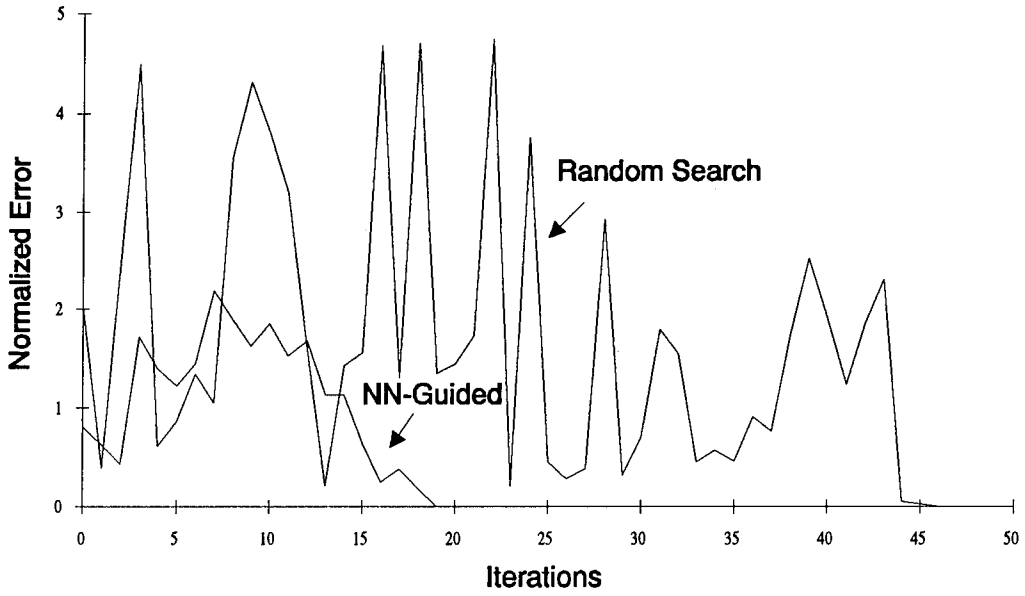


Figure 5 The convergence rates of the neural network (NN) guided and random search simulated annealing modeling process.

IV. ACKNOWLEDGMENT

The authors wish to sincerely thank Professor Clifton Fonstad of Massachusetts Institute of Technology, Cambridge, Massachusetts, and his semiconductor device research group for providing HBT devices for measurements. These measurements are invaluable to the past and future progress of this research project.

REFERENCES

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

[2] P. H. Ladbrooke, *MMIC Design: GaAs FETs and HEMTs*, Artech House, Boston, 1989.

[3] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, pp. 4-22, April, 1987.

[4] M. Vai, S. Prasad, and H. Wang, "A Smith chart represented by a neural network and its applications," *IEEE MTT-S International Microwave Symposium Digest*, pp. 1565-1568, June, 1992.

[5] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 81, 1984, pp. 3088-3092.

[6] H. Sato, J. C. Vlcek, C. G. Fonstad, B. Meskoob, and S. Prasad, "InGaAs/InAlAs/InP collector-up microwave heterojunction bipolar transistors," *IEEE Electron Device Lett.*, vol. EDL-11, pp. 457-459, 1990.

Fabrication, performance and characterization of Si delta-doped FET grown by MBE

Q. Chen and M. Willander

Division of Electronic Devices, IFM
Linköping University, S-581 83 Linköping, Sweden

Abstract

Si delta-doped FETs have been fabricated using low-energy Sb ion doping technique during Si MBE growth to realize delta-doping and a low temperature device processing budget. The FETs with a gate length of 2 μm showed a gate breakdown voltage of 3.8-4.7 V and a maximum extrinsic transconductance of 28 mS/mm, corresponding to an intrinsic transconductance of 42 mS/mm. Those results are the best reported so far in Si δ-FETs of similar gate length. The gate charge-control and low-field drift-mobility in Si δ-FETs were measured using C-V and I-V techniques. Measured charge-control had a strong nonlinearity and was in good agreement with theoretical calculation. We observed, for the first time, that the drift-mobility in the delta-doped FET was decreased with the increasing gate voltage.

I. Introduction

Recently, the delta (δ) doping (or planar doping) technique has been developed for both GaAs and Si. This doping technique can provide high-density quasi-two-dimensional electron or hole gas in a V-shaped quantum well created by high density dopants in highly localized doping sheet. The delta-doping technique is being used in different field-effect transistor (FET) structures, e.g., MESFET [1-5] and MOST [5,6] to improve the performance of the device. A delta-doping layer is used as an electron or hole conduction channel in the delta doped MESFET (δ-FET). Compared with conventional MESFET and single channel MODFET, the δ-FET has a large gate breakdown voltage, high drain current capability, high transconductance and easy control of the threshold voltage. Most of the works on the δ-FET published so far were performed in GaAs [1-3]. Using solid phase epitaxy (SPE) method to prevent segregation of dopants in MBE growth and then to realize δ-doping in Si, a few results have been reported on Si δ-FETs [4,5]. In this paper, we present the experimental results on growth, fabrication,

0-7803-0894-8/93 $3.00 © 1993 IEEE

performance and characterization of Si delta-doped FETs grown by Si MBE and low-energy Sb ion doping technique [7].

II. Growth and fabrication

Epitaxial growth of Si films was performed in a VG-80 Si MBE system. N-type delta-doping was carried out using an electron-impact, single grid, low-energy ion source providing an accelerated Sb ion beam. This doping technique overcomes well the severe problems of thermal coevaporation doping, induced by surface segregation of dopants in Si MBE. The test delta-doping samples were analyzed by capacitance-voltage (C-V) profiling, high resolution cross-sectional transmission electron microscopy (XTEM) and secondary-ion spectrometry (SIMS). XTEM lattice images and C-V measurements showed that the full width at half maximum (FWHM) concentration of delta-doping profiles was $\leq$ 2 nm and 3 nm in test samples of a wide range sheet doping density in delta-doping layer [7]. The δ-FET layer structures were grown on p-type (100) Si substrate of doping concentration about 10^{15}-10^{16} cm^{-3}. After growth of a 300-nm-thick undoped Si (of residual n-type doping density about low-10^{15} cm^{-3}) buffer layer at a substrate temperature $T_S \approx 750$ °C, the Si deposition was interrupted. The sample surface was exposed, at $T_S \approx 620$ °C, to the low-energy Sb-ion source of a given flux for the time duration which was chosen such that the desired sheet doping density of 8×10^{12} cm^{-2} was obtained. An Sb-ion acceleration potential of 250 V was used during the exposure. An undoped Si cap layer, 30 nm thick, was then grown at the same temperature.

The thermal budget allowance should not exceed that used in material growth, to ensure that the characteristics of delta-doping layer are retained. A low temperature processing budget, without using ion implantation for source and drain contacts and then the high temperature post-implantation annealing, was used in fabrication of the Si δ-FETs. All processing steps were below 450 °C for which Sb out-diffusion was negligible. Following the mesa etching of 400-nm-depth which was down to the p-type substrate, a 400-nm-thick SiO_2 film (thick oxide), serving as device isolation and passivation, was deposited by low-pressure chemical vapour deposition (LPCVD) in the mixture of SiH_4 and O_2 at 400 °C. After etching completely the oxide on the top of the then formed island, a 100-nm-thick oxide (thin oxide) was deposited. The source and drain contact windows were opened. A 1000-nm-thick Al film for source and drain contacts was deposited, defined and alloyed at 450 °C in N_2 for 10 min. At this stage the non-selfaligned source and drain ohmic contacts to the delta-doped channel were formed by exposing two opposing edges of the mesa to Al film while the other two edges were kept covered with oxide. No high temperature post-implantation annealing were used. The Schottky

gates were formed by the deposition of a 100-nm-thick Pt film through the gate window on thin oxide and then by the gate definition. Fig.1 shows a schematic diagram of the Si δ-FET.

III. Performance of the device

Fig. 2(a) shows a typical I-V output characteristic of 2 μmx500 μm gate devices. The device reaches a near pinch-off at the gate voltage of -2.5 V. However, as reported previously [1,5], the pinch-off is somewhat not completed at high drain voltage, due to the punchthrough. The breakdown characteristic of Schottky gate is shown in Fig. 2(b). The turn-on voltages of the gate diodes are typically 0.8 V, in accordance with the Schottky barrier height of Pt on Si. The breakdown voltages of the gate diodes (defined at a gate current of 0.3 mA/mm) measured on various FETs across the wafer are between 3.8 V and 4.7 V.

A maximum transconductance of 28 mS/mm was obtained at V_{gs}=0.0 V and V_{ds}=3.0 V , which is better than the best result of 20 mS/mm reported so far for the Si δ-FETs of the same gate length [5]. This is attributed to the higher sheet doping density in delta-doping layer used in our structure than that in [5]. The peak intrinsic transconductance at V_{ds}=3.0 V is evaluated to be as high as 42 mS/mm, using an intrinsic transconductance extraction method [8]. The source and drain series resistances have not been optimized carefully in the non-selfaligned device structure and low temperature processing budget.

Fig. 3 shows the measured extrinsic transconductance as a function of the gate voltage for a drain voltage of 3.0 V. The transconductance shows a broad plateau around its peak of FWHM>1.8 V, which is typical of delta-doped FETs, in contrast to MODFETs which show a sharp peak of FWHM<0.5-0.7 V in the transconductance vs the gate voltage because of the onset at high gate biases of parallel conduction in high-bandgap barrier layers of low mobility. The presence of a broad plateau in transconductance in the δ-FET is due to that the carriers are confined in a V-shaped potential well while high forward gate bias can be applied since the gate is on the undoped layer above the delta-doping sheet.

IV. Charge-control and low-field mobility in Si δ-FETs

In studying operation of a FET, there are two important physical parameters which have to be taken into account carefully [9]: the charge-control process of the gate, i.e., how the carrier density in the channel n_s changes with the gate voltage V_{gs}; and the low-field *drift-mobility* of the channel carriers and its possible dependence on the gate voltage. No effort so far has been devoted to investigate the charge-control in the δ-FET. As for mobility, all published results

were for *Hall mobility* in delta-doped *bulk* structures without the gate and no gate voltage modulation effect on mobility which exists possibly in the δ-FET could be measured. The gate charge control and the low-field drift-mobility in Si δ-FETs is study experimentally by using C-V and I-V techniques to the "fat" devices of a 10 μm-long and at least 500 μm-wide gate.

The gate-to-channel capacitance C_{gc} vs the gate voltage V_{gs} were measured by tiring source with the drain. Shown in Fig.4 is an example of the C_{gc}-V_{gs} characteristics of Si δ-FETs measured at 300 K. The gate-to-channel capacitance in Si δ-FETs increases with the increasing gate voltage and has no saturation trend over the gate voltage range used. This is quite different from the C_{gc}-V_{gs} characteristics of MODFETs [10] and MOSFETs [11] in which an almost constant C_{gc}, corresponding to a near-linear charge-control, could be measured in the strong inversion regime. The gate charge-control can be obtained experimentally by [12]

$$qn_s(V_{gs})=\int_{-\infty}^{V_{gs}} C_{gc}(V)\, dV \ . \qquad (1)$$

Fig.5 gives n_s vs V_{gs} for the device used in Fig.4. There is clearly a strong nonlinearity in measured charge-control process. The solid line in Fig.5 is the fitting of a quadratic model we suggested previously [13,14]

$$n_s(V_{gs})=A[(V_{gs}-V_{th})+B(V_{gs}-V_{th})^2] \ , \qquad \text{for } V_{gs}>V_{th} \qquad (2)$$

to measured data. Here A, B and V_{th} are constants and V_{th} is so-called threshold voltage. It is obtained from data fitting that in our Si δ-FETs $A\approx 6.21\times10^{11}$ $cm^{-2}V^{-1}$, $B\approx 0.59$ V^{-1} and $V_{th}\approx$ -2.45 V. (2) fits well experimental result. This verifies our previous suggestion. A comparison of the charge-control between experiment and a calculation based on an exact model in [13] was made and shown in Fig.6. The agreement between them is quite well when V_{gs} is far above V_{th}. The agreement becomes poor as V_{gs} is close to or below V_{th} because the distributive nature of the large channel resistance deforms strongly measured C_{gc}-V_{gs} characteristics.

After n_s vs V_{gs} of a FET was experimentally determined, the drain currents I_{ds} or the drain output conductances $g_d=\partial I_{ds}/\partial V_{ds}$ were measured in the small drain bias region (typically less than 40~50 mV). The low-field drift-mobility μ_a at a specific V_{gs} can be found from [12]

$$\mu_a=\frac{L}{W}\frac{I_{ds}}{qn_sV_{ds}}=\frac{L}{W}\frac{g_d}{qn_s} \ , \qquad (3)$$

where L is the gate length and W is the gate width. We used g_d in extracting μ_a to minimize measurement error. g_d at a given V_{gs} was obtained by linearly fitting I_{ds} vs V_{ds} measured for

$V_{ds} \leq 40$ mV. Fig.7 shows a typical result of the measured μ_a vs V_{gs} in Si δ-FETs. There is clearly a gate voltage dependence. For $V_{gs}>-2.0$ V (a value close to V_{th}), μ_a decreases as V_{gs} increases. In contrast, μ_a decreases drastically with the decreasing V_{gs} for $V_{gs}<-2.0$ V. However, we should pointed out that the later trend *may* be only an artifact. There is an error, as discussed above, in measuring n_s from measured C_{gc}-V_{gs} characteristics as V_{gs} is close to or below V_{th}. In this gate voltage region the overestimated carrier density n_s, as seen from the comparison with the theoretical curve in Fig.6, would result in an underrated drift-mobility μ_a.

In fact the drift-mobility shown in Fig.7 is an average value

$$\mu_a = \frac{\sum_{i,k} \mu_{i,k}\, n_{i,k}}{n_s}\ , \qquad (4)$$

where $n_s = \sum_{i,k} n_{i,k}$, i is the index of subbands, k=1, 2 is the index of the set of subbands corresponding to two different electron effective masses along the [100] direction [13], $n_{i,k}$ and $\mu_{i,k}$ are the sheet density and mobility of electrons occupying in the subband of indices i and k, respectively. Therefore the gate voltage dependence of mobility in Fig.7 reflects the shifts of both the subband occupancy and subband mobility [15]. Using the concept of the average mobility, all the channel carriers available for conduction are like having an unified mobility. It is, therefore, very convenient to use the average mobility in device and circuits CAD. We propose the following simple linear form

$$\mu_a = \mu_0 [\, 1-\theta(V_{gs} - V_{th})] \qquad \text{for} \quad V_{gs} > V_{th}\ , \qquad (5)$$

to fit measured mobility data in Si δ-FETs. Here μ_0 and θ (>0) are two constants. (5) fits well experimental data, as shown by the solid line in Fig.7. μ_0=172 cm^2/Vs and θ=0.15 V^{-1} were obtained from the curve fitting.

V. Summary

Si delta-doped FETs have been fabricated using low-energy Sb ion doping technique to realize delta-doping profile during Si MBE growth and a low temperature budget in device processing. The FETs with a gate length of 2 μm showed a gate breakdown voltage of 3.8-4.7 V and a maximum extrinsic transconductance of 28 mS/mm, corresponding to an intrinsic value of 42 mS/mm. Those are the best reported so far in Si δ-FETs of similar gate length. The gate charge-control and low-field drift-mobility in Si δ-FETs were measured using C-V and I-V

techniques. Measured charge-control showed a strong nonlinearity and was in good agreement with theoretical result based on an "exact" quantum-mechanical model. We observed, for the first time, that the drift-mobility in the δ-FET was decreased with the increasing gate voltage. Both the nonlinearity in the charge-control and the gate voltage dependence of the drift-mobility would have large influences on dc and high-frequency performance of the device.

Acknowledgments

We would like to thank J. Carter, E.H. Thaki and E.R.A. Evans of Southampton University, U.K. for their collaboration in device processing, as well as W.-X. Ni and G.V. Hansson for Si MBE growth of the device structure. The work was supported by the Swedish National Board for Industrial and Technical Development (NUTEK).

References

1. E.F. Schubert et al, "Self-aligned enhancement-mode and depletion-mode GaAs field-effect transistors employing the δ-doping technique," *Appl. Phys. Lett.*, **49**, pp.1729-1731, 1986.
2. W.P. Hong et al, "DC and AC characteristics of Delta-doped GaAs FET," *IEEE Electron Device Lett.*, **EDL-10**, pp. 310-312, 1989.
3. K. Board and H.C. Nutt, "Double delta-doped FETs in GaAs," *Electron. Lett.*, **28**, pp. 469-471, 1992.
4. H.P. Zeindl et al, "Delta-doped MESFET with MBE-grown Si," *J. Electrochem. Soc.*, **136**, pp. 1129-1131, 1989.
5. K. Nakagawa et al, "Atomic-layer doped field-effect transistor fabricated using Si molecular beam epitaxy," *Appl. Phys. Lett.*, **54**, pp. 1869-1871, 1989.
6. R.G. Biswas et al, "Silicon boron delta doped FET: growth and fabrication," *Electron. Lett.*, **28**, pp. 667-669, 1992.
7. W.-X. Ni et al, "δ-function-shaped Sb-doping profile in Si (001) obtained using a low-energy accelerated-ion source during molecular-beam epitaxy," *Phys. Rev.B*, **46**, pp. 7551-7558, 1992.
8. Q. Chen and M. Willander, "Relationship between measured and intrinsic transconductances for FETs with gate conduction," *Solid-St. Electron.*, **35**, pp. 513-514, 1992.
9. S.M. Sze, *Physics of Semiconductor Devices*, 2nd ed., Wiley, New York, 1981, Chap.6.
10. M. Shur, *GaAs Devices and Circuits*, Plenum, New York, 1987, Chap. 10.
11. Y. Tsividis, *Operation and Modeling of the MOS Transistor*, McGraw-Hill, New York, 1987, Chaps. 4-6.

12. C.G. Sodini, T.W. Ekestedt, and J.L. Moll, "Charge accumulation and mobility in thin dielectric MOS transistors," *Solid-St. Electron.* **25**, 1982, p. 833-840.
13. Q. Chen and M. Willander, "Analysis of charge-control in Si delta-doped field-effect transistors," *J. Appl. Phys.*, **69**, 1992, pp. 8233-8236.
14. Q. Chen and M. Willander, "Analytical d.c. and small-signal a.c. modeling of Si delta-doped field-effect transistors," *Solid-St. Electron.* **35**, 1992, pp. 687-693.
15. Q. Chen et al, "Charge-control and mobility in Si delta-doped FETs grown by MBE," *Solid-St. Electron.*, **36**, 1993, in press.

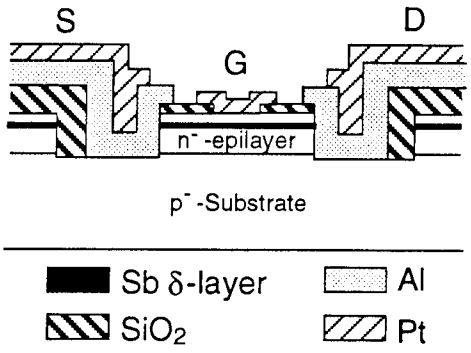


Fig.1 A schematic structure diagram of Si δ-FETs.

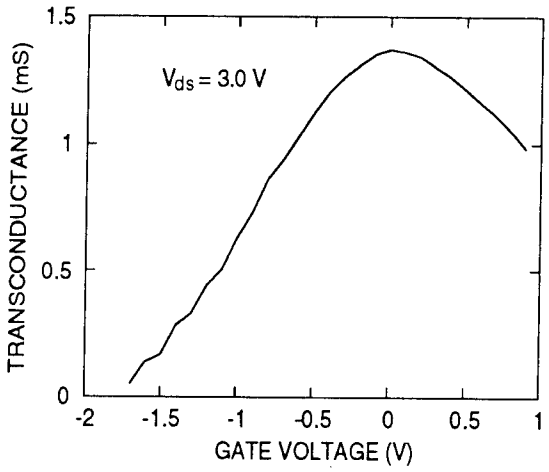
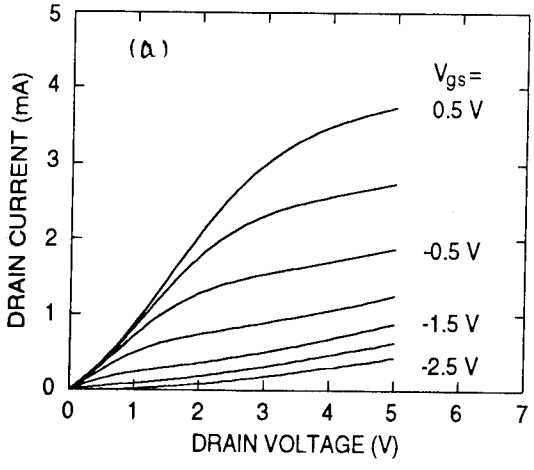


Fig.3 Measured extrinsic transconductance vs gate voltage for a drain voltage of 3.0 V.



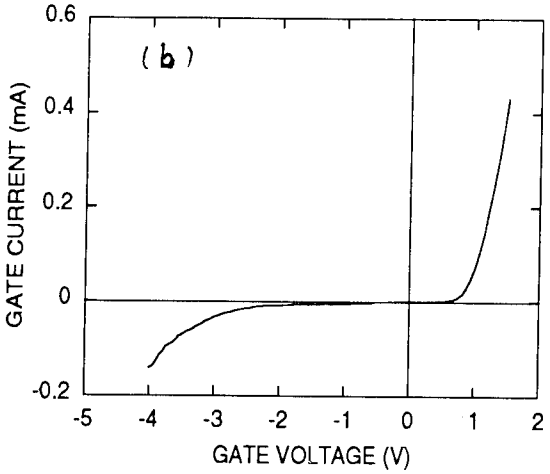


Fig.2 Typical drain output (a) and gate-source (b) current-voltage characteristics of 2 μm x 500 μm Si delta-doped FETs measured at 300 K.

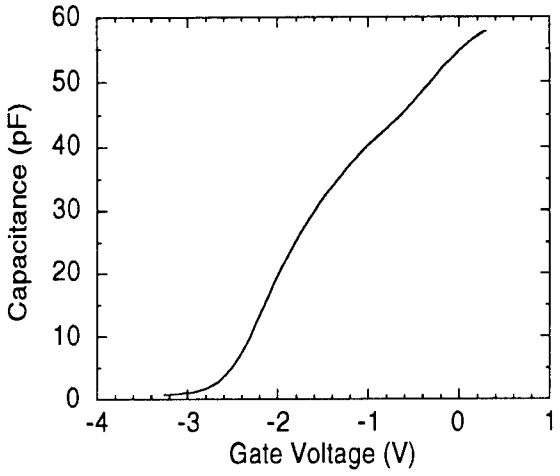


Fig.4 The gate-to-channel capacitance vs the gate voltage of Si δ-FETs measured at 300 K.

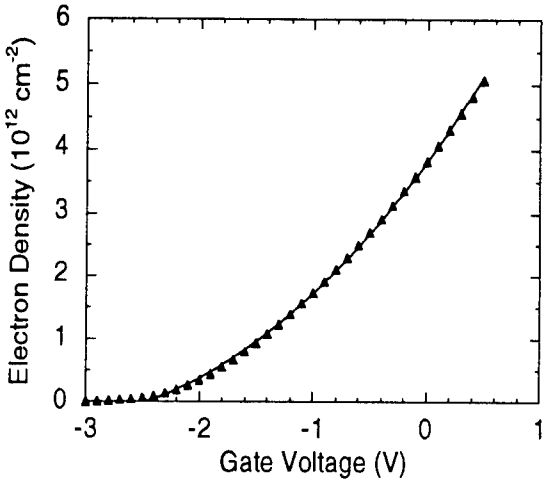


Fig.5 Electron sheet density vs the gate voltage obtained from the measurement in Fig.2. Solid line is the quadratic fitting to measured data using equation (2).

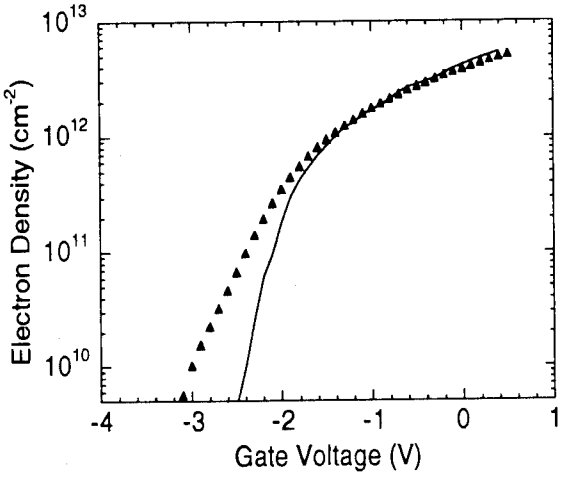


Fig.6 Comparison of measured (marks) and calculated (solid line) electron sheet densities vs the gate voltage.

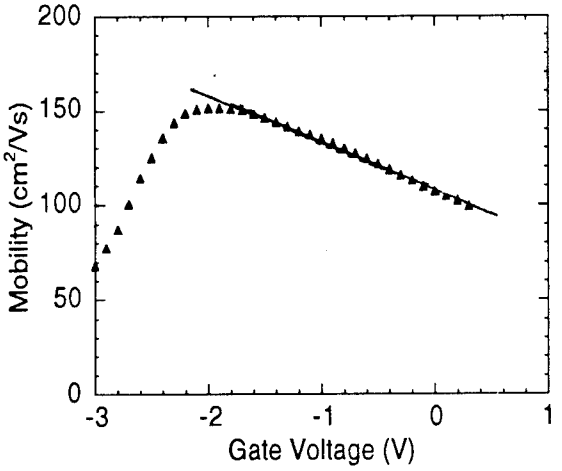


Fig.7 The average low-field electron drift-mobility vs the gate voltage. Solid line represents the linear fitting by $\mu_a=\mu_0[1-\theta(V_{gs}-V_{th})]$, for $V_{gs}>V_{th}$.

Simulation Of Optical Switches Using The Monte Carlo Method

†G.M. Dunn, †A .B. Walker,
‡J. H. Jefferson and ‡D. C. Herbert.

†School Of Physics,
University Of East Anglia,
Norwich, NR4 7TJ, U.K.
‡Electronics Division,
Defence Research Agency,
St. Andrews Rd.,
Malvern,
Worcestershire, WR14 3PS, U.K.

July 29, 1993

Abstract

We have calculated the response of a sub-micron GaAs optical switch to light pulses of wavelengths 799 nm, 734 nm and 633 nm.

It was found that the response of the total current (including displacement) at 799nm was characterized by an approximately exponentially decaying current, whilst at 633nm the current first rose to a peak before falling exponentially again. For the intermediate wavelength, the response was broadly exponential, though interrupted by a peak corresponding to the peak electron photocurrent.

The response of the electron photocurrent component was characterized by an initial rapid rise to a peak after about 2 picoseconds, followed by an exponential decay. The initial pulse reached its peak later and was broader for the higher energy radiation, the delayed response being explained by the larger population of electrons raised to the lower symmetry heavier bands at the higher creation energies. The hole response was largely unaffected by the energy of the incident radiation.

0-7803-0894-8/93 $3.00

1 Introduction

Optical devices are playing an increasingly important role in the electronics industry. Ever faster baud rates require faster optical switches and as such much attention is being devoted toward the optimization of these devices.

Relatively little theoretical work has been done in this field, largely because of the scarcity of bipolar simulation codes of the Monte Carlo variety. Of those that have been done Rosenzweig et al (1990) do not include displacement currents in their analysis and Peterson (1987) uses a drift diffusion method applied to silicon. Drift diffusion codes are unfortunately not well suited to the simulation of sub-micron GaAs devices [1] [2] [9] due to transient transport phenomena and as such only an exact solution of the Boltzmann transport equation will supply reliable solutions.

Of the various forms of switch, one of the most attractive is the GaAs Metal-Semiconductor-Metal (M.S.M.) type with interdigitated Schottky contacts which are well suited to monolithic integration and possess the qualities of a small capacitance and high carrier clearance rate (a 105 Ghz band width has been measured by Van Zeghbroek et al (1988)).

Although the favoured wavelength of the light used in this type of switch is toward the low energy part of the visible spectrum $\approx 850nm$, we believe it of interest to explore the effects of using different wavelengths and to this end we have used a 2 dimensional bipolar Monte Carlo simulation code to study the response a sub-micron optical switch to different energies of radiation.

2 Method

Monte Carlo simulation is a useful means of simulating the behaviour of small semiconductor devices. Unlike conventional (Drift Diffusion) simulation methods which assume the existence of equilibrium carrier transport conditions throughout the device, the Monte Carlo method provides an essentially exact solution of the Boltzmann transport equation and is prone only to statistical errors.

The method simulates the motion of charge carriers through a semiconductor device by following the progress of several tens of thousands of "super particles". Each super particle represents a number of real particles such that the charge applied to the super particles will reproduce the correct charge density. These particles are propagated classically between collisions according to their velocity, effective mass and the prevailing field. The selection of

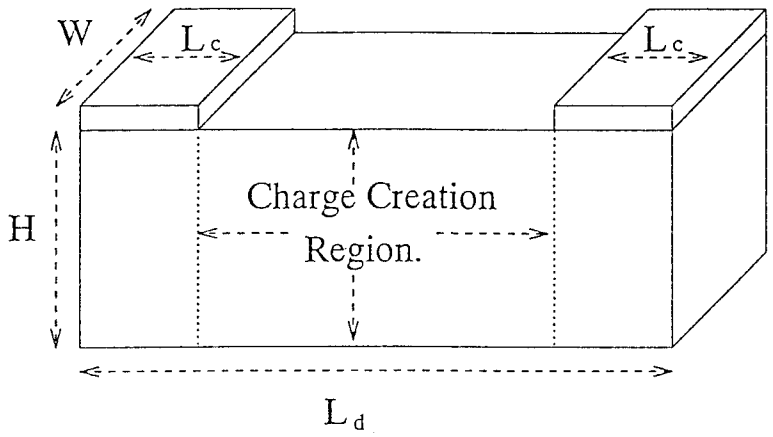


Figure 1: The Idealized Device Geometry of the simulation indicating the charge creation region. $W = \infty$, $L_d = 0.9\mu m$, $L_c = 0.2\mu m$ and $H = 0.5\mu m$.

the propagation time, scatter mechanism and other related quantities is done by generating a random number and using this number to select, for example a scatter mechanism, from a predetermined table of probabilities. A thorough discussion of the Monte Carlo method may be found in Fawcett, Boardman and Swain (1970). The model employed consisted of a 2 dimensional code simulating electron transport in GaAs in a three spherical band approximation and hole transport in a two warped band approximation. Details of the scatter rate formula and parameters used for the holes may be found in [3] and [12] and for the electrons in [7] [8] .

The model simulated 15000 electron-hole pairs created according to an assumed Gaussian pulse of radiation from a laser with a standard deviation of 144 fs. The incident radiation (and hence the created charge) was assumed to be attenuated with a decay length λ of 0.234 microns [15].

$$n = n_m exp\left(-\frac{x}{\lambda}\right) \tag{1}$$

The pulses were assumed to be of sufficient intensity to create a number of particles (n_m) equivalent to a density of $0.375 \times 10^{23} m^{-3}$. The potential was updated every 15 femto seconds using an exact matrix solution method of the L.U. type [13] on a mesh of 0.015 by 0.015 microns and a cloud in cell charge assignment and force interpolation scheme. The contacts were simulated using conventional "Schottky contacts" [11]. Recombination was not included in the simulation as the timescale of the "sweep out" of the device took place over the order 10^{-11} seconds, many orders of magnitude below the recombination time.

We calculated the response of the device to pulses of wavelengths 799 nm, 734 nm and 633 nm (equivalent to creation energies of 0.005 Hartrees, 0.01 Hartrees and 0.02 Hartrees under a bias of 0.1 Hartrees (2.72 Volts).

3 Results

Our results are best understood by first looking at the electron photocurrent component of the total current before adding the displacement current.

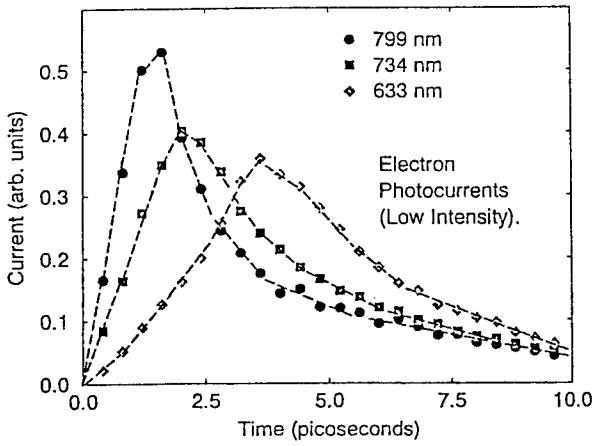


Figure 2: The electron photocurrents for the three wavelengths of incident radiation.

Figure 2 displays the the electron photocurrents at the three wavelengths of 799nm, 734nm and 633nm. The electron current response curves are broadened and the rise time of the initial peak substantially lengthened by light of shorter wavelength. The reason for this can be seen in figure 3 which shows the electron band populations at the wavelengths of 799 and 633nm respectively. Clearly, a substantially greater proportion of electrons created at the higher energy (corresponding to the lower wavelength) are quickly promoted out of the light Γ valley (band 1 in the figure) to the heavier lower symmetry bands (bands 2 and 3), thus delaying the response of the electrons. The anode and cathode total currents, broken down into their components are displayed in figure 4. As can be seen, the first picosecond of the total current is dominated by the displacement current which is caused by an increase of electric field at the anode and cathode. This is due to electrons and holes moving under the initial electric field causing highly charged regions to either side of the charge

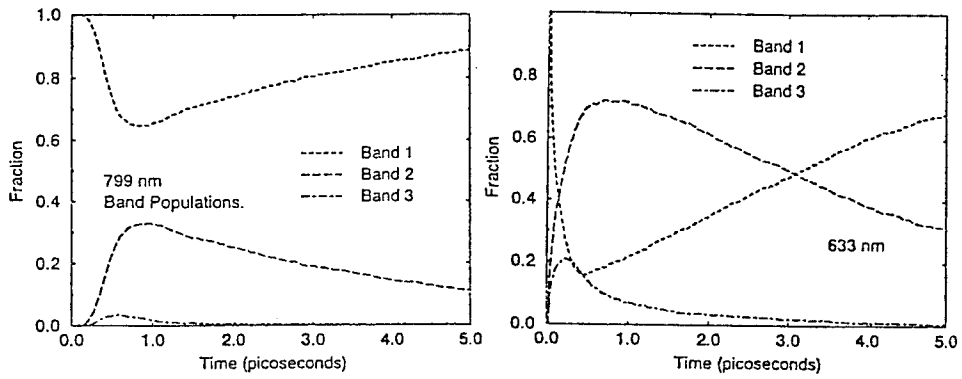


Figure 3: The electron band populations at the wavelengths of 799 and 633nm.

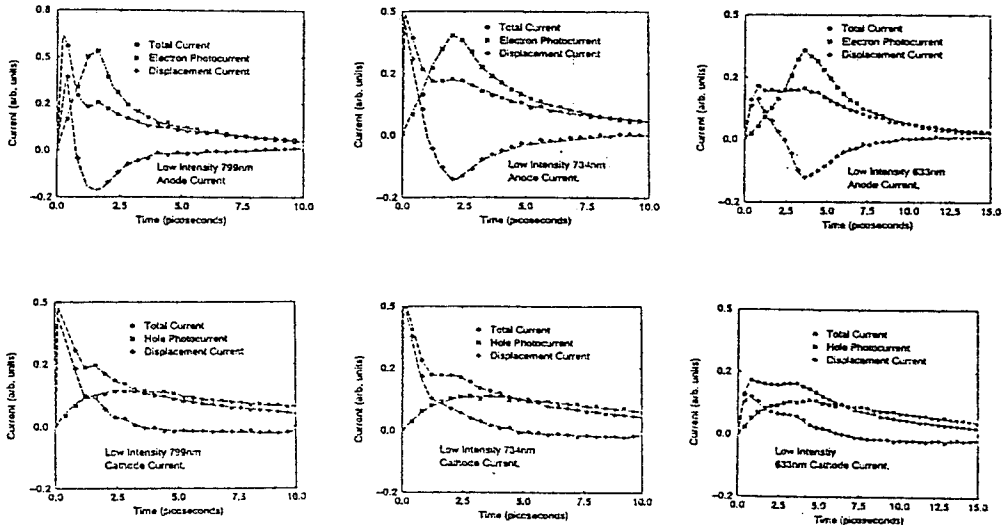


Figure 4: The anode and cathode currents at each incident wavelength.

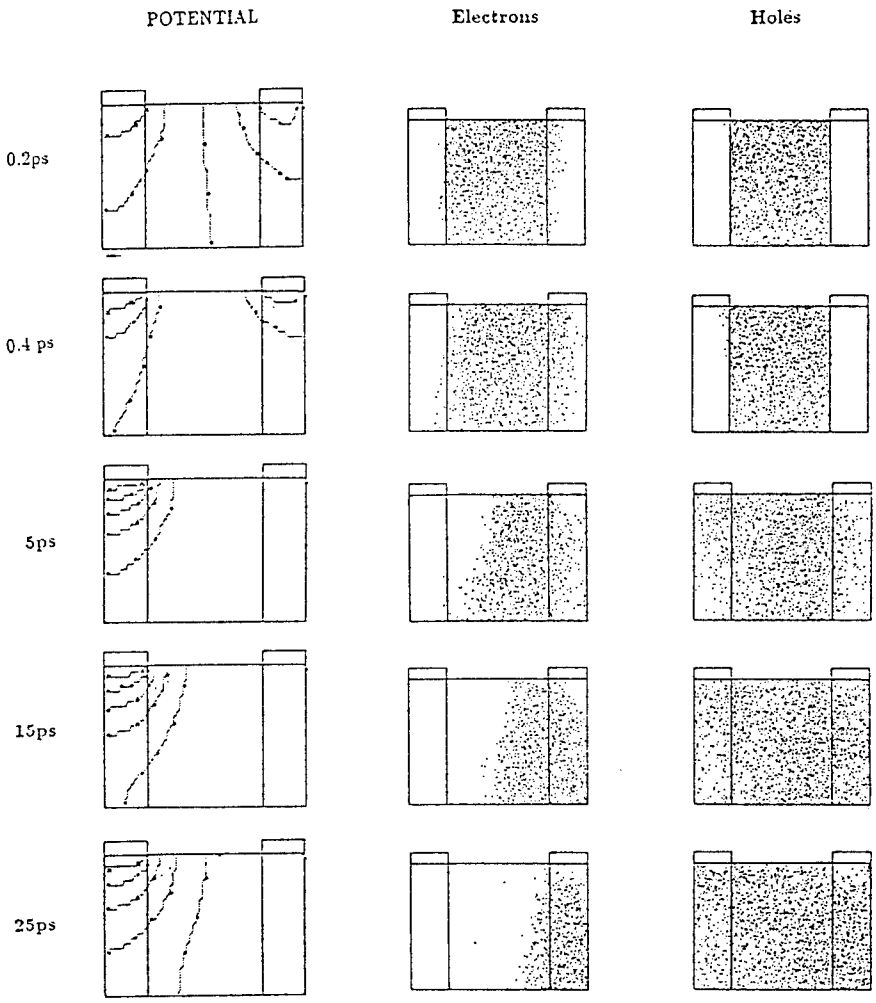


Figure 5: The evolution of the potential (contours at intervals of 0.54 V) and electron and hole densities (each point corresponds to the position of a superparticle in the simulation). The anode is to the right.

creation region. This movement is much slower in the 633nm case than the 799nm case due to the higher effective mass of the electrons, and so explains the difference in the initial rise time of the current.

As the electrons start to pass through the anode, the total charge density on the anode side begins to reduce (both because electrons are being lost through the contact and because holes are also diffusing into the region (see figure 5)). This in turn causes a reduction in the electric field at the anode with a corresponding displacement current in the opposite direction and similar magnitude to the electron photocurrent. Eventually the electric field at the anode becomes very small with most of the potential drop occurring over the positively charged cathode side.

The current is then dominated by the photocurrent on each contact with electrons diffusing out of the anode and holes being driven out of the cathode under the high electric field. As the electrons are lost, the extent of the charge neutral region on the anode side reduces (see figure 5 at 15 and 25ps) with corresponding changes of potential. In the final stages (after about 50ps) when most of the electrons are lost from the device, the current on the anode side once again has a significant displacement current component as the electric field on the anode returns to its' initial value.

3.1 Errors

Figure 6 shows the charge and current through each contact compared at 734nm. During the initial rapid transient the error in the charge conservation is approximately 7% but this quickly falls to much less than 1 % non-charge conservation. Much of this error is due to the numerical integration of the electric field over the contact which in effect meant that the number of points were the electric field was known was limited by the mesh size (14 points across the contact). Smaller sources of error arise from the lack of consideration of magnetic field components in the simulation and the coarseness of the super-particle to real particle ratio in the simulation which in the case of the lower hole currents reveals itself as "roughness" in the current curves.

It should be stressed that the simulation was 2 dimensional and as such assumed that the depth (H in figure 1) of the swit ı was infinite.

4 Conclusion

The initial current is dominated by the displacement current, the rise time is determined by how quickly highly charged regions can be formed to either

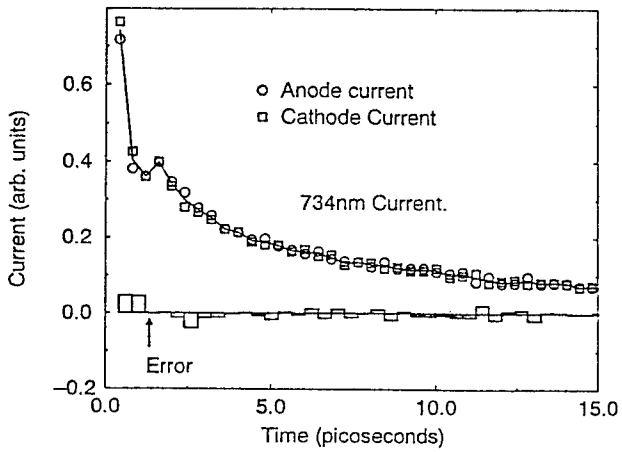


Figure 6: The current through each contact compared for at 799nm showing the error in charge conservation.

side of the charge creation region.

We can see no (speed) advantage in using higher energy radiation as this causes promotion of the electrons to higher energy and heavier bands which, as just mentioned, slows the initial displacement response and also slows the passage of the electron photocurrent. In addition the slower photocurrent can cause undesirable structure to appear in the current pulse due to the late arrival of electrons at the anode, rather than merging smoothly with the initial displacement current pulse.

An overall faster response (in terms of clearing the device of charge carriers) can be achieved by lowering the intensity of the incident radiation, thus reducing the amount of screening at the cathode and so maintaining a strong electric field over the entire device.

References

[1] Baccarani G., Jacobini C. and Mazzone A. M., Solid State Electronics, **20** (1977), p. 5. MC elec transport accross base compared to DD.

[2] Baccarani G. and Wordeman M. R., Solid State Electronics **28**, (1985), p. 407. DD overshoot in silicon - fudged model.

[3] Brudevoll T., Lund B. and Fjeldly T. A.,J. Appl. Phys. **71**, (1992), p. 4972. MC warm and hot hole drift in GaAs

[4] Brudevoll et al., J. Appl. Phys. **67**, (1990), p. 737.

[5] Costato M., Jacoboni C. and Reggiani L., Phys. Stat. Sol. B, **52**, (1972), p. 461. MC Hole Transport in Polar Semicond.

[6] Costato M. and Reggiani L., Phys. Stat. Sol. B, **58**, (1973) p. 471. MC Hole Transport in cubic semiconduct.

[7] Fawcett W., Boardman A. D. and Swain S., J. Phys. Chem. Solids, **1970**, p. 1963. MC elec trans. in GaAs.

[8] Fischetti M. V., IEEE Trans. Elec. Dev., **38**, (1991), p. 634. MC elec trans - parameters part 1.

[9] Herbert C. Semicond. Sci. Technol. **6** (1991) p. 405. Ballistic corrections to base transit.

[10] Hinckley J. M. and Singh J., Appl. Phys. Lett., **53**, (1988), p. 785. Hole transport in strained GaAs.

[11] Hockney R. W., and Eastwood J. W., "Computer Simulations Using Particles", Adam Hilger, Bristol and New York. (1989).

[12] Jacoboni C. and Lugli P., Ed. Selberherr S., "The Monte Carlo Method For Semiconductor Device Simulation", Springer-Verlag Wien New York.

[13] "Numerical Recipes", Press E.H., Flannery B.P., Teukolsky S. A. and Vetterlin W.T., Cambridge University Press (1987).

[14] Reggiani L., J. Phys. Chem.Solids, **37**, (1976), p. 293.

[15] Rosenzweig et al, SPIE **1362** Phys. Concep. of Mat. for Nov. Opto. Dev. App.: Dev. Phys. and App., (1990), p.168.

[16] Schockley W. and Read W. T., Phys. Rev., **87**, (1952), p. 835. Elec hole Recombination

[17] Van Zeghbroek B. J., Patric W., I dbout J. M. and Vettiger P., IEEE Elec. Dev. Lett, **9**, (1988), p. 527.

Application of the First-Order Bipolar Model to Harmonic Distortion Analysis of HBT's

Bahman Meskoob and Sheila Prasad
Department of Electrical and Computer Engineering
Northeastern University, Boston, Massachusetts, 02115

ABSTRACT

Although large-signal models for GaAs MESFETs have been developed extensively, the same is not true for heterojunction bipolar transistors (HBT's). A first-order model, based on first-order device physics, is used for harmonic distortion analysis of an inverted HBT. It is shown that the first-order model is adequate at medium power levels and inadequate at higher power levels where the device is driven into hard saturation and cutoff.

I. INTRODUCTION

With advances in the fabrication of heterojunction bipolar transistors (HBT's), there have been numerous articles covering various aspects of the HBT performance [1]-[2]. Kim et. al. [1] have reviewed the GaAs device fabrication and IC technology in detail. An advantage of HBT's is their high linearity. Third-order intermodulation distortion points (IP3) up to 35 dBm have been reported [1]. Efforts in both linear and nonlinear modeling of HBT's have been ongoing [3]-[7].

This paper presents the application of the first-order bipolar model to the simulation of HBT nonlinearity. The quasi-static method based on small-signal measurements and modeling at various bias points is used. Equations governing nonlinearity of the intrinsic elements are deduced and used in a harmonic balance simulator. The simulated second and third harmonics are then compared to corresponding measurements. Although, by virtue of its size and I-V characteristics, the device used here is suitable for small-signal operation; the large-signal modeling presented here demonstrates the limitations of the first-order model.

II. DEVICE MODELING

The device used here is a 5 x 10 μm^2 InGaAs/InAlAs/InP inverted HBT with an f_T of 24 GHz. The layer structure and the I-V characteristics of the device are shown in Figures 1 and 2, respectively. The fabrication and performance of this device have been reported in [2]. Small-signal modeling efforts and investigation of the bias-dependence of the intrinsic elements have also been reported [5], [8].

0-7803-0894-8/93 $3.00

For large-signal simulations, the quasi-static modeling technique based on small-signal measurements is used. Small-signal s-parameters are measured over a wide frequency band (0.05 to 26 GHz) and bias conditions from cutoff to saturation. Thirty nine bias points are chosen within the range of base current (I_b) from 0 to 250 μA and collector-emitter voltage (V_{ce}) from 0.6 to 2 V. The small-signal model, shown in Fig. 3, is fitted to the measured s-parameters at each bias point. This is done using the commercial software Touchstone [9]. The parasitic elements, C_{be}, C_{bc}, C_{ce}, L_b, L_e, L_c, and contact resistors, R_b, R_c, R_E, and the intrinsic base resistor (included in R_b) are assumed to be linear and invariant with bias. The bias-dependent elements are R_e, C_e, C_c, α_o, and g_o. It has been shown that consideration of the bias variation of these elements is sufficient for accurate small-signal modeling [8]. The maximum normalized errors of the small-signal fits averaged over 39 bias points are 6% for S_{11}, 11% for S_{21}, 8% for S_{12}, and 8 % for S_{22} at the fundamental, second, and third harmonic frequencies of 6, 12 and 18 GHz. The details of modeling efforts over the bias region can be found in [8].

The following equations are used for the characterization of intrinsic elements

$$R_e = \frac{v_t}{i_e} \qquad \text{where} \quad i_e = i_o \left(e^{v_{be}/v_t} - 1\right)$$

$$C_e = c_{je} + \tau_o \cdot g_e \qquad \text{where} \quad g_e = \frac{1}{R_e}$$

$$C_c = \frac{c_{jc}}{\sqrt{1 + \frac{v_{cb}}{v_o}}}$$

$$g_o = \frac{i_e}{v_e}$$

$$\alpha_o = a_0 + a_1 \cdot v_{ce} + a_2 \cdot i_e$$

where all voltages and currents are those of the intrinsic device. The first three equations are given by first-order device physics. The fourth equation is also well known, where v_e is the Early-voltage of the device. Given the fact that the device model neither accounts for saturation nor avalanching, the last equation is included in order to account for variations of α_o with bias. These equations are fitted to the bias variation of the intrinsic element values and the corresponding constants are determined.

During small-signal measurements, the base current (I_b) and the external collector-emitter voltage (V_{ce}) constitute the sources while the collector current (I_c) and the external base-emitter voltage (V_{be}) are measured. Given the external bias voltages and currents, the intrinsic voltage and current values need to be determined. Since the element values deduced from small-signal s-parameters are not consistent with the measured external DC bias, a hybrid approach is used in which the bias values of I_b and V_{ce} and the values of α_o and g_o as determined from small-signal modeling are taken as independent variables. The two parameters v_t and i_o are then adjusted for a fit between the measured and modeled R_e and the measured and modeled V_{be}. Given v_t and i_o,

the intrinsic bias values are determined and used for determination of the model parameters c_{je}, τ_o, c_{jc}, v_o, v_e, a_0, a_1, a_2. The model parameters are listed in Table 1. Figures 4 through 8 present plots of the intrinsic element values, R_e, C_e, C_c, g_o, α_o, as determined by small-signal modeling together with those determined from equations of the first-order model. The two parameters in Figures 4 through 8 are the base current (I_b) and the external collector-emitter voltage (V_{ce}) with increments of 50 μA and 0.2 V respectively. These figures show that the first-order equations are not adequate for modeling C_e and C_c over their entire bias range. In addition, a higher order variation for α_o with respect to i_e should be considered.

III. SIMULATION OF HARMONIC DISTORTION

A harmonic balance simulator was developed for a personal computer incorporating the model equations. A simulation was performed at 6 GHz and bias of V_{ce} = 1.2 V and I_b = 150 μA corresponding to I_c of 4.5 mA . Figure 9 shows the measured and simulated fundamental, second, and third harmonics. The maximum deviation between the measurements and simulations are: fundamental, 0.5 dB, second harmonic, 3.4 dB, and third harmonic, 7.3 dB. It is evident that at output power levels exceeding -3 dBm, corresponding to 4.5 mA of peak swing across the output 50 Ω load, the simulated second harmonic response begins to deviate from the measurement. This demonstrates the limitations of the model as it does not account for hard operation within the saturation and cutoff region. A more elaborate model taking into account the forward operation of the base-collector junction should result in an improved simulation. There is only limited success in simulating the third harmonic. The slopes of the simulated and measured third harmonic distortion are not the same. Although the slope of the simulated third harmonic distortion is 3, the measured distortion has a slope of 2.5. However, at higher power levels the third harmonic simulation does approach the measurement.

IV. CONCLUSION

The use of the first-order bipolar model for harmonic distortion simulation was investigated. The values of the intrinsic elements as deduced from small-signal modeling were fitted to and compared with those of the first-order bipolar model. The device modeling technique was outlined. A harmonic balance simulation was performed demonstrating the use and limits of the model. In summary, the first-order bipolar model is adequate for modeling the second harmonic distortion up to power levels where the device is driven into saturation or cutoff; and its use for modeling the third harmonic distortion is limited to higher power levels.

ACKNOWLEDGMENT

The authors wish to thank Dr. Michael E. Kim for providing a grant from TRW, Inc. which partially supported this research. They also wish to thank Prof. C. G. Fonstad of the

Massachusetts Institute of Technology for providing the devices and the use of laboratory facilities for this work.

REFERENCES

[1] M. E. Kim, A. K. Oki, G. M. Gorman, D. K. Umemoto, J. B. Camou," GaAs heterojunction bipolar transistor device and IC technology for high performance analog and microwave applications," IEEE Trans. Microwave Theory Tech., vol. MTT-37, pp. 1286-1303, Sept. 1989.

[2] H. Sato, J. Vlcek, C. G. Fonstad, B. Meskoob, S. Prasad," InGaAs/InAlAs/InP collector-up microwave heterojunction bipolar transistors," IEEE Electron Device Lett., vol. EDL-11, pp. 457-459, Oct. 1990.

[3] M. E. Hafizi, C. R. Crowell, M. E. Grupen," The DC characteristics of GaAs/AlGaAs heterojunction bipolar transistors with application to device modeling," IEEE Trans. Electron Devices, vol. ED-37, pp. 2121-2129, Oct. 1990.

[4] D. R. Pehlke, D. Pavlidis," Direct calculation of the HBT equivalent circuit from measured s-parameters," 1992 IEEE MTT-S Int. Microwave Symp. Dig., pp. 735-738.

[5] B. Meskoob, S. Prasad, M-K. Vai, C.G. Fonstad, J.C. Vlcek, H. Sato, C. Bulutay," A small-signal equivalent circuit for the collector-up InGaAs/InAlAs/InP heterojunction bipolar transistor," IEEE Trans. Electron Devices, vol. ED-39, pp. 2629-2632, Nov. 1992.

[6] S. Maas, B. Nelson, D. Tait," Intermodulation in heterojunction bipolar transistors," 1991 IEEE MTT-S Int. Microwave Symp. Dig., pp. 91-93.

[7] A. Samelis, D. Pavlidis," Analysis and optimization of third order intermodulation distortion mechanisms in AlGaAs/GaAs heterojunction bipolar transistors," 1992 IEEE MTT-S Int. Microwave Symp. Dig., pp. 1587-1590.

[8] B. Meskoob, S. Prasad, M-K. Vai, J.C. Vlcek, H. Sato, C.G. Fonstad," Bias-dependence of the intrinsic element values of InGaAs/InAlAs/InP inverted heterojunction bipolar transistor," IEEE Trans. Microwave Theory Tech., vol. MTT-40, pp. 1012-1014, May 1992.

[9] "Touchstone, Version 2.1," EEsof, Inc., 1990.

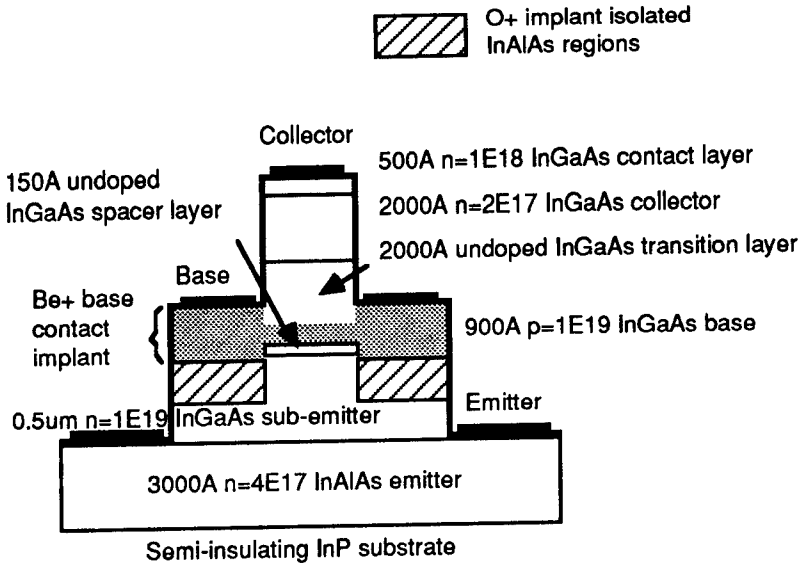


Fig. 1 Vertical structure of the 5x10 μm^2 InGaAs/InAlAs/InP inverted HBT.

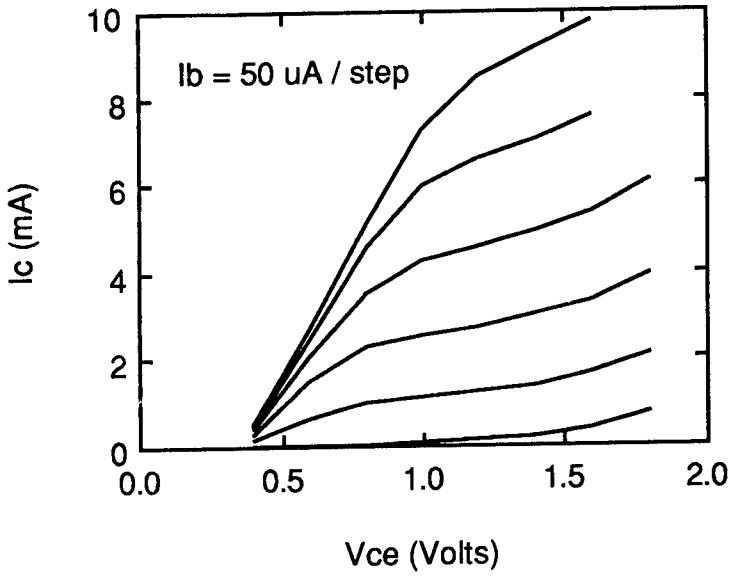
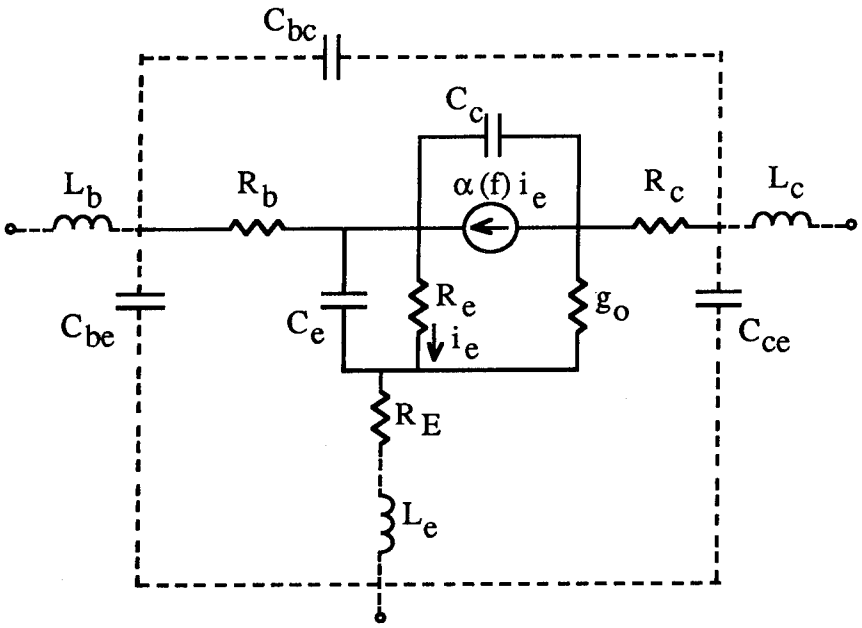


Fig. 2 Collector current (I_c) versus collector-emitter voltage (V_{ce}) transfer characteristics of the HBT.



R_b	170 Ohms
R_c	17 Ohms
R_E	50 Ohms
τ_c	1.42 pS
τ_b	0.72 pS
L_b	30 pH
L_c	49 pH
L_e	10 pH
C_{be}	40 pF
C_{ce}	31 pF
C_{bc}	6.3 pF

$$\alpha(f) = \alpha_0 \frac{\sin(\omega\tau_c/2)}{\omega\tau_c/2} \frac{e^{-j\omega(\tau_c/2+m\tau_b)}}{1+j\omega\tau_b}, \quad m = 0.22$$

Fig. 3 Equivalent circuit model of the HBT.

Table 1 Parameters of the first-order model's intrinsic elements.

v_t	25 mV	v_o	1.0 V
i_o	1e-16 A	v_e	2.0 V
c_{je}	0.13 pF	a_0	0.915
τ_o	3 pS	a_1	0.03 V^{-1}
c_{jc}	0.033 pF	a_2	4.0 A^{-1}

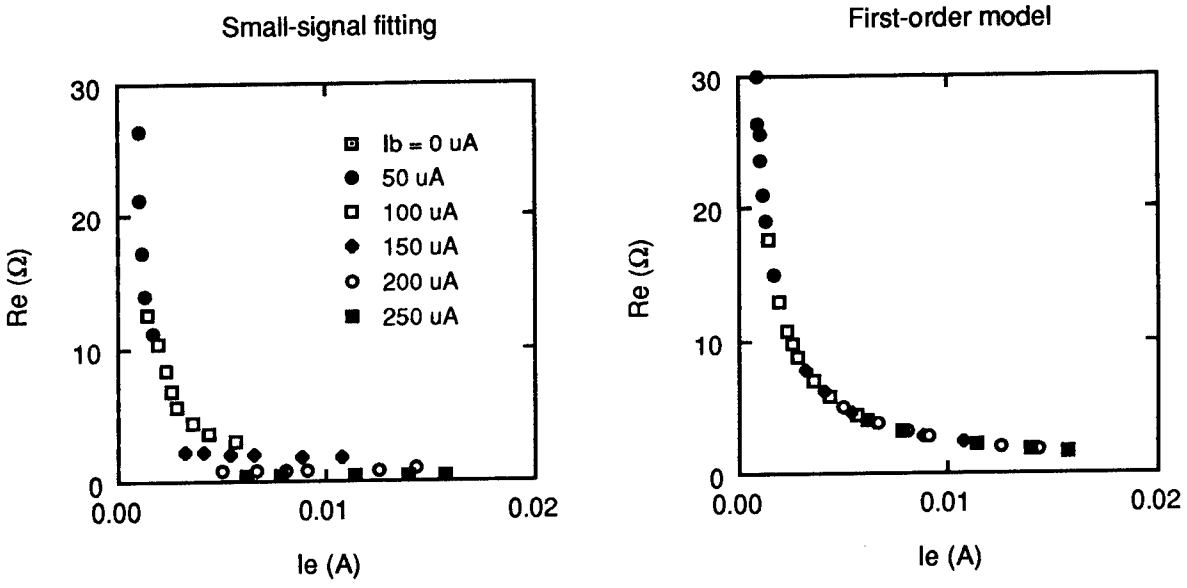


Fig. 4 Plot of R_e versus the internal emitter current (I_e) with the base current (I_b) and the external collector-emitter voltage (V_{ce}) as parameters.

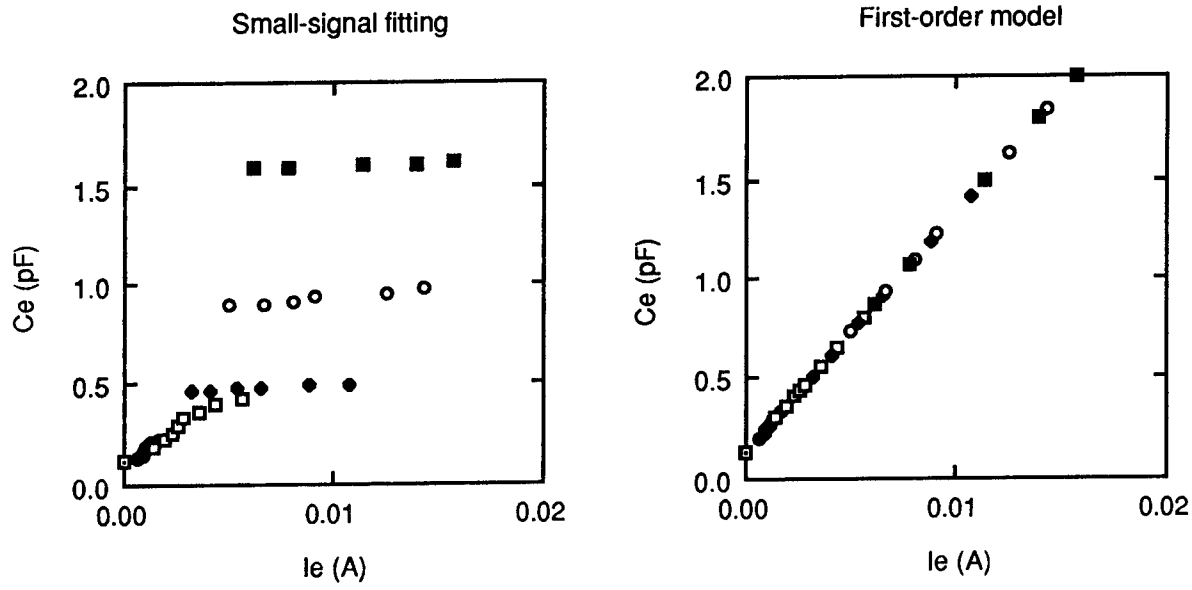


Fig. 5 Plot of C_e versus the internal emitter current (I_e) with the base current (I_b) and the external collector-emitter voltage (V_{ce}) as parameters.

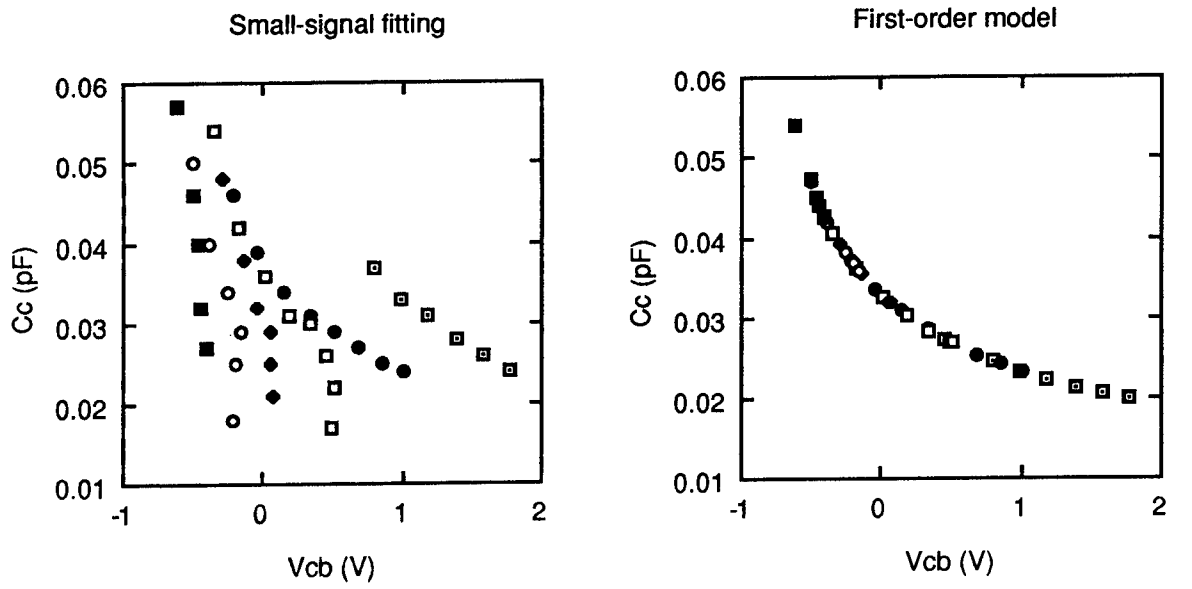


Fig. 6 Plot of C_c versus the internal collector-base voltage (V_{cb}) with the base current (I_b) and the external collector-emitter voltage (V_{ce}) as parameters.

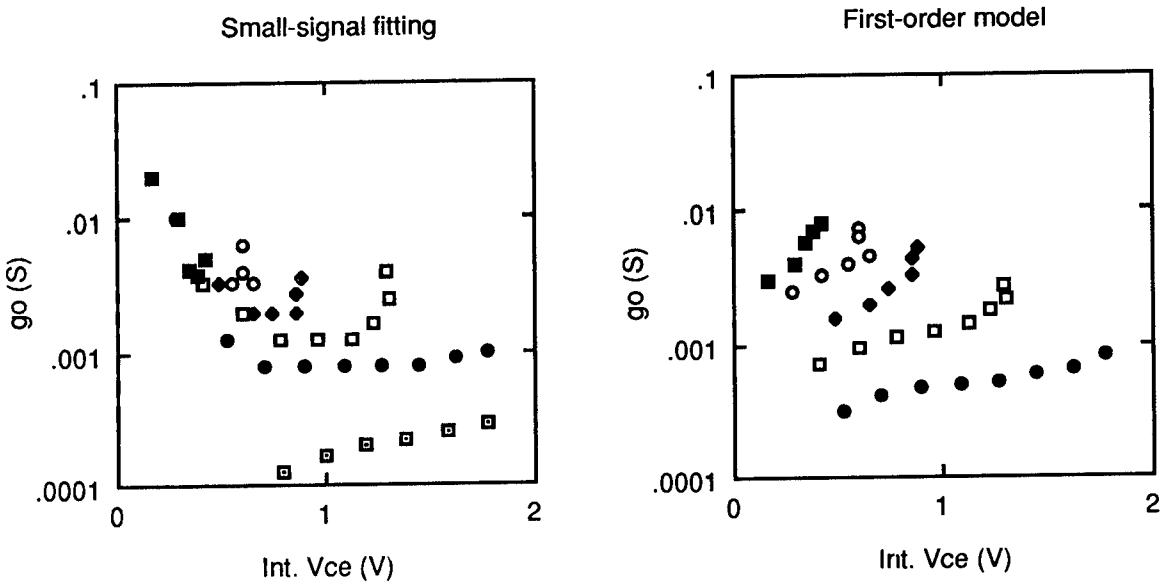


Fig. 7 Plot of g_o versus the internal collector-emitter voltage (Int. V_{ce}) with the base current (I_b) and the external collector-emitter voltage (V_{ce}) as parameters.

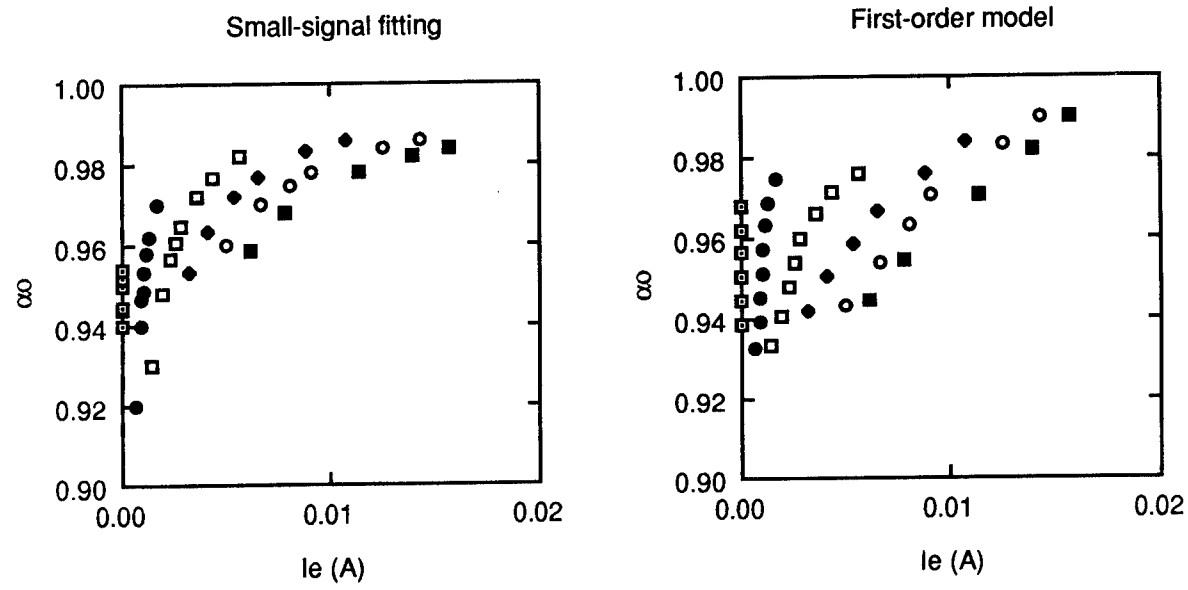


Fig. 8 Plot of α_o versus the internal emitter current (I_e) with the base current (I_b) and the external collector-emitter voltage (V_{ce}) as parameters.

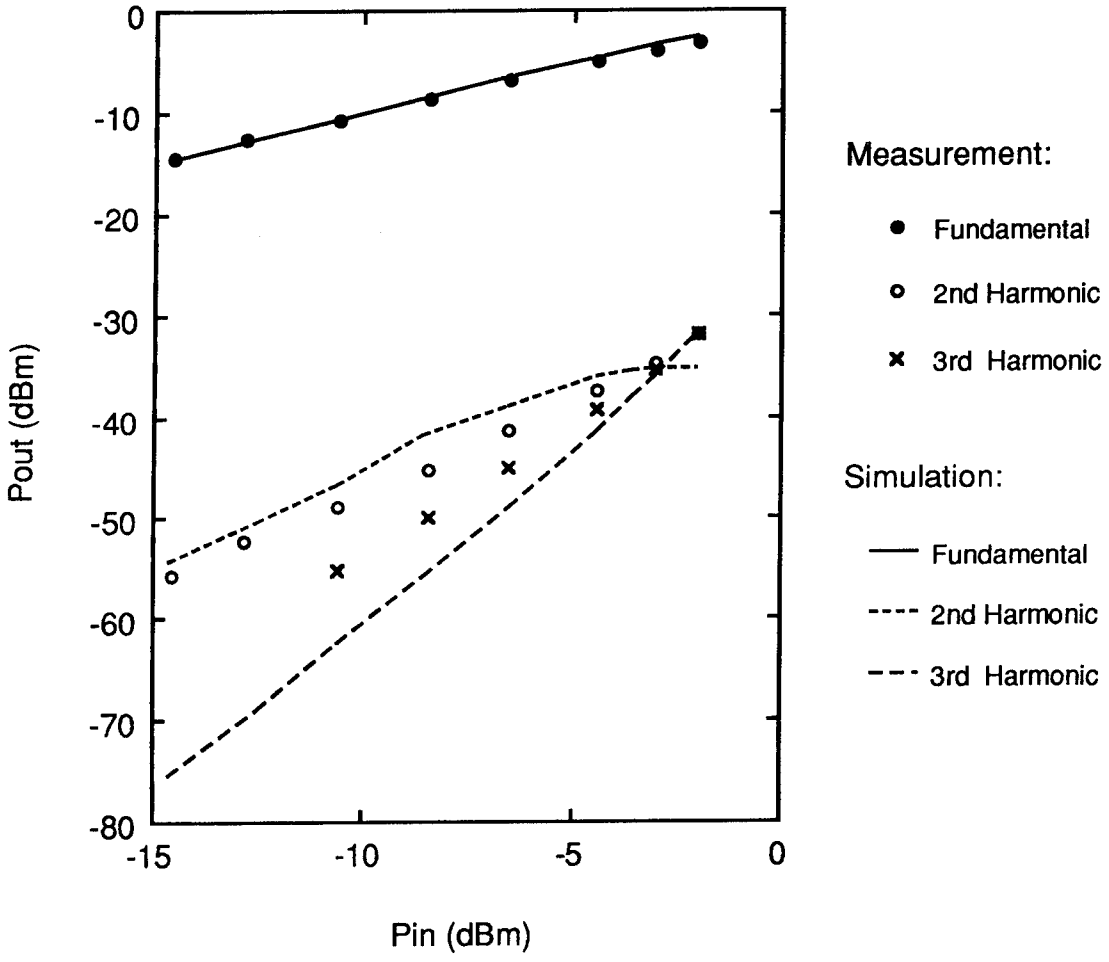


Fig. 9 Measured and simulated fundamental, second and third harmonic distortion of the HBT.

Thermal Management in Strongly Non-Planar Microwave HBTs

Jürgen Schneider, Hermann Schumacher and Uwe Erben

Dept. of Electron Devices and Circuits
University of Ulm
D-89069 Ulm
Germany

Abstract

We have systematically investigated technological parameters of $Al_{0.3}Ga_{0.7}As/GaAs$ power heterojunction bipolar transistors (HBTs) using finite-element modelling, taking into account the temperature dependence of the semiconductor materials involved and a realistic nonplanar mesa structure. Our calculations include a multilayer structure and raise the question of the magnitude of heat transfer via top metal contacts. Equally important, for the case of multielement power HBTs with close element spacing, the lateral spread of the heat flow is inhibited by the neighbouring elements and the dependence of the thermal resistance on element spacing and substrate thickness can be approximated by a very simple analytical relationship.

I. INTRODUCTION

Power amplification at microwave frequencies using heterojunction bipolar transistors has received significant publicity of the past few years. Most noteably, HBTs offer high output power per unit length and excellent power-added efficiencies. Contrary to FETs which are largely limited by maximum current and breakdown voltage, HBTs are limited by thermal constraints – their full potential can only be utilized under pulsed, low-duty-cycle operation [1]. The management of heat currents inside HBT structures becomes a serious concern in the design of microwave power HBTs when maximized average output power is the goal. Most researchers have, up to now, approached the thermal simulation of these devices as a *planar*, single- or multilayer problem [2]–[4]. Real microwave HBTs, however, have a strongly *nonplanar* mesa structure composed of vertical sections of semiconductor materials with varying composition and doping. The thermal conductivity of III-V semiconductor materials like $GaAs$ and InP is much worse than that of Silicon and, as the latter, decreases significantly with increasing temperature [5]. In case of ternary semiconductors like $AlGaAs$, the thermal conductivity is also a strong function of the stochiometry [6].

As our numerical calculations show, the lateral thermal flow in a one–emitter HBT and in a multi–cell HBT leads approximately to a homogeneous heat transfer through the substrate. This homogeneous heat flow can be discribed by a simple analytical formula, so that predictions about the thermal behavior of a power HBT for different substrate materials, for different lateral dimensions and for different substrate thicknesses can be made. In case of the one–emitter HBT, when the geometry is small compared to the substrate thickness and chip dimensions, the heat will diffuse laterally and vertically until the chip borders are reached, at which point the flow will become near-homogeneous in the vertical direction.

To achieve large output powers from a single transistor chip, practical power HBTs combine many transistor cells in close proximity [7]. In such a case, the dissipated heat from the neighbouring elements decreases the lateral temperature gradient and consequently the lateral flow. This means that the cross-section of substrate through which the heat can be passed gets drastically smaller for all but the boundary elements, leading to an increase in the thermal resistance which can also be described by an analytical formula for an homogenous heat flow through the substrate.

0-7803-0894-8/93 $3.00

Owing to the rather poor thermal properties of popular microwave semiconductors, means to improve heat-sinking are in great demand. Besides the usual thinning of the substrate, placement of thin-film devices on substrates of superior thermal conductivity and thermal transfer via thick surface metal contacts have been investigated in the context of this work.
Collector–up heterojunction bipolar transistor structures have been the scope of work by several authors [8],[9]. They possess an inherently small base-collector capacitance, significantly increasing the maximum frequency of oscillation over emitter-up transistors with the same base sheet resistance. This property makes them attractive also for microwave power applications. As a demonstrator for the techniques employed in our calculation method, we will show how, in case of $Al_{0.3}Ga_{0.7}As/GaAs$ devices, the inverted structure adversely affects the junction temperature.

II. GENERAL CONSIDERATIONS

As a twodimensional model for the unit cell of a typical power HBT we chose the layer structure and geometry depicted in Fig. 1. As can be seen, the emitter width was $2\,\mu m$, the base-emitter spacing $0.2\,\mu m$, the base width $100\,nm$, the collector width $1\,\mu m$, and the substrate thickness $100\,\mu m$, to highlight the more important parameters.

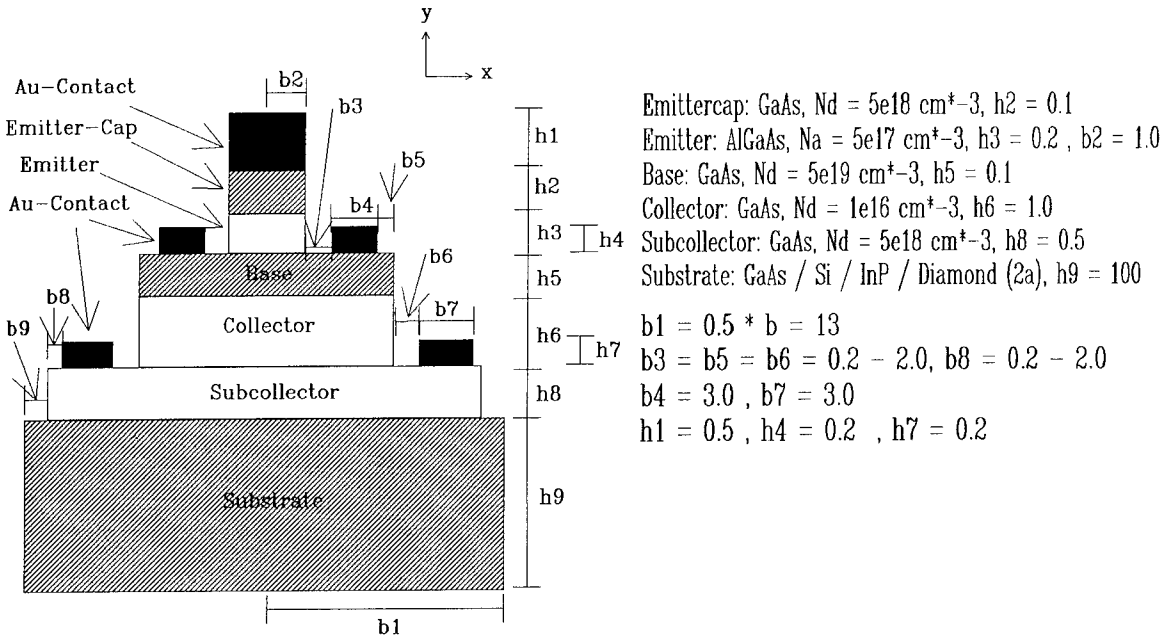


Figure 1: Twodimensional model for one cell of an $AlGaAs/GaAs$–HBT. All dimensions are in μm.

The main quantity in a thermal modell for a semiconductor device is the thermal conductivity λ. In semiconductors, λ shows approximately a $1/T$ temperature dependence. This temperature dependence results from heat transport mainly by acoustic phonons. An increase in doping concentration leads to an increased scattering of acoustic phonons by the quasifree electrons and holes, so that λ decreases [5]. In the ternary semiconductor $Al_xGa_{1-x}As$, where typically $x = 0.3$ in the emitter of an HBT, one observes an important reduction of the thermal conductivity compared to $GaAs$. This observation can be explained by the increased disorder in the $GaAs$ crystal lattice due to the substitution of Ga by Al, leading to a remarkable mass difference scattering of the acoustic phonons [6], and reducing the thermal conductivity

as a consequence

$$\lambda_{GaAs}(300K) \approx 3.6 \cdot \lambda_{Al_{0.3}Ga_{0.7}As}(300\,K)\,. \tag{1}$$

So for thermal considerations the HBT described above represents an inhomogeneous body whose local thermal conductivity is influenced by the varying material ($Al_{0.3}Ga_{0.7}As$, $GaAs$, Au metallizations) and the varying doping in the semiconductor regions. As a consequence λ gets an explicit dependence on the position $\mathbf{x}$ and the nonlinear heat equation for our semiconductor device can be written in the form

$$-\left(\nabla(\lambda(\mathbf{x},T))\,\nabla T + \lambda(\mathbf{x},T)\,\Delta T\right) = \eta_J(\mathbf{x},T)\,, \tag{2}$$

where η_J represents the dissipated power density. In our modell we assumed abrupt interfaces between the individual regions, so that the gradient of λ at the interfaces is no longer defined, and as a consequence we must split the nonlinear heat equation (2) into an eleven-dimensional system of partly nonlinear partial differential equations of the second order, which are coupled via the boundary conditions:

$$\lambda_{Au,e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)T = 0 \tag{3}$$

$$\frac{\partial}{\partial x}\left(\lambda_{ec}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{ec}(T)\frac{\partial T}{\partial y}\right) = 0 \tag{4}$$

$$\frac{\partial}{\partial x}\left(\lambda_{e}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{e}(T)\frac{\partial T}{\partial y}\right) = 0 \tag{5}$$

$$\frac{\partial}{\partial x}\left(\lambda_{b}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{b}(T)\frac{\partial T}{\partial y}\right) = 0 \tag{6}$$

$$\lambda_{Au,b1}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)T = 0 \tag{7}$$

$$\lambda_{Au,b2}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)T = 0 \tag{8}$$

$$\frac{\partial}{\partial x}\left(\lambda_{c}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{c}(T)\frac{\partial T}{\partial y}\right) = -\eta_J \tag{9}$$

$$\frac{\partial}{\partial x}\left(\lambda_{sc}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{sc}(T)\frac{\partial T}{\partial y}\right) = 0 \tag{10}$$

$$\lambda_{Au,sc1}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)T = 0 \tag{11}$$

$$\lambda_{Au,sc2}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)T = 0 \tag{12}$$

$$\frac{\partial}{\partial x}\left(\lambda_{su}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda_{su}(T)\frac{\partial T}{\partial y}\right) = 0\,. \tag{13}$$

The thermal conductivities $\lambda_{Au,e}$, λ_{ec}, λ_e, λ_b, $\lambda_{Au,b1}$, $\lambda_{Au,b2}$, λ_c, λ_{sc}, $\lambda_{Au,sc1}$, $\lambda_{Au,sc2}$ and λ_{su} describe the thermal behaviour in the emitter contact, in the emittercap–layer, in the emitter, in the base, in the base contacts, in the collector, in the subcollector, in the subcollector contacts and in the substrate, respectively. Further we assumed that the dissipated power density η_J is evenly distributed across the intrinsic collector depletion region V_{ci} and obtained

$$\eta_J = \frac{P_J}{V_{ci}}\,, \quad P_J := U_{bc}\,I_c\,. \tag{14}$$

The boundary conditions are a very important part of any thermal model. The thermal resistance at the interface between the individual layers was assumed negligible, and hence the temperature is constant across the interface: $T_i = T_{i+1}$. In contrast to previous work, we allowed the metallizations to conduct

heat away from the device, according to Newton's law of cooling, which is an equation linking the normal derivative (n) of temperature and the temperature at the surface of the metallization ($surm$)

$$-\lambda_{Au}\left(\frac{\partial T}{\partial n}\right)_{surm} = h\,(T_{surm} - T_{amb})\,. \tag{15}$$

The heat transfer coefficient h depends crucially on the device geometry and the environment it is embedded in (e.g. distance to bond pad, bond wires) and is a subject of our present theoretical and experimental investigations.
The bottom of the substrate was assumed to be on a constant ambient temperature ($T_S = T_{amb} = 293K$) – a fair approximation considering that the thermal resistances within the device here are much larger than the case–ambient thermal resistances of commercially available packages.
The free surfaces ($surf$) of the device are assumed to be thermally isolated, as heat dissipation by radiation can be neglected according to our calculations. Here, the gradient of the temperature field disappears:

$$\left(\frac{\partial T}{\partial n}\right)_{surf} = 0\,. \tag{16}$$

The problem as described above was solved using the finite-element program **ANSYS**. The mesh had to be chosen very dense ($0.1\,\mu m$) in the vicinity of the heat source, to assure convergence in the presence of large temperature gradients there. Fig. 2 shows an example for a temperature distribution and a mesh density in the vicinity of the heat source.

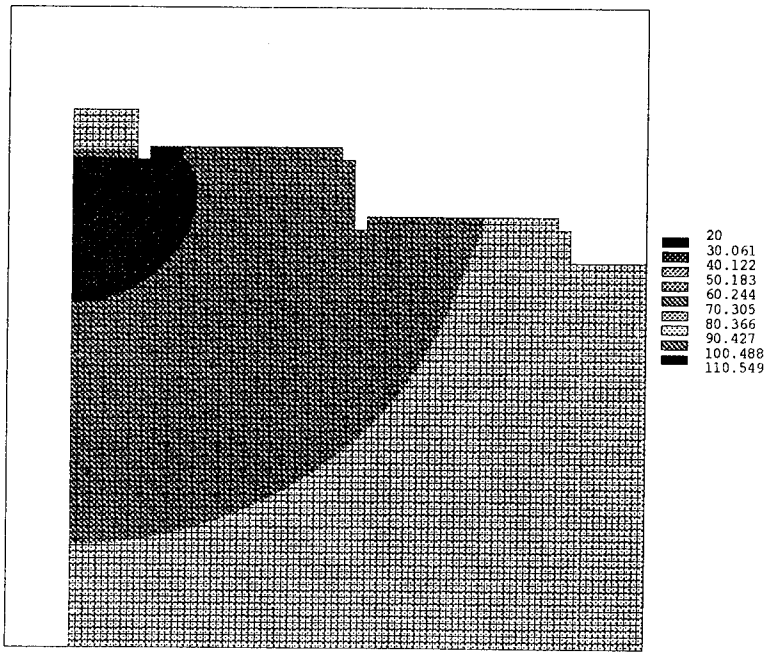


Figure 2: Temperature distribution in °C inside a nonplanar $Al_{0.3}Ga_{0.7}As/GaAs$–HBT structure for a power dissipation $P = 2,5\,W/mm$ and for heat transfer coefficients $h_e = 10^{-5}\,W/(K\,\mu m^2)$ at the emitter, and $h_{bc} = 10^{-7}\,W/(K\,\mu m^2)$ at the base and collector contacts.

To preserve memory and CPU time, the mesh density was gradually decreased with increasing distance from the heat source.

The finite-element method adopted here allows us to model more realistically the effects of the metal contacts and the stochiometry-dependent thermal properties of the semiconductor layers, particularly the $Al_{0.3}Ga_{0.7}As$ emitter layer. Note that the inclusion of the metal contacts is important even if no heat transfer is allowed to the outside ($h_e = h_{bc} = 0$), because they redistribute heat within the device, as can be seen in Fig. 2.

While the one-emitter HBT is well suited to describe the method employed in our calculations, the numerical results presented below relate to a more realistic multi–element power HBT formed by placing many one-emitter cells side by side in close proximity on one substrate, with pitch $b = 2b_1$. We assume that the cells in the middle of the multi–element power HBT dissipate the same power. No electro-thermal feedback mechanism is included in our model at this point. Then, for symmetry reasons, the elements are separated by adiabatic planes (a) where

$$\left(\frac{\partial T}{\partial n}\right)_a = 0 \tag{17}$$

and hence no heat is passed laterally between adjacent cells.

III. NUMERICAL RESULTS

For the multielement HBT outlined above, we consider the effect of various substrate materials, assuming that the HBT subcollectors sit directly on the substrate of choice. Such a structure can be fabricated using epitaxial lift-off [10]. For the emitter length l_e we choose 20 μm in our calculations. Not surprisingly, diamond fares best among the substrates investigated, see Fig. 3.

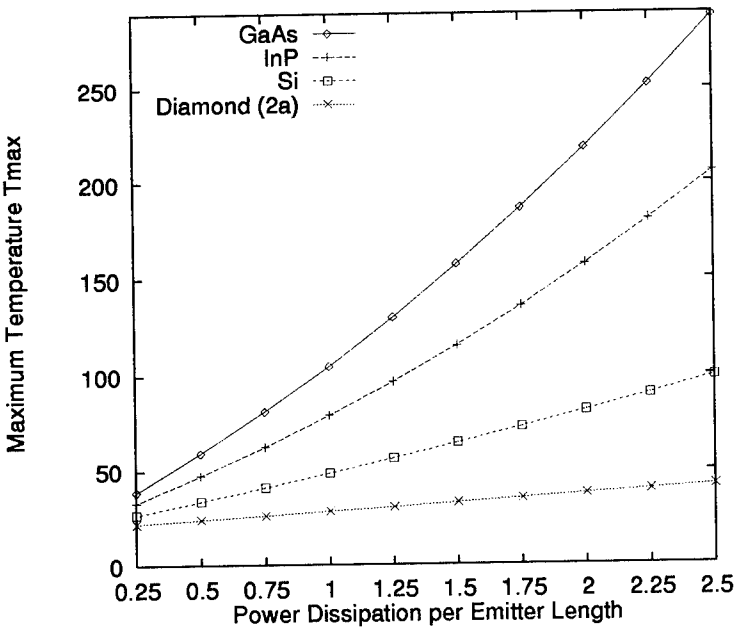


Figure 3: Effect of various substrate materials on the junction temperature T_{max} in ^oC, vs. dissipated power in W/mm.

The significant decrease in junction temperature obtained when choosing InP or Si over $GaAs$ should be noted. The decrease of thermal conductivity with increasing temperature creates the nonlinear increase in junction temperature with increasing dissipated power for $GaAs$ and InP, whereas for Si and diamond, a

linear relationship closely approximates the data. These calculations neglect the effect of heat transfer via surface metal contacts.
To evaluate the effect of substrate thickness on the thermal resistance

$$R_{th} := \frac{T_{max} - T_{amb}}{P_J} , \tag{18}$$

we have to distinguish between single- and multielement devices. In a single-element device with one heat source of small geometry on substrates like $GaAs$ with low thermal conductivity, lateral heat spreading prevails until the diffusive thermal flow reaches the chip boundaries. Substrate thickness, then, has a reduced effect on the junction temperature. In the more relevant case of many closely spaced transistor elements, this situation changes, as outlined above. The cross-section available for heat conduction towards the substrate becomes a linear function of the element spacing b.
As Fig. 4 shows, the effect of substrate thickness becomes the more pronounced, the smaller the element spacing is. The lower curve, for an element spacing of $b = 236\,\mu m$ (=single HBT), represents a case with little thermal interaction, close to the single-element case. The top curve, representing a spacing of $36\,\mu m$, shows a much stronger dependence.

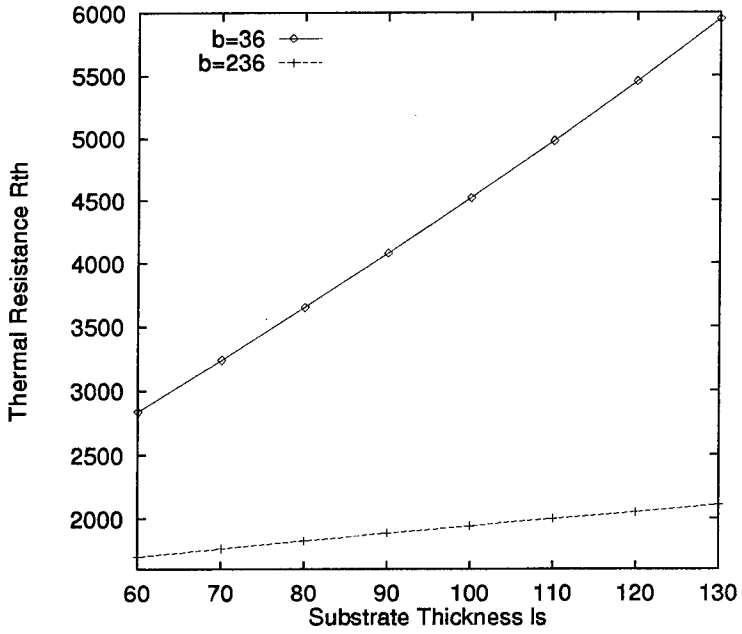


Figure 4: Effect of substrate thickness l_s in μm on the thermal resistance R_{th} in K/W, for two different element spacings in μm, and a dissipated power of $1,5\,W/mm$.

In both cases, the dependence of the thermal resistance on the substrate thickness l_s is *approximately* linear

$$R_{th} \approx R_o + \frac{1}{\bar{\lambda}_2} \cdot \frac{l_s}{b} , \quad \bar{\lambda}_2 := \bar{\lambda} \cdot l_e \tag{19}$$

$$R_0 \quad : \quad \text{thermal resistance of collector and subcollector} ,$$

which is typical for a homogeneous heat flow through the substrate and so the derivative of R_{th} is directly related to the inverse of the element spacing b

$$\frac{dR_{th}}{dl_s} \propto \frac{1}{b} . \tag{20}$$

Here, $\bar{\lambda}$ is the mean substrate thermal conductivity, averaged over the temperature range in the substrate.

So far, we have assumed that no heat is being transfered away from the device via the metal contacts on top, in accordance with the literature. Let us, next, try to assess the validity of this assumption. Lacking experimental data, heat transfer coefficients have been calculated for the emitter, base, and collector contacts. In each case, the metal thermal conductivity was assumed to be constant (a very good approximation) and the heat flow within the metal to be homogeneous. For the base and collector contacts, we used a thickness of $0.5\,\mu m$, leading to

$$h_{bc} = 10^{-7} \frac{W}{K\,\mu m^2} \cdot$$

For the emitter contact, however, a very thick metal layer of $10\,\mu m$ was assumed, in order to study the heat-sinking effect. While this number is about a factor of three larger than a typical airbridge, our results indicate nevertheless that heat-sinking via emitter airbridge contacts cannot be neglected. We obtain for the emitter contact heat transfer coefficient

$$h_e = 10^{-5} \frac{W}{K\,\mu m^2} \cdot$$

Even larger coefficients should be possible by flipchip mounting of the device, as has been suggested.
As Fig. 5 clearly shows, the heat transfer via the contacts will have a very significant effect on the maximum junction temperature, owing largely to the heat-sinking effect of the emitter contact.

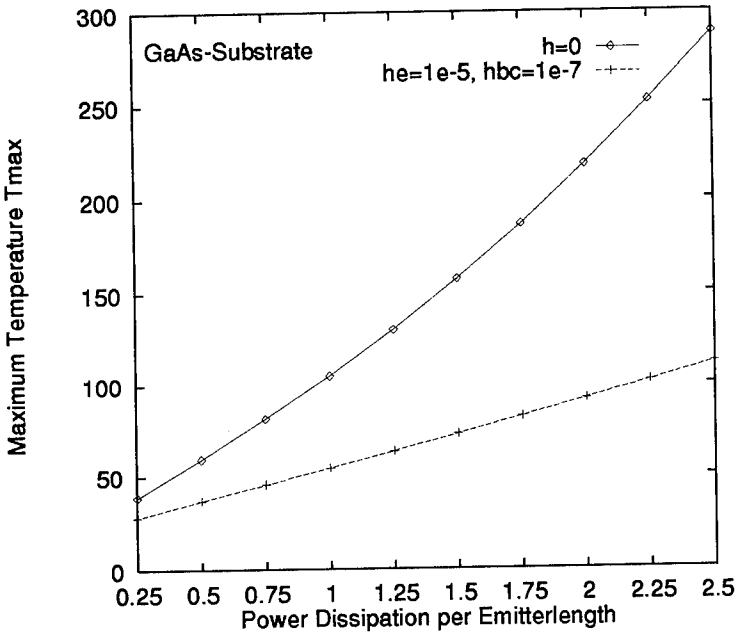


Figure 5: Effect of heat transfer via emitter, base, and collector contacts on the maximum junction temperature T_{max} in $^{\circ}C$ for various magnitudes of power dissipation in W/mm.

The influence of heat transfer via the contacts can also be dramatic when we consider the effect of element spacing on the maximum junction temperature for a given dissipated power. From both a chip real-estate and a microwave design point of view, close spacing is very desirable. If we neglect heat transfer over the contacts, we find that the maximum temperature T_{max} is inversely proportional to the spacing parameter b as Fig. 6 shows.
This dependence can again be explained by a nearly homogeneous heat flow through the substrate. With the definition of the thermal resistance (18) and with equation (19) one gets immediatly for T_{max}

$$T_{max} \approx T_{amb} + P_J\,R_o + \left(\frac{P_J\,l_s}{\lambda_2}\right) \frac{1}{b} \cdot \tag{21}$$

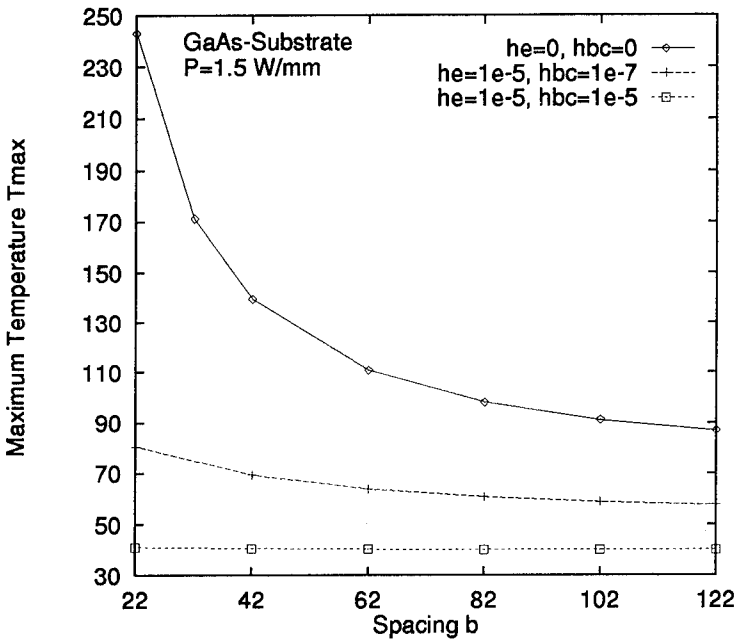


Figure 6: Effect of emitter spacing b in μm on the maximum junction temperature T_{max} in $^{\circ}C$ for various magnitudes of the heat transfer coefficient h in $W/(K\,\mu m^2)$.

Using our estimated heat transfer coefficients as outlined above, the dependence becomes much weaker due to a significiant heat conduction through the contact metallizations. If we allow $h = 10^{-5}\,W/(K\,\mu m^2)$ for all contacts (see Fig. 6, lowest curve), the temperature becomes independent of spacing. This seemingly surprising fact is due to the collector contacts being placed between individual cells in our assumed structure, thermally isolating the individual transistor cells if the heat transfer coefficient of the collector contact is large enough. Such a structure, as desirable as it might be, might be difficult to fabricate, however.
Finally, let us consider an inverted HBT structure where the collector is on top and the $Al_{0.3}Ga_{0.7}As$ emitter is buried beneath the transistor structure. The main advantage of such a structure is the absence of any extrinsic base-collector capacitance, leading to a much improved maximum frequency of oscillation over a more conventional emitter-up device with comparable geometric design rules and base sheet resistance. In the context of power HBTs, however, this structure has the disadvantage that the heat flow from the base-collector region to the substrate now has to pass through the poorly heat-conducting $Al_{0.3}Ga_{0.7}As$ layer. The top curve in Fig. 7a shows that, in the absence of heat-sinking in the collector contact, we observe a significant increase in junction temperature.
If we now allow substantial heat flow through the collector contact ($h_c = 10^{-5}W/(K\,\mu m^2)$), Fig. 7b, the junction temperature of the collector-up structure is reduced to that of the emitter-up device. Effective collector contact heat-sinking is therefore essential for collector-up power HBTs.

IV. CONCLUSION

We have demonstrated finite-element modelling as a means to investigate the thermal behaviour of strongly nonplanar, inhomogeneous HBT structures. In multielement devices with close element spacing, lateral heat spreading is prohibited and the thermal transport in the substrate is near–homogeneous. The dependence of their thermal resistance on element spacing and substrate thickness is reduced to a simple, analytic formula. The heat transfer coefficient of the metal contacts on top of the structure has been calculated

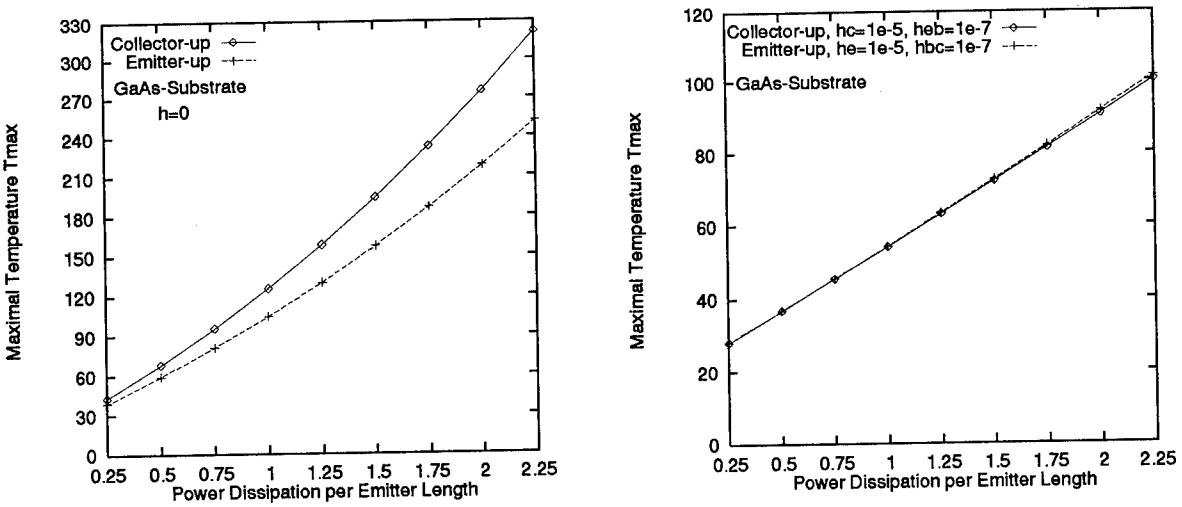


Figure 7: Maximum temperature T_{max} in $°C$ for a collector-up and emitter-up HBT structures vs. dissipated power in W/mm, for (a) no heat-sinking via contacts and (b) heat-sinking via contacts allowed.

using realistic assumptions. It was found that increasing the heat transfer coefficient of the emitter contact to $10^{-5}W/(K\,\mu m^2)$, e.g. by thick airbridges or flip-chip mounting on a heatsink, can markedly improve the thermal behaviour of power HBTs.
Future work in this area will have to be directed towards experimental verification of the contact heat transfer coefficients, as well as establishing an electro-thermal feedback mechanism into the framework of our simulations.

ACKNOWLEDGMENTS

This work has been supported by a grant from the German Research Council. The authors gratefully acknowledge stimulating discussions with and valuable assistance from Dr. A. Trasser, University of Ulm, Dr. M. Passlack, now with AT&T Bell Labs, and from Drs. Leier and Dämbkes of the Daimler-Benz Research Center, Ulm.

*

References

[1] M.G. Adlerstein, M.P. Zaitlin, G. Flynn, W. Hoke, J. Huang, G. Jackson, P. Lemonias, R. Majarone, and E. Tong, *High power density x-band heterojunction bipolar transistors*, Electron. Lett. **27**, pp. 148–149, 1991

[2] G. Gao, M. Wang, X. Gui, H. Morkoç, *Thermal Design Studies of High–Power Heterojunction Bipolar Transistors*, IEEE Transaction on Electron Devices, **36**, pp. 854–862, 1989

[3] P. Webb, I. Russell, *Thermal Simulation of Transients in Microwave Devices*, IEE Proceedings–G, **138**, pp. 329–334, 1991

[4] X. Gui, P. Webb, G. Gao, *Use of the Three–Dimensional TLM Method in the Thermal Simulation and Design of Semiconductor Devices*, IEEE Transactions on Electron Devices, **39**, pp. 1295–1301, 1992

[5] C. Bhandari, D. Rowe, **Thermal Conduction in Semiconductors**, Wiley Eastern, 1988

[6] M. Afromowitz, *Thermal Conductivity of $Ga_{1-x}Al_xAs$ Allows*, J. Appl. Phys., **44**, pp. 1292–1294, 1973

[7] G. Gao, H. Morkoç, M. Chang, *Heterojunction Bipolar Transistor Design for Power Applications*, IEEE Transactions on Electron Devices, **39**, pp. 1987–1996, 1992

[8] H. Kroemer, *Heterojunction Bipolar Transistors and Integrated Circuits*, Proceedings of the IEEE, **70**, pp. 13–25, 1982

[9] S. Yamahata, Y. Matsuoka, T. Ishibashi, *High-f_{max} Collector-Up AlGaAs/GaAs Heterojunction Bipolar Transistors with a Heavily Carbon-Doped Base Fabricated Using Oxygen-Ion Implantation*, IEEE Electron Device Letters, **14**, pp. 173–175, 1993

[10] E. Yablonovitch, T. Gmitter, J. Harbison, R. Bhat, *Extreme selectivity in the lift-off of epitaxial GaAs films*, Appl. Phys. Lett., **51**, pp. 2222–2224, 1987

A NEW RESONANT TUNNELING TRANSISTOR FABRICATED BY CLEAVED EDGE OVERGROWTH

Ç. Kurdak, D. C. Tsui, S. Parihar, H. Manoharan, S. A. Lyon, and M. Shayegan

Department of Electrical Engineering, Princeton University
Princeton NJ 08540

ABSTRACT

A new three-terminal resonant tunneling transistor is realized by the molecular beam epitaxial cleaved edge overgrowth technique in GaAs/AlGaAs system. In addition to negative differential resistance, this device exhibits both positive and negative transconductance behavior which can find useful applications in high speed logic circuits.

I. INTRODUCTION

Three-terminal resonant tunneling devices can provide many useful applications in future electronics. Over the years, this motivated many researchers; different resonant tunneling transistors were proposed[1, 2, 3] and some such new devices have been experimentally realized[4, 5, 6]. Most of the work on resonant tunneling has been on bulk devices where three-dimensional electrons tunnel through the subbands of a quantum well. Luryi *et al.* first proposed a novel three-terminal resonant tunneling transistor in which two-dimensional electrons tunnel through the subbands of a quantum wire[1]. In this structure, unlike bulk resonant tunneling devices, tunneling current can be controlled by a gate.

We have realized this three-terminal resonant tunneling device, which we dubbed the surface resonant tunneling transistor (SRTT), by using the molecular beam epitaxial (MBE) cleaved edge overgrowth technique. The device exhibits negative and positive transconductance as well as negative differential resistance (NDR). SRTT can find useful applications in future ultra large scale integration (ULSI) and can be a potential alternative to current CMOS technology.

II. GROWTH AND FABRICATION

Fabrication of SRTT requires two crystal growth sequences. First, a double barrier structure is grown by standard MBE on a semi-insulating (001) GaAs substrate with the following parameters: 8000 Å n^+-GaAs contacting layer (Si-doped to $4 \times 10^{17} cm^{-3}$), 1.5 μm

0-7803-0894-8/93 $3.00

undoped GaAs spacer layer, 70 Å $Al_{0.4}Ga_{0.6}As$ barrier, 90 Å GaAs well, 75 Å $Al_{0.4}Ga_{0.6}As$ barrier, 1.5 μm undoped GaAs spacer layer, and 4000 Å n^+-GaAs top contact layer (Si-doped to $4 \times 10^{17} cm^{-3}$). Then, 4 $mm \times 8$ mm size pieces from this wafer are loaded back to the MBE chamber for the second growth. Substrates are cleaved in the MBE chamber a few seconds before the growth. The second growth on the (110) surface consists of 20 Å AlAs and 250 Å $Al_{0.3}Ga_{0.7}As$ undoped spacer layers, a 300 Å $Al_{0.3}Ga_{0.7}As$ layer containing ten equally spaced Si delta doping layers with a doping concentration of $1 \times 10^{12} cm^{-2}$, and a 60 Å GaAs cap layer. We grew at a lower temperature, $T = 480\ C$, and at high arsenic beam equivalent pressure, $P = 2 \times 10^{-5}\ Torr$, which has been shown to give a better quality of growth on (110) GaAs surface[7].

The sample structure of SRTT is schematically shown in Fig. 1. Ohmic contacts to the top and bottom contacting layers, to be called source and drain in the rest of the paper, are done by alloying AuGeNi and InSn metals, respectively. Subsequently, 100 Å Ti is evaporated on the overgrown surface to serve as a gate. The channel width of the device is 300 μm. We also fabricated a two-terminal control device with no gate for comparison.

III. RESULTS AND DISCUSSIONS

Fig. 2 shows the I-V characteristics of the two-terminal control device. The device exhibits nonlinear I-V including multiple NDR regions. Sizes and positions of the features are consistent with the picture of resonant tunneling of two-dimensional electrons through the subbands of a quantum wire.

I-V characteristics of SRTT measured at $T = 4.2\ K$ are shown in Fig. 3 (a) for common-source configuration. In addition to NDR the device exhibit negative transconductance. In the saturation regime, drain current first increases, then decreases, and then increases again with increasing gate-source voltage, V_{GS}. This negative transconductance behavior is more clearly seen in the data of Fig. 3 (b). As indicated by the arrows, the transconductance actually exhibits more than one dip as a function of V_{GS}.

This negative transconductance behavior can be explained in two ways. Firstly, in the saturation regime, voltage drop on the double barrier depends on the voltage at the pinch-off point which is only a function of gate voltage. In this way, by just changing the gate voltage, we can achieve resonance condition and resonances from the different subbands of the quantum wire would lead to the features observed in the transconductance data of Fig. 2 (b). Secondly, gate voltage causes a fringing electric field between two-dimensional electron layers and the quantum wire which can also lead to negative transconductance behavior[1]. For this particular device, in the saturation regime, the negative transconductance behavior is mainly due to the former effect.

In conclusion, we have fabricated a new resonant tunneling transistor exhibiting negative and positive transconductance behavior.

ACKNOWLEDGEMENTS

We would like to thank Y. Katayama for many useful discussions, A. Zaslavsky for designing the MBE wafer, and L. N. Pfeiffer and K. W. Baldwin for sharing their experience with the molecular beam epitaxial cleaved edge overgrowth technique. This work is supported by the Office of Naval Research.

References

[1] S. Luryi and F. Capasso, "Resonant tunneling of two-dimensional electrons through a quantum wire: A negative transconductance device," *Appl. Phys. Lett.* vol. 47, pp. 1347-1349, 1985; see also the Erratum, *ibid.* vol. 48, pp. 1693, 1986.

[2] S. Luryi, "Coherent versus incoherent resonant tunneling and implications for fast devices," *Superlattices and Microstructures*, vol. 5, pp. 375-382, 1989.

[3] F. Capasso and R. A. Kiehl, "Resonant tunneling transistor with quantum well base and high energy injection," *J. Appl. Phys.* vol. 58, pp. 1366-1368 ,1985.

[4] M. A. Reed, W. R. Frensley, R. J. Matyi, J. N. Randall, and A. C. Seabangh, "Realization of a three-terminal resonant tunneling device: The bipolar resonant tunnelling transistor," *Appl. Phys. Lett.* vol. 54, pp. 1034-1036, 1989.

[5] K. F. Longenbach, Y. Wang, and W. I. Wang, "Two-dimensional electron gas modulated resonant tunnelling transistor," *Appl. Phys. Lett.* vol. 59, pp. 967-969, 1991.

[6] A. Zaslavsky, D. C. Tsui, M. Santos, and M. Shayegan, "Resonant tunneling of two-dimensional electrons into one-dimensional subbands of a quantum wire," *Appl. Phys. Lett.* vol. 58, pp. 1440,1442, 1991.

[7] L. Pfeiffer, H. L. Stormer, K. West, and K. W. Baldwin, "Quantum wire structures by MBE overgrowth on a cleaved edge," *J. Crys. Gro.* vol. 111, pp. 333-338, 1991 and references therein.

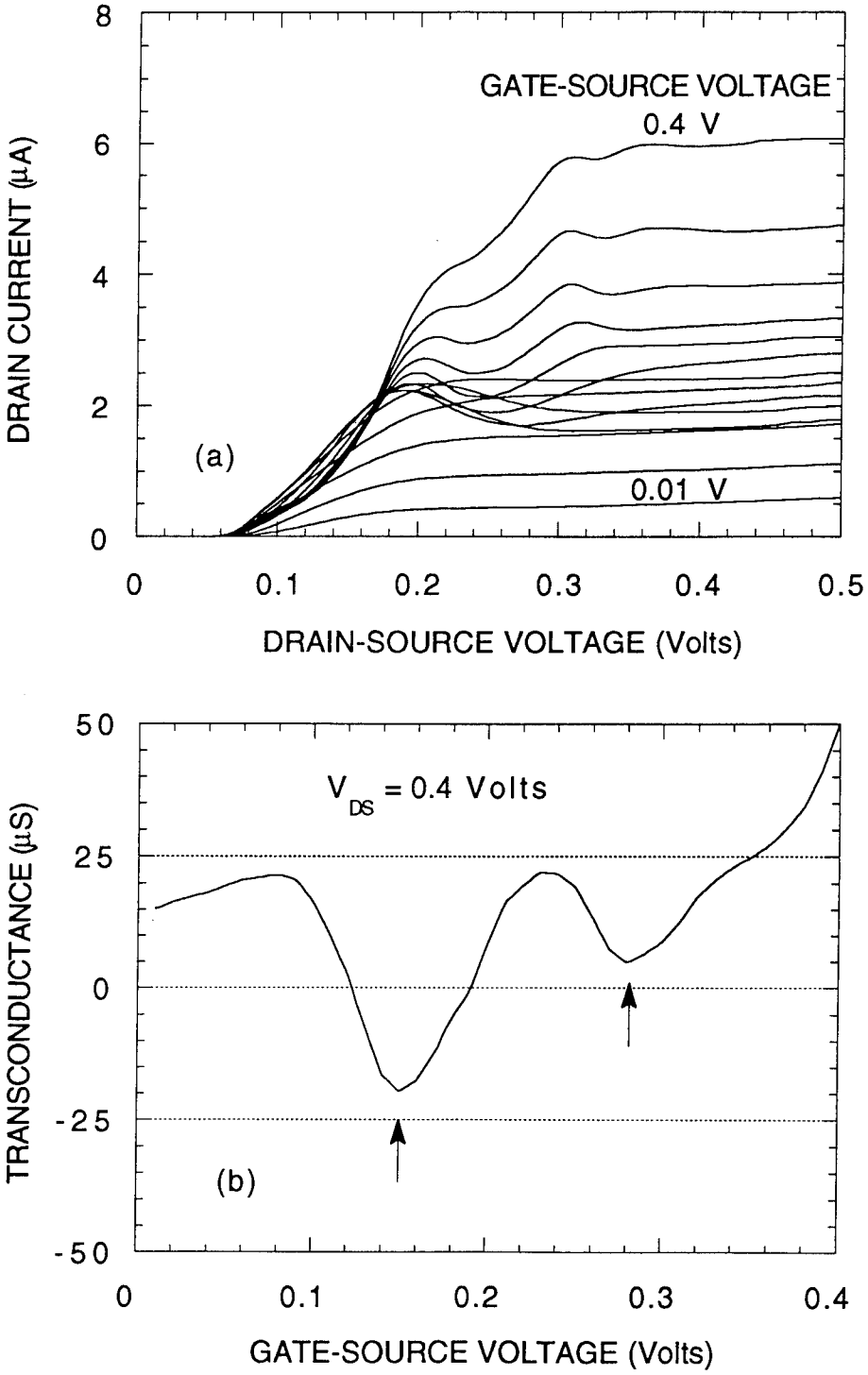


Figure 3: (a) I_D-V_{DS} characteristics for common-source configuration at $T = 4.2\ K$. The curves are given for V_{GS} changing from 0.01 V to 0.4 V with even increments of 0.03 V. (b) Transconductance versus V_{GS} in the saturation regime for $V_{DS} = 0.4\ V$.

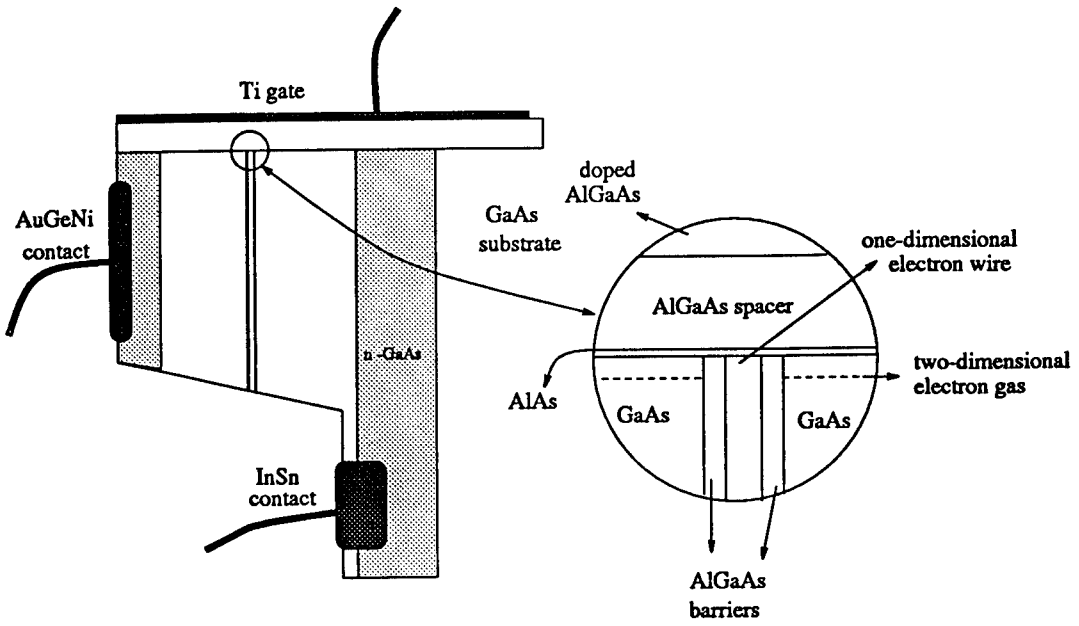


Figure 1: Schematic cross-sectional diagram of SRTT (not to scale).

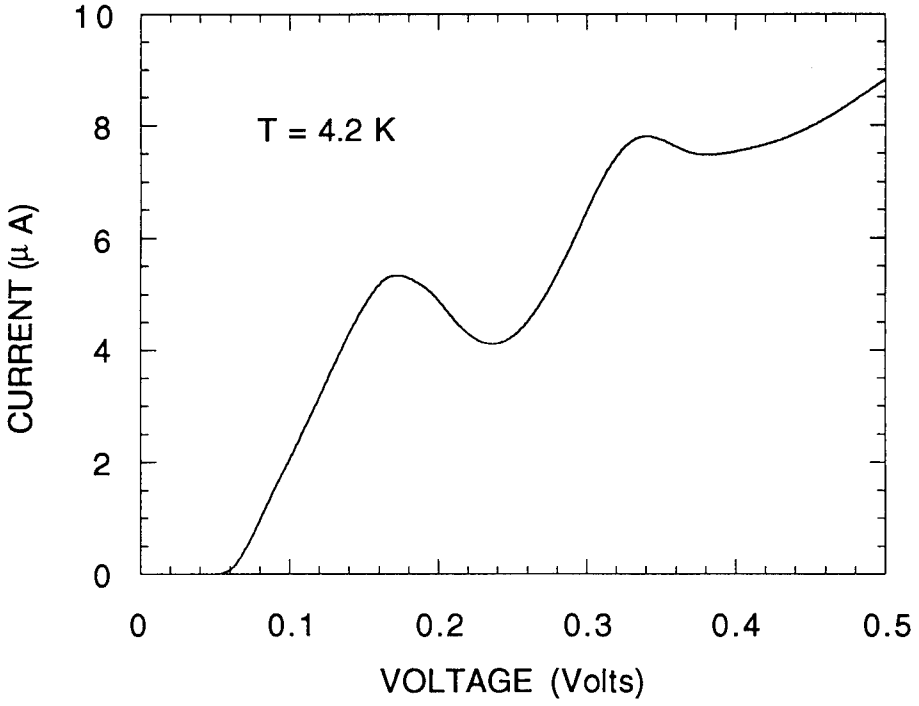


Figure 2: Current versus voltage for the two-terminal control device at $T = 4.2\ K$.

Effect of High Field Electron Transport Characteristics on Transistor Performance in InGaAs/InAlAs Heterostructures‡

John Zahurak†*, Agis A. Iliadis*, Stephen A. Rishton†, and W. Ted Masselink†

‡ Work Performed at the IBM T.J. Watson Research Center, Yorktown Heights, NY 10598

* Department of Electrical Engineering, University of Maryland, College Park, MD

† IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598

Abstract: — We have studied electron transport in a variety of doped and modulation-doped InGaAs/AlInAs quantum wells within the context of FET performance in devices fabricated from the same structures. Electron velocities in the structure were characterized using Hall, geometric magnetoresistance, and high-frequency velocity-field measurements. Peak velocities of 1.5×10^7 cm/s were measured for various delta-doped wells, while a peak velocity of 2.0×10^7 cm/s was measured for the MODFET structure. The effect of various well widths on the transport characteristics was also studied. FETs with 1.8μm gate lengths and uniformly doped channels had transconductances of 267 mS/mm while the MODFET had g_m = 338 mS/mm. MODFETs with 0.5 μm gates yielded g_m = 590 mS/mm. In general, device performance was well correlated to the peak electron velocity and only somewhat correlated to the low-field mobility.

Introduction

The InGaAs/AlInAs quantum well system is of considerable interest due to its suitability for high speed transistor applications and for long wavelength opto-electronic applications. The advantages of this material system over other material systems such as the GaAs/AlGaAs system include higher electron mobilities, higher peak electron velocities, larger gamma-L valley separation, and a larger conduction band discontinuity at the heterojunction. Currently, pseudomorphic InGaAs/AlInAs Modulation Doped Field Effect Transistors (MODFETs) grown on InP hold most FET performance records [1]. However, the MODFET structure has several drawbacks. These include deep traps in the AlInAs donor layer [2] and an upper limit on the number of carriers that can be induced into the channel for transport.

These problems can be avoided by introducing the dopants directly into the quantum well. Doped-channel FETs are viewed as promising devices and have recently attracted considerable attention [3-6]. Heavily doping the channel, of course, degrades the low-field mobility of the structure. As device sizes continue to shrink into the deep sub-micron regime, however, the importance of low-field transport is minimized due to more of the channel being a region of very high electric fields [7].

0-7803-0894-8/93 $3.00 © 1993 IEEE

In this paper we study electron transport in bulk InGaAs, in doped InGaAs/AlInAs quantum wells, and in InGaAs/AlInAs modulation-doped heterostructures. The effects of various doping profiles and well widths on electron transport and transistor performance are investigated.

All structures are grown lattice matched on InP:Fe substrates using gas-source MBE. Each quantum well structure consists of a 400-Å back $Al_{0.48}In_{0.52}As$ barrier layer, an n-doped InGaAs quantum well, a 250-Å AlInAs gate layer, and a 50-Å n+ $In_{0.52}Ga_{0.47}As$ layer to facilitate ohmic contact formation (see Figure 1). To study the effect of the doping profile on electron transport and transistor performance, structures with 300-Å quantum wells were grown with both uniform and delta-doping profiles. The position of the delta-doping is located either at the center of the well (center delta-doped), 50-Å from the top of the well (top delta-doped), or 50-Å from the bottom of the well (back delta-doped). To investigate the effect of well width on performance, center delta-doped quantum wells were grown with well widths of 150-Å, 75-Å, and 40-Å. Finally, a delta-doped MODFET structure was grown by

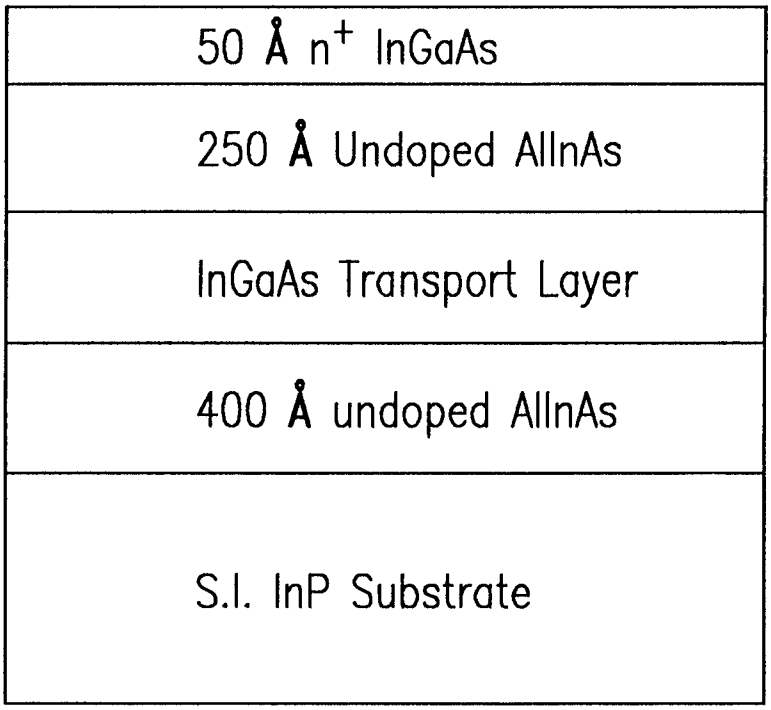


Figure 1. : Schematic of the structures used for velocity-field measurements and transistors. The InGaAs quantum well width and doping profile were varied.

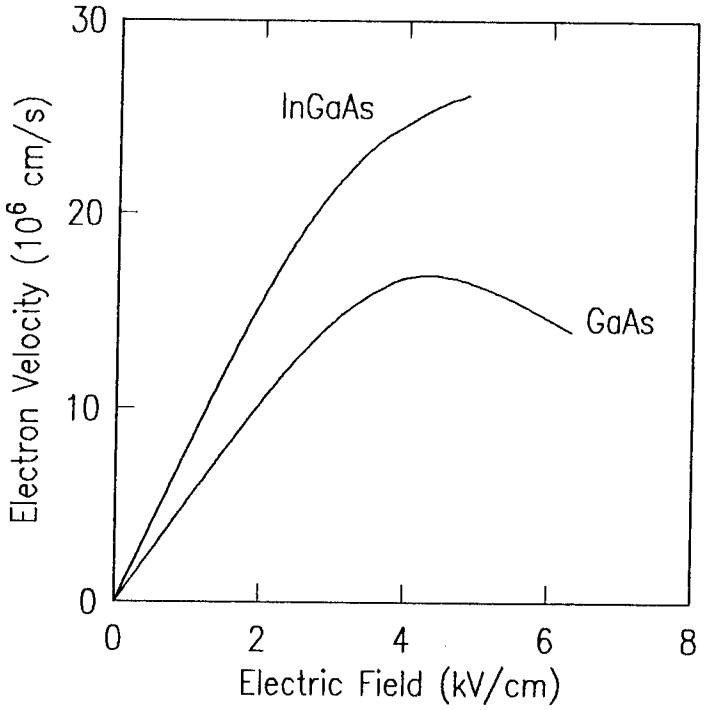


Figure 2. : Electron velocities measured in bulk n-$In_{0.53}Ga_{0.47}As$ and bulk n-GaAs. The InGaAs sample clearly demonstrates superior low-field mobility and peak electron velocity. Both samples are doped 3×10^{16} cm^{-3}.

moving the delta-doping from the quantum well into the top AlInAs layer. In this case, the thickness of the spacer layer between the delta-doping and the top of the 300-Å well is 50-Å.

Transport Results

Electron transport in these structures was characterized using low-field mobility measurements, microwave measurements, and when appropriate, geometric magnetoresistance (GMR) measurements [8]. Low-field mobilities were measured as a function of temperature using the van der Pauw technique. At higher electric fields, domain formation occurs and electron velocities can not be accurately measured using dc techniques. Therefore, a microwave measurement technique was used to characterize high-field transport. This technique utilizes a large 35 GHz sinusoidal field superimposed on a small dc field as described in reference [9]. Previously this technique has been used to measure electron velocities in bulk GaAs, in modulation doped GaAs/AlGaAs heterostructures, and in lightly doped GaAs/AlGaAs quantum wells [10]. Figure 2 shows a comparison between measured

electron velocities in bulk n-InGaAs and bulk n-GaAs control samples. The samples were doped 3.0×10^{16} cm^{-3}. The InGaAs sample had a low field mobility of 8097 cm^2/Vs and a peak electron velocity of 2.6×10^7 cm/s, while the GaAs sample had a low field mobility of 5260 cm^2/Vs and a peak electron velocity of 1.7×10^7 cm/s. Clearly, the InGaAs demonstrates both superior mobility and peak electron velocity.

To evaluate the effect of various doping profiles on electron transport, four 300-Å quantum well structures with 2-DEG concentrations of 2×10^{12} cm^{-2} were compared: uniformly doped, top delta-doped, center delta-doped, and back delta-doped. Due to the increased overlap of their electron wavefunction with the ionized impurities, all of the delta-doped samples had lower mobilities than the uniformly doped sample: $\simeq$3500 cm^2/Vs as compared to 4059 cm^2/Vs (at 300 K). Of the delta-doped structures, the back-doped sample had the lowest mobility of 3050 cm^2/Vs. This is attributed to the additional scattering from the inferior quality of the inverted interface. We also note that due to the degenerate doping levels, none of the samples demonstrated carrier freeze-out at low temperatures. Most samples had mobility maxima at temperatures around 77 K.

The center and the back delta-doped samples had a peak electron velocity of $\simeq 1.5 \times 10^7$ cm/s. The back delta-doped and the uniformly doped samples had slightly lower peak electron velocities of $\simeq 1.4 \times 10^7$ cm/s (see Figure 3). Previous transport measurements on lightly doped GaAs quantum wells had indicated that the peak electron velocity in center delta-doped quantum wells could greatly exceed that observed in uniformly doped quantum wells [10]. This is due to a decrease in the ionized impurity scattering probability as electrons leave the ground state and populate the first excited state of the quantum well. Due to the much higher sheet electron concentration present in actual transistor structures, this effect is largely masked, and we observed only a slight enhancement in peak electron velocity. The peaks in the velocity-field curves of the various structures occurred at various fields in accordance with the samples low-field mobility. Samples with lower low-field mobilities required significantly higher electric fields to achieve the same peak electron velocities as the higher mobility samples (5.8 kV/cm compared to 3.8 kV/cm).

As the well width was decreased from 300-Å to 40-Å, we observed a monotonic decrease in mobility (Mobility (40-Å) = 1232 cm^2/Vs) and a similar decrease in peak electron velocity (0.89×10^7 cm/s).

The MODFET structure, while having excellent low field mobility (9411 cm^2/Vs) reached a peak electron velocity of only 2×10^7 cm/s at a field of 2 kV/cm. This peak value is in good agreement with that measured by Hong et. al. [11] using a pulsed current- voltage technique. While this velocity is higher than that of the doped quantum wells, it is considerably less than the value expected from undoped InGaAs. We measure the peak velocity in a moderately doped (3×10^{16} cm^{-3}) bulk InGaAs sample to be 2.6×10^7 cm/s. This result indicates that the primary factor limiting the velocity of electrons in MODFETs is the real space transfer

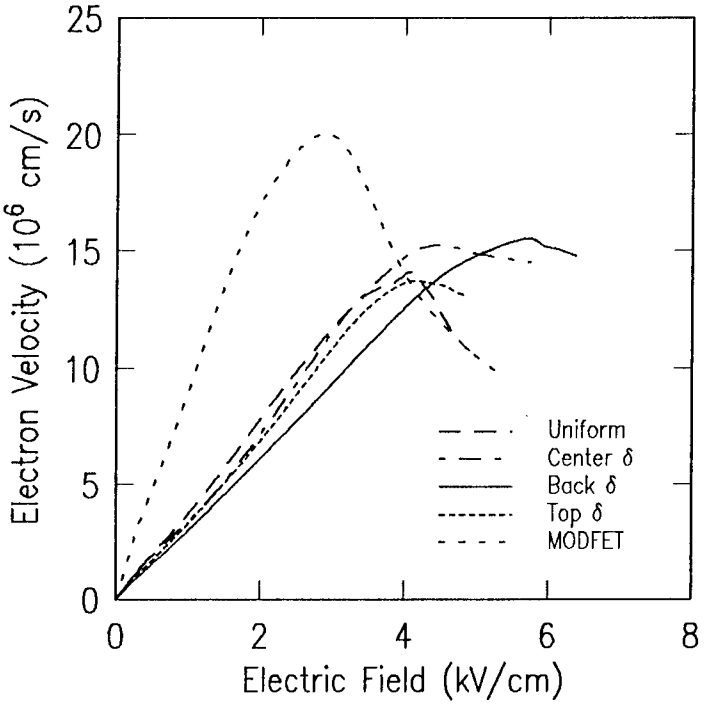


Figure 3. : Electron velocities in 300 Å thick quantum wells. The doping profiles in the well were uniform, top delta- doped, center delta-doped, or back delta-doped. A MODFET structure is also included.

of electrons from the InGaAs gamma valley into the gamma valley of the AlInAs barrier layers. GMR measurements were also performed on the MODFET structure. These results were found to be in good agreement with the low-field microwave data.

Transistor Fabrication and Results

Transistors with a gate length of 1.8 μm were fabricated under similar conditions using standard photolithography. First AuGe/Au metalizations were deposited. After lift-off, the contacts were annealed to become ohmic. This was followed by mesa isolation using a $H_3PO_4 : H_2O_2 : H_2O$ (1 : 1 : 38) etch. Next, the gate recesses were formed using a selective etch of citric acid : H_2O (1 : 1) [12]. Finally, Ti/Au gates were deposited (using electron beam evaporation) and lifted off.

All structures exhibited excellent I-V curves with complete pinch-off and good output conductances . Shown in Figure 4 is a typical family of curves for a 300-Å uniformly doped

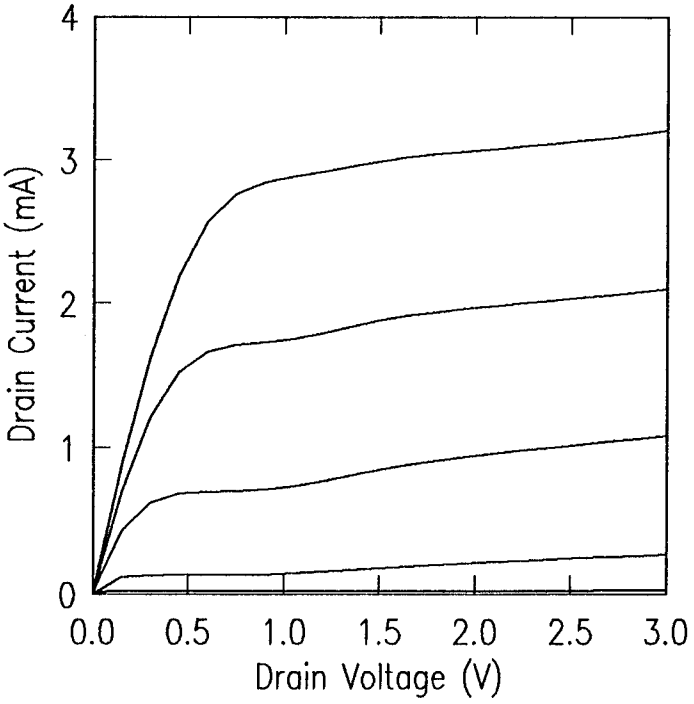


Figure 4. : FET characteristics at 300 K for a device with $L_g = 1.8\mu m$. The active channel is a 300-Å doped InGaAs quantum well lattice matched to InP. Gate voltages range from 0.0 to -1.6 V in -0.2 V increments.

quantum well FET. Table 1 contains the low-field mobility, peak electron velocity, extrinsic transconductance, output conductance, and saturation current (gate unbiased) for all of the 300-Å quantum well structures. As can be seen, the MODFET structure displayed the highest transconductance. However, the magnitude of the increase in the transconductance of the MODFET over the doped channel structures is much less than that expected from a comparison of the low field mobility data. Rather, the increase corresponds much more closely with the increase observed in the peak electron velocity.

Of the doped quantum well structures, the uniformly doped well had slightly higher transconductance than the others but with lower current levels. We attribute the higher transconductance to superior modulation of the carriers by the gate, while the lower current levels are a result of the slightly inferior high field transport properties. It should be noted however, that preliminary results on sub-micron (0.5 μm) devices indicate that as gate lengths are reduced, the delta-doped wells demonstrate both higher current levels and higher

transconductances than the uniformly doped wells (446 mS/mm compared to 315mS/mm) [13].

Doping Profile	Low-Field μ (cm^2/Vs)	$v_{(peak)}$ (cm/s)	g_m^{ext} (mS/mm)	g_o (mS/mm)	$I_{DS\ sat}$ (mA)
Uniform	4059	$1.41x \times 10^7$	267	3.39	3.21
Top δ	3424	$1.37x \times 10^7$	214	5.42	4.2
Center δ	3654	$1.52x \times 10^7$	230	12.2	6.0
Back δ	3051	$1.55x \times 10^7$	200	5.76	6.38
MODFET	9411	$2.00x \times 10^7$	338	6.3	6.0

Table 1. Samples with 300Å InGaAs/AlInAs quantum wells with various doping profiles.

In general, the position of the delta-doping in the quantum well had little effect on transistor performance. The exception to this was the back delta-doped structure. As the distance between the gate and the delta-doping increased, a corresponding loss of control of the gate over the carriers results in lower transconductances. Also, placing the dopants close to the inverted interface degraded the ability of the device to totally pinch-off. It is possible that this leakage current through the back interface could be reduced by replacing the 400-Å InGaAs layer with a low temperature buffer layer [14].

Well Width	Low-Field μ (cm^2/Vs)	$v_{(peak)}$ (cm/s)	g_m^{ext} (mS/mm)	g_o (mS/mm)	$I_{DS\ sat}$ (mA)
300 Å	3654	$1.52x \times 10^7$	230	12.2	6.0
150 Å	2768	$1.29x \times 10^7$	199	5.80	3.1
75 Å	2193	$1.35x \times 10^7$	210	4.1	2.1
40 Å	1232	$0.89x \times 10^7$	126	1.4	0.88

Table 2. Samples with center delta-doped InGaAs quantum wells.

As the quantum well width was reduced from 300-Å to 40-Å, we see a sharp decrease in the transconductance from 230 mS/mm to 126 mS/mm. The drain saturation current (for Vg = 0) also dropped drastically from 6.0 mA to 0.88 mA (see Table 2). Bahl et. al. observed a similar decrease in g_m and $I_{D\ sat}$ in their study of the effect of well width on breakdown voltage [15]. Several factors are responsible for this decrease. First, as the quantum well width is reduced, the scattering probability between the electrons and the ionized impurities increases. Also, scattering due to any roughness at the heterojunction interfaces will become

more important. Second, we observed an apparent decrease in the number of carriers available for transport. Although the dopant concentration remained constant, the measured 2-DEG concentration dropped from 2.5×10^{12} cm^{-2} for the 300-Å well to 1.6×10^{12} cm^{-2} for the 40-Å quantum well. This is explained by the increased separation in energy between the donor levels and the allowed states in the well as the width of the well is reduced. Also, parallel conduction in the AlInAs layers occurs due to the electron wavefunction penetrating into the AlInAs barrier layers. One additional problem associated with the narrower quantum wells is that of contacting them electrically. As the well width decreased from 300-Å to 40-Å, the contact resistance increased from 0.3 ohm-mm to 1.3 ohm-mm. This contact resistance is the resistance to electrically contact the conducting well. After the ohmic contacts were formed, we selectively etched away the top n+ InGaAs layer. The resistance was then measured using the standard transmission line model (TLM).

Conclusion

We have studied the effect of various doping profiles and quantum well widths on electron transport and transistor performance. We find that while the doping profile has a strong impact on low field mobility, the peak electron velocity in each of the doped 300-Å wells is very close to 1.5×10^7 cm/s. The MODFET structure, while demonstrating superior low-field transport, did not reach the high-field peak velocity expected from bulk InGaAs measurements. Rather, its peak value was limited by real space transfer into the AlInAs layers.

For 1.8μm gate length transistors, the MODFET structure offers slightly better performance than the doped quantum well structures. In general, transistor performance followed the trends predicted by the transport data. The exception to this was the better than expected performance of the uniformly doped quantum well structure. Preliminary results on sub-micron devices, however, indicate that the superior performance of this structure disappears as gate lengths enter the sub-micron region. In general, the position of the delta-doping had little effect on transistor performance. Reducing the width of the quantum well strongly degraded both transport and transistor performance.

Acknowledgements—The authors benefited from discussions with Khalid Ismail.

References

[1] L.D. Nguyen, A.S. Brown, M.A. Thompson and L.M. Jelloian, "50-nm Self-Alligned-Gate Pseudomorphic AlInAs/InGaAs High Electron Mobility Transistors," *IEEE Trans. on Electron Devices*, vol. 39, pp. 2007-2014, 1992.

[2] W-P. Hong, S. Dhar, P.K. Bhattacharaya and A. Chin, "Deep Levels and a Possible D-X-Like Center in Molecular Beam Epitaxial InAlAs," *J. Electronic Materials*, vol. 16, pp. 271-274, 1987.

[3] S.R. Bahl, B.R. Bennett and J.A. del Alamo, "Doubly Strained InAlAs/n"+"-InGaAs HFET with High Breakdown Voltage," *IEEE Electron Device Letters*, vol. 14, pp. 22-24, 1993.

[4] J. Dickmann, "Influence of the Delta Doping Position in the Channel on the Device Performance of AlGaAs/InGaAs Modulation-Doped Field-Effect Transistors," *Appl. Phys. Lett.*, vol. 60, pp. 88-90, 1992.

[5] M. Matloubian, A.S. Brown, L.D. Nguyen, M.A. Melendes, L.E. Larson, M.J. Delaney, J.E. Pence, R.A. Rhodes, M.A. Thompson and J.A. Henige, "High-Power V-Band AlInAs/GaInAs on InP HEMTs," *IEEE Electron Device Letters*, vol. 14, pp. 188-189, 1993.

[6] Y-J. Chan and D. Pavlidis, "Single and Dual p-Doped Channel InAlAs/InGaAs FETs and the Role of Doping," *IEEE Trans. on Electron Devices*, vol. 39, pp. 466-472, 1992.

[7] J. Dickmann, "The Backside Pulse Doped Channel Heterostructure Field-Effect Transistor: Design, DC, and RF Performance," *Jpn. J. Appl. Phys.*, vol. 32, pp. 17-25, 1993.

[8] W.T. Masselink, W. Kopp, T. Henderson and H. Morkoç, "Measurement of the Electron Velocity-Field Characteristic in Modulation Doped Structures Using the Geometrical Magnetoresistance Method," *IEEE Electron Device Lett.*, vol. 6, pp. 539-541, 1985.

[9] W.T. Masselink, "Electron Velocity in GaAs: Bulk and Selectively Doped Heterostructures," *Semicond. Sci. Technol.*, vol. 4, pp. 503-512, 1989.

[10] W.T. Masselink, "High-Differential Mobility of Hot Electrons in Delta-Doped Quantum Wells," *Appl. Phys. Lett.*, vol. 59, pp. 694-696, 1991.

[11] W-P. Hong and P.K. Bhattacharya, "High-Field Transport in InGaAs/InAlAs Modulation Doped Heterostructures," *IEEE Trans. on Electron Devices*, vol. 34, pp. 1491-1495, 1987.

[12] M. Tong, K. Nummila, A. Ketterson, I. Adesida, C. Caneau and R. Bhat, "InAlAs/InGaAs/InP MODFET's with Uniform Threshold Voltage Obtained by Selective Wet Gate Recess," *IEEE Electron Device Lett.*, vol. 13, pp. 525-527, 1992.

[13] J. Zahurak, A. Iliadis, S. Rishton and W.T. Masselink, (To be Published).

[14] A.S. Brown, U.K. Mishra, C.S. Chou, C.E. Hooper, M.A. Melenedes, M. Thompson, L.E. Larson, S.E. Rosenbaum and M.J. Delaney, "AlInAs/GaInAs HEMT's Utilizing Low-Temperature AlInAs Buffers Grown by MBE," *IEEE Electron Device Lett.*, vol. 10, pp. 565-567, 1989.

[15] S.R. Bahl and J.A. del Alamo, "Breakdown Voltage Enhancement from Channel Quantization in InAlAs/n"+"-InGaAs HFETs," *IEEE Electron Device Letters*, vol. 13, pp. 123-125, 1992.

Planar Cold Cathodes

Umesh K. Mishra and Wei-Nan. Jiang
Department of Electrical and Computer Engineering, University of California
Santa Barbara, CA 93106

INTRODUCTION

There is a great need for electron sources for a variety of applications, such as klystrons, Traveling Wave Tubes (TWTs), planar panel displays, vacuum triodes, etc.. The cold cathode electron emitters under investigation currently can be basically divided into two types: field emitters and planar emitters. Field emitters are based on the field emission mechanism which was first presented by Fowler and Nordheim in 1928[1]. A field emitter can be basically described as a triangular tip which can be made from semiconductors or metals by deposition and micro-fabrication techniques as shown in Fig. 1. When a positive bias is applied to the anode which is placed very close to the tip which is grounded, a very high electric field could be established at the tip so that electrons inside the semiconductors or metals would have considerable probability to tunnel through the vacuum barrier and be collected by the anode. Extensive research has been done to fabricate reliable field emitter arrays with high emission current density, and various types of the field emitters, including the vertical and the lateral types, have been developed so far with great success. Microwave amplifiers and flat panel displays based on field emitter arrays have also been reported recently[2,3,4]. The main challenges which field emitters are facing are uniformity and reliability. The former is due to their sensitivity to fabrication processes, and the latter is due to the ion bombardment of the residual gases under operation which has been reported by many researchers.

In planar emitters, electrons are accelerated inside solid state materials by internal electric field. Those electrons which reach the surface with enough kinetic energy to overcome the surface work function could be emitted into vacuum as shown in Fig. 2. A lot of work has been done in this area and many planar emitters based on different hot electron generation mechanisms have been described in the past 30 years. To obtain considerable emission efficiency and current density, the solid state surface is usually coated with low work function material (on the order of one mono-layer), such as cesium and cesium oxide.

This paper reviews the development of planar cold cathodes and discusses the characteristics of AlGaAs/GaAs planar-doped-barrier electron emitters (PDBEEs) based on the study of their temperature dependent current-voltage (I-V) characteristics.

HISTORICAL PERSPECTIVE

Since 1960, several types of planar cold cathode electron emitters have been demonstrated. Some of them showed very promising results and are still under investigation today.

1. Tunneling electron emitters

Mead[5] reported his experiments on tunneling emission using a metal-insulator-metal structure based on his finding of tunneling of electrons occurring at a metal-insulator interface when a high electric field is present within the insulator[6]. Shortly, several groups reported their emission from tunneling emitters [7,8,9]. However, the results were not very encouraging. The emission current density and efficiency were too low for practical

0-7803-0894-8/93 $3.00

application, and the emission was not stable and reproducible, mainly because the fabrication techniques were inadequate to form high quality ultra-thin metal and insulator layers at that time. Recently, Yokoo *et al.*[10] reported electron emission from Si-SiO_2-metal structures with an efficiency of 0.7%. But, the emission current density is low and the deterioration of the oxide at high fields results in poor stability of the emitters.

2. Negative-electron-affinity (NEA) electron emitters

A report by Scheer and Laar[11] demonstrated that the work function could be below the edge of the conduction band of a cesiated p-type GaAs. Several electron emitters based on this negative electron affinity were then proposed. The use of a Schottky barrier at a metal-semiconductor contact to inject electrons through a thin metal film was first proposed by Atalla[12] and electron emission from such a structure has been reported by using cesiated ZnS/Pt[13] and ZnS/Pd[14].

Negative-electron-affinity (NEA) cold cathodes which consist of a forward biased p-n junction with a cesiated surface were proposed by Geppert[15], and has been demonstrated from GaAs[16,17,18], Si[19,20], GaAsP[21,22], and GaP[23,24]. An efficiency o f 4% and an emission current density of over 1 A/cm^2 have been reported[25,26]. Several designs of vidicon-type tubes have also been demonstrated[27]. However, they usually suffered from operational instability and short life time mainly due to the loss of the activation material from NEA surface. Although some improvements have been made to stabilize the activation material[28], to the best of author's knowledge, devices based on NEA cold cathodes have not reached practical applications.

3. Si and GaAs Avalanche electron emitters

In 1957, Burton[29] reported his observation of electron emission from a Si p-n junction reverse-biased into avalanche breakdown. The electron emission he described was from a cesiated cleavage surface which intersects the p-n junction. Following the discovery of Burton, a lot of work has been done to look into the characteristics of this new phenomenon[30,31,32]. Besides silicon, silicon carbide (SiC) was also investigated[33]. However, all those early works didn't lead to any practical applications because of the low emission efficiency, low current density and poor reproducibility. Not too much progress was made in the 1970's until Van Gorkom and Hoeberechts at Philips Research Laboratories[34] demonstrated their cesiated Si avalanche cold cathodes (SACCs) with the help of modern semiconductor technologies. After continuous research of more than ten years, the state of art silicon avalanche cold cathodes can emit electrons into vacuum with efficiency of 1 to 8 % and current densities of 300 to 3000 A/cm^2 [35]. Small experimental TV tubes using silicon avalanche cold cathodes have also demonstrated[36,37]. Emission from non-cesiated silicon avalanche cold cathodes has been reported with a much lower efficiency on the order of 10^{-5} [38,39]. van Zutphen[40] applied same device structure to GaAs and reported emission with an efficiency of 1.1×10^{-6} (non-cesiated surfaces).

PLANAR-DOPED BARRIER ELECTRON EMITTERS (PDBEEs)

Recently, we reported electron emission from GaAs[41] and AlGaAs PDBEEs[42]. A planar-doped-barrier (PDB) structure basically consists of n^+-undoped (*i*)-n^+ layers with a p-type delta-doped layer inserted within the undoped region giving rise to a triangular barrier structure as shown in Fig. 3. When a positive bias is applied to the surface (the back is grounded), the p-i-n junction on the right of the barrier (acceleration region) is reversed-biased first until the acceptors of p^+ layer are fully ionized (if they are not fully ionized under equilibrium), much like a reach-through transistor[43] After the reach-through point, any further bias is distribute between two junctions so that the p-i-n junction on the

left of the barrier (injection region) is forward-biased and the acceleration region is further reverse-biased. Electrons are then injected over the triangular barrier into the high field acceleration region and accelerated towards the surface. Electrons gain kinetic energy from the field and lose kinetic energy through various scattering processes, such as phonon scattering, impact ionization, electron-electron scattering, etc.. Those electrons reaching the surface with kinetic energy higher than the work function can be emitted into vacuum.

The early applications of PDB diodes were based on their unique characteristics of adjustable barrier heights and asymmetrical rectifying characteristics. For those applications, PDB structures have relative low barrier heights (less than the half of the band gap) and thick undoped regions (on the order of thousand Angstroms). Experiments have shown that the current across those PDB diodes are governed by the thermionic emission over the triangular barrier. In a semi-log plot, I-V curves are almost straight lines. Observed non-linearity of the I-V curves has been attributed to the diffusion of carriers from n^+ layers into the undoped regions[44].

According to thermionic emission theory, the barrier height of a PDB diode can be experimentally determined by temperature dependent I-V measurements. However, to our surprise, on a semi-log plot, the I-V curves of the PDBEE, which was the first reported GaAs based electron emitter[45], are weakly dependent on temperature, and the curves are essentially parallel to each other as shown in Fig. 4. A negative differential resistance (NDR) feature at low bias is also observed at low temperature. This clearly indicates that the carrier transport across this PDBEE is not governed by the thermionic emission but rather by other mechanisms.

TEMPERATURE DEPENDENT I-V CHARACTERISTICS

To understand the electron transport across PDBEEs, which is essential to the understanding of device performance, a systematic study on the temperature dependent I-V characteristics of PDB diodes has been carried out. The diodes under investigation have the same layer and doping structure as listed in Table I. The only difference among them is the doping density of the p^+ layer which is listed in Table II. Different from real PDBEEs, these test PDB diodes have a thicker uniformly doped n^+ layer instead of a very thin n^+ channel as the top layer in order to exclude the possible effects caused by the reflection of electrons at the surface.

Fig. 5 shows the measured temperature dependent I-V characteristics of five PDB diodes listed in Table I and II. Among them, Fig. 5(c) is a good point to start our discussion. It corresponds to a p^+ doping density of $8.4\text{x}10^{12}$ cm^{-2}. The I-V curves shown in this figure can be divided into three regions: (*i*) a low bias region which has weak temperature dependence in both magnitude and slope, (*ii*) a medium bias region which has a strong temperature dependence in both magnitude and slope, and (*iii*) a high injection region which is easily recognized as a resistive region. The cross-over voltage (V_{cross}), at which regions (*i*) and (*ii*) intersect, increases with decreasing of the temperature as indicated by the dashed line in the figure. When Σ increases to $1.0\text{x}10^{13}$ cm^{-2} as shown in Fig. 5(b), region (*ii*) which is clearly shown in Fig. 5(c) can be barely seen and the complete I-V characteristics is weakly dependent on temperature. In addition to the weak dependence on temperature, there is a NDR feature at $V_a \approx 0.5$ volts which becomes more visible at low temperature. As Σ increases further to $2.0\text{x}10^{13}$ cm^{-2}, I-V curves are still weakly dependent on temperature and the NDR feature at low bias becomes more prominent as shown in Fig. 5(a). On the other hand, when Σ decreases, V_{cross} shifts to lower bias and region (ii) becomes more dominant as shown in Figs. 4(d) and (e).

A schematic band diagram of a PDB diode under a bias of V_a is shown in Fig. 6. Several possible current mechanisms are indicated in the figure as arrows. J_{th} is the thermionic emission term over the triangular barrier (contributed by electrons with kinetic energy larger than the barrier), J_{fe} is the current contributed by electrons with energy less than the barrier tunneling through the triangular barrier (thermionic field emission), J_{bb} is the tunneling current of electrons from the valence band to the unoccupied states in the conduction band through the energy gap of the acceleration region, J_r and J_g are the recombination and generation currents in forward bias and reverse bias junctions respectively, and J_h is the hole diffusion current. From simple diode theory[46], it is not difficult to find out that J_h and J_g are less important in the PDB diodes under investigation and will be ignored during later discussions.

As well known, thermionic emission currents depends strongly on temperature with the slope of semi-log I-V curves inversely proportional to temperature. On the other hand, tunneling current across the energy gap depends weakly on temperature both in the magnitude and slope. This suggests that the current transport may be governed by thermionic emission in region (*ii*) and by tunneling in region (*i*). With increasing Σ, the tunneling component (J_{bb}) becomes more dominant and finally the complete I-V characteristic is governed by the tunneling as shown in the Fig. 5(a). On the other hand, with decreasing Σ, the tunneling becomes less important and eventually the carrier transport will be dominated by thermionic emission (as observed in earlier PDB diodes).

To obtain a quantitative understanding of the carrier transport across the PDB diodes, a simulation which takes into account the quantum mechanical transmission probability through the barrier and tunneling through the forbidden energy gap has been carried out. The details of the simulation have been published elsewhere[47]. Fig. 7 shows the result of the simulation for diode C at room temperature (300K). The open squares represent the tunneling current (J_{bb}), the solid squares represent the transmission current (J_{trans}) and the solid line represent the thermionic emission current (J_{th}). It is very clear that the tunneling of electrons with kinetic energy less than the barrier plays an important role in carrier transport even at room temperature. Since J_{th} is almost parallel to J_{trans}, the transmission current can be approximated by the thermionic emission theory with a lower zero bias barrier height. In this way, we can still label the region for $V > V_{cross}$ as the thermionic emission dominated region. Fig. 8 shows the simulation results for diode C at different temperatures, from 300K to 100K. The solid lines are simulation results and open squares are experimental results. The agreement between the theory and experiment is very good.

It is now clear that regions (*i*) and (*ii*) in Fig. 5(c) are tunneling and thermionic emission dominated regions, respectively. The cross-over voltage of these two different currents V_{cross} depends on Σ and temperature. For diode A, resistive region is reached before reaching the cross-over point, and carrier transport is governed by tunneling through the band gap. For diode B, V_{cross} is about 2.2 volts at 300K and the thermionic emission region can be barely seen. In both diodes A and B, a NDR feature has been observed which is more prominent at low temperature. It can be explained in a similar way to the forward characteristics of a Esaki diode[48]

EMISSION CHARACTERISTICS

In a PDB electron emitter, for a given surface work function and transit distance (the distance between the peak of the barrier to the surface), the emission efficiency is basically dependent on the maximum kinetic energy electrons can gain from the field. Because electrons tunneling through the band gap of the acceleration region have much lower kinetic

energy than those injected over the barrier, the emission efficiency will be low if the carrier transport across the device is governed by the tunneling. Therefore, it is essential for high performance emitters that devices should be operated in thermionic dominated region.

A direct way of suppression the tunneling current J_{bb} relative to the thermionic emission current is by using materials with larger band gap than GaAs, since tunneling probability decreases drastically with increasing of the band gap. Fig. 9 shows the temperature dependent I-V characteristics of a $Al_{0.3}Ga_{0.7}As$ PDBEE which has a similar barrier structure to the GaAs PDBEE[49]. The strong temperature dependence of its I-V characteristics indicates that the carrier transport across this $Al_{0.3}Ga_{0.7}As$ PDBEE is governed by the thermionic emission. Fig. 10 shows the emission performances of two PDBEEs. An efficiency of 1 % and an emission current density of 2.64 A/cm^2 have been obtained from the $Al_{0.3}Ga_{0.7}As$ PDBEE, which are about three and six times higher than obtained from the GaAs PDBEE with a similar structure.

As described earlier, in PDBEEs, electrons are injected into a high field region which is biased below the avalanche breakdown, rapid electron thermalization due to the impact ionization is absent. High efficient, high emission current density cold cathodes based on PDB structures is expected by using materials with large bandgap and reasonable good transport characteristics. Electron emission without cesiation is also possible. Possible materials include InGaAs/GaAs, AlGaN, etc..

SUMMARY

A systematic study on the temperature dependent I-V characteristics of PDB diodes with high accelerating field has been carried out. In PDBEEs, the field inside the acceleration region is high so that electrons tunneling through the energy band gap becomes an important mechanism and even governs the carrier transport under high enough accelerating field. Since electrons tunneling through the energy band gap of the acceleration region have much low kinetic energy than those injected over the barrier, it is very important to operate PDBEEs in the thermionic emission region to obtain high performance electron emitters. High efficient, high emission current density cold cathodes based on PDB structures is expected by using materials with large bandgap and reasonable good transport characteristics. Electron emission without cesiation is also possible with the use of large bandgap materials.

REFERENCES

1 R.H. Fowler and L.W. Nordheim, Proc. Roy. Soc., Ser. A, vol. 119, p. 173, (1928)

2 C.A. Spindt, C.E. Holland, A. Rosengreen, and I. Brodie, "Field-emitter-array development for high-frequency operation", J. Vac. Sci. Technol. B **11**(2), 468 (193)

3 I. Brodie, "Advanced technology: flat cold-cathode CRTs", Information Display, No.1, 17 (1989)

4 Betsui, Digist of sixth International Vacuum Microelectronics Conference, July, 12-15, 1993, Newport, RI, p. 137

5 C.A.Mead, "Operation of Tunneling-Emission Devices", J. Appl. Phys., **32**, 646 (1961)

6 C.A.Mead, Proc. Inst. Radio Engrs., **48**, 359, 1478 (1960)

7 J. Cohen, "Tunneling Emission into Vacuum", J. Appl. Phys., **33**, 1999 (1962)

8 H Kenter and W.A. Feibelman, "Electron Emission from Thin Al-Al_2O_3-Au Structures", J. Appl. Phys., **33**, 3580 (1962)

9 M.Silver, D.G. Onn, P. Smejtek, and K. Masuda, Phys. Rev. Lett., **19**, 626 (1967)

[10] K. Yokoo, H. Tanaka, S. Sato, J. Murota, and S. Ono, "Emission characteristics of metal-oxide-semiconductor electron tunneling cathode", J. Vac. Sci. Technol. **B11**(2), 429 (1993)
[11] J.J. Scheer and J. van Laar, "GaAs-Cs: A new type of photoemitter", Solid State Communication, **3**, 189 (1965)
[12] M.M. Atalla, U.S. Patent No. 3,121,809, February 18, 1964
[13] R. Williams and C.R. Wronski, "Electron Emission from The Schottky Barrier Structure ZnS:Pt:Cs", Appl. Phys. Lett., **13**, 231 (1968)
[14] R.K. Swank, "Characteristics of a ZnS:Pd:Cs_2O Cold Cathode", J. Appl. Phys., **41**, 778 (1970)
[15] D.V. Geppert, "A proposed p-n junction cathode:, Proceedings of IEEE, **54**, 61 (1966)
[16] B.F. Williams and R.E. Simon, "Electron emission from a 'cold-cathode' GaAs p-n junction", Appl. Phys. Lett., **14**, 214 (1969)
[17] H. Schade, H. Nelson, and H. Kressel, "Novel GaAs-(AlGa)As cold-cathode structure and factors affecting extended operation", Appl. Phys. Lett., **20**, 385 (1972)
[18] H. Kressel, E.S. Kohn, J.J. Tietjen, and L.R. Weisberg, "An optoelectronic cold cathode using an $Al_{1-x}Ga_xAs$ heterojunction structure", Appl. Phys. Lett., **16**, 359 (1970)
[19] E.S. Kohn, "Cold-cathode electron emission from silicon", Appl. Phys. Lett., **18**, 272 (1971)
[20] E.S. Kohn, IEEE Trans. Electron. Devices, **ED-20**, 321 (1973)
[21] K.R. Faulkner, R.A. Astridge, J.R. Howorth, and R.K. Surridge, Appl. Phys. Lett. 23, 298 (1973)
[22] J.R. Howorth, R.K. Surridge, and I.C. Palmer, Sixth Symposium on Photoelectronic Image Devices, London, 1974, Advances in Electronics and Electron Physics (Academic, New York, 1976), p. 463
[23] E. Stupp, A. Pelissier, M. Kidder, and A. Milch, "GaP negative-electron-affinity cold cathodes: a demonstrated and appraisal", J. Appl. Phys. **48**, 4741 (1977)
[24] H. Kan, T. Nakamura, H. Katsuno, M. Hagino, and T. Sukegawa, "New structure GaP-GaAlP heterojunction cold cathode", IEEE Trans. Electron Devices, **ED-26**, 1759 (1979)
[25] H. Schade, H. Nelson, and H. Kressel, "Novel GaAs-(AlGa)As cold-cathode structure and factors affecting extended operation", Appl. Phys. Lett., **20**, 385 (1972)
[26] H. Fressel, H. Schade, and H. Nelson, "Heterojunction cold-cathode electron emitters of (AlGa)As-GaAs", J. Luminescence, **7**, 146 (1973)
[27] Y. Hatanaka, T. Kawai, T. Sukegawa, T. Ando, and R. Nishida, "A camera tube with NEA cold cathodes", IEEE Trans. Electron. Devices, **ED-28**, 72 (1981)
[28] H. Kan, H. Katsuno, T. Nakamura, M. Hagino, and T. Sukegawa, "New activation methods for long-life and highly stable GaP-GaAlP heterojunction cold cathodes", J. Appl. Phys. **51**(6), 3404 (1980)
[29] J.A. Burton, "Electron emission from avalanche breakdown in silicon", Phys. Rev. **108**, 1342 (1957)
[30] R.J. Hodgkinson, "Hot-electron emission from semiconductors", Solid-State Electron., **5**, 269 (1062)
[31] D.J. Bartelink, J.L. Moll, and N.I. Meyer, Phys. Rev. **130**, 972 (1963)
[32] M. Waldner, "Hot electron emission from silicon surfaces", J. Appl. Phys., **36**, 188 (1965)

[33] R.V. Bellau, and A.E. Widdowson, "Some properties of reverse-biased silicon carbide p-n junction cold cathodes", J. Phys., **D5**, 656 (1972)
[34] G.G.P. van Gorkom and A.M.E. Hoeberechts, "Electron emission from depletion layers of silicon p-n junctions", J. Appl. Phys. **51**, 3780 (1980)
[35] A.M.E. Hoeberechts, "Novel silicon avalanche diode as a direct modulated cathode with integrated planar electron-optics", Technical Digest of IEDM, 1990 p. 155
[36] A.M.E. Hoeberechts, "Novel silicon avalanche diode as a direct modulated cathode with integrated planar electron-optics", Technical Digest of IEDM, 1990 p. 155
[37] G.G.P. van Gorkom and A.M.E. Hoeberechts, "Silicon cold cathodes", Phillips Tech. Rev. **43**, 49 (1987)
[38] G.G.P. van Gorkom and A.M.E. Hoeberechts, "An efficient silicon cold cathode for high current densities", Philips J. Res., 39, 51 91984)
[39] J.Y. Ea, D. Zhu, Y. Lu, B. Lalevic, and R. J. Zeto, "Silicon avalanche cathodes and their characteristics", IEEE Trans. Electron. Devices, **ED-38**, 2377 (1991)
[40] T. van Zutphen, "Development of a GaAs avalanche electron-emitting diode cold electron emitter", IEEE Trans. Electron. Devices, ED-36, 2715 (1989), and T. van Zutphen, "Electron emission from a GaAs avalanche electron emission diode", Microelectronic Engineering, **11**, 333 (1990)
[41] W. N. Jiang, D. J. Holcombe, M. Hashemi and U. K. Mishra, "GaAs Planar Doped Barrier Vacuum Microelectronic Electron Emitters", IEEE Elec. Dev. Lett., vol. **14**, p. 143 (1993)
[42] W. N. Jiang and U. K. Mishra, "1 % efficiency AlGaAs Planar Doped Barrier Electron Emitters", to be published in Electron. Lett.
[43] S.Z. Sze, D.J. Coleman, and A. Loya, "Current transport in metal-semiconductor-metal (MSM) structures", Solid State Electron., **14**, 1209 (1971)
[44] R.F. Kazarinov and S. Luryi, "Carrier injection over triangular barriers in unipolar semiconductor structures", Appl. Phys. Lett., **38**, 15 (1981)
[45] GaAs PDBEE
[46] S.M. Sze,"Physics of semiconductor devices" 2nd edition, (John Wiley & sons, 1981)
[47] W. N. Jiang and U. K. Mishra, "Current transport mechanisms in GaAs planar-doped-barrier diodes with biuilt-in figh fields", to be published in J. Appl. Phys., Nov. (1993)
[48] L. Esaki, Phys. Rev., **109**, 603 (1958)
[49] Although the AlGaAs PDBEE has the same designed structure as that of GaAs one, the actual barrier structures could be different due to the fact that Be atoms tend to diffuse and segregate towards the surface during the MBE growth.

Table I Structure of PDB diodes

Layer	Material	Thickness	Dopant	Doping
Channel*	GaAs	400nm	Si	$5x10^{18}cm^{-3}$
	GaAs	15nm†		
Barrier			Be	see Table II
	GaAs	25nm†		
Injector	GaAs	400nm	Si	$5x10^{18}cm^{-3}$

* We have used the term of *Channel* to name the top n^+ layer in real PDBEEs

† These are designed parameters. The real dimensions of these layers may be different from the designed ones due to the diffusion and segregation of Be during the growth.

Table II Be delta-doping density

Device ID	Σ ($x10^{12}cm^{-2}$)
A	20.0
B	10.0
C	8.4
D	6.7
E	5.0

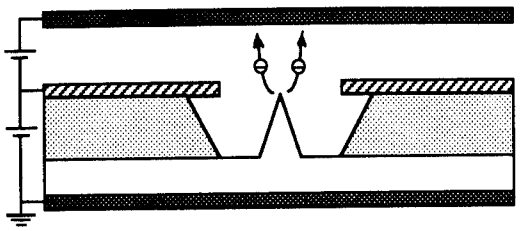


Fig. 1
Schematic illustration of a field emitter

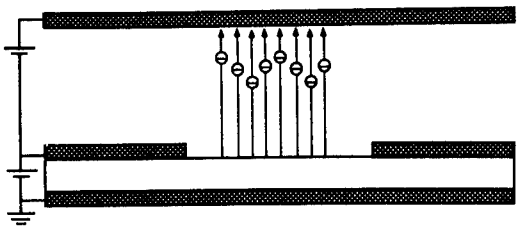
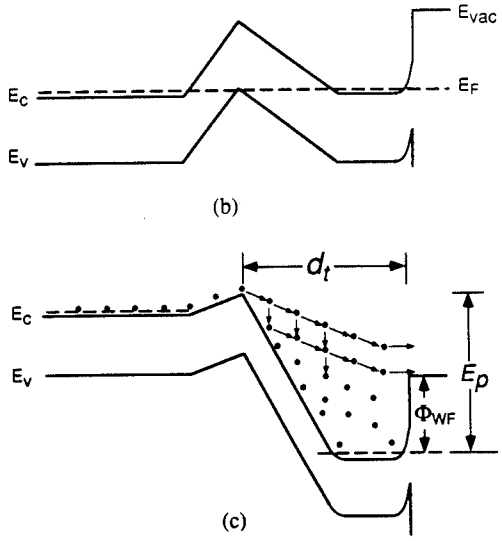


Fig. 2
Schematic illustration of a planar emitter



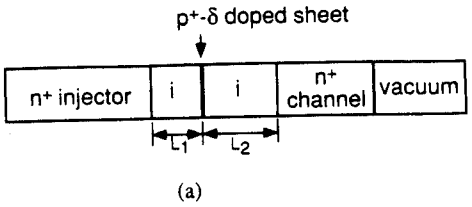


Fig. 3
Schematic diagram of a PDBEE showing (a) layer sequence, (b) band diagram under zero-bias, (c) band diagram under forward bias

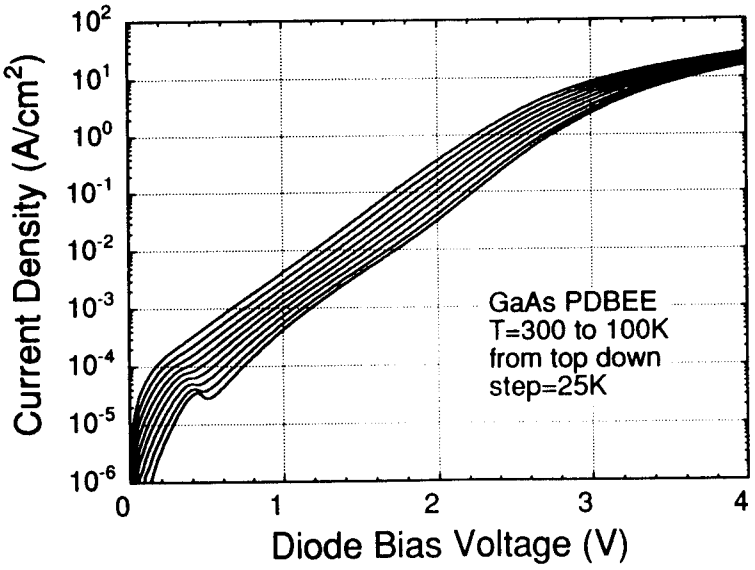
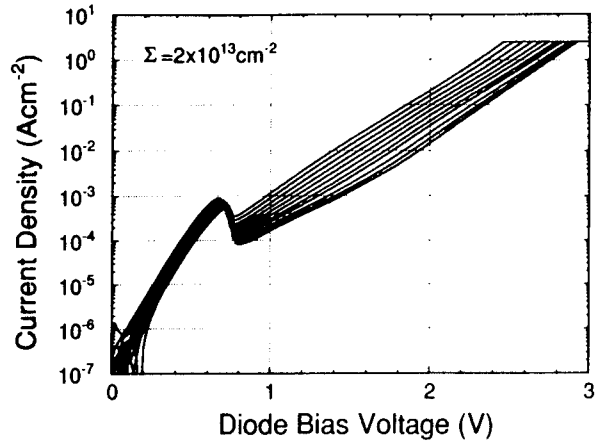
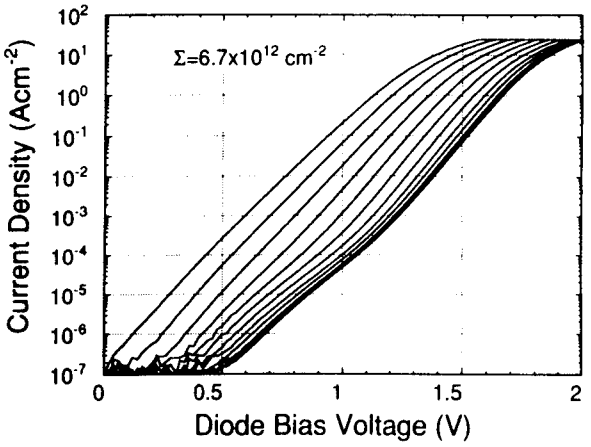


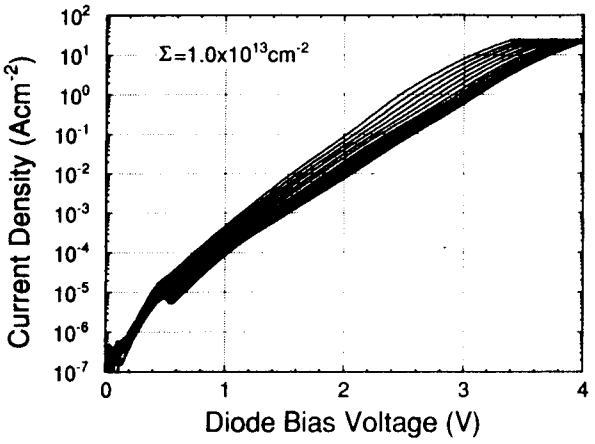
Fig. 4 Temperature dependent I-V characteristics of the GaAs PDBEE



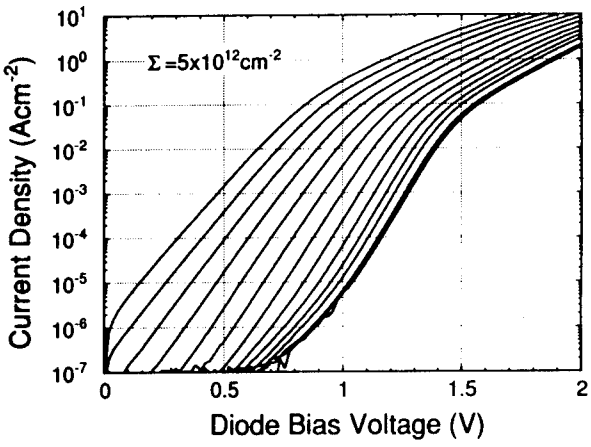
(a)



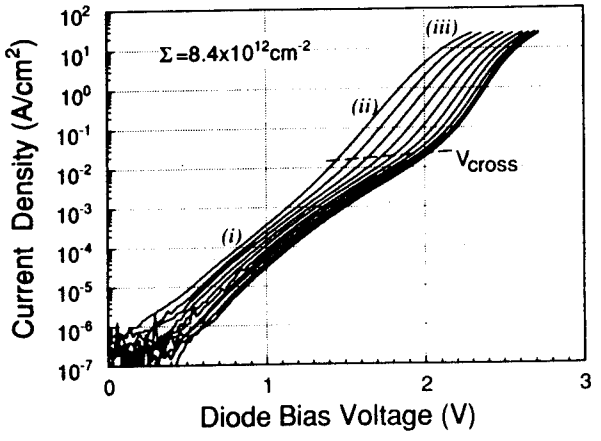
(d)



(b)



(e)



(c)

Fig. 5 Temperature dependent I-V characteristics of five test GaAs PDB diodes. with the p+ doping density of (a) $2x10^{13}cm^{-2}$, (b) $1x10^{13}cm^{-2}$, (c) $8.4x10^{12}cm^{-2}$, (d) $6.7x10^{12}cm^{-2}$, and (e) $5x10^{12}cm^{-2}$.

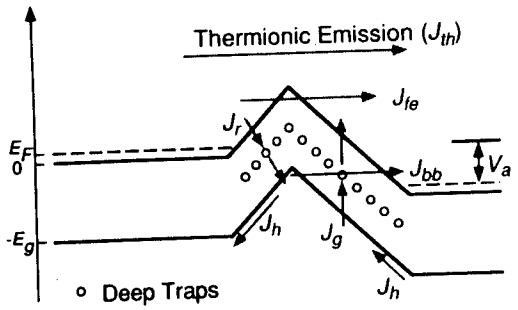


Fig. 6 Schematic illustration of the band diagram under forward bias showing possible carrier transport mechanisms.

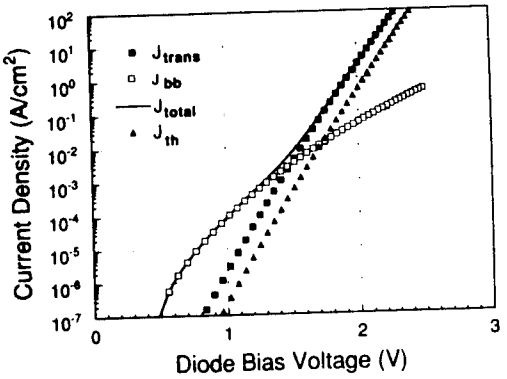


Fig. 7 Simulation results showing different current components across the test PDB diode C.

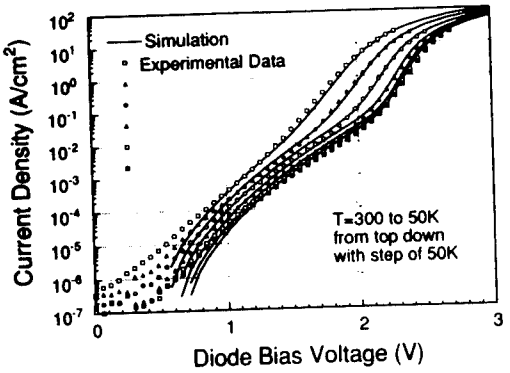


Fig. 8 Comparison between the simulation and experimental data at different temperatures. The open squares are experimental data, and solid lines are simulated results.

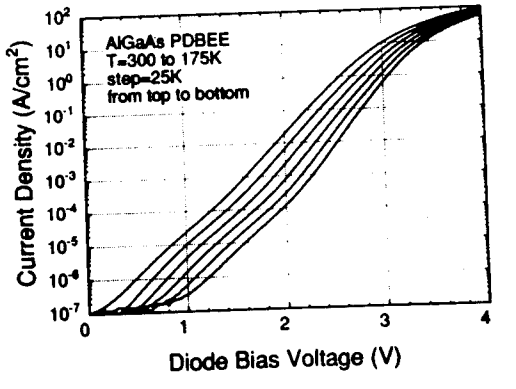
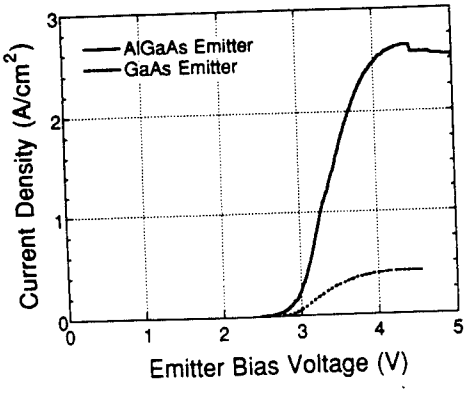
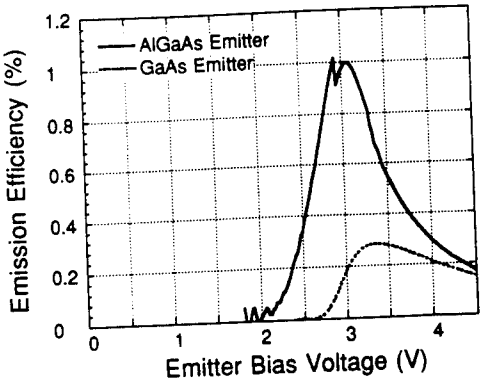


Fig. 9 Temperature dependent I-V characteristics of the Al0.3Ga0.7As PDBEE.



(a)



(b)

Fig. 10 Emission characteristics of the GaAs and Al0.3Ga0.7As PDBEEs showing (a) emission current density vs. diode bias and (b) emission efficiency vs. diode bias.

InGaAs/AlAs/InGaAsP Resonant Tunneling Bipolar Transistors Grown by Chemical Beam Epitaxy

W. L. Chen, G. O. Munns, D. Knightly, J. R. East, and G. I. Haddad

Solid-State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109-2122

Abstract

Resonant tunneling bipolar transistors (RTBT's) have been systematically studied using chemical beam epitaxy (CBE) for the first time. The RTBT structure studied is a InP-based transistor, consisting of single or multiple AlAs/$In_{0.75}Ga_{0.25}As$/AlAs RTD's in the emitter layer of a conventional heterojunction bipolar transistor (HBT) and an InGaAs or InGaAsP collector layer. Using the InGaAsP collector layer, the RTBT showed an improvement of breakdown voltage from 4 V to 10 V. The averaged dc β's are around 10 and 20 at 300 K and 77 K, respectively. In the transfer I-V characteristics, the RTBT showed 1 to 4 negative differential transconductance (NDT) peaks with peak-to-valley current ratios of 1.5 to 5.28 at 300K. Using such NDT peaks, several RTBT digital functions were demonstrated at room temperature, including a frequency multiplier and exclusive NOR gate.

I. INTRODUCTION

Resonant tunneling bipolar transistors (RTBT's) [1],[2],[3] have attracted a great deal of interest because of their low power and high speed capability for applications in digital circuits. Using the resonant tunneling mechanism, resonant tunneling transistors (RTT's) have negative differential transconductance (NDT) in the output current, which offers the possibility of greater circuit density, higher speed, and increased number of logical functions per device over conventional transistor technology [4]. For example, a full adder was demonstrated using only 7 to 10 resonant tunneling hot electron transistors (RHET's) [5] at cryogenic temperature. The RTBT's studied can be realized at room temperature, which provides advantages over the RHET's and hot electron transistors (HET's).

To date, there has been few investigations of RTBT's using CBE, which is capable of growing P-containing compounds and InP lattice-matched materials, providing materials with good transport and breakdown characteristics. For example, compared to GaAs, InGaAs is a better base material for RTBT's because it provides less surface Fermi level pinning to reduce the surface depletion depth in the extrinsic base region, and it also has a larger Γ to L band

0-7803-0894-8/93 $3.00

separation (0.56 eV) to avoid intervalley scattering. Using InP as a collector material also has several advantages over using InGaAlAs or AlGaAs, including (1) a larger Γ to L band separation (0.69 eV) (2) a larger saturation velocity (2.4×10^7 cm / sec), (3) a lower oxygen reactivity and (4) a higher chemical etching selectivity over InGaAs. Furthermore, chemical beam epitaxy (CBE) is capable of growing P-containing compounds and InP lattice-matched materials with abrupt heterointerfaces for resonant tunneling structures. Such capabilities do not exist simultaneously either in a MBE or MOCVD machine.

The purposes of this study are to demonstrate: (1) the growth capability of CBE for RTBT's; (2) high speed and high breakdown RTBT's and (3) RTBT digital functions for reducing the complexity of conventional transistor technology. Therefore, the paper starts from a brief description of the growth optimization of resonant tunneling structures using CBE. After the growth, a novel self-aligned process will be described for fabricating the RTBT's studied. An InGaAsP collector layer is used to increase the breakdown voltage of the RTBT's. Then device performance will be reported at 77 K and 300 K, respectively. Finally, some RTBT digital functions at room temperature are discussed.

II. Experiments

(a) Material growth

The semiconductor samples used in this study were grown by using an Intevac (Varian) Gen II Chemical Beam Epitaxy (CBE) reactor. All the structures were grown on (100) Fe-doped semi-insulating InP substrates at a substrate temperature of around 500° C. The substrates were prepared using a standard cleaning procedure, then loaded into the CBE reactor for growth. The source materials were trimethylindium (TMI), triethylgallium (TEG), trimethylamine alane (TMAA), 100% phosphine (PH_3) and 100% arsine (AsH_3) for the $In_{0.53}Ga_{0.47}As$, $In_{0.75}Ga_{0.25}As$, 1.1 μm InGaAsP lattice matched to InP and AlAs. Si and Be were used for the n-type and p-type dopants.

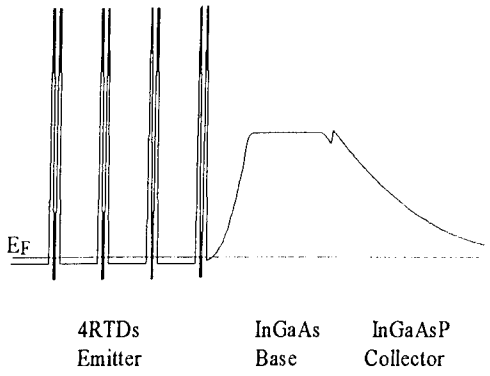
The first structure grown was $In_{0.53}Ga_{0.47}As$/AlAs/$In_{0.75}Ga_{0.25}As$/AlAs/$In_{0.53}Ga_{0.47}As$ RTD structure to optimize the growth conditions for obtaining reasonable PVR's and suitable sub-band energy for avoiding Γ-to-L intervalley scattering. Different growth temperatures and growth interrupts were studied. It was found that (1) high substrate temperature induced indium segregation and interface roughness problems, which prevented resonant tunneling; (2) appropriate growth interrupts stabilized the gas switching sequences and smoothed the AlAs/$In_{0.75}Ga_{0.25}As$/AlAs highly strained interfaces. Furthermore, using hydride drying filters could reduce the oxygen incorporation in the AlAs barriers, resulting in about 2 times improvement in PVR's of the RTD's grown. A detailed discussion can be found in Ref. [6].

To produce multiple resonant tunneling peaks for multistate logic, several RTD's were integrated together and separated by a 500 Å n^+ InGaAs layer to decouple the RTD's. Then the

multiple RTD's, including barriers and wells, were uniformly doped at $5\times10^{17}\,cm^{-3}$ to reduce the series resistance and hysteresis effects.

To increase the breakdown voltage, an InGaAsP layer was incorporated into the collector of the RTBT. The quality of such layer is critical to device performance, thus 1.1 μm InGaAsP is calibrated for good surface morphology, mobility and lattice matching. The bulk materials calibrated showed good qualities from X-ray and low temperature PL measurements. The double crystal X-ray rocking curve of a typical 1.1 μm InGaAsP indicated clear Pendellösung fringes and lattice matching to InP within 70 arc seconds ($|\Delta a / a| < 2.6\times10^{-4}$). The PL measurement was performed at 18K, showing a FWHM of 5.6 meV.

After optimization of RTD's and InGaAsP, several RTBT's were grown. The typical RTBT structure, as shown in Fig.1, consists of single or multiple RTD's in the InGaAs emitter layer ($5\times10^{17}\,cm^{-3}$), an InGaAs base doped with Be ($1.5\times10^{19}\,cm^{-3}$), and an InGaAsP or InGaAs collector layer ($2\times10^{16}\,cm^{-3}$). There is a 500 Å InGaAs layer between the base and the serial RTD regions to assure the potential profiles of the RTD's are the same at zero bias condition.



(a)

	cm-3		nm	
n+	5E18	InGaAs	200	
n	5E17	InGaAs	50	
n	5E17	AlAs	2	↑
n	5E17	InGaAs	4.5	
n	5E17	AlAs	2	x 4
n	5E17	InGaAs	50	↓
undoped		InGaAs	3	
p+	1.5E19	InGaAs	100	
undoped		InGaAs	30	
n-	2E16	InGaAs(P)	300	
n+	5E18	InGaAs	600	

Fe doped InP substrate

(b)

Fig. 1 A typical InGaAsP RTBT structure. (a) band diagram; (b) material structure.

(b) Device fabrication and measurement

The RTBT's were fabricated by a novel self-aligned process. A Ni/Ge/Au/Ti/Au metal system was used as emitter/collector contacts and a Pd/Ge/Ti/Au system was used as a shallow ohmic base contact. Selective wet etching, succinic acid, was adopted for base recess. Moreover, a low temperature PECVD SiO_2 was deposited for interconnect dielectrics and for preventing the lateral conduction of emitter mesa and base metal. Finally, base mesa etching was achieved using a combination of RIE and chemical etching to avoid undercut and damage. A typical processed RTBT is shown in Fig. 2.

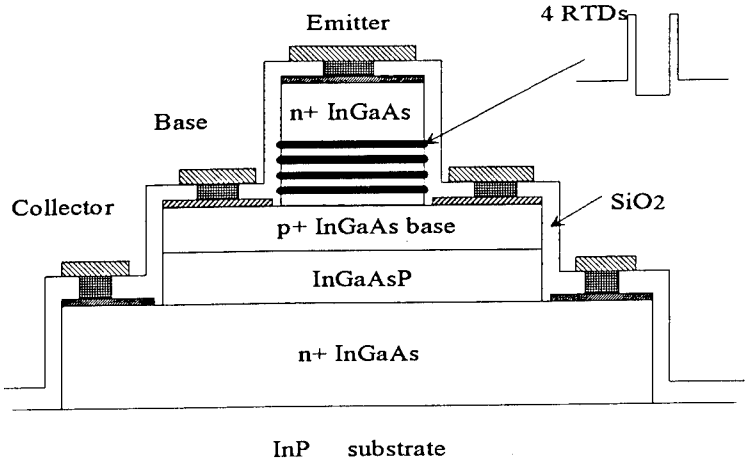


Fig. 2 A typical processed RTBT.

Microwave and dc measurements were performed using self-aligned and non self-aligned devices at 77 K and 300 K, respectively. For demonstration of the functionality of RTBT's, external resistors and power supplies were connected to the device. All the digital functions were obtained at room temperature.

III. Results and discussion

(a) Sub-band energy

The sub-band energies of the RTD determine the electron injection energy into the base of a RTBT. Such an energy should be large enough to obtain reasonable current gain and should be smaller than the upper valley energy to avoid intervalley scattering. Moreover, in order to make the RTBT's and RTD's useful for circuit applications, high PVR's are required, together with a current density greater than $10^4 A / cm^2$

Experimentally, the RTD optimization was carried out with different well thicknesses and identical AlAs barrier thickness (20 Å). After optimization, the best RTD obtained had a PVR of 8 and 20 at 300 K and 77 K, respectively, with a current density of $5 \times 10^4 A / cm^2$ and a resonant voltage of 0.35 V. However, the RTBT's fabricated showed the highest PVR of 5, indicating that some of the injected hot electrons were scattered while traveling through the long, heavily doped base region. Such electrons contributed to the valley current, resulting in a decrease of PVR in the transfer I-V characteristics of the RTBT.

(b) RTBT's with an InGaAsP collector layer

After the optimization of RTD's and bulk materials, several RTBT's were grown and fabricated. The first RTBT study was to investigate the effect of the InGaAsP collector layer. Thus two identical RTBT's were grown with a single serial RTD in the emitter. The difference is that the first RTBT has an InGaAs collector and the second one has 1.1 μm (Eg ~ 1.1 eV) InGaAsP layer lattice matched to InP as a collector layer. A comparison of the device results is shown in Fig. 3. It is shown that the first RTBT shows a dc current gain (β) of 10 and a V_{CE0} breakdown voltage of 4 V. The other one has a smaller β of 4 and a larger V_{CE0} breakdown voltage of 10 V.

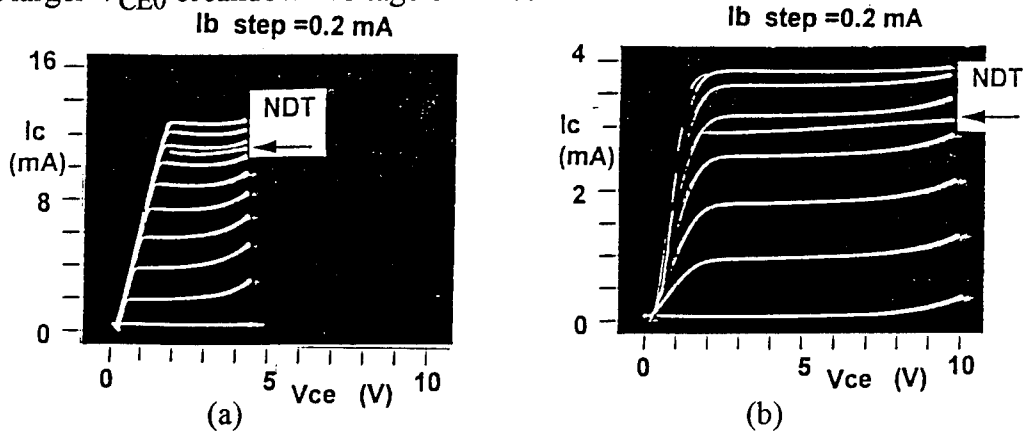


Fig. 3 The RTBT common emitter I-V characteristics at 300 K. (a) with InGaAs collector; (b) with InGaAsP collector.

The operation of the RTBT behaves like a regular transistor if the V_{BE} (I_B) bias is less than the built-in voltage of the InGaAs p-n junction, resulting in most of the bias voltage drop across this junction. In this region, the collector current increases with V_{BE} (I_B) until the base-emitter reaches the flat band condition. Beyond the flat-band, additional V_{BE} will start to fall across the RTD structure, causing the quenching of resonant tunneling and a decrease of collector current [7]. From the common emitter I-V characteristics, the first RTBT starts to exhibit NDT performance at I_B greater than 1.8 mA, Fig. 3(a), and the other one starts to exhibit NDT at I_B greater than 1.4 mA, Fig. 3(b).

The transfer I-V characteristics are shown in Fig. 4 for the RTBT with an InGaAsP collector layer. In this configuration, I_C is measured against V_{BE} with a fixed V_{CE} at 300 K. The peak current is 3.7 mA at V_{BE}= 1.8 V and the valley current is 0.7 mA at V_{BE}= 3.6 V, representing a PVR of 5.28. Consequently, such a RTBT is constructed with external resistors to form an inverter gate. An increase of input voltage (Vin) causes a decrease of output voltage (Vout) until the device reaches the NDT region, where an increase of Vout is observed with Vin. As shown in Fig. 5, a frequency multiplier, constructed using the same circuit, can increase the frequency of sine waves by a factor of three.

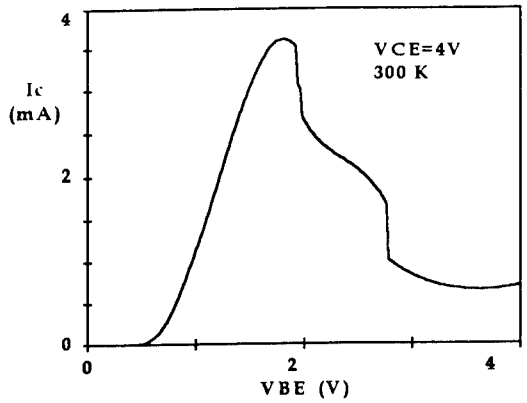


Fig. 4 The RTBT transfer I-V characteristics.

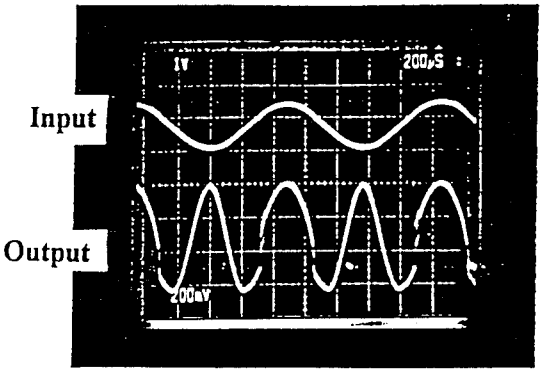
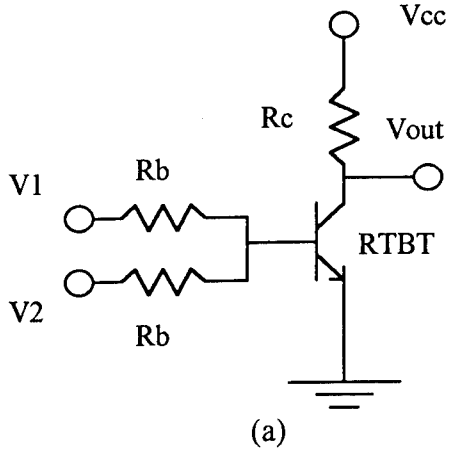


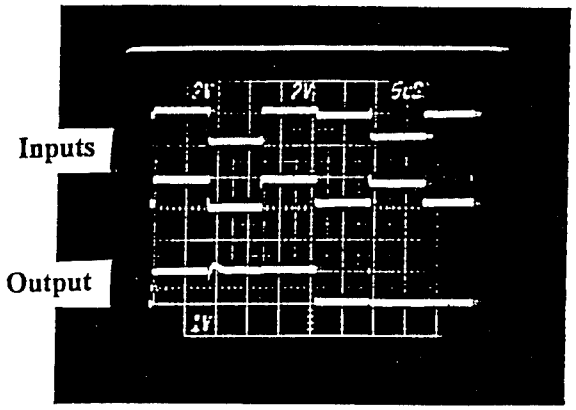
Fig. 5 Frequency response of the RTBT frequency multiplier.

c) Room temperature XNOR

An exclusive NOR function was also demonstrated at room temperature using the RTBT circuit shown in Fig. 6(a). Basically, the XNOR circuit can be treated as an "inverter" circuit with two voltage inputs, V1 and V2. As shown in Fig. 6(b), if V1 equals to V2, both are in high or low input states, the "inverter" circuit sees an input voltage, corresponding to a high Vout state. On the other hand, if V1 is not equal to V2, one is in high and the other is in low input state, resulting in a low output state. This logic behavior demonstrates an XNOR function. A stable RTBT with a sufficient PVR is critical to realize room temperature XNOR. Normally, more than 10 conventional transistors are required to realize such a function. Thus it is believed that using RTBT can reduce the complexity of conventional transistor technology in digital circuit applications.



(a)



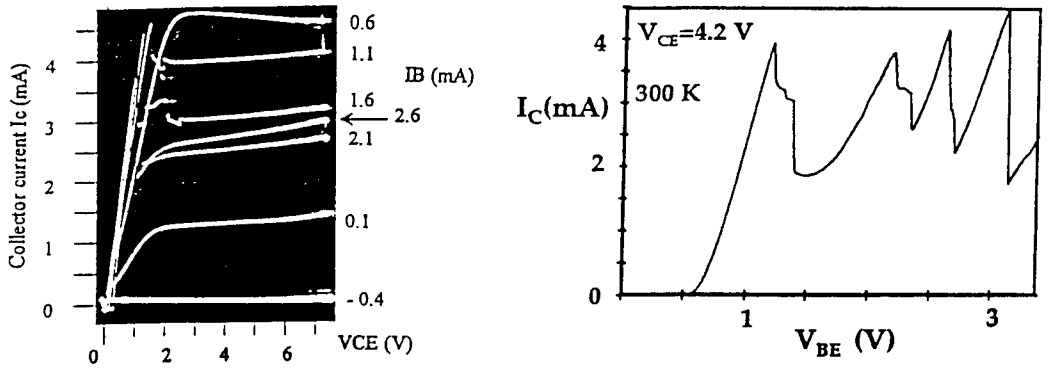
(b)

Fig. 6 The RTBT XNOR at 300 K.

An S-parameter measurement was also conducted to estimate the microwave performance of the InGaAsP RTBT. For a self-aligned transistor with an emitter area of $5 \times 20 \mu m^2$, the highest cutoff frequency (f_T) and fmax observed are around 15 and 9 GHz.

(d) RTBT's with multiple NDT peaks

In order to realize more sophisticated digital functions using a fewer number of transistors, RTBT's with multiple NDT peaks were studied. A uniform doped RTD structure was used to reduce the hysteresis problem. The common emitter I-V characteristics are shown in Fig. 7(a) for a RTBT with 4 stacked RTD's in the emitter and an InGaAsP collector layer. The RTBT behaves like a regular transistor when I_B < 1.1 mA, and starts to show NDT characteristics when I_B > 1.1 mA. This transistor has dc β's of 10 and 20 at 300 K and 77 K, respectively, with a V_{CE0} breakdown voltage of 10 V. Moreover, there are 4 NDT peaks observed from the transfer I-V characteristics with PVR's of 1.5 to 2.7 at room temperature, as shown in Fig. 7(b). It is believed that the 2-step current drop in the first two NDT peaks are due to the circuit instability.

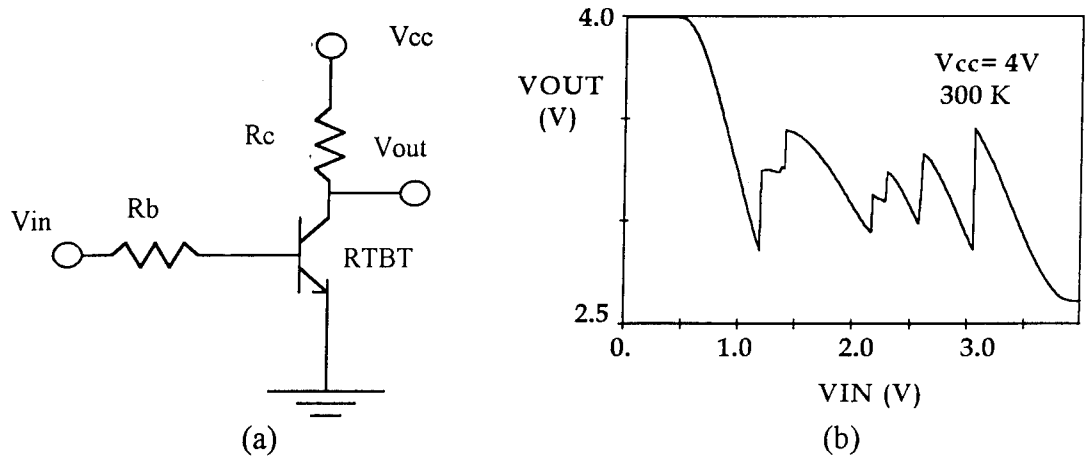


(a) (b)

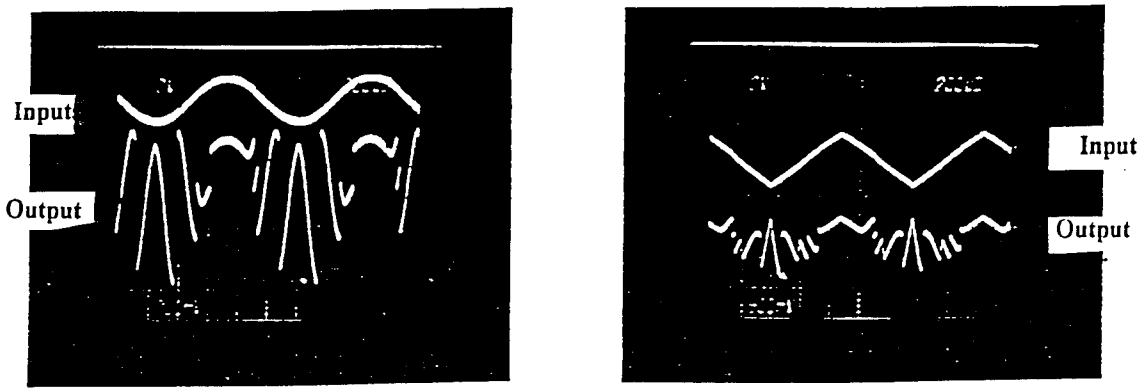
Fig. 7 (a) The common emitter I-V characteristics for the RTBT with 4 stacked RTD's in the emitter and an InGaAsP collector layer; (b) the transfer I-V characteristics at 300 K.

(e) Frequency multiplier

Using the RTBT with 4 NDT peaks, an inverter circuit was constructed with the Vin-Vout characteristics as shown in Fig. 8(a), (b). Using such a frequency multiplier, triangular input waves are multiplied by a factor of 7 and sine input waves by a factor of 9, as shown in Fig. 8(c) and (d).



(a) (b)



(c) (d)

Fig. 8 (a) The invertor circuit using the RTBT shown in Fig. 7; (b) the Vin-Vout plot for the invertor; (c) the response of the RTBT frequency multiplier for sine waves; (d)the response for triangular waves.

IV. Conclusions

RTBT's with multiple NDT peaks have been realized using CBE for the first time. Both reasonable dc and microwave performance were obtained at room temperature. These RTBT's were used to demonstrate digital functions at room temperature, including frequency multiplier and XNOR. It is believed that using RTBT's can reduce the complexity of conventional transistor technology significantly.

Acknowledgment: This work is supported by the U. S. Army Research Office under the URI program Grant No. DAAL03-92-G-0109. The authors would also like to thank Solti Peng and John Cowles for microwave measurements.

References

[1] F. Capasso, S. Sen, A. C. Gossard, A. L. Hutchinson, and J. H. English, "Quantum well resonant tunneling bipolar transistor operating at room temperature," *IEEE Electron Dev. Lett.* **EDL-7**, 573 (1986).

[2] T. Futatsugi, Y. Yamaguchi, K. Ishii, K. Imamura, S. Muto, N. Yokoyama, and A. Shibatomi, "A resonant tunneling bipolar transistor (RBT): A proposal and demonstration for new functional device with high current gains," *Tech. Dig. IEDM*, p.286 (Dec. 1986).

[3] S. Sen, F. Capasso, A. Y. Cho, and D. L. Sivco, "Multiple state resonant tunneling bipolar transistor operating at room temperature and its application as a frequency multiplier," *IEEE Electron Dev. Lett.* **9**, 533 (1988).

[4] A. C. Seabaugh, W. R. Frensley, Y. C. Kao, J. N. Randall, and M. A. Reed, "Quantum-well resonant tunneling transistors'" the proceedings of the 1989 IEEE Cornell conference, Ithaca, N. Y. p. 255.

[5] K. Imamura, M. Takatsu, T. Mori, T. Adachihara, H. Ohnishi, S. Moto, and N. Yokoyama, "A full adder using resonant-tunneling hot electron transistors (RHETs)," *IEEE Trans. Electron Dev.* **39**, 2707 (1992).

[6] W. L. Chen, G. O. Munns, L. Davis, P. K. Bhattacharya, and G. I. Haddad, "The growth of resonant tunneling hot electron transistors using chemical beam epitaxy," the 4th International Conference in Chemical Beam Epitaxy, section S-7, July 1993.

[7] F. Capasso, S. Sen, and F. Beltram, "Quantum-effect devices", *High-Seed Semiconductor Devices* ed. by S. M. Sze, (John Wiley & Sons, New York, 1990) Chapter 8.

RESONANT INTERBAND AND INTRABAND TUNNELING IN InAs/AlSb/GaSb DOUBLE BARRIER DIODES

J.L. Huber and M.A. Reed

Department of Electrical Engineering, Yale University
New Haven, Connecticut 06520

G. Kramer and M. Adams

Motorola Inc., Phoenix Corporate Research Laboratories
Tempe, Arizona 85284

C.J.L. Fernando and W.R. Frensley

University of Texas at Dallas
Richardson, Texas 75083

ABSTRACT

We have realized a series InAs/AlSb/GaSb tunneling structures in which both interband and intraband tunneling occur, dependent on injection energy. The baseline structure consists of a single InAs well with GaSb barriers which serve as quantum wells for interband tunneling and barriers for intraband tunneling. At low biases, interband tunneling occurs through a coupled double well structure in the GaSb valence bands. At higher biases, intraband tunneling occurs through the InAs quantum well. The addition of a thin AlSb barrier at different points in the structure changes both the strength and number of peaks in the I-V/G-V characteristics.

I. INTRODUCTION

Resonant tunneling diodes (RTDs) usually consist of a double barrier, single quantum well semiconductor heterostructure where the center region has quasi-bound quantum well states accessible to electron (or hole) transport [1]. RTDs are interesting from a device standpoint because of the possibility of fast response time for both oscillators and bistable devices [2]. Resonant interband tunneling (RIT) devices differ from conventional RTDs in that tunneling occurs from an electron-like state to a hole-like state (or vice versa).

The InAs/AlSb/GaSb material system is suitable for fabrication of RIT devices because of the type II nature of the InAs/GaSb interface which has the GaSb valence band edge sitting ~150 meV above the InAs conduction band edge (Figure 1). This allows device structures where electrons from an InAs emitter are injected directly into a quasi-bound quantum state in the GaSb valence band [3], [4]. We have investigated the tunneling characteristics of a series of double barrier InAs/AlSb/GaSb structures. The structures were designed such that both intraband and interband tunneling occur, depending on injection energy.

0-7803-0894-8/93 $3.00

II. DEVICE FABRICATION

The structures were grown in a Varian (Intevac) Gen II solid-source MBE system. The growth rates of GaSb, AlSb, and InAs were 1.0, 1.0, and 0.8 μm/h, respectively. Before growth the wafers were heated to 610 °C for 15 minutes to desorb the oxide. The first epi layer, 0.2 μm of GaAs, was deposited at T_{sub} = 580 °C to smooth the wafer surface. A 0.5 μm superlattice buffer region, consisting of a GaSb/AlSb superlattice followed by 0.4 μm of AlGaSb, was then grown at a T_{sub} of 530 °C. The T_{sub} was then lowered to ≈ 500 °C for the growth of the active region of the structure, which consisted of 1μm n+ InAs ($2x10^{18}$ cm^{-3}), 50 nm of lightly doped InAs ($2x10^{16}$ cm^{-3}), the double barrier layer sequence, 50 nm of lightly doped InAs ($2x10^{16}$ cm^{-3}) and then a cap layer of 0.25 μm of n+ InAs ($2x10^{18}$ cm^{-3}).

Mesa devices were formed using standard lithographic and wet etch techniques with non-alloyed Ni/Ge/Au contacts. The surface of the sample was covered with an insulating layer of silicon nitride, with via holes to allow contact between the device contacts and bonding pads.

III. RESULTS AND DISCUSSION

A series of four structures, based on a single design, was investigated. The baseline structure consisted of two 9.5 nm GaSb barriers and a 5.5 nm InAs quantum well (all nominally undoped). A thin AlSb barrier was added, in various positions, to the remaining three structures. Self-consistent band diagrams, not including quantum effects, for the structures are shown in Figure 2a, 3a, 4a, 5a. In structure #2, the AlSb barrier is 1.5 nm thick and each InAs well layer is 3.5 nm thick, while in structure #3 and #4 the AlSb barrier is 2.5 nm thick. A 10 nm undoped InAs spacer layer was placed between both n+ contacts and the double barrier structure for all devices.

At low biases, interband tunneling occurs through a coupled double well structure in the GaSb valence band. The GaSb layer thicknesses were chosen such that for a single GaSb layer, only one light hole quantum well state should exist in the GaSb valence band above the InAs conduction band edge [5]; selection rules forbid coupling between conduction band states and heavy hole states. The n = 2 light hole states lie well below the window to the InAs conduction band and therefore, are not available for interband tunneling. The InAs well thickness was chosen such that lowest energy quantum well state would be above the GaSb valence band edge. Thus, at higher biases, intraband tunneling occurs through a single InAs quantum well state.

We have used the band profiles generated by a self-consistent Poisson solver to assist in the labeling of the various NDR regions. The energy levels of the InAs quantum well states were then calculated using a single band model. This method is precise for modeling the intraband resonant state energies because the InAs quantum well states lie above the window between the InAs conduction band and the GaSb valence band. A rough estimate of the interband resonances can be made using the single band model by solving for the energy levels of the valence band states. However, this is not precise because of the coupling between the electron-like states in the InAs and the hole-like states in GaSb, and should only be treated as an approximation.

For all the structures, with one exception, the intraband resonances occurred at a slightly lower voltage than predicted by the model. The exception was for the only case when intraband tunneling occurred initially thorough a thin AlSb barrier (reverse bias direction of structure #3). For this case, the resonance occurred where predicted. This implies that there is some interaction between the GaSb valence band states and the InAs conduction band state. For the case where the resonance occurred where predicted, the AlSb barrier reduces this interaction.

Basline structure

The I-V/G-V curve for structure #1, measured at 77 K, from a 32 μm square device is shown in Figure 2b. The I-V is symmetrical around zero bias, therefore only one bias direction is shown. The NDR region at 0.62 V corresponds to conventional intraband tunneling through the confined state in the InAs well (this resonance occurs at a slightly larger than expected bias, due to parasitic contact resistance). The NDR region near 0.09 V corresponds to interband tunneling through the confined state created by the near GaSb barrier. In addition to these NDR regions, a small feature appears in the G-V curve near 0.05 V, which may correspond to tunneling into the far GaSb well.

AlSb barrier in the center of the InAs well

Structure #2 is similar to the baseline structure (#1), with the addition of a 1.5 nm thick AlSb barrier in the center of the InAs well, creating, in effect, a double well (the total InAs well thickness is also 0.5 nm larger than that of structure #1). This should add a second intraband tunneling resonance. Because the AlSb barrier lies in the center of the well, little or no change to the interband tunneling characteristics are expected. The I-V/G-V curve at 4 K for structure #2 is shown in Figure 3b. The mesa size was 16 μm square.

The intraband resonance from the near InAs well is visible at 0.58 V. A weaker feature, assumed to be caused by tunneling into the far InAs well, appears at 0.46 V. A strong NDR appears at 0.17 V, consistent with interband tunneling into the near GaSb valence band well. A number of smaller unidentified features in the G-V curve are evident; at 0.26 V, and below 0.17 V. One of these features (occurring below V = 0.17 V) may be due to interband tunneling into the far GaSb well. The others do not correspond to any expected intraband resonances. They are also visible at 77 K, while similar features for structure #1 were not visible. This implies that they are the result of the addition of the AlSb barrier. At k = 0, selection rules allow coupling to the light hole band only. However, away from k = 0, coupling to the heavy-hole band is allowed [6]; thus the additional features may be due to interband tunneling through heavy hole states.

AlSb barrier outside the GaSb barrier

Structure #3 is similar to the baseline structure except for the addition of a 2.5 nm AlSb barrier on the outside of one of the GaSb barriers. The I-V/G-V curve from structure #3, measured at 4 K, from a 32 μm square device is shown in Figure 4b. Negative bias voltages correspond to the case where electrons initially tunnel into the AlSb barrier.

The intraband resonances for the structure appear at -0.90 V and 0.68 V (there are also small features, at -0.96 V and 0.89 V slightly above each of the intraband resonances). The interband resonances into the near GaSb well appear at -0.07 V and 0.18 V. Additional features occur below the large interband tunneling resonance, similar to those in structure #2. For the negative bias interband resonance, the peak-to-valley ratio is higher and the peak current smaller than for the forward bias resonance. This is expected since electrons initially tunnel through the AlSb barrier, consistent with the intraband peak current (29 mA @ +0.68 V, 13 mA @ -0.90 V), and with previous observations for other InAs/AlSb/GaSb tunneling structures [7].

AlSb barrier inside the GaSb barrier

Structure #4 is similar to the baseline structure except for the addition of a 2.5 nm AlSb barrier on the inside of one of the GaSb barriers, rather than on the outside. The I-V/G-V curve from structure #4, measured at 4 K, from a 32 μm square device is shown in Figure 5b. Positive

applied biases correspond to electrons initially encountering the GaSb barrier associated with the AlSb barrier.

In the positive bias direction, the intraband resonance occurs at 0.70 V, in the negative bias direction, the intraband resonance occurs at 0.52 V. Strong NDR regions, assumed to be interband resonances, occur at -0.04 V and -0.18 V. These two features are qualitatively similar, each with a small peak at a slightly higher bias (at -0.13 V and -0.23 V). Another small feature of similar strength is located at -0.31 V. Weaker structure occurs at 0.21 V, 0.49 V, and 0.96 V, and very weak structure occurs at 0.17 V, 0.28 V, and 0.59 V. Another strong NDR region occurs at -0.36 V, also with a weaker feature at the slightly larger bias (-0.47 V).

Based on the assumption that there is only coupling between the conduction band and light hole band, we expect to see one intraband resonance and two interband resonances for our structures. However, for the structures with the thin AlSb barrier, particularly structure #4, there are more resonances than are expected. Based on these results, it appear that a multiband model (# of bands > 2) is necessary to identify the various resonances.

IV. SUMMARY

We have realized InAs/AlSb/GaSb heteroepitaxial tunneling structures in which both interband and intraband tunneling occur, dependent on injection energy. The structures consist of a single InAs well with GaSb barriers which serve as quantum wells for interband tunneling and barriers for intraband tunneling. The addition of a single, thin AlSb barrier variably throughout the double barrier structure can greatly effect the nature of the features in the I-V/G-V characteristics. Intraband resonance can be identified using a single band model, when well separated from interband resonances. Though interband resonances can be approximated using a single band model, a multiband model is necessary to accurately model the resonant energies. The number of observed interband features is greater than can be accounted for by simple conduction band-light hole band tunneling, and may be due to heavy hole band coupling.

ACKNOWLEDGMENT

We would like thank Saied Tehrani and Herb Goronkin for helpful discussions, and the NNF at Cornell University for assistance with mask fabrication. This work was partially supported by a State of Connecticut Goodyear grant, NSF, and was partially performed under the management of FED (the R&D Association for Future Electron Devices) as a part of the R&D of Basic Technology for Future Industries supported by NEDO (New Energy and Industrial Technology Development Organization, Japan).

REFERENCES

[1] L.L. Chang, L. Esaki, and R. Tsu, *Appl. Phys. Lett.*, "Resonant tunneling in semiconductor double barriers," vol. 24, pp. 593-595, 1974.

[2] T.C.L.G.Sollner, W.D. Goodhue, P.E. Tannenwald, C.D. Parker, D.D. Peck, "Resonant tunneling through quantum wells at frequencies up to 2.5 THz," *Appl. Phys. Lett*, vol. 43, pp. 588-590, 1983.

[3] L.F. Luo, R. Beresford, and W.I. Wang, "Interband tunneling in polytype GaSb/AlSb/InAs heterostructures," *Appl. Phys. Lett*, vol. 55, pp. 2023-2025, 1989.

[4] J.R. Söderström, D.H. Chow, and T.C. McGill, "New negative differential resistance device based on resonant interband tunneling," *Appl. Phys. Lett*, vol. 55, pp. 1094-1096, 1989.

[5] R.Q. Yang and J.M. Xu, "Bound and quasibound states in leaky quantum wells," *Phys. Rev. B*, vol. 46, pp. 6969-6974, 1992.

[6] D.Z.Y. Ting, E.T. Yu, and T.C. McGill, "Role of heavy-hole states in interband tunnel structures," *Appl. Phys. Lett*, vol. 58, pp. 292-294, 1991.

[7] J.F. Chen, L. Yang, M.C. Wu, S.N.G. Chu, and A.Y. Cho, "On the effect of the barrier widths in the InAs/AlSb/GaSb single-barrier interband tunneling structures," *J. Appl. Phys*, vol. 68, pp. 3451-3455, 1990.

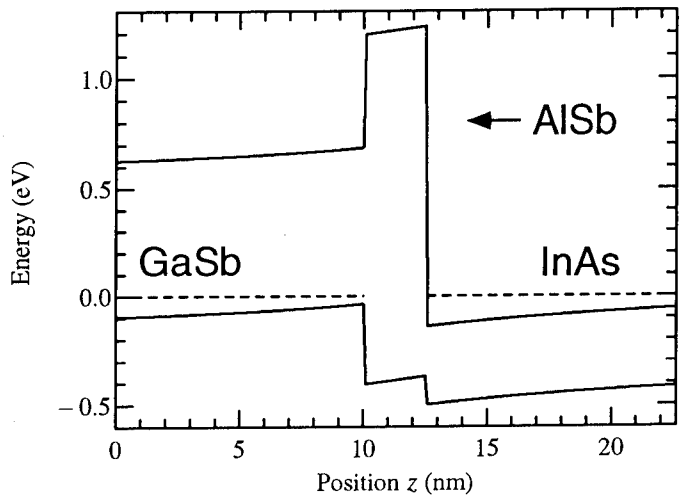
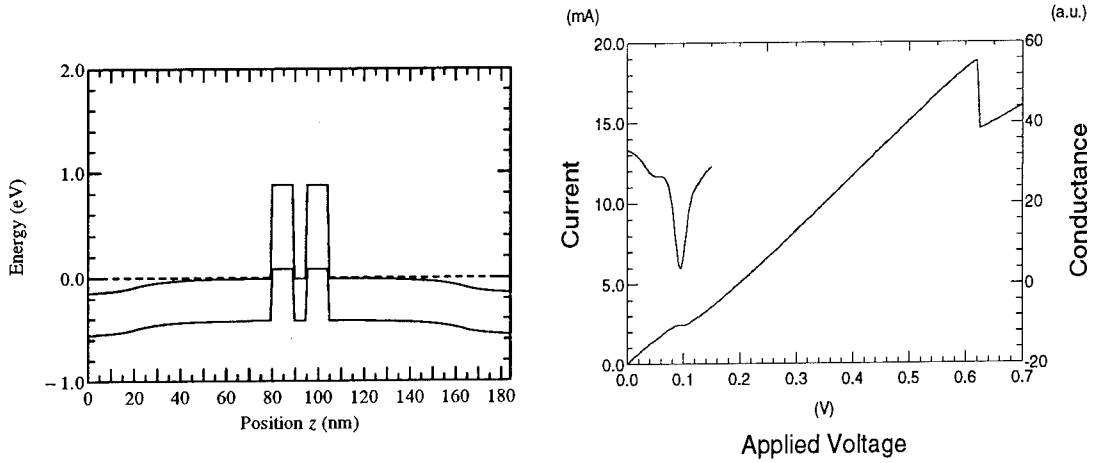
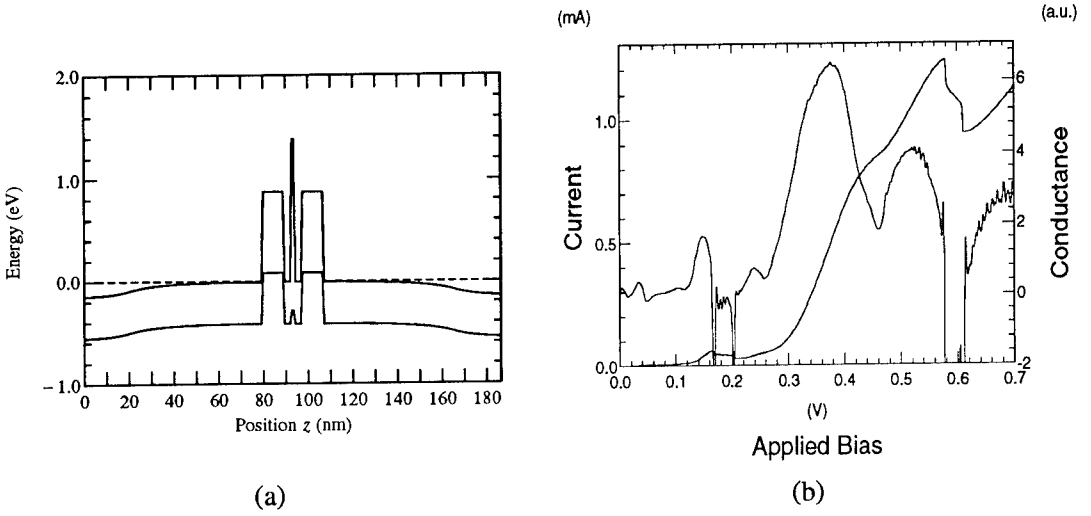


Figure 1. Self-consistent band diagram of a simple InAs/AlSb/GaSb structure. The GaSb valence band edge lies ~150 meV above the InAs conduction band edge allowing injection from the InAs conduction band into the GaSb valence band.



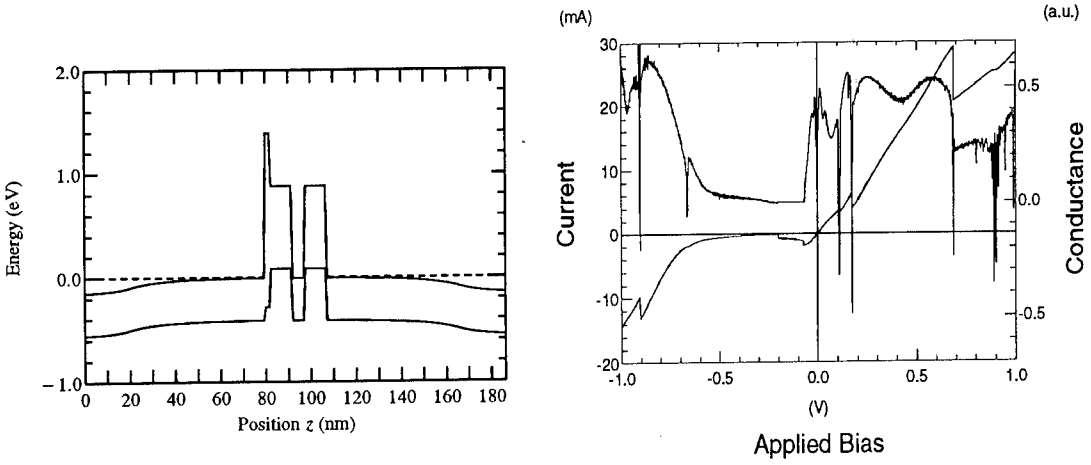
(a) (b)

Figure 2. (a) Self-consistent band diagram for structure #1. (b) I-V/G-V for structure #1. The G-V curve is not shown above V = 0.15 V.



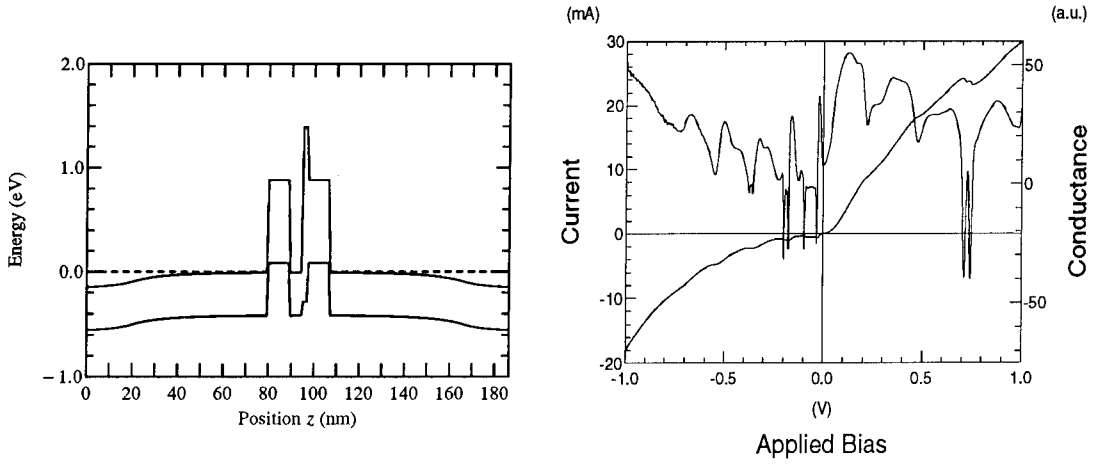
(a) (b)

Figure 3. (a) Self-consistent band diagram for structure #2. (b) I-V/G-V for structure #2.



(a) (b)

Figure 4. (a) Self-consistent band diagram for structure #3. (b) I-V/G-V for structure #3. The G-V curve between -0.07 V and -0.20 V has been approximated with a straight line for clarity.



(a) (b)

Figure 5. (a) Self-consistent band diagram for structure #4. (b) I-V/G-V for structure #4.

Ultrafast Operation of Heterostructure Bipolar Transistors Resulting from Coherent Base Transport of Minority Carriers

Serge Luryi

AT&T Bell Laboratories, Murray Hill, NJ 07974

Introduction

The possibility of using transit-time phase shifts in a bipolar junction transistor to extend the active transistor operation to higher frequencies has been discussed by a number of researchers. The discussion has been centered on the negative dynamic output impedance that can be achieved due to the phase delay in carrier drift across the collector space-charge region. This idea, proposed originally by Wright [1], is attractive for millimeter and submillimeter wave applications. It has gained a lot of attention in the context of a conventional junction transistor implementation [2,3], heterostructure bipolar transistors (HBT) [4-6], unipolar ballistic transistors [7], and field-effect transistors [8]. Unfortunately, in transistors with diffusive base transport, the practical possibility of extending the active transistor behavior beyond f_T is severely limited [6] by parasitic effects. Utilization of transit-time resonances, arising from carrier drift across the collector junction in a conventional GaAs/AlGaAs HBT, requires a reduction of the base and collector resistances, as well as the collector capacitance, by factors of 10 from estimates, based on the state-of-the-art technology.

This problem was recently analyzed by Grinberg and myself [9]. Like in transit-time diodes, the negative dynamic output impedance in transistors arises from a combination of the injection phase delay $\varphi = \varphi_E + \varphi_B$ (the total transit angle of the emitter and the base) and the drift delay θ in the base-collector junction. Transit-time effects are swamped by the device parasitics because in diffusive transistors the available phase φ is too small. Exponential decay of the magnitude of the base transport factor α_B with the phase φ_B, acquired in a diffusive propagation across the base, implies that the overall phase of the collector current ($\varphi_B + \varphi_E + \theta$) must be acquired almost entirely at the expense of the collector phase delay θ. On the other hand, the magnitude of the negative dynamic resistance also decreases with increasing θ and for $\theta \approx \pi$ it becomes too small to overcome the influence of parasitic elements.

Transit time resonances can withstand the parasitics only if they can be achieved while keeping the collector delay $\theta \lesssim 1$, i.e. if a sufficient phase $\omega\tau_B \equiv \varphi_B \gtrsim \pi$ is acquired in the base transport alone – without a significant penalty in the magnitude of the base transport factor α_B. In general, with increasing frequency ω, the base transport factor $\alpha_B(\omega) = |\alpha_B| \exp(-\omega\tau_B)$ describes an inbound spiral $|\alpha_B| \to 0$ in the complex plane [10]. We shall use the term *coherent* speaking of any base transport mechanism that provides a sufficiently *slow* spiraling in of $\alpha_B(\omega)$, say, such that $\alpha_B \gtrsim 0.5$ for $\varphi_B \gtrsim \pi$.

One possibility to achieve coherence arises from ballistic transport at cryogenic temperatures [9]. In the ballistic case the modulated signal injected at the base-emitter interface is washed out because of the thermal spread in the normal velocities of injected carriers, leading to a variance $\delta\tau$ in their base propagation time. The latter process is analogous to the Landau damping of density waves in collisionless plasmas and has an effect similar to diffusion. Collisionless base propagation by itself does not imply coherence.

Coherent regime in a ballistic transistor arises when $\delta\tau \ll \tau_B$, i.e., when the injected electrons form a collimated and monoenergetic beam. A good approximation to such a beam results from the passage of electrons across an abrupt heterointerface at low temperatures. For electrons traveling with a velocity v perpendicular to the base layer, a periodic injection modulation at the emitter interface

0-7803-0894-8/93 $3.00

with a frequency f sets up an electron density wave of wavelength $\lambda = v/f$. For $\delta\tau \ll \tau_B$, this wave does not appreciably decay over the entire base, resulting in a transport factor

$$\alpha_B = e^{-(\omega\delta\tau)^2/2}\, e^{-i\omega\tau_B} \qquad (1)$$

As discussed in detail in Ref. [9], a coherent transistor can have both the common-emitter current gain h_{21} and the unilateral power gain U exceeding unity at frequencies far above the usual f_T. The transistor speed is limited not by the base propagation time τ_B but rather by its variance $\delta\tau$. Calculations carried out for exemplary heterostructures, implemented at the state-of-the-art rules of technology, indicate that active transistor behavior can be extended to about 1 THz.

It is clear, however, that the ballistic coherent operation requires cryogenic temperatures. For a collisionless transport to hold over the entire base width, the electron kinetic energy Δ (the injection energy, corresponding to a conduction-band discontinuity in the base-emitter junction [11]) should not exceed the optical phonon emission threshold, $\Delta \lesssim \hbar\omega_{opt}$. On the other hand, to achieve coherence, the injection energy Δ must substantially exceed the thermal spread, whence we need $\hbar\omega_{opt} \gtrsim \Delta \gg kT$. This means that the implementation of base transport coherence in a ballistic HBT requires at least liquid N_2 temperatures. Moreover, the ballistic coherent operation is limited to ultra-high frequencies. The concept cannot even be tested at lower frequencies, because the first resonance in U appears only at $f_\pi \simeq \pi f_T = 1/2\tau_B$, which must evidently be higher than the collision rate $1/\tau_{coll}$ governing the momentum relaxation of ballistic electrons at energy Δ.

Another idea for achieving base transport coherence (which is not limited to ultra-high frequencies and also works at room temperature) can be traced back to the famous paper by Shockley [12]. In that paper he introduced the concept of transit-time diodes and suggested that the delay in minority-carrier transit across a transistor base can lead to an active device at extended frequencies. A necessary condition for this to occur is that the directed transport across the base be much faster than the diffusive transport, that tends to wash out a modulated structure of the injected distribution. Shockley suggested [12] that this condition can be met in a minority-carrier delay diode with a variable doping in the base. However, because of a limited range of the potential variation available with an exponentially graded doping, the feasibility of this approach is marginal and it has never been realized. Recently, Shockley's argument was reconsidered [13] in the context of HBT with a graded alloy base composition. It was shown that coherence of the base transport in such devices is feasible and may lead to useful applications. In addition to transistors with a linearly graded base, we discussed a new a structure in which the base bandgap narrows down toward the collector in N discontinuous steps of minimum height larger than $\hbar\omega_{opt}$. In this structure the minority transport occurs by a *strongly accelerated diffusive process*, adequate for achieving the transit time resonance both at ultra-high and conventional frequencies. Our discussion below is based on Ref. [13].

Step Base Transistor

Consider a HBT whose base bandgap narrows down toward the collector in a stepwise fashion, Fig. 1. Let the band edge for minority carriers in the base consist of N steps W_j which are high enough, $\Delta_j \gg kT$, that carriers are effectively forbidden to return once they have fallen off a particular step. Moreover, let us assume that all the memory of the previous journey is lost in each step. The latter condition is reasonable if the steps are higher than the threshold for a rapid inelastic process, e.g., optical phonon emission: $\Delta_j \gtrsim \hbar\omega_{opt}$. With this condition fulfilled, we can treat the transport in each step individually and characterize it by a step transport factor α_j,

$$\alpha_j(\omega) = \frac{1}{\cosh\left[(2i\varphi_j)^{1/2}\right]}, \qquad (2a)$$

where $\varphi_j \equiv \omega\tau_j$ is the phase acquired on step j and $\tau_j = W_j^2/2D$ is the step propagation time by diffusion. We assume that the steps are narrow enough that $\varphi_j \ll 1$. In this case, to within quadratic in φ_j terms, Eq. (2a) reduces to the form

$$\alpha_j(\omega) \approx e^{-\varphi_j^2/3}\, e^{-i\varphi_j} . \qquad (2b)$$

We see that at small values of φ_j, the magnitude $|\alpha_j|$ deviates from unity *quadratically* in φ_j. This observation enables us to construct a coherent α_B with a large phase and little decay in magnitude. Indeed, the total base transport factor is the product of α_j's:

$$\alpha_B(\varphi) = \prod_{j=1}^{N} \alpha_j \approx e^{-\varphi^2/3N}\, e^{-i\varphi} , \qquad (3)$$

where $\varphi = \sum \varphi_j$ is the overall phase acquired in the base transport; we have assumed for simplicity that all steps are equal and $\varphi = N\varphi_j$.

The intrinsic current gain $\beta_B = \alpha_B(1-\alpha_B)^{-1}$ will have peaks $|\beta_B| > 1$ at $\varphi = 2\pi m$, provided

$$|\alpha_B(2\pi m)| > 0.5 . \qquad (4)$$

The first peak occurs at frequency $f = 2\pi f_T$. According to Eq. (3), this requires $N > N_{2\pi} = (2\pi)^2/3\ln 2 \approx 19$.

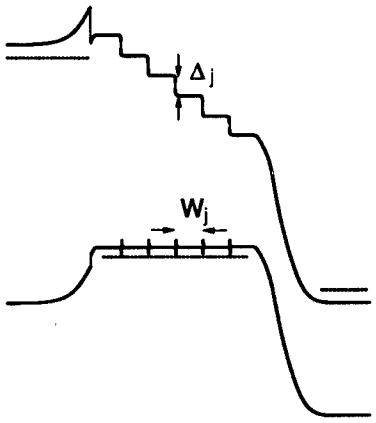


Figure 1. Schematic band diagram of a step-base heterostructure bipolar transistor.

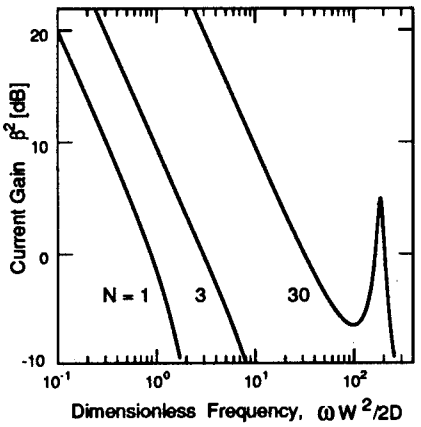


Figure 2. Common emitter current gain β of HBT's with the same width W and different number N of steps in the base.

Physically, the described effect originates from an enhancement of the *forward* diffusion transport. The essential condition is carrier thermalization at every step, which provides the independence of α_j's and restricts particles from returning to a preceding step. The resultant "stepped up" diffusion is a rather peculiar, at first glance even counter-intuitive, process. Thus, in a *static* regime, the minority concentration in the base is a *periodic* (rather than decreasing) function of the distance. Indeed, the steady-state current in each step is

$$J = \frac{e[n(0_j) - n(W_j)]}{D\, W_j} \approx \frac{N\, e\, n(0)}{D\, W} , \qquad (5)$$

where $W = NW_j$ is the total base thickness and $n(0_j)$ is the concentration at the beginning of step j. From the current continuity in the absence of recombination, we have $n(0_j) = n(0)$, independent of j. Equation (5) shows that the diffusion flux is enhanced by a factor of N and so is the effective diffusion velocity, which becomes $2D/W_j$. If we increase the number of steps keeping W constant, the conventional cutoff frequency f_T will increase in proportion to N, cf. Fig. 2. Clearly, the recombination limited static current gain will also scale in proportion to N. This effect has in fact been observed by Ohishi et al. [14] in a InGaAsP/InP HBT with a double-layer ($N=2$) base.

For high enough N, we would see a peaked structure in the current gain $|h_{21}|$ and for $N > N_{2\pi}$ the peak value is greater than unity, cf. the curve in Fig. 2 for $N = 30$. However, this situation is difficult to realize in practice, because each step must be high enough ($\Delta_j \gtrsim \hbar\omega_{opt}$) to ensure the no-return condition, while the overall potential drop is limited by the bandgap difference.

Power gain

Conditions for the transistor oscillation activity are more relaxed. Consider the intrinsic case first. An extended-frequency peak in the unilateral power gain U appears when the real part of the common-emitter output impedance $r_{22} \equiv \mathcal{R}e(z_{22}^e)$ changes sign [5,9]. In an intrinsic transistor we have $r_{22} = R_\varphi + R_E$, where

$$R_\varphi = \frac{\cos(\varphi) - \cos(\varphi + \theta)}{\omega C_C \theta} \; |\alpha_B| \; \approx \frac{|\alpha_B| \; \sin(\varphi+\theta')}{\omega C_C} \; , \qquad (6)$$

R_E is the emitter resistance, and $\theta \equiv \omega\tau_C = 2\theta'$ is the collector transit angle. The phase φ includes the (usually small) delay $\varphi_E = \omega R_E C_E$, due to emitter capacitance C_E. The approximate relation in the right-hand side of Eq. (6) corresponds to $\theta \lesssim 1$. In this case a power gain peak occurs at $\varphi \approx \pi$, provided $|\alpha_B| > \omega C_C R_E$.

The latter condition is relatively easy to accommodate. However, as pointed out by Tiwari [6], considerations of the phaseshift effects on unilateral gain are rather meaningless without including extrinsic resistances and capacitances. A careful analysis [9] shows that when these "parasitics" are included, the condition for a peak in U near $\varphi = \pi$ is

$$|\alpha_B(\pi)| \; > \; \omega\,\tau_X \; , \qquad (7)$$

where τ_X is a parasitics-limited time constant. For example, in a model which includes extrinsic emitter, base, and collector resistances (R_{Ex}, R_{Bx}, and R_{Cx}, respectively, cf. the equivalent circuit [9] of an abrupt-junction HBT) but neglects an extrinsic collector capacitance C_{Cx}, the expression for U is

$$U = \frac{|\alpha_B \alpha_C|^2}{4\omega^2 C_C^2 (R_B + R_{Bx})} \; \frac{1}{R_\varphi + R_X} \; , \qquad (8)$$

where $R_X \equiv R_E + R_{Ex} + R_{Cx} + R_{Cx}(R_E + R_{Ex})/(R_B + R_{Bx})$ and $\alpha_C \equiv (\sin\theta'/\theta')\,e^{-i\theta'}$ is the collector transport factor. The corresponding τ_X in Eq. (7) is $\tau_X = C_C R_X$. Besides the "low" frequency regime, where $U > 1$, the transistor will be active in the range of frequencies, where

$$|\alpha_B| \; \sin(\varphi + \theta') + \omega\tau_X \leq 0 \; . \qquad (9)$$

In this range $U < 0$ and hence one can obtain $U \gg 1$ by adding a series resistance. Obviously, inequality (9) can only be obtained if $|\alpha_B| > \omega\tau_X$; from Eq. (3) this means that the number of steps in the base must exceed the value $N_\pi = (\pi^2/3)\,|\ln(\omega\tau_X)|^{-1}$. For a transistor not overdamped by the parasitics, say $\omega\tau_X \lesssim 0.5$, we need $N \gtrsim 5$. The solid line in Fig. 3 displays the gain $|U|$, calculated for a model HBT with 5 steps in the base. The two peaks correspond to a vanishing denominator in (8); between the peaks U is negative.

It should be emphasized that we are not considering any "ballistic" boosts in the particle speed at the step edges. Such effects may in fact be beneficial; they are not expected to qualitatively modify the result (3). If we increase N at the expense of making the step height smaller, $\Delta_j < \hbar\omega_{opt}$, then the necessary energy relaxation will not occur at each step and hot electrons will diffuse backward as well as forward. In order to treat this regime quantitatively, a Boltzmann transport model for minority carriers has been developed [15]. It includes the effects of a finite optical phonon scattering time τ_{op} and finite exit velocity at each step. Preliminary results obtained in this model indicate that for $\Delta_j > \hbar\omega_{opt}$ the present simplified approach is valid, provided that τ_{op} is shorter than the elastic collision time of hot carriers exiting a step.

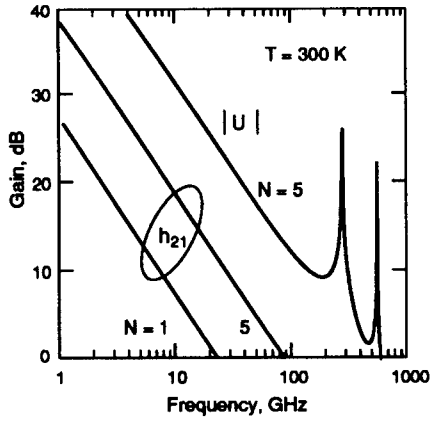


Figure 3. Common-emitter current gain $|h_{21}|$ and the unilateral gain $|U|$ of HBT's with the same base width $W = 0.25\ \mu m$, the same diffusivity $D = 50\ cm^2/s$, the same collector transit time $\tau_C = 1$ ps, and different number N of steps in the base.

Assumed parameters: $R_B = R_{Bx} = 50\ \Omega\cdot\mu m$, $R_E = 10\ \Omega\cdot\mu m$, $R_{Ex} = R_{Cx} = 20\ \Omega\cdot\mu m$, $C_C = 0.25$ fF/μm, and $C_C = 10$ fF/μm.

Comparison with Linear Grading

In the limit of $N \to \infty$ and $\Delta_j, W_j \to 0$ the base structure becomes equivalent to that in a linear-graded HBT [16-18] with $\nabla E_G = \lim \Delta_j / W_j$. In such a device, the phase of α_B is acquired with the drift velocity $v \equiv -\mu \nabla E_G / e$, while the magnitude $|\alpha_B|$ is attenuated due to spreading by diffusion. An equation analogous to (3) can be derived by solving the continuity equation $\nabla \cdot J = e(\partial n/\partial t)$ for the minority current taken in the drift-diffusion form $J = n\mu\nabla E_G + eD\nabla n$. Using the phasor notation

$$n(x, t) \equiv \bar{n}(x) + \hat{n}(x)\, e^{i\omega t},$$

the continuity equation for the dynamic component $\hat{n}$ reduces to

$$i\omega\hat{n} = -v\,\hat{n}' + D\,\hat{n}''. \tag{10}$$

Normalizing the coordinate x to the base width W, we find that the solution of Eq. (10), satisfying $\hat{n}(1) = 0$, is of the form

$$\hat{n} = A\, e^{rx} \sinh[\lambda(x-1)], \tag{11}$$

where $\lambda \equiv \sqrt{r^2 + 2i\omega\tau_D}$ and r is the ratio of the characteristic diffusion time $\tau_D = W^2/2D$ to the drift time $\tau_B = W/v$, viz.

$$r \equiv \frac{\tau_D}{\tau_B} = \frac{Wv}{2D} \approx \frac{\Delta E_G}{2kT}, \tag{12}$$

where ΔE_G is the total bandgap variation. The last equation in the right-hand side of (12) results form Einstein's relation $eD = \mu kT$ (valid for not too high bandgap gradients).

From Eq. (11) we find the base transport factor $\alpha_B \equiv J(1)/J(0)$ in the form

$$\alpha_B(\omega) = \frac{\exp(r)}{\cosh(\lambda) + (1 + 2i\omega\tau_B/r)^{-1/2} \sinh(\lambda)}, \tag{13}$$

In the absence of a grading, $v \to 0$, Eq. (13) reduces to $\alpha_B = \cosh^{-1}[(2i\omega\tau_D)^{1/2}]$, which corresponds to Eq. (2a) extended to the entire base. For a large grading, $r \gg 1$, one has, asymptotically, $\lambda \approx r + i\omega\tau_B + (\omega\tau_B)^2/2r$ and Eq. (13) reduces to

$$\alpha_B(\varphi) = e^{-\varphi^2/2r}\, e^{-i\varphi}, \tag{14}$$

where $\varphi \equiv \omega\tau_B(1 - 1/2r)$. The minority carrier drift effects are qualitatively similar to those resulting from the enhanced forward diffusion. Parameter $2r$ in Eq. (14) plays the same role as $3N$ in Eq. (3).

[In the ballistic case, Eq. (1), the parameter equivalent to r is $r_B \equiv (\tau_B/\delta\tau)^2$.] For our example of 5 steps (Fig. 3) we need $\Delta E_G = 5\Delta_j \gtrsim 5\hbar\omega_{opt} \approx 180\,\text{meV}$ (in GaAs/AlGaAs). Precisely the same effect will be achieved with a graded-gap base with $r = 7.5$, requiring $\Delta E_G = 15\,kT \approx 380\,\text{meV}$ at room temperature. As means for achieving the extended frequency operation at $\varphi \gtrsim \pi$, both approaches appear equally feasible.

The step-base approach offers additional flexibility in the design of ultrahigh speed heterostructure transistors. Perhaps its most obvious practical application is to relieve the stringent trade-off between the requirements of low base resistance and short base propagation time. As discussed above, this can be accomplished even with a small number of steps, e.g., $N = 2$. For higher number of steps, $N \gtrsim 5$, one can obtain an active behavior of the transistor at an extended frequency $f \approx \pi f_{max}$, provided (cf. Eq. 7) the device at this frequency is not overdamped by extrinsic elements.

Conclusion

We have discussed coherent base propagation effects that can be obtained in stepwise graded and linearly graded bandgap base structures of heterostructure bipolar transistors. These effects open up one or several bands of frequencies above f_T, where the transistor is active. Physically, the active behavior of a coherent transistor results from the phaseshift between the collector current and voltage, acquired during the minority-carrier transit across the base.

It is instructive to consider the following question: why is it that the phaseshift, obtained in a coherent base transport, is advantageous compared to that arising in the saturated-velocity transport across the base-collector space-charge region? Mathematically, the answer is evident from a comparison of expressions for the base and collector transport factors: the disadvantage of a base-collector propagation results from the factor $\sin(\theta')/\theta'$ which diminishes the magnitude of the complex transport factor α_C at high frequencies. Physically, the difference is rooted in the nature of screening of a dynamic charge moving *inside* or *outside* a conductor. Carriers moving in the depleted layer induce currents of equal magnitude in both the base and the collector electrodes [19]. In contrast, minority carriers moving inside the base induce a base current only to the extent that their total charge varies in time.

Acknowledgements

This paper reviews the work carried out in collaboration with Anatoly Grinberg and Vera Gorfinkel, to whom I am deeply grateful for all the insights they provided. I am also grateful to Marc Dignam for useful discussions and to John Cowles for drawing my attention to Ref. [14].

References

1. G. T. Wright, "Transistor transit-time oscillator", *Electron. Lett.* **3**, pp. 234-235 (1967).
2. G. T. Wright, "Small-signal theory of the transistor transit-time oscillator (translator)", *Solid-St. Electron.* **22**, pp. 399-403 (1979).
3. Y. Druelle, Y. Crosnier, and G. Salmer "Large-signal behaviour and capability of the transistor transit-time oscillator", *IEE Proc.* **127**, Pt. I, pp. 72-77 (1980).
4. N. Dagli, W. Lee, S. Prasad, and C. G. Fonstad, "High-frequency characteristics of inverted-mode heterojunction bipolar transistors", *IEEE Electron Device Lett.* **EDL-8**, pp. 472-474 (1987).
5. N. Dagli, "Physical origin of the negative output resistance of heterojunction bipolar transistors", *IEEE Electron Device Lett.* **EDL-9**, pp. 113-115 (1988).
6. S. Tiwari, "Frequency dependence of the unilateral gain in bipolar transistors", *IEEE Electron Device Lett.* **EDL-10**, pp. 574-576 (1989).

7. N. Dagli, "A unipolar transistor with negative output resistance", *Solid-St. Electron.* **33**, pp. 831-836 (1990).

8. P. Roblin and S. C. Kang, "Unilateral power gain resonances and roll-off with frequency for the velocity-saturated MOSFET/MODFET wave equation", *IEEE Trans. Electron Devices* **ED-39**, pp. 1490-1495 (1992).

9. A. A. Grinberg and S. Luryi, "Coherent Transistor", *IEEE Transactions Electron Devices* **ED-40**, pp. 1512-1522 (1993).

10. A. A. Grinberg and S. Luryi, "Ballistic versus diffusive base transport in the high-frequency characteristics of bipolar transistors," *Appl. Phys. Lett.* **60** pp. 2770-2772 (1992).

11. H. Kroemer, "Heterostructure bipolar transistors and integrated circuits", *Proc. IEEE* **70**, pp. 13-25 (1982).

12. W. Shockley, "Negative resistance arising from transit time in semiconductor diodes", *Bell. Syst. Tech. J.* **33**, pp. 799-826 (1954).

13. S. Luryi, A. A. Grinberg, and V. B. Gorfinkel, "Heterostructure Bipolar Transistors with Enhanced Diffusion of Minority Carriers", *Appl. Phys. Letters* **63** (September 13, 1993).

14. T. Ohishi, Y. Abe, H. Sugimoto, K. Ohtsuka, and T. Matsui, "Ultra-high current gain InGaAsP/InP heterojunction bipolar transistor", *Electron. Lett.* **26**, pp. 392-393 (1990).

15. M. M. Dignam, A. A. Grinberg and S. Luryi, "Boltzmann equation studies of minority-carrier transport in the base of a heterojunction transistor", unpublished.

16. H. Kroemer, "Quasi-electric and quasi-magnetic fields in nonuniform semiconductors", *RCA Rev.* **18**, pp. 332-342 (1957).

17. D. L. Miller, P. M. Asbeck, R. J. Anderson, and F. H. Eisen, "(GaAl)As/GaAs heterojunction bipolar transistors with graded composition in the base", *Electron. Lett.* **19**, pp. 367-368 (1983).

18. J. R. Hayes, F. Capasso, A. C. Gossard, R. J. Malik, and W. Wiegmann, "Bipolar transistor with graded band-gap base", *Electron. Lett.* **19**, pp. 410-411 (1983).

19. W. Shockley, "Currents to conductors induced by a moving point charge", *J. Appl. Phys.* **9**, pp. 635-636 (1938).

Novel Metal/2-DEG Junction Transistors

W.C.B. Peatman, H. Park, B. Gelmont, M. Shur
Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903

P. Maki, E.R. Brown
MIT Lincoln Laboratory, Lexington, MA 02173

and M.J. Rooks
National Nanofabrication Facility, Cornell University, Ithaca, NY 14853

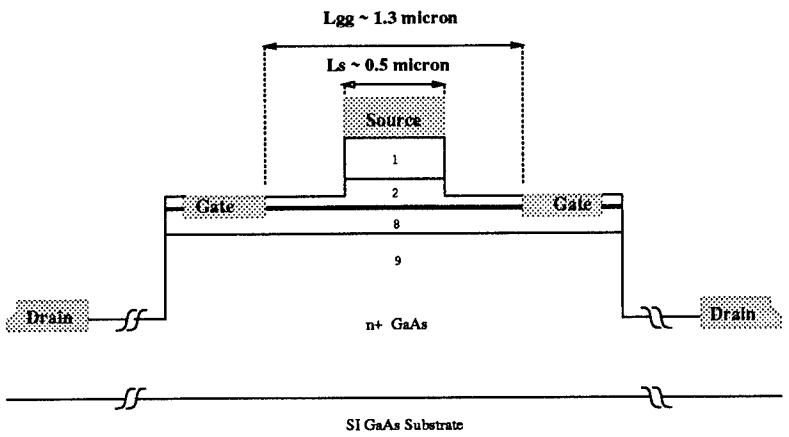
ABSTRACT

We present here the research on new metal/2-dimensional electron gas (2-DEG) Schottky contacts. This new high speed contact has unique characteristics which are particularly promising for applications in the fields of millimeter wave electronics and high speed, low power integrated circuits. We describe here two new transistors which utilize a side-gate formed by plating gate metal into a trench etched through the plane of the 2-DEG. The first transistor is the Schottky-gated resonant tunneling transistor (SG-RTT) which has demonstrated high transconductance and novel switching properties at room temperature. The second device is a novel 2-dimensional metal-semiconductor field effect transistor (2-D MESFET) which is particularly promising for low power, high speed integrated circuit applications. We also briefly discuss several applications of these new transistors.

I. INTRODUCTION: HETERODIMENSIONAL METAL/2-DEG SCHOTTKY CONTACTS

Heterodimensional Schottky contact devices represent a new class of devices in which the Schottky barrier is formed between the metal and a 2-d or 1-d electron gas [1]. These junctions are formed by electroplating gate metal into a trench which is etched through the plane of a compound semiconductor (e.g. AlGaAs/GaAs) heterostructure. The electrical characteristics of these devices are dominated by the Schottky junction with the electron gas. These devices have unique electronic properties which are fundamentally different from other metal/semiconductor junctions, including low specific capacitance, very high speed transport in the electron gas, quantization effects related to lower dimensional electronic systems, and high breakdown voltage. The first heterodimensional Schottky device investigated was the metal/2-DEG Schottky diode [2-4]. This device, formed using a pseudomorphic AlGaAs/InGaAs/GaAs modulation-doped heterostructure, exhibited excellent room temperature Schottky diode characteristics including ideality factor of 1.26, breakdown voltage of 15 V, series resistance of 18 Ω and junction capacitance ranging from about 30 fF (at zero-bias) to about 5fF (at -12 V). These devices were operated at 225 GHz as frequency triplers, delivering 500 μW output power with 0.5 percent efficiency [4]. We are optimizing these metal/2-DEG multiplier devices and expect to achieve much higher efficiencies at higher frequencies in the near future. In this paper, we discuss two new transistors in which the gate is formed using this metal/2-DEG junction technology developed at the University of Virginia. The first is the Schottky-gated resonant tunneling transistor (SG-RTT) which is discussed next.

0-7803-0894-8/93 $3.00 © 1993 IEEE



GaAs	2×10^{18} cm^{-3}	2000 Å
GaAs	1×10^{17} cm^{-3}	1000 Å
GaAs	4×10^{15} cm^{-3}	100 Å
$Al_{0.7}Ga_{0.3}As$	4×10^{15} cm^{-3}	22 Å
GaAs	4×10^{15} cm^{-3}	50 Å
$Al_{0.7}Ga_{0.3}As$	4×10^{15} cm^{-3}	22 Å
GaAs	4×10^{15} cm^{-3}	100Å
GaAs	1×10^{17}cm^{-3}	1000 Å
GaAs	2×10^{18} cm^{-3}	10,000 Å
SI GaAs Substrate		

Fig. 1. Schematic of the Schottky-Gated Resonant Tunneling Transistor (SG-RTT). The SG-RTT AlGaAs/GaAs double-barrier resonant tunneling heterostructure specifications are also shown (right).

II. THE SCHOTTKY-GATED RESONANT TUNNELING TRANSISTOR

The Schottky-gated Resonant Tunneling Transistor (SG-RTT) is fabricated by etching a trench through the thin layers of a double-barrier resonant tunneling structure and electroplating the gate metal onto the sidewall of the gate trench, as shown in Fig. 1. By limiting the contact area to the low doped layers near the tunneling layers, the gate characteristics are dominated by the junction between the metal and the accumulation layer which forms above the top barrier layer. The voltage applied to the gate modulates the quasi two-dimensional electron accumulation layer in a fashion analogous to the gate modulation of bulk semiconductor in conventional vertical field effect transistors [5]. The advantages of using the double barrier resonant tunneling structure include more efficient modulation of the 2-DEG and ballistic transport across the depletion layer. Furthermore, the presence of the negative differential resistance region in the SG-RTT naturally leads to applications in high speed, highly functional digital logic circuits and high speed switching applications.

Recently, we fabricated SG-RTT devices using the $Al_{.3}Ga_{.7}As$/GaAs double-barrier resonant tunneling structure, shown in Fig. 1 (right), which was grown by molecular beam epitaxy on a semi-insulating (SI) GaAs substrate [6]. The fabrication was performed using UV contact lithography for the drain, interconnect and isolation levels. Electron beam lithography was used to define the source and gate features. Chemically-assisted ion beam etching was used with the source metal as mask to form the source mesa, thereby removing the n^+ GaAs prior to plating the gate. A second, shallow CAIBE etch was performed through the gate PMMA to expose the sidewall of the low-doped double-barrier structure, and Pt/Au metals were pulse-electroplated into the trench. Interconnect metals were evaporated using the lift-off technique and a final wet etch was performed to the semi-insulating substrate, underetching the interconnects leads, thereby isolating the source and gate contact pads from the drain ohmic contacts. A scanning electron micrograph of the SG-RTT is shown in Fig. 2. The final device dimensions on the mesa (see Fig. 1) were $W = 16\ \mu m$, $L_{gg} = 1.3\ \mu m$, $L_s \approx 0.5\ \mu m$ and $L_{dd} = 6\ \mu m$.

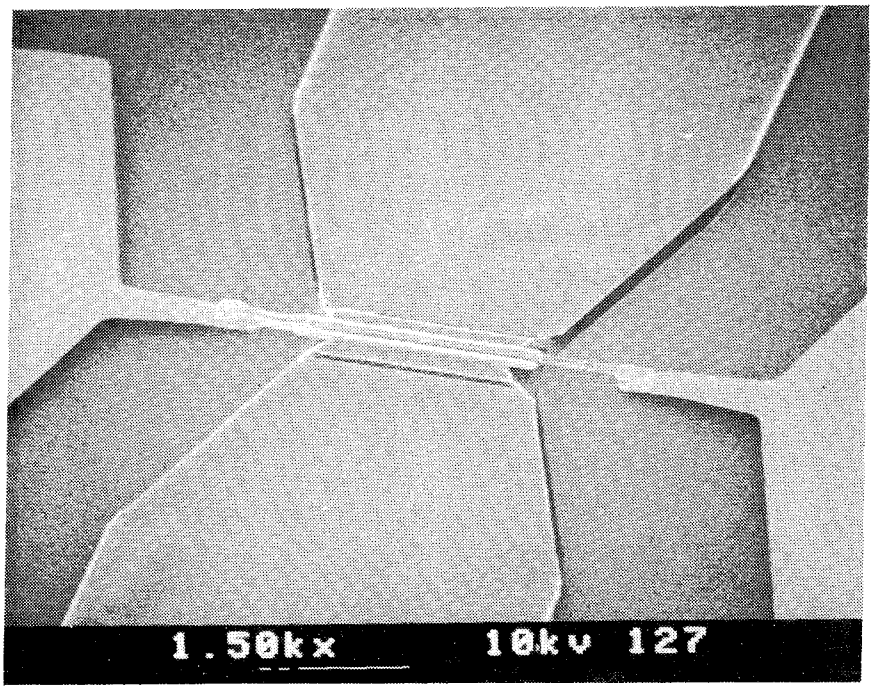


Fig. 2. Scanning electron micrograph of the SG-RTT. The mesa (center) is shown with drain pad (top and bottom), gate pad (left) and source pad (right) contacts.

Room temperature I-V characteristics of the SG-RTT are shown in Fig. 3. For these measurements, an HP4145 semiconductor parameter analyzer was used and the gate compliance was 5 μA (left) and 10 μA (right). The gate voltage was varied from 0.0 to 1.0 V in 0.25 V steps (left) and from -1.0 V to 1.0 V in 0.5 V steps (right). The specific extrinsic (small signal) transconductance, g_m, had a broad peak below the peak voltage, exceeding 100 mS/mm near V_{ds} = 1 V. Other devices had g_m in excess of 200 mS/mm [6]. These g_m values are the highest reported to date in RTT devices having Schottky gate modulation. The peak current density was 2.7×10^4 A/cm^2 at zero-gate bias and higher under forward gate-bias. The gate current was intentionally limited to small values to illustrate the "current gain", defined as $\Delta I_d/\Delta I_g$ where ΔI_g is the gate compliance. This current gain is of order 100 in these devices. We also note that the peak voltage shifted to higher V_{ds} values at high gate bias. We investigated the current switching near the peak voltage (between peak and valley voltage) and measured current changes of up to 1.25 mA (78 mA/mm) over only 1 mV change in gate voltage [6]. This switching characteristic of the SG-RTT may be useful for high speed square-wave generation where the dc biased transistor amplifies a (e.g. sinusoidal) input signal through a load resistance.

Although these prototype devices were not optimized for high frequency operation, it is interesting to estimate the f_t of these devices. We model C_{gs} similar to capacitance of the metal/2-DEG diode [1]. Assuming an accumulation layer carrier density of 1×10^{12} cm^{-2}, we estimate C_{gs} to be roughly 100 fF/mm. The gate-drain capacitance, C_{gd}, is modeled as a parallel-plate capacitance of roughly 500 fF/mm for a gate length of about 0.5 μm. Thus, we crudely estimate $f_t \approx g_m/2\pi(C_{gs} + C_{gd}) \approx 25$ GHz in in this prototype device. By reducing the gate length to 0.1 μm, by reducing the gate-to-gate length L_{gg} to about 0.5 μm and by using higher peak current density material, f_t values of order 500 GHz should be possible.

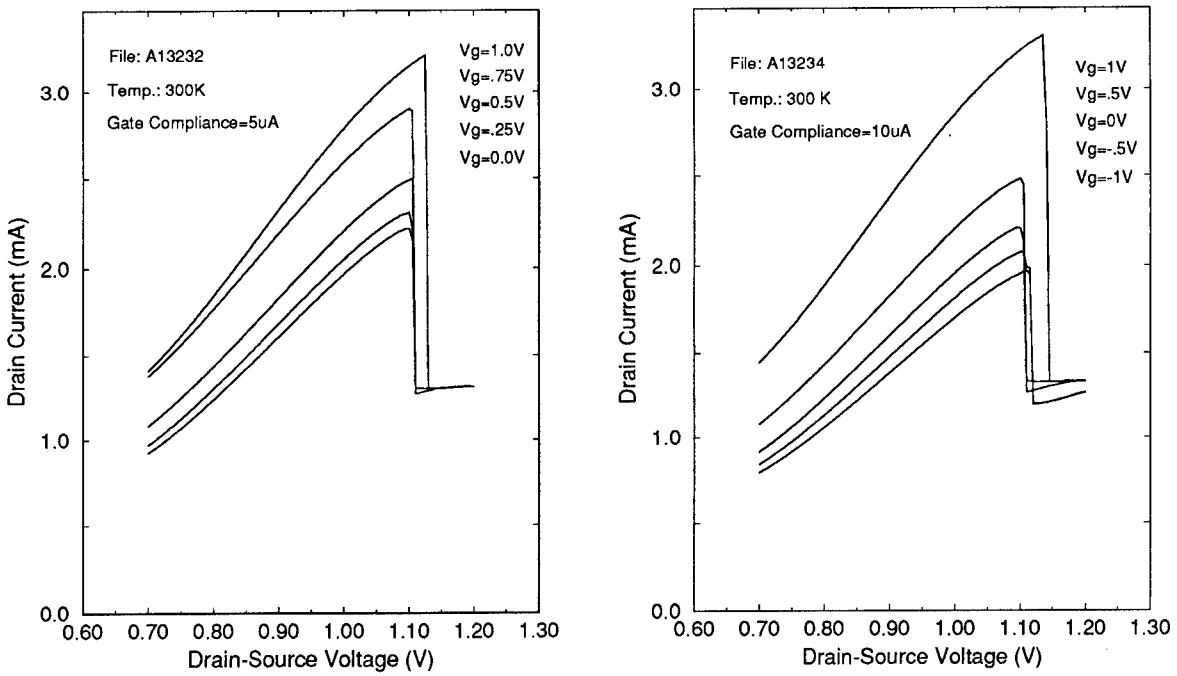


Fig. 3. Room temperature $I_d - V_{ds}$ characteristics of the SG-RTT for different gate voltage ranges and gate compliance.

III. NOVEL 2-D METAL-SEMICONDUCTOR FIELD EFFECT TRANSISTOR

A schematic of the 2D-MESFET is shown in Fig. 4. The $Al_{.27}Ga_{.73}As$/GaAs modulation-doped structure was grown by molecular beam epitaxy on a semi-insulating GaAs substrate. The 2-DEG forms at the top of a 2 μm undoped GaAs layer and is separated from a 300 Å n^+ AlGaAs layer ($n_s = 1.5 \times 10^{18}$ cm^{-3}) supply layer by a 30 Å undoped $Al_{.27}Ga_{.73}As$ spacer layer. A 200 Å GaAs cap layer of doping density $n_s = 1 \times 10^{17}$ cm^{-3} serves as the top gate contact layer. The 2-DEG sheet concentration and mobility at room temperature were 7×10^{11} cm^{-2} and 7700 cm^2/V-s, respectively. The source and drain Ni/Ge/Au metalization was performed using standard evaporation and lift-off techniques. The side gate was formed by etching a trench through the 2-DEG layer and electroplating Pt/Au into the trench. The next step involved Cr/Au metalization to form pads for device probing. In this metalization step, a top (MODFET-like) gate is also formed and serves to modulate the 2-DEG concentration. The final fabrication step is an isolation etch to the SI substrate. This etch forms "air-bridge" leads between gate Schottky metal and the gate contact pad. The nominal 2-DEG channel dimensions (see Fig. 4) were $W_0 = 3$ μm and $L_{sd} = 20$ μm. The gate length L_g was 10 μm and the gate-to-gate distance was about 1 μm.

The room temperature I-V characteristics were measured using an HP4145B semiconductor parameter analyzer. First, the top gate was biased over the range -0.5 V to 0.3 V in 0.1 V steps while the side gate was held constant at 0 V. These "MODFET-like" characteristics are shown in Fig. 5 (left). Pinch-off occurred at $V_{g,side} \approx -0.5$ V. Next, the top gate voltage was held constant at $V_{g,top} = 0.3$ V and the side gate voltage was varied from 0 V to -7.0 V, resulting in the characteristics shown in Fig. 5 (right). When the top gate was biased at 0 V, the side gate achieved pinch-off at about -7 V, due to lower electron sheet density beneath the top gate. The

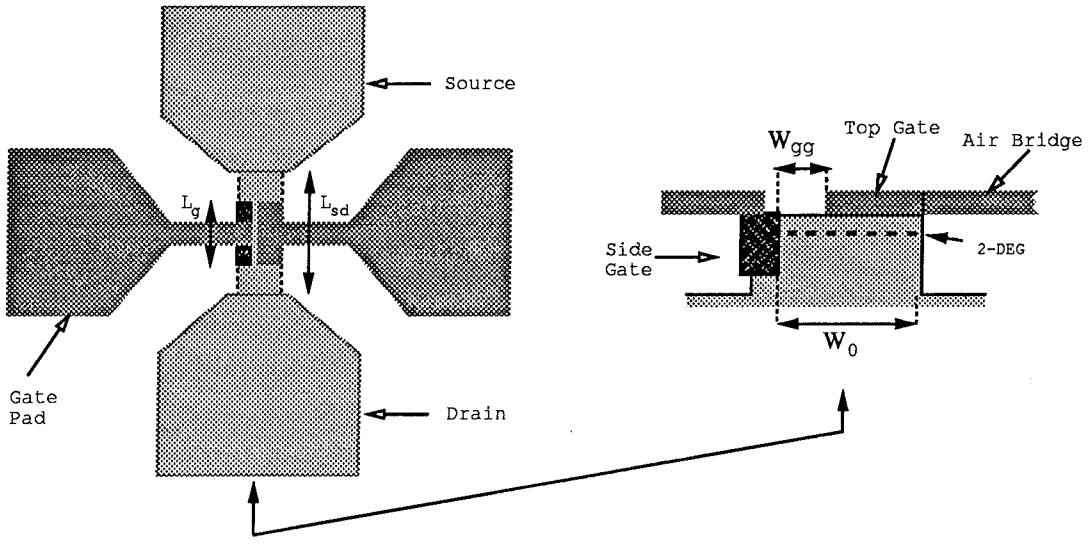


Fig. 4. Schematic of 2D-MESFET with novel metal/2-DEG junction side gate. Cross-sectional view (right) illustrates how the side-gate modulates the channel width.

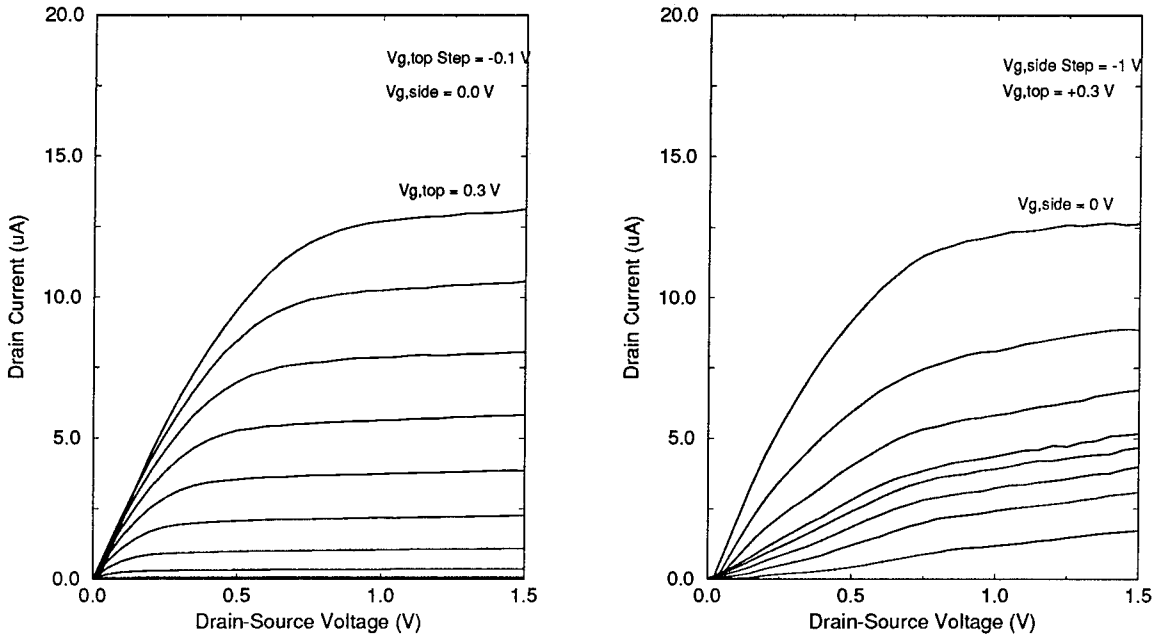


Fig. 5. 2D-MESFET $I_d - V_{ds}$ characteristics at 300 K. (Left) $V_{g,top}$ ranges from 0.3 V to -0.5 V (-0.1 V step) with side gate bias, $V_{g,side} = 0$ V. (Right) $V_{g,side}$ ranges from 0 to -7 V (-1 V step) with top gate bias, $V_{g,top} = 0.3$ V.

current modulation in Fig. 5 (right) is in good agreement with our charge control model for the new side gate contact. Finally, we compared the subthreshold slope for the case when the side gate was allowed to float and the case when the side gate was biased simultaneously with the top gate. We found that the subthreshold slope was larger in the latter case, indicating that the side gate has a better pinch-off characteristic. This effect will be much more dramatic in scaled devices and should permit operation of the 2-D MESFET at much lower currents, resulting in a large reduction in the power consumption in integrated circuits.

IV. SUMMARY

We have reported results for two new field effect transistors based on the heterodimensional metal/2-DEG junction gate. This junction is formed using an electroplate technique developed at the University of Virginia. In the Schottky-gated Resonant Tunneling Transistor (SG-RTT), the gate modulates the area of the quasi-two-dimensional electron accumulation layer above the top barrier in a AlGaAs/GaAs double-barrier resonant tunneling heterostructure. In this fashion, the gate modulates the tunneling current. The SG-RTT exhibited high transconductance in excess of 100 mS/mm at room temperature and also exhibited a positive shift in the peak voltage at high forward gate bias. This latter characteristic may be useful for high speed switching applications. We also presented first results of a novel 2-dimensional metal-semiconductor field effect transistor (2-D MESFET) in which the 2-DEG channel width is modulated by a metal/2-DEG side gate. We observed better pinch-off by the side gate compared to that for a MODFET-like gate in the same device. With appropriate scaling, this device should operate at much lower current than conventional transistors, leading to much lower power consumption in integrated circuits.

ACKNOWLEDGEMENTS

This research was partially supported by the Office of Naval Research and by the U.S. Air Force. The authors thank S. Svensson of Martin Marietta Labs for the Van der Pauw data and MBE material used in the 2-D MESFET.

REFERENCES

[1] B. Gelmont, W. Peatman, and M. Shur, "Heterodimensional Schottky Metal-Two delivering 500 μW output power with 0.5 percent efficiency. We are Dimensional Electron Gas Interfaces," *20th Conf. Phys. and Chem. of Semic. Interfaces,* Williamsburg, VA, (January 1993), to be published, *J. Vac. Sci. Tech.* B, vol 11, no. 4, July/August, 1993.

[2] W.C.B. Peatman, T.W. Crowe, and M. Shur, "Design and Fabrication of Heterostructure Varactor Diodes for Millimeter and Submillimeter Wave Multiplier Applications," *Proc. 1991 IEEE/Cornell Conf. Advanced Concepts in High Speed Semic. Dev. and Circuits*, Cornell University, Ithaca, NY, pp. 49-57, Aug. 1991.

[3] W.C.B. Peatman, T.W. Crowe, and M. Shur, "A Novel Schottky/2-DEG Diode for Millimeter and Submillimeter Wave Multiplier Applications," *IEEE Electron Device Lett.*, vol. 13, pp. 11-13, Jan. 1992.

[4] W.C.B. Peatman, T.W. Crowe, M. Shur, and B. Gelmont, "A Schottky/2-DEG Varactor Diode for Millimeter and Submillimeter Wave Multiplier Applications," *Proc. 3rd. Int'l. Conf. Space THz Tech.*, University of Michigan, Ann Arbor, MI, pp. 93-109, March 1992.

[5] Y.H. Won, K. Yamasaki, T. Daniels-Rice, P.J. Tasker, W.J. Schaff, and L.F. Eastman, "A High Voltage-Gain GaAs Vertical Field-Effect Transistor with an InGaAs/GaAs Planar-Doped Barrier Launcher," *IEEE Elect. Dev. Lett.*, vol. 11, pp. 376-378, Sept. 1990.

[6] W.C.B. Peatman, E.R. Brown, M.J. Rooks, P. Maki and M. Shur, "A Vertical Schottky-gated Resonant Tunneling Transistor with High Transconductance at Room Temperature," Submitted to *IEEE Elect. Dev. Lett.*, June 1993.

Performance and Reliability Characteristics of GaAs Planar Doped Barrier Detector Diodes Using Carbon-Doped Acceptor Spikes Grown By Molecular Beam Epitaxy

Roger J. Malik[1], Yogi Anand[2], Miro Micovic[1],
Michael Geva[3], and Robert W. Ryan[1]

[1]AT&T Bell Laboratories, Murray Hill, NJ 07974

[2]M/A-COM, Inc., Burlington, MA 01803

[3]AT&T Bell Laboratories Solid State Technology Center,
Breinegsville, PA 18031

ABSTRACT

Significant improvements in the reliability as well as yield and reproduceability of GaAs planar doped barrier (PDB) microwave detector diodes have been achieved by substituting carbon-doped acceptor regions in place of conventional beryllium acceptor dopant. The superior characteristics of the C-doped PDB diodes are thought to be related to the hyperabrupt and stable carbon acceptor doping spikes (10-60A) obtained in the n^+-i-p^+-i-n^+ doping profile grown by molecular beam epitaxy (MBE). Mesa geometry PDB diodes (10-20 micron diameter) in micropill packages were RF tested at 10 and 35GHz. Excellent tangential sensitivity up to -58dBm was measured with a detector video impedance of 2-50kohms depending upon barrier height. The electrostatic discharge (ESD) failure threshold voltages were found to be much higher for the PDB diodes (3500V) in comparison to Schottky detector diodes (300V).

I. INTRODUCTION

Planar doped barrier (PDB) diodes[1-3] are promising new devices for replacement of point contact and Schottky diodes for microwave detector and mixer applications. PDB diodes have demonstrated excellent performance as detectors up to 94GHz[4] and as subharmonic mixers at 140GHz[5]. The key advantages of the PDB diode are that the barrier height can be readily designed for particular applications (i.e., for low drive mixers or zero-bias detectors) and the asymmetry in the current rectification can also be varied. In addition, PDB diodes have been reported to have lower 1/f noise and higher resistance to electrostatic discharge (ESD) failure than Schottky diodes[6]. Recent work has focussed on the reliability and manufacturability of these devices[7]. This paper discusses the performance and reliability of GaAs PDB diodes using carbon-doped acceptor regions.

II. PDB DIODE STRUCTURE

The doping profile and the conduction band energy diagram for a PDB diode is shown in Figure 1(a) and 1(b), respectively. The PDB structure consists of an n^+-i-p^+-i-n^+ doping profile in which an extremely thin (1-10nm) acceptor region is placed within an intrinsic (nominally undoped) region

0-7803-0894-8/93 $3.00

bounded by two donor contact regions. The typical doping densities for the n^+ and p^+ regions are 1-$5x10^{18}cm^{-3}$. The thin acceptor spike is fully ionized forming a negative space charge region which is compensated by positive space charge induced at the edges of the n^+ contact regions. This results in a triangular conduction band energy profile as shown in Figure 1(b) whose height and shape are determined by the sheet acceptor concentration and its positioning within the intrinsic region. A highly rectifying PDB diode is made by placing the acceptor spike closer to one of the n^+ contact regions. The barrier height, Φ_B, and capacitance, C, of the PDB diode are given approximately by the following expressions:

$$\Phi_B = \frac{qN_AX_A}{E}\frac{L_1 L_2}{(L_1+L_2)} \qquad (1) \qquad\qquad C = \frac{E\,a}{(L_1+L_2)} \qquad (2)$$

where N_AX_A is the acceptor region doping-thickness product, L_1 and L_2 are the intrinsic region widths, a is the diode area, q is the electronic charge, and E is the dielectric permittivity. Molecular beam epitaxy (MBE) is used to grow these structures at slow growth rates (approximately 2.8 Angstroms/sec) in order to achieve precise control over the thicknesses and doping levels in the PDB diode.

A schematic cross-section of a PDB diode is illustrated in Figure 2. Shallow mesas are fabricated using wet chemical etching with diameters of 10-25 microns. AuGeNi-Au contacts are formed to the top anode contact and the sidewalls are passivated with silicon nitride. The wafer is thinned to 100 microns and a backside contact of AuGeNi-Au is formed. The wafer is diced and the individual diodes are thermo-compression bonded in ceramic micropill packages.

III. CHARACTERIZATION OF C AND Be ACCEPTOR DOPANTS

The first GaAs PDB diodes were grown using beryllium and silicon for the acceptor and donor dopants, respectively. It was found experimentally that it was very difficult to reproduce the barrier heights and I-V characteristics of PDB diodes using Be. Furthermore the diodes were found to degrade rapidly under high temperature reverse burn-in. Similar phenomena has also been seen in AlGaAs/GaAs heterojunction bipolar transistors (HBT) and has been attributed to Be diffusion under current stressing causing p-n junction shifts. Carbon evaporated from a heated graphite filament[8-9] has been used as an acceptor dopant in the MBE growth of GaAs. It has been shown that the C diffusion coefficient in GaAs is much lower than Be presumably due to the strong tetrahedral Ga-C bonds (diamond-like) in the crystal.

A comparison of the diffusion coefficients of C and Be during growth was analyzed using a test sample. Two delta-doped spikes of both C and Be (deposited during growth interruption) were placed in the middle of 500A thick $Al_{0.3}Ga_{0.7}As$ layers grown at 580C. A piece of this sample was cleaved off and underwent a rapid thermal anneal (RTA) of 900C for 15 seconds. A SIMS analysis of the samples was done using an O^+ ion beam at 3keV energy to obtain high resolution depth profiles. The resultant impurity profiles for C and Be are shown in Figure 3(a) and 3(b), respectively. It is seen that although the nominal sheet acceptor concentrations are the same, the peak Be concentration is approximately 5 times lower than for C. Also the leading edges of the C profiles are parallel with the Al signal. Together this implies that the C profile is atomically abrupt (as is known for the Al profile) and that significant diffusion of Be has occurred during MBE growth. After the RTA, there is no indication of C diffusion. In contrast, after the RTA there is further additional diffusion of Be leading to a broader

profile. This data suggests that C is preferred over Be as the acceptor dopant in GaAs PDB diodes since hyperabrupt and stable doping spikes can be obtained.

A SIMS depth profile of a typical PDB diode is shown in Figure 4. The C-doping spike is 35A wide and the undoped region widths are 100A and 2000A resulting in a highly asymmetric rectifying diode. The finite width of the C-doping profile is due to SIMS instumental broadening associated with the sputtering rate and knock-on. Also to be noted is the large unintentional C contamination at the epilayer-substrate interface. This can be minimized by using ozone cleaning of the substrate prior to epitaxial growth.

The control of the C-doping concentration and barrier height was examined. A series of PDB diodes were grown using a fixed C-doping level and varying the growth time of the acceptor spike region. The measured log I-V characteristics are shown in Figure 5. The barrier height, ϕ_B, can be determined from the extrapolated saturation current at zero-bias, I_0. From the inset in Figure 5, it is seen that the calculated barrier heights are directly proportional to the growth time of the acceptor spike regions. This is expected from Equation (1) where the barrier height is proportional to the doping-thickness product of the acceptor spike region. From this data, it is estimated that the C-doping level is controlled to within +/- 2% which corresponds to +/- 1^oC control in the graphite filament temperature. This level of doping control is very important to obtain PDB diodes with reproduceable I-V characteristics.

IV. DC AND RF PERFORMANCE

Zero-bias PDB detector diode chips were mounted in standard ceramic pill packages (ODS-119) and tested at DC and for RF performance at X and Ka band frequencies. Table 1 shows the typical DC characteristics of these diodes which exhibit low forward turn-on voltages (0.06V @ 10uA) due to the low designed barrier height (ϕ_B=0.2V). The temperature dependence of the forward voltage (V_f @ 100uA) for zero-bias PDB and high barrier Si Schottky diodes is illustrated in Figure 6. The zero-bias PDB diodes are seen to be less sensitive to temperature variation than Si Schottky diodes. The detected output voltage as a function of input power level at 10 and 35GHz is given in Figure 7. Table 2 shows the video impedance, tangential sensitivity, and detected output voltage (@-10dBm) at 10 and 35GHz. The zero-bias PDB diodes display excellent tangential sensitivity up to -58dBm and a video impedance of 2-50kohms depending upon the barrier height. The video bandwidth was 2kHz to 2MHz and the amplifier noise figure was 2dB. No matching circuit or tuner was used to optimize the measurement.

Electrostatic and RF burnout are important parameters for microwave diodes in many applications. The very small active areas of PDB and Schottky diodes limit their power handling capability and make them vulnerable to failure caused by accidental electrostatic discharge (ESD). Zero-bias PDB and equivalent Si and GaAs Schottky diodes were subjected to electrostatic pulses using a manual ESD simulator, IMCS Model-700. The ESD pulser simulates electrostatic discharge pulses through the human body. The circuit consists of a 100pF capacitor in parallel with a 1000ohm resistor and discharges in less than 100ns. The diodes were also tested for RF burnout at 10GHz by gradually increasing the CW power to the diodes. The results of these tests are shown in Table 3. It is seen that the Schottky diodes can withstand 800-1500V in the forward direction but only 200-300V in the reverse direction. The zero-bias PDB diodes can withstand 3500V in both directions making them far more rugged against ESD failure compared to Schottky diodes. The CW burnout level for zero-bias PDB diodes is similar to high barrier Si and GaAs Schottky diodes and is limited to the maximum power dissipation of the diodes resulting in metal contact failure.

V. SUMMARY

Zero-bias GaAs planar doped barrier detector diodes with excellent DC and RF performance at 10 and 35GHz have been demonstrated. The previous problem of poor reliability and reproduceability of the PDB diode I-V characteristics was solved by substituting carbon for the conventional beryllium dopant in the thin acceptor spike regions grown by molecular beam epitaxy. SIMS analysis has shown that C-doping spikes are extremely abrupt and are highly stable against diffusion compared to Be. The threshold voltages for electrostatic discharge (ESD) failure of PDB diodes were found to be much higher (3500V) than for equivalent Schottky diodes (300V) offering improved reliability in many systems applications.

ACKNOWLEDGEMENTS

The authors would like to thank J. Park for device fabrication, B. Brown, W.D. Martin, J. Mullin, R. Nelson, and T. Bosia for DC and RF characterization, and J. Nagle, A. Torabi, and R. Sachs for growth of some MBE samples.

REFERENCES

1. R.J. Malik, K. Board, L.F. Eastman, C.E.C. Wood, T.R. AuCoin, and R.L. Ross, "Planar Doped Barriers in GaAs by Molecular Beam Epitaxy", *Electronic Lett.*, **16**, 836, (1980).

2. I. Dale, S. Neylon, A. Candle, and M.J. Kearney, "Planar Doped Barrier Mixer and Detector Diodes as an Alternative to Schottky Diodes for Both Microwave and Millimeter Wave Applications", *IEEE MTT-Symposium Digest*, **1**, 467, (1989).

3. Y. Anand, J. Hillson, A. Torabi, and J. East, "Millimeter Wave Planar Doped Barrier Detector Diodes", *Proc. Second Intl. Symposium on Space TeraHertz Technology*, 467, February (1991).

4. M.J. Kearney, A. Condie, and I. Dale, "GaAs Planar Doped Barrier Diodes for Millimetre-Wave Detector Applications", *Electronic Lett.*, **27**, 721 (1991).

5. U. Guttrich, K.M. Strohm, and F. Schaffler, "D-Band Subharmonic Mixer with Silicon Planar Doped Barrier Diodes", *IEEE Trans. Microwave Theory Tech.*, **39**, 366, (1991).

6. M.J. Kearney and I. Dale, "GaAs Planar Doped Barrier Diodes For Mixer and Detector Applications", *GEC J. Research*, **8**, 1, (1990).

7. Y. Anand and R.J. Malik, " Reliability Comparison of Beryllium versus Carbon for GaAs Planar Doped Barrier Diodes", *Proc. GaAs Reliability Workshop*, October (1992).

8. R.J. Malik, R. N. Nottenberg, E.F. Schubert, J.F. Walker, and R.W. Ryan, "Carbon doping in molecular beam epitaxy of GaAs from a heated graphite filament", *Appl. Phys. Lett.*, **53**, 2661, (1988).

9. R.J. Malik, J. Nagle, M. Micovic, R.W. Ryan, T. Harris, M. Geva, L.C. Hopkins, J. Vandenberg, R. Hull, R.F. Kopf, Y. Anand, and W.D. Braddock, "Properties and Applications of carbon-doped GaAs and $Al_xGa_{1-x}As$ layers grown by MBE with a pyrolytic graphite filament", *J. Crystal Growth*, **127**, 686, (1993).

DC CHARACTERISTICS OF ZERO BIAS PDB DIODES
(10 GHz-40 GHz PILL PACKAGE)

V_f @ 10 μA	0.06 volts
V_f @ 100 μA	0.14 volts
V_b @ 10 μA	1.0 volts
C_j (Calculated)	60 fF
R_s	15 ohms

Table 1

RF PERFORMANCE OF ZERO BIAS PDB DODES
(10 GHz-40 GHz PILL PACKAGE)

Frequency	10 GHz	35 GHz
Video Impedance R_v	2-50 kohms	2-50 kohms
Tangential Sensitivity T_{ss}	–50 to –58 dBm	–50 to –58 dBm
Voltage Output @-10 dBm	80 mV	65 mV

Table 2

ELECTROSTATIC AND CW BURNOUT AT 10 GHz

	Low Barrier Si-Schottky	Med. Barrier Si-Schottky	High Barrier Si-Schottky	GaAs Schottky	GaAs PDB (Zero Bias)
V_f @ 1 mA (volts)	0.4	0.55	0.7	0.72	0.2
V_b @ 10 μA (volts)	3	5	7	7	1
C_t (pF)	0.1	0.1	0.1	0.1	< 0.1
Electrostatic Pulses (3)					
Forward (volts)	800	1100	1500	900	3500
Reverse (volts)	300	300	300	200	3500
CW (watts)	0.2-0.4	0.3-0.5	0.4-0.6	0.4-0.6	0.4-0.6

Table 3

GaAs PDB DIODE

(a)

(b)

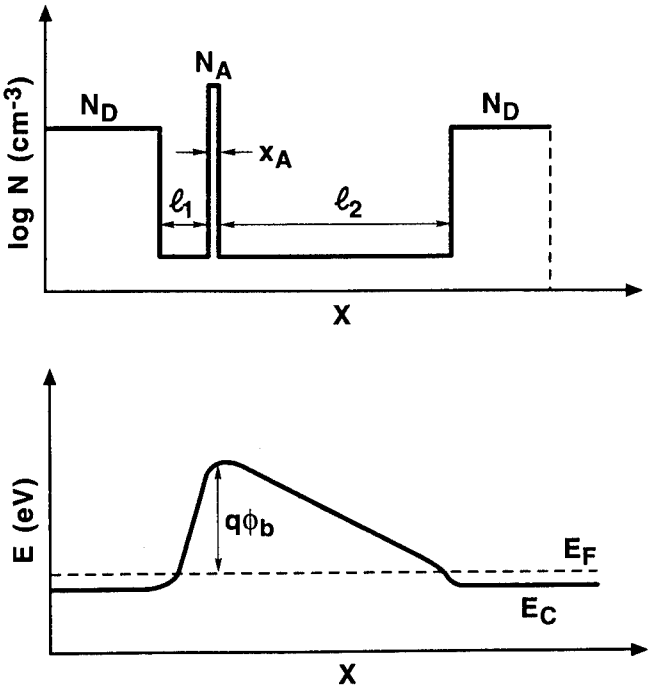


Figure 1

GaAs PDB DIODE

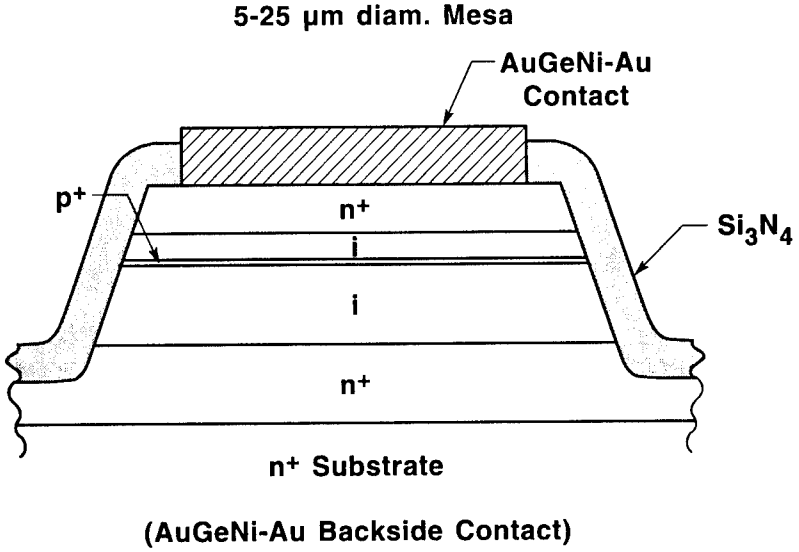
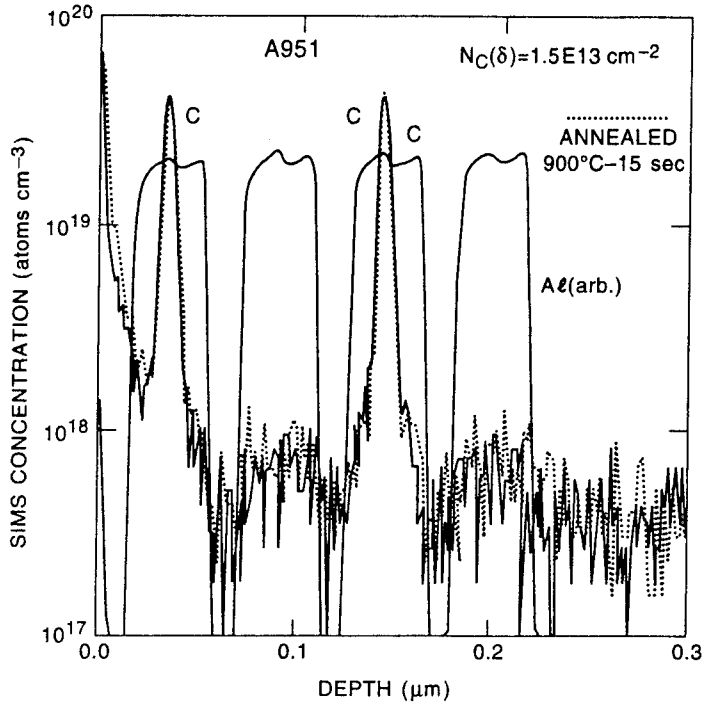


Figure 2

(a)


DIFFUSION OF C IN Aℓ0.3Ga0.7As
A951
NC(δ)=1.5E13 cm−2
C
C
C
ANNEALED
900°C–15 sec
Aℓ(arb.)
SIMS CONCENTRATION (atoms cm−3)
10^20
10^19
10^18
10^17
0.0
0.1
0.2
0.3
DEPTH (μm)

(b)

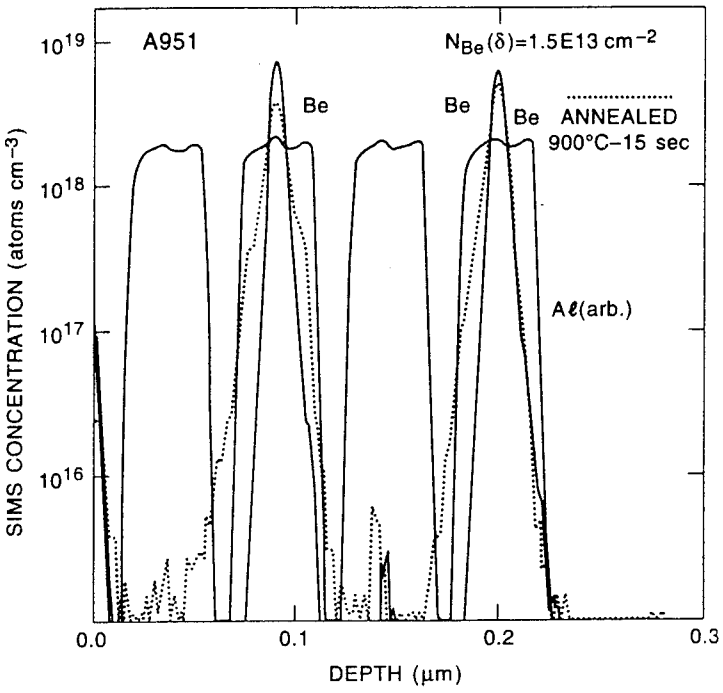

DIFFUSION OF Be IN Aℓ0.3Ga0.7As
A951
NBe(δ)=1.5E13 cm−2
Be
Be
Be
ANNEALED
900°C–15 sec
Aℓ(arb.)
SIMS CONCENTRATION (atoms cm−3)
10^19
10^18
10^17
10^16
0.0
0.1
0.2
0.3
DEPTH (μm)

Figure 3

SIMS – PLANAR DOPED BARRIER DIODE

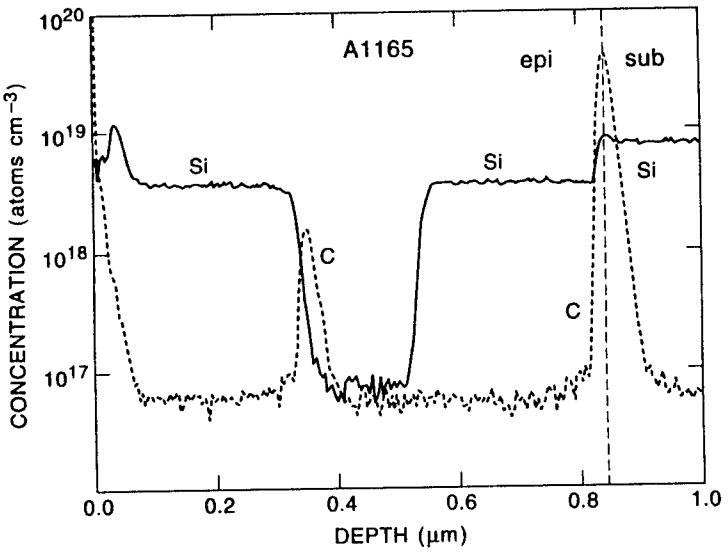
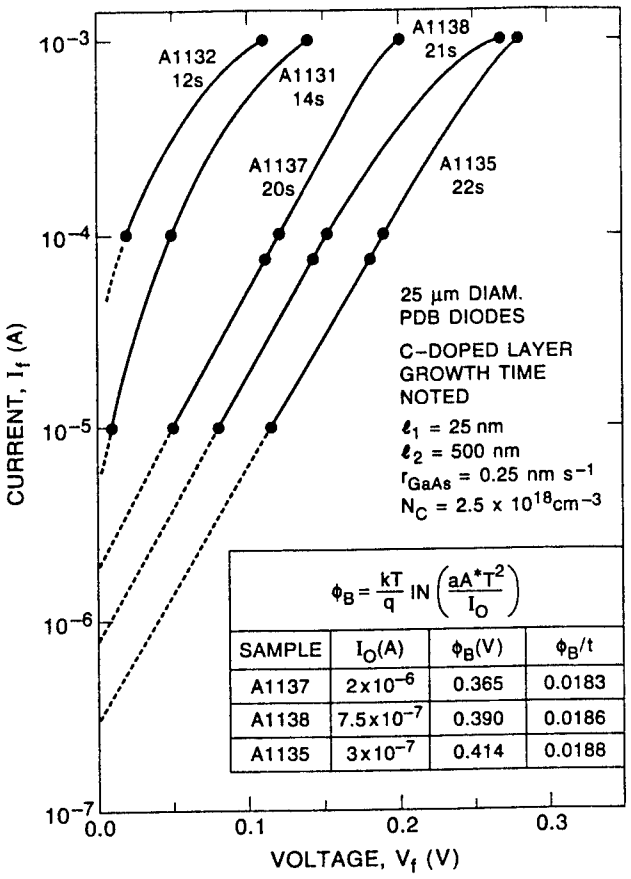


Figure 4

CARBON DOPED PDB DIODES



$$\phi_B = \frac{kT}{q} \text{IN}\left(\frac{aA^*T^2}{I_O}\right)$$

SAMPLE	I_O(A)	ϕ_B(V)	ϕ_B/t
A1137	2×10^{-6}	0.365	0.0183
A1138	7.5×10^{-7}	0.390	0.0186
A1135	3×10^{-7}	0.414	0.0188

Figure 5

TEMPERATURE DEPENDENCE OF GaAs PDB AND Si SCHOTTKY DIODE OUTPUT VOLTAGE

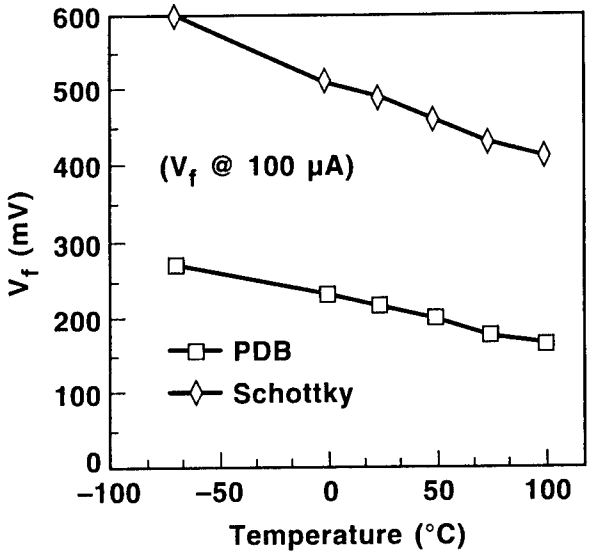


Figure 6

GaAs ZERO BIAS PDB DETECTOR DIODE Output Voltage vs RF Power

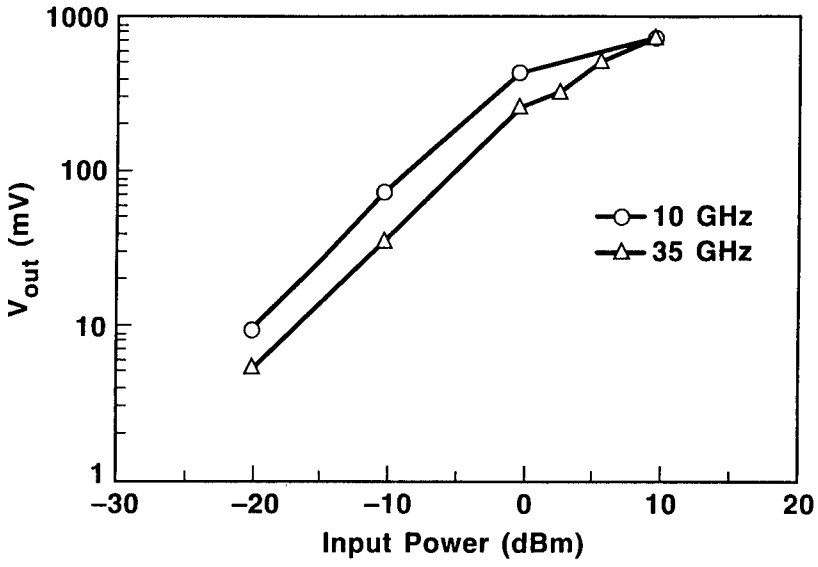


Figure 7

A NOVEL GaSb/AlSb/InAs HIGH EFFICIENCY RECTIFYING DIODE

Robert J. Ferro

The Aerospace Corporation
Mail Stop M4-991 P. O. Box 92957
Los Angeles, CA 90009-2957

Ronald R. Robson, William W. Hooper, Thomas C. Hasenberg,
Leslie D. Bailey, Scott M. Newell

Hughes Research Laboratories, Malibu CA 90265

ABSTRACT

We have fabricated a novel type of interband resonant tunnel diode by MBE growth of a superlattice, each period of which consists of layers of GaSb, AlSb and InAs. The diode is designed to conduct in forward bias by resonant interband tunneling, so that the entire stack appears to be a short circuit. Reverse bias current is impeded by band gap blocking. The reverse breakdown voltage varies with the number of periods, with each period capable of blocking about 0.4 V. A four period diode conducts 5000 A/cm^2 at less than 0.3 V of forward bias.

INTRODUCTION

It is well known that the current-voltage characteristic of tunnel diodes and backward diodes can exhibit greater curvature than that of thermal diodes such as PN junction, schottky barrier and PIN diodes [1, 2], and can thus give a lower turn-on voltage. However tunnel diodes cannot block more than about 1 V in the off state because of thermal injection current. We have fabricated superlattice interband resonant tunnel diodes designed to use resonant interband tunneling to maintain the large IV curvature and low turn-on voltage of a tunnel diode while using band gap blocking to provide moderate reverse breakdown voltages.

0-7803-0894-8/93 $3.00

The diode is made from repeated layers of InAs, AlSb and GaSb. This material system has been of interest recently because the valence band edge of GaSb is 0.14 eV above the conduction band edge of InAs, and a number of devices based on interband tunneling between InAs electrons and GaSb light holes have been made [3-8]. The energy band diagram for one of the diodes we fabricated is shown in Figure 1. Conduction occurs in the on state by tunneling of electrons from the GaSb valence band through the AlSb barrier into the InAs conduction band, followed by recombination with light holes in the next layer of GaSb. The widths of the quantum wells in the InAs conduction band and the GaSb valence band are designed so that the InAs electron and GaSb light hole quasi-stationary states line up with no voltage applied across the diode. This allows the diode to take advantage of resonant interband tunneling. Conduction in the off state is prevented by band gap blocking. A bias that tends to make InAs conduction band electrons tunnel across the AlSb to recombine with GaSb holes causes misalignment of the quasi- stationary states and lines up the InAs conduction band with the band gap of GaSb to prevent tunneling.

We imposed two conditions to obtain the widths of the two quantum wells in our initial design: 1) the sum of the energies of the quasi-stationary light hole state E_h and the electron state E_e must equal the 0.14 eV difference between the GaSb valence band edge and the InAs conduction band edge, and 2) the slopes of the wave function envelopes in the InAs conduction band and GaSb valence band satisfy the boundary condition at the InAs/GaSb heterojunction:

$$| (1/m_h^*) (d\psi_h/dx) | = | (1/m_e^*) (d\psi_e/dx) | \qquad (1)$$

where the subscripts h and e refer to GaSb light holes and InAs conduction band electrons respectively. The effective masses of 0.056 m_0 for light holes in GaSb and 0.023 m_0 for electrons in InAs with equation (1) lead to

$$w_2 = 2.435\, w_1$$

for the second condition, where w_1 and w_2 are the widths of the GaSb and InAs quantum wells respectively. We assumed a quadratic relation between well width and energy level and used the bound state energies of 0.18 eV for light holes in a 3 nm GaSb quantum well [6] and 0.05 eV for electrons in a 15 nm InAs quantum well [7] to arrive at w_1 = 5.01 nm and w_2= 12.20 nm. The energies E_1 and E_2 of the GaSb and InAs quasi-stationary states are 0.065 eV and 0.075 eV respectively. The AlSb barrier thickness was chosen to be 2.5 nm in order to give forward bias current densities of several hundred A/cm^2 [6].

EPITAXIAL GROWTH AND DIODE FABRICATION

All of the diode wafers were grown by molecular beam epitaxy on InP substrates in a Perkin-Elmer 430 MBE system. We grew a thick buffer layer between the InP substrate and the active diode layers in order to relieve the 4% strain. The buffer was made by the sequential growth of the following: 500 nm of N+ $Ga_{0.47}In_{0.53}As$, a five period superlattice each period of which contained two monolayers of $Ga_{0.47}In_{0.53}As$ followed by two monolayers of InAs, and a thick layer of InAs N doped at 10^{18} cm^{-3}. We then grew the active diode layers (superlattice) followed by a 200 nm N+ InAs cap layer. Diode wafer no. 1 was fabricated with two periods, wafer no. 2 with four periods and wafer no. 3 with eight periods.

The diodes were fabricated with a wet-etch mesa process. The etch bath consisted of H_2O_2, H_2SO_4 and H_2O in a 1:1:100 solution. The bottom N+ layer and the bottom ohmic contact metal was common to all of the diodes. The anode and cathode contact metallizations both consisted of 100 nm of AuGe, deposited by e- beam evaporation and patterned by lift-off. The contacts were not alloyed.

A non-self aligned process was used for wafer no. 1 which resulted in a resistance between the diode mesa and the cathode contact metal of about 100 Ω that dominated the IV characteristics (Figure 2). Wafers nos. 2 and 3 were processed in a self-aligned manner, using the anode metal as the mask for the cathode metal deposition.

ELECTRICAL MEASUREMENTS

The IV characteristics and MBE profile of a 500 μm^2 diode from wafer no. 1 are shown in Figure 2. The reverse current for the diode shows two negative differential resistance (NDR) regions. The first NDR region is caused by band gap blocking in the vicinity of the bottom AlSb layer as indicated in Figure 2c. After the reverse current is reduced by bandgap blocking, further increases in reverse bias voltage cause the reverse current to again increase with bias. A reverse bias voltage increment of about 0.4 V is required to overcome the band gap blocking. This voltage is comparable in magnitude to the InAs room temperature energy gap of 0.36 eV and indicates that thermal injection of holes from the GaSb and AlSb valence bands into the InAs valence band is the mechanism by which the reverse current increase is resumed.

The second NDR for the diodes in wafer no. 1 occurs at a larger current density. The band gap blocking that causes the second reverse bias cutoff occurs at the uppermost AlSb layer in the diode, as indicated in Figure 2c. The current before cutoff is large because there is no quasi-stationary state in the top InAs layer.

A region of NDR is also observed in forward bias. This NDR is due to blocking at the upper InAs/GaSb heterojunction [8].

Figures 3a and 3b show the IV characteristics for four period diodes (wafer no. 2) and eight period diodes (wafer no. 3) respectively. The diodes have anode contacts that are 100 μm^2. The IV curves for the diodes on wafers 2 and 3 are dominated by the intrinsic diode characteristics rather than parasitic resistance. The lower resistance in wafers 2 and 3 is due to the self-aligned process that was used. However, some resistance remains because of contact resistance between the tungsten probe tips and the gold metallization and also because of the wet etch undercut of the anode contact metal. We estimated the probe contact resistance to be 20 Ω and the spreading resistance due to undercut to be 20 Ω for wafer 2 and 60 Ω for wafer 3.

The four period diode gives a forward current of 5 mA at a forward voltage of 0.3 V. Most of the voltage drop is due to contact and spreading resistance. The eight period diode gives a forward drop of 0.75 V. We believe the parasitic resistance of the eight period diodes was large because of severe undercutting of the anode contact during the mesa etch. An anisotropic etching process is needed to optimize the diode structure.

The reverse characteristics of the four and eight period diodes are similar in nature. The reverse current increases with reverse bias until band-gap blocking at one of the GaSb/AlSb/InAs sections prevents InAs conduction band electrons from entering the GaSb valence band and causes NDR to occur. A negative bias increase of about 0.4 V then causes injection of holes into the InAs valence band to overcome the bandgap blocking and the current increases until bandgap blocking occurs at another section of the diode. Since all the sections are similar in design, the bandgap blocking occurs at similar currents, giving the reverse diode characteristic the sawtooth appearance seen in Figures 3a and 3b. The sawtooth characteristic has potential applications for high speed circuits [9,10] The looping seen in Figure 3 is caused by interaction of the diode NDR with the inductance of the curve tracer leads.

CONCLUSION

We have fabricated GaSb/AlSb/InAs superlattice rectifying diodes on InP substrates with a self-aligned wet etch process. The diodes can block more than 1.5 V in the reverse direction while conducting 5,000 A/cm^2 in the forward direction with a forward voltage drop of less than 0.3 V. We have also fabricated diodes capable of blocking more than 3 V of reverse bias. The forward voltage drops are larger than is desirable, but improvement can be expected from more heavily doped contact layers, anisotropic etching for self-aligned contacts, thinner barrier layers and doped InAs and GaSb quantum wells. Reductions in the reverse current can be expected from reductions in the InAs and GaSb quantum well thicknesses.

REFERENCES

1. H. C. Torrey and C. A. Whitmer, Crystal Rectifiers, McGraw- Hill, New York, 1948, Ch. 8.

2. J. Karlovsky, "The Curvature Coefficient of Germanium Tunnel and Backward Diodes," Solid-State Electron., vol. 10, p. 1109, 1967.

3. M. Sweeny and J. Xu, "Resonant Interband Tunnel Diodes," Appl. Phys. Lett., Vol. 54, No. 6, pp. 546-548, Feb. 6, 1989.

4. J. R. Soderstrom, D. H. Chow, T. C. McGill, "New Negative Differential Resistance Device Based on Resonant Interband Tunneling," Appl. Phys. Lett., Vol. 55, No. 11, pp. 1094-1096, Sep. 11, 1989.

5. L. F. Luo, R. Beresford, W. I. Wang, "Interband Tunneling in Polytype GaSb/AlSb/InAs Heterostructures," Appl. Phys. Lett., Vol. 55, No. 19, pp. 2023-2025, Nov. 6, 1989.

6. J. F. Chen, L. Yang, A. Y. Cho, "Investigation of the Influence of the Well and Barrier Thicknesses in GaSb/AlSb/GaSb/AlSb/InAs Double-Barrier Interband Tunneling Structures," IEEE Electron Device Letters, Vol. 11, No. 11, pp. 532-534, November 1990.

7. R. Beresford, L. F. Luo., K. F. Longenbach, W. I. Wang, "Resonant Interband Tunneling through a 110 nm InAs Quantum Well," Appl. Phys. Lett. 56(6),551-553, February 5, 1990.

8. K. F. Longenbach, L. F. Luo, W. I. Wang, "Resonant Interband Tunneling in InAs/GaSb/AlSb/InAs and GaSb/InAs/AlSb/GaSb Heterostructures," Appl. Phys. Lett. 57 (15), pp. 1554-1556, 10/8/90.

9. T. Kuo, H. Lin, R. Potter, D. Shupe, "A Novel A/D Converter Using Resonant Tunneling Diodes," IEEE J. Solid-State Circuits, Vol. 26, No. 2, pp. 145-149, February, 1991.

10. S. Sen, F. Capasso and A. Cho, "Integration of GaAs/AlAs Resonant Tunneling Diodes for Digital and Analog Applications with Reduced Circuit Complexity," 1987 GaAs IC Symposium Technical Digest, pp. 61-64, Portland, OR, October 13-16, 1987.

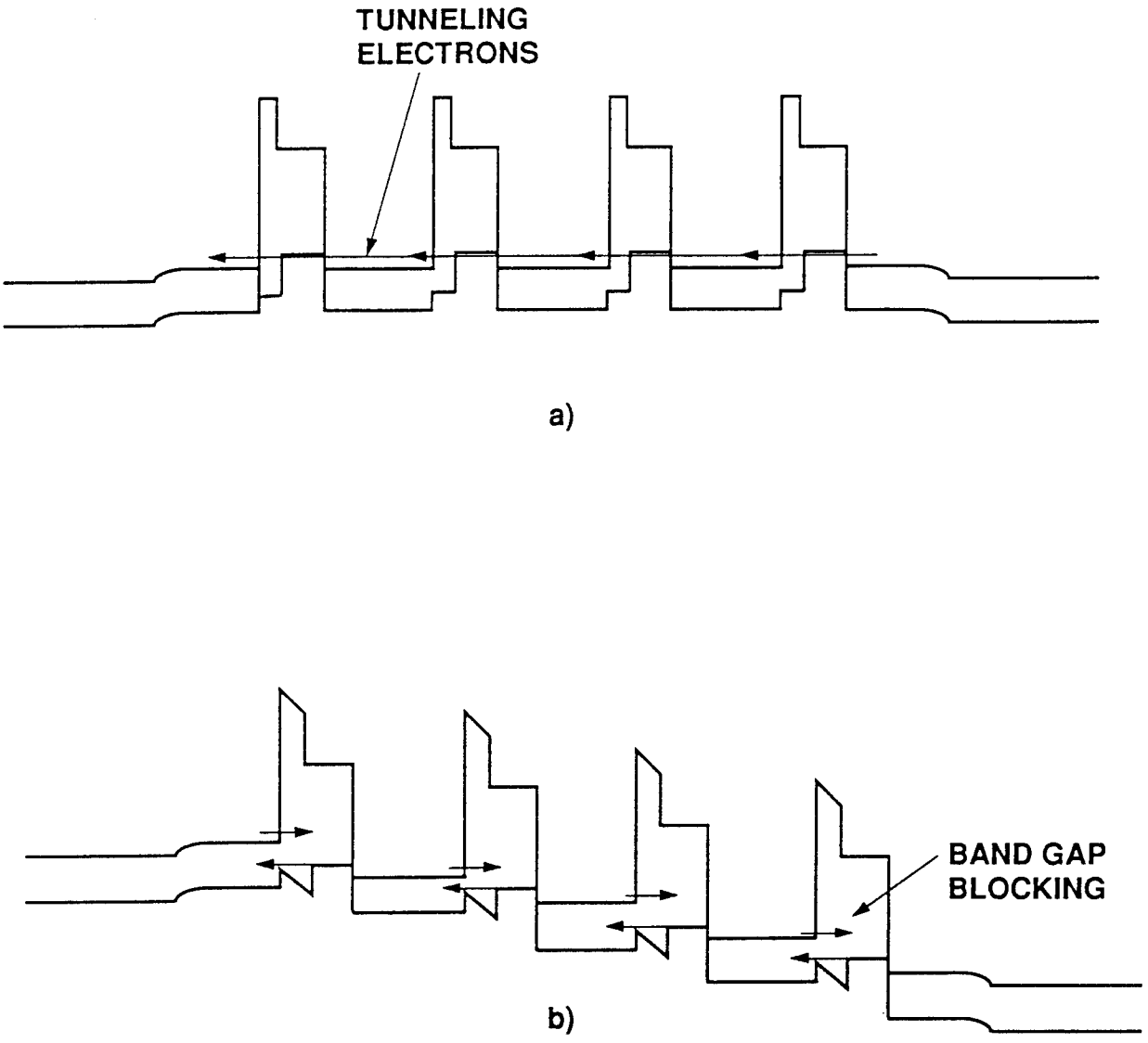
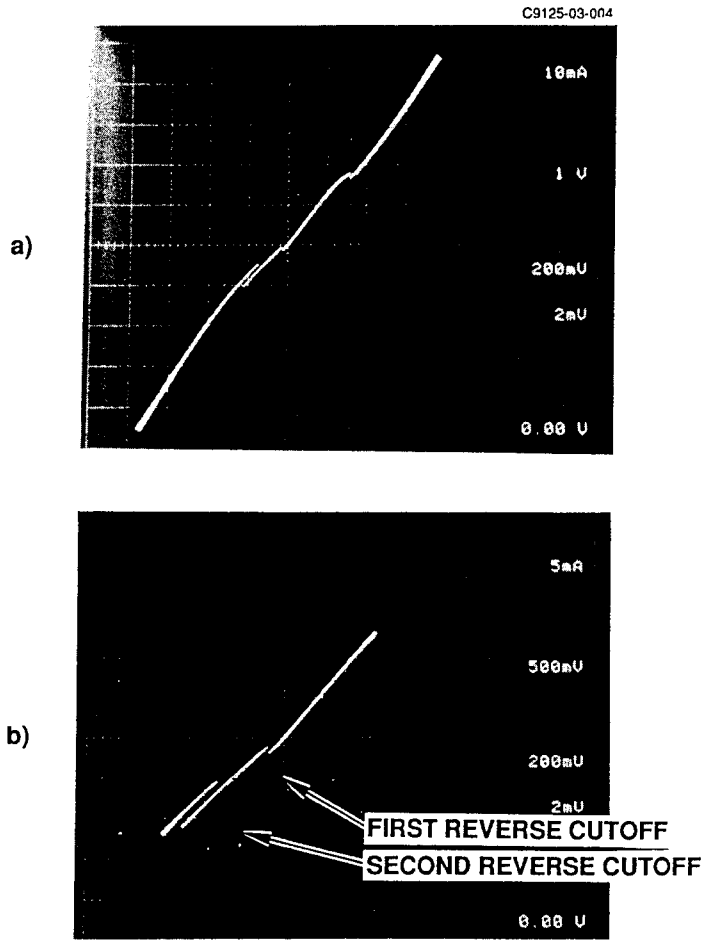


Figure 1 Energy Band Diagrams for Four Section InAs/AlSb/GaSb Diode in a)Forward Bias, and b)Reverse Bias



c)

	N+	200 nm InAs CAP	
		12.2 nm InAs	BAND GAP BLOCKING FOR SECOND
FORWARD BIAS ELECTRON DIRECTION ↑	REVERSE BIAS ELECTRON DIRECTION ↓	2.5 nm AlSb	← REVERSE NDR OCCURS HERE
		5.0 nm GaSb	BAND GAP BLOCKING FOR FORWARD
		12.2 nm InAs	← NDR
		2.5 nm AlSb	← BAND GAP BLOCKING FOR FIRST
		5.0 nm GaSb	REVERSE NDR OCCURS HERE
		12.2 nm InAs	
	N+	1.0 nm InAs	

Figure 2. a) IV Characteristic of a Two Section InAs/AlSb/GaSb Diode Showing Band Gap Blocking in Forward and Reverse Direction, b) Magnified Scale IV Curves from (a) Showing First and Second Cutoff in Reverse Bias. Second Cutoff Occurs at Higher Current Density because of Lack of InAs Stationary State; c) Epitaxial Structure of Two Period Diode.

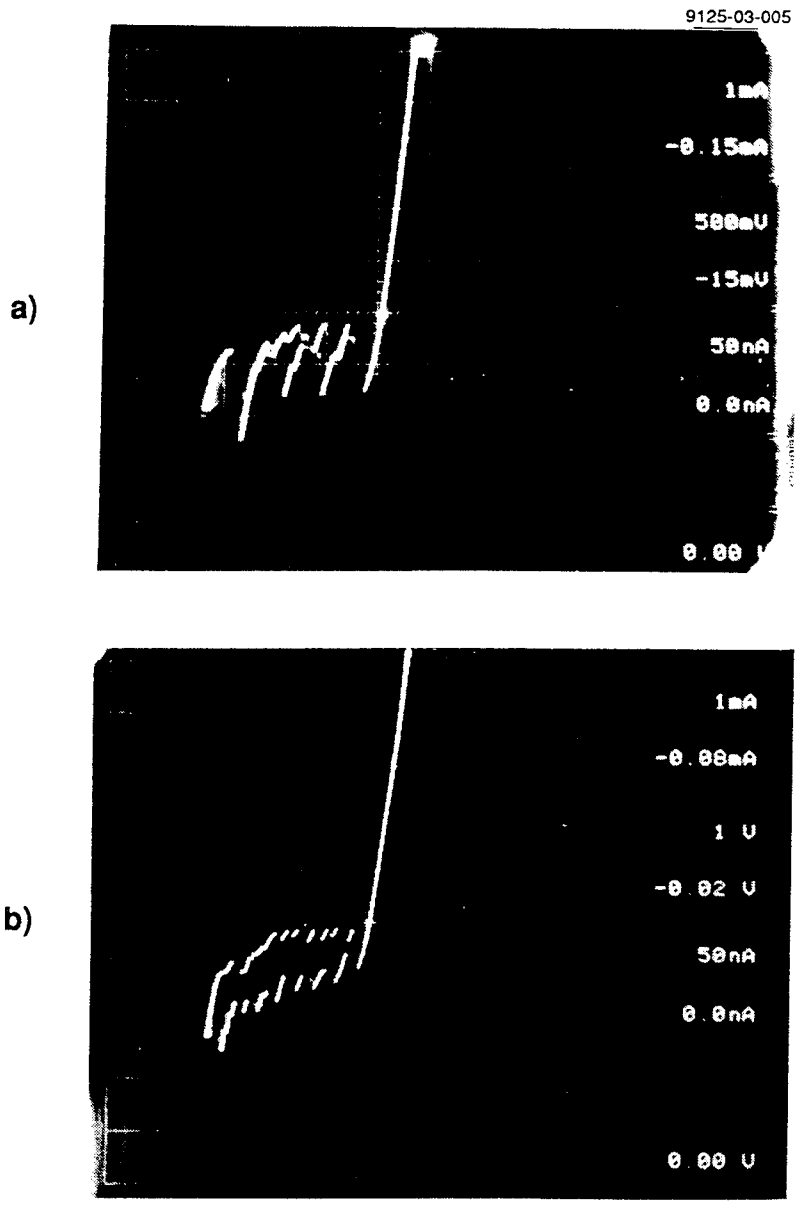


Figure 3. IV Characteristics for a) Four Section, and b) Eight Section InAs/AlSb/GaSb Rectifying Diodes. Note that Voltage Scale is Different for (a) from (b).

DEVELOPMENT OF AN APPROPRIATE MODEL FOR THE DESIGN OF D-BAND InP GUNN DEVICES

R. Kamoua*, H. Eisele†, G. I. Haddad†, G. Munns†, and M. Sherwin‡

* Department of Electrical Engineering,
State University of New York at Stony Brook,
Stony Brook, NY 11794-2350

† Solid State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109-2122

‡ Sandia National Laboratory
Albuquerque, NM 87185-5800

Abstract

The potential of InP Gunn devices as power sources in the fundamental mode at D-band frequencies (110 GHz - 170 GHz) is investigated. A self-consistent ensemble Monte Carlo model has been developed to design and identify suitable structures for operation in this frequency range. Using this model with typical InP material parameters found in the literature will be shown to give results inconsistent with experiment. Based on experimental results from a 1.7 μm long Gunn structure, more realistic material parameters were estimated. The resulting model is then used to design various structures with active regions in the 1 μm range. In particular, two structures, one with a flat doping profile and the other with a linearly graded doping profile, were fabricated and tested. State-of-the-art performance from these structures operating in the fundamental mode was obtained at frequencies ranging from 108.3 GHz to 155 GHz. The flat structure yielded optimum results at 108.3 GHz with a power level of 33 mW while the graded structure gave 20 mW at 120 GHz, 17 mW at 133 GHz, 10 mW at 136 GHz, and 8 mW at 155 GHz. These results will be compared with the model predictions.

1 Introduction

Sources for power generation in the W-band frequency region often consist of two terminal solid state devices. Among these devices Gunn and IMPATT diodes are the most important with Gunn devices suited for low noise, moderate power applications, and IMPATT devices suited for high power applications. At higher frequencies, adequate low-noise sources are not

0-7803-0894-8/93 $3.00

readily available. However, this is not due to a lack of effort but rather to the many challenging problems encountered. These problems include fabricating and packaging devices with increasingly smaller features, designing more complex circuits, and more importantly overcoming fundamental limitations of the carrier response in semiconductors at these frequencies.

In this paper, InP Gunn devices will be shown to be capable of generating adequate amounts of power in the D-band frequency region and in particular in the 140 GHz atmospheric window. Toward this goal, a physical model based on the Monte Carlo technique has been developed to design, simulate, and predict the performance of InP Gunn structures at high frequencies. The validity of the model is found to be strongly dependent on the use of appropriate InP material parameters. These parameters were estimated from a comparison between the model predictions and experimental results at frequencies in the W-band region. Two InP Gunn structures were then designed for operation in the D-band region. The corresponding experimental results will be presented and compared with model predictions

2 Simulation Model

The conversion efficiency and the output power of the InP Gunn structures are estimated by considering the device in a resonant circuit represented by a load resistance and a resonating inductance. The effects of contact resistances, substrate resistances, and skin effect losses are represented in the resonant circuit by a series resistance. The Gunn device itself is represented by a nonlinear admittance consisting of a negative conductance and a capacitive reactance. The admittance is obtained from a characterization of the internal transport dynamics in response to the applied voltage across the device. For this analysis, a model based on the self-consistent ensemble Monte Carlo technique is employed [1]. A sinusoidal RF voltage superimposed on a dc voltage is applied across the device. The current response over many RF periods is then obtained from the Monte Carlo simulation. The current is Fourier analyzed and the fundamental component is used to estimate the device admittance. The next section discusses the important issue of the relation between InP material parameters and the accuracy of the Monte Carlo model.

3 Overview of InP Material Parameters

Before proceeding to device simulation, It is appropriate to overview the currently used InP material parameters. Table 1 lists some parameters relevant to the Monte Carlo method as found in the literature. There is a wide range of values for many parameters important to the Gunn effect. In particular, there is more than an order of magnitude uncertainty in the Γ-to-L intervalley coupling constant. A method often employed to determine these parameters consists of fitting experimental data to the Monte Carlo generated velocity versus electric field characteristics. As pointed out by Fischetti [6], some of the experimental data are relatively insensitive to some of the scattering parameters, such as Γ-to-L deformation potential based

on the velocity-field characteristics. On the other hand, the Gunn effect is very sensitive to the rate of intervalley transfer and therefore to the intervalley deformation potential. This suggests that more accurate estimation of these parameters is possible by comparing experimental data from Gunn devices at high frequencies with model predictions.

Reference		[2]	[3]	[4]	[5]	[6]	this work
Energy Separation (eV)	Γ-L	0.54	0.6	0.6	0.6	0.832	0.4
	Γ-X	0.775	0.8	0.3		1.5	0.775
Effective Mass ($\frac{m^*}{m_0}$)	Γ	0.078		0.068	0.08	0.082	0.082
	L	0.26	0.4	0.4	0.4	0.353	0.5
	X	0.325	0.4	0.4		0.46	0.5
Nonparabolicity factor ($\frac{1}{eV}$)	Γ	0.83			0.635	0.85	0.83
	L	0.23			0.0	0.33	0.23
	X	0.38					0.38
Intervalley Coupling Constant (10^9 eV/cm)	Γ-L	1.0	0.44	0.1	2.5	0.506	1.0
	Γ-X	1.0	0.43	1.0		0.498	1.0
	L-X	0.9	0.367			0.468	0.468
	L-L	1.0	0.246		0.41	0.575	0.575
	X-X	0.9	0.46			0.28	0.28
Acoustic Deformation Potential (eV)	Γ	8			7	5	5.0
	L	8			12	5	5.0
	X	8				5	5.0
LO Phonon Energy (eV)	Γ	0.043			0.043	0.0424	0.043
	L	0.0423					0.0423
	X	0.0416					0.0416
Static Dielectric Constant		12.35		12.37	12.35	12.61	12.61
Optical Dielectric Constant		9.52		9.61	9.52	9.61	9.61

Table 1: InP Material Parameters from the literature and this work.

4 More Appropriate InP Material Parameters

The doping profile of an InP Gunn structure to be used for such a comparison is shown in figure 1. It has a 1.7 μm long active doped at 1×10^{16} cm^{-3}, a 0.1 μm long cathode region doped at 3×10^{17} cm^{-3}, and a 0.2 μm long anode region doped at 3×10^{17} cm^{-3}. An InP wafer with this structure has been processed. Devices with various sizes have been mounted on copper heat sinks. Tapered leads were thermocompression bonded to the diode and to four metallized quartz standoffs. For some devices, metallized quartz rings were used instead of the standoffs. A 50-μm device was tested in a W-band resonant cavity with the following

results: 40.0 mW output power at an oscillation frequency of 80.0 GHz and an efficiency of 1.6 %. The bias voltage was 5.0 V and the dc current was 500 mA. The structure shown in figure 1 is simulated using the Monte Carlo model. The dc bias voltage is set to 5.0 V and the temperature to 450 K. A starting set of material parameters is taken from reference [6] in table 1. No oscillations were obtained with these values for frequencies ranging from 75 GHz to 120 GHz. Upon examining these parameters, it appears that the $\Gamma - L$ intervalley energy separation of 0.832 eV is too large. A systematic procedure for changing the values of the different parameters is adopted. In particular the values used for the intervalley energy separation and the effective mass in each valley are targeted. It is expected that the occurrence of oscillations will be enhanced if the electron effective mass in the satellite valleys is increased, the intervalley energy separation is reduced, and the scattering rates to the satellite valleys are increased. The combined effect of these changes is to increase the transfer to the satellite valleys for the same bias voltage and reduce the average electron velocity at high electric fields. As a consequence a larger negative differential mobility is obtained which is more favorable for nucleating space charge layers. Upon making the above changes a set of parameters, listed in the last column of table 1, is obtained which will be shown to yield good agreement with experimental results.

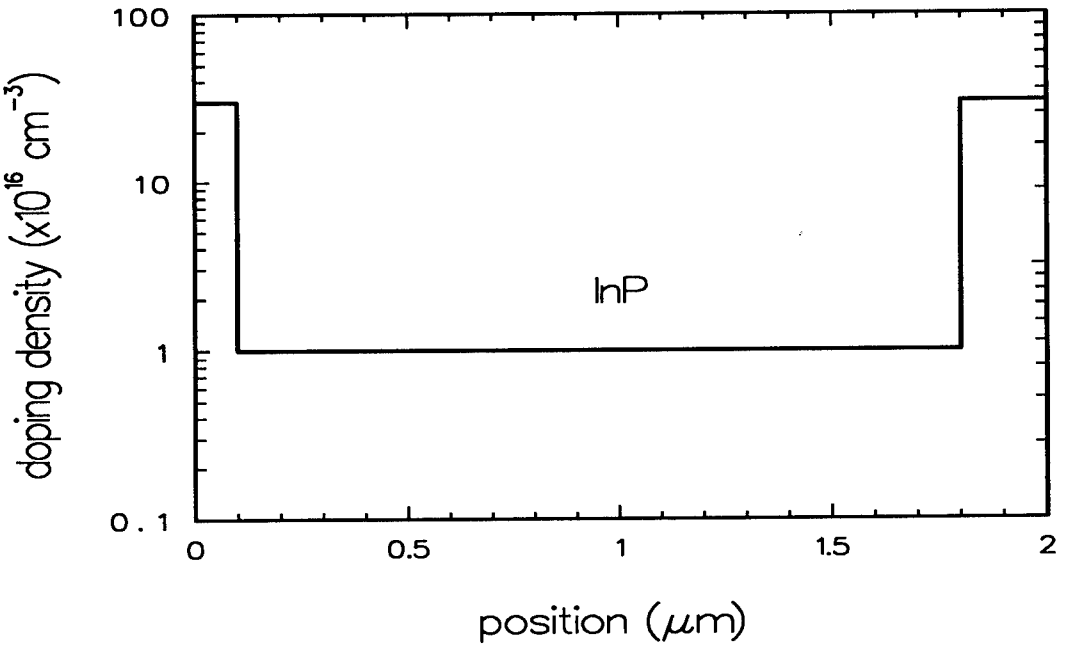


Figure 1: Doping profile of the W-band InP Gunn structure.

As mentioned earlier, we expect the electron transfer to the upper valleys to be enhanced as a result of the changes made to the material parameters. This is clearly revealed in figure 2 which displays the ratio of electrons in the L valley to the total number of electrons as a function of position across the device. The two curves are taken at a phase of $\frac{\pi}{4}$ in one RF cycle and correspond to Monte Carlo results employing the material parameters taken from the literature and the new material parameters used in this work. The latter set of parameters

results in a much higher ratio of electrons in the L valley and in particular closer to the cathode contact (position = 0). The simulation of the structure in figure 1 at 80 GHz resulted in an average dc current density of 26603 A.cm^{-2} and an equivalent admittance Y_D = (-2114 + j 2057) (mho.cm^{-2}). In order to compare the simulation results with the experimental data, the diameter of the device is set to 50 μm which requires a load resistance of 12.4 Ω. The predicted output power is 45.8 mW with an efficiency of 1.79 % and a dc current of 522 mA. These results are in very good agreement with the experimental data obtained from a similar Gunn structure.

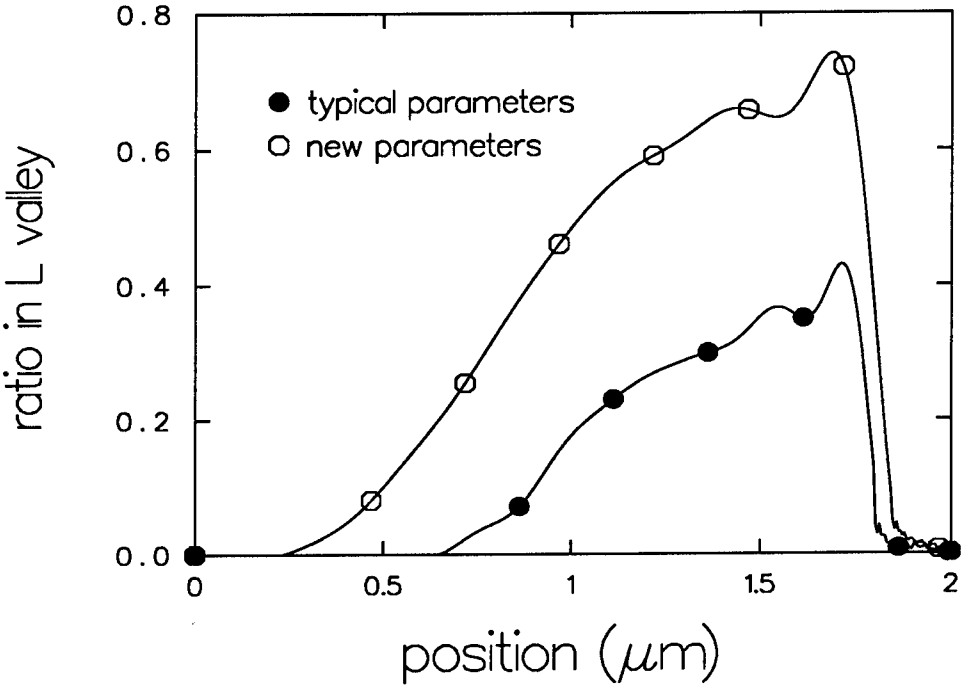


Figure 2: Ratio of electrons in the L valley as a function of position at $\omega t = \frac{\pi}{4}$.

In the next section, two InP Gunn structures designed for operation in the D-band will be considered. Simulation results based on the above Monte Carlo model will be compared with experimental data.

5 InP Gunn Structures for D-band Operation

For operation in the D-band, it is necessary to decrease the device length and increase the doping level in the active region. Two InP structures with different doping profiles are considered. The first has a 1-μm-long flat doping profile while the second has a 1-μm-long linearly graded doping profile in the active region.

5.1 Flat Doping Profile

This structure consists of a 1-μm-long active region doped at 2.5 $\times$ 10^{16} cm^{-3}. A wafer was grown by CBE (Chemical Beam Epitaxy). The wafer was processed and devices with a nominal diameter of 35 μm were packaged and tested in a W-band cavity with a resonant cap on the device. Oscillations were obtained at 108.3 GHz with 33 mW CW output power and 1.87 % efficiency. The device was biased at 4.1 V and has a dc current of 430 mA.

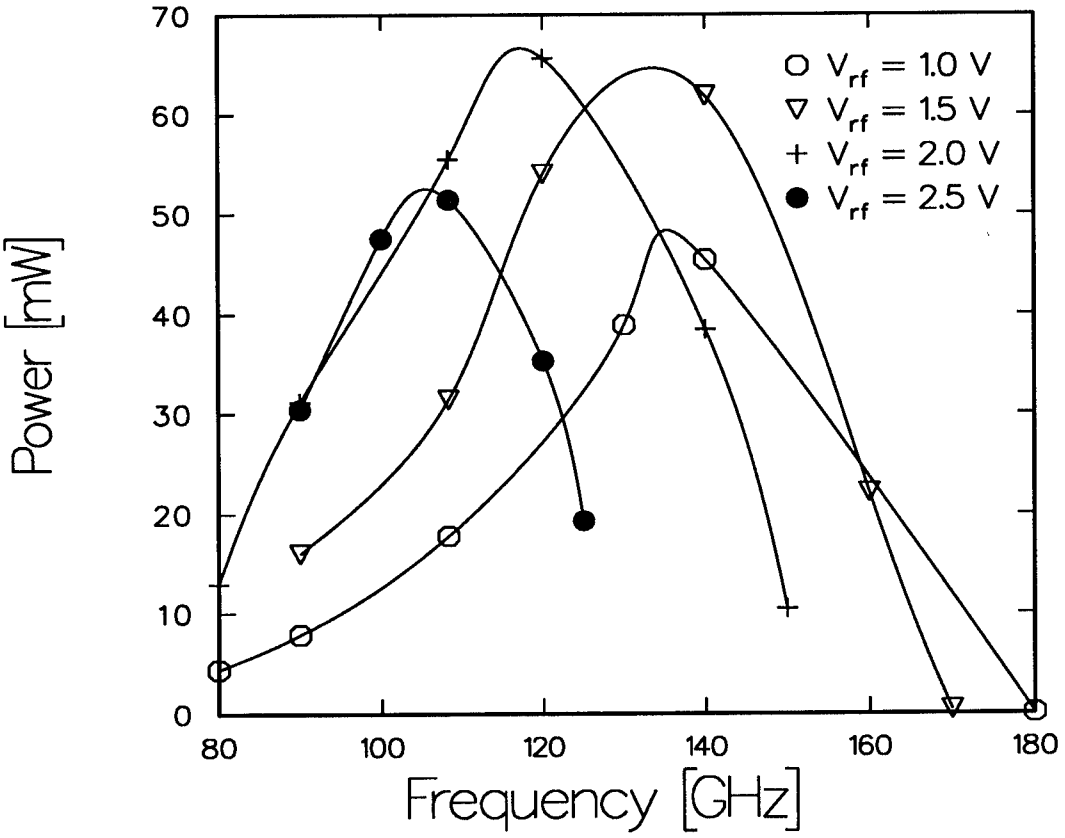


Figure 3: Predicted output power as a function of frequency for a 1-μm-long Gunn structure with a flat doping profile.

To simulate this structure, the bias voltage was set to 4.1 V and the operating temperature to 480 K. The actual device operating temperature is estimated to be close to 500 K since most of the devices failed as the bias voltage is increased beyond 5 to 5.5 V. The performance was evaluated as a function of frequency and for different RF voltages (1.0 V, 1.5 V, 2.0 V, and 2.5 V). Figure 3 shows a plot of the RF power versus frequency. More than 60 mW output power is predicted near 120 GHz. For an RF voltage of 2.5 V, the power peaks near 108 GHz and the oscillation bandwidth becomes smaller compared with lower RF voltages. The predicted power at 108.3 GHz in this case is approximately 50 mW. This value is in good

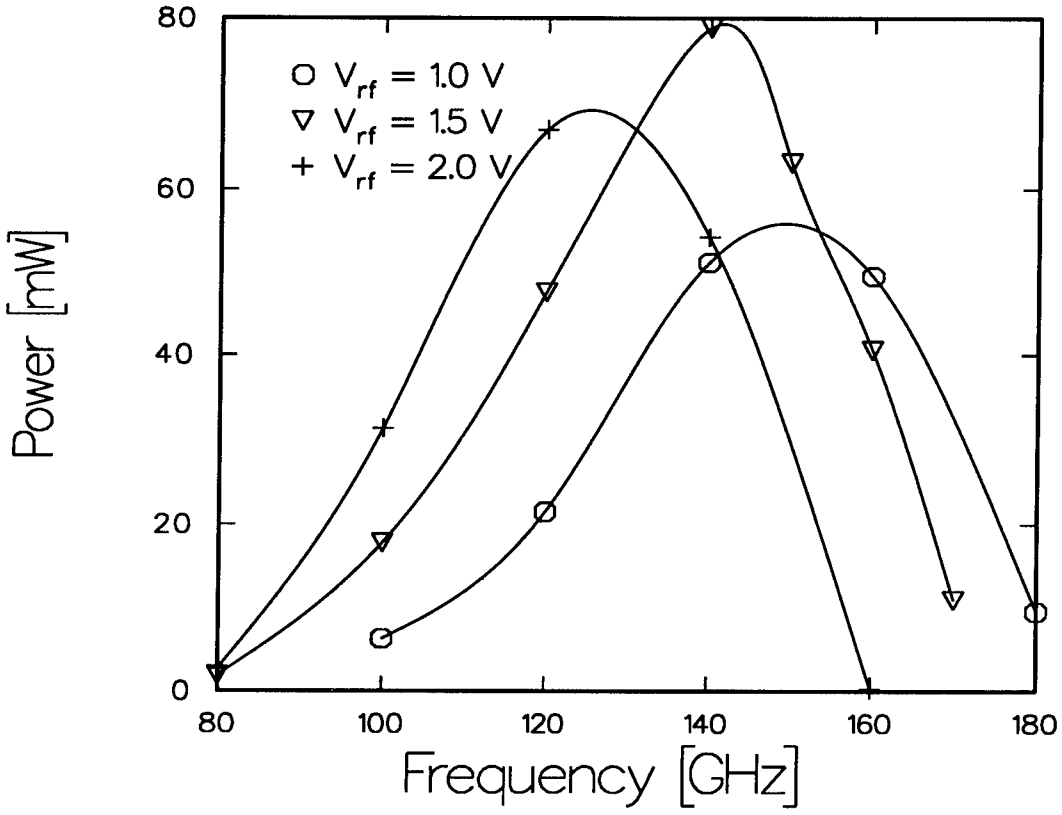


Figure 4: Predicted output power as a function of frequency for a 1-μm-long Gunn structure with a linearly graded doping profile.

agreement with the experimental results when taking into account cavity losses and the fact that the experimental results are obtained from a single wafer.

5.2 Graded Doping Profile

In order to further enhance the performance of InP Gunn devices in the D-band, a nonuniform doping profile in the active region is considered. The advantages of a linearly doping profile include reducing the peak electric field, lowering the current density, and improving the efficiency and output power. A wafer with a graded doping profile was designed and grown by MOCVD. The structure consisted of a 1-μm-long active region with a doping linearly increasing from 7.5×10^{15} cm^{-3} at the cathode side of the active region to 2.0×10^{16} cm^{-3} at the anode side. Samples were processed with integrated heat sinks consisting of gold and silver layers. Devices with a nominal diameter of 45 μm were mounted on copper heat sinks and tested in a D-band waveguide cavity. Oscillations were obtained at 120 GHz with a power of 20 mW, at 136 GHz with a power of 10 mW, at 133 with a power of 17 mW, and at 155 GHz with a power of 8 mW. These devices are believed to operate in the fundamental mode since the

measured Q values were between 30 and 100 around 120 GHz using a self-injection locking method. These Q values are not characteristic of a harmonic mode operation [7, 8]. Even though these results are preliminary, they represent the best performance from Gunn devices operating in the fundamental mode reported at these frequencies.

Simulations were carried out on a similar structure using the Monte Carlo model. The dc bias voltage was set to 4.1 V, the operating temperature at 400 K, and the device diameter was fixed at 45 μm. Figure 4 shows the output power as a function of frequency for different RF voltages. Oscillations occur at frequencies ranging from 80 GHz to 180 GHz. At the two extremes however, the oscillation condition requires load resistances less than 1 Ω. The output power peaks near 140 GHz where 80 mW is predicted for an RF voltage of 1.5 V. The simulation predicts oscillations in a frequency range corresponding to what has been observed experimentally, but with better performance. As mentioned above, the experimental results are preliminary and more work need to be done to further improve the performance of the graded structures at D-band frequencies.

6 Conclusions

A self consistent ensemble Monte Carlo model has been developed for the simulation of InP Gunn devices at high frequencies. When typical InP material and physical parameters were employed, the model could not predict our experimental results. Appropriate parameters were determined by comparing simulation results and experimental data obtained from an InP Gunn structure operating at 80 GHz. The new parameters are characterized by low intervalley energy separation and high deformation potentials compared to typical values used in the literature. Possible reasons for these trends include the high operating temperature of the Gunn device which alters the band structure, and the lower density of states of nonparabolic bands compared to the actual band structure.

Based on this model, simulation results have shown that it is possible to operate fundamental mode InP Gunn devices in the D-band frequency region. Two structures have been designed, modeled, fabricated, and tested. The first structure consisted of a 1-μm-long active region uniformly doped at 2.5×10^{16} cm^{-3}. The second structure had a linearly graded doping profile increasing from 7.5×10^{15} cm^{-3} at the cathode side to 2.0×10^{16} cm^{-3} at the anode side. Preliminary experimental results were very encouraging and represent the state of the art performance from Gunn devices operating in the fundamental mode at these frequencies. Both structures operated over roughly the same frequency range, however the graded structure yielded better performance at the high frequency end. Improving the experimental results requires optimizing the doping profile, reducing further the contact resistances, and developing better heat sinks and packaging techniques.

Acknowledgements

This work was supported by the Center for Space Terahertz Technology at the University of Michigan under Contract No. NAGW-1334.

References

[1] R. Kamoua, J.R. East, and G.I. Haddad, "Ensemble Monte Carlo Characterization of Graded $Al_xGa_{1-x}As$ Heterojunction Barriers," **Journal of Applied Physics**, 68(3):1114-1122, August, 1990.

[2] K. Brennan, K. Hess, J. Y. Tang, and G. J. Iafrate, "Transient Electronic Transport in InP Under the Condition of High-Energy Electron Injection," **IEEE Trans. on Electron Dev., ED-30**, No. 12, pp. 1750-1753, Dec. 1983.

[3] D. C. Herbert, W. Fawcett, and C. Hilsum, "High Field Transport in Indium Phosphide," **J. Phys. C: Solid State Phys.**, Vol. 9, pp. 3969-3975, 1976.

[4] G. H. Glover, "Study of Electron energy Relaxation Times in GaAs and InP," **J. Appl. Phys., 44**, No. 3, pp. 1295-1301, March 1973.

[5] T. J. Maloney, and J. Frey, "Transient and Steady-State Electron Transport Properties of GaAs and InP," **J. Appl. Phys., 48**, No. 2, pp. 781-787, Feb. 1977.

[6] M. V. Fischetti, "Monte Carlo Simulation of Transport in Technologically Significant Semiconductors of the Diamond and Zinc-Blende Structures-Part I: Homogeneous Transport," **IEEE Trans. on Electron Dev., ED-38**, No. 3, pp. 634-649, March 1991.

[7] H. Barth, "A High Q Cavity Stabilized Gunn Oscillator at 94 GHz," **IEEE MTT-S Digest**, p. 179, 1986.

[8] I. G. Eddison, and D. M. Brookbanks, "Operating Modes of Millimeter Wave Transferred Electron Oscillators," **Electronics Letters, 17**, No. 3, p. 112, 1981.

QUANTUM WIRE LASERS

Eli Kapon, Bellcore

PAPER UNAVAILABLE FOR PUBLICATION

0-7803-0894-8/93 $3.00

RELIABILITY OF STRAINED QUANTUM WELL LASERS

H. Yoon, Y. C. Chen, L. Davis, H. C. Sun, K. Zhang, J. Singh, and P. K. Bhattacharya

Solid State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109-2122

ABSTRACT

Strained quantum well lasers have demonstrated remarkably improved characteristics compared to unstrained quantum well lasers. For extracting the highest level of performance, the required strain may be large. An important factor in the use of strained quantum wells is the long-term stability of the pseudomorphic active region and the associated reliability of the device. The effect of strain on reliability is investigated, in particular, for $In_xGa_{1-x}As/GaAs$ (x=0.2, 0.25, and 0.3) multiple quantum well lasers in 64mW/facet constant output power tests at 85°C for 40 hours. Laser characteristics such as the operating currents (I_{op}), the threshold currents (I_{th}), and the slope efficiencies (dL/dI) are measured during the test and serve as useful degradation parameters. The average changes in I_{op} are 15, 9.9, and 0.22%, and the average changes in I_{th} at 85°C are 21, 8.7, and -1.2% for x=0.2, 0.25, and 0.3, respectively. The average changes in dL/dI at 85°C are -19, -14, 1.5%, respectively. Defect migration into the pseudomorphic active region is verified to be the dominant mechanism of degradation observed in these lasers. Hence, to account for the strain-induced reliability improvement, it is necessary to study the propagation of defects in semiconductor heterostructures. A theoretical model is constructed based on the the linear theory of elasticity, and relevant experiments are conducted for its support. Strain energy considerations show that defect propagation across a strained layer is unfavorable. The nonradiative defect densities in the $GaAs/Al_{0.4}Ga_{0.6}As$ quantum wells with and without the surrounding pseudomorphic $In_{0.2}Ga_{0.8}As$ layers are compared by measuring the photoluminescence intensities after intentionally creating defects and enhancing their diffusion. The structures with pseudomorphic $In_{0.2}Ga_{0.8}As$ layers consistently show much higher quantum well photoluminescence intensity by as much as 130 times, thereby confirming our model. These results clearly account for the observed reliability improvement in quantum well lasers with increased strain in the well.

I. INTRODUCTION

With the development of highly sophisticated fiber communication technologies, optoelectronic systems have demanded increasingly high-performance semiconductor light sources. Accordingly, technological innovations and breakthroughs have been continually made in achieving a device of superior characteristics. In particular, InGaAs/GaAs strained quantum well lasers have become the most important devices available as pumping sources for erbium-doped fiber amplifiers and as transparent substrate vertical cavity lasers for space communication applications. Undoubtably, the device stability and reliability are of prime importance. The utilization

0-7803-0894-8/93 $3.00

of innovative design, material, and production technologies for the enhancement of laser operating characteristics must therefore be carefully examined in proper perspective regarding their impact on practical device reliability.

The use of strain in InGaAs/GaAs quantum well lasers has shown benefits of low threshold current [1], high modulation bandwidth [2], and high power capability [3]. Although promising reliability [4] of these lasers has been reported as well, the underlying mechanism of improved reliability in strained lasers has not been carefully studied. In this paper, strain-induced reliability improvement is experimentally verified, and suitable theoretical and experimental models are constructed to account for the result.

II. EXPERIMENTAL DETAILS

In this work, device reliability is examined in constant output power degradation tests. To establish a sound basis of comparison among devices of different strains, all testing conditions and degradation-enhancing stress parameters are kept identical. Then, the strain-related degradation mechanisms, which will manifest themselves as relative changes among different strain systems, can be analyzed by monitoring suitable degradation parameters. Changes in sensitive signatures of laser characteristics such as the operating current (I_{op}), the threshold current (I_{th}), and the slope efficiency (dL/dI) will delineate the strain's effect on the device reliability.

The tested devices are $In_xGa_{1-x}As/GaAs$ (x=0.2, 0.25, and 0.3) multiple quantum well separate confinement heterostructure (MQW-SCH) lasers grown by the molecular beam epitaxy (MBE). A simple device structure, a stripe geometry broad area laser, is chosen in order to ease the analysis and prevent processing-related variations from obfuscating the results. The structure is shown in Fig. 1. Four 50Å $In_xGa_{1-x}As$ quantum wells with 250Å GaAs barriers are surrounded by a 0.1μm GaAs guiding layer and a 1μm doped $Al_{0.6}Ga_{0.4}As$ cladding layer on each side. A 0.6μm n^+ GaAs buffer layer and a 0.1μm p^+ GaAs cap layer are included.

0.1 μm	GaAs	p^+ (1×10^{19} cm^{-3})
1.0 μm	$Al_{0.6}Ga_{0.4}As$	p (5×10^{17} cm^{-3})
0.1 μm	GaAs	i
4 MQWs	$In_xGa_{1-x}As$ -- 50 Å GaAs -- 250 Å	i
0.1 mm	GaAs	i
1.0 μm	$Al_{0.6}Ga_{0.4}As$	n (5×10^{17} cm^{-3})
0.6 μm	GaAs buffer	n^+ (5×10^{18} cm^{-3})
	GaAs (100) substrate	n^+

Fig. 1. $In_xGa_{1-x}As/GaAs$ MQW-SCH laser structure.

The samples are grown on (001)-oriented substrates. Facets are formed by cleavage and are left uncoated. A total of six 300 μm x 75μm stripe geometry broad area lasers, two from each strain system, are randomly selected for the test.

Prior to a 40-hour aging test, several pre-test characterization measurements are made, including emission wavelength (λ_g) and light versus current (L-I) measurements at several temperatures. During the test, the heat sink temperature is maintained at 85°C to within 1°C, and the bias was pulsed at 0.4μs pulse width and 5KHz repetition rate. Every two hours during the test, an L-I measurement is taken for I_{th} and dL/dI extraction. At the end of the aging test, similar post-test characterization measurements are made for the full-fledged comparison.

To elucidate the effect of a strained layer in defect propagation, further experimentation is conducted in which the quantum well photoluminescence (QW-PL) of two structures grown by MBE is observed. As shown in Fig. 2, their structures are identical but for the addition of two thin strained $In_{0.2}Ga_{0.8}As$ layers (guarding layers) surrounding the quantum well in one structure. Samples of each structure with the $Al_{0.4}Ga_{0.6}As$ layers grown at three different temperatures (600°C, 640°C, and 670°C) are prepared in order to monitor the growth temperature-related effects. Defects of various types are intentionally created by using substrates of high etch pit densities (EPD, 5x10cm^{-2}) and/or by implanting Si ($R_p+\Delta R_p \sim 0.19\mu$m, $N_p = 10^{18}$cm^{-3}) for the structure with the $Al_{0.4}Ga_{0.6}As$ layers grown at 670°C. The diffusion of these defects is enhanced by the high growth temperature and subsequent rapid thermal annealing (RTA, at 800°C for 15 seconds), respectively. The QW-PL intensities of both the as-grown and the implanted, RTA-treated samples are contrasted. Care is taken in designing an implantation profile such that the charge transfer into the QW is insignificant, and sufficient time is allowed to insure that substantial defects may have entered the QW region. Since the accumulated defects in the quantum well will lower its optical efficiency, a comparison of defect densities between the two structures can be made by measuring their QW-PL intensities.

150 Å	GaAs
2000 Å	$Al_{0.4}Ga_{0.6}As$
80 Å	$In_{0.2}Ga_{0.8}As$
400 Å	$Al_{0.4}Ga_{0.6}As$
100 Å	GaAs QW
400 Å	$Al_{0.4}Ga_{0.6}As$
80 Å	$In_{0.2}Ga_{0.8}As$
300 Å	$Al_{0.4}Ga_{0.6}As$
	GaAs (100) substrate

(a)

150 Å	GaAs
2400 Å	$Al_{0.4}Ga_{0.6}As$
100 Å	GaAs QW
700 Å	$Al_{0.4}Ga_{0.6}As$
	GaAs (100) substrate

(b)

Fig. 2. GaAs/$Al_{0.4}Ga_{0.6}As$ QW structures
(a) with and (b) without strained $In_{0.2}Ga_{0.8}As$ layers.

III. RESULTS AND DISCUSSION

There is clear evidence of a reduced degradation rate of the lasers for higher strained systems. The gradual degradation can be quantified by percentage changes in I_{op}, I_{th}, and dL/dI. Fig. 3 is the plot of operating current (I_{op}) versus time. The plot of L-I curves in Fig. 4 shows degradation of I_{th} and dL/dI over time. The results are summarized in Table I. In all cases, minute changes in the lasing wavelength assures that the pseudomorphic regions have not relaxed.

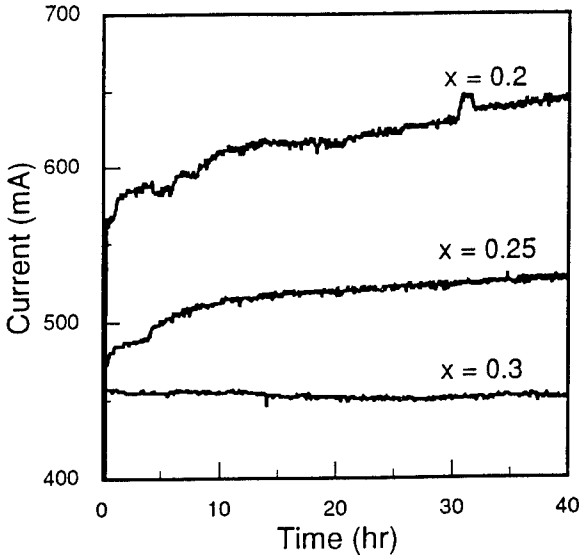


Fig. 3. Operating current vs. time from $In_xGa_{1-x}As/GaAs$ laser constant output power tests.

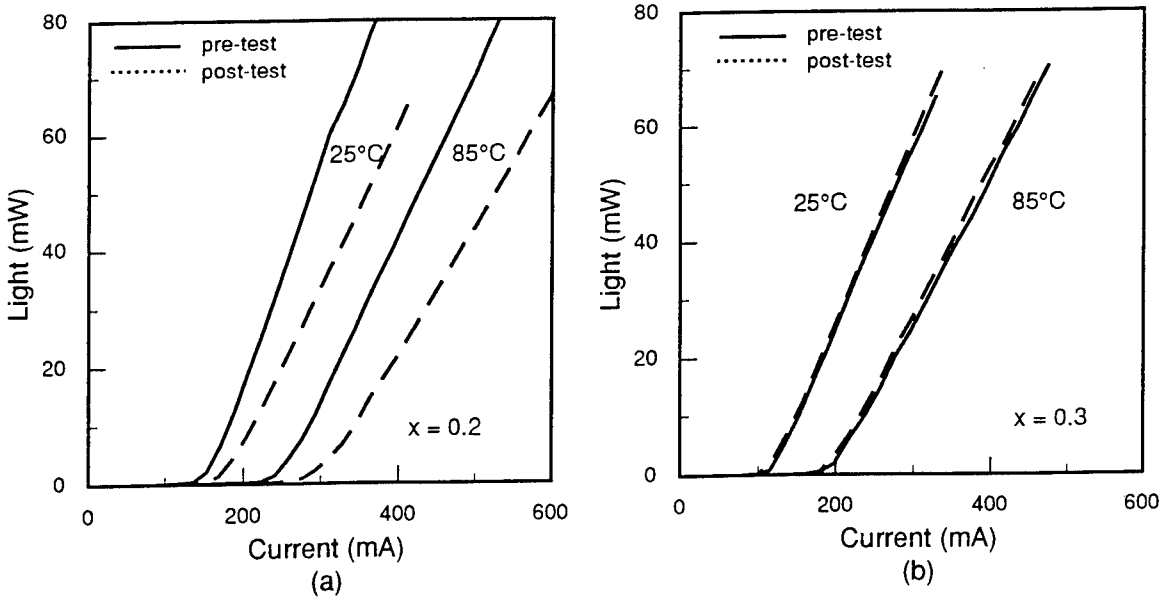


Fig. 4. L-I curves for (a) x=0.2 and (b) x=0.3 before and after the degradation test.

Table I. Degradation of $In_xGa_{1-x}As/GaAs$ MQW-SCH lasers by percentage changes

	Average change [%]		
	x=0.2	x = 0.25	x = 0.3
Iop [mA]	15.3	9.89	0.224
Ith[mA]	16.2	11.5	-1.28
Ith[mA]	20.7	8.71	-1.23
λ_g [μm]	-0.0250	-0.254	0.0462
dL/dI per facet [W/A] @ 25 °C	-13.6	-13.5	0.075
dL/dI per facet [W/A] @ 85 °C	-18.7	-13.6	1.50

The strain is also found to have the beneficial effect of preserving the optical quality of quantum wells in the structures with guarding layers. The ratios of the QW-PL peak intensities for the two as-grown structures with $Al_{0.4}Ga_{0.6}As$ layers grown at 600°C, 640°C, and 670°C are 3, 5, and 17, respectively. The exponential dependence of the diffusivities of defects on the temperature is reflected. In addition, the low temperature PL spectra measured from the implanted and RTA-treated structures with and without the guarding layers show the QW electron-to-heavy-hole (e-hh) transition intensities differing by ~130 as shown in Fig. 5. These large changes are due to the combined effect of the increased number of available defects, the increased diffusivity of defects, and the annealing of the internal defects inside the QW.

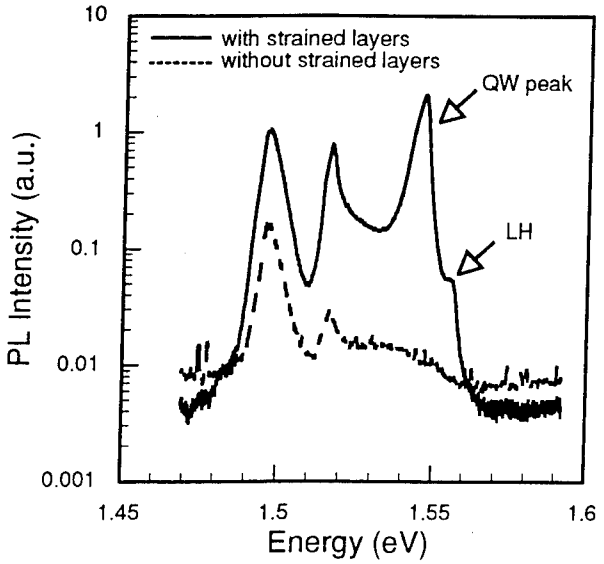


Fig. 5. Low-temperature PL spectra of $GaAs/Al_{0.4}Ga_{0.6}As$ QW structures with and without strained layers after implantation and RTA

Several degradation mechanisms may be observed during aging of lasers: (i) defect formation and growth in the active region, (ii) defect formation and growth in the confining region, and (iii) facet degradation. The objective is to connect these mechanisms to variations in measurable degradation observables. The actual degradation under device operation is quite complex in that more than one type of degradation can occur simultaneously. The fact that the confining layers in all strain systems are exactly identical has two important implications. First, the comparable defect densities in the confining regions should lead to similar degradation rates of these regions under the low-level injection. Second, the rates of facet degradation should also be comparable since most optical confinement takes place within the large cladding regions with the QW regions contributing only a minute portion of the total optical confinement. Therefore, degradation rates due to the facet and confinement region deterioration will approximately be the same for all strain systems. Thus, varying rates of related degradation observables can be solely attributed to the different amount of defect accumulation in the active regions of the three strain systems. From the notably insignificant degradation rates for the $In_{0.3}Ga_{0.7}As/GaAs$ lasers, the degradation rates due to the mechanisms (i) and (ii) universally found in all strain systems are indeed negligible. Therefore, the defect propagation into the active layer can be deduced to be the most critical mechanism in determining the overall laser degradation rate.

The dark line defects (DLD) preferentially propagating along <100> and growing by recombination enhanced defect motion have been identified as the dominant cause of rapid degradation in InGaAs/GaAs lasers [5]. For gradual degradation, point defects such as interstitials and vacancies play the dominant role. In the active and guiding regions, they will form nonradiative recombination centers and increase the carrier recombination rate. Details of the various defects are quite diverse and complicated, but for a rudimentary analysis, they can be treated as a collection of identical point defects. The carrier lifetime, τ, may be expressed as

$$\frac{1}{\tau} = \sigma v_{th} N_{nr} + \frac{1}{\tau_r} \tag{1}$$

where τ_r, σ, v_{th}, and N_{nr} are the radiative carrier lifetime, the capture cross section of the defect, the thermal velocity of the carriers, and the density of defects in the active region, respectively. Since the threshold current can be written as

$$I_{th} = \frac{qVn_{th}}{\tau} + I_L \tag{2}$$

where V, n_{th}, and I_L are the active volume, the threshold carrier density, and the leakage current away from the active layer, respectively [6], it is clear that from (1) and (2) that an increase of the defect density will cause the threshold current to increase. Also, the output power per facet can be written as

$$L = \frac{h\nu}{2q} \nu_g \alpha_m \tau_p (I - I_{th}) \tag{3}$$

where ν_g, α_m, and τ_p are the group velocity, the mirror loss, and the photon lifetime, respectively [6]. From (3), it is easy to see that

$$\frac{dL}{dI} = \frac{h\nu}{2q} \nu_g \alpha_m \tau_p. \tag{4}$$

From (1)-(4), it can be seen that defects induce reductions in the carrier and photon lifetimes, producing an increased threshold current and a reduced slope efficiency from the leakage current and the scattering and absorption loss of photons, respectively. These equations conform very well with the experimental results if the defect density in the active region of higher strained system is to be lower. Therefore, to account for the observed strain-induced laser reliability improvement, the strain must have provided a definite means of hindering defect propagation into the pseudomorphic active layer.

IV. ANALYSIS

To establish the physical basis for the strain-related degradation mechanism, a model is proposed where the defect propagation from lattice matched layers into pseudomorphic layers is suppressed due to an unfavorable strain energy barrier between the layers [7]. Within the limit of the compositionally dependent critical thickness, the strain in the pseudomorphic layer can be accommodated elastically. For a small deformation such as that caused by a point defect, the linear theory of elasticity provides a good approximation of the stress-displacement relation.

The strain energy density can be written in the Einstein convention as

$$w = \frac{1}{2} c_{ijkl} \epsilon_{ij} \epsilon_{kl} \tag{5}$$

where c's and ϵ's are the elastic coefficients and the strain tensor coefficients, respectively [8]. For a cubic crystal, there are only three nonzero elastic constants $c_{iiii} \equiv c_{11}$, $c_{iijj} \equiv c_{12}$, and $c_{ijij} \equiv c_{44}$ for $i \neq j$, and (5) can be simplified as

$$w = \frac{1}{2} c_{11}(\epsilon_{11}^2 + \epsilon_{22}^2 + \epsilon_{33}^2) + c_{12}(\epsilon_{11}\epsilon_{22} + \epsilon_{22}\epsilon_{33} + \epsilon_{33}\epsilon_{11}) + c_{44}(\epsilon_{12}^2 + \epsilon_{23}^2 + \epsilon_{31}^2). \tag{6}$$

In an epitaxial semiconductor layer grown on (001) substrate, it can be shown that $\epsilon_{ij} = 0$ for $i \neq j$. So the last term in (6) is equal to zero. Furthermore, the strain tensor can be written as

$$\epsilon_{11} = \epsilon_{22} = \epsilon_{\parallel} \equiv \epsilon = \frac{a_s - a_o}{a_s}, \tag{7}$$

$$\epsilon_{33} = \epsilon_{\perp} = \frac{-2c_{12}}{c_{11}} \epsilon \tag{8}$$

where a_s and a_o are the substrate and epi-layer lattice constants, respectively [9].

Now consider an excess distortion above equilibrium ($\delta\epsilon_{11}$) producing an incremental strain energy in the strained and the lattice matched layers. In the pseudomorphic layer (region A), the excess strain energy density can be obtained from (6) and (8) to be

$$\delta w_A = c_{11}\epsilon_{11}\delta\epsilon_{11} + c_{12}\epsilon_{22}\delta\epsilon_{11} - 2\frac{{c_{12}}^2}{c_{11}}\epsilon_{11}\delta\epsilon_{11} + \frac{1}{2}c_{11}\delta\epsilon_{11}^2. \tag{9}$$

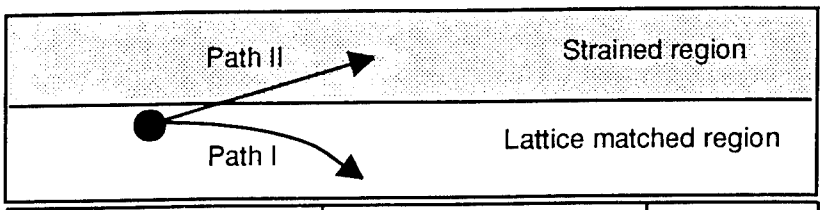
In the lattice matched layer (region B), the energy density in excess of the equilibrium energy density (which equals zero) is

$$\delta w_B = \frac{1}{2} c_{11} \delta\epsilon_{11}^2. \tag{10}$$

As an approximation, the same elastic coefficients are used for both layers.

To find the fate of each defect near the heterostructure, the path of minimum strain energy must be determined. This necessitates a case study on the epitaxial strain and the defect-induced strain by their types, i.e, either compressive or tensile. Together with the fact that $\delta\epsilon_{11} \ll \epsilon_{11}$ and that $c_{11} \approx 2c_{12}$ [9], the preferred path of defect propagation can be summarized as in Table II by using (9) and (10). Although the precise nature of the defects is quite difficult to identify, if the compressive-type defects predominate in the lattice matched region of our lasers, it can be seen from Table II that the defect movement will be met by much less resistance in the lattice matched layer than in the compressively strained layer. Consequently, the defect has to surmount a large energy barrier if it is to propagate into the pseudomorphic layer.

Table II. Preferred path for defect propagation.



Strain in the pseudomorphic layer	Strain induced by defect	Preferred path
Compressive ($\varepsilon < 0$)	Compressive ($\delta\varepsilon < 0$)	I
Compressive ($\varepsilon < 0$)	Tensile ($\delta\varepsilon > 0$)	II
Tensile ($\varepsilon > 0$)	Compressive ($\delta\varepsilon < 0$)	II
Tensile ($\varepsilon > 0$)	Tensile ($\delta\varepsilon > 0$)	I

V. CONCLUSION

The strain-induced laser reliability improvement in InGaAs/GaAs MQW-SCH lasers is verified. The experimentation on the photoluminescence of QW structures reinforces the strain's reliability benefit. The presented theoretical model shows that a strained layer can effectively impede point defects from penetrating through the heterostructures. A reduction of defects and an understanding of the relationship between the nonradiative recombination process and changes in device characteristics are seen to be very important to improve reliability. Through intensive material analysis and reliability testing, much improved as well as reliability-assured device design, material, and production technologies can be achieved. The suppression of defect propagation by the pseudomorphic guarding layers has important implications in MBE growth of semiconductor heterostructures. Instead of smoothing buffer layers, strained guarding layers can be used not only to improve device reliability but also to save growth time and source materials. This aspect is being currently investigated.

VI. ACKNOWLEDGMENT

This work is supported by the Office of Naval Research under Grants N00019-89-J1519 and N00014-90-J1831.

REFERENCES

[1] S. D. Offsey, et al., "Strained-layer InGaAs-GaAs-AlGaAs graded-index separate confinement heterostructure single quantum well lasers grown by molecular beam epitaxy," Appl. Phys. Lett., **54**, 2527 (1989).

[2] I. Suemune, et al., "Extremely wide modulation bandwidth in a low threshold current strained quantum well laser," Appl. Phys. Lett., **53**, 1378 (1988).

[3] A. Larsson, et al., "A 980 nm pseudomorphic single quantum well laser for pumping erbium-doped optical fiber amplifiers," IEEE Photon. Tech. Lett., **2**, 540 (1990).

[4] S. E. Fischer, et al., "Long-lived InGaAs quantum well lasers," Appl. Phys. Lett., **54**, 1861 (1989).

[5] R. G. Waters, et al., "Inhibited dark-line defect formation in strained InGaAs/AlGaAs quantum well lasers," IEEE Photon. Tech. Lett., **2**, 531 (1990).

[6] G. P. Agrawal and N. K. Dutta, *Long-Wavelength Semiconductor Lasers*, New York: Van Nostrand Reinhold, 56 (1986).

[7] Y. C. Chen, et al., "Suppression of defect propagation in semiconductor by pseudomorphic layers," to be published in J. Appl. Phys, **74**, Sept. (1993).

[8] J. P. Hirth and J. Lothe, *Theory of Dislocations*, New York: McGraw-Hill, Chap. 2, (1968).

[9] J. Singh, *Physics of Semiconductors and Their Heterostructures*, New York: McGraw-Hill, Chap. 7, (1993)

FABRICATION AND PERFORMANCE OF A MONOLITHIC OPTOELECTRONIC TRANSISTOR

K.B. Nichols, B.F. Aull, P.A. Maki, S.C. Palmateer, E.R. Brown, T.A. Lind,
B.F. Gramstorff, Massachusetts Institute of Technology

PAPER UNAVAILABLE FOR PUBLICATION

0-7803-0894-8/93 $3.00

Inversion Channel HFET with unity current gain frequency of 14GHz and Surface Emitting Laser from a Single Epitaxial Growth

P.A.Kiely, P.A.Evaldsson, P. Cooke and G.W.Taylor
AT&T Bell Laboratories, Holmdel, N.J. 07733 and Geo Centers Fort Monmouth

Abstract

High speed capability up to 14GHz is reported for the inversion channel HFET utilizing the inversion channel device structure. The transistor is fabricated from the same epitaxial growth and with the same fabrication technology as previously used to demonstrate the vertical cavity Double Heterostructure Opto-Electronic Switching (DOES) Laser.

1. INTRODUCTION

The interest in optoelectronic integration has recently increased significantly with the improvement in performance of vertical cavity surface emitting lasers (VCSEL)[1-4]. In order to fulfill the expectations of these technologies to form smart pixel arrays[5] and high density integrated circuits with optical inputs and outputs, the surface emitting laser growth and fabrication technology must also be capable of producing a high performance transistor. In addition, to achieve significant levels of integration and thus functionality, a FET logic base is desired since the power consumption of bipolar circuits relegates them to relatively low levels of integration. The first monolithic integration of an FET and a VCSEL was demonstrated recently by Lockheed[6], where a MESFET was formed on top of the VCSEL using the highly resistive multilayer top reflector as an isolation layer.

The inversion channel technology[7] is a recently reported approach to optoelectronic integration which provides a way of merging a Heterostructure Field Effect Transistor (HFET) and a VCSEL structure such that high performance from both components is achieved. The central feature of the technology is an inversion channel formed by modulation doping of a heterostructure interface which is produced in a single epitaxial growth. The quantum wells form the channel for the FET and simultaneously the active region of the laser. One fabrication sequence is able to produce both components. Recently[8] the Double Heterostructure Opto-Electronic Switching (DOES) laser was reported in which a dielectric stack

0-7803-0894-8/93 $3.00

was used with post growth mode positioning to optimize the threshold current in the DOES VCSEL. The lasing characteristics of these devices are shown in figure 1. In this paper we report high speed performance of the HFET fabricated from the same material.

2. RESULTS

The material was grown by MBE on a semi-insulating GaAs substrate and the active region consisted of three 80Å strained $In_{0.2}Ga_{0.8}As$ quantum wells. The bottom mirror of the laser structure is a 27.5 period distributed Bragg reflector and the total thickness of the laser cavity was designed to be 5λ to incorporate the full inversion channel structure and allow HFET fabrication and electrical operation of the DOES without sacrificing switching parameters. The detailed doping and thickness parameters of the DOES structure including the barrier and charge sheet dopings have been reported earlier[8]. HFET fabrication began with implantation of Mg and P to create a highly doped p-type surface region. Tungsten was deposited and patterned using a tri-level technique to produce 1μm gate features. Note that the Tungsten gate forms an ohmic contact to the semiconductor. The semiconductor was then etched using Reactive Ion Etching (RIE) to within 1500Å of the quantum wells. Then Si was implanted (75 keV, $2\times10^{13}cm^{-2}$, 150 keV, $4\times10^{13}cm^{-2}$) to form the source and drain regions. The implants were activated using a Rapid Thermal Anneal at 950°C for 10s. Souce and drain contacts were formed by lift-off of AuGe/Ni/Au. A well is etched using RIE and a collector contact is formed in this region by lifting off Cr/AuZn. The contacts were alloyed at 450°C for 12s. The transistors were isolated by hydrogen implant to the semi-insulating substrate in order to eliminate parasitic capacitances. A dielectric was deposited to allow for metal cross-overs: this can be the same dielectric as was used for the top reflector in the DOES VCSEL[8], but in this case it is cured Polyimide. Holes were patterned in the Polyimide and bond pads were then lifted off to allow for high speed testing.

The output characteristics of 1×50μm gate length HFETs are shown in figures 2a,b. The threshold voltage was +0.4V (enhancement) which is ideal for enhancement/depletion logic applications. The transconductance (g_m) in figure 2a exhibits a maximum of 130mS/mm at a gate voltage of 1.2V. The gate conduction at 1.5V on the gate is 0.04mA (< 1mA/mm), which is negligible compared to the

maximum drain current of 160mA/mm as seen in figure 2b. This voltage for gate conduction agrees well with the predicted value[9]. The drain breakdown voltage for the HFETs was about 10V. The h21 parameter for the HFETs were measured using an HP8510 network analyzer and the results are shown in figure 3 indicating an uncorrected unity current gain cut-off frequency of 14GHz. This compares well with measurements of HEMTs and MESFETs of similar gate lengths[10]. To further increase the circuit speed of the HFET the drain capacitance must be reduced. This is achieved by reducing the collector doping. Due to Be doping in the p-doped active layer of the DOES and the subsequent RTA some Be doping in the collector is unavoidable but may be reduced considerably by novel p doping techniques in the MBE. The dc parameters show good matching to the DOES VCSEL parameters. The laser threshold is 5mA for a 14um diameter device and delivers an output power of 1.5mW for 3mA above threshold. Therefore, a 1μm HFET of 50μm width would provide ample drive current (8mA) for the laser. Reducing the dimensions of the laser would allow even smaller HFET dimensions.

In summary a f_t of 14GHz in the inversion channel HFET has been obtained from the same epitaxially grown material as a switching laser having a 5mA threshold current, a switching voltage of 10V, and a holding current of 0.5mA. Both devices using essentially identical processing sequences.

REFERENCES

[1] A. Scherer, J. L. Jewell, M. Walther, J. P. Harbison, and L. T. Florez, "Fabrication of Low Threshold Voltage Microlasers," Electron. Lett., 1992, 28, (13), pp. 1224-1226.

[2] K. L. Lear, S. A. Chalmers, K. P. Killeen, "A Very Low Voltage and Current Density Threshold Vertical Cavity Surface-Emitting Laser," IEEE LEOS meeting Boston 1992, Post-Deadline paper PD3.

[3] M. G. Peters, F. H. Peters, D. B. Young, J. W. Scott, B. J. Thibeault, and L. A. Coldren, "High Wallplug Efficiency Vertical-Cavity Surface-Emitting Lasers using Lower Barrier DBR Mirrors," Electron. Lett., 1993, 29, (2), pp. 170-172.

[4] R. A. Morgan, G. D. Guth, M. W. Focht, M. T. Asom, K. Kojima, L. E. Rogers, and S. E. Callis, "Transverse Mode Control of Vertical-Cavity Top-Surface-Emitting Lasers," IEEE Photon, technol. Letters, 1993, 4, (5), pp. 374-377.

[5] J. Cheng, P. Zhou, S. Z. Sun, S. Hersee, D. R. Myers, J. Zolper, and G. A. Vawter, "Surface-Emitting Laser-Based Smart Pixels for Two-Dimensional Optical Logic and Reconfigurable Optical Interconnections," IEEE J. Quantum Electron., 1993, 29, (2), pp. 741-756.

[6] Y. J. Yang, T. G. Dziura, T. Bardin, S. C. Wang, and R. Fernandez, "Monolithic integration of a vertical cavity surface emitting laser and a metal semiconductor field effect transitor," Appl. Phys. Lett., 1993, 62, (6), pp. 600-602.

[7] G. W. Taylor, P. A. Evaldsson, P. A. Kiely, T. Vang, P. R. Claisse, S. L. Daryanani, D. P. Docter, S. K. Sargood, and P. W. Cooke, "Integrated Inversion Channel Optoelectronic Devices and Circuit Elements for Multifunctional Array Applications," IEEE J. Quantum Electron., 1993, 29, (2), pp. 785-800

[8] P. A. Evaldsson, G. W. Taylor, P. W. Cooke, S. K. Sargood, P. A. Kiely, and D. P. Docter, "A High-Efficiency Vertical-Cavity Surface-Emitting Switching Laser Fabricated with Post Growth Cavity Mode Positioning," IEEE Photon. Techn. Lett., June 1993,

[9] G. W. Taylor and P. A. Kiely, "Theoretical and Experimantal Results for the Inversion Channel Heterostructure Field Effect Transistor," IEE Proceedings G, To be published September 1993.

[10] A. A. Ketterson, W. T. Masselink, J. S. Gedymin, J. Klem, C-K. Peng, W. F. Kopp, H. Morkoc, K. R. Gleason, "Characterization of InGaAs/AlGaAs Pseudomorphic Modulation-Doped Field-Effect Transistors," IEEE Trans. Electron Dev., 1986, ED-33, (5), pp. 564-571.

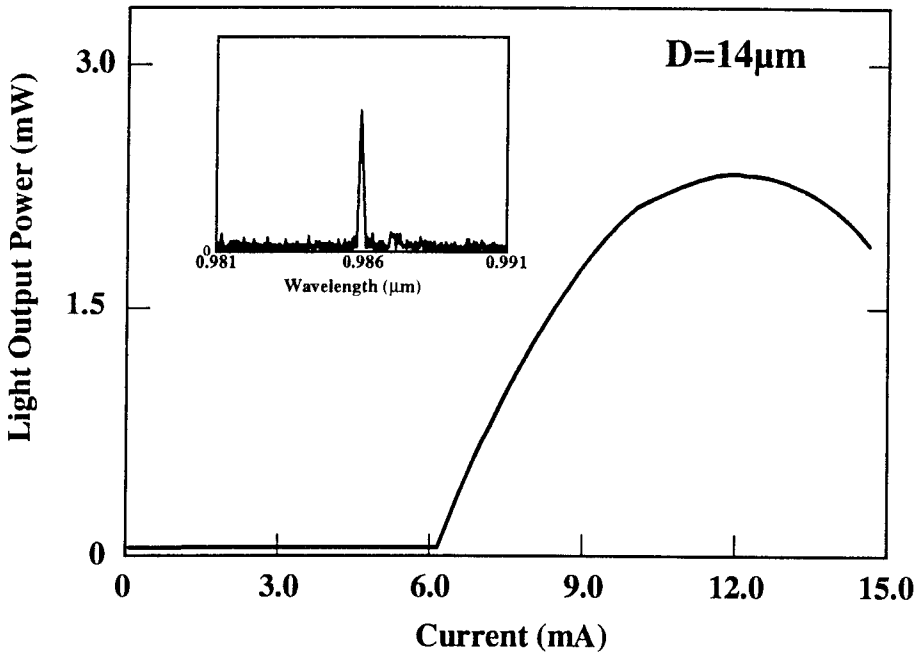


Figure 1 Lasing characteristics o a 14μm diameter DOES VCSEL

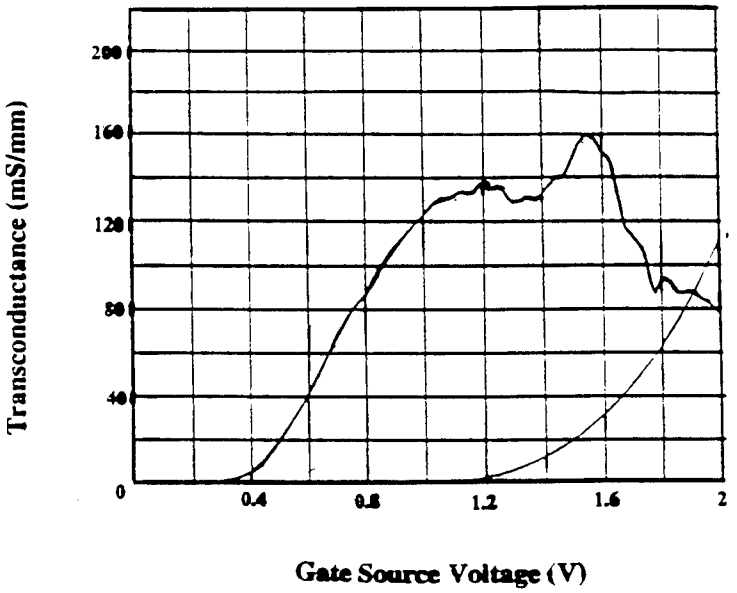


Figure 2a Transconductance variation with gate voltage of a 1μmx50μm HFET

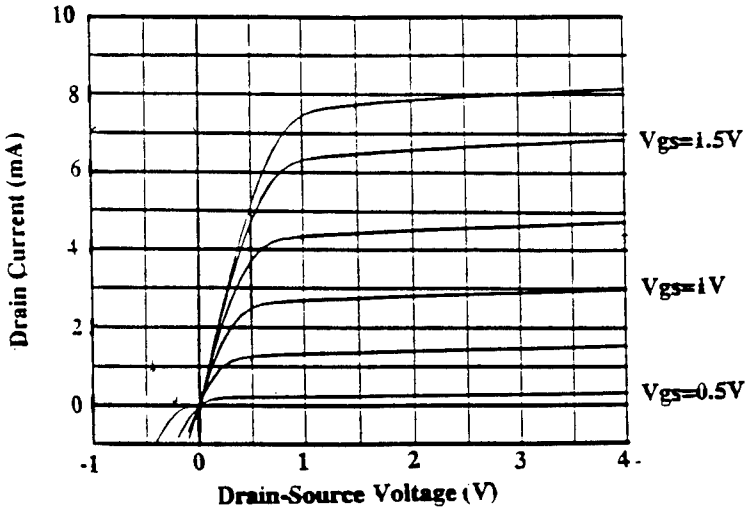


Figure 2b Drain output characteristics of a 1μmx50μm HFET

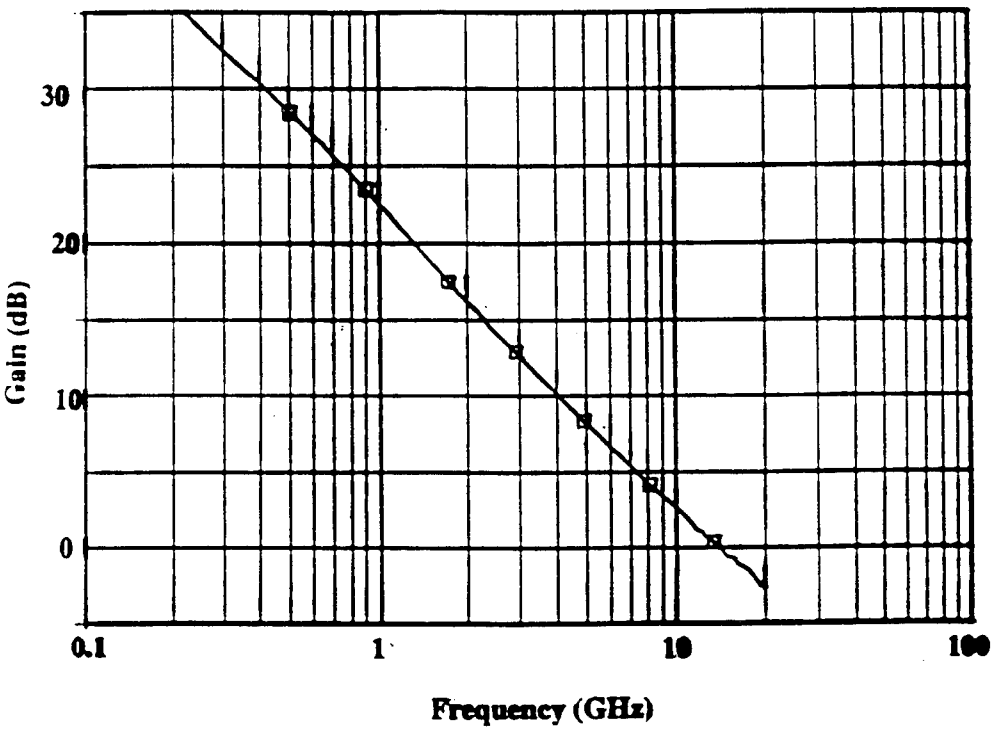


Figure 3 Current gain versus frequency for a 1μm HFET

Direct Optical Injection Locking of InP-based MODFET and GaAs-based HBT Oscillators

M. Karaküçük, W. Li, D. Yang, P. Freeman, J. East,
G.I. Haddad and P. Bhattacharya
Solid State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109-2122

ABSTRACT

We have investigated direct optical injection locking and tuning of high frequency oscillators made with GaAs/AlGaAs Heterojunction Bipolar Transistors (HBT's) and 0.25 μm gate InGaAs/InAlAs Modulation Doped Field Effect Transistors (MODFET's). We have developed an HBT technology using transparent indium-tin-oxide (ITO) emitter contacts for convenient optical access. Optical injection locking and tuning experiments have been performed on 6 GHz microwave oscillators made with the HBT's. Locking ranges up to 2.5 MHz and tuning ranges up to 25 MHz have been measured with the injection of the optical RF power at 30 dB below the oscillator power level. Similarly, direct subharmonic optical injection locking at 10 and 19 GHz has been achieved with the MMIC InGaAs/InAlAs MODFET oscillators.

1 Introduction

The control of microwave circuits by optical signals has received wide attention because of the wide bandwidth of the optical control signals and the inherent isolation from RF signals. Optical injection locking of microwave oscillators is particularly promising for applications in phased-array radars and cellular communication. Optical control of GaAs FET oscillators has already been reported[1]. Recently, we demonstrated direct optical injection locking of GaAs/AlGaAs HBT oscillators[2] and InGaAs/InAlAs MODFET oscillators[3]. The FET devices have the advantages of easier light coupling through the gate-source and gate-drain openings and their mature device technology. On the other hand, HBT devices have large internal optical gain and, in addition, their superior $1/f$ noise behavior is especially attractive for low phase noise microwave oscillators. However, it is difficult to couple light into the HBT devices because of the self-aligned metallizations currently used for microwave HBTs. Although lateral coupling of light via an on-chip waveguide is possible, this would require a second epitaxial growth and extra processing steps.

In this paper, we describe a new HBT technology using transparent indium-tin-oxide (ITO) emitter contacts for convenient optical access. Microwave oscillators with frequencies of 6 - 10 GHz have been made using these HBTs and direct optical injection

0-7803-0894-8/93 $3.00

locking experiments have been performed at 6 GHz. Experimental results of optical injection locking of MMIC InGaAs/InAlAs MODFET oscillators are also presented.

2 Experimental Results

2.1 Optical injection locking of GaAs/AlGaAs HBT oscillators with transparent ITO emitter contacts

Indium Tin Oxide (ITO) is optically transparent (T $\geq$ 90%) with negligible absorption and and has a conductivity of $3 \times 10^3 \Omega^{-1}\text{cm}^{-1}$ at room temperature. ITO films have been extensively studied for use in solar cells[5][6], and recently, photodetectors[7][8]. In most of the previously reported applications, a Schottky contact was formed between the ITO and the semiconductor. We have developed and optimized an ITO ohmic contact for both n-type and p-type GaAs with typical specific contact resistances of $1 \times 10^{-6}\Omega\text{cm}^{-2}$ and $2 \times 10^{-7}\Omega\text{cm}^{-2}$, respectively, which are comparable with those of alloyed metal contacts. The simplicity and the reliability make the ITO contact attractive for optoelectronic device applications.

The ITO film was prepared by RF magnetron sputtering of a nominally 91 mol.% In_2O_3 : 9 mol.% SnO_2 composite target onto unheated substrates. Post-deposition rapid thermal annealing (RTA) in Ar was used to remove structural defects and to improve the electrical and optical characteristics. For a typical 2500Å thick ITO film, a sheet resistance $R_H \sim 2.5\Omega/\square$ and transmittance of T= 90% at λ = 850nm have been achieved. A lift-off technique using photoresist and optical lithography was employed to define sputtered ITO patterns as small as 1.0μm in size. Normally, ITO forms a Schottky-like rectifying contact with p- or n-type GaAs. To circumvent this, a 50 - 100Å layer of either silver or indium was deposited before sputtering the ITO. This pre-layer, which is also optically transparent, serves as an intermediary layer that helps form low resistance ohmic contacts. Indium is preferentially used in ITO-n-GaAs contacts. After the post-deposition annealing, it forms a thin layer of InGaAs which readily forms an excellent ohmic contact with ITO. A contact resistance as low as $2 \times 10^{-7}\Omega\text{cm}^{-2}$ was obtained with 50Å In and 2500Å ITO on Si-doped GaAs ($n = 5 \times 10^{18}\text{cm}^{-3}$). Silver is used in ITO-p-GaAs contact to prevent formation of a p-n junction between the p-GaAs and the ITO film which is inherently an n-type semiconductor.

The HBT structure, shown in Table 1, was grown by molecular beam epitaxy (MBE). Self-aligned HBT devices were fabricated. Device isolation was provided by single dose, high energy O^+ implantation (5 MeV and $O^+ = 1.7 \times 10^{15}$ cm^{-2}) using 3 μ m thick electro-plated Au masks and appropriate annealing (RTA 560°C in Ar for 20 seconds). After removing the Au masks, the sample was patterned with photoresist to delineate the emitters, then a 60Å thick indium layer was deposited by evaporation, followed by 2500Å of ITO deposited by RF sputtering. $2 \times 20\mu\text{m}^2$ emitters were defined by the subsequent lift-off process. The sample was then RTA annealed at 450°C in Ar for 20 seconds to form the ohmic contacts. ITO was only used for emitter contacts for two reasons: First, the contact resistance between ITO and n-type GaAs is lower than that of p-GaAs, and secondly, base resistance is a primary factor in determining the high

Layer	Composition	Thickness (μm)	Type	Doping (cm^{-3})
Cap	GaAs	0.02	n^+	5.0×10^{18}
Emitter	$Al_{0.25}Ga_{0.75}As$	0.30	n	5×10^{17}
Base	GaAs	0.10	p^+	1.0×10^{19}
Collector	GaAs	0.60	n	2×1016
Subcollector	GaAs	0.50	n^+	5.0×10^{18}
Substrate	GaAs	300	-	S.I.

Table 1: Layer structure of the AlGaAs/GaAs heterojunction bipolar transistor

frequency performance of the HBT as,

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_B C_{BC}}} \tag{1}$$

where f_T is the current gain cutoff frequency, R_B the base resistance, and C_{BC} the base-collector capacitance. For these reasons, a standard metallic ohmic contact was used for the base.

Base contacts were formed by wet-etching using the patterned photoresist and the ITO emitter as masks followed by evaporation of Zn/Pd/Au and lift-off. The emitter ITO was covered with photoresist to block metal deposition during this process. The collector contacts were made using a similar etching and lift-off process. Finally, coplanar transmission line patterns for microwave probing were formed using electro-plating, and connection to the devices was completed with an air-bridge. HBT's with emitters of $2 \times 10\mu m^2$ and interdigitated base contacts were fabricated. A photomicrograph of the fabricated HBT is shown in Figure 1.

Common-emitter I-V measurements show a uniform current gain (h_{fe}) of 10 - 20. Gummel plot measurements show junction ideality factors of 1.01 and 1.43 for the base-collector and emitter-base junctions, respectively. Microwave performance of the devices was measured from 45 MHz - 26.5 GHz using an HP-8510 network analyzer. The short circuit current gain ($|H_{21}|$) and the maximum available power gain (G_{max}) were computed from the measured S-parameters. As shown in Fig. 2, f_T = 18 GHz and f_{max} = 20 GHz were obtained at I_C = 10mA and V_{ce} = 3V

To characterize the photoresponse of the HBTs, short optical excitation pulses from a mode-locked dye laser (λ = 853nm) pumped with a Nd:YAG laser were used. The nominal pulse width is $\sim$ 7ps and the repetition rate is 76MHz. The response of the phototransistor at V_{ce} = 4V and $I_b = 50\mu$A has a full width at half maximum (FWHM) of 80ps, as shown in Fig. 3. Taking into account the optical pulse width and the S4 Tektronix sampling head response (FWHM=35ps), the intrinsic device response time is $\cong$ 60ps.

Microwave oscillators were made using these devices. The oscillator circuit was built on a 15 mil thick Duroid substrate using microstrip lines for impedance matching. The

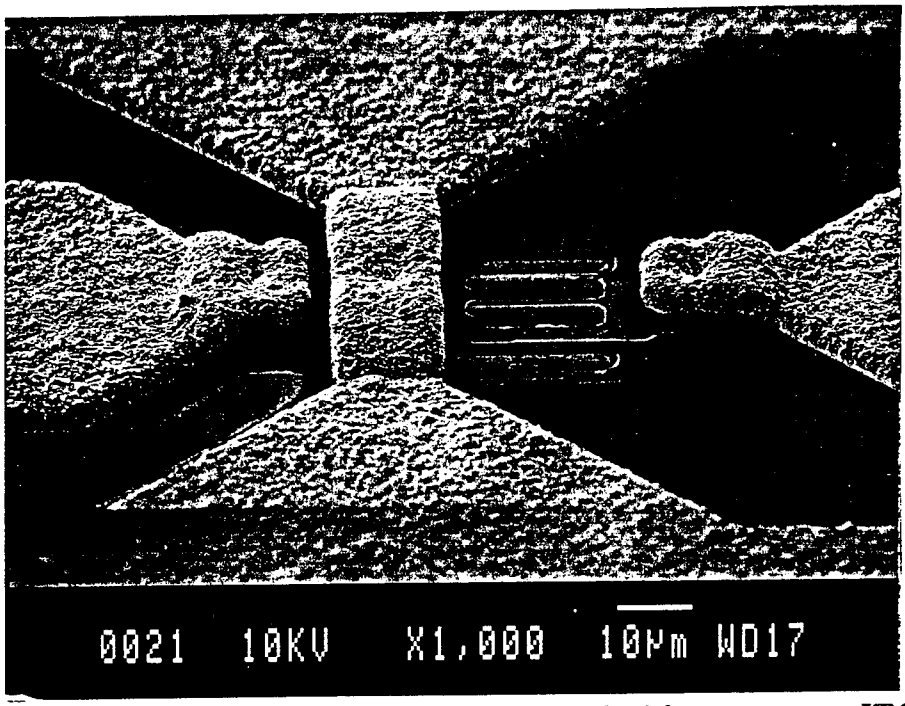


Figure 1: Microphotograph of the fabricated HBT with transparent ITO emitter ohmic contact.

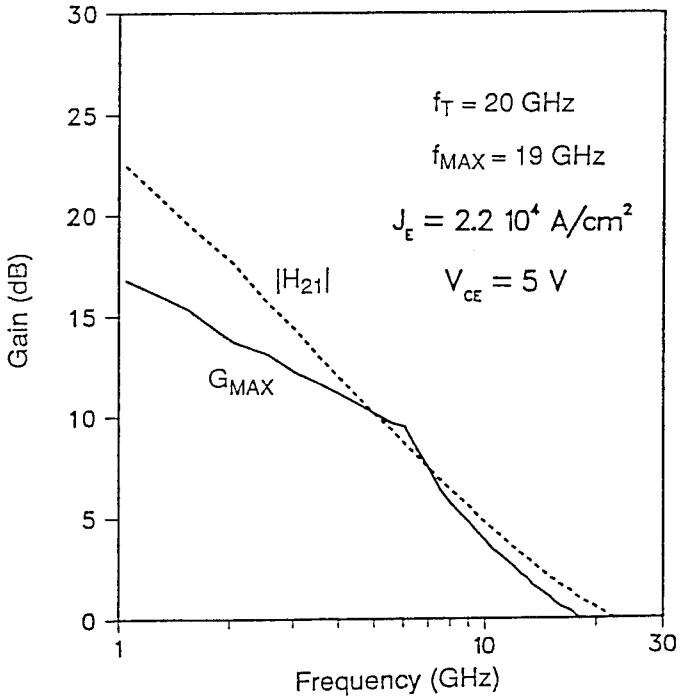


Figure 2: Measured current gain (h_{21}) and power gain (G_{max}) as a function of frequency for HBT with ITO emitter contact.

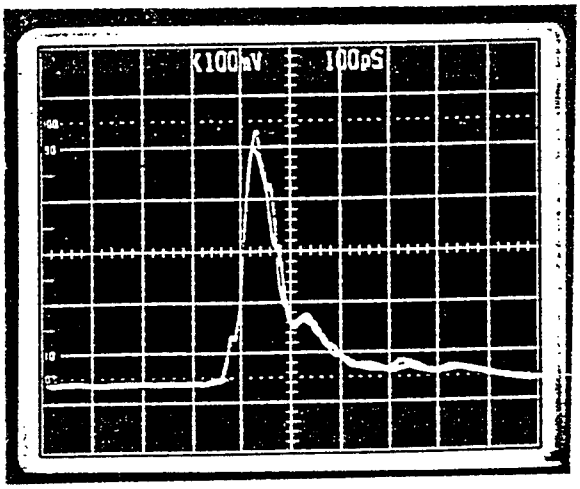


Figure 3: Pulse response of the GaAs/AlGaAs HBT with transparent ITO contact at $V_{ce} = 4$V and $I_b = 50\mu$A. The excitation pulses ($\lambda = 853$nm) have a nominal pulse width of 10 ps.

transistor was wire bonded onto the circuit. The active device was biased externally through bias tees at $V_{CE} = 5$ V and $I_C = 4.5$ mA for maximum power output. Output power up to 10 mW was obtained at the oscillation frequency of 5-8 GHz.

An AlGaAs laser diode ($\lambda = 830$ nm) was used as the light source for direct optical tuning and injection locking experiments. The average optical power was adjusted by a DC current source. RF power from a synthesizer was superimposed over the DC bias current to produce modulated light output. The modulated light was focused directly onto the transparent ITO emitter contacts of the HBT. The oscillator output spectra were monitored and recorded with a spectrum analyzer. When the oscillator was unlocked but the frequency of the injected signal ($f_{inj.}$) was close to the oscillating frequency ($f_{opr.}$), the oscillator output spectra contained many single-sided sidebands. This phenomena is very similar to that observed in the case of microwave injection. When $f_{inj.}$ was tuned to within the locking range, all sidebands disappeared and the phase noise was significantly reduced, as expected for an injection-locked oscillator. The locking range as a function of RF generator output power is also plotted in Fig. 4. In this experiment, a locking range 2.5 MHz was obtained. This locking range is mainly limited by the available optical RF power. Due to the limitation of the high frequency response of the laser diode, the injected optical RF power is about 30 dB lower than the oscillator output. We also performed a microwave injection locking experiment on the same oscillator and a locking range up to 10 MHz has been obtained at the injection power of -10 dBm. In both cases, locking range as a function of RF generator power agrees with the theoretically predicted square-root dependence.

For optical tuning, CW light from the laser diode was directly focused onto the HBT. The optical power was varied by changing the DC bias current of the laser. The oscillator frequency always increased as the optical power increased, and a frequency tuning range of up to 25 MHz was measured using this technique.

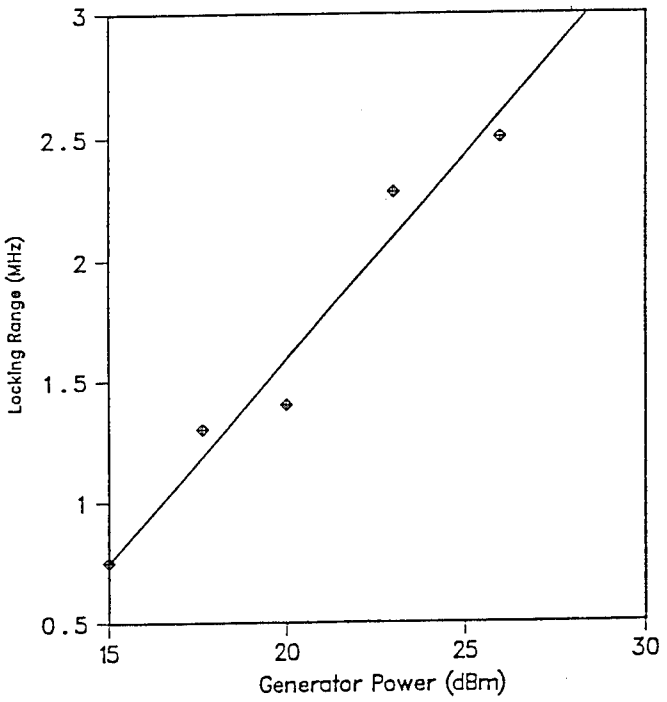


Figure 4: Locking range as a function of RF generator output power.

2.2 Direct optical injection locking of MMIC $In_{0.53}Ga_{0.47}As$/ $In_{0.52}Al_{0.48}As$ MODFET oscillators

The $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As/InP$ material system is attractive for optoelectronic applications because the circuits can operate at wavelength rangings from 1.3 to 1.7μm. Pseudomorphic InGaAs/InAlAs MODFET's have demonstrated excellent microwave performance. In addition, X and W-band MMIC oscillators using pseudomorphic MODFET's have been demonstrated. We achieved direct optical subharmonic injection locking of an InGaAs/InAlAs MODFET-based MMIC oscillator for the first time.

Lattice-matched $In_{0.53}Ga_{0.47}As$/ $In_{0.52}Al_{0.48}As$ MODFET structures were grown by molecular beam epitaxy (MBE) on semi-insulating InP substrates. An InGaAs/ InAlAs superlattice was grown first, followed by a 3000Å undoped InAlAs. The channel consists of 550Å undoped $In_{0.53}Ga_{0.47}As$. This is followed by a 50Å undoped InAlAs spacer, a 150Å n^+-InAlAs layer Si-doped at $2.0 \times 10^{18}cm^{-3}$, and a 250 Å undoped InAlAs barrier layer. Finally, a 150 Å n^+-InGaAs cap layer was grown to facilitate ohmic contact formation.

The oscillator design employs a comon-source circuit configuration. The negative resistance is achieved by parallel inductive and capacitive feedback elements. The MMIC oscillators were fabricated using the following steps: mesa isolation, source-drain ohmic contacts formation, gate definition, dielectric film deposition, interconnection, and air-bridge formation. The MODFET's have 0.25μm gate stripes which were defined using e-beam lithography. Two oscillator were used in the injection locking measurements.

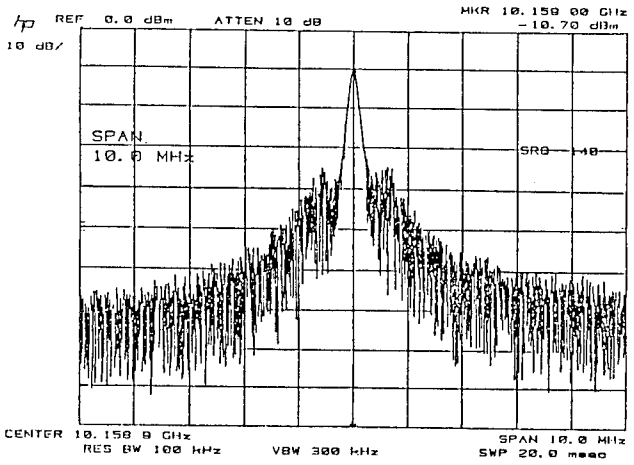


Figure 5: Injection locked spectrum of 10.159 GHz oscillator.

The oscillation frequency and output power were adjusted by tuning the drain and gate biases. Oscillator A was biased at $V_g = 0$V and $V_d = 0.308$V. The oscillation frequency was 10.164 GHz and the output power was -12.8 dBm. A maximum output power of 1.2 dBm can be obtained at higher drain bias. Oscillator B oscillates at 9.523 GHz and 19.034 GHz with power levels of -14.3 dBm and -29.7 dBm, respectively. Both oscillators were optically injection locked at the fundamental and subharmonic frequencies. The nonlinearity in the laser diode and the MODFET were used to obtain frequency multiplication.

A high-speed Ortel SL1020 AlGaAs laser diode ($\lambda = 840$ nm) modulated by a microwave signal source was used as light source for the injection locking experiment. For the injection locking of oscillator A, the laser diode was modulated at 4 dBm. The injection locked spectrum of the oscillator at 10.159 GHz is shown in Fig. 5. The stability of the oscillator was improved and the sideband FM noise reduced. The fundamental and subharmonic locking range and FM noise degradation are summarized in Table 2. An optimum locking range of 4.8 MHz was found at a subharmonic factor of 1/2. Locking range was increased for subharmonic injection locking. The FM noise reduction at 100kHz offset the peak is about -10 dBc in the injection locked oscillator output compared with that of free-runing oscillator. The locking range decreased when subharmonic factor decreased below 1/6 due to comparatively lower modulation power. For oscillator B, injection locking was achieved at both oscillation frequencies. Only subharmonic modulation signals were used for 19 GHz injection locking since the laser operation frequency was limited by its relaxation oscillation frequency. The subharmonically injection-locked spectrum of the oscillator at 19.033GHz is shown in Fig. 6. The subharmonic factor was 1/4. A summary of measured data is shown in Table 3.

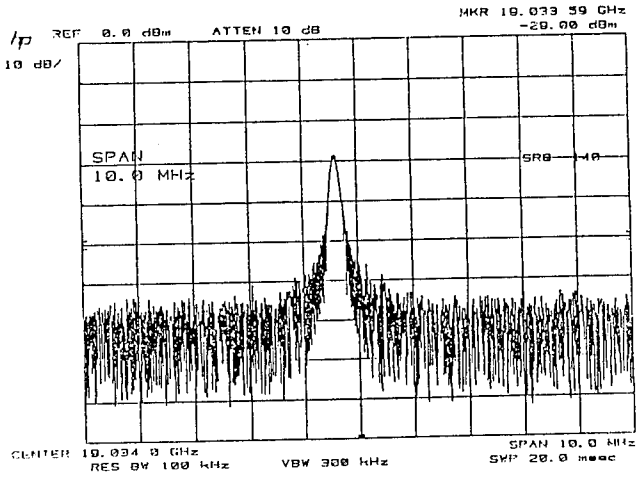


Figure 6: Subharmonicaly injection locked spectrum of MODFET oscillator at 19.03 GHz.

Subharmonic factor	Frequency (GHz)	Locking range (MHz)	FM noise reduction at 100 kHz offset (dBc))
1	10.160	3.0	-10.0
1/2	5.080	4.8	-11.0
1/3	3.387	3.6	-11.3
1/4	2.540	2.8	-9.7
1/6	1.571	0.5	-6.0

Table 2: Locking Range and FM Noise Reduction of the Optical Subharmonic Injection Locking of Oscillator A at 10.159 GHz

Subharmonic factor	Frequency (GHz)	Locking range (MHz)	FM noise reduction at 100 kHz offset (dBc))
1/2	9.517	1.6	-11.0
1/4	4.758	1.6	-10.0
1/6	3.72	1.8	-9.3

Table 3: Locking Range and FM Noise Reduction of the Optical Subharmonic Injection Locking of Oscillator B at 19.033 GHz

3 Conclusion

Two different schemes for optically controlled integrated microwave oscillators have been demonstrated using GaAs/AlGaAs heterojunction bipolar transistors (HBTs) and InGaAs/InAlAs MODFETs. We have developed an optical GaAs/AlGaAs HBT technology using transparent indium-tin-oxide (ITO) emitter contacts for convenient optical access. Optical tuning and injection locking experiments have been performed on 6 GHz microwave oscillators made of these devices. Tuning ranges up to 25 MHz and locking ranges up to 2.5 MHz have been demonstrated with the injection of optical RF power at 30 dB below the oscillator power level. Similarily, monolithically integrated 0.25μm-gate InGaAs/InAlAs MODFET oscillators demonstrate direct optical injection locking at 10 GHz and 19 GHz.

Acknowledgements

The authors would like to thank Dr. Victor Rotberg of the Michigan Ion Beam Laboratory for performing the high energy ion implantation and Dr. A. Paolella and Dr. D. Sturzebecher at U.S. Army LABCOM for their help during some of the measurements. providing the measurement and useful discussion. This work is supported by the Advanced Research Project Agency under contract F30602-92-C-0087.

References

[1] R.D. Esman, L. Goldberg, and J.F. Weller, "Optical Phase Control of Optically Injection-Locked FET Microwave Oscillator," *IEEE TMTT,* **37**, pp. 1512-1518, Oct. 1989.

[2] M. Karaküçük, W.Q. Li, P.N. Freeman, J.R. East, G.I. Haddad and P.K. Bhattacharya,"A Direct Optically Injection Locked 2.6 GHz HBT Oscillator," *Optical and Microwave Technol. Lett.*, **6**, pp. 609-611, August 1993.

[3] D. Yang, P.K. Bhattacharya, and T. Brock, "Direct optical injection locking of monolithically integrated $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ MODFET oscillators," submitted for publication to *Electron. Lett.*, April 1993.

[4] K.L. Chopra, S. Major, and D.K. Pandya, "Transparent Conductors — a status review," *Thin Solid Films*, **102**, pp. 1-46, 1983.

[5] S. Nassen, and T.J. Coutts, "Indium tin oxide/gallium arsenide solar cells," *J. Appl. Phys.*, **58**, pp. 4463-4463, Dec. 1985.

[6] X. Li, M.W. Wanlass, T.A. Gessert, K.A. Emery, and T.J. Coutts, "High-efficiency indium tin oxide/indium phosphide solar cells," *Appl. Phys. Lett.*, **54**, pp. 2674-2676, June 1989.

[7] M. Zirngible, Y. Hu, R. Sachot, and M. Ilegems, "Characterization of a top-illuminated p-i-n diode with indium tin oxide contact," *Appl. Phys. Lett.*, **54**, pp. 2076-2078, 1989.

[8] J.-W. Seo, A.A. Ketterson, D.G. Ballegeer, K.-Y. Cheng, I. Adesida, X. Li, and T. Dessert, "A comparative study of metal-semiconductor-metal photodetectors on GaAs with indium-tin-oxide and Ti/Au electrodes," *IEEE Photon. Technol. Lett.*, **4**, pp. 888-890, Aug. 1992.

[9] S.J. Pearton, F. Ren, J.R. Lothian, T.R. Fullowan, A. Katz, P.W. Wisk, C.R. Abernathy, R.F. Kopf, R.G. Elliman, M.C. Ridgway, C. Jagadish, and J.S. Williams, "Use of MeV O^+ ion implantation for isolation of GaAs/AlGaAs heterojunction bipolar transistors," *J. Appl. Phys.* , **71**, pp. 4949-4954, May 1992.

HIGH SPEED MONOLITHICALLY INTEGRATED PIN-MODFET TRANSIMPEDANCE PHOTORECEIVERS

A.L. Gutierrez-Aitken, P. Bhattacharya, Y.C. Chen, D. Pavlidis, and T. Brock

Solid State Electronics Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109-2122

ABSTRACT

The performance characteristics of transimpedance and transimpedance/voltage photoreceivers using an $In_{0.53}Ga_{0.47}As$ p-i-n photodiode integrated with a 0.1 μm gate length regrown pseudomorphic $In_{0.60}Ga_{0.40}As$ MODFET realized by MBE have been investigated. The photodiode exhibits very low dark current and a responsivity of 0.6 A/W at λ=1.55 μm. The regrown MODFETs have extrinsic transconductance values as high as 610 mS/mm and channel currents up to 350 mA/mm at a drain bias of 1.5 V. The measured temporal response of the photoreceiver with a transimpedance amplifier exhibits a FWHM value of 90 ps and a transimpedance gain of 45 dBΩ with a 800 Ω feedback resistor. This photoreceiver demonstrated 24 dB optical-to-electrical conversion gain at 9.3 GHz. The photoreceiver with a transimpedance/voltage amplifier shows a FWHM of 200 ps and a transimpedance gain of 63 dBΩ with a 1.2 KΩ feedback resistor.

I. INTRODUCTION

High speed, high sensitivity photoreceivers are required for long-distance transmission and interconnection optical systems. Monolithically integrated photoreceivers are projected to have advantages because of their reduced parasitic elements, small size, compactness and high reliability[1]. The fabrication of a high performance monolithically integrated photoreceiver usually demands different optimized layer structures for the photodetector and the amplifying device. A variety of photoreceivers have been reported using either of two approaches to choose the devices to be integrated. One approach is to use the same structure for both devices[2]-[4], compromising to some extent the individual performance in order to have a compatible fabrication process. The second approach is to use independently optimized structures, which in general will require a more involved fabrication process but with the advantage of high integrated performance. One way to integrate independently optimized structures is to use single epitaxial growth with stacked structures where the epilayer consists of one structure on top of the other[5]-[7]. An alternative approach is to use two-step epitaxial growth, where the first structure is selectively etched away before the second structure is grown on top of the substrate[8],[9]. With this method, parasitic elements between different structures are eliminated and a more planar surface is obtained. However, the wafer preparation before the second epitaxial growth is critical for device and circuit performance.

0-7803-0894-8/93 $3.00 © 1993 IEEE

Many integrated front-end photoreceivers using a transimpedance configuration based on compatible or stacked structures have been reported but none have a regrown transimpedance amplifier. We report in this work the fabrication and characterization of InP-based monolithically integrated p-i-n photodiode and 0.1 μm gate length regrown MODFET photoreceiver using transimpedance and transimpedance/voltage amplifiers for the first time.

II. CIRCUIT DESIGN AND FABRICATION

Figure 1(a) shows the circuit diagram of the photoreceiver with a transimpedance amplifier. It consists of a p-i-n photodiode (D_1), a common source amplifier (Q_1) with its output dc coupled to a source-follower buffer/level-shift stage (Q_2) and an output buffer (Q_3) to provide 50 Ω matching output impedance. The use of a 800 Ω feedback resistor (R_F) allows a large bandwidth operation without limiting the sensitivity of the photoreceiver due to thermal noise. The design includes a 3.6 nH input inductor and a 4 pF capacitor in parallel with the level shifter resistor to improve the high frequency response. The simulated 3 dB bandwidths without and with L_1 and C_{LS} are 2.1 GHz and 10.8 GHz respectively. The inductor also reduces the effective noise contribution from R_F to the total output noise power density within the electrical bandwidth of the circuit. Figure 1(b) shows the circuit diagram of the photoreceiver with a transimpedance/voltage amplifier. This photoreceiver circuit incorporates a two-stage voltage amplifier between the transimpedance input stage and the output buffer. The feedback resistor, input inductor and capacitors are 1.2 KΩ, 4.6 nH and 1 pF respectively. The simulated 3 dB bandwidths for this circuit without and with L_1, C_{LS1}, C_{LS2} and C_{LS3} are 1.3 GHz and 3.6 GHz respectively.

Figures 2(a) and (b) show the photodiode and MODFET layer structures used in our design. The p-i-n photodiode with a 0.8 μm i-$In_{0.53}Ga_{0.47}As$ region sandwiched between $In_{0.52}Al_{0.48}As$ window and contact layers is grown first by solid-source molecular beam epitaxy. The thickness of the *i* region is optimized considering the responsivity of the photodiode, transit-time speed limitation and the 3 dB bandwidth related to the frequency response of the electrical circuit. The wafer is then patterned with a SiO_x mask and the p-i-n structure is selectively etched away down to the S.I. InP substrate. Next, the wafer is prepared and cleaned thoroughly before the 150 Å strained $In_{0.60}Ga_{0.40}As$ channel MODFET is regrown. The wafer is then processed conventionally to fabricate the photoreceiver circuits. Ohmic contacts to the source and drain regions of the MODFET and the n^+ layer of the photodiode were made by electron-beam evaporation of Ni/Ge/Au/Ti/Au. The p^+ ohmic contact for the photodiode was made by evaporating Pd/Zn/Pd/Au. The MODFETs have a drain-source spacing of 2 μm and a π-gate layout to reduce gate resistance. The 0.1 μm length T-gate is fabricated using a bilayer e-beam photoresist process to lift-off Ti(500 Å)/Au(2500 Å). The resistors are realized by evaporating and lifting-off a thin layer of Titanium (Ti) with sheet resistance of R_s = 20 $\Omega/\Box$. The capacitors are fabricated using the gate and the interconnection metallizations as planar electrodes separated by a ~1500 Å sputtered SiO_x dielectric layer. The inductors are realized using the interconnection metallization in a rectangular spiral layout. Finally, pillars and air-bridges are plated to complete the circuit.

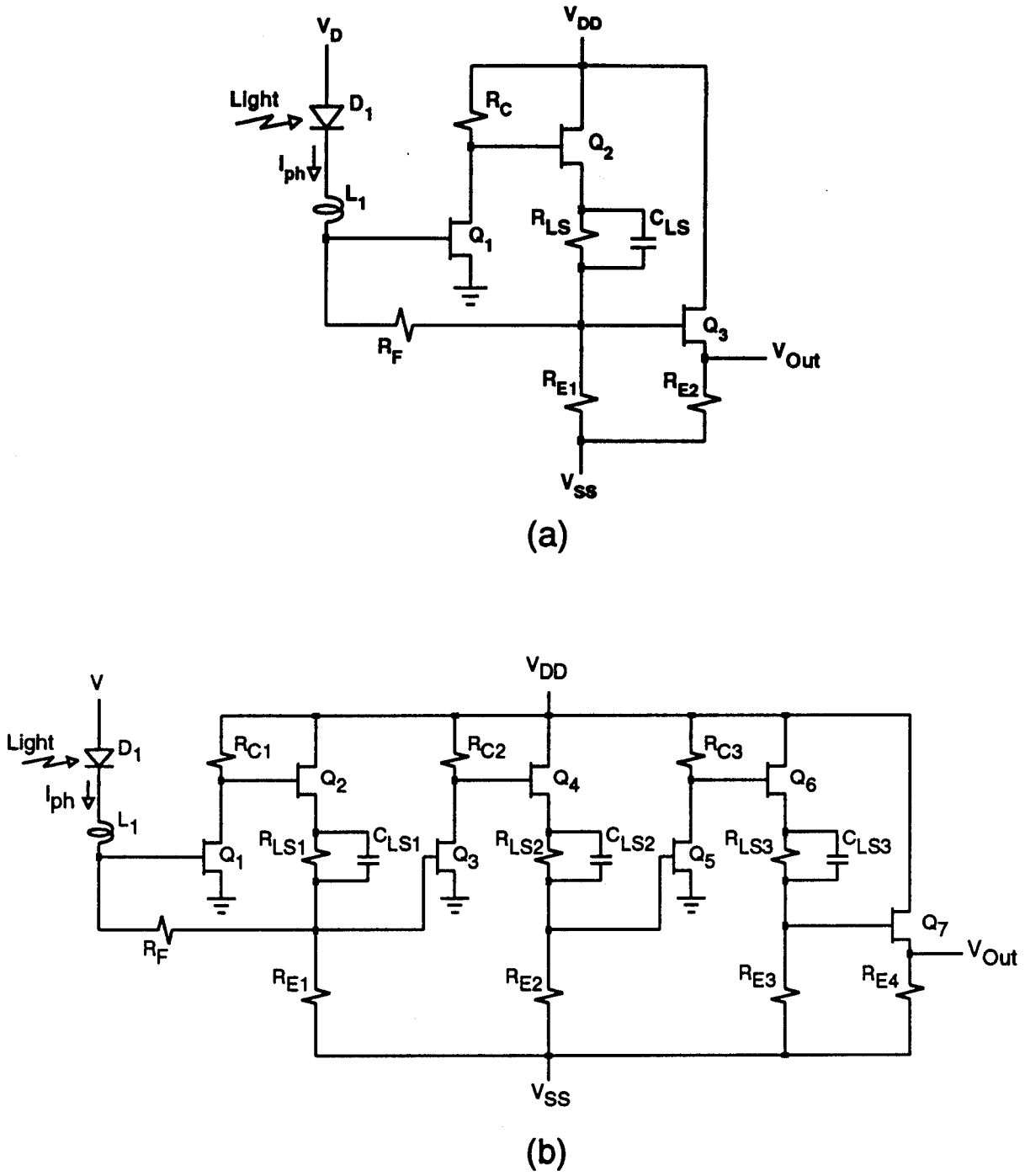


Figure 1: Circuit diagram of the (a) transimpedance, and (b) transimpedance/voltage photoreceivers.

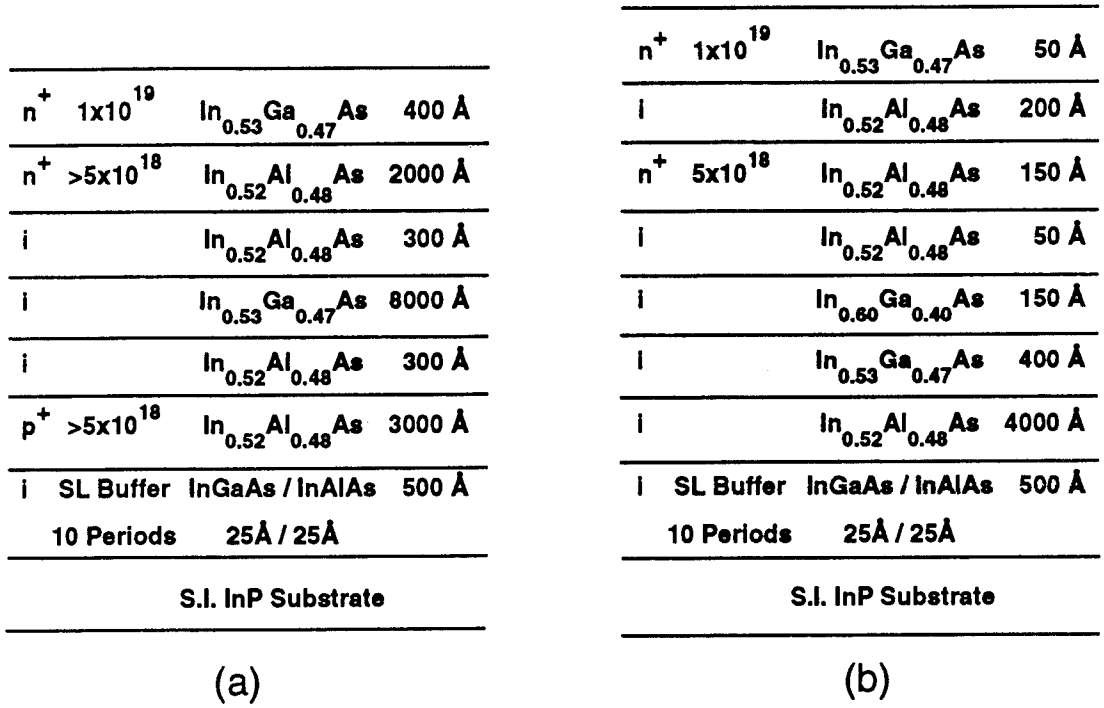


Figure 2: (a) PIN photodiode, and (b) MODFET structures showing the epitaxial layer characteristics.

III. MEASUREMENT AND EXPERIMENTAL RESULTS

For the resistors an average value of R_s = 24.6 Ω/□ is achieved for 600 Å of Ti. Values of 50 Ω to 2400 Ω are obtained consistently with tolerances of ±15 % across the sample. The capacitors show leakage currents of the order of 10 nA for voltage biases of ±1 V and the capacitance value is within ±20 % for DC voltage biases from -4 V to +4 V. The 630 μm^2 p-i-n photodiodes have breakdown voltages of 16-18 V, a leakage current of ≤35 nA at -10 V, and a responsivity of 0.6 A/W at λ=1.55 μm with antireflecting coating.

The I-V and transfer characteristics of a regrown 0.1 μm gate length MODFET are shown in Figs. 3(a) and (b) respectively. The peak extrinsic transconductance (g_m) of isolated as-grown MODFETs is 810 mS/mm with channel current densities up to 460 mA/mm at a drain bias of 1.5 V, whereas the regrown MODFETs have a peak transconductance of 610 mS/mm with current densities up to 350 mA/mm at a drain bias of 1.5 V. To our knowledge these are the highest g_m values reported for regrown 0.1 μm gate length InP-based MODFETs.

The response speed of the photoreceiver was measured using 15 ps (FWHM) pulses at λ=1.55 μm from a color center laser coupled to a single mode fiber. The as-cleaved fiber tip was placed vertically on top of the photodiode at a distance of ~4 μm using an optical probe. The circuits were diced and bonded to a coplanar waveguide test fixture. The photodiode was

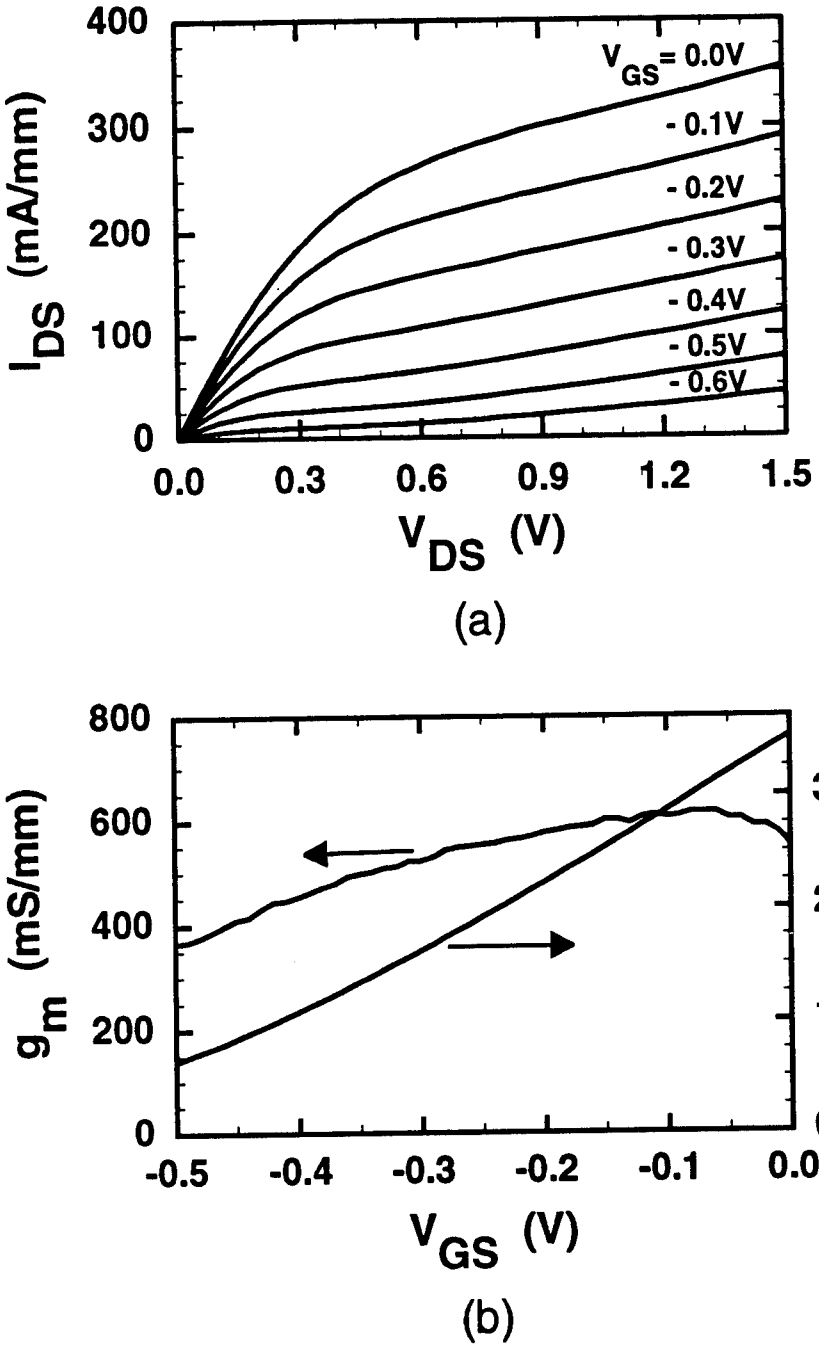


Figure 3: (a) Current-voltage, and (b) transfer characteristics of a 0.1 μm gate length regrown MODFET.

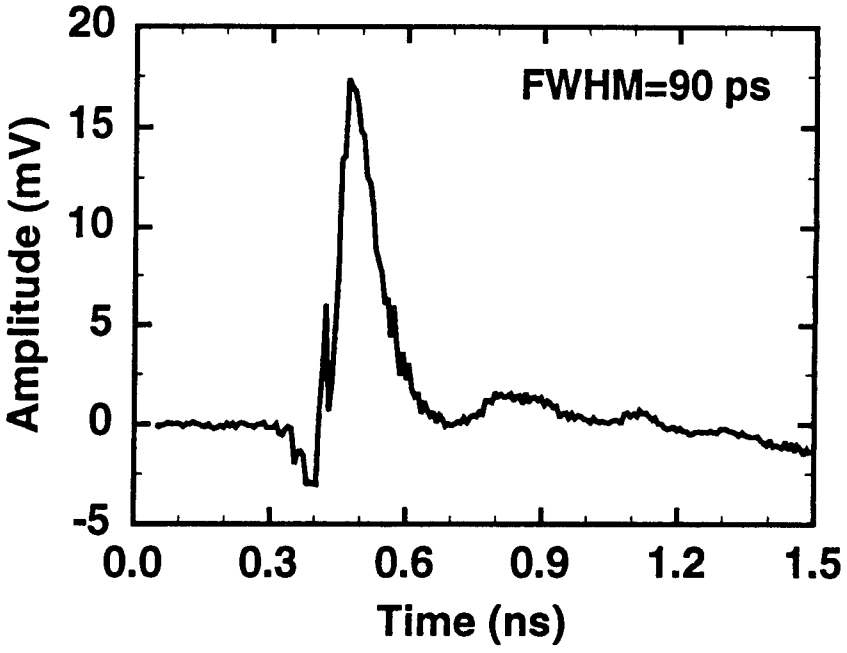


Figure 4: Measured temporal response of the photoreceiver with a transimpedance amplifier.

biased at -6 V using a microwave bias-T network. The output was taken using another bias-T to a sampling head of $\leq$25 ps rise time. The deconvolved impulse response of the photoreceiver with a transimpedance amplifier with bias optimized for speed is shown in Fig. 4. A FWHM of 90 ps and a rise time of 60 ps with a transimpedance gain of 45 dBΩ were obtained. This photoreceiver demonstrated 24 dB of optical-to-electrical conversion gain at 9.3 GHz.

The temporal response of the photoreceiver with a transimpedance/voltage amplifier is shown in Fig. 5. A FWHM of 200 ps and a rise time of 110 ps with a transimpedance gain of 63 dBΩ were obtained. The 3 dB bandwidth for this circuit, obtained from Fourier transformation of the temporal response, is 2.5 GHz.

IV. CONCLUSION

In conclusion, we have successfully fabricated and characterized high speed PIN-MODFET integrated photoreceivers with transimpedance and transimpedance/voltage amplifiers based on regrown 0.1 μm gate length MODFETs. The PIN photodiode showed leakage currents of $\leq$ 35 nA at -10 V, and a responsivity of 0.6 A/W. The regrown MODFETs demonstrated high transconductance with peak values of 610 mS/mm and current densities up to 350 mA/mm at a drain bias of 1.5 V. The measured temporal response of the photoreceiver with transimpedance amplifier showed a FWHM of 90 ps with a transimpedance gain of 45 dBΩ. This photoreceiver showed 24 dB of optical-to-electrical conversion gain at 9.3GHz. The photoreceiver with a

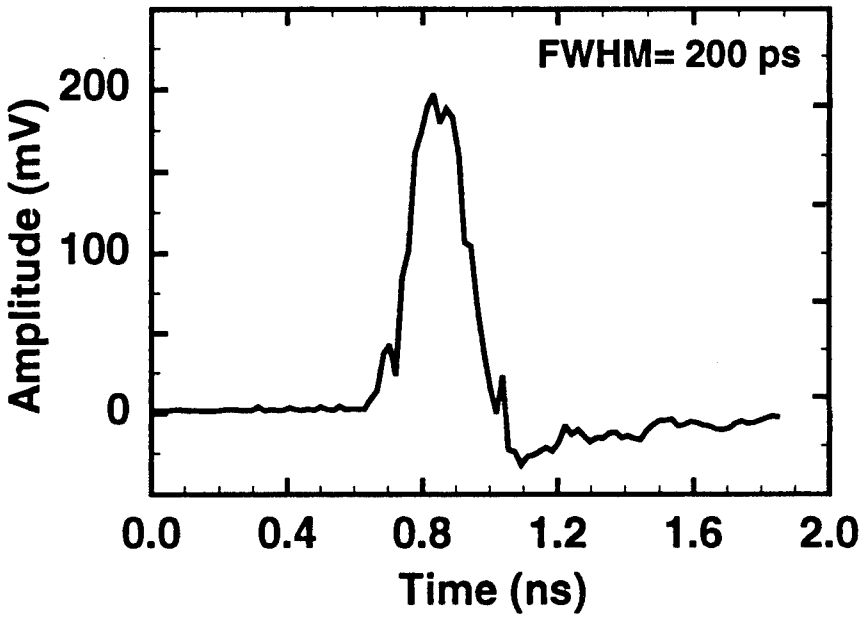


Figure 5: Measured temporal response of the photoreceiver with a transimpedance/voltage amplifier.

transimpedance/voltage amplifier exhibited a FWHM of 200 ps and a bandwidth of 2.5 GHz with a transimpedance gain of 63 dBΩ. These results verify the feasibility of OEICs using regrowth to integrate independently optimized structures.

Acknowlegment

The authors wish to thank L. Davis for his help during the temporal response measurements. The measurement of the optical-to-electrical frequency response of the photoreceiver with a transimpedance amplifier was done in cooperation with J.P. Nicolazzi and R. Lefevre at CNET-Bagneux, France. This work is supported by the U.S. Army Research Office (URI program) under Grant No. DAAL03-92-G-0109.

References

[1] O. Wada, T. Sakurai, and T. Nakagami, "Recent Progress in Optoelectronic Integrated Circuits (OEIC's)," *IEEE J. Quantum Elecron.*, vol. QE-22, p. 805, 1986.

[2] O. Wada, H. Hamaguchi, M. Makiuchi, T. Kumai, M. Ito, K. Nakai, T. Horimatsu, and T. Sakurai, "Monolithic Four-Channel Photodiode/Amplifier Receiver Array Integrated on a GaAs Substrate," *J. Lightwave Tech.*, vol. LT-4, pp. 1694-1702, 1986.

[2] O. Wada, H. Hamaguchi, M. Makiuchi, T. Kumai, M. Ito, K. Nakai, T. Horimatsu, and T. Sakurai, "Monolithic Four-Channel Photodiode/Amplifier Receiver Array Integrated on a GaAs Substrate," *J. Lightwave Tech.*, vol. LT-4, pp. 1694-1702, 1986.

[3] C.S. Harder, B.V. Zeghbroeck, H. Meier, W. Patrick, P. Vettiger, "5.2-GHz Bandwidth Monolithic GaAs Optoelectronic Receiver," *IEEE Elec. Dev. Lett.*, vol. EDL-9, pp. 171-173, 1988.

[4] C.-L. Cheng, R.P.H. Chang, B. Tell, S.M. Zima Parker, Y. Ota, G.P. Vella-Coleiro, R.C. Miller, J.L. Zilko, B.L. Kasper, K.F. Brown-Goebeler, and V.D. Mattera, Jr., "Monolithically Integrated Receiver Front End: $In_{0.53}Ga_{0.47}As$ p-i-n Amplifier," *IEEE Trans. Elec. Dev.*, vol. 35, pp. 1439-1444, 1988.

[5] Y. Zebda, P.K. Bhattacharya, D. Pavlidis, and J.P. Harrang, "Performance characteristics of $In_{0.6}Ga_{0.4}As/In_{0.52}Al_{0.48}As$ modulation-doped field-effect transistor monolithically integrated with InGaAs p-i-n photodiodes," *J. Appl. Phys.* vol. 68, pp. 1918-1920, 1990.

[6] S. Chandrasekhar, B.C. Johnson, E. Tokumitsu, A.G. Dentai, C.H. Joyner, A.H. Gnauck, J.S. Perino, and G.J. Qua, "A Monolithic Long Wavelength Photoreceiver Using Heterojunction Bipolar Transistors," *J. Quantum Electron.*, vol. QE-27, pp. 773-777, 1991.

[7] Y. Akahori, Y. Akatsu, A. Kohzen, and J. Yoshida, "10-Gb/s High-Speed Monolithically Integrated Photoreceiver Using InGaAs p-i-n PD and Planar Doped InAlAs/InGaAs HEMT's," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 754-756, 1992.

[8] R. Lai, P.K. Bhattacharya, D. Pavlidis, and T. Brock, "Monolithically Integrated Planar Front-End Photoreceiver with 0.25 μm Gate Pseudomorphic $In_{0.6}Ga_{0.4}As/In_{0.52}Al_{0.48}As$/InP Modulation-Doped Field-Effect Transistors," *Electron. Lett.*, vol. 27, pp. 364-366, 1991.

[9] P.R.Berger, N.K.Dutta, D.A.Humphrey, P.R.Smith, S.-J.Wang, R.K.Montgomery, D.Sivco, and A.Y.Cho, "1.0 GHz Monolithic p-i-n MODFET Photoreceiver Using Molecular Beam Epitaxial Regrowth," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 891-894, 1992.

MSM Waveguide Photodetectors Optimized for Monolithic Integration with HEMTs

M.H. Leary and J.M Ballantyne

School of Electrical Engineering, Phillips Hall
Cornell University, Ithaca NY 14853

ABSTRACT

This presentation describes the design of two MSM waveguide photodetectors which share their layer structure with high electron mobility transistors. This design makes these detectors particularly suited to integration with HEMTs in OEIC's. Resistive loss at the electrode surface, radiation loss into the substrate, and scattering loss are all accounted for in the optimization for internal detection efficiency of these 50GHz detectors.

I. INTRODUCTION

Waveguide photodetectors are well suited for use in optoelectronic integrated circuits. Their optical waveguide structure makes them compatible with optical signal routing and processing. MSM waveguide detectors have the added advantage that their planar electrodes are well suited to easy integration with high speed transistors. Those same planar electrodes have a very low capacitance, especially compared to the parallel plate electrodes of waveguide p-i-n detectors, which allows long absorption lengths and high detection efficiencies. Many others have reported MSM waveguide photodetectors suitable for use in OEIC's[1,2,3], however, we believe our design is unique because we utilize the same layer structure as a HEMT for these detectors, thus enhancing the suitability of these detectors for OEIC's. Once we committed to the transistor layer structure, in which the absorber is a thin InGaAs quantum well, the MSM waveguide structure became the only way to achieve good responsivity and high speed simultaneously. We will discuss the design of two detectors, with the structures shown in Fig 1. The first we call the longitudinal detector, because the electrodes run along the waveguide ridge. The second we call the transverse detector because the interdigitated electrode fingers run perpendicular to the waveguide. In the interdigitated design, the waveguide ends just before the detector, and the fingers lie on a flat surface. This geometry prevents carriers being generated in the deep fringing electric field under the ridge.

First we will discuss design considerations applicable to both versions of the detector. These include the layer structure, quantum well disordering, bandwidth constraints, and parasitic losses. Then we will show how these considerations were applied to the longitudinal and transverse detector designs.

0-7803-0894-8/93 $3.00

II. COMMON DESIGN CONSIDERATIONS

LAYER STRUCTURE

On the left in Fig 2 is the layer structure of a standard HEMT[4]. From the top down, the layers are an n^+ GaAs ohmic contact layer, an AlGaAs barrier layer, a thin GaAs smoothing layer, an InGaAs pseudomorphic channel, a GaAs subchannel, an AlGaAs/GaAs superlattice smoothing buffer, and the GaAs substrate. In this form, this structure is unsuitable for waveguide photodetectors because it does not support any guided modes; no part of the structure acts as a lower optical cladding. On the right is the layer structure we designed to support both HEMTs and MSM photodetectors. All the layers from the n^+ contact layer through the InGaAs channel remain untouched, since these are critical to the performance of the transistors. To support transverse waveguiding, we increased the thickness w of the SL buffer to act as a lower cladding. The GaAs subchannel now acts as the primary optical guiding layer; we changed its thickness, d, to optimize the sensitivity of the detectors. We also added a GaAs layer on top to form the waveguide ridge. The thickness t of that ridge was also a parameter in our design. Below the ridge layer is a thin AlGaAs layer that acts as an etch stop layer, allowing the etching of the ridges and uncovering of the undisturbed transistor contact layer in one simple step. The InGaAs pseudomorphic channel is used as the absorber for the detectors. We expect the absorption edge to be at 0.98 μm, at which wavelength the other materials in the structure are transparent.

At the bottom of Fig 2 is a scale drawing of the detector structure showing the approximate modal intensity distribution in the layers. The structure forms a highly asymmetric waveguide. Since the QW absorber is close to the low-index upper cladding, the major design challenge is to maximize the light intensity in the absorber while minimizing parasitic losses. One way we elimitated parasitic loss is by selectively disordering the quantum well, as indicated in the figure. By heating the material in the presence of a silicon dioxide cap, the diffusion of group III elements greatly increases, turning the InGaAs in the QW into a quarternary with higher bandgap[5]. We employ this technique to eliminate absorption under the electrodes where the electric field is weak, and also in the waveguides leading to the detector.

BANDWIDTH CONSIDERATIONS

These detectors were designed with a 3dB frequency of 50 GHz for use in a 44GHz integrated preamp. There are two limitations to the bandwidth of a photodetector: the transit time of photogenerated carriers and the capactitance of the electrodes. Unlike conventional bulk MSM detectors in which the carriers follow the curved electric field lines, the carriers in our detectors are confined in the quantum well, and thus follow straight paths as in a p-i-n diode. For this case, the transit time limited bandwidth can be expressed as $0.45\ v/L$ [6],where v is the saturation velocity of the carriers and L is the transit distance. With a saturation velocity of 10^7 cm/s, the transit distance must be at most 0.9 μm in order to obtain a 50 GHz bandwidth. The spacing of the electrodes in our design is thus 0.9 μm.

We calculated the capacitance of the edge coupled electrodes with a standard conformal mapping technique[7]. In order to obtain the capacitance limited bandwidth, we calculated the charging time into a standard 50Ω load, and inverted it to get a bandwidth.

PARASITIC LOSSES

We considered four sources of parasitic loss. Absorption by the QW in undesireable regions, such as under the electrodes and in the waveguides leading to the detector, we have eliminated with selective quantum well disordering. Scattering out of the waveguide is caused by roughness in the surfaces of the structure, and is mostly a processing issue. Typical values for this loss reported on structures similar to ours were about 1 cm^{-1} [8,9], which is small compared to other losses in the system, so we ignored this loss.

Since the lower optical cladding has a finite thickness, some optical energy will leak into the substrate. Because our lower cladding is formed of many many layers of superlattice, it was important to keep the layer as thin as possible. Fig 3 shows the attenuation coefficient due to leakage into the substrate as a function of the lower clad thickness w for various values of the guide layer thickness d. The reference line is 1 cm^{-1}, the approximate scattering loss. Our detectors use 0.3 μm guides, for which a 1.5 μm lower clad gives a loss a decade less than the scattering loss. That thickness required a reasonable 150 periods of the superlattice, so we chose w=1.5 μm and ignored substrate radiation loss in future calculations.

Real metal electrodes have finite resistivity, which causes loss where the optical field meets the metal surface. This loss is quite substantial, and plays a major role in our design tradeoffs. Fig 4 shows attenuation coefficient due to resistive loss at the electrodes as a function of d, the guide thickness. We calculated this loss by inserting the complex dielectric constant of the metal, and solving for the imaginary part of the propagation constant of the optical mode[10]. These losses are fairly large, especially for thin guides, for which the optical mode tends to spread into the metal surface.

III. LONGITUDINAL DETECTOR

Fig 5 is a plot of the optical intensity fraction in the QW absorber as a function of d for various ridge thicknesses t. As expected, thicker ridges give better lateral confinement, increasing the intensity fraction in the absorber. Thinner guide layers give higher intensity fractions since the absorber is near the top of the guide, so spreading the mode out actually increases the intensity fraction in the absorber. For combinations of t and d above those shown here, the waveguides became multi mode. Because the distribution of power among the modes of a multimode waveguide is virtually impossible to know *a priori*, and the detection efficiency of the structure varies for the different modes, we constrained ourselves to single mode waveguides, allowing calculation and optimization of the detection efficiency.

With the same parameters, Fig 6 shows the maximum detection efficiency, calculated as the ratio of power absorbed in the QW to total power absorbed. This would be the detection efficiency of an infinitely long detector. The best achievable efficiency is about the same for many combinations of d and t. We chose a 0.3 μm guide with a 0.1 μm ridge; this design has sufficiently large absorption that a 250 μm long detector gives 90% detection efficiency. The device could be made about four times longer before its response would be capacitance limited. We restricted ours to 250 μm so it would be short compared to the modulation wavelength at 44GHz.

IV. TRANSVERSE DETECTOR

In order to discuss the design of the transverse, interdigitated detector, we must first define some terminology. The maximum internal efficiency is the same as defined above. The lateral efficiency is the fraction of the lateral optical intensity distribution covered by the electrode fingers. This quantity is related to the length of the electrode fingers. The internal efficiency is the net detection efficiency, accounting for lateral efficiency and finite length.

Fig 7 shows maximum internal efficiency vs d for two values of electrode finger width. In both cases, the space between the electrode fingers was 0.9 μm. The narrower fingers give better efficiency because there is less resistive loss at the electrode surfaces. 0.1 μm fingers are difficult to fabricate reliably, however, so we used 0.25 μm fingers in our design. The electrodes lie on a flat surface in this design, so the ridge thickness t is not a parameter. Although the maximum internal efficiency is practically independent of the guide thickness, the total absorption coefficient is largest for the smallest d, as we saw earlier. Unlike the longitudinal detector, this detector is capacitance limited, so it was necessary to use a 0.3 μm guide layer to get the smallest absorption length.

With the absorption maximized, the detection efficiency of this detector was maximized by varying the initial length of the electrode fingers and the spreading angle of the light exiting the waveguide. The spreading angle is controlled by varying the lateral width of the waveguide mode. Fig 8 represents an optimization of those two parameters for maximum internal efficiency as a function of lateral efficiency. The best performance comes at a surprisingly large 98.5% lateral efficiency, and the internal efficiency is significantly degraded from $\eta_{internal,max}$ because capacitance limits the length of the device.

V.CONCLUSIONS

Both the longitudinal and transverse detectors share the layer structure of HEMTs, modified from the original HEMT structure to include a 0.3 μm optical guiding layer and a 1.5 μm optical lower cladding. Both are very efficient, and short compared to the modulation wavelength at 44GHz. Both employ quantum well disordering to eliminate absorption in the waveguides leading to the detectors.

The longitudinal detector exhibits a particularly high detection efficiency of 90%, and uses a GaAs waveguide ridge 0.9 μm wide and 0.1 μm thick. It uses quantum well disordering to eliminate carrier generation under the electrodes. The capacitance of the device is about 1/4 of the capacitance that would limit the response. All the dimensions of the device are large enough to be fabricated entirely by optical lithography.

The transverse detector has a somewhat lower internal efficiency, and uses interdigitated electrodes in a fan shape 5 μm long at one end and 15 μm long at the other. The mode shape required for best efficiency is realizable with a waveguide ridge made from AlGaAs and SiO_2. The 0.25 μm electrode fingers with 0.9 μm spaces used in this design would need to be fabricated by electron beam lithography.

We have demonstrated that it is possible to design very efficient high speed MSM waveguide photodetectors based on a layer structure compatible with high electron mobility transistors.

ACKNOWLEDGEMENT

This work was supported by Rome Laboratory under contract F30602-91-C-0063.

REFERENCES

[1]Gammel and Ballantyne, "An integrated photoconductive detector and waveguide structure," *Appl. Phys. Letters*, **36**(2) p. 149 (1980).

[2]Hong, Chang, Bhat, Nguyen, and Koza, "Monolithically Integrated Waveguide-MSM Detector-HEMT Amplifier Reciever for Long Wavelength Lightwave Systems," *IEEE Photonics Technology Letters*, **3**(2) p.156 (1991).

[3]Soole, Shumacher, LeBlanc, Bhat, and Koza, "Butt-Coupled InGaAs metal-semiconductor-metal waveguide photodetector formed by selective area regrowth," *Applied Physics Letters*, **56**(16) p. 1518 (1990).

[4]Yoo, Burm, Park, Mandeville, Schaff, and Eastman, "AlGaAs/InGaAs Double Modulation-Doped Field Effect Transistor with a Graded Composition in the InGaAs Quantum Well," *to be published.*

[5]Suzuki, Iwamura, and Mikami, "TE/TM mode selective channel waveguides in GaAs/AlAs superlattice fabricated by SiO_2 cap disordering," *Applied Phys. Let.*, 56(1), Jan 1990, p.19.

[6]Bowers, John E. and Burrus, Charles A., "Ultrawide-Band Long-Wavelength p-i-n Photdetectors," *J. Lightwave Tech*, LT-5, p.1341, 1987.

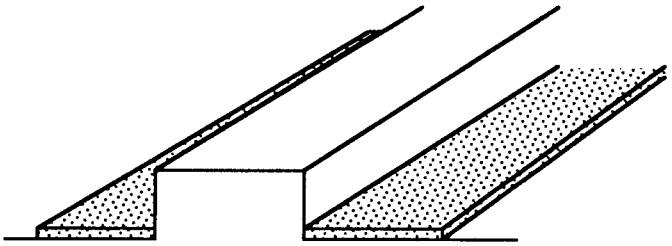
[7]Lim and Moore, "Properties of Alternately Charged Coplanar Strips by Conformal Mappings," *IEEE Transactions on Electron Devices*, v. ED-15, p.173.

[8]Deri, Kapon, Schiavone, "Scattering in low-loss GaAs/AlGaAs rib waveguides," *Applied Phys Let.*, 51(11) Sept 1987, p. 789,

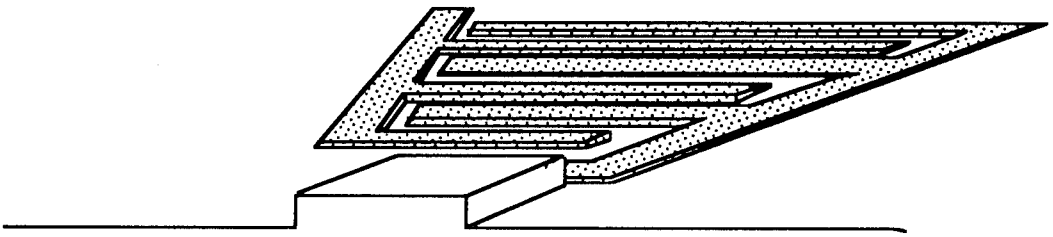
[9]Stern, M. S., Kendall, P. C., Hewson-Browne, R.C., and Robson, P.N., "Scattering loss from rough sidewalls in semiconductor rib waveguides," *Electronics Letters*, 1989.

[10]Garmire and Stoll, "Propagation losses in metal-film-substrate optical waveguides," *Journ. Quantum Electronics*, **QE-8** (10), p. 763.

MSM Waveguide Detector Structures

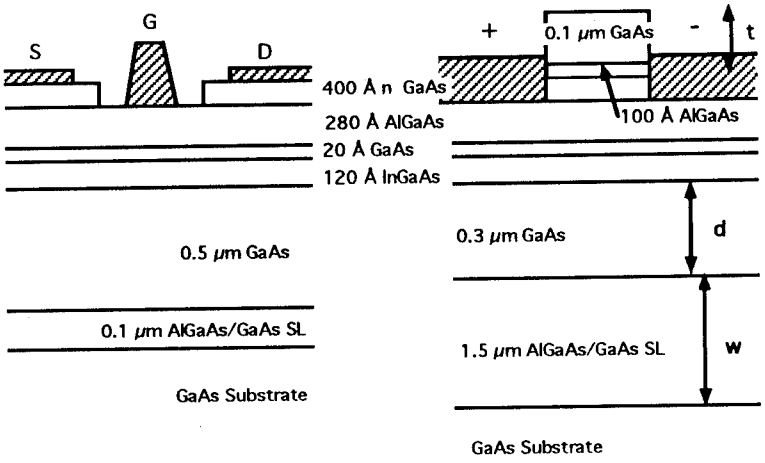


(a) Longitudinal Detector



(b) Transverse Detector

Fig 1



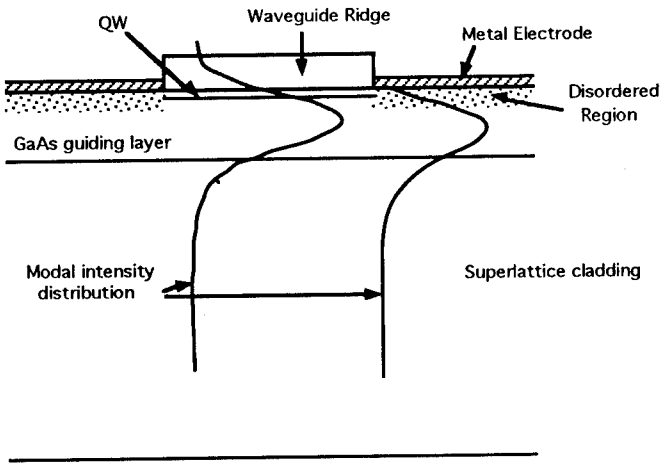


Fig 2

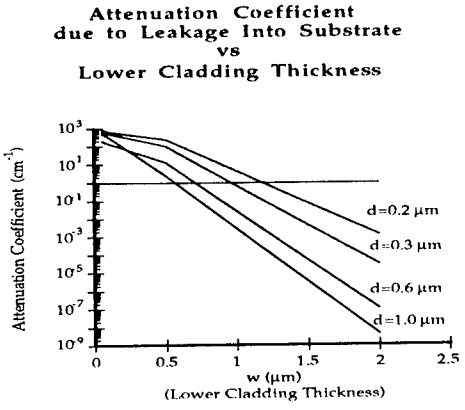


Fig 3

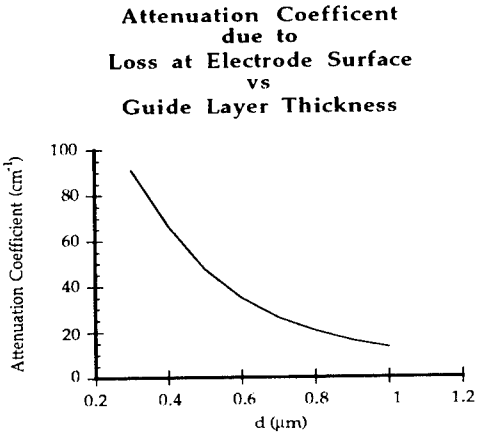


Fig 4

Longitudinal Detector

Intensity Fraction in the QW
vs
Guide Layer Thickness

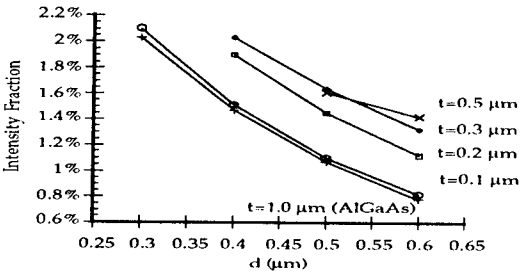
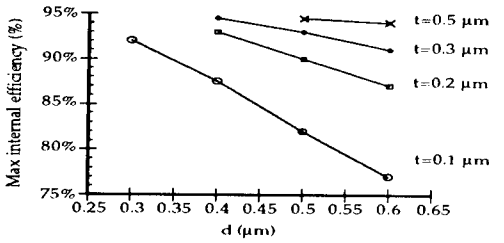


Fig 5

Longitudinal Detector

Maximum Internal Efficieny
vs
Guide Layer Thickness



$$\eta_{internal,max} = \frac{\text{Detectable Power Absorbed}}{\text{Total Power Absorbed}}$$

Fig 6

Transverse Detector

Maximum Internal Efficiency
vs
Guide Layer Thickness

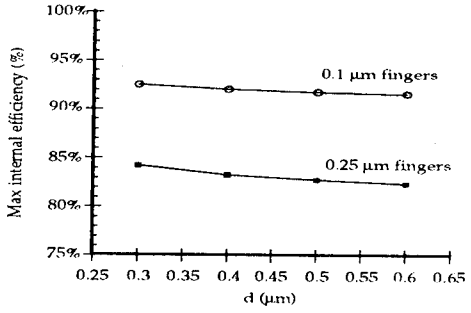


Fig 7

Transverse Detector

Overall Internal Efficiency
vs
Lateral Efficiency

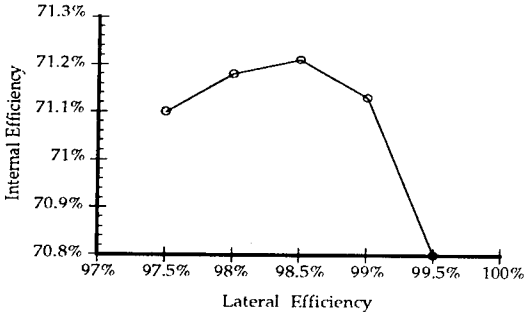


Fig 8

DELAY OF FIELD COLLAPSE IN PHOTOCONDUCTIVE GAPS FABRICATED ON GaAs/AlGaAs MODFET MATERIAL

J. A. Sheridan, B. A. Nechay, D. M. Bloom,
E. L. Ginzton Laboratory, Stanford University
Stanford CA 94305

P. M. Solomon,
IBM Watson Research Laboratory,
Yorktown Heights NY 10598

Y. C. Pao,
Litton Solid State Division
Santa Clara CA 95054

ABSTRACT

We report the fabrication of a 1 μm photoconductive gap voltage step generator with constant output voltage of 0.10 V for 100 ps and a 2 ps rise time. The photoconductor is fabricated on GaAs/AlGaAs MODFET material and uses the two-dimensional electron gas in the FET channel as the conductive medium. It is fully process-compatible with MODFETs. In devices with standard ohmic contacts, the output current stayed constant for up to 50 ps, even though the electric field in the gap is expected to collapse in less than 1 ps. Two-dimensional device simulations show that, although the field in the gap does collapse on the expected time scale, the current is initially determined by the contacts. A new photoconductor design, in which the ohmic contacts are laterally recessed from the etched gap region, further delayed field collapse, extending constant voltage operation to 100 ps.

I. INTRODUCTION

Most research on photoconductors has been directed toward their use as high-speed sampling gates [1], high-speed, high-voltage switches [2], and generation and detection of free-space radiation [3]. For these applications, much effort has been concentrated on reducing the time duration of the photoconductor response to an impulse function on the picosecond time scale. However, it is also important to generate a voltage step function with a picosecond rise time in order to study the step response of electronic devices. A step rather than a pulse is required when measuring the large-signal nonlinear response of the device. Such measurements will yield information about nonequilibrium carrier transport phenomena, such as velocity overshoot, which may substantially enhance device performance. These transport effects, which have been studied extensively by Monte Carlo simulations [4-7], are expected to be seen in GaAs on the picosecond time scale, but have not been directly measured in electronic devices [8]. To pursue this study, it is necessary to fabricate a switch that can turn on the electronic device under test with a fast rise time (~1 ps), and can also yield a constant output voltage during the entire measurement window (~40 ps). It is desirable to

0-7803-0894-8/93 $3.00

monolithically integrate this switch with the device under test. In this paper, we present the fabrication and high-speed characterization of a photoconductive gap picosecond voltage step generator fabricated on GaAs/AlGaAs MODFET material.

II. MEASUREMENT PRINCIPLES

Figure 1 shows a measurement scheme in which a photoconductive gap provides the picosecond voltage step function required for testing the step response of a modulation-doped field-effect transistor (MODFET). In this configuration, the MODFET is initially biased into pinch-off. When light shines on the photoconductive gap, it will provide a fast rise time signal to switch on the FET. The output current of the FET is then measured with a high-speed measurement technique such as electro-optic sampling, which will be described in detail in Section V. If parasitics are minimized, the output current of the device will be indicative of the electron velocity inside the device. In order to measure the step response of the FET, it is necessary for the photoconductor output voltage to remain constant over the entire measurement window.

It is advantageous to use a photoconductive (PC) gap to provide the voltage step function to the FET for several reasons. First, a rise time of approximately 1 ps is needed, limiting the choices to photoconductive gaps [1], resonant tunneling diodes [9], or nonlinear transmission lines [10]. Also, to avoid signal distortion and parasitics associated with bond wires, it is desirable to monolithically integrate the switch with the FET. It is possible to design a photoconductive gap that can be monolithically integrated with a MODFET and requires no additional epitaxial layers or compromises in the FET design. Also, since an optical measurement technique is used, it is beneficial that the step function generator be switched-on optically. This is because the same laser pulse may be used to switch on the step generator (and hence the FET) and to measure the FET response. Thus, the two are inherently synchronized, eliminating loss of temporal resolution due to timing jitter encountered when synchronizing electronic sources to the laser.

III. PHOTOCONDUCTOR DESIGN

First, consider a photoconductor in which the ohmic contacts are flush with the etched gap edges, referred to as the "ohmic PC." It was expected that the current out of such a photoconductor would decay very rapidly even though both the recombination lifetime and the transit time may be long. This is because photogenerated holes that transit away from the anode cannot be replenished through the n-type ohmic contacts. This imbalance of electron charge near the anode causes a highly resistive space charge region to form, and most of the applied voltage is dropped across this region. This is referred to as "field collapse," because the electric field is reduced to zero across the remainder of the gap.

Field collapse has been measured experimentally inside photoconductors [11] and was expected to switch off the PC response in less than 1 ps. To delay field collapse and to allow a constant output voltage over the 40 ps measurement window, we designed a new type of photoconductor, which we call the "recessed-ohmic (RO) PC." In this design, shown in Figure 2, the ohmic contacts are laterally recessed from the etched gap region, leaving 2 μm of unetched material between the gap and either ohmic contact. It is expected that the ionized donors in the unetched regions will provide positive charge to compensate the otherwise unbalanced electron concentration, delaying formation of the space charge region and hence the onset of field collapse. After the holes transit into the etched gap region, the device is expected to behave like an ohmic PC.

IV. EPITAXIAL LAYERS AND DEVICE PROCESSING

The photoconductors were fabricated on MBE-grown, modulation-doped material, optimized to obtain good characteristics for the MODFETs with which the photoconductors will be integrated in the future. The dopants are provided by atomic planar doping with a sheet charge density of 4.0 x 10^{12} /cm^2, and the $Al_{0.25}Ga_{0.75}As$ spacer layer between this doping and the undoped GaAs channel is 20 Å. Above the doping is a 300Å undoped AlGaAs layer, followed by a 100 Å undoped GaAs layer. The mobility at room temperature is approximately 6000 cm^2/Vsec, and simulations indicate that the charge in the channel is approximately 1.2 x 10^{12}/cm^2.

The process flow used to fabricate the photoconductors is shown in Figure 3. First, ohmic contacts are defined by depositing and annealing Au:Ge:Ni contacts. Then, the material between the ohmics is etched using an H_3PO_4:H_2O_2:H_2O etch, until the two-dimensional election gas in the FET channel is depleted. This etch depth is determined by monitoring the resistance of the material in an etched region. The etched gap sizes range from 1 x 10 μm to 4 x 10 μm for both ohmic and RO PC's. Finally, a Ti:Au interconnect layer is deposited to define 50 Ω coplanar waveguide (CPW) transmission lines. This process is completely compatible with a standard MODFET fabrication process.

V. EXPERIMENTAL SET-UP

The laser system used to test the photoconductors is a mode-locked Nd:YAG laser followed by two fiber-grating pulse compressors, which produces 350 fs pulses at 1.06 μm wavelength and 82 MHz repetition rate. In order to excite the GaAs photoconductors, the 1.06 μm light is frequency-doubled to 532 nm in a potassium titanyl phosphate (KTP) crystal. The 532 nm light is focused to a 4 μm spot size on the photoconductive gap, and a DC bias is applied. The CPW on the output side of the device is contacted by a microwave probe, and the PC output voltage is monitored on a sampling oscilloscope. These scope measurements yield the amplitude, fall time, and time duration of constant voltage output, or "plateau time," of the PC response.

Because the resolution of the sampling oscilloscope is limited to 10 ps, the rise time and the first 10 picoseconds of the PC response are measured using an electro-optic (EO) sampling [12-14] technique. This method is based on the electro-optic effect. That is, in certain types of materials, including GaAs, the presence of an electric field in the material will induce birefringence in it. Thus, if an optical beam traverses this material, the polarization of the beam will be altered due to the presence of the field. An experimental EO sampling set-up is shown in Figure 4. As shown, the 1.06 μm light that is not used for frequency conversion is used as the non-invasive probe beam. The 532 nm pulse excites the photoconductor, and, as the photoconductor output voltage travels down the CPW line, it creates an electric field in the GaAs substrate. The 1.06 μm probe pulse is transmitted through the backside of the wafer, reflects off the signal line of the CPW, is passed through a polarizing beam splitter cube and detected, measuring the voltage on the transmission line at a particular time. The time delay between the 532 nm pump pulse and the 1.06 μm probe pulse is controlled by a corner-cube reflector on a stepper motor that varies the optical path difference. This allows the probe beam to sample the device response at a frequency measurable by a conventional oscilloscope. The measurement bandwidth of our system exceeds 300 GHz.

VI. EXPERIMENTAL RESULTS

Contrary to expectation, the voltage output of the ohmic photoconductors did not shut off in less than 1 ps. In fact, the ohmic PC's were able to sustain 80 mV output voltage (± 2%) for 50 ps, with a 1/e fall time of 600 ps, as shown in Figure 5a. Two-dimensional device simulations have been done to elucidate this prolonged response, and will be described in the following section. The recessed ohmic photoconductors did achieve our goal of delaying the field collapse. These devices were able to output constant voltage of 0.11 V for an average of 100 ps, a factor of 2 longer than the ohmic PC's, as shown in Figure 5b. With 75 μW incident average optical power, a typical RO PC response was 0.10 V constant voltage output over 100 - 120 ps with a 1/e fall time of 500 - 800 ps. In both types of devices, the output voltage increased with both bias and incident optical power, while plateau times decreased. With a bias variation, there is a trade-off between the amplitude of the output voltage and the plateau time. The RO PC outputs ranged from 25 mV constant for 500 ps at 1V bias to 200 mV constant for 40 ps at 8V bias.

An EO sampling trace of a recessed-ohmic photoconductor is shown in Figure 6a. In this figure, the sampling calibration is that 1 ms corresponds to 1.1 picoseconds in real time. It can be seen that the photoconductor response is flat over our 40 ps measurement window. The rise time is 2 ps, longer than the expected rise. No dependence of the rise time on optical power was found, suggesting that the rise is not simply caused by a charging time, which should decrease with an increase in optical power. A possible explanation for this slow rise may be that carriers are excited into lower-mobility satellite valleys by the high photon energy of the 532 nm light. Monte Carlo calculations have shown that the rise time of the output current increases when the photon energy of the exciting light exceeds the bandgap energy between the L-valley minimum and the top of the valence band, 1.75 eV (0.7 μm) in GaAs [15]. At high bias, a voltage-dependent dip was seen in the EO time trace, primarily in 1 μm devices, as shown in Figure 6b. This dip was seen to occur at earlier times when the optical power was increased or the applied voltage was increased. The dip may be another manifestation of hot carrier excitation into satellite valleys of GaAs, specifically the "velocity undershoot" effect predicted by 2D Monte Carlo simulations [15] for GaAs under 532 nm excitation. According to these simulations, the average velocity of electrons will actually decrease after several picoseconds, and then return to its equilibrium value due to the dynamics of carrier transfer between the valleys. In our structure, an additional possibility is the real space transfer of electrons by the light into the AlGaAs. Additional Monte Carlo modeling for our conditions and device geometry is needed in order to verify these conjectures.

VII. SIMULATIONS

Two-dimensional device simulations have been done to investigate why the photocurrent in the ohmic devices remained constant beyond the time the electric field in the device was expected to collapse. Silvaco's optoelectronic/heterostructure program was used, assuming field-independent mobilities for electrons and holes. The potential as a function of time across a vertical cross-section of the device may be seen in Figure 7a. These simulations show the expected time-dependent voltage drop near the anode as the photogenerated holes transit out of this region. However, even though the field starts to collapse at 500 fs, the contacts limit the current through the device until the gap resistance becomes comparable to the contact resistance. To illustrate this, the output current as a function of time for the same modeled ohmic PC with two different contact resistances is shown in Figure 7b. It can be seen that the device with larger contact resistance, which is similar to our devices, has a longer plateau region. In addition, it is seen that the voltage drop across the contacts reduces the field in the gap for a given voltage, decreasing the hole velocity and further delaying the field

collapse. The simulations also show that, even after the electric field collapses, current is still carried by diffusion until the carriers recombine, giving a long tail to the PC response. Thus, while a time-resolved measurement of the collapse of the electric field within a photoconductive gap has been made[11], it cannot be assumed that the output current of the device will decay on that time scale. At present, the simulations predict plateau regions that are much shorter than those we have observed in our ohmic photoconductors. We are trying to explain this discrepancy by including in the model surface traps, which may trap holes and prevent their rapid transit out of the anode region.

As explained, we had previously tried to combat the expected field collapse by laterally recessing the ohmic contacts from the etched gap region, leaving unetched material between the gap and the contacts. Simulations show that the field in this device does not collapse until the holes transit into the etched gap region, delaying the field collapse as expected. In addition, the heavy doping in the unetched region reduces the resistivity compared with the etched gap region. This decreases the hole velocity and further delays field collapse compared with the ohmic PC's.

VIII. CONCLUSION

We have fabricated a 1 μm photoconductive gap picosecond voltage step generator on GaAs/AlGaAs MODFET material. In photoconductors with standard ohmic contacts, prolonged responses have been seen, beyond the expected current decay due to electric field collapse in the gap. Two-dimensional device simulations have shown that, although the field in the gap does collapse on the expected time scale, the current is initially limited by the contacts, and several picoseconds elapse before the current becomes sensitive to the internal field in the photoconductor. Presently, the numerical model does not predict plateau times as long as those seen in our data. In addition, we have specially designed a new type of photoconductor, the recessed-ohmic photoconductor, in which the ohmic contacts are laterally recessed from the etched gap region to delay field collapse. These devices were able to sustain constant voltage for over 100 ps. We have also seen an interesting dip in response of these devices that may be experimental evidence of the velocity undershoot effect predicted by Monte Carlo calculations for GaAs devices excited with high photon energy light.

ACKNOWLEDGMENTS

The authors wish to thank Robert Cottle of Silvaco International for simulation advice, and Rashit Nabiev and Lars Eng for helpful technical discussions. J. A. Sheridan gratefully acknowledges an AT&T Bell Laboratories graduate fellowship. This work was funded by ARPA through the Mayo Foundation on contract #N66001-89-C0104.

REFERENCES

[1] D. H. Auston, "Picosecond photoconductors, physical properties and applications," in *Picosecond Optoelectronic Devices*, edited by C. H. Lee (Academic, New York, 1983), pp. 73-117.

[2] T. Motet, J. Nees, S. Williamson, and G. Mourou, "1.4 ps rise time high-voltage switching," *Appl. Phys. Let.t.*, vol. 59, pp. 1455-1457, 1991.

[3] D. H. Auston, K. P. Cheung and P. R. Smith, "Picosecond photoconducting Hertzian dipoles," *Appl. Phys. Lett.*, vol. 45, pp. 284-286, 1984.

[4] J. G. Ruch, "Electron dynamics in short channel field-effect transistors," *IEEE Trans. Electron Devices*, vol. ED-19, no. 5, pp. 652-654, 1972.

[5] T. J. Maloney and J. Frey, "Transient and steady-state electron transport properties of GaAs and InP," *Journ. Appl. Phys.*, vol. 48, pp. 781-787, 1977.

[6] A. E. Iverson, G. M. Wysin, D. L. Smith, and A. Redondo, "Overshoot in the response of a photoconductor excited by subpicosecond pulses," *Appl. Phys. Lett.*, vol. 52, pp. 2148-2150, 1988.

[7] M. V. Fischetti, "Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures--Part I: Homogeneous transport," and M. V. Fischetti and S. E. Laux, "Part II: Submicrometer MOSFETs," *IEEE Trans. Electron Devices*, vol. 38, pp. 634-660, 1991.

[8] Transient effects have been investigated in photoconductors, in the form of photoconductivity overshoot and temporal changes in electric field, including:
C. V. Shank et al, *Appl. Phys. Lett.*, vol. 38, pp. 104-105, 1981;
K. Meyer et al, *Appl. Phys. Lett.*, vol. 53, pp. 2254-2256, 1988; and
W. H. Knox et al, *Picosecond Electronics and Optoelectronics* digest, p. 40-45, 1991.

[9] E. Ozbay, D. M. Bloom, D. H. Chow, and J. N. Schulman, "1.7 ps, microwave integrated circuit compatible InAs/AlSb resonant tunneling diodes," to be published in *Electronics Letters*, Aug. 1993.

[10] C. J. Madden, R. A. Marsland, M. J. W. Rodwell, D. M. Bloom, and Y. C. Pao, "Hyperabrupt-doped GaAs nonlinear transmission line for picosecond shock wave generation," *Appl. Phys. Lett.*, vol. 54, pp. 1019-1021, 1989.

[11] W. Sha, J. Rhee, T. B. Norris, W. J. Schaff, "Transient carrier and field dynamics in quantum well transport: from ballistic to the quasi-equilibrium regime," *IEEE Journ. Quantum Elect.*, vol. 28, pp. 2445-2455, 1992.

[12] K. J. Weingarten, M. J. W. Rodwell, and D. M. Bloom, "Picosecond Optical Sampling of GaAs Integrated Circuits," *IEEE J. Quantum Electron.*, vol. QE-24, pp. 198-220, 1988.

[13] J. A. Valdmanis, "Electro-optic measurement techniques for picosecond materials, device sand integrated circuits," *Measurements of High-Speed Signals in Solid State Devices*, R. B. Marcus, Ed., in *Semiconductors and Semimetals*, R. K. Willardson and S. C. Beer, Eds., vol. 28, Academic Press, Orlando FL, 1990.

[14] J. M. Wiesenfeld, "Electro-optic sampling of high-speed devices and integrated circuits," *IBM J. Res. Develop.*, vol. 34, pp. 141-161, 1990.

[15] E. Sano, "Two-dimensional ensemble Monte Carlo calculation of pulse responses of submicrometer GaAs metal-semiconductor-metal photodetectors," *IEEE Trans. Elect. Devices,* vol. 38, pp. 2075-2081, 1991.

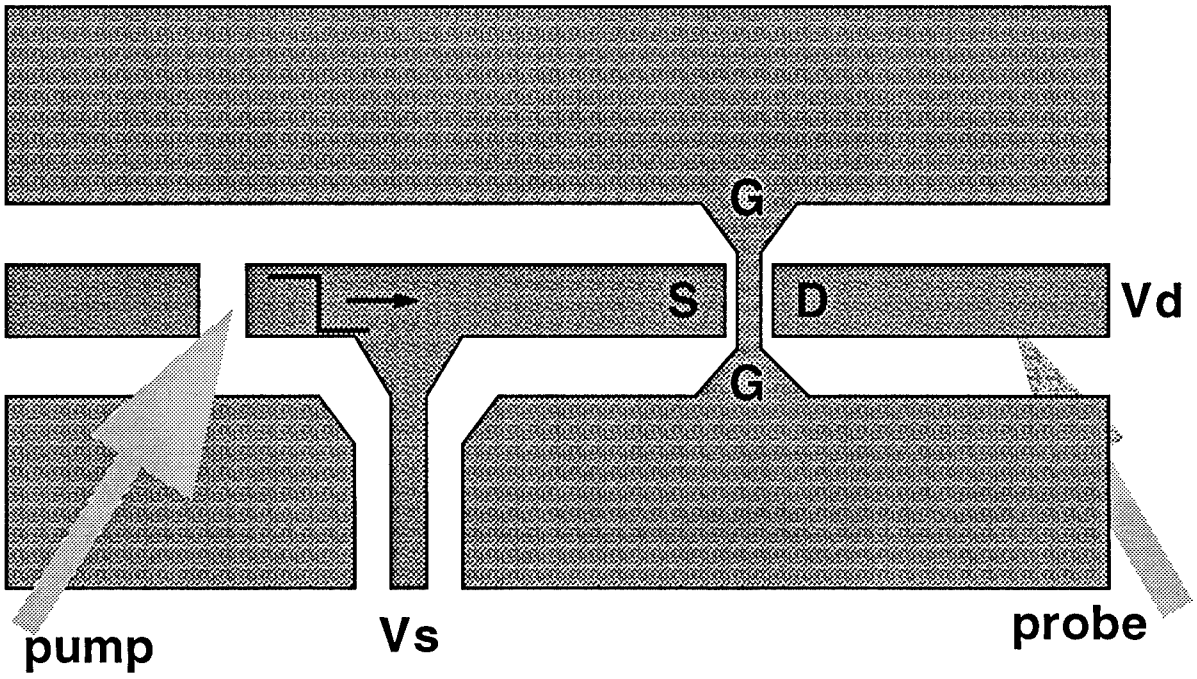


Figure 1: Integrated photoconductor/MODFET for the investigation of the step response of MODFETs on a picosecond time scale. S = source, G = gate, D = drain.

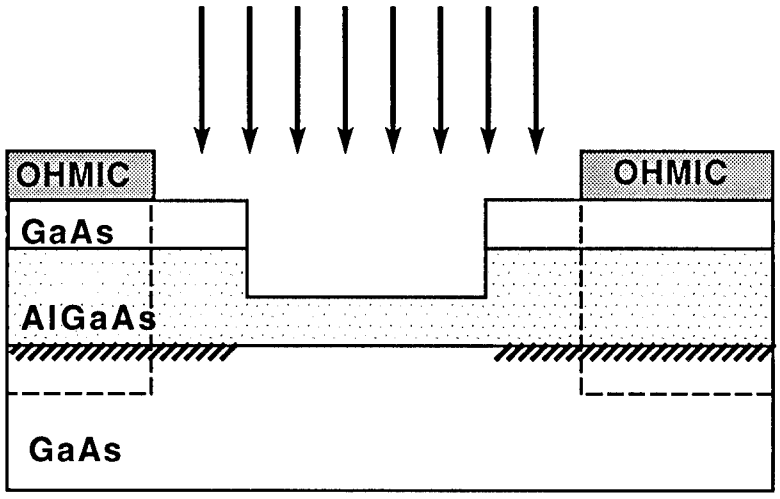


Figure 2: Schematic of recessed ohmic photoconductor.

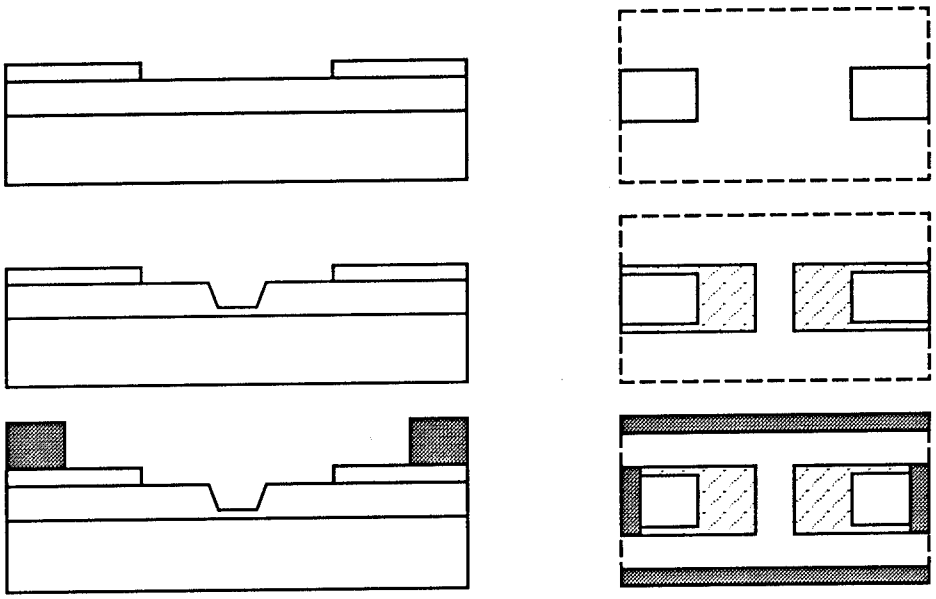


Figure 3: Photoconductor process flow: ohmic contacts, photoconductor etch, and interconnect metal.

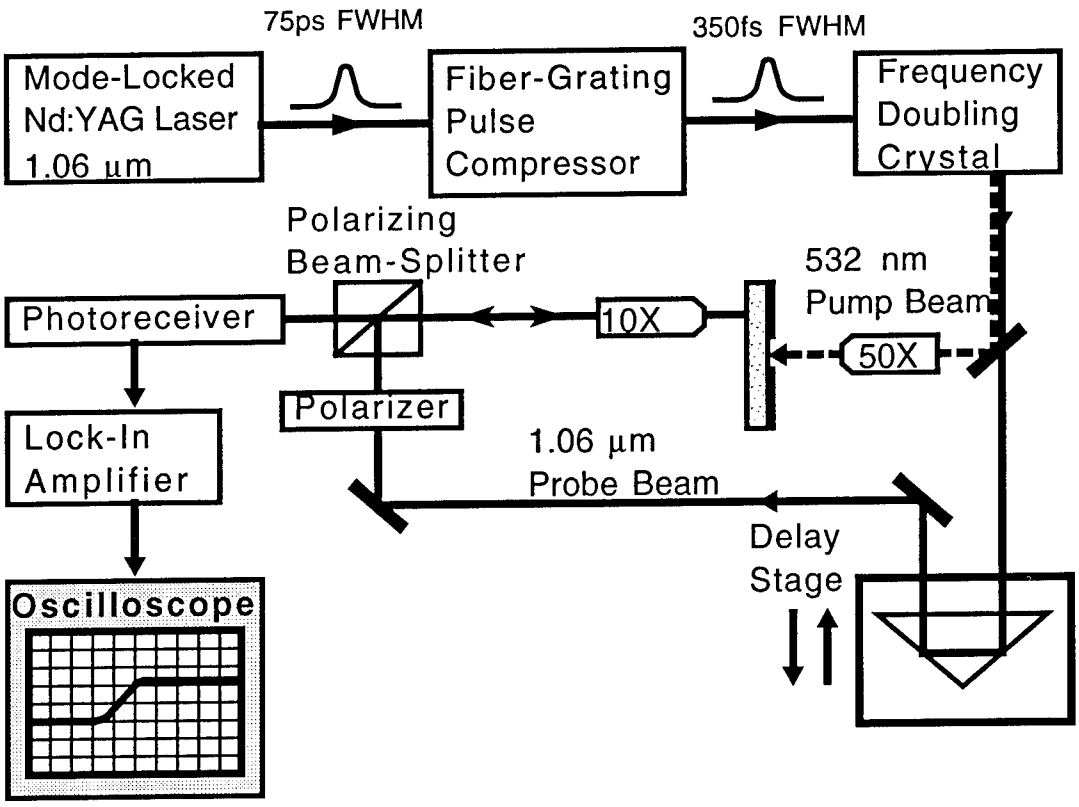
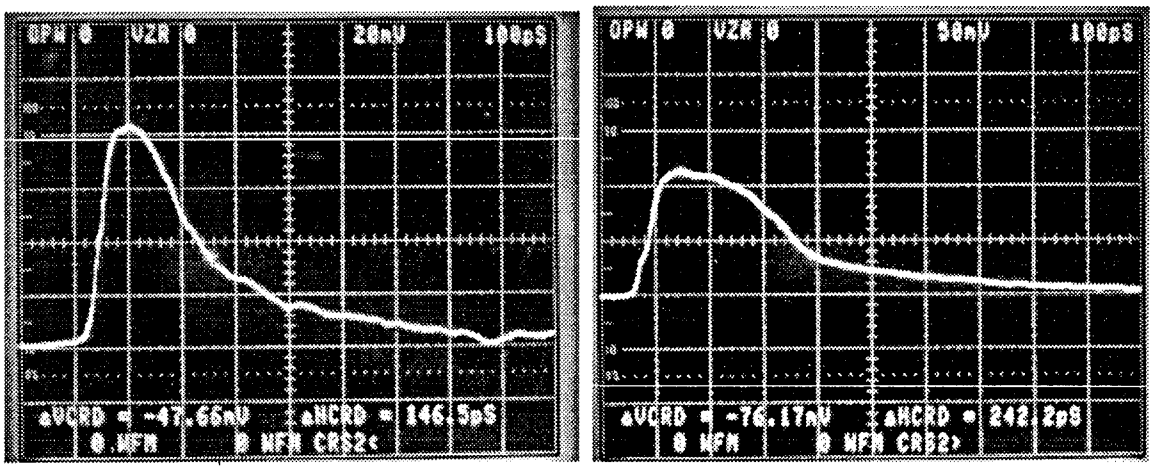
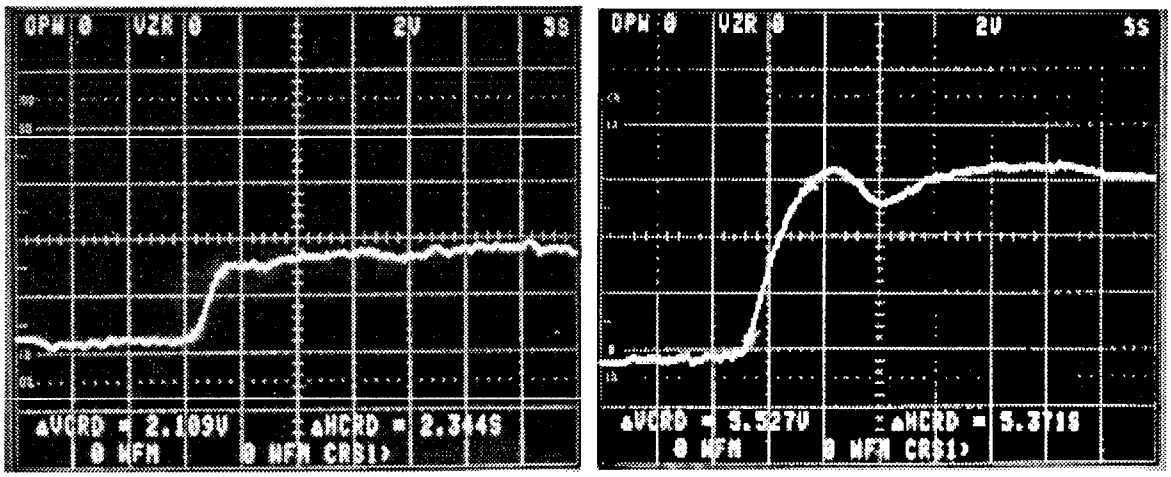


Figure 4: Pump/probe electro-optic sampling set-up.



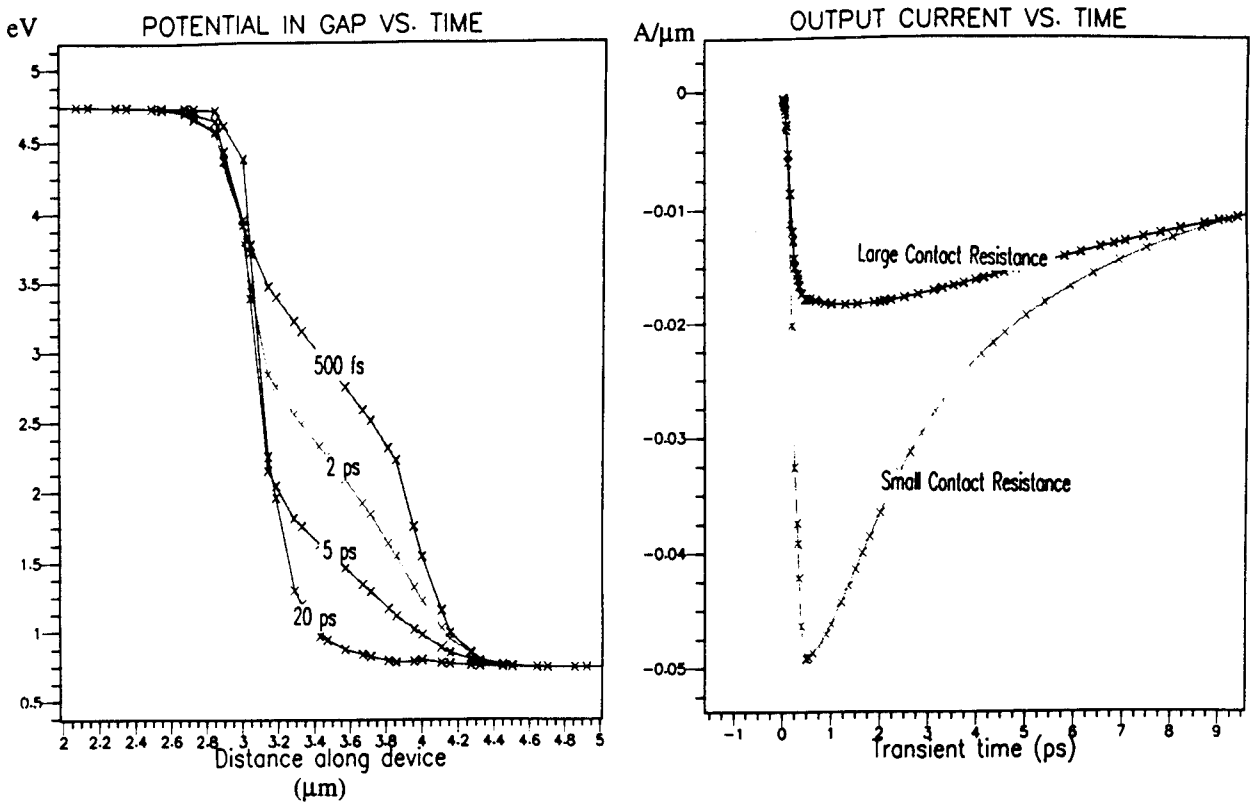
(a) (b)

Figure 5: (a) Response of ohmic photoconductor, vertical scale is 20 mV/div, and horizontal scale is 100 ps/div. (b) Response of recessed ohmic photoconductor, vertical scale is 50 mV/div, horizontal scale is 100 ps/div.



(a) (b)

Figure 6: (a) EO sampling trace of recessed ohmic photoconductor, vertical scale is equivalent 80 mV/div, horizontal scale is equiv. 5.5 ps/div. (b) EO sampling trace showing dip, vertical scale is equiv. 50 mV/div, horizontal scale is equiv. 5.5 ps/div.



(a) (b)

Figure 7: (a) Potential (in eV) across an ohmic photoconductor as a function of time. Horizontal scale is distance across the device in μm. Between 3 and 4 μm is the etched gap region. (b) Output current (in A/μm) as a function of transient time (in ps) for the same modeled device with small and large contact resistance

An Improved Gate Breakdown Model for Studying High Efficiency MESFET Operation

T.A. Winslow, A.S. Morris and R.J. Trew

High Frequency Electronics Laboratory
Department of Electrical and Computer Engineering
North Carolina State University
Box 7911
Raleigh, NC 27695-7911

Abstract

A gate breakdown model for MESFETs is presented. The model is based on quantum tunneling initiated avalanche ionization. The breakdown model is incorporated into a physics based MESFET model. Forward and reverse gate conduction are shown to be the main factors in determining the power and efficiency of MESFET amplifiers.

I. Introduction

The fundamental limitations of power MESFET amplifiers are the gate contact breakdown and forward conduction mechanisms[1]. Forward gate conduction limits the peak gate voltage and drain current swing. Gate-drain breakdown causes an increase in rectified gate and drain current which shifts the device quiescent bias point and initiates premature power saturation. Both mechanisms limit the maximum power and efficiency an amplifier can achieve. High efficiency and power performance of FET amplifiers is important for satellite and cellular communication systems, phased array radars, or any application where prime power and thermal considerations are important.

Breakdown models are typically based on the avalanche ionization mechanism, but recently the tunneling aspect of breakdown has gained attention[2][3]. Breakdown models based on the avalanche ionization mechanism alone cannot predict the decrease in breakdown voltage with increasing gate voltage from pinch-off[4][5] except under certain conditions[6][7], or the decrease in breakdown voltage with increasing temperature which has been experimentally correlated with device burn-out[21]. The avalanche model cannot predict the characteristic 'twist' in the gate current seen in MESFETs[4][5]. The temperature dependence of the breakdown voltage indicates that there is an important tunneling component to the breakdown mechanism. Although the gate breakdown process is affected by device geometry, material parameters and processing steps, the breakdown mechanism results from a combination of tunneling and avalanche generation. The strong electric field at the drain side of the gate initiates avalanche and field-assisted tunneling which dominate the device characteristics at high drain terminal voltages under pinch-off conditions. Previous studies of Schottky diodes demonstrate that field emission[8], thermionic-field emission[9], and thermionic emission[10] processes can dominate the diode current depending on the temperature, doping, and bias (and tunneling effective mass). The surface of the MESFET is an important part of

0-7803-0894-8/93 $3.00 © 1993 IEEE

the breakdown process due to the strong disorder in the bonding arrangements and loss of stochiometry at the semiconductor/passivant interface. The disorder at the device surface induces interface traps and distorts the surface band structure, thereby affecting the breakdown process. In this paper, a breakdown model for MESFETs, combining the tunneling and avalanche processes and surface trapping effects, is presented. This work provides the first breakdown model for MESFETs that combines the tunneling and avalanche processes and incorporates thermal effects in the breakdown mechanism. The model described in this work can successfully explain the breakdown effects on large-signal MESFET operation.

II. Breakdown Model

Proper characterization of electron emission from the gate is critical for a breakdown model. In this work, the electron transmission probability is determined **exactly** for a Schottky barrier without using the Wentzel-Kramers-Brillouin (WKB) approximation. A new formulation for the transmission probability is found because the majority of the electron flux crossing a Schottky barrier occurs near the peak of the barrier where neither thermionic nor thermionic-field emission is entirely dominant. The WKB approximation is inadequate for studying typical MESFET Schottky barrier operation. The breakdown model presented in this section generalizes the gate electron emission formulation and represents a distinct departure from purely avalanche based breakdown models[15].

To describe the tunneling component of the breakdown mechanism, the tunneling probability for a Schottky barrier is determined. Accurate quantum characterization of the barrier region requires the **exact** solution of the Schrodinger equation. In this model, the gate Schottky barrier is approximated by a three region linear potential energy barrier, shown in Figure 1 (image force effects are neglected). Region I is the gate metal region. Region II is a region of linear energy variation of width W and Region III is a constant potential region (Regions II and III are within the semiconductor). Region III is at a potential V_a relative to the metal Fermi level ($\mathcal{E}_f$). The linear barrier approximation is used since, as seen in Figure 1, there is only a small difference between the typically used quadratic potential solution[8][9][10] and the linear solution (near the transition between Region II and Region III) at energy levels relative to the metal Fermi energy such that significant tunneling does not take place. The Schrodinger equation is solved in one dimension using the linear potential model $V(x) = \phi$

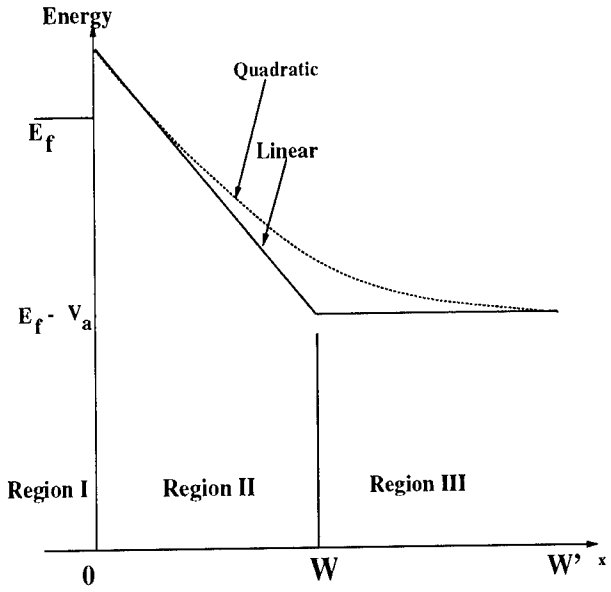


Figure 1: Linear and Quadratic Approximation to the Schottky Potential Barrier

$_b(E_g)$ - qE_gx where E_g is the barrier electric field strength and $\phi_b(E_g)$ is the Schottky barrier height. To determine the transmission probability, the wave function solutions and derivatives at the region boundaries are equated and the integration constants are determined. The trial solutions in region I and III are plane wave solutions to the Schrodinger equation. The solution of the Schrodinger equation in Region II is

$$\Psi_{II}(x) = C \cdot Ai\left(-\left(\frac{2E_g m_t q}{\hbar^2}\right)^{\frac{1}{3}} \frac{\mathcal{E} - \phi(E_g) + qE_g x}{qE_g}\right) + D \cdot Bi\left(-\left(\frac{2E_g m_t q}{\hbar^2}\right)^{\frac{1}{3}} \frac{\mathcal{E} - \phi(E_g) + qE_g x}{qE_g}\right) \quad (1)$$

where *Ai* is the Airy function A and *Bi* is the Airy function B (C and D are integration constants). The condition of the MESFET surface affects the energy band characteristics and increases the tunneling effective mass of electrons during the tunneling process which affects the breakdown voltage. The wave vectors in Region I and III are defined as

$$k_I = \frac{\sqrt{2m_m \mathcal{E}}}{\hbar} \quad (2)$$

and

$$k_{III} = \frac{\sqrt{2m_t(\mathcal{E} - \mathcal{E}_f + V_a)}}{\hbar} \quad (3)$$

where V_a is the applied potential across the Schottky barrier and $\mathcal{E}$ is the tunneling energy. The metal and tunneling effective masses are m_m and m_t, respectively, and are assumed constant. The wave vector in Region II is spatially dependent and is incorporated into the $\Psi_{II}(x)$ wave function solution in Eqn. 1. The transmission probability of the linear approximation to the Schottky barrier is based on the ratio of transmitted waves to incident waves and is thus $T(\mathcal{E}) = \frac{k_{III}}{k_I} T_o$ where T_O is defined by

$$T_o^{-1} = \frac{\left[\begin{array}{l}\left(A_0' B_W' - AB\right)^2 + \left(A_0 B_W' - A_W' B_0\right)^2 k_I^2 + \\ \left(A_0' B_W - A_W B_0'\right)^2 k_{III}^2 + \left(A_0 B_W - A_W B_0\right)^2 k_I^2 k_{III}^2 + \\ \left(A_W B_W' - A_W' B_W\right) 2k_I k_{III} + \left(A_0 B_0' - A_0' B_0\right)\end{array}\right]}{\left(4k_I \left(A_W' B_W - A_W B_W'\right)^2\right)} \quad (4)$$

where $\alpha = \left(\frac{2m_t qE_g}{\hbar^2}\right)^{\frac{1}{3}}$, $\beta = \frac{\mathcal{E} - \phi_b(E_g)}{qE_g}$, $\eta_0 = -\alpha\beta$, $\eta_W = -\alpha(\beta + W)$, $A_0 = Ai(\eta_0)$,

$B_0 = Bi(\eta_0)$, $A_0' = -\alpha \frac{\partial Ai(\eta_0)}{\partial x}$, $B_0' = -\alpha \frac{\partial Bi(\eta_0)}{\partial x}$, $A_W = Ai(\eta_W)$, $B_W = Bi(\eta_W)$,

$A_W' = -\alpha \frac{\partial Ai(\eta_W)}{\partial x}$, and $B_W' = -\alpha \frac{\partial Bi(\eta_W)}{\partial x}$. The transmission probability T(ℰ) is **exact** for the linear potential barrier and is valid over all positive tunneling energies.

The quantum transmission probability T(ℰ) represents the fraction of incident carriers at an energy level that penetrate the Schottky barrier. The source of electrons incident on the metal and semiconductor side of the barrier are assumed to be in thermal equilibrium and maintain a Fermi-Dirac distribution in energy. The approximation for the metal flux source is quite good. This model assumes that electrons penetrating the Schottky barrier from the semiconductor come from a Fermi-Dirac distribution of electrons at a given scattering distance L_S from the metal/semiconductor interface. Electrons that transport across the barrier are not in equilibrium with the lattice until sufficient scattering events occur to randomize the electron momentum and equilibrate its energy. The length L_S is the average distance traversed by an electron scattered from its equilibrium Fermi-Dirac distribution in the semiconductor across the barrier into the metal. The semiconductor flux source is similar to the metal flux source except for the relative position of the Fermi level and the density of states in the semiconductor. The incident flux of electrons from the metal on the Schottky barrier is found by integrating the velocity distribution of the source electrons. Following [11], the total flux is defined as

$$\vec{N} = \frac{1}{4\pi^3} \int_{k_x} \int_{k_y} \int_{k_z} \vec{v} f_n dk_x dk_y dk_z \tag{5}$$

where f_n is the occupation probability of a state at energy ℰ. Converting to an integration over momentum using $\vec{p} = \hbar \vec{k}$ to find the incident flux (x direction momentum) with momentum in the range p_X to $p_X+\partial p_X$, and transforming to polar coordinates yields the total number of incident electrons on the Schottky barrier per unit time $N_{ms}(\mathcal{E}_x) = C_m \ln\left(1+\exp\left[-\frac{\mathcal{E}_x - \mathcal{E}_f}{kT}\right]\right)$ where $C_m = \frac{m_m kT}{2\pi^2\hbar^3}$.

To obtain the total current density (dropping the x direction dependence) transporting across the barrier from the gate metal, the incident penetrating flux is integrated over the energy distribution to yield $J_{ms} = \int_0^\infty N_{ms}(\mathcal{E}) T(\mathcal{E}) d\mathcal{E}$. The incident flux of electrons from the semiconductor is similar to the metal flux source except the Fermi level is at a potential V_a relative to the metal Fermi level and the density of states is zero below the conduction band. This model neglects tunneling effects from the valence band. The incident source flux from the semiconductor is

$N_{sm}(\mathcal{E}) = C_s \ln\left(1+\exp\left[-\frac{\mathcal{E}_x - \mathcal{E}_f + V_a}{kT}\right]\right)$ where $C_s = u(\mathcal{E} - \phi_b(E) + qEL_s)\frac{m_t kT}{2\pi^2\hbar^3}$ and u(ℰ) is a step function. The current density from the semiconductor to the metal is thus $J_{sm} = q\int_0^\infty N_{sm}(\mathcal{E}) T(\mathcal{E}) d\mathcal{E}$. The total net gate current density $J_t = J_{ms} - J_{sm}$ is the difference between the (reverse J_{ms}) metal to semiconductor and the (forward J_{sm}) semiconductor to metal current density. The avalanche ionization model used in the next

section is based on the formula by Okuto and Crowell[12]. It is field and temperature dependent and fits the theoretical results of Baraff[13] and measurements.

III. MESFET Model

The gate tunneling model is applied to a state-of-the-art quasi-two dimensional, physics based MESFET model[14] which solves the Poisson and drift-diffusion equations (in the gate region of the device) using a minimum number of approximations. For simplicity, it is assumed that MESFET breakdown occurs at the gate edge and within the gate depletion region where the highest electric fields exist. Breakdown in the substrate of a properly designed device should be minimal.

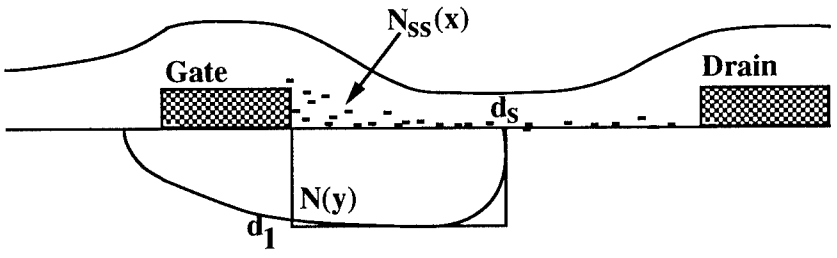


Figure 2: Cross section of a typical MESFET showing depletion extension region approximation, the depletion depth d_1, the depletion extension length d_S, and surface charge $N_{SS}(x)$.

Avalanche generation (and holes) that occurs at the drain edge is also neglected in this work. Breakdown in the gate region of typical MESFETs is the focus of this work and is assumed to be the dominant breakdown mechanism in MESFETs. The combined MESFET and gate breakdown model are coupled to a general purpose matching circuit by the harmonic balance technique[16]. The new model predicts the DC and RF performance of typical MESFETs. The model calculates the average channel temperature based on DC power dissipation and the device thermal resistance. Simulated DC-IV characteristics allow varying channel temperature with drain current and bias voltage. RF (and RF-IV for viewing dynamic load lines) simulations determine the channel temperature based on the drain bias and average DC drain current which can increase with strong RF power drive. The model assumes that surface traps maintain a constant occupation level under RF operation and does not predict dispersion effects. The surface trap occupation does not vary during RF simulation and is dependent upon the MESFET DC bias and operating temperature[17]. The effects of surface Fermi level pinning in the region beyond the gate surface depletion region are neglected for two reasons: (1)some of the trapped surface charge locally couples to the donor ions in the surface deplation region and will not directly effect the drain to gate potential (and thus the breakdown voltage) and (2) a properly passivated MESFET will have a channel resistance dominated by the gate depletion region and not the interelectrode resistance caused by the surface depletion. The trapped surface charge within the gate depletion region is assumed to affect the strength of the electric field (slope of the barrier) in the barrier region but not the triangular shape of the barrier.

To incorporate the tunnel breakdown model, the region between the gate and drain is included to determine the effect of the depletion extension and surface charge on the electric field strength at the gate edge. To include this

region in the MESFET model, the extended depletion region shown in Figure 2 is approximated by a rectangular area. The depletion depth d_1 and the depletion extension width d_s determine the size of the region and are drain and gate bias dependent. The depletion depth d_1 and the depletion extension width d_s are affected by the surface trapped charge and are determined in [17]. The occupied surface charge density $N_{ss}(x)$ varies along the surface due to the injection and subsequent capture of tunneling electrons by surface trapping states[18]. Gauss' law is used to determine the electric field that terminates on the fraction f_G of the gate length L_g due to the net charge in the extended depletion region and the trapped surface charge. Applying Gauss' law to this region, and assuming the electric field is distributed uniformly at the gate edge where the majority of the tunneling occurs, yields

$$q\int_0^{W_s}\int_0^{d_1}\int_0^{d_s} N_D(y) - N_{ss}(x)\delta(y)\,dxdydz = \varepsilon\int_0^{W_s}\int_0^{f_g L_g} E_g\,dydz \tag{6}$$

where W_g is the total gate width. The solution of Gauss' law assumes that most of the electric field diverging from the enclosed square Gaussian surface in Figure 2 terminates on the edge of the gate and is proportional to the density of surface and gate depletion charge. The strong two-dimensional nature of the fields is contained within the integral solution. Simplifying the solution based on the previous approximations, the electric field at the gate edge E_g is $E_g = \frac{q}{\varepsilon f_G L_g}\left[d_s\int_0^{d_1} N_D(y)dy - \int_0^{d_s} N_{ss}(x)dx\right]$. The integrals $\int_0^{d_1} N(y)dy$ and $\int_0^{d_s} N_{ss}(x)dx$ are easily computed from the MESFET model formulation[14][17]. The electric field strength at the gate edge affects the tunneling probability of the barrier. The total tunneling current flowing from the gate edge is $I_t = f_G L_g W_g J_t(E_g) = I_{tf} - I_{tr}$. Tunneling electrons are injected into the gate depletion region and can initiate avalanche ionization. The density of the tunnel injected charge $n_t = \frac{J_t(E_g)}{qv_s}$ is determined from the tunnel current density $J_t(E_g)$ and the velocity of the carriers v_s. The avalanche generated current density due to charge tunneling into the strong field of the depletion extension region between the gate and drain is found from the quasi-static drift-diffusion equation where the avalanche ionization coefficient[12] $\alpha_{ii}(E_g)$ is dependent on the field at the gate edge. The intrinsic carrier concentration n_i, and bulk lifetime τ are included. For simplicity, the avalanche field and the tunnel injected charge density are assumed to be nearly uniform in the linear barrier region. The avalanche generated current density is

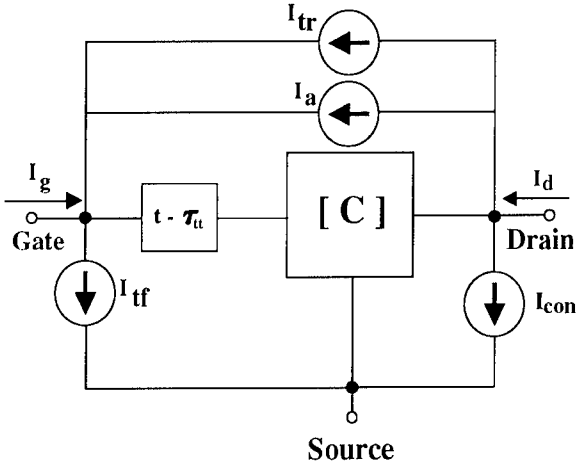


Figure 3: The large-signal representation of the MESFET including the forward I_{tf} reverse I_{tr} tunneling sources and the avalanche generation source I_a. The drain conduction current I_{con}, gate charging time τ_{tt}, and the capacitance matrix [C] are included.

approximated by $I_a \approx qd_s \frac{n_t - n_i}{\tau} - q(n_t + n_i)v_s d_s \alpha_{ii}(E_g)$. The coupled tunnel and avalanche breakdown mechanisms are represented as current sources on the new equivalent, large signal MESFET model [17] in Figure 3. The forward (I_{tf}) and reverse (I_{tr}) components of the tunnel current density and the avalanche generated (I_a) currents are indicated as current sources between the gate, drain and source.

IV. Results

High efficiency MESFET operation results from the ability to optimally shape the drain current and voltage waveforms. This is typically accomplished by harmonic tuning. Proper tuning of the MESFET adjusts the dynamic load line within the IV plane to allow a maximum drain current and voltage. Reverse and forward gate conduction place limits on the allowed loadline operating area within the IV plane. To examine the effects of breakdown and forward gate conduction, two experimental power MESFETs, fabricated with the same specifications but taken from different wafers, are simulated. Both devices (utilizing buried channel profile with recessed gate) were tuned for maximum efficiency under Class AB conditions (Vd = 9V) at 9.5GHz. The RF performance of the MESFETs was simulated using the MESFET model with the new tunnel avalanche breakdown mechanism. The model parameters were extracted from the measured data provided by the foundry. Table 4.1 in [17] presents a comparison and discussion of the measured and simulated date. Data supplied consisted of doping profiles, DC-IV characteristics, and RF power performance[17]. The parameter extraction routine is based on a new quasi-newton minimization algorithm[19] that adjusts the physical model parameters, within constrained and physically reasonable bounds, to minimize the difference between measured and simulated DC and RF data. Some of the main adjustable parameters include device parasitics, doping profile, surface charge density, tunneling effective mass, matching circuit impedances, channel mobility and saturation velocity. The high efficiency performance of experimental Device A, shown in Figure 5, is greater than the performance of experimental Device B, shown in Figure 4. The average DC gate currents for device A and B are compared in Figure 7 and Figure 6, respectively. The negative gate current indicates[15] that there is substantially more breakdown occurring in Device B than in Device A. Decreased breakdown voltage in Device B reduces the maximum efficiency the amplifier can attain by causing premature saturation. Measurd and simulated DC gate breakdown characteristics are compared in Figure 8 and 9. Device A has a higher DC gate breakdown voltage in Figure 9 as compared to Device B as seen in Figure 8. The Device A DC breakdown voltage is lower than Device A, but also the RF-IV (which are under constant temperature) breakdown voltage is also lower as can be seen form the dynamic load lines at the peak operating efficiency power level. The simulated dynamic RF-IV and load line for Device A at the input power level (at peak efficiency) of 21dBm is shown in Figure 10. The device operation is dominated by the balance between the forward gate conduction mechanism and reverse breakdown as can be seen from the position of the load line. Device B dynamic RF-IV and loadline is shown in Figure 11 at 21dBm input power (at peak efficiency). Device B RF performance is dominated by reverse gate breakdown as demonstrated by the position of the load line. The load line extends significantly into the breakdown region. The

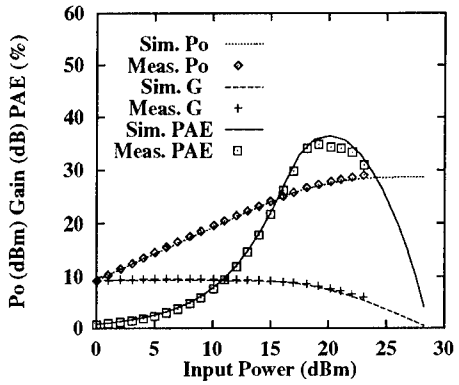


Figure 4: The RF output power P_O, power added efficiency PAE, and gain measured and simulated results of experimental MESFET B.

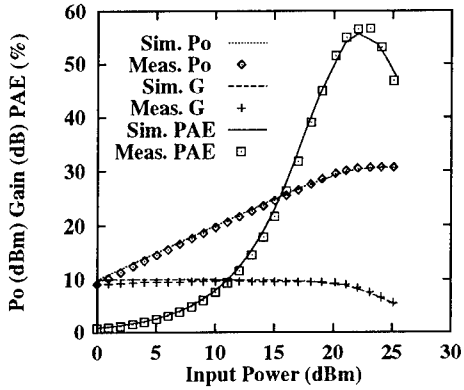


Figure 5: The RF output power P_O, power added efficiency PAE, and gain measured and simulated results of experimental MESFET A.

Device B load line encounters the breakdown region at lower input power levels than in Device A, which identifies the main difference between the two devices. The excellent match between measured and simulated data presented demonstrates the effectiveness of the new breakdown model.

V. Conclusion

A new gate breakdown model for studying high efficiency MESFET operation is presented. The breakdown model is based on tunnel initiated avalanche ionization in the high field region of the gate depletion region. This is the first breakdown model that couples the avalanche and tunneling mechanisms with temperature effects which can be used to perform RF and DC simulations. The effects of distorted surface band structure on the tunneling effective mass and breakdown voltage are seen for the first time. The coupled breakdown and MESFET models are used to explain the complex behavior of large-signal MESFET operation. Good agreement was obtained from the simulated and measured results.

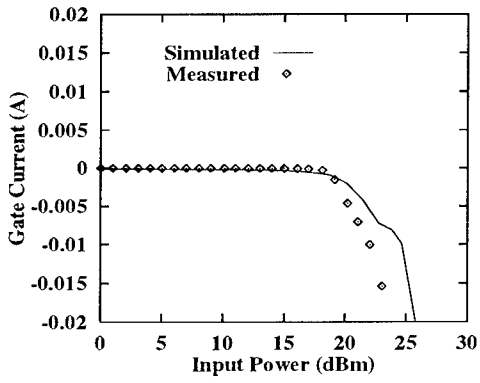


Figure 6: The RF average gate current measured and simulated results of experimental MESFET B.

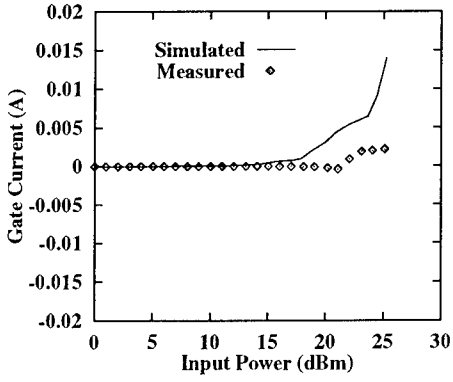


Figure 7: The RF average gate current measured and simulated results of experimental MESFET A.

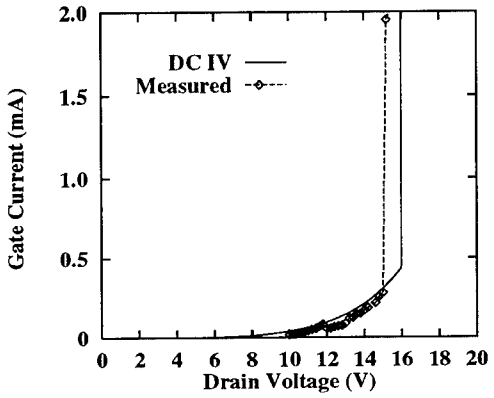


Figure 8: Measured and simulated DC gate current for Device B showing breakdown characteristics under pinch-off conditions.

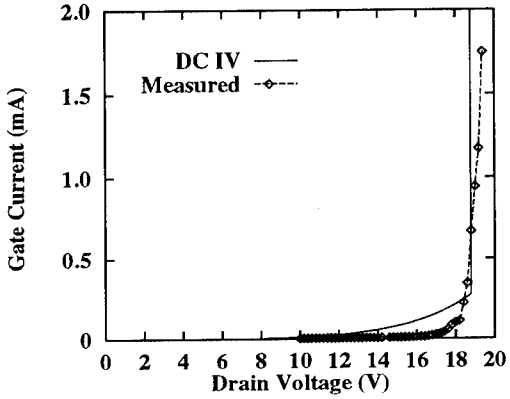


Figure 9: Measured and simulated DC gate current for Device B showing breakdown characteristics under pinch-off conditions.

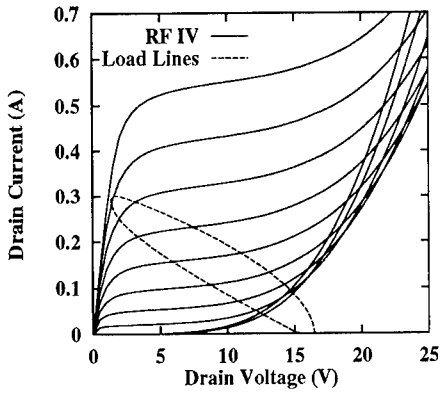


Figure 10: The dynamic RF-IV and load line for Device B at the input power of 21dBm (maximum efficiency of 36%).

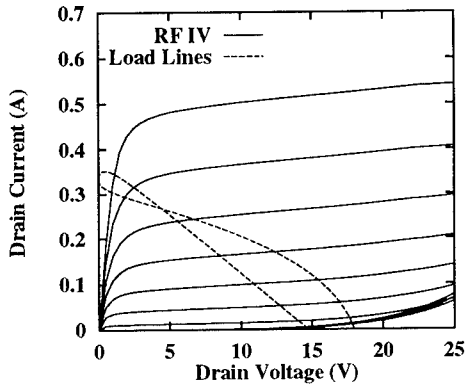


Figure 11: The dynamic RF-IV and load line for Device A at the input power of 21dBm (maximum efficiency of 56%)

References

[1] T.A. Winslow and R.J. Trew, "An Investigation of Efficiency Limiting Mechanisms in GaAs MESFET Amplifiers," 1992 European International Microwave Conference Digest, pp. 1195-1200.

[2] R.J. Trew and U.K. Mishra, "Gate Breakdown in MESFETs and HEMTs," IEEE Electron Device Letters, vol. 12, no. 10, pp. 524-526, Oct. 1991.

[3] R.J. Trew and T.A. Winslow, "A gate breakdown mechanism in MESFETs and HEMTs," Mat. Res. Soc. Symp. Proc., vol. 241, pp. 193-198, 1992.

[4] M. Fukuta, T. Mimura, H. Suzuki, and K. Suyama, "4-GHz 15-W Power MESFETs," IEEE Trans. on Electron Devices, vol. 25, no. 6, pp. 559-562, June 1978.

[5] R. Yamamoto, A. Higashisaka, and F. Hasegawa, "Light Emission and Burnout Characteristics of GaAs Power MESFETs," IEEE Trans. on Electron Devices, vol. 25, no. 6, pp. 567-573, June 1978.

[6] J. Ashworth, P. Lindorfer,"Analysis of the Breakdown Phenomena in GaAs MESFETs," Proc. ESSDERC, pp. 241-244, 1990.

[7] Y. Wada, M. Tomizawa, "Drain Avalanche Breakdown in Gallium Arsenide MESFETs," IEEE Trans. on Electron Devices, vol. 35, no. 11, Nov 1988.

[8] R. Stratton, "Theory of field Emission from semiconductors," Physical Review, vol. 125, no. 1, pp.67-82, 1962.

[9] F.A. Padovani and R. Stratton, "Field and thermionic-field emission in Schottky barriers," Solid-State Electronics, vol. 9, pp. 695-707, 1966.

[10] C. R. Crowell and S. M. Sze, "Current Transport in Metal-semiconductor barriers," Solid State Electronics, vol. 9, pp. 1035-1048, 1962.

[11] W.F. Leonard and T.L. Martin, Electronic Structure and Transport Properties of Crystals. NY: Krieger Publishing, 1980.

[12] Y. Okuto, and C.R. Crowell, "hreshold Energy Effect on Avalanche Breakdown Voltage in Semiconductor Junctions," Solid State Electron., vol. 18, pp. 161-168, 1975.

[13] G.A. Baraff, "Distribution functions and ionization rates for hot electrons in semiconductors," Physical Review, vol. 128, pp. 2507-2517, 1962.

[14] M.A. Khatibzadeh and R.J. Trew, "A large-signal analytic model for the GaAs MESFET," IEEE Trans. Microwave Theory and Techniques, vol. 36, pp.231-238, 1988.

[15] W.R. Frensley, "Power Limiting Breakdown Effects in GaAs MESFETs," IEEE Trans. on Electron Devices, vol. 28, no. 8, pp. 962-970, August 1981.

[16] M.S. Nakhla and J. Vlach, "A Piecewise Harmonic-Balance Technique for Determination of Periodic Response of Nonlinear Systems," IEEE Trans. Circuits Syst., Vol. CAS-23, pp. 85-91, Feb. 1976.

[17] T.A. Winslow, "A Gate Breakdown Model for GaAs MESFETs," Ph.D. Dissertation, North Carolina State University, 1993.

[18] T.M. Barton and P.L. Ladbrooke, "The Role of the Device Surface in the High Voltage Behavior of the GaAs MESFET," Solid State Electronics, vol. 29, no. 8, pp.807-813, 1986.

[19] T.A. Winslow, R.J. Trew, P. Gilmore, and C.T. Kelley, "Simulated Performance Optimization of GaAs MESFET Amplifiers," Proc. of the 1991 IEEE-Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits, pp. 393-402.

[20] C. Tsironis, "Prebreakdown Phenomena in GaAs Epitaxial Layers and FETs," IEEE Transactions on Electron Devices, vol. 27, no. 1, pp. 277-282, Jan., 1980.

[21] S. Wemple, W. Niehous, H. Fukui, J.C. Irvin, H.C. Cox, J.C.M. Hwang, J.V. DiLorenzo, and W.O. Schlosser, "Long-Term and Instantaneous Burnout in GaAs Power FETs: Mechanisms and Solutions," IEEE Transactions on Electron Devices, vol. 28, no.7, pp.834-840, July 1981.

High Field Drift Domains in GaAs and InP Based Heterostructure Field Effect Devices

E. Kohn, S. Strähle, D. Geiger, U. Erben

University of Ulm
Dept. of Electron Devices and Circuits
D-89069 Ulm, Germany

Abstract

Modelling the high field drift region of HFET's as a drift capacitance between gate and drain in series with the gate capacitance allows to estimate the extension of the drift region, which determines feed back and output conductance, thus relating the microwave power gain to the device structure. The technique is applied to various GaAs and InP based FET structures.

I Introduction

To reach the RF-performance limits of high speed devices, the electrical equivalent circuit model parameters need to be directly correlated with the physical design parameters as the geometrical layout, the heterostructure configuration, and the charge carrier transport properties. High switching speeds have been achieved by optimizing the H_{21} cut-off frequency. Because of the short circuited output condition however, the feedback behavior and output conductance, which are related to the gate-drain high field drift region, enter only marginally.

However, the gate to drain high field region conditions dominate power gain, breakdown and overall signal delay. The benefit of the high field drift region and domain formation has been recognized since long, and its impact on power gain and breakdown has been extensively analysed analytically and numerically in the past [1,2]. Due to the 2D-nature of this region, the models used are generally limited to rather specific device structures especially in respect to the recess configuration. In addition, the domain formation depends on the electron dynamics in the specific heterostructure. Thus, an effective drift region length as figure of comparison might be very helpful for the comparison of various design concepts. The analytical treatment of the drift region as space charge layer leads to the lateral spreading model [3]. The treatment as stationary high field domain leads to a domain capacitance of the extension of the lateral domain space charge [4]. Thus, it is suggested to model the drift region by its space charge capacitance.

II Feed-Back Circuit Model

In the following an effective drift region length is defined from S-parameter measurements and standard small signal parameter extraction related to the equivalent circuit. For our purpose, the feed back capacitance C_{gd} is split into a

0-7803-0894-8/93 $3.00

series connection of two capacitors, the input capacitor and a drift capacitor, which is now related to the drift region between gate and drain:

$$1/C_{gd} = 1/C_{gs} + 1/C_d \qquad \text{eq.1}$$

Evidence for this series circuit model comes from the linear feed-back correlation:

$$C_{gs}/C_{gd} = 1 + \beta\, g_m/G_{ds} \qquad \text{eq.2}$$

which has been experimentally verified for a variety of HFETs over a wide range of bias conditions [5,6]. The confinement factor is found to be β=0.3 for optimum cases. Dividing eq.2 by C_{gs} yields the feed-back series circuit of eq.1 with:

$$C_d = G_{ds}C_{gs}/\beta\, g_m \qquad \text{eq.3}$$

a capacitance associated with the extension of the drift region. Thus, here the drift capacitance is directly linked to the output conductance. The physical interpretation is shown in fig.1.

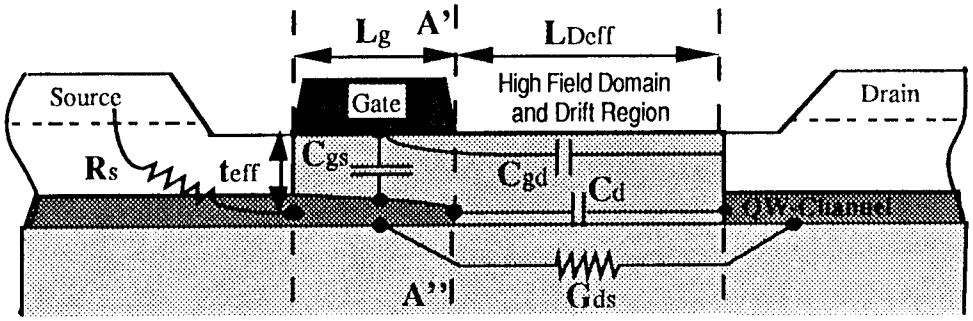


Fig.1:
Active area device cross section with physical interpretation of drift capacitance

The devices considered are short gate length FET's with a high aspect ratio to suppress 3D-effects. For such QW-channel structures C_{gs} may be approximated as parallel plate capacitor with t_{eff} as the gate to channel separation. Assuming that the lateral channel field concentrates between the bottom of the channel and the surface, C_d can also be considered a lateral parallel plate capacitor. Evidence for this is found from modelling the higher valley hot electron distribution in HEMT's [7]. Then the two capacitors may be approximated as follows:

$$C_{gs} = \frac{dQ_{source}}{dV_{gs}} \approx \frac{\varepsilon\, w\, L_g}{t_{eff}} \qquad \text{eq.4}$$

$$C_d = \frac{dQ_{drain}}{dV_{drift}} \approx \frac{\varepsilon\, w\, t_{eff}}{L_{D_{eff}}} \qquad \text{eq.5}$$

Here, V_{drift} represents the internal voltage across the drift region, and L_{Deff} marks the lateral extension of the drift capacitor towards the drain in the parallel plate capacitor model. This allows to extract a drift region length, expressed in terms of the input and feed back capacitance, knowing only the geometrical layout of the device and the average refractive index of the heterostructure:

$$L_{Deff} = L_g \left(\frac{\varepsilon w}{C_{gs}} \right)^2 \left(\frac{C_{gs}}{C_{gd}} - 1 \right) \qquad \text{eq.6}$$

In cases, where eq.2 applies, L_{Deff} can also be related to the output conductance:

$$L_{Deff} = L_g \left(\frac{\varepsilon w}{C_{gs}} \right)^2 \left(\beta \frac{g_m}{g_{ds}} \right) \qquad \text{eq.7}$$

Eq.7 leads to a direct proportionality between the extension of the drift region and the output resistance (at constant g_m and C_{gs} in this first order model). Thus, L_{Deff} can indeed serve as figure of merit for the intrinsic power gain performance.

Two critical assumptions need to be made for this model:

(1) The gradual channel region is considered an intermediate node due to the small internal voltage drop in the gradual channel region, and

(2) The vertical line (A´)-(A´´) has to represent an equi-potential line for C_d, but at the same time needs to consume the V_g-potential drop. This approximation will hold, if the influence of the gate-source voltage on the drift region is small in comparison to the gate-drain bias, thus if the gate field acts only as fringing field on C_d.

III PM-GaAs HFET Analysis

IIIA Drift Region Characteristics

The development of the drift region can be monitored for the entire field of operation analogue to the FET output characteristics. This is illustrated in fig.2 for a high current recessed gate GaAs doped channel pseudomorphic HFET with highly doped cap layer and 0,25µm gate length [9]. One can see, that in this model the drift region starts to develop already in the linear regime, because already here C_{gs} becomes non-symmetrical in respect to source and drain. In saturation, the drift region reflects two different conditions:

At low current levels the drift region acts as a lateral space charge layer; the space charge being the donor sheet charge in the channel and active part of the supply layer. Opening the channel and allowing current to flow at saturated velocity, the donor charge will in part be compensated by the electronic charge and L_{Deff} will

increase with I_D. On the other hand, the device will be deeper in saturation at low current level, increasing L_{Deff}. At small drain bias the latter effect will be pronounced and L_{Deff} decreases with I_D, at high drain current the trend should be reversed and L_{Deff} should increase with I_D. At high current levels, the drift region reflects a stationary domain with a laterally varying domain charge.

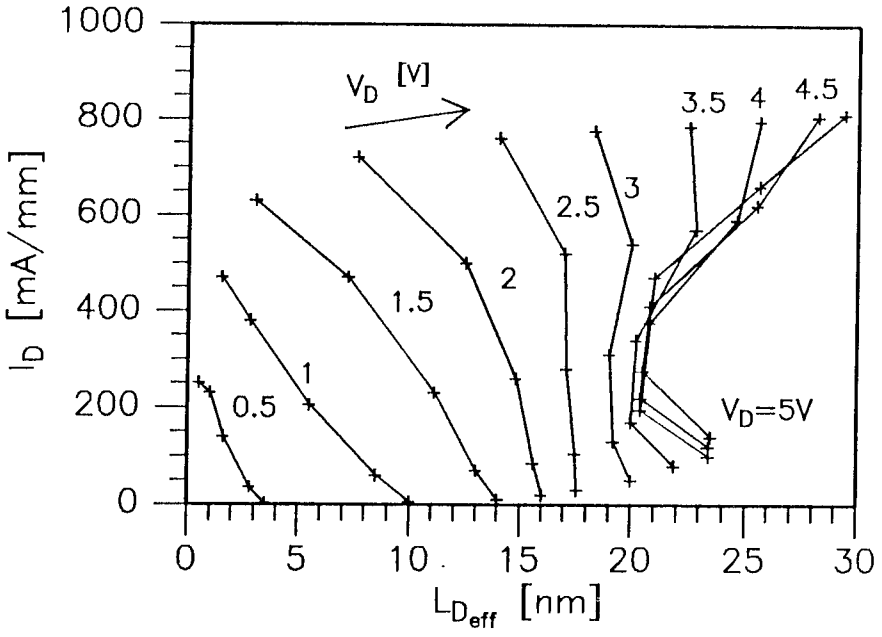


Fig.2:
Drift region extension L_{Deff} as function of output current and applied drain bias for 0.25µm gate length device. Gate bias parameters as in fig.3.

These trends seem to be observed, however, care has to be taken with this interpretation, because only part of the extrinsic voltages act on the drift region. Therefore, the intrinsic voltages are determined at each bias point (using RF, cold-FET and DC evaluation) and an intrinsic drift region characteristics is plotted in fig.3.

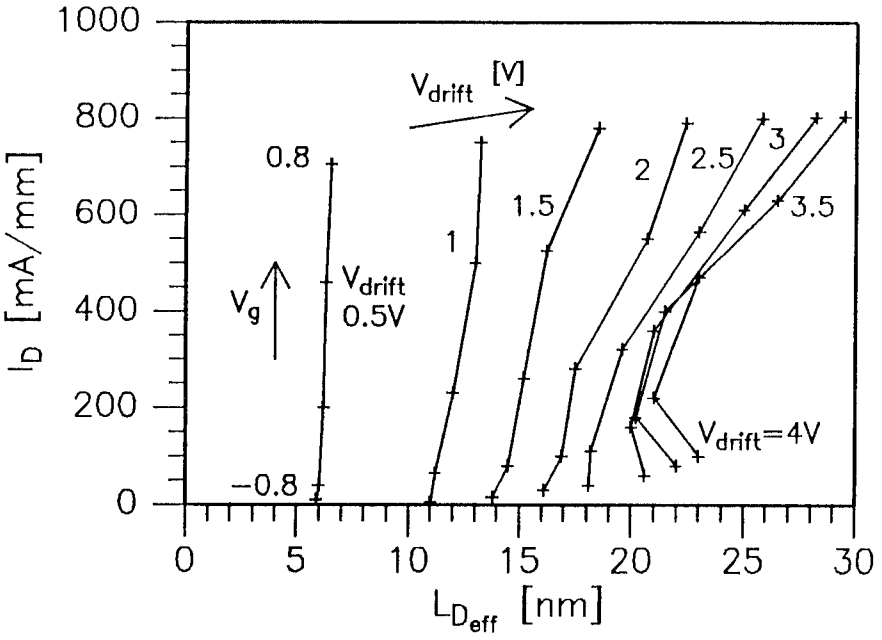


Fig.3:
Development of L_{Deff} as function of drain current with intrinsic bias conditions as parameter.

Changes are pronounced at high currents, where L_{Deff} is limited due to the influence of the parasitic series resistances. This leads to a decrease in L_{Deff} and increase in domain capacitance (besides a reduction in g_m). The following simple model was used to estimate the drift region voltage:

$$V_{drift} = V_D - (R_S+R_D)I_D - E_cL_g + \alpha(V_g - R_SI_D) \qquad \text{eq.8}$$

The linear regime is (largely) taken care of by substracting E_cL_g (E_c=4kV/cm); and $\alpha(V_g - R_SI_D)$ is the fraction of V_g acting laterally on the drift region. Therefore, α may be interpreted as a fringe field factor. The effective drift region length is determined by eq.6.
α is the only fitting parameter in this model taking care of the 2D-field distribution. It is determined in the following way: The lateral spreading model [3] predicts a $V_{drift} \propto N_S L_{Deff}^2$-relationship. A large α however tends to linearize this behavior. Plotting the data in a double logarithmic scale, α is varied to fit a linear slope with s=2 near pinch-off (where the electronic domain charge is still neglectible) and in the region, where the space charge layer is fully developed ($L_{Deff} \geq 10$nm). This is however only possible with a small $\alpha \leq 0.1$. (see fig.4).

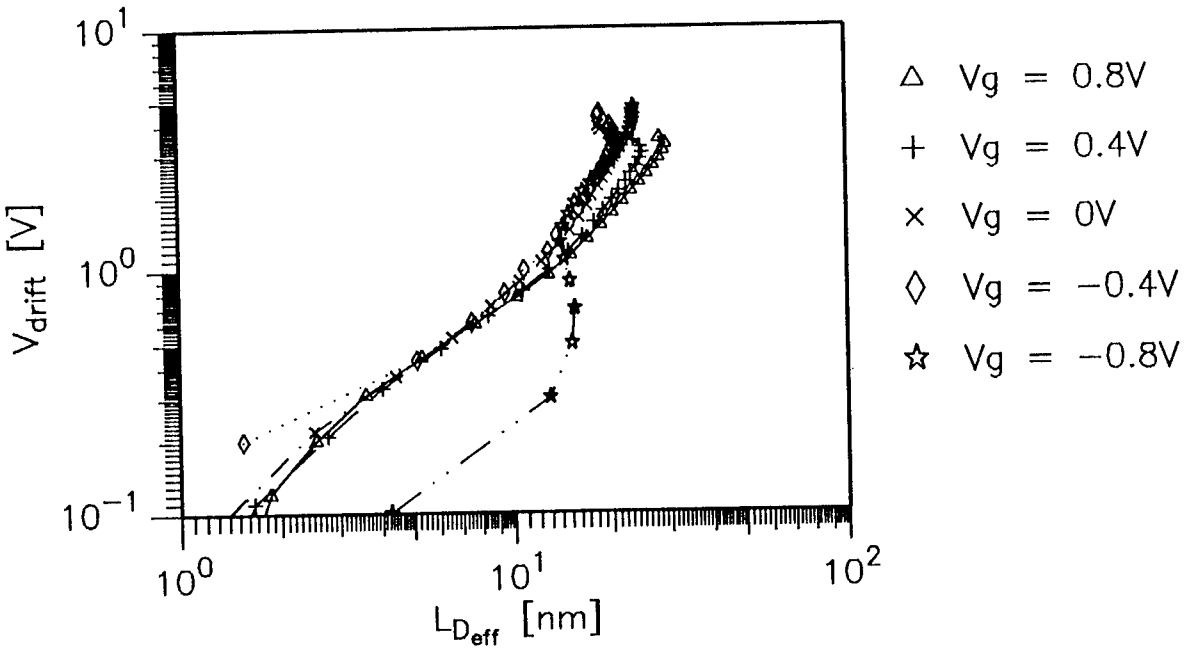


Fig.4
L_{Deff} vs. V_{drift} (α=0.1 in eq.8) in double logarithmic presentation. The curve for Vg=-0.8V is in pinch-off with noticeable leakage current for $V_{drift} \geq 1$V.

At high currents L_{Deff} stretches out further (at the same V_{drift}) consistent with the influence of the electronic domain charge. The slope decreases slightly, possibly due to the lateral domain charge distribution (indicating a higher density near the gate). The extension is limited at 20 to 30nm, where breakdown occurs. It seems that the drift region touches the gate recess edge at this point. Examining

the structure by SEM shows a gate to recess edge separation of approx. 40nm. Taking into account a gate fringe capacitance and current crowding at the edge, the first order model data seem to be in reasonable agreement with the structural parameter.

IIIB Effective Channel Velocity

Four main parameters determine the average electron channel transit velocity:
(a) the intrinsic overshoot and high field velocity regions,
(b) the charge control and modulation efficiency,
(c) the delay and capacitive loading of the drift region, and
(d) the parasitic elements, which have been taken care of by the equivalent circuit model. Of specific interest here is the influence of the drift region on v_{eff}. Therefore, the differential f_T-approach after [8] has been used to monitor v_{eff} with L_{Deff} (see fig.5).

In the linear regime v_{eff} increases, C_{gs} becomes non-symmetric in respect to source and drain and an effective drift region starts to develop. From fig.2 is seen that the simple model leads here to a sudden development of L_{Deff} with V_D. Reaching saturation v_{eff} peaks and decreases only moderately. This means that the overshoot regime dominates the performance up to breakdown in this narrow gate recessed device. The charge control determines the value of v_{eff} by g_m.

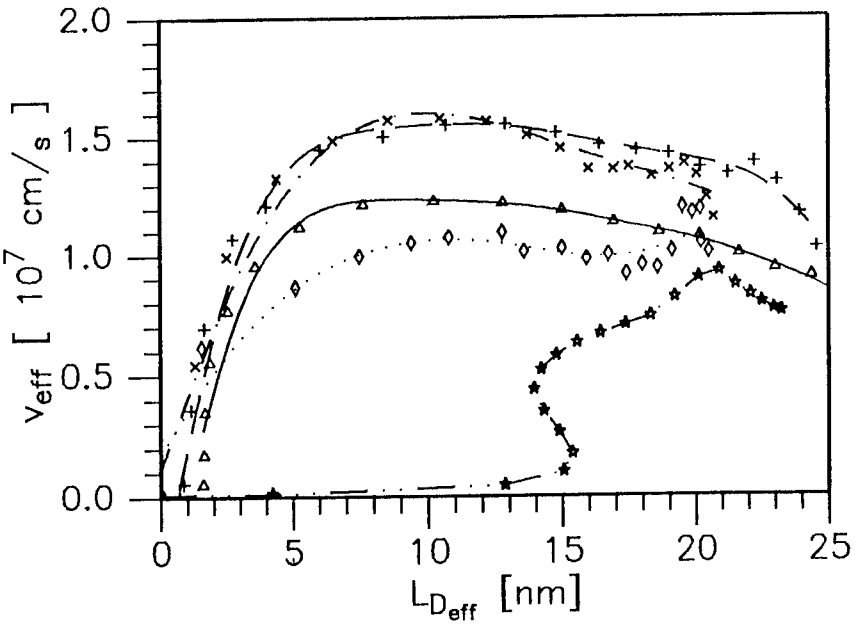


Fig.5:
Variation of the effective channel electron transit velocity with drift region length and gate bias as parameter.
Symbols as in fig.4.

Near pinch-off v_{eff} decreases rapidly. However, at high drain bias the channel becomes conductive again, which is also seen in the RF behavior until breakdown is reached.

IIIC Output Conductance

Eqs.1, 2 and 7 link the output conductance to the feed back capacitance and drift region length. From these equations a linear relationship is expected and indeed seen between $1/G_{ds}$ and L_{Deff} (see fig. 6).

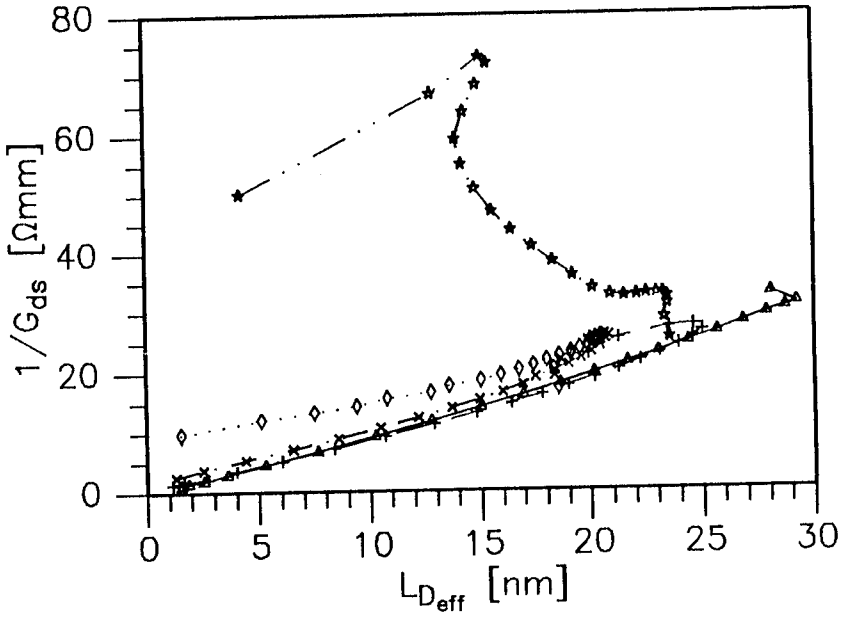


Fig.6:
Correlation of output resistance $1/G_{ds}$ with the drift region length for various gate bias conditions. Symbols used as in fig.4.

Near pinch-off the output conductance is essentially suppressed. However, once channel current starts to flow at high drain bias $1/G_{ds}$ reduces rapidly onto the general curve.

IV Comparison with InP based HEMT Devices

InP based devices have shown considerably higher f_{max}/f_T-ratios at high f_T than GaAs based devices. This fact has widely been associated with the high InGaAs channel electron mobility and peak velocity. In addition, it has been suggested that the drift region extends further under comparable bias conditions and reduces the feed back capacitance and output conductance [10]. Thus, the above technique has been applied to such structures to verify the latter point. In essence, it implies that the hot carriers follow mainly the broad velocity/field relationship of the InAlAs or InP barrier material, rather than the steep curve of the InGaAs channel. Two InP based device structures are considered in this experiment and compared to the GaAs based HFET presented in section III:

(1) A structure with InGaAs channel and InGaAs/InAlAs superlattice sub-channel and buffer. The 2DEG-density in this structure is $N_S=2.8*10^{12}cm^{-2}$; the gate length is 0.3µm. Details of this structure have been published elsewhere [10].

(2) An InGaAs-InP stepped QW-channel sandwiched between two InAlAs barriers. The 2DEG-density located in the InGaAs sub-channel is $N_S=2.2*10^{12}cm^{-2}$; the gate length is 0.6µm. Details of this structure are given in [11].

As a basis for comparison gate bias conditions near the maximum f_T-condition have been chosen with drain bias as extrinsic and drift voltage as intrinsic variable. The development of L_{Deff} and $1/G_{ds}$ with V_{drift} is shown below.

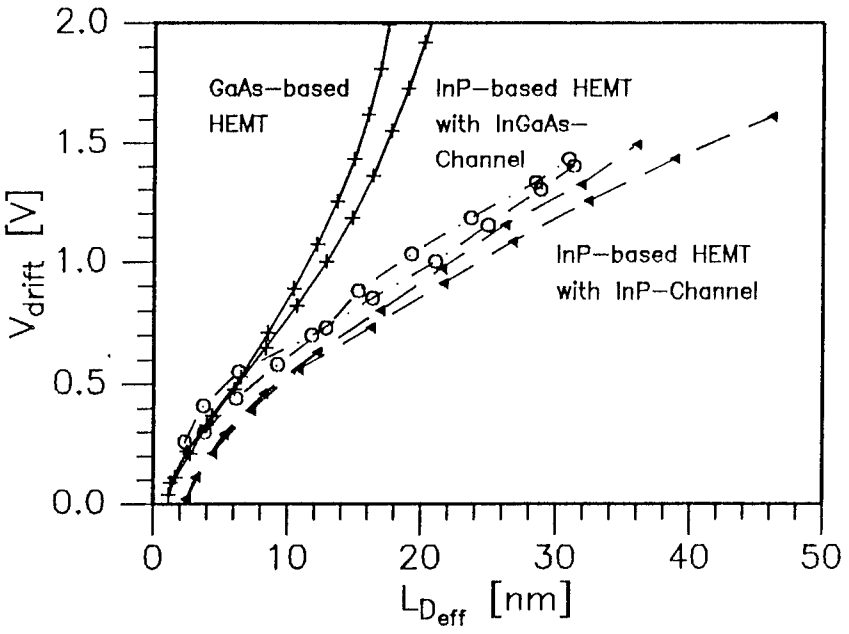


Fig.7:
Development of drift length with intrinsic drift voltage for GaAs and InP based HFETs as described in the text. In each case two gate bias conditions near the maximum f_T bias point are presented.

It can be seen from fig.7 that the drift region in both InP based devices develops essentially deeper than in the case of the GaAs device. Their drift region develops linear even by fitting the curves with a small α=0.1 (in a procedure as described in section IIIA). An explanation within the frame of this model for this distinctly different behavior to the GaAs based devices is a lateral space charge distribution in the channel (hyperbolically decreasing towards the drain) due to the electron velocity distribution in the overshoot regime in both heterostructures.

The output conductances are compared in fig.8.

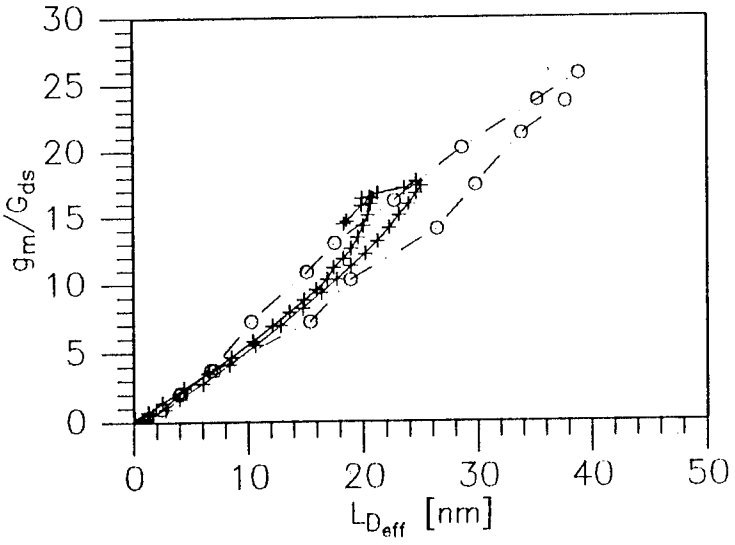


Fig.8:
Output conductance vs. drift region length for the GaAs based device and the InP based device (1) at the bias points of fig.7.

In this case an identical behavior is observed. No benefit from the deep InGaAs/InAlAs QW is seen. This is indeed expected from a more detailed analysis, which shows that the ratio g_m/G_{ds} is only dependent on the product of the structural aspect ratio and the drift region aspect ratio and independent of the material basis and heterostructure configuration [6].

Thus, the high f_{max}/f_T ratio of InP based HFET's can essentially be linked to the wider extension of their drift region. At the same time a high $f_T L_g$-product is observed. This indicates a high average transit velocity in this region. An analysis in this respect may follow that of section IIIB.

V Conclusion

A novel concept has been presented relating C_{gd} and G_{ds} (which are the key intrinsic parameters for the power gain cut-off frequencies) of QW-channel FET devices to the intrinsic gate-drain high field drift region. The drift region is described by a drift capacitor and with the help of a simple geometrical approximation a bias dependent effective drift region length can be defined and used as figure of merit to compare a variety of devices. This is possible regardless of the technology used, because only S-parameter measurements and standard equivalent circuit element extraction are needed besides the physical gate length and width of the device. A pronounced difference in the development of the drift region is seen between GaAs based and InP based devices:

In the case of the GaAs device investigated, the drift region can be described as a lateral space charge layer, the extension being determined by the active channel donor charge up to rather high current levels. The extension of the drift region follows therefore largely a $L_{Deff} \propto (V_{drift})^{1/2}$-relationship, limiting the extension of the drift region at high drain fields. To utilize high overshoot velocities a narrow recess is therefore essential.

In the contrary, for InP-based devices a linear relationship $L_{Deff} \propto V_{drift}$ is seen even for high drain fields. One might speculate that in this case the drift region extension is determined by the lateral profile of the electronic domain charge. The straight proportionality would then lead to a hyperbolically decreasing net charge density towards the drain. It seems therefore that the high f_{max}/f_T-ratio at high f_T obtained with the InP based devices is mainly linked to a high velocity at high fields. Therefore it might be suggested that the hot carriers traverse the drift region mainly in the InAlAs and InP barrier material.

Correlating the extension of the drift region with the output conductance shows that for both cases a linear relationship holds. It seems that the mechanisms causing parallel conduction can be well linearized in the region of interest. No dependence of the output resistance on the specific QW-configuration is observed (in the open channel high field region) and the dominating influence of the aspect ratio product of the device is confirmed.

Considering both feed-back and output conductance, it can be concluded that the main reason for the high intrinsic gain cut-off frequencies of InP based HFET devices is the large drift region extension in conjunction with a high transit velocity.

Acknowledgement

The devices, which have been used in this study, have in part been fabricated with the help of the Daimler Benz Research Center Ulm. Many thanks are therefore to C. Woelk for providing the GaAs-based MBE grown heterostructures and B. Maile for the 0.25µm gate e-beam lithography. Many thanks are also to U. Spitzberg for assistance in the S-parameter measurements and data processing.

References

[1] R.W.H. Engelmann, C.A. Liechti; Bias Dependence of GaAs and InP MESFET Parameters; IEEE Trans. EL. Dev., ED-24 (1977), 1288-1296

[2] R.Wroblewski, G. Salmer, Y. Crosnier; Theoretical Analysis of the DC Avalanche Breakdown in GaAs MESFET's; IEEE Trans. El. Dev., ED-30 (1983), 154-159

[3] S.H. Wemple, W.C. Niehaus, H.M. Cox, J.V. DiLorenzo, W.O. Schlosser; Control of Gate-Drain Avalanche in GaAs MESFET's; IEEE Trans. El. Dev., ED-27 (1980), 1013-1018

[4] T.A. Fieldy; Analytical Modeling of the Stationary Domain in GaAs MESFETs; IEEEE Trans. El. Dev., ED-33 (1986), 874-880

[5] E. Kohn; A Semi-empirical Comparison of Key Dynamic Properties of GaAs and InP based HFET Structures; 17th Int. Sympos. of GaAs and Related Compounds 1990, I.o.P. Conf. Series 112, 411-416

[6] J. Dickmann, E. Kohn, S. Strähle, A. Wirsch, H. Künzel, H. Lee, H. Nickel; RF-Current De-Confinement in III-V- HFET's; 1991 IEEE/Cornell Conf. on Advanced Concepts in High Speed Semiconductor Devices and Circuits, Proc. 435-444

[7] J.L. Thobel, L. Baudry, P. Bourel, F. Dessenne, M. Charef; Monte Carlo Modeling of High Field Transport in III-V Heterostructures; GaAs Simulation Group VII Meeting, University of Leeds, April 22-23, 1993

[8] M. Feng, C.L. Lau, V. Eu; An Experimental Determination of Electron Drift Velocity in 0.5µm Gate-Length Ion-Implanted GaAs MESFET's; IEEE Electron Dev. Lett., 12 (1991), 40-41

[9] S. Strähle, M. Poese, X. Zhang, U. Erben, H. Thomas, D. Geiger, E. Kohn; "Structure and Performance of GaAs PM-HFET with High Current Densities"; European Heterostructure Technology Workshop 1992, Gregynon Hall (Wales), Sept 1992

[10] E. Kohn, J. Dickmann; Carrier Confinement and Feed Back Correlation in InAlAs/InGaAs HEMT's on InP Substrates; 3rd Int. Conf. on InP and Related Compounds; Cardiff (UK), 1991, Proc. 296-299

[11] S. Strähle, B. Henle, E. Mittermeier, U. Erben, P.K. Rees, E. Kohn; Two Layer Stepped-QW Channel HFET's on InP Substrate; This IEEE/Cornell Conf. Paper 7.2

High Temperature Operation of N-Type 6H-SiC and P-Type Diamond MESFETs

M. W. Shin, G. L. Bilbro and R. J. Trew

High Frequency Electronics Laboratory
Department of Electrical and Computer Engineering
North Carolina State University,
Raleigh, NC 27695 - 7911

ABSTRACT

Large signal RF and DC performance of n-type 6H-SiC and p-type diamond MESFETs has been simulated at various operating temperatures by a large-signal RF simulator using the harmonic balance technique and the two-dimensional device simulator, PISCES-IIB. The RF performance of SiC MESFET is predicted to be optimal in a temperature range slightly higher than room temperature. At room temperature the simulated SiC MESFET exhibits an output power of 3.5 W/mm for an operating frequency of 8 GHz with 16.5 dB gain and 44 % power-added efficiency at 24 dBm input power. In contrast to the SiC MESFET, the RF performance of the diamond MESFET is improved with temperature, but the current level is much lower than that in SiC in the entire temperature region investigated. The very different temperature dependencies of DC and RF performance in SiC and diamond MESFETs are attributed to the significant difference in the dopant ionization energies in SiC and diamond.

Introduction

Wide band-gap semiconductors have been given significant attention as potential materials for high-frequency, high-power, and high-temperature applications due to the high saturation velocity, high breakdown voltage, and high thermal conductivity [1]. It has been predicted that SiC and diamond possess the materials characteristics that may permit RF electronic devices with performance greater than what is available from devices fabricated from Si and GaAs [1]. It has been shown that diamond and SiC can provide substantial improvements in the power-handling capability when compared to conventional semiconductors [2]. A more critical evaluation based on existing packaging technology indicated that the device size of a high voltage 6H-SiC MOSFET can be significantly reduced compared to a Si device at high temperature operation [3].

The temperature dependence of the drain current level and thus the RF performance of the MESFET structure is mainly determined by : a) number of activated carriers in channel ; b) saturation velocity of the carrier; and c) mobility of the carrier. In particular, consideration of the dopant ionization process is critical in diamond and SiC for an accurate theoretical prediction on the high temperature device performance. This is due to the relatively high ionization energies of dopants in these materials. With the help of recent significant research efforts in SiC and diamond, the materials parameters and temperature dependencies of current transport mechanisms for these materials are better understood.

0-7803-0894-8/93 $3.00

In this paper, we report DC and RF characteristics of simulated SiC diamond MESFET over a wide temperature range from room temperature up to 1000 K. The present-day materials parameters were embedded in the simulation. Temperature dependencies of saturation drift velocity and mobility were also considered at various operating temperatures. Very importantly, incomplete dopant ionization was employed to calculate the activated carriers in the channel. A preliminary comparison is made based on the same device structure for diamond and SiC MESFETs.

Simulation Experiments

The DC and RF performance of n-type 6H-SiC and p-type diamond MESFETs was investigated by a harmonic balance circuit simulator which utilizes the accuracy of the two-dimensional numerical device simulator, PISCES-IIB. The harmonic balance simulator extracts the MESFET gate and drain current and capacitance characteristics which are calculated by PISCES-IIB and then performs large and small signal analysis. The large-signal RF simulator is a NCSU physics based GaAs MESFET model and has demonstrated excellent accuracy in predicting the RF performance of a variety of industrial devices fabricated from GaAs[4].

The exact role of various scattering mechanisms that govern the temperature dependence of mobility is not presently known in both 6H-SiC and diamond. For 3C-SiC, the dependence has been reported to have a form of $T^{-\alpha}$ where α ranges from 1.2 to 1.4[5]. For p-type diamond, the mobility degrades rapidly with temperature. The temperature dependence has been reported to be $T^{-\beta}$ where β is between 2.2 and 2.8[6, 7]. We assumed values of 1.3 and 2.8 for n-type 6H-SiC and p-type diamond, respectively. The dependences of temperature on the drift velocity of the carrier were assumed to vary as $T^{-0.87}$ and $T^{-0.76}$ for SiC and diamond, respectively. Table I lists the major material parameters used in the simulation of SiC and diamond MESFETs at room temperature.

Table I. List of the major material parameters (room temperature) used in the simulation.

Materials Parameters	Diamond (p-type)	6H-SiC (n-type)
Activation Energy (eV)	0.35	0.10
Energy Gap (eV)	5.4	2.86
Low-field Mobility (cm^2/V-s)	1200	340
Saturation Velocity (cm/s)	1.7×10^7	2×10^7
Density of States (Nc) (/cm^3)	1.8×10^{19}	1.04×10^{19}
Density of States (Nv) (/cm^3)	5.0×10^{19}	2.8×10^{19}
Relative Permittivity	5.7	10.0
S.R.H. Life Time (hole) (s)	1×10^{-9}	1×10^{-9}
S.R.H. Life Time (electron) (s)	1×10^{-9}	1×10^{-9}

Several parameters such as intrinsic carrier concentration, density of states, and band gap energy exhibit temperature dependence and are included in the simulation.The effect of temperature on free carrier activation is modeled using Fermi-Dirac statistics for incomplete ionization. The activation energy of the dopant has the most sensitive dependence on temperature. Fortunately, the activation energies for n-type SiC(Nitrogen doped) and p-type diamond(Boron doped) are well established which are 0.10 eV[8] and 0.35 eV[9], respectively. Due to the high thermal conductivity of these two materials, the possible dissipated dc energy was not taken into account in the RF simulation. In typical RF devices fabricated from conventional materials such as Si and GaAs, the dissipated energy causes a temperature rise in the conducting channel and severely degrades the DC and RF performance.

DC and RF Performance of P-Type Diamond MESFET

The present-day materials parameters have been implemented for the DC simulation of a p-type diamond MESFET at several temperatures. A cross sectional sketch of the investigated device structure is shown in Fig.1. The workfunction of the gate metal is 5.1 eV. The contact resistance of ohmic contacts (source and drain) is assumed to be 1 x 10^{-5} Ω-cm^2.

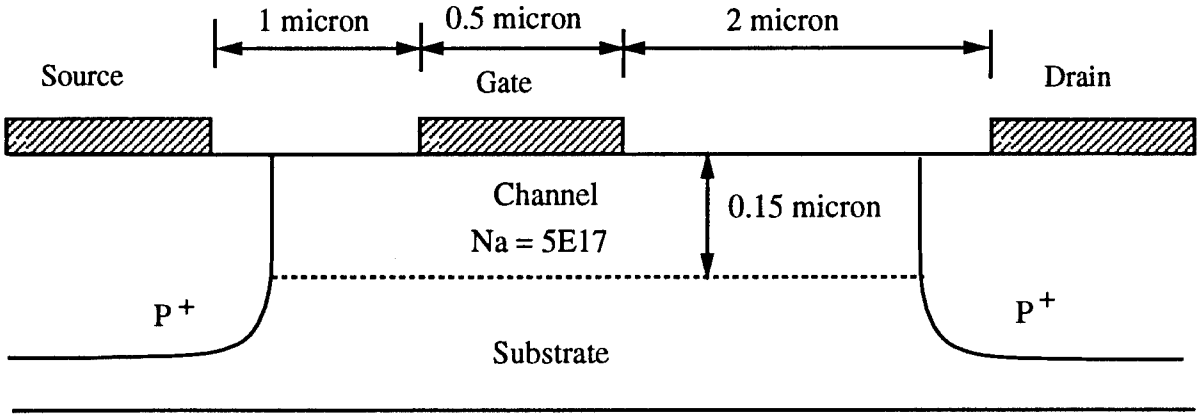


Fig. 1 Schemetic cross sectional sketch of a p-type diamond MESFET. The structure of a n-type SiC MESFET is same. The gate width of the device is 1 mm.

Figs.2 (a) - (b) show the DC I-V characteristics obtained at room temperature and 650 C. The gate voltage ranges from 0 to 20 V. The drain-source current at 923 K is substantially higher than that at room temperature. The magnitude of channel current is an indicator of the power capability of the MESFET. For this reason, the DC performance of the device was simulated at various temperatures. The result is shown in Fig.3(a). The peak drain current increases with temperature almost up to 950 K and starts to drop at this temperature. The influence of temperature on drain current is significant. The peak drain current at 923 K is approximately 25 times higher than that at room temperature. For simplicity, the peak drain current in Fig.3 is defined as the drain current at Vgs = 0 and Vds = 50 V.

The RF performance for different temperatures was performed using a harmonic balance simulator. Figs.4 (a) - (c) compare the RF characteristics of the device at different operating temperatures. The device was characterized at an operating frequency of 5 GHz with Vgs = 10 V and Vds = 40

V. At higher temperatures, there is an improvement in RF output power as well as gain and power-added efficiency (PAE). This result is consistent with the channel current variation which is shown in Fig.3. The PAE at 773 K has approximately 33 % peak at an input power of 22 dBm. The linear gain is about 8.5 dB and the output power is approximately 0.75 W/mm at this temperature.

The improvement of DC and RF performance with temperature in a diamond MESFET can be attributed to a high value of activation energy (0.35 eV) for the ionization of Boron. The high activation energy also explains the significant variation of drain current with temperature as seen in Fig.3(a). It is believed that the magnitude of channel current and thus the RF performance is dominated by the ionization process until 950 K. The current level beyond this point begins to fall due to degraded mobility and drift velocity at high temperature.

DC and RF Performance of n-Type 6H-SiC MESFET

DC and RF characteristics of a 6H-SiC MESFET were investigated for the same device structure as in the diamond MESFET discussed above. The workfunction of gate metal is 5.7 eV. The contact resistance of ohmic contacts is assumed to be 1×10^{-5} Ω-cm^2. This number is a realistic value that can be achieved by highly doping(1×10^{19}/cm^3) under the source and drain contact. Fig.3(b) shows the peak drain current of the SiC MESFET in a temperature range of 300 K to 1000 K. Compared to the diamond MESFET, the DC-IV characteristics of a SiC MESFET are very different. The magnitude of drain current at the entire temperature range is much higher. And the maximum drain current level is at about 373 K, which is much lower. The variation of drain current with temperature is much smaller.

The temperature dependence of IV characteristics in SiC MESFET results from relatively low ionization energy of donor(nitrogen) in SiC, which is about 0.1 eV. This number is much larger compared to 0.35 eV in diamond. Also the smaller temperature dependence of mobility in SiC contributes to less sensitivity of the drain current to temperature after the peak current maximum(373 K). Figs.5(a)-(b) show the DC-IV characteristics of SiC MESFET at 300 K and 923 K. The pinch-off voltage in SiC (about -9 V) is less than that in diamond due to the higher permittivity of SiC.

The higher drain current level in SiC MESFET leads to a better RF performance compared to the diamond MESFET. The RF performance of SiC MESFET was simulated at 300 K, 573 K, 773 K and 923 K. The result is shown in Figs.6(a)-(c). Our investigation suggests that the best RF performance of SiC MESFET is obtained at a temperature range slightly higher than room temperature. At 300 K, the maximum output power of 3.5 W/mm with power-added efficiency of 44 % and linear gain of 16.5 dB was obtained at an operating frequency of 8 GHz with Vgs = -3.5 V and Vds = 40 V. At 923 K, the device produced about 1.8 W/mm with efficiency of 32 % and gain of 12.5 dB under the same bias condition. Based on the specific device structure under present investigation, the RF performance of SiC MESFET is superior to that of diamond MESFET. It must be pointed out, however, that thermal effects and breakdown of devices were not taken into account in the present study. The thermal conductivity of diamond is known to be about 7 times higher than that of SiC. The higher thermal conductivity will lead to a higher breakdown voltage in diamond than in SiC at very high temperatures. The higher breakdown voltage expands the operating area for the dynamic load line in large signal performance [10] and reduces the advantage of SiC over diamond.

Conclusion

The DC and RF performance of n-type 6H-SiC and p-type diamond MESFETs has been simulated using the present-day materials parameters at various temperatures. For SiC MESFET, the RF performance is predicted to be optimal at near room temperature and to be degraded as temperature is increased. At room temperature the simulated SiC MESFET exhibits an output power of 3.5 W/mm for an operating frequency of 8 GHz with 16.5 dB gain and 44 % power-added efficiency. In contrast, the RF performance of a diamond MESFET is improved with temperature from room temperature up to 950 K due to the very high activation energy of dopants for ionization. The current level of the diamond MESFET is much lower than that of the SiC MESFET, particularly at the low temperature region investigated. At 773 K, the diamond MESFET produces about 0.75 W/mm at 5 GHz and 33 % power-added efficiency at 21 dBm input power. From a preliminary analysis based on a simple device structure, it may be concluded that the RF performance of a n-type 6H-SiC MESFET is superior to that of p-type diamond MESFET. Further research that implements several other factors such as thermal effects, breakdown voltage, and optimized device structure is needed for a more objective comparison between the SiC and diamond devices.

ACKNOWLEDGMENTS

The SiC research was funded by the Westinghouse Electronic Co. and the Office of Naval Research (BMDO/IST) under contracts N00014 - 88 - K - 0341. The diamond research was funded by Kobe Steel USA, Inc. and the Office of Naval Research (BMDO/IST) under contracts N00014 - 90 - J - 1604. MWS would like to thank to Dr. T. A. Winslow for valuable discussion

REFERENCES

[1] R. J. Trew, J. B. Yan, and P. M. Mock, "The potential of diamond and SiC electronic devices for microwave and milimeter-wave power application, Proc. of IEEE, vol. 79, no. 5, pp. 598, 1991.

[2] K. Shenai, R. C. Scott, and B. J. Baliga, "Optimum semiconductors for high-power electronics", IEEE Trans. Electron Devices, vol. 36, no. 9, pp. 1811, 1989.

[3] M. Bhatnagar and B. J. Baliga, IEEE Trans. Electron Devices, vol. 40, no. 3, pp. 645, 1993.

[5] M. A. Khatibzadeh and R. J. Trew, "A large-signal analytic model for the GaAs MESFET", IEEE Trans. Microwave Theory and Techniques, vol. 23, no. 6, pp. 545, 1976.

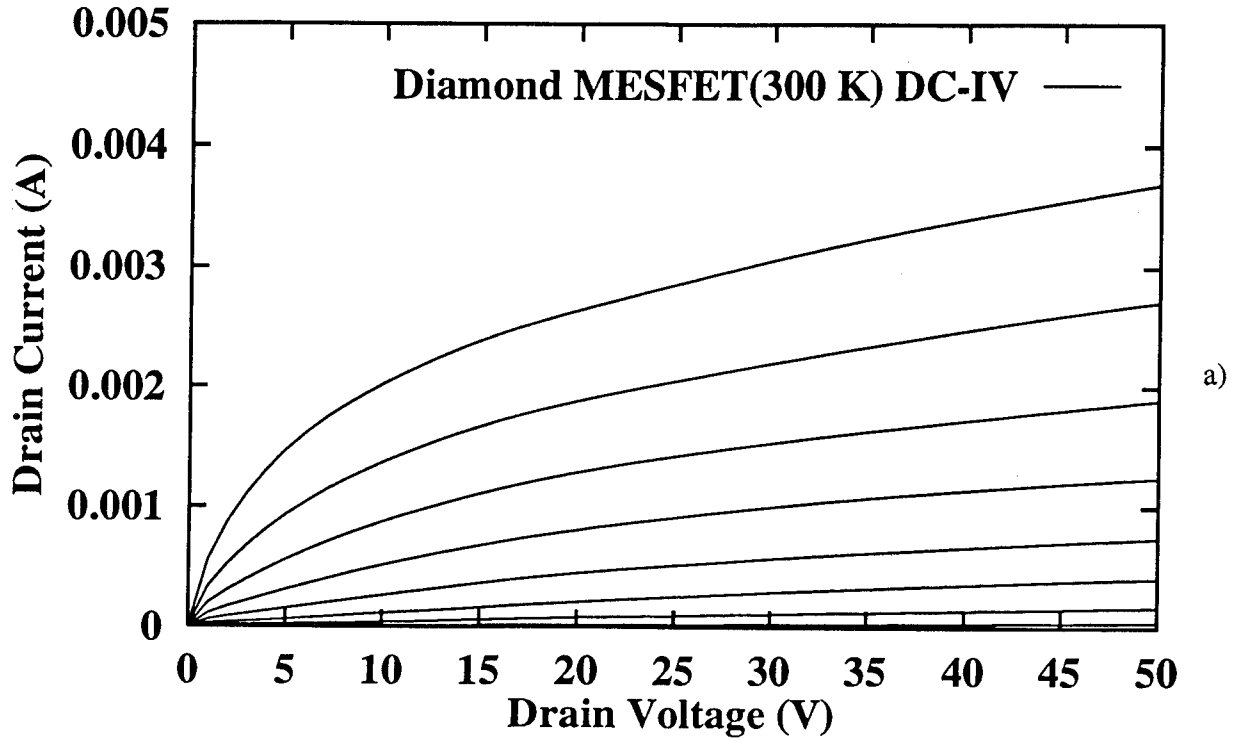
[6] N. Fujimori, H. Nakahata, and T. Imai, "Properties of boron-doped epitaxial diamond films, Jap. J. Appl. Phys. vol. 29, pp. 824, 1990.

[7] D. L. Dreifus, Private communication.

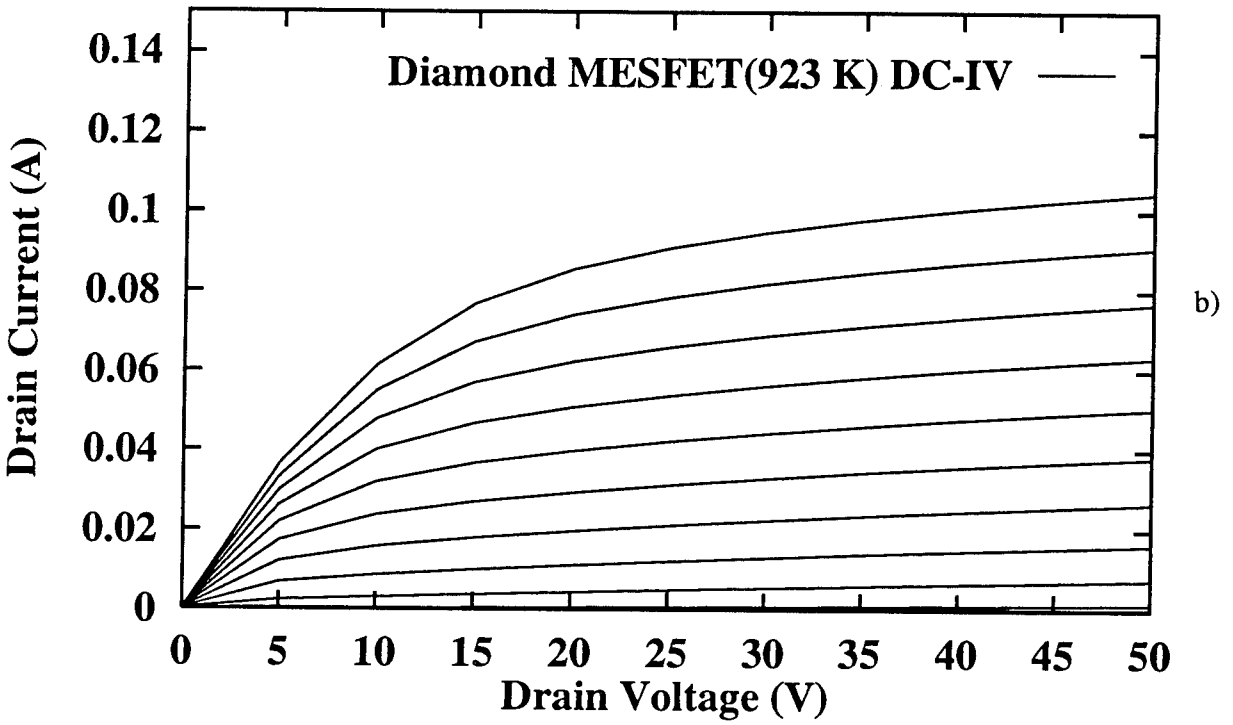
[8] T. P. Chow and R. Tyagi, "Calculation of specific contact resistivity of metals on wide band-gap semiconductors", presented at Tri-service SiC workshop, Washington D. C., 1993 and private communication with T. P. Chow.

[9] J. E. Field, "Properties of diamond", Academic Press, London, 1979.

[10] T. A. Winslow, "A gate breakdown model for GaAs MESFETs", Ph.D. dissertation, North Carolina State University, 1993.

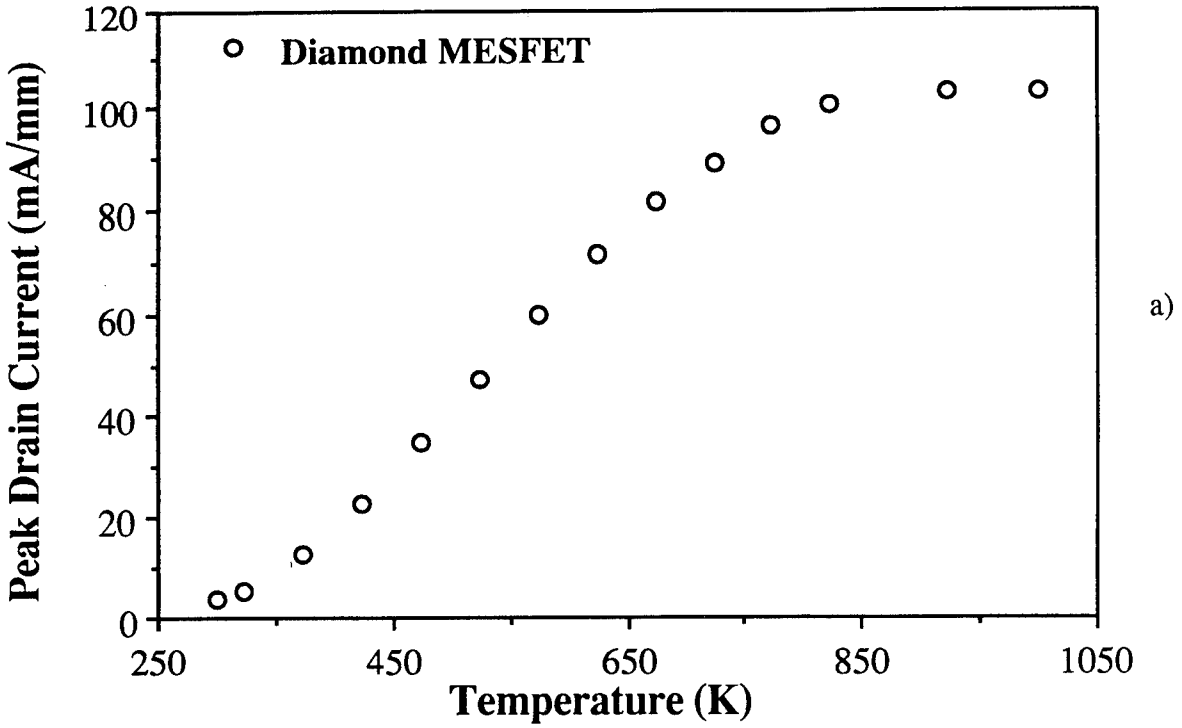


a)

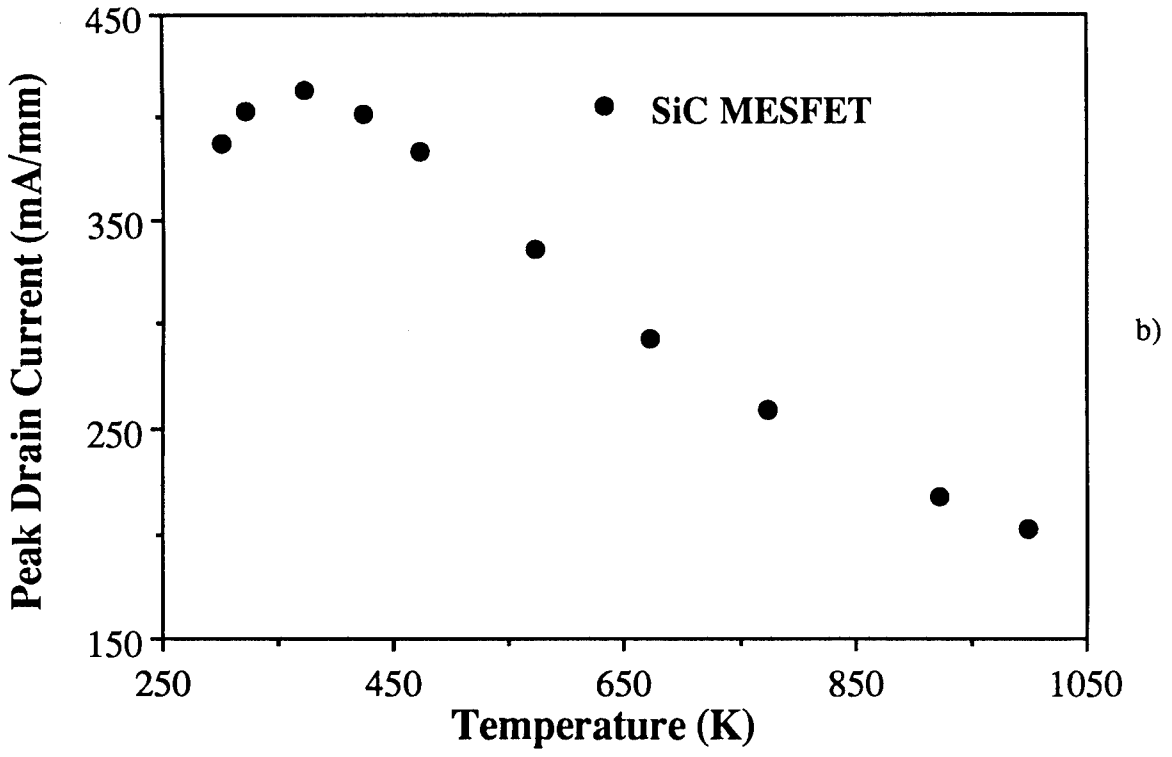


b)

Fig.2 I-V characteristics for a p-type diamond MESFET (W = 1 mm) at a) 300 K and at b) 923 K. The gate voltage ranges from 0 to 20 V in steps of 2 V. Ids is very low at 300 K.



a)



b)

Fig.3 Peak drain current (Ids at Vgs = 0 and Vds = 50 V) versus device operating temperatures for a) p-type diamond MESFET and b) n-type 6H-SiC MESFET.

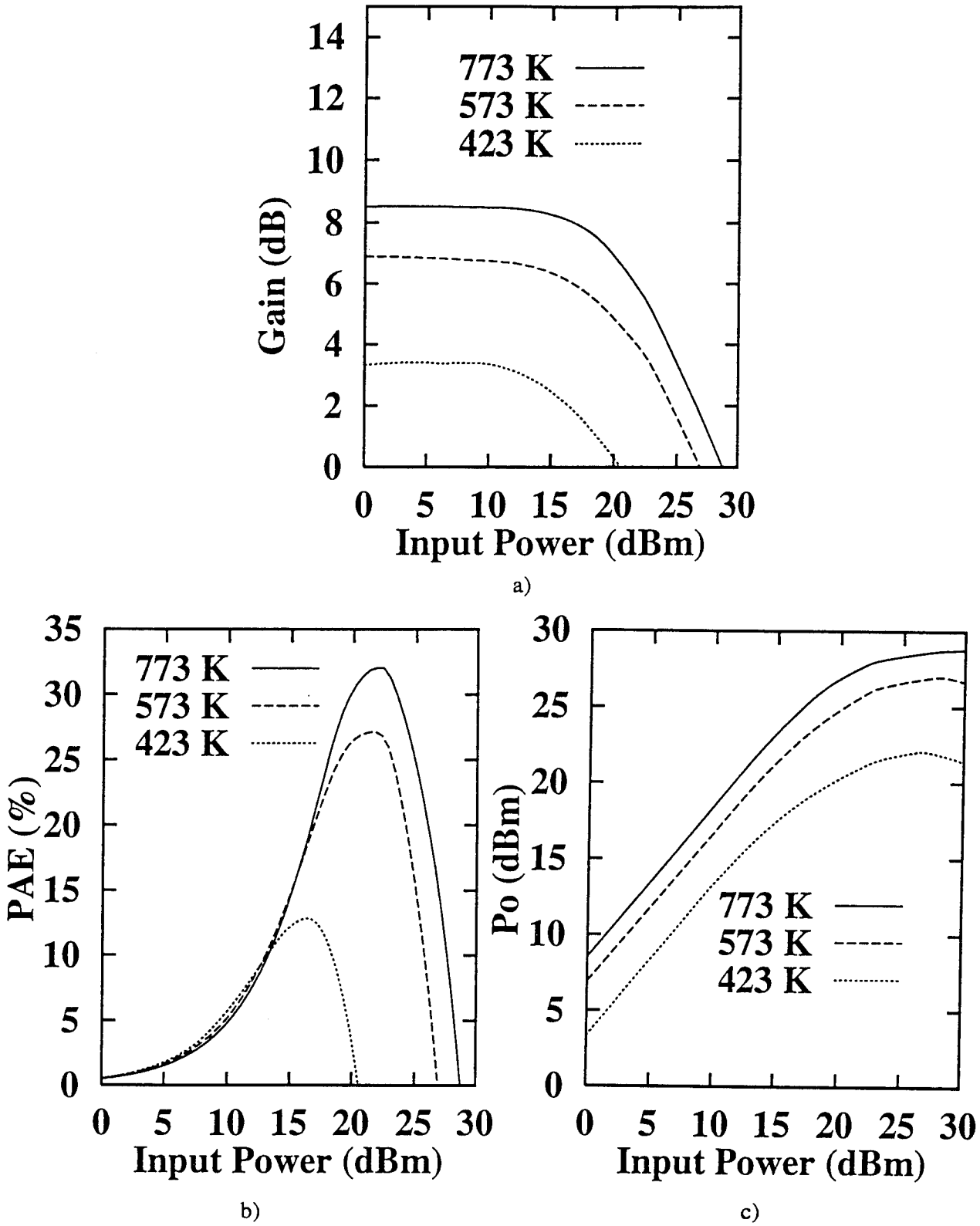
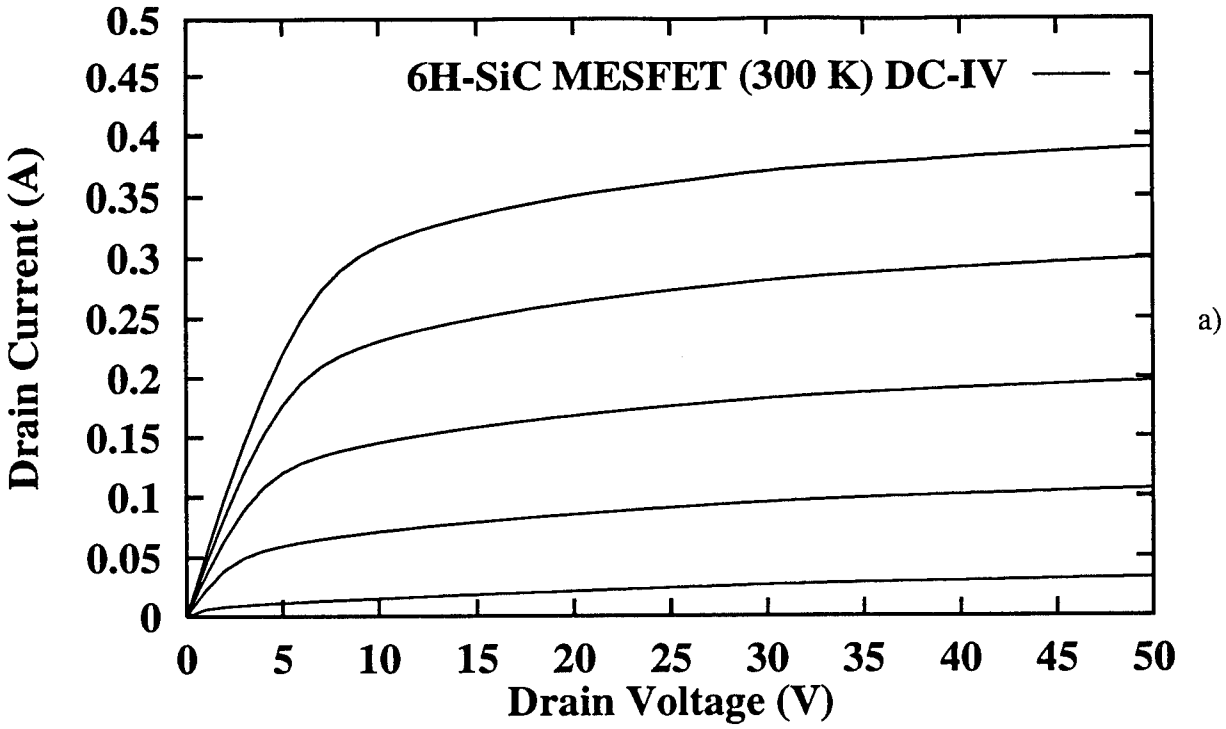
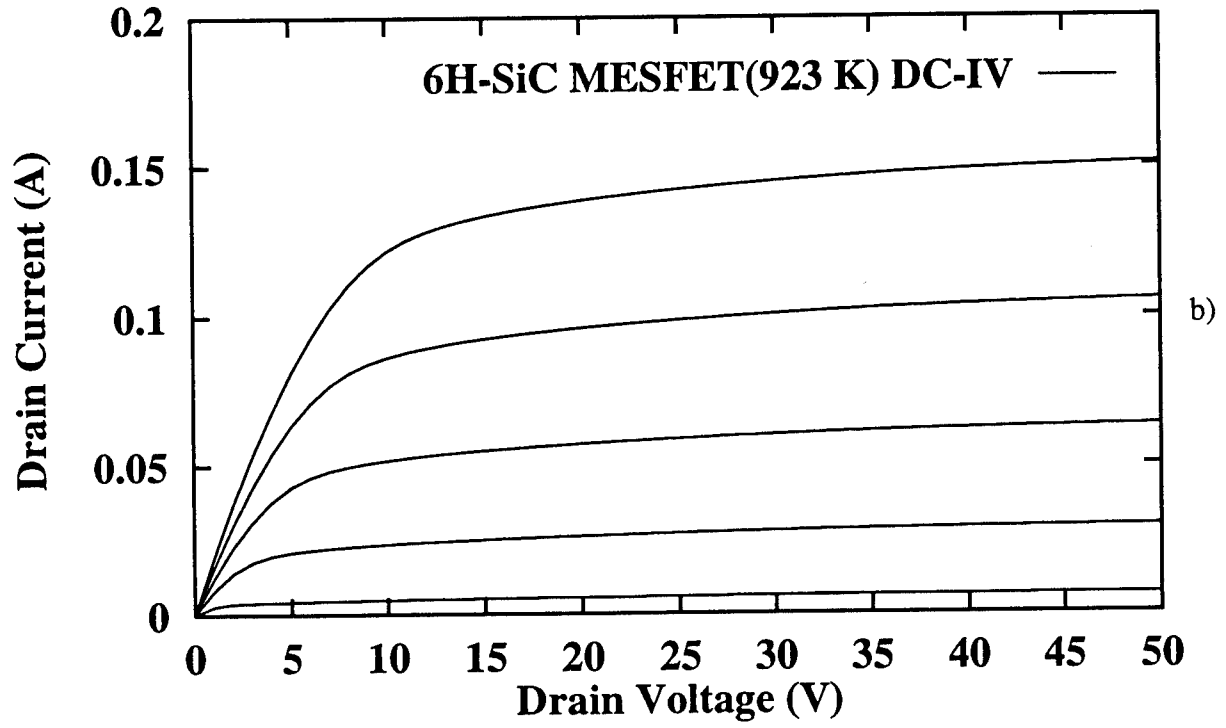


Fig.4 RF performance of a p-type diamond MESFET at different operating temperatures: a) gain versus input power b) power-added efficiency versus input power and c) output power versus input power (F = 5 GHz, W = 1 mm, Class A).



a)



b)

Fig.5 I-V characteristics for a n-type 6H-SiC MESFET (W = 1 mm) at a) 300 K and at b) 923 K. The gate voltage ranges from 0 to -10 V in steps of -2 V.

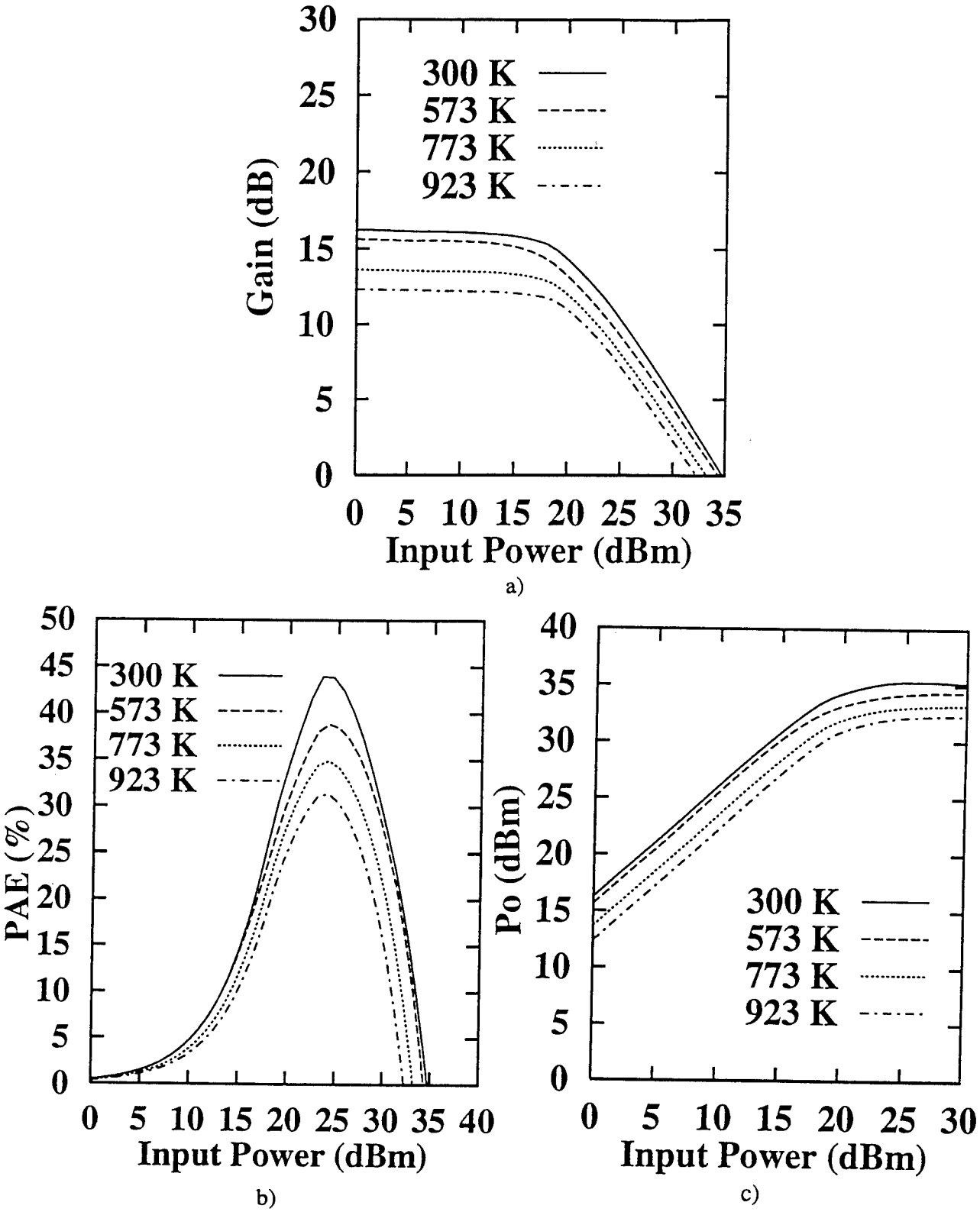


Fig.6 RF performance of a n-type 6H-SiC MESFET at different operating temperatures: a) gain versus input power b) power-added efficiency versus input power and c) output power versus input power (F = 8 GHz, W = 1 mm, Class A).

SELF-CONSISTENT CALCULATIONS OF [111]-ORIENTED GaAs AND InP BASED PSEUDOMORPHIC HEMT's.

J.L.Sánchez-Rojas and E.Muñoz

Departamento de Ingeniería Electrónica. E.T.S.I. Telecomunicación. Ciudad Universitaria S/N. 28040. Madrid. Spain.

ABSTRACT

Band structure calculations, charge distribution and charge control properties in InGaAs-based pseudomorphic modulation doped [111]-FET's are presented. Compressive strain in AlGaAs/InGaAs/GaAs structures, and both tensile and compressive strains in AlInAs/InGaAs/InP devices are used to generate internal electric fields via the piezoelectric effect. Normal and inverted HEMT's with maximum charge in the well and minimized parallel conduction have been designed using this piezoelectric internal field. Improvements in charge concentration (up to 50% more) and its distribution in the well (almost centered) are achieved in this new orientation. Transconductance and gate capacitance versus gate voltage are compared for the [100] and [111] orientations. We conclude that the device performance is improved when the [111] substrate is used.

I. INTRODUCTION

There has recently been much interest in the properties and potential device applications of strained layers, piezoelectrically active in semiconductor heterostructures grown on polar substrates [1]. A number of optical devices have been proposed [1,2], and Snow et al [3] have recently suggested that the use of a strained barrier grown in the [111] direction could be used to achieve inversion without the need for extrinsic doping. The use of InGaAs pseudomorphic channels in high electron mobility transistors (PMHEMT's), has led to considerable improvements in device performance due to the increased ΔE_c, which allows high electron densities even when low Al content barriers are used [4].

Calculations showing enhanced carrier densities and device performance in [111]-oriented AlGaAs/InGaAs/GaAs HEMT structures, where the piezoelectric field is placed within the active region, have been reported [5]. In this paper we present a comparative analysis of [111] and [100] PMHEMT's grown on GaAs or InP substrates. Important device parameters, such as transconductance and gate capacitance, are estimated from the channel charge versus gate voltage behavior.

0-7803-0894-8/93 $3.00

($N_d^+(z)$ and $N_a^-(z)$ for donors and acceptors respectively) and the free carriers (in our case we consider an n-type HEMT):

$$\rho(z) = q\left(N_d^+(z) - N_a^-(z) - n(z)\right) \qquad (4)$$

Assuming a standard depletion approximation in the GaAs substrate, $N_a^-(z)$ is equal to the residual acceptor density, $N_a(z)$. This assumption considers that the Fermi level is close to the conduction band in the rest of the structure and all the acceptor levels are filled. The free electron concentration is obtained adding the populations of all the subbands up to a maximum of ten (usually, no more than three have electrons), with a spatial probability factor given by:

$$n(z) = \sum \frac{m^* KT}{\pi \hbar^2} \ln\left\{1 + \exp\left[(E_f - E_i)/KT\right]\right\} |\psi_i(z)|^2 . \qquad (5)$$

where E_f is the electron quasi-Fermi level and E_i stands for the electron quantum eigenvalues.

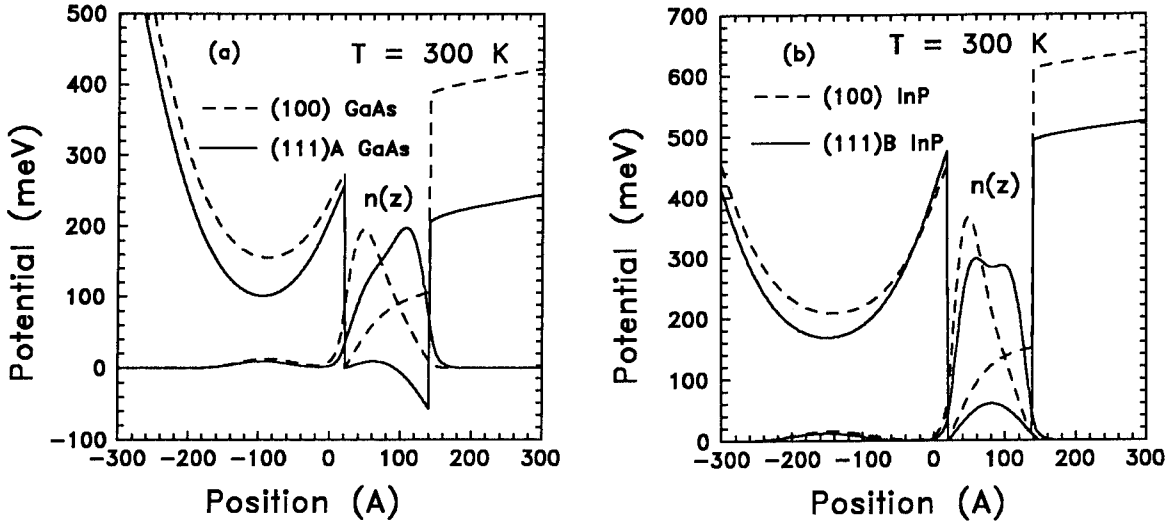


Fig. 1. Conduction band and charge distribution for [100] (dashed lines) and [111] (solid lines) devices on GaAs (a) and InP (b).

It is clearly seen, from equations (3) to (5), that the wavefunctions are needed to solve the Poisson equation in $V_h(z)$, which is also included in the Hamiltonian. Therefore, a self-consistent solution has to be obtained. In order to solve the above system of equations two boundary conditions have to be considered:

II. SELF-CONSISTENT MODEL

In this paper new ideas for InGaAs PMHEMT's are presented. [111]-oriented substrates of GaAs or InP are used, allowing for compressive or tensile strain in the InGaAs layers. The model presented here follows the envelope-wave-function formalism in the effective mass approximation [6]. This method has the advantage that it is easy to take into account the effect of any potential profile superimposed on the lattice potential, the latter being described with parameters such as the electron and hole effective masses and energy gaps. The Hamiltonian for electrons is given by:

$$H = \left\{ \frac{-\hbar^2}{2\, m_0} \frac{d}{dz} \left(\frac{1}{m^*(z)} \frac{d}{dz} \right) + V(z) \right\} \qquad (1)$$

and the subsequent eigenvalue problem:

$$H\, \psi = E\, \psi \ , \qquad (2)$$

where m_o is the free electron mass and $m^*(z)$ the electron effective mass in the growth direction, is solved for a typical multilayer system containing an undoped InGaAs QW layer grown on a semi-insulating layer (substrate or buffer), and cladded with a selectively doped barrier. The electron effective mass can change significantly from the InGaAs QW to the AlGaAs or AlInAs barrier [7], and this fact is included in eq. (1) by means of a soft step-like function, suchas that assumed by Stern et al [8]. The abruptness of the effective mass change is increased until no changes in the simulation results (wavefunctions and eigenenergies) are observed.

According to the modulation doping concept, a spatially distributed charge density, corresponding to electrons transferred into the well and to ionized donors in the barrier, is obtained. As a result, each electron in the system "sees" the electrostatic influence of all the others. This picture poses a very complicated, many-body problem, which is simplified using a one-electron approximation corrected for the exchange-correlation interaction. We use the simple analytic parameterization proposed by Hedin and Lindqvist as in ref. [8].

The resulting potential, V(z), that is introduced in the Schrödinger equation, contains three terms: $\Delta E_c(z)$, $V_h(z)$, and $V_i(z)$, corresponding to the discontinuities in the conduction-band, the electrostatic (Hartree) potential and the correlation term, respectively. The term $V_h(z)$ is derived from the solution of the Poisson equation:

$$\nabla\left(\varepsilon(z)\nabla V_h(z)\right) = q\rho(z) \qquad (3)$$

where the total charge density, $\rho(z)$, includes the ionized impurities

i) The substrate is large enough to contain the space charge region due to the residual ionized impurities.

ii) The potential $E_c - E_f$ at the top interface is given by the Schottky barrier.

The Schrödinger equation is now solved for an initial V(z) function using a "finite element" method with a basis of cubic β-splines [9]. Then the Fermi level is obtained iteratively until the charge neutrality condition is fulfilled, i.e.:

$$Q_s = \int_o^\infty \rho(z)\, dz \qquad (7)$$

Q_s being the surface charge density at the metal gate which is related to the electric field at this point. Once the Fermi level is known, the integration of Poisson's equation generates the electrostatic potential that is introduced again into the Hamiltonian. For convergence purposes, the scheme proposed in ref. [10] is followed.

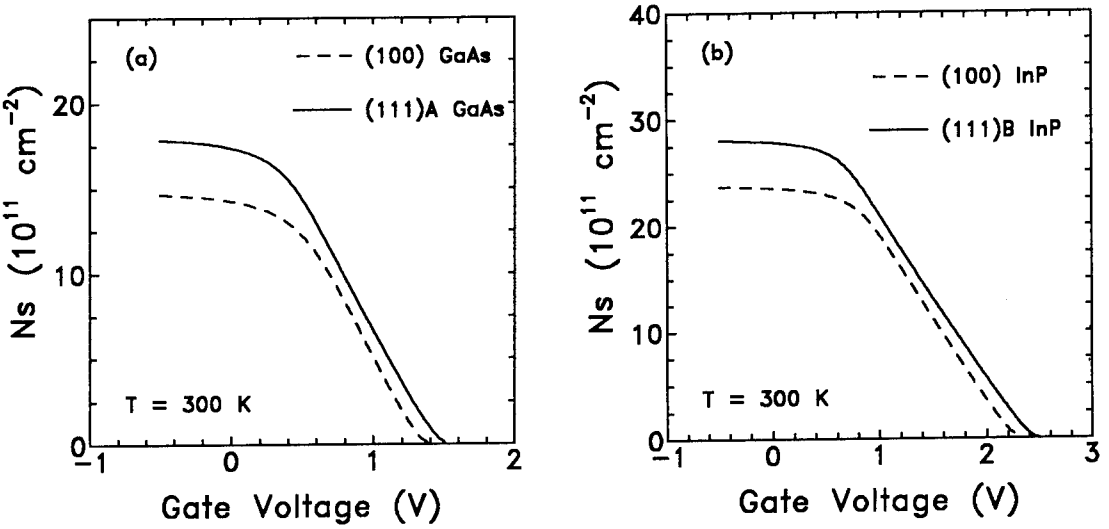


Fig. 2: Charge in the well versus reverse gate voltage for GaAs (a) and InP (b) based devices.

III. PIEZOELECTRIC EFFECTS

It is well known that strain in zincblend lattices generates polarization fields [11]. We assume that the lattice constant mismatch is accommodated by internal biaxial strain in the QW, rather than by the generation of misfit dislocations (we are below the critical thickness). Common III-V semiconductors have a negative piezoconstant. Thus, the sign of the polarization vector in the active layer, if present, depends on the sign of the strain (difference of lattice constants between layer and

substrate), and on the kind of face (terminated with the anion, B, or cation, A) on which the device is grown.

Polarization in the strained well is given by

$$P = \sqrt{3}\, e_{14}\, \epsilon_{xy} \tag{8}$$

where e_{14} is the piezoelectric constant in InGaAs. ϵ_{xy} is the symmetrized off-diagonal component of the strain tensor, and is given by

$$\epsilon_{xy} = - \frac{(C_{11} + 2C_{12})}{2C_{44}}\, \epsilon_{xx} \tag{9}$$

with

$$\epsilon_{xx} = \frac{4C_{44}}{C_{11} + 2C_{12} + 4C_{44}} \left[\frac{a_1}{a_2} - 1 \right] \tag{10}$$

where a_1 and a_2 are the lattice constants of the substrate and well, respectively, and the C_{ij} are the elastic constants. All the constants for InGaAs are linearly interpolated from the values for the binary constituents. In our calculations, the dielectric constant $\varepsilon = \varepsilon_o \varepsilon'$ is assumed the same for the well and barriers, giving a discontinuity in the electric field of P/ε. This boundary condition has to be included in the self-consistent loop described above.

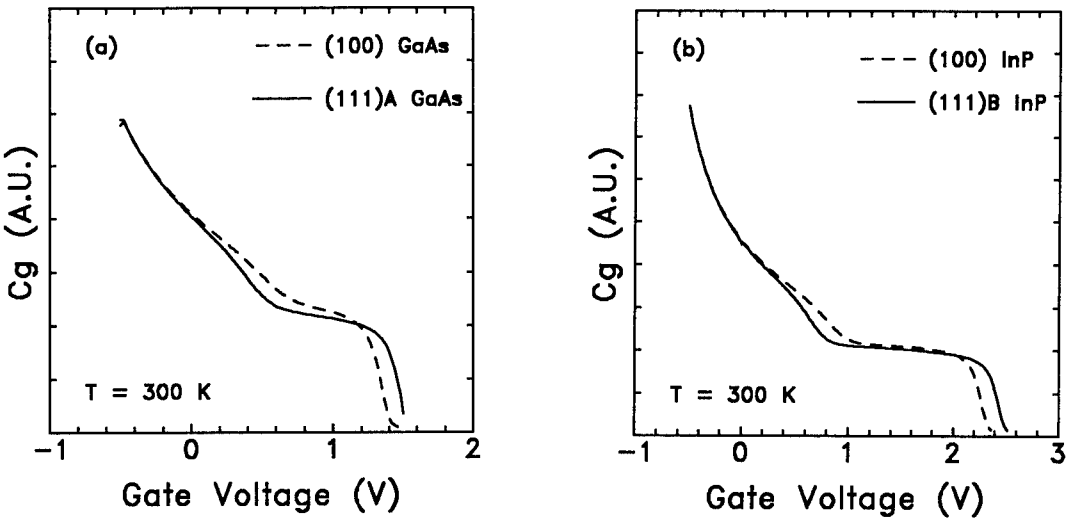


Fig. 3. Gate capacitance versus reverse gate voltage.

For the structures analyzed in this work, the following practical rule is obtained: compressive strain and (111)A orientation give rise to an electric field pushing the electrons towards the substrate. The same is verified when tensile strain and (111)B faces are considered. An electric field of opposite direction is obtained when one of the combinations compressive/(111)B or tensile/(111)A is present.

The above rules express also how to achieve the compensation of the electrostatic potential which pushes the electrons towards the doped-barrier interface. Compressive strains should be convenient to fabricate normal HEMT's grown on (111)A, while inverted devices could be obtained on (111)B orientations. On the other hand, tensile strain (only possible with the InP system) should be used in normal (inverted) HEMT's grown on (111)B ((111)A) substrates.

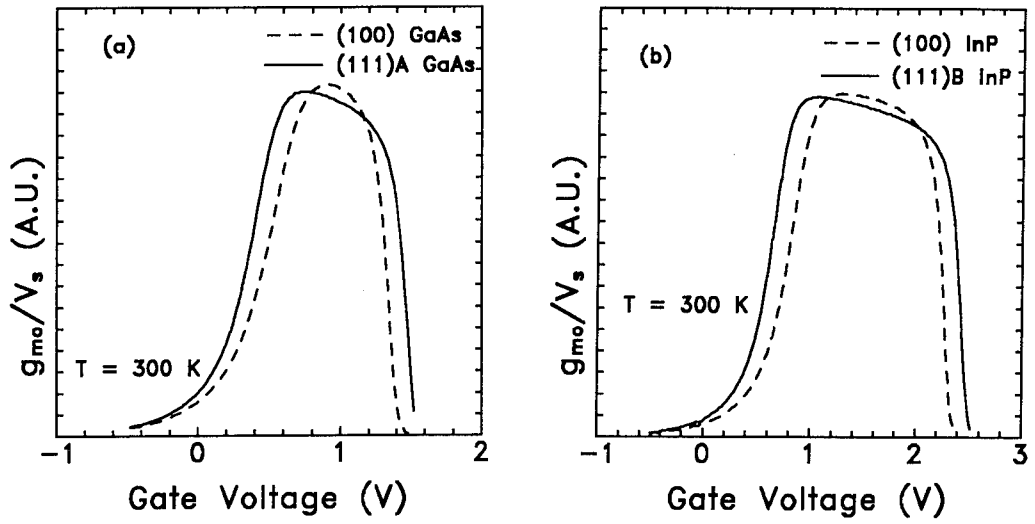


Fig. 4. Normalized intrinsic transconductance as a function of the reverse gate voltage.

IV. RESULTS AND DISCUSSION

Two structures reflecting the above ideas have been studied in detail. Figure 1a shows the conduction band profile and charge distribution, n(z), of a normal $Al_{22}GaAs/In_{12}GaAs/GaAs$ HEMT. The calculations were performed for room temperature operation in a structure composed of a 350Å AlGaAs doped barrier (2.5×10^{18} cm^{-3}), a spacer layer 20Å thick, an InGaAs well 120Å thick, and the GaAs buffer/substrate. An acceptor doping level of 1×10^{15} cm^{-3} is taken for the spacer, the QW and the GaAs barrier. We have assumed a Schottky barrier height of 1eV at the metal semiconductor interface, corresponding to Al on AlGaAs. The profiles for both, the [111]A (solid lines) and the [100] (dashed lines) orientations are shown. From this comparison, we observe that the piezoelectric field is slightly stronger than that corresponding to the quasi-flatband condition. In spite of that, the average distribution of electrons is improved as compared to the [100] case [5]. In the normal [111]A HEMT, the electrons are closer to the center of the well for a wide range of well widths, whereas in the [100] structure the wavefunction is peaked much closer to the AlGaAs spacer interface.

The self-consistent calculations for devices on InP substrates give the profiles shown in fig. 1b. An $Al_{48}InAs/In_{37}GaAs/Al_{48}InAs$ structure has

been studied. The widths of the different layers and the doping levels are assumed to be the same as for the GaAs-based structures. The piezoelectric constant is lower for the InGaAs QW in the InP-based structure, thus giving an exact cancelation of the Hartree potential on it, while its strain is similar to that for the GaAs-based device.

Large channel carrier densities are required in power devices. A charge control analysis of the above piezoelectric HEMT's has been performed and results have been compared with those grown on conventional [100] substrates. Figures 2a and 2b show the sheet electron concentration (N_s) at 300 K for structures grown on GaAs and InP, respectively. Solid and broken lines denote [111] and [100] orientations, respectively. An enhanced carrier density for the [111] case is deduced. In fact, as it has been previously reported for the GaAs case, a 50% improvement in charge concentration is predicted as compared to [100] structures [5].

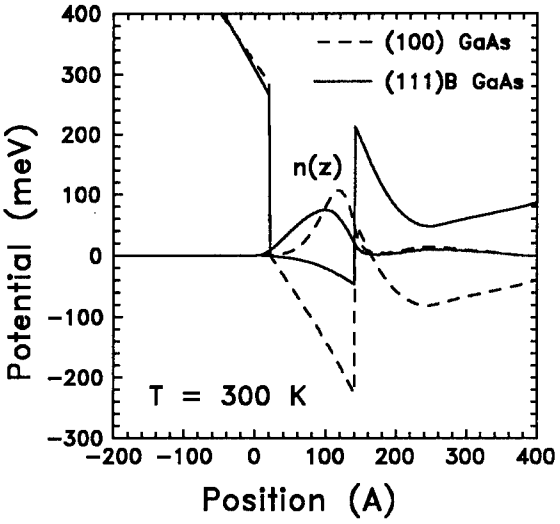


Fig. 5. Potential profile and charge distribution.

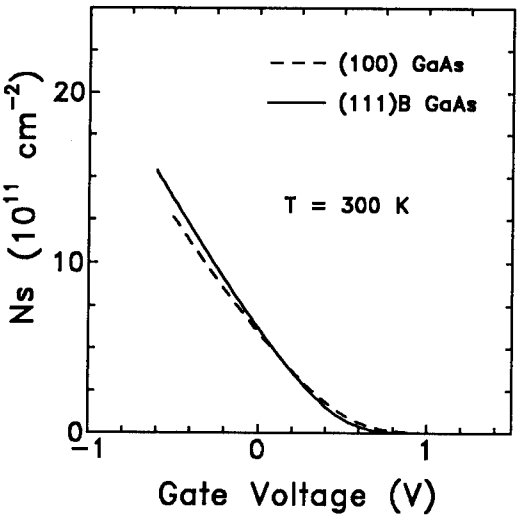


Fig. 6. Charge in the well vs. reverse gate voltage.

The gate capacitance (C_g) is calculated by differentiating the charge in the gate with respect to the gate applied voltage, and results are shown in figures 3a and 3b. Near the threshold voltage, donors in the doped barrier are fully ionized, being electrons far from the gate, resulting in a small C_g. An increase in the gate voltage (V_g) gives rise to confined electrons that do not shift noticeably. Thus, C_g remains almost constant. Finally, the total capacitance increases rapidly when the gate is forward biased, as donor neutralization and parallel conduction become marked. By comparing the [111] and the [100] orientations no significant differences are found, except that for the [111]-oriented devices C_g is almost constant for a wider range of gate voltages.

We have also estimated the intrinsic transconductance under the short channel approximation as

$$g_{mo} = C_w v_s \qquad (11)$$

where v_s is the saturation velocity for the electrons, and C_w is the capacitance due to electrons in the channel. This is calculated as a function of V_g for both the [100] and the [111] cases (figures 4a and 4b in GaAs and InP, respectively). First, the maximum value of C_w is almost the same for both structures, but it is slightly larger for the [100]-oriented. We assume the same v_s for [100] and [111] orientations, thus leading us to identify C_w and g_{mo}. A more important point is that g_{mo} takes the highest values through a wider range of voltages in the [111] structure than in the conventional [100] orientation.

Finally, self consistent calculations have been performed for inverted HEMT's grown on [111]B and on [100] GaAs substrates. Conduction band profiles and charge distributions are shown in figure 5. Once again the piezoelectric field allows us to have the electrons at the center of the QW, while a very symmilar number of carriers is simultaneously obtained, as it is shown in figure 6.

V. CONCLUSIONS

We have compared normal and inverted HEMT's grown on [100] and [111] substrates. Both GaAs and InP-based structures have been considered. The effects of the strain-induced piezoelectric field have been evaluated by using a unidimensional self-consistent model. Charge control, gate capacitance and intrinsic transconductance have been compared for the several structures. It is concluded that [111] devices present improved characteristics for HEMT's applications.

ACKNOWLEDGMENTS

The authors would like to thank J. Sánchez-Dehesa and P. Houston for fruitful discussions and suggestions. We would also like to thank A. Sacedón, F. Calle and E. Calleja for encouraging this work. Financial support from the spanish CICYT proyect TIC 93-0025 is also acknowledged.

REFERENCES

[1] D. L. Smith and C. Mailhiot, "Theory of semiconductor superlattice electronic structure", *Rev. Mod. Phys.*, No. 62, pp. 173-234, 1990.

[2] K. W. Goossen, E. A. Caridi, T. Y. Chang, J. B. Stark, D. A. B. Miller, and R. A. Morgan, " Observation of room-temperature blue shift and bistability in a strained InGaAs-GaAs <111> self-electro-optic effect device", *Appl. Phys. Lett.*, vol. 56, pp. 715-717, 1990.

[3] E. S. Snow, B. V. Shanabrook, and D. Gammon, "Strain-induced two-dimensional electron gas in [111] growth-axis strained-layer

structures", *Appl. Phys. Lett.*, vol. 56, pp. 758-760, 1990.

[4] H. Morkoç and H. Unlu, "Factors affecting the performance of AlGaAs/GaAs and AlGaAs/InGaAs Modulation-doped FET: Microwave and Digital Applications", *Semicond. and Semimetals*, vol. 24 pp. , 1987

[5] J. Sánchez-Dehesa, J.L. Sánchez-Rojas, C. López and R. J. Nicholas, "Enhanced carrier densities and device performance in piezoelectric pseudomorphic HEMT structures", *Appl. Phys. Lett*, vol. 61, pp. 1072-1074, 1992.

[6] G. Bastard, "Wave mechanics applied to semiconductors heterostructures", *Edition de Physique, Les Ulis, France*, 1988.

[7] Landölt-Bornstein, "Numerical Data and Functional Relationships in Science and Technology", *Vol. 17 Springer Verlag*, 1990.

[8] F. Stern and S. Das Sarma, "Electron energy levels in GaAs-AlGaAs heterojunctions", *Phys. Rev. B*, vol. 30, pp. 840-845. 1984.

[9] C. De Boor, "A Practical Guide to Splines", *Springer Verlag*, 1978

[10] F. Stern, "Iteration methods for calculating self-consistent fields in semiconductor inversion layers", *Journal of Computational Physics*, vol. 6, pp. 56-67, 1970.

[11] D. L. Smith, "Strain-generated electric fields in [111] growth axis strained-layer superlattices", *Solid state commun.*, vol. 57, pp. 919-921, 1986.

A Non-Quasi-Static Modular Model for HBTs *

Arthur S. Morris, Robert J. Trew,
C. T. Kelley†, and Gerard J. Hayes

Department of Electrical and Computer Engineering
†Department of Mathematics
North Carolina State University
Raleigh, N.C. 27695-7911
(919)-737-2336

Abstract

Heterojunction bipolar transistors(HBTs) show promise as a high speed and high power density device for many circuit applications. However, the quasi-static models found in standard circuit simulation tools can not treat fast transients in HBTs properly. This leads to inaccurate simulations at high frequency and of strongly non-linear operation. To properly account for the charge in transit through the device, non-quasi-static(NQS) models must be used.

This work presents a model for the bipolar transistor formed from regional modules. Each module is a NQS solution to a specific region of the transistor and uses material and geometry inputs. These modules are solved for physical consistency during non-linear circuit simulation. The modularization allows appropriate approximations for each region to yield analytic solutions.

The input parameters for the model reflect the physical structure of the device as much as possible to provide intuitive results and verifiability. This allows direct device optimization since all parameters are either uncorrelated or their correlations can be derived from process parameters. Thus the device can be optimized in its circuit environment.

The model provides for many effects which previously required numerical simulation for accurate results. These include forward and reverse Early, Webster/Rittner, and Kirk/quasi-saturation effects. By following the modular modeling scheme, these effects are simply the result of varying boundary conditions on each of the regional solutions. The modular model provides much of the physical insight of numerical models but with computational requirements on the same order as conventional circuit models.

1 Introduction

The heterojunction bipolar transistor has proven itself in many high-speed applications. Circuit design with HBTs requires accurate models over the entire useful frequency range of the devices. Standard bipolar models include the Ebers-Moll coupled diode model[1], the Gummel-Poon integral charge control model[2], and their extensions[3, 4]. These models are known as *quasi-static* since the internal device solutions are obtained under static conditions. Time dependence is then added through charging elements such as capacitances. For transistors where the response is limited by RC time constants, either in the embedding circuit or in

*This work was supported in part by ONR contract N00014-89-C-0181, NSF contract DMS-9024622, and by Texas Instruments.

the transistor itself, the quasi-static models are sufficient. If the transistor and circuit have very low parasitic capacitances and resistances, then the actual propagation of charge through the device must be treated. This is known as *non-quasi-static* modeling. NQS modeling is especially important when linearity is of issue. Inaccuracies in signal-mixing products well out of the operating frequency range will affect the products in-band and may interfere with the convergence of the simulator.

NQS modeling of the quasi-neutral regions in the base and emitter of bipolar transistors has been performed previously[5, 6, 7, 8, 9]. Their effort was concentrated on deriving lumped approximations to the full NQS solution. In our model, we make no such approximations.

In an optimized HBT the emitter capacitance, emitter minority storage, and base transit time are greatly reduced over that for a comparable homojunction transistor. Thus the transit time through the depleted collector becomes a significant fraction of the total device delay. Choosing the collector epi-layer thickness involves considering f_t, f_{max}, and collector avalanche breakdown voltage. If breakdown voltage is not the limiting factor, this leads to a tradeoff between collector transit delay and collector capacitance. Usually, the choice is made so that f_{max} is approximately twice f_t for microwave amplifiers. This requirement leads to wider collector regions and thus to transit delays dominated by the collector. If breakdown is the limiting element, then the collector will need to be even thicker leading to a larger dominance by the collector delay.

During a voltage transient, the collector space-charge-region thickness and thus the collector "delay" are modulated. An accurate large-signal HBT model should include this collector modulation. Quasi-static modeling including drifting charge into the nonlinear collector-base capacitance has been used for intermodulation distortion calculations[12]. However, the transient nature of the drifting charge and the modulation of the collector depletion region requires a NQS analysis. A NQS model for space-charge-regions such as the collector has been developed[10].

Section 2 outlines the modules in the model and details the simulation strategy. Section 3 contains the model for the space-charge regions including drift charge and boundary fields. Section 4 presents the NQS generalization of Ramo's theorem for moving boundaries. Section 5 presents the NQS quasi-neutral region model. These models allow any combination of charge injection and boundary motion.

The analysis is applied to a typical HBT structure in section 6 and DC simulations performed.

2 Modular Model

The bipolar transistor is divided into five natural regions. These are the emitter, base, and collector quasi-neutral regions which are separated by the emitter and collector space-charge regions. At each boundary between any two regions the potential, electric field, current, and carrier concentrations must be continuous (only graded junctions are treated). This is diagrammed in figure 1. A small subset of these variables is chosen that completely determines the state of the transistor and allows the direct analytic calculation of all regions. These are the four boundary positions between quasi-neutral and space-charge regions (X_e, X_{b1}, X_{b2}, X_c) and the minority densities at both ends of the quasi-neutral base (N_{b1}, N_{b2}).

One-dimensional analysis is used to simplify the calculations. This is a very good assumption for HBTs due to their low base sheet resistance. A sixth module is used to account for the nonlinear base resistance. This is considered to be purely one-dimensional orthogonal to the rest of the transistor. It is calculated from the instantaneous quasi-neutral base width and majority carrier concentration.

These six modules form the modular model of the bipolar transistor. A quasi-Newton nonlinear equation solver using GMRES[14] is used

to obtain solutions when the model is embedded in a circuit which includes sources. The solver solves the usual Kirchoff circuit constraints simultaneously with the field and charge density continuity conditions.

Time-variation is handled differently in each module depending on its non-linear behavior. The space-charge regions (strongly amplitude non-linear) are described in the time domain while the quasi-neutral regions and the collector/emitter current generator(modulated delays) are described in the frequency domain. Modern commercial simulators are able to perform both steady state and transient solutions with both device descriptions through the use of FFTs and dynamic convolution.

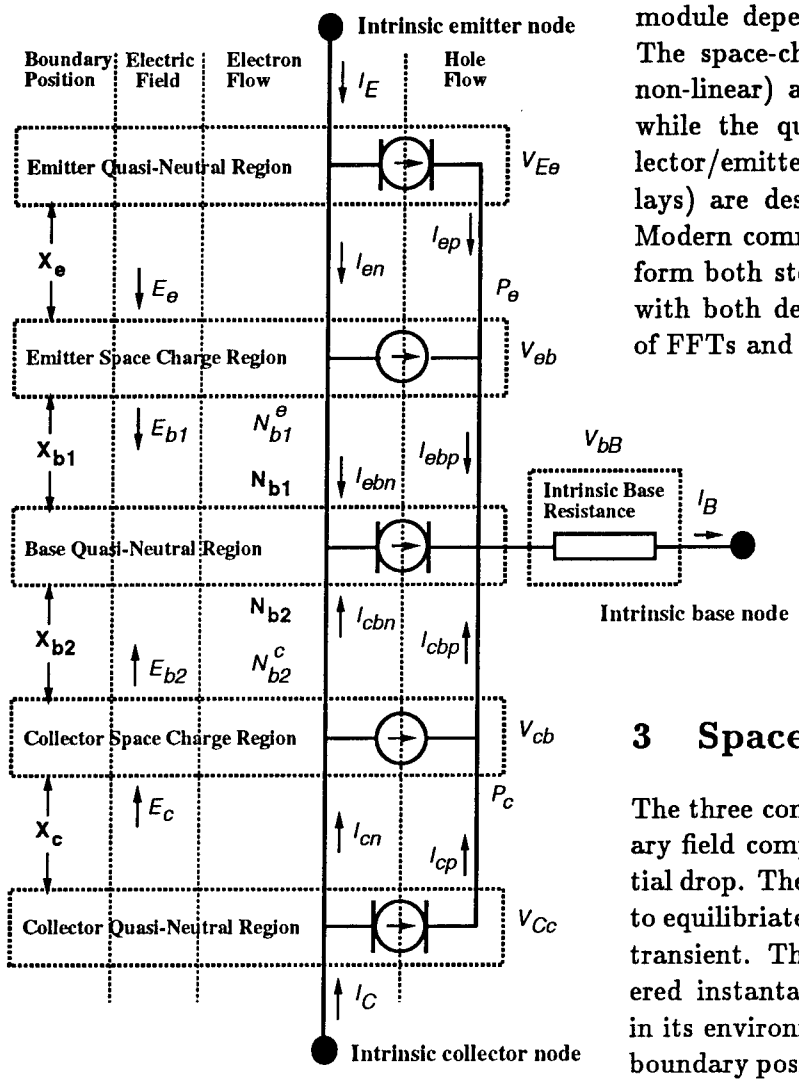


Figure 1: Block Diagram of Modular Model.

3 Space-Charge Region

The three components of this model are boundary field computation, Gauss' Law, and potential drop. The space-charge regions are assumed to equilibriate much faster than any operational transient. This allows the model to be considered instantaneous in its response to changes in its environment, i.e. given the instantaneous boundary positions and free charge distribution, the electric field and potential are determined.

The boundary fields are computed assuming significant voltage drop across the region. This yields boundary fields given by

$$\mathcal{E} = \frac{kT}{qL_D} \tag{1}$$

where L_D is the local Debye length computed from the local properties of the quasi-neutral semiconductor. One requirement is that the

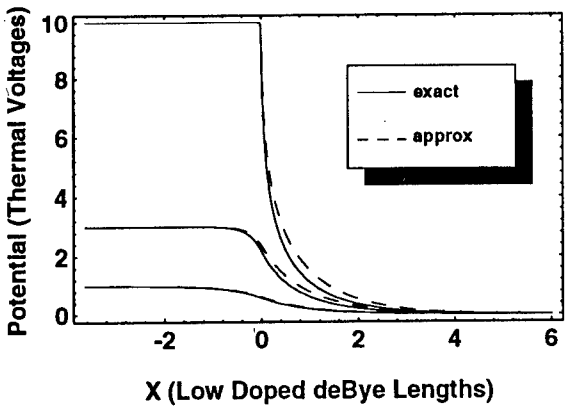


Figure 2: Exact and Approximate Potentials.

majority carrier distribution be known in all regions that may be neutral including those with steep doping gradients such as the subcollector-collector junction or for the base-collector junction at high current densities. Exact solutions for this case require solving the extremely non-linear Poisson-Boltzmann equation[15]. We have obtained an approximate solution for the case of step doping profiles in high-low junctions. For the higher carrier concentration side(assume electrons and $x < x_j$), we obtain the simple Debye length scaling

$$\phi = \phi_0 - A \exp\left(\frac{x - x_j}{L_D^+}\right). \qquad (2)$$

For the lower carrier side($x > x_j$), the potential varies as

$$\phi = \phi_1 + \frac{kT}{q} \log(\coth(\left(\frac{x - x_j + B}{2L_D^-}\right)^2)). \qquad (3)$$

The constants A and B are determined prior to simulation from the structure to yield potential and electric field continuity. A normalized comparison of these expressions with an exact numerical calculation is shown in figure 2. Note that the approximation slightly underestimates the peak electric field. This is necessary to ensure that non-physical results where $\nabla \cdot \mathcal{D} > \rho$ do not occur on the boundaries.

The boundary fields are continued into the space-charge region using Gauss' law and similarly for potential using the integral form of Poisson's equation.

$$\epsilon(x)\mathcal{E}(x) = \epsilon(x_b)\mathcal{E}(x_b) + \int_{x_b}^{x} \rho(x' - x_b)dx' \qquad (4)$$

$$\phi(x) = \phi(x_b) - \epsilon(x_b)\mathcal{E}(x_b)\int_{x_b}^{x} \frac{1}{\epsilon(x')}dx' - \int_{x_b}^{x} \frac{1}{\epsilon(x')} \int_{x_b}^{x'} \rho(x'' - x_b)dx''dx' \qquad (5)$$

These integrals are computed analytically for various doping profiles and assuming a constant drift velocity in the collector.

Thus given the boundary positions and minority current waveforms, we compute the voltages V_{eb} and V_{cb} and the electric field discontinuities. The voltage and minority current are used to compute the boundary minority densities through Fletcher boundary conditions[13] together with the scattering limited velocity charge density. This yields a collector boundary with non-zero minority density in forward active operation. The discontinuities in minority density and electric field are eliminated during the circuit solve to yield a consistent solution.

4 SCR Current Generator

The derivation of the NQS collector current generator model[10] begins by deriving the current induced between two moving conducting plates by a moving charge sheet between them. This solution is then integrated to provide the solution for a continuous charge distribution in a collector.

The collector current generator can be viewed as discharging the base charge storage or base diffusion capacitance. Thus the majority charge, which was neutralizing the minority transit charge while it was in the base, is discharged as the minority charge is swept across the depletion region.

We derived the solution for continuous charge between the plates by assuming an arbitrary

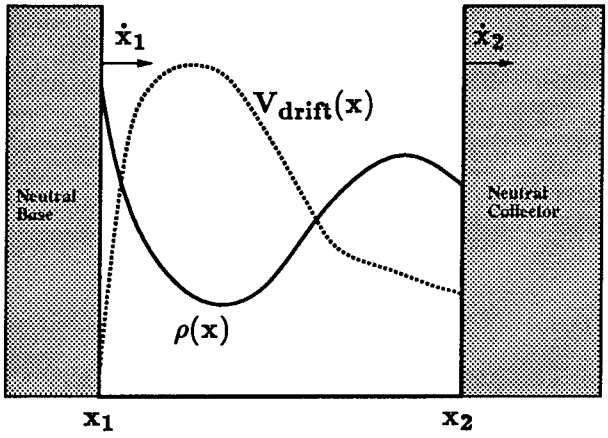


Figure 3: Geometry of current generator problem.

free charge distribution $\rho(x,T)$ and an arbitrary velocity distribution $V(x,T)$. This geometry is shown in Figure 3. The current generator for any time T is of the form

$$\begin{aligned} I_{gen} &= \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} V(x)\,\rho(x)\,dx \\ &+ \frac{x_1\dot{x}_2 - \dot{x}_1 x_2}{\{x_2 - x_1\}^2}\int_{x_1}^{x_2}\rho(x)\,dx \\ &- \frac{\dot{x}_2 - \dot{x}_1}{\{x_2 - x_1\}^2}\int_{x_1}^{x_2} x\,\rho(x)\,dx. \end{aligned} \quad (6)$$

The first term is Ramo's theorem[11] and the other two terms incorporate the effects of the moving boundaries. Note that we get one term proportional to the the total free charge in the depletion layer and another proportional to the first moment of the free charge. Equation 6 is the general solution for the moving boundary current generator.

Equation 6 is formulated in the time domain and has *memory* since the instantaneous collector current depends on the terminal conditions for a backwards interval in time(as would be expected for a NQS model). Quasi-static models only depend on the instantaneous terminal conditions and their first time derivatives. In a time-domain simulator, implementation of the NQS model would require storage of the model state for at least the transit time. In a frequency-domain harmonic balance simulator, since we are examining periodic behavior, we have access to the entire terminal time-domain waveforms from the coefficients of the Fourier expansion. This will provide the necessary information to implement a NQS simulation without significant changes to the simulator. The same could be accomplished in a waveform balance simulation in the sampled time-domain.

It has been shown [10] that motion of the space-charge region boundaries generates intermodulation products. In addition, the phase of the modulation can have a significant effect on fundamental output.

5 Quasi-Neutral Regions

The quasi-neutral minority current is modeled using drift aided diffusion assuming a constant electric field[9, 13]. The solution for two boundaries is

$$\begin{pmatrix} J_1 \\ J_2 \end{pmatrix} = \frac{qD_m\zeta}{\sinh(\zeta)W}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \quad (7)$$

where

$$A_{11} = \frac{\eta\sinh(\zeta)}{\zeta} + \cosh(\zeta) \quad (8)$$

$$A_{12} = \exp(-\eta) \quad (9)$$

$$A_{21} = \exp(\eta) \quad (10)$$

$$A_{22} = \frac{-\eta\sinh(\zeta)}{\zeta} + \cosh(\zeta) \quad (11)$$

and

$$\zeta = \sqrt{\eta^2 + (\frac{W}{\mathcal{L}})^2} \quad (12)$$

$$\mathcal{L} = \sqrt{\frac{D_m\tau_m}{1 + j\omega\tau_m}} \quad (13)$$

$$\eta = \frac{q\mathcal{E}W}{2kT} \quad (14)$$

$\mathcal{L}$ is the frequency dependent diffusion length, W is the width of the region, m is the excess minority density, $\mathcal{E}$ is the electric field (positive for aiding field), and J is the resulting current. This is easily reduced for the long base case.

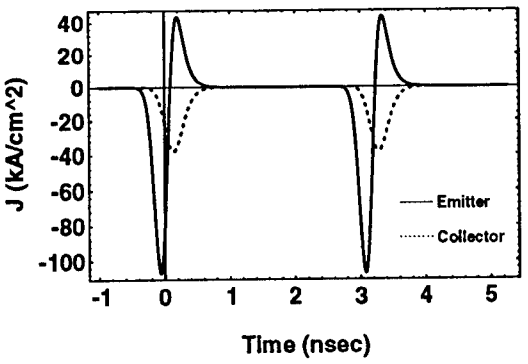


Figure 4: NQS Class C Base Edge Currents

Previous NQS models[5, 6, 7, 8, 9] have truncated these results in various ways. Since this module is in the frequency domain, no truncation is required. Simulation of these equations for class C operation yields minority currents as shown in figure 4. Note that as minority currents cross neutral regions their higher frequency components are strongly attenuated. Also note the positive diffusion current as carriers are extracted back through the emitter junction.

The majority current is considered to be purely drift and is modeled in the time domain using a lumped non-linear resistor approximation. The resistance is calculated from the boundary positions and from the distribution of majority current(from the inverse relation between majority and minority currents). Thus we obtain the voltages V_{bB}, V_{Cc}, and V_{Ee}.

	Dim (Å)	% Al	N (cm^{-3})
Emitter Cap	10000	0	10^{18}
Emitter	10000	30	10^{17}
Base	1000	0	10^{18}
Collector	10000	0	10^{16}
Subcollector	10000	0	10^{18}

Figure 5: HBT Structure Parameters

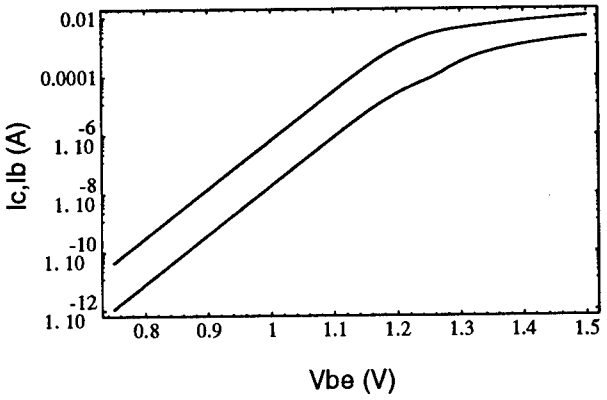


Figure 6: Gummel Plot

of the corresponding collector space-charge region boundaries are shown in figures 7 and 8.

The output characteristics shown in figure 9 taken around the critical current of 3.2 mA clearly show the "quasi-saturation" distortion of the knee between forward active and saturation. Figure 9 also illustrates a slight Early effect, reverse operation, and predicts a offset voltage of .13 volts.

6 Simulations

The model has been integrated into a non-linear simulator. DC simulations were performed on a typical AlGaAs/GaAs HBT structure shown in figure 6. A round number of 10^5 m/s for the scattering limited velocity and an emitter stripe dimension of 2 x 10 μm^2 were used. A model evaluation requires $<$ 10 ms and a converged DC solution requires $\approx$ 200 ms on a DecStation 3100. The Gummel plot shown in figure 6 shows the turnover associated with resistive effects and the current gain falloff from base pushout. Plots

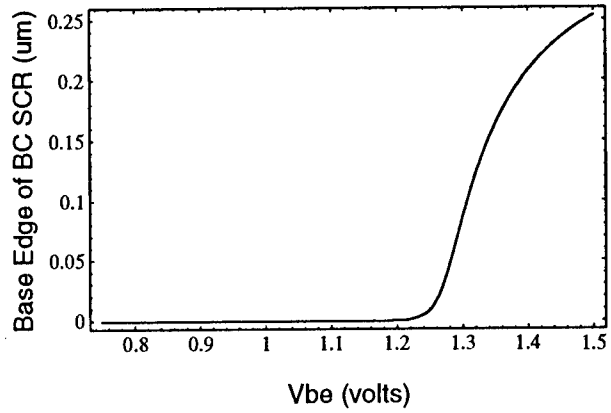


Figure 7: Base edge of collector

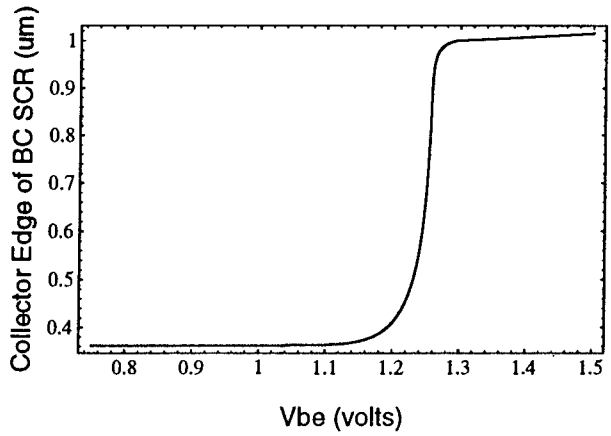


Figure 8: Sub-collector edge of collector

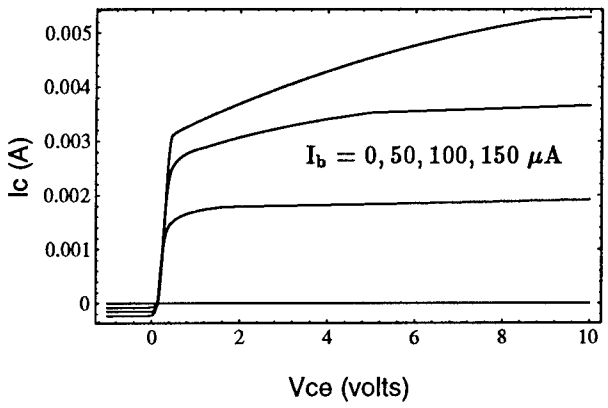


Figure 9: Output Characteristics

Expansion of the simulator to yield RF behavior is in progress. This involves implementing harmonic and waveform balance using the same non-linear equation solver.

7 Conclusions

A large-signal non-quasi-static model of a bipolar transistor has been developed. It is formulated in regional modules to allow moving boundaries and to provide clear division of charge storage and transit mechanisms. Charge transit is simulated in the frequency domain and the space-charge regions are simulated in the time domain. The modules are easily modified. This modeling strategy will allow device engineers more flexibility than conventional circuit based models.

The model parameters are the device geometry and material parameters thus providing both predictive capability and process optimization capability.

Each module has been tested in isolation. The NQS nature of each has been demonstrated. NQS modeling is imperative for very fast nonlinear devices. NQS effects include delay, harmonic filtering and generation, and reverse currents.

The modules have been integrated into a simulator and DC simulations have been performed. The DC results illustrate behavior such as the Kirk Effect that previously required use of numerical simulators.

The NQS modular bipolar model is an example of an alternative approach to solid state device modeling that can include many physical effects not contained in standard circuit models and can be computationally efficient. The NQS nature of the modules should provide more accurate simulations and ease convergence difficulties.

References

[1] J. J. Ebers and J. L. Moll,"Large Signal Behaviour of Junction Transistors", *Proceedings*

of IRE, vol. 42, 1954.

[2] H. K. Gummel and H. C. Poon, "An Integral Charge Control Model for Bipolar Transistors", *Bell System Technical Journal*, vol. 49, 1970.

[3] I. E. Getreu, *Modeling the Bipolar Transistor*, Elsevier 1978.

[4] P. Antognetti and G. Massobrio, *Semiconductor Modeling with Spice*, McGraw-Hill, 1988.

[5] B. S. Wu and F. A. Lindholm, "One-Dimensional Non-Quasi-Static Models for Arbitrarily and Heavily Doped Quasi-Neutral Layers in Bipolar Transistors", *IEEE Trans. on Electron Devices*, vol. ED-36, pp. 727-737, April 1989.

[6] B. S. Wu and F. A. Lindholm, "Non-Quasi-Static Models Including All Injection Levels and DC, AC, and Transient Emitter Crowding in Bipolar Transistors", *IEEE Trans. on Electron Devices*, vol. ED-38, pp. 167-177, January 1991.

[7] J. S. Hamel and C. G. Selvakumar, "The General Transient Charge Control Relation: A New Charge Control Relation for Semiconductor Devices", *IEEE Trans. on Electron Devices*, vol. ED-38, pp. 1467-1476, June 1991.

[8] J. G. Fossum and S. Veeraraghavan, "Partitioned-Charge-Based Modeling of Bipolar Transistors for Non-Quasi-Static Circuit Simulation", *IEEE Electron Device Letters*, vol. EDL-7, pp. 652-654, December 1986.

[9] M-K Chen, F. A. Lindholm, and B. S. Wu, "Comparison and Extension of Recent One-Dimensional Bipolar Transistor Models", *IEEE Trans. on Electron Devices*, vol. ED-35, pp. 1096-1106, July 1988.

[10] A. S. Morris and R. J. Trew, "A Non-Quasi-Static Collector Model for HBTs", *Proceedings of IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits*, pp. 366-373, 1991.

[11] Z. Djurić, M. Smiljanić, B. Radjenović, "The Application of Ramo's Theorem to the Impulse Response Calculation of a Reach-Through Avalanche Photodiode", *Solid State Electronics*, vol. 27, pp. 833-835, 1984.

[12] H. F. F. Jos, "A Model for the Non-Linear Base-Collector Depletion Layer Charge and its Influence on Intermodulation Distortion in Bipolar Transistors", *Solid State Electronics*, vol. 33, pp. 907-915, 1990.

[13] R. L Pritchard, *Electrical Characteristics of Transistors*, pp. 328-371, McGraw-Hill, 1967.

[14] Y. Sad and M. H. Schultz, "A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comput.*, vol. 7, pp. 856-869, 1986.

[15] R. M. Warner and B. L. Grung, *Transistors - Fundamentals for the Integrated Circuit Engineer*, pp. 347-384, Kreiger, 1990.

Monte Carlo Simulation of Wide AlGaAs Barriers

R.J. Mills, G.M. Dunn and Alison B. Walker
School Of Physics, University Of East Anglia,
Norwich, NR4 7TJ. ENGLAND

M.E. Daniels, P.J.Bishop, K.O. Jensen and B.K. Ridley.
Dept. of Physics, University of Essex, Wivenhoe Park,
Colchester, CO4 3SQ. ENGLAND

D.C. Herbert, J.H. Jefferson
Electronics Division, Defence Research Agency,
St. Andrews Rd., Malvern, Worcestershire. WR143PS. ENGLAND

July 29, 1993

Abstract

An theoretical study has been made of electron transport over a wide AlGaAs barrier with graded interfaces between GaAs contact layers doped at $1x10^{18}cm^{-3}$. Drift Diffusion theory has been used giving excellent agreement with experimental current - voltage curves over the temperature range 100-200K. We also present a fully self Consistent 1-D Monte Carlo simulation in which the change in alloy composition in the graded interfaces is accounted for by position dependent scattering. We present a method used in the Monte Carlo simulation which allowed the modelling of the heavily doped contact regions even though the electron density could change by up to 8 orders of magnitude in the device.

1 Introduction

An experimental investigation of electron transport over wide AlGaAs barriers, which is the precursor to an investigation of transport over quantum wells, has been made [1]. In order to be able to explain the results of both these systems it was necessary to develop

0-7803-0894-8/93 $3.00

suitable models. The use of A Drift Diffusion model, which is discussed briefly here, was in excellent agreement with the experimental results at low bias. However, at high bias, the drift diffusion results differed as ballistic effects became more significant. To model the transport properties of the device at high fields requires the use of a Monte Carlo Simulation. With the Monte Carlo simulation we have to overcome 2 major problems. Firstly, the change in Alloy composition means the electron experiences different band structure, consequently different scatter rates as it moves through the structure. Secondly, the presence of the barrier causes the electron density in the device to change by many orders of magnitude, consequently nearly all of the simulated particles reside in the contacts.

In this paper we will present a Monte Carlo method which incorporates position dependent scattering to accommodate the change in alloy composition. Also, we introduce a method used in the Monte Carlo which allows the simulation of devices where the number density can change by many orders of magnitude with no increase in computational time.

2 Sample Structure

The samples were prepared by molecular beam epitaxy (MBE) on $< 100 >$ - orientated, heavily Si-doped (n^+) GaAs substrates. The samples were grown at the Defence Research Agency (DRA) Malvern.

A $1\mu m$ n^+ GaAs layer ($1x10^{18}cm^{-3}$) was deposited on the substrate. The nominally undoped, symmetrical, AlGaAs barrier was then grown in three stages (a) graded layer with x increasing from 0 to 0.25 over 500Å, (b) central $Al_{0.25}Ga_{0.75}As$ region with a sample dependent width between 700-2100Å, (c) second graded layer with x decreasing from 0.25 to 0 over 500Å. Finally, a $1\mu m$ n^+ GaAs layer ($1x10^{18}cm^{-3}$) was deposited.

The materials were processed into mesas using standard photolithograhic techniques. The mesas were round, with diameters between 45-375μm. AuGeNi contacts were deposited onto both the base and the top of the mesa

3 Drift Diffusion Model

The Drift Diffusion model used assumes that transport is collision dominated rather than ballistic, that there is no tunneling and that the electrons are always completely thermalised at the temperature of the lattice. These assumptions are valid at low electric fields in the steady state.

Here, a steady-state, one dimensional, self-consistent Drift Diffusion model, which includes all aspects of a variable composition structure [2] [3], is solved using a relaxation method [4]. The method iteratively improves on an initial guessed solution to the four coupled linear differential equations,

$$\frac{dJ}{dx} = 0 \tag{1}$$

$$J(x) = -e\mu(x)n(x)\frac{d\psi}{dx} \tag{2}$$

$$\frac{d\phi}{dx} = -E \tag{3}$$

$$\frac{dE}{dx} = \frac{e}{\epsilon(x)}(N(x) - n(x)) - \frac{1}{\epsilon(x)}E\frac{d\epsilon}{dx} \tag{4}$$

where J is the current density, μ is the electron mobility, $n(x)$ is the electron density, ψ is the Fermi energy, ϕ is the electrostatic potential, E is the external electric field, ϵ is the dielectric constant and N is the net doping density, $N = N_D - N_A$ where N_D and N_A are the donor and acceptor densities.

4 Monte Carlo Simulation

Monte Carlo simulation is a method of solving the Boltzmann Equation, without any priori assumptions of the form of the electron distributions. It is possible to study non-equilibrium effects in sub-micron devices in regimes where the Drift Diffusion model breaks down, for example at high fields where transport is increasingly dominated by ballistic effects. The Monte Carlo technique is reviewed in Fawcett, Boardman and Swain (1970) [5].

In the present Monte Carlo we use a spherical, non-parabolic 3-band approximation. We include acoustic phonon, polar optic phonon, ionised impurity, inter valley and alloy scattering. Material parameters have been taken from [6]. Velocity-field curves for $Al_xGa_{1-x}As$ obtained from the Monte Carlo were in good agreement with experimental results [7].

An outstanding problem in Monte Carlo simulation is the inability to include heavily doped contact regions in the simulation of devices where the electron density in some regions can be many orders of magnitude less than in the contacts. Traditionally authors have avoided this problem by using suitable boundary conditions to mimic the action of the contacts [8] [9]. This approximation can, in certain circumstances, be physically incorrect [10].

Results from the present Drift Diffusion Model (fig. 1) show that at low bias the electron density can change by 8 orders of magnitude. This means that the number of simulated particles in the low density region will be negligible. This results in extremely poor statistics in the Monte Carlo resulting in impracticable amounts of CPU at moderate fields and failure at low fields.

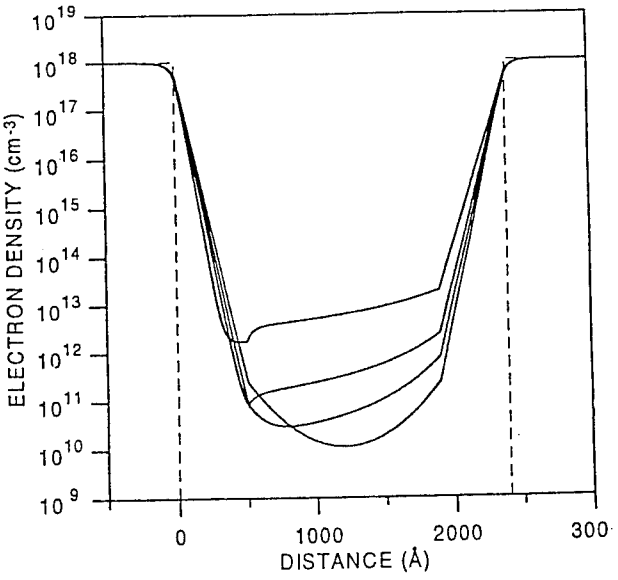


Fig. 1 Electron density as a function of distance across the structure. Plots are shown for applied voltages, from bottom to top at x=1200Å of 0,100,200,400mV. The dashed line represents the doping profile.

In order to combat this problem Philips and Price [11] developed a method which was used by Fischetti et al [12] and Sangiori et al [13] in the simulation of the MESFET. The concept is shown schematically in Fig. 2 The low density region is assigned an enhancement factor, ν. Every particle that enters the region is replicated ν times, the charge on each of the replicated particles being reduced by the same factor ν. The main difficulty with this technique was that the optimum value of ν was only known "a posteriori". Also the changes in carrier density through the MESFET were moderate, when compared to those in the present structure, requiring an enhancement factor of the order of 10.

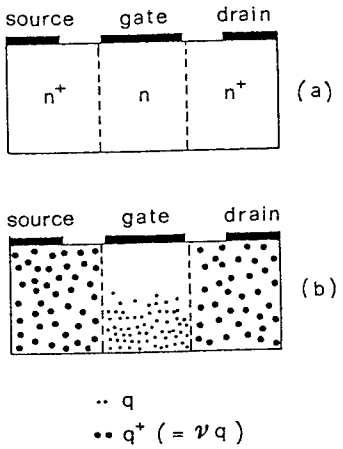


Fig. 2 Pictorial description of the multiplication technique (b) described in the text, applied to the MESFET structure shown in (a) [14]

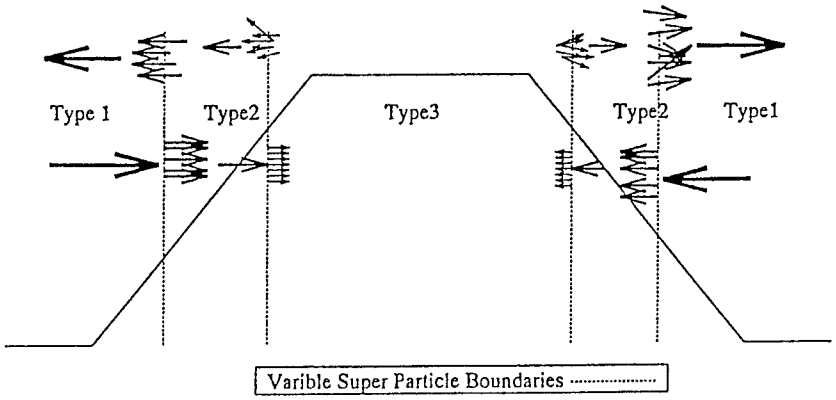


Fig. 3 Schematic Diagram showing position of the points of enhancement for the AlGaAs barrier. Type 1 refers to particles with ν times the charge of type 2 particles, similarly type 3 refer to particle with charge $\frac{1}{\nu}$ times the charge of type 2.

4.1 Self Adjusting Super Particle Region

In the simulation of the AlGaAs barriers we have seen that the change in electron density can be 8 orders of magnitude necessitating enhancement factors of 10^8. Consequently it was necessary to extend the previous method.

From Fig. 3 we used 4 points where the enhancement occured, called Super Particle Boundaries. Thus an overall enhancement of 10^8 could be achieved with an enhancement ratio of 10^4 at each Super Particle Boundary. With such large enhancement factors the placement of these Super Particle Boundaries are critical. If one is set too close to the high density region the number of particles created will results in an intractable simulation.

To combat this problem we first set a limit on the number of simulated particles. If this limit is exceeded in the duration of the simulation the Super Particle Boundaries move to reduce the total number. The simulation thus finds the optimum configuration for the Super Particle regions resulting in an equivalent number electrons in the low and high density regions. This final configuration, which cannot be determined a priori, is not only dependent on the geometry of the device but on the value chosen for the enhancement factor.

4.2 Results of Variable Super Particle Regions

We now compare standard Monte Carlo with the new Variable Super Particle technique. We chose a bias such that the change in number density is not too great for the standard Monte Carlo. We have chosen a enhancement factor of 100 for the Variable Super Particle Monte Carlo.

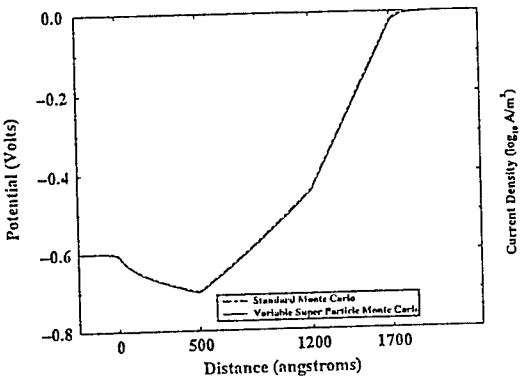


Fig. 4

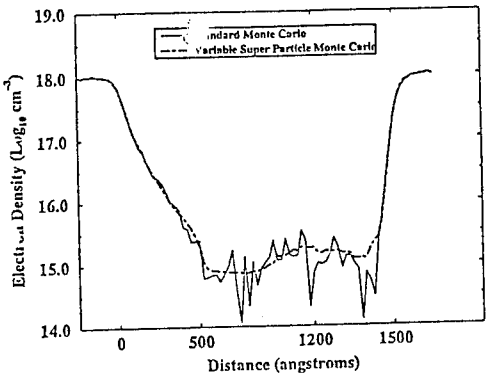


fig. 5

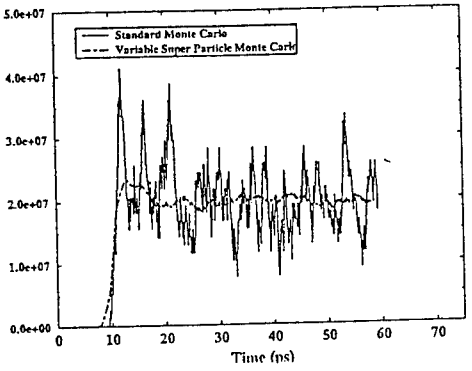


Fig. 6

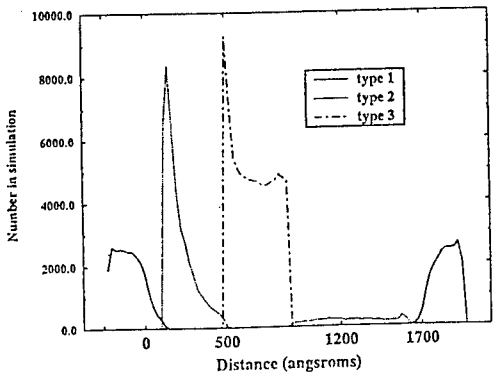


fig. 7

We have applied a Bias of 0.6V which effectively reduced the barrier height and thus the change in number density from the heavily doped contact to the central barrier region is reduced (see Fig. 4). In Fig. 5 we can see the electron density as a function of position throughout the device. The standard Monte Carlo and the Variable Super Particle Method are in good agreement though the fluctuations on the standard MC are significantly higher. The same can be said about Fig. 6 the plot of current density versus time through the

device. From Fig. 7 we can see that the number of type3 electron in the central barrier region is of the same order of magnitude as the number of type1 and type2. The significance of this result is shown if we reduce the applied bias thus raising the effective barrier height. The standard Monte Carlo then gives a zero current as none of the simulated particles are able to surmount the barrier in the time scale of the simulation. However, we have the facility to increase the enhancement factor to 1000 effectively allowing changes in electron density of 6 orders of magnitude and 10,000 for 8 orders of magnitude. In principal the Variable Super Particle Monte Carlo can cope with arbitrary potential profiles which give arbitra electron density profiles by a suitable choice of the enhancement factor.

In a homogeneous system when using Monte Carlo the choice of scatter process chosen for each carrier is determined by its energy and position in k-space. Once these quantities have been determined the selection of scatter process is chosen randomly, the choice weighted according to the relative probability of the processes occuring. In a system where the alloy concentration changes the relative probabilities of the competing scatter mechanisms change with position. Thus in the present Monte Carlo we have included position dependent scattering where the individual scatter rates have been calculated for discreet values of the Al concentration. As the electron moves through the device we determine the scatter process by the position in real space as well as in k-space by the usual Monte Carlo random number selection.

5 Comparison With Experimental Results

From Figure 8 we see that the drift diffusion model is in good agreement with the experimental results at low bias. However, at higher bias the assumption of the electron transport being collision dominated breaks down as ballistic effects become increasingly important. We would like at this point to say that the full details of the Drift Diffusion model and the experimental results can be found in Daniels et al [1]

As we can see in Fig. 9 In order to fit the I-V current at low bias we needed to reduce the barrier height This was to compensate for the fact that the Fermi Energy is 43meV at density $1x10^{18}$ in GaAs by 43mev. This resulted in an over estimation of the current at high bias. We can explain this discrepancy by considering the effects of using classical statistics in Monte Carlo.

5.1 Inclusion of Degenerate Satistics

From Fig. 10 we can see that the difference in energy distribution between the Degenerate and Non-Degenerate contact will be significant a low bias. The degenerate electrons distribution, having a higher mean energy, will effectively see a lower barrier. With the exponential dependence of the current on the effective barrier height, we can see that the

use of degenerate statistics will give a greater current than with non-degenerate statistics. We suggest that the necessity of reducing the barrier height by 40meV to fit at low bias will not be necessary if we use the correct statistics.

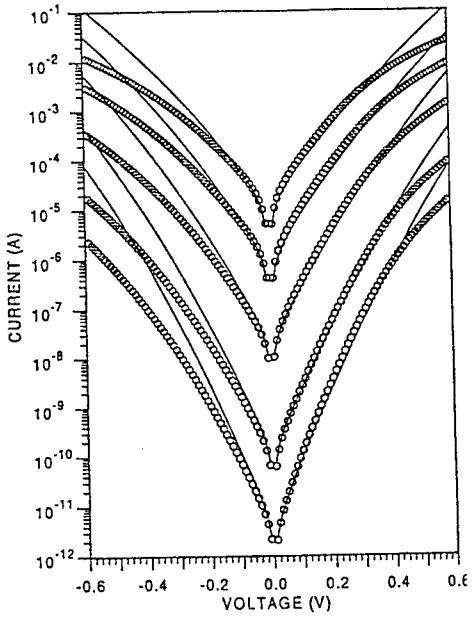


Fig. 8 shows the comparison of the Drift Diffusion model to the experimental results. This was for a central $Al_{0.25}Ga_{0.75}As$ region of 1400Å at temperatures, from top to bottom, of 200K, 180K, 160K,140K, 120K and 100K.

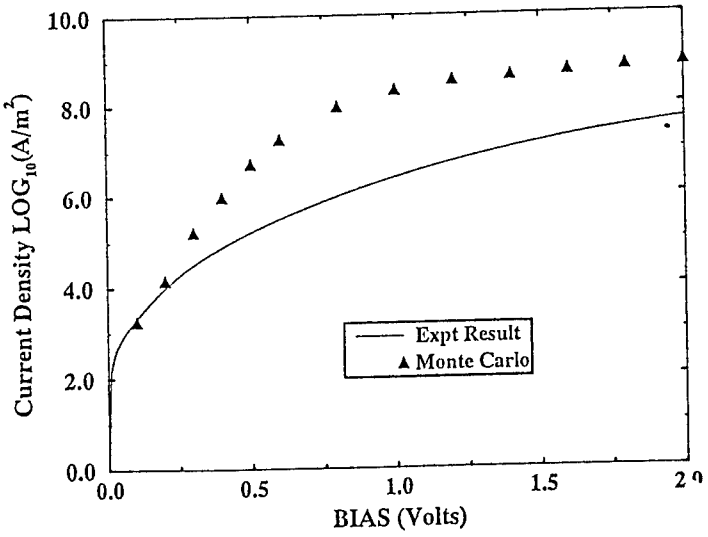


Fig. 9 Comparison Of I-V curve at 200K for central $Al_{0.25}Ga_{0.75}As$ region of 700Å.

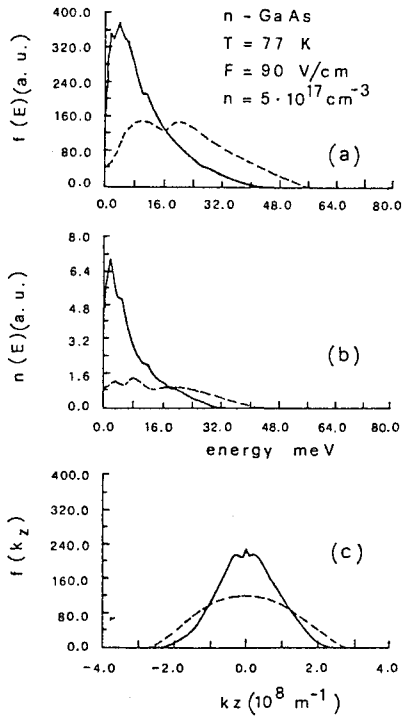


Fig. 10 Effect of exclusion principle in electron dynamics, as obtained with MC calculation. Figure shows distribution function of energy in GaAs at 77K with an applied electric Field of 90 V/cm. [15]

6 Conclusion and Future Work

We have present a new method which allows the heavily doped contact to be simulated in Monte Carlo. We have also allowed for the change in alloy concentration by including position dependent scattering. The discrepancy between Monte Carlo and experimental results is believed to be caused by incorrect statistics in the doped contact region. In future publications this will be corrected.

We will also be looking at the inclusion of quantum wells in the central region. Experimental results have been gained on such systems and results are being presented at the 8^{th} International Conference on Hot Carriers in Semi-conductors. August 1993, Oxford, England. We will attempt to model the trapping of thermally excited electrons both on the Drift Diffusion and Monte Carlo simulation. We would also like to look at the effects of sequential tunneling in the central region where many wells are present.

References

[1] M.E. Daniels, P.J. Bishop, K.O. Jensen, B.K. Ridley, D.A. Ritchie, M. Grimshaw, E.H. Linfield, G.A.C. Jones and G.W. Smith. to be published in J. Appl. Phys.

[2] J.E. Sutherland and J.R. Hauser, IEEE Trans. Elec. Dev. **24** 363 (1977).

[3] A.H. Marshak and C.M. van Vliet, Proc. IEEE **72** 148 (1984).

[4] Computer Simulation Using Particles. R.W. Hockney and J.W. Eastwood, **Adam Hilger** (1989). (Oct. 1993).

[5] W. Fawcett, A.D. Boardman and S. Swain. J. Phys. Chem. Solids. **31** pp1963-1990 (1969).

[6] S. Adachi, J. Appl. Phys. **58** 3 R1-R28(1985).

[7] M.V. Fischetti, IEEE. Trans. Elec. Dev. **38** 3 pp634-649 (1991).

[8] G. Baccarani C. Jacoboni and A. M. Mazzone, Solid-State Elec. **20** p.5-10 (1977).

[9] C.M. Maziar, M.E. Klausmeier-Brown, S. Bandyopadhyay, M.S. Lundstrom and S. Datta, IEEE Trans. Elec. Dev. **ED-33** 7 p881-888 (1986).

[10] R.J. Mills, D.C. Herbert, J.H. Jefferson and Alison B. Walker, Semicond. Sci. Technol., **8** 9 (1993).

[11] A. Philips, P.J. Price. Appl. Phys. Lett. **30** 528 (1977).

[12] M.V. Fischetti, S.E. Laux. Phys. Rev. B **38** 9271 (1988).

[13] E. Sangiori, B. Ricco, F. Venturi. IEEE. Trans. Computer-Aided Design CAD-7,259 (1988).

[14] Y. Yamada, S. Ikeda, N. Shimojok : In Proc., XVII European Solid State Device Research Conference ESSDERC87 p.111. Bologna (1987).

[15] C. Jacaboni and P. Lugli. 'The Monte Carlo Method for Semiconductor Device Simulation'. **Springer-Verlag** (ed. S. Selberherr) 1989.

Model-Based Comparison of RF Noise in Oscillating Diamond and SiC MESFETs

G. L. Bilbro M. W. Shin R. J. Trew
North Carolina State University, Raleigh, NC 27695-7911

A. N. Riddle
Macallan Consulting, 1583 Pinewood Way, Milpitas, CA 95035

July 23, 1993

Abstract

We report a simulation study of RF oscillator noise including $1/f$ noise and incomplete activation in diamond and silicon carbide MESFETs oscillating in resonant circuits that maximize power at the load.

1 Introduction

We have developed a new system to predict RF power and noise performance in MESFET oscillators by combining a large-signal circuit simulator, a two-dimensional semiconductor device model, a microwave oscillator design procedure, and a general theory of oscillator noise.

The NCSU large-signal MESFET model uses harmonic balance to simulate the interaction between a linear circuit including parasitics and an active nonlinear device represented as a table of quasi-static capacitances, currents, and delays for relevant biases[1].

The resulting large-signal simulation is used to determine an oscillator circuit using conventional techniques. A maximum output optimal resonant filter of a simple fixed shunt topology is calculated from large-signal RF simulations[2].

The complete oscillator circuit is characterized with additional RF large-signal simulations and the effects of broadband additive noise and $1/f$ gate noise on the oscillator amplitude are calculated[5]. The resulting noise spectra for amplitude S_A and phase S_ϕ are then compared for diamond and SiC.

This report is presented as a first step toward a practical method for noise simulation suitable for computer aided design of microwave oscillator circuits using diamond or SiC MESFETs.

0-7803-0894-8/93 $3.00

2 Oscillator Noise Theory

In this section, the general theory of oscillator noise due to Riddle will be reviewed[5, 6, 7]. The resulting relation among input noise, transfer, and tapped output will be used to obtain an expression for output noise spectra in terms of input noise sources.

The general theory is formulated in terms of the dimensionless inverted transfer function D. Negative resistance, negative conductance, feedback, and reflection oscillators can all be realated to D. D describes the circuit at the large signal operating point. The equation

$$n = DA \tag{1}$$

relates the circuit noise to the output signal through D. This approach is taken becasue it easily handles all oscillator configurations and can account for additve and upconverted noise[5]. It describes how the output signal is balanced by the large gain of the circuit and the finite input noise power.

The inverse transfer function $D = U + jV$ is written in terms of its real and imaginary parts, which are expanded in a Taylor series,

$$D = U_0 + \delta A\frac{\partial U}{\partial A} + j\delta A\frac{\partial V}{\partial A} + \delta\omega\frac{\partial U}{\partial \omega} + j\delta\omega\frac{\partial V}{\partial \omega} + \delta\epsilon\frac{\partial U}{\partial \epsilon} + j\delta\epsilon\frac{\partial V}{\partial \epsilon} \tag{2}$$

where jV_0 is omitted because the net phase shift around loop vanishes. The quantities

$$\begin{aligned} \delta A &= A - A_0 \\ \delta\omega &= \omega - \omega_0 = \omega_m \\ \delta\epsilon &= \epsilon - \epsilon_0, \end{aligned} \tag{3}$$

are small deviations around the mean operating conditions $A_0, \omega_0, \epsilon_0$ for the oscillator. All partials are evaluated at the mean operating conditions. In this theory the oscillator poles do not lie on the $j\omega$ axis. Previous theories encountered sigularities because of the assumption of $j\omega$ poles[3]. The finite noise power in the oscillator, $\overline{n^2}$, leads to a finite loop gain, $1/U_0$, and a finite output power given by $A_0^2 = \overline{n^2}/(U_0^2)$ If n were zero then infinite gain would be required to give a finite output power — the $j\omega$ axis poles would produce infinite gain. The finite loop gain allows the $j\omega$ poles to be slightly to the left of the $j\omega$ axis. An intersting implication of the interaction between nonlinearities, loop gain and noise in the oscillator is that the oscillator poles become statistical variables.

Additive input is modeled as

$$n = (n_i + jn_q)e^{j\omega_0 t}. \tag{4}$$

The oscillator output is modeled as pure carrier A_0 with small perturbations in amplitude and phase

$$A(t) = A_0(1 + m)cos(\omega_0 t + \beta) \tag{5}$$

so that

$$A \approx A_0(1 + m + j\beta)e^{j\omega_0 t} \tag{6}$$

where $m_0 \cos \omega_m t$ is amplitude modulation and $\beta_0 \cos \omega_m t$ is phase modulation.

Equations 1 and 2 together with the expressions for n and A can be used to obtain in-phase part of n

$$\frac{n_i}{A_0} = \left[U_0 + A_0 \frac{\partial U}{\partial A} + j\omega_m \frac{\partial V}{\partial \omega}\right] m + \left[j\omega_m \frac{\partial U}{\partial \omega}\right] \beta + \left[\delta\epsilon \frac{\partial U}{\partial \epsilon}\right] \tag{7}$$

and the quadrature part of n

$$\frac{n_q}{A_0} = \left[A_0 \frac{\partial V}{\partial A} - j\omega_m \frac{\partial U}{\partial \omega}\right] m + \left[U_0 + j\omega_m \frac{\partial V}{\partial \omega}\right] \beta + \left[\delta\epsilon \frac{\partial V}{\partial \epsilon}\right] \tag{8}$$

Equation 1 is split into Equations 7 and 8 so that amplitude and phase modulations can be obtained directly. Equations 7 and 8 involve stochastic variables for additive noise $n = n_i + jn_q$, modulation $m + j\beta$, and low frequency perturbation $\delta\epsilon$, as well as deterministic coefficients for the means $A_0, \omega_0, \epsilon_0$, and the partials of U, V with respect to A, ω, ϵ which are certain.

Equations 7 and 8 can be solved for m, the amplitude modulation, and β, the phase modulation, in terms of n_i, n_q, and $\delta\epsilon$. From the resulting expression together with the statistics of n_i, n_q, and $\delta\epsilon$, the phase amplitude noise spectral densities

$$\begin{aligned} S_\phi(f_m) &\equiv \overline{\beta^2} \\ S_A(f_m) &\equiv \overline{m^2} \end{aligned} \tag{9}$$

can be obtained for the oscillator at offset f_m from the carrier. Here the overline indicates expectation including correlations between n_i and n_q if the additive noise is not white, as well as correlations with $\delta\epsilon$. Even though the theory is general, this paper will address the case of noise for which correlations other than $\overline{n^2}$ and $\overline{\epsilon^2}$ are insignificant.

Given many assumptions this theory reduces to a simple and familiar model for oscillator noise. The device noise changes the device phase through small changes in the device bias (this is the $\partial V/\partial \epsilon$ term and the result is phase noise). The nearness of the poles to the $j\omega$ axis then accounts for the high noise gain and the $1/f^2$ response in the phase noise. The complete theory is required to account for the affects of amplitude saturation and the interaction of device and circuit load lines on the output noise. These load lines were often shown graphically in earlier theories[3].

For a feedback oscillator $D = 1 - KH$ with amplification K and filter H, but at microwave frequencies, it is convenient to characterize the two-port active GaAs MESFET device with large-signal $\hat{S}$ parameters and the passive two-port resonant filter subcircuit including the load with $\tilde{S}$. Then

$$\begin{aligned} D = 1 &- \hat{S}_{21}\tilde{S}_{21} - \tilde{S}_{11}\hat{S}_{22} - \hat{S}_{11}\tilde{S}_{12}\tilde{S}_{21}\hat{S}_{22} \\ &- 2\hat{S}_{11}\tilde{S}_{22} - \tilde{S}_{11}\hat{S}_{12}\hat{S}_{21}\tilde{S}_{22} \\ &- \hat{S}_{11}\hat{S}_{12}\tilde{S}_{12}\hat{S}_{21}\tilde{S}_{21}\tilde{S}_{22} + 2\hat{S}_{11}\tilde{S}_{11}\hat{S}_{22}\tilde{S}_{22} \\ &+ \hat{S}_{11}^2\tilde{S}_{12}\tilde{S}_{21}\hat{S}_{22}\tilde{S}_{22} + \hat{S}_{11}^2\tilde{S}_{22}^2 \\ &+ \hat{S}_{11}\tilde{S}_{11}\hat{S}_{12}\hat{S}_{21}\tilde{S}_{22}^2 - \hat{S}_{11}^2\tilde{S}_{11}\hat{S}_{22}\tilde{S}_{22}^2 \end{aligned} \tag{10}$$

Parameter	units	Diamond	SiC
a	μm	0.15	0.2
$Doping$	$10^{17} cm^{-3}$	4.0	1.3
v_{sat}	$10^6 cm/sec$	8.6	8.7
μ	cm^2/sec	85	99
$E_{activation}$	eV	0.37	0.18
ϵ_r		5.7	10.0

Table 1: Parameters used in simulation of p-type diamond and n-type SiC MESFETs at 500 C.

The device parameters $\hat{S}$ depend most strongly on bias and amplitude. The resonant filter is linear so that the filter parameters $\tilde{S}$ depend only on ω.

3 Oscillator Noise for Diamond and SiC

Each device was characterized by tables of currents and capacitances determined by DC and low frequency simulations with Pisces-IIB, which treats incomplete carrier activation and other thermal effects in the channel and at the metal contacts, as well as detailed device geometry[4] as shown in Table 1 for $T = 500\ C$, $L_g = 0.5\ \mu m$, $L_{gs} = L_{gd} = 1\ \mu m$ and $W = 1000\ \mu m$. The resulting IV curves are shown in Figures 1 and 2.

Inspection of the DC IV curves provides biases for for subsequent RF simulations. the p-type diamond device was biased at $V_{gs} = 10\ V$ and $V_{ds} = -50\ V$ and and the n-type SiC device was biased at $V_{gs} = -3\ V$ and $V_{ds} = 50\ V$.

Harmonic balance was used obtain the MESFET $\hat{S}$ parameters by simultaneously solving the preceding DC device representation with typical values for parasitic elements for 500Ω load and source. The MESFETs were simulated at 4 GHz with harmonics at 8 GHz and 12 GHz over a range of input powers. The diamond circuit oscillates at 24 dBm and the SiC circuit at 18 dBm. The amplitude of oscillation was adjusted to maximize added power produced by each device and the S parameters and their partial derivatives with respect to frequency, power, and gate bias were estimated at those maxima. The results are shown in Figures 3 and 4. These two devices differ in doping and channel depth. For these material and circuit parameters, the diamond device has lower gain, but higher output power, Pout, and power added efficiency, PAE. For MESFETs with equal geometry and doping, SiC exhibits substantially higher Pout.

For each device, a resonant filter was designed (as a convenient example) to maximize power delivered to the load (rather than to minimize oscillator noise) using a conventional technique[2] for the shunt topology of Figure 5. The $\tilde{S}$ parameters and their partial derivatives with respect to frequency were calculated for the optimal filters.

The partial derivatives for filter and device were combined to compute the partial derivatives of Equation 2 using Equation 10. The resulting partials were used to evaluate the expressions for noise spectra in Equation 9.

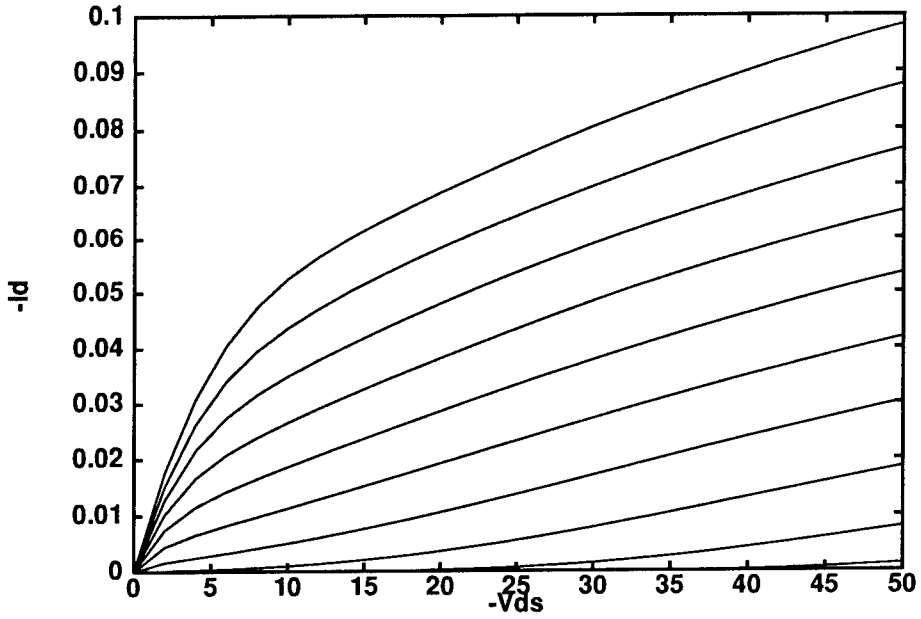


Figure 1: DC IV curves for Diamond MESFET for $V_{gs} = 0, 2, ..., 20$.

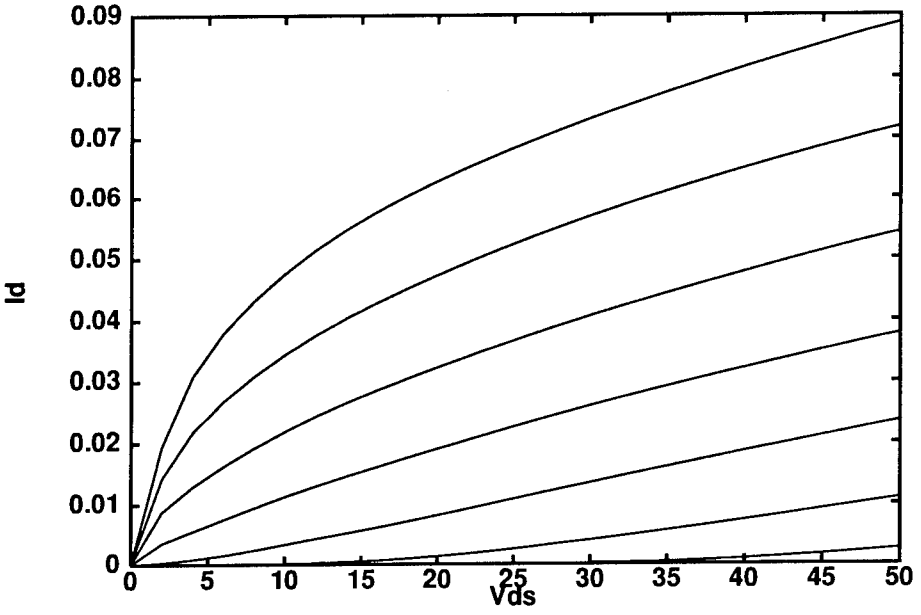


Figure 2: DC IV curves for SiC MESFET for $V_{gs} = -0, -1, ..., -10$.

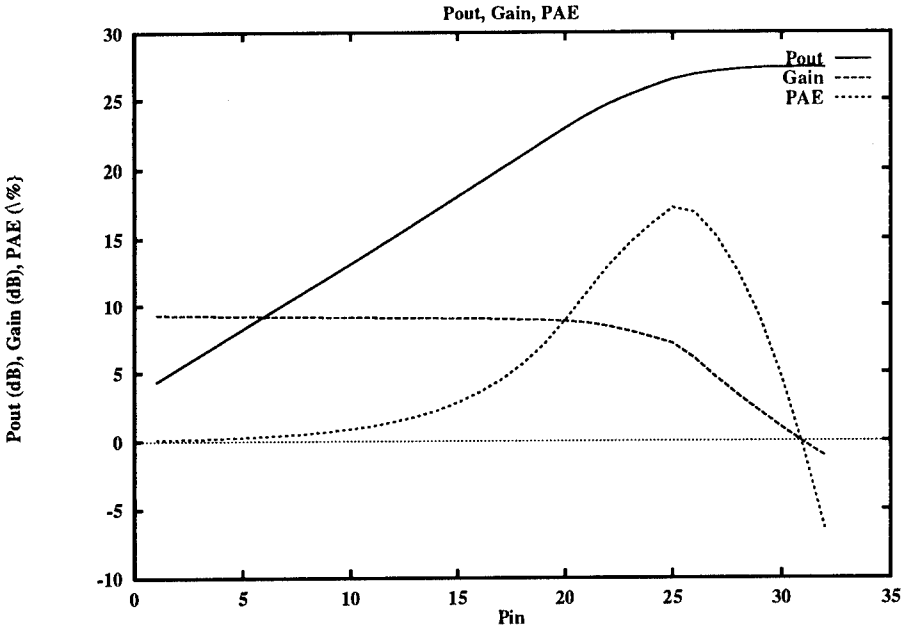


Figure 3: RF performance of Diamond MESFET.

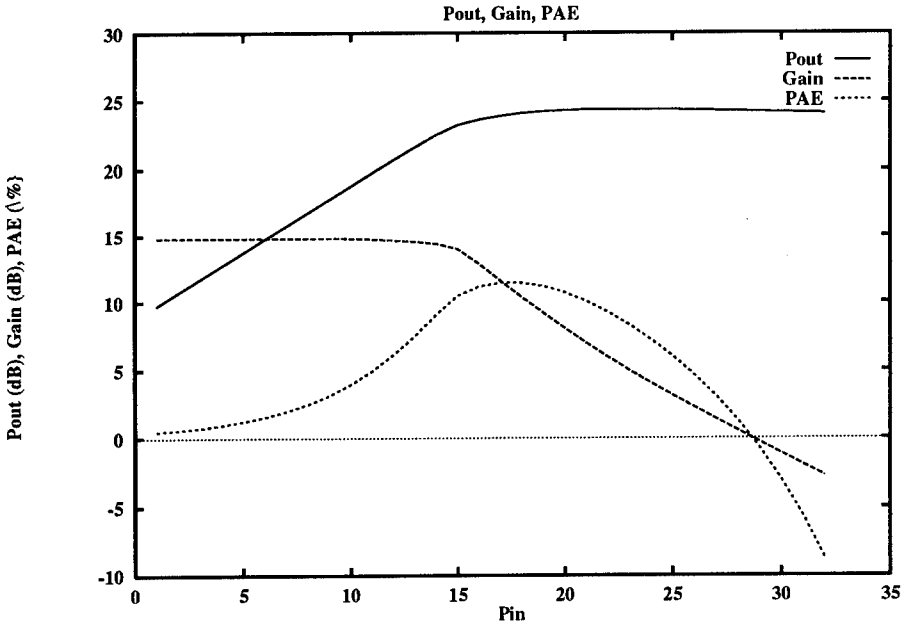


Figure 4: RF performance of SiC MESFET.

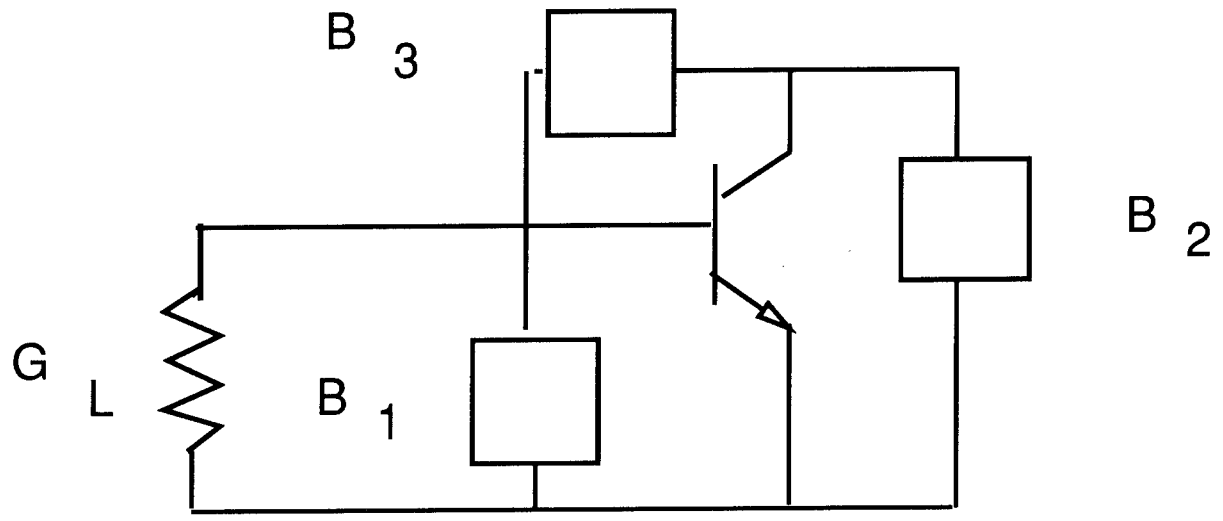


Figure 5: Topology of resonant circuit.

The results are plotted in Figures 6 and 7 for additive noise amplitude $1.0 \times 10^{-9}\ V/\sqrt{Hz}$ and modulation noise amplitude proportional to $1/f$ for small offset and equal to $1.0 \times 10^{-9}\ V/\sqrt{Hz}$ above the corner frequency $1.0 \times 10^{+7}\ Hz$. Until data exists to determine the statistics of the noise sources for SiC or diamond MESFETs, we use identical values for both materials. Simulation results are shown for 4 GHz at 500 C. The noise spectral power curves were calculated with the same additive and modulative input noise statistics but near the carrier, both S_ϕ and S_A are about 3 dB lower for the diamond device.

4 Conclusion

A general procedure for oscillator noise simulation has been presented. The general theory was used to implement a noise simulator for MESFET oscillators by integrating a physics-based RF simulation of the active device with conventional oscillator design techniques.

Electrical characteristics of the active devices were simulated with a large-signal MESFET model which combines the physics of the FET channel with harmonic balance to predict large-signal RF performance. The resonant circuit is determined with conventional oscillator design techniques to maximize power at the load in a simple shunt topology. The inverse transfer function D is computed at four neighboring operating points to estimate partial derivatives of D with respect to frequency, power, and gate bias. These three partials are used to predict noise due to upconversion of baseband noise and to crossconversion of inband noise.

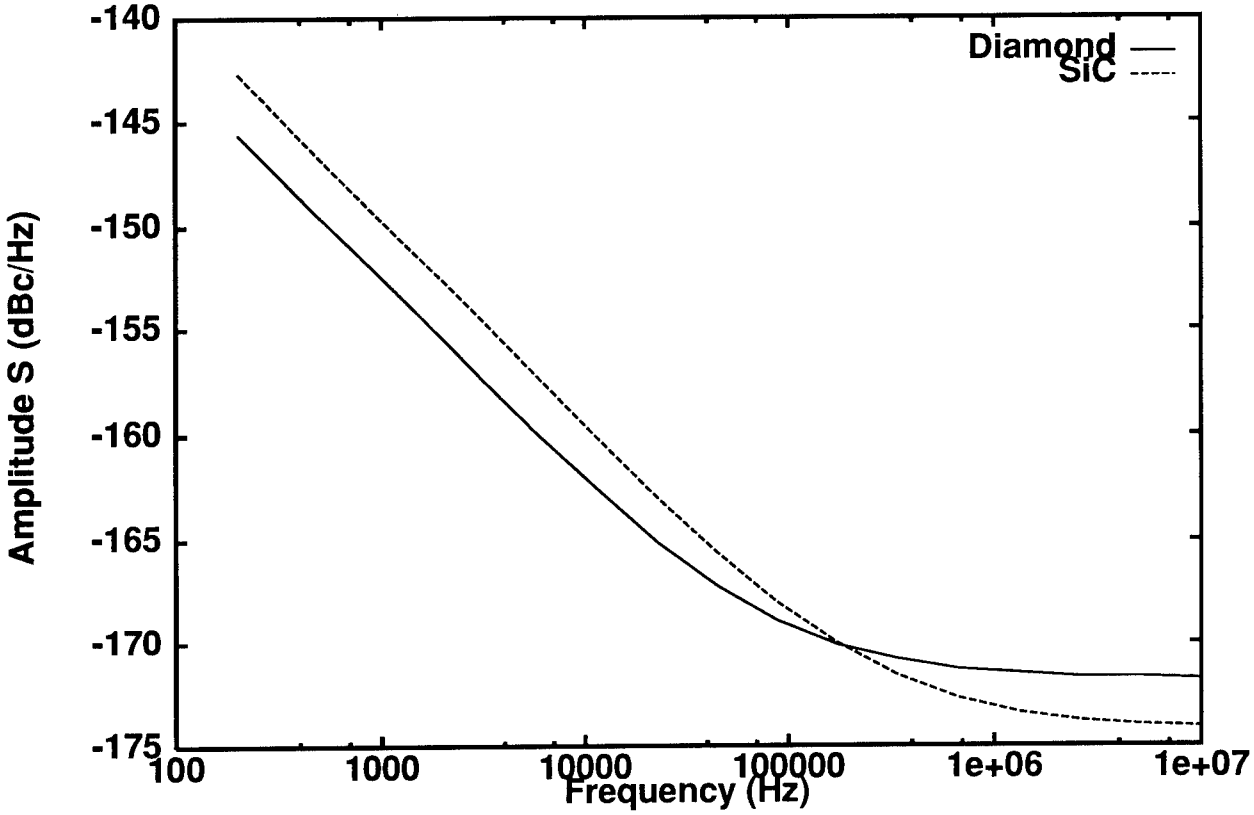


Figure 6: Amplitude noise spectra.

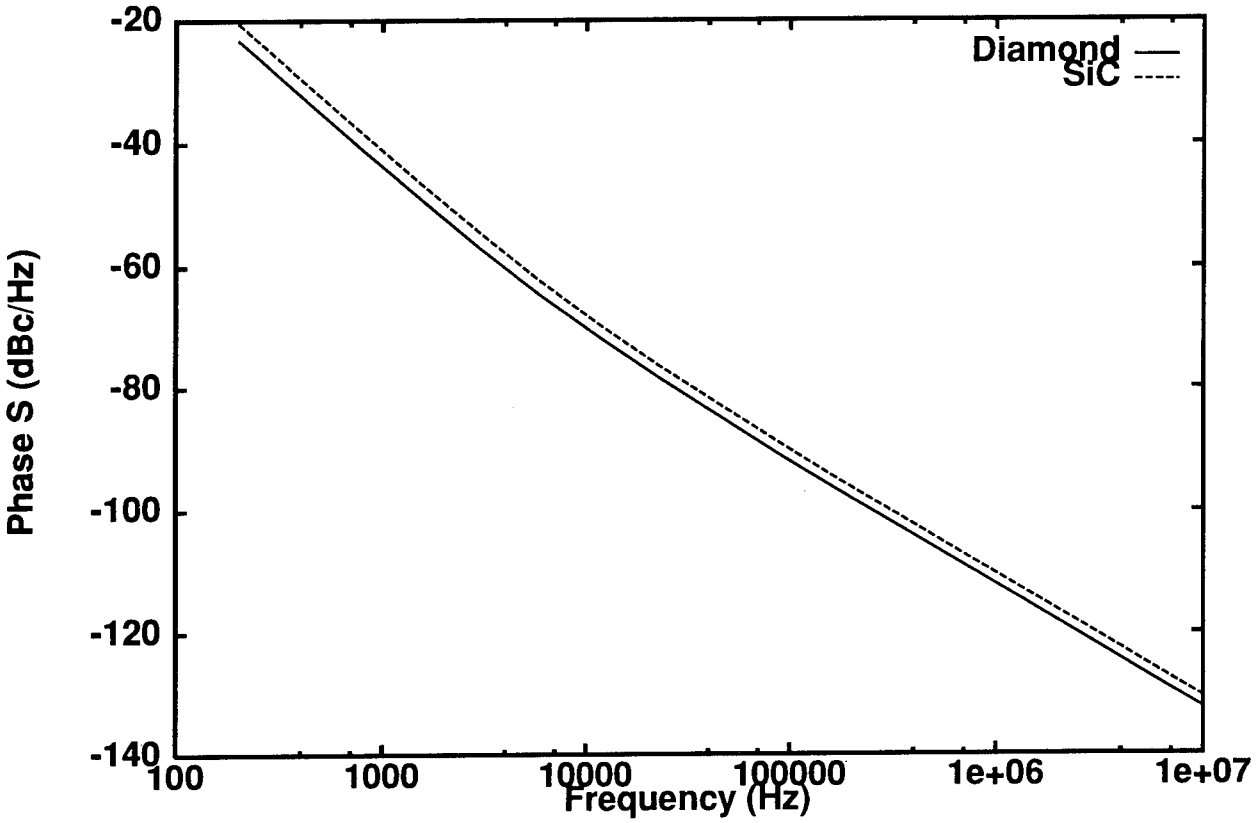


Figure 7: Phase noise spectra.

References

[1] M. A. Khatibzadeh and R. J. Trew. A large-signal, analytic model for the GaAs MESFET. *IEEE Trans. Microwave Theory and Techniques*, MTT-36:231–238, 1988.

[2] K. L. Kotzebue and W. J. Parrish. The use of large signal S-parameters in microwave oscillator design. In *Proc 1975 IEEE Intl. Symp. Circuits and Systems*, 1975.

[3] K. Kurokawa. Noise in synchronized oscillators. *IEEE Trans. Microwave Theory and Techniques*, pages 234–240, April 1968.

[4] M. R. Pinto, C. S. Rafferty, H. R. Yeager, and R. W. Dutton. Pisces-II. Technical report, Stanford Electronics Laboratories, Stanford University, Stanford, CA 94305, 1985.

[5] A. N. Riddle. *Oscillator noise: theory and characterization.* PhD thesis, North Carolina State University, 1986.

[6] A. N. Riddle and R. J. Trew. A new approach to low phase noise oscillator design. In *IEEE/Cornell Conference digest*, pages pp 309–311, 1985.

[7] A. N. Riddle and R. J. Trew. A new measurement system for oscillator noise characterization. In *IEEE MTT-S digest*, pages pp 509–512, 1987.

Thermal Runaway Analysis of High Power AlGaAs/GaAs Heterojunction Bipolar Transistors

L. L. Liou, B. Bayraktaroglu and C. I. Huang
Solid State Electronics Directorate
Wright Laboratory
Wright-Patterson Air Force Base, OH 45433-7323

ABSTRACT

Thermal runaway study of multiple emitter AlGaAs/GaAs power HBTs using an analytical electro-thermal model is described. Thermal runaway causes the fatal destruction of the device under the voltage modulation mode of operation, and thermally-induced current instability of the device under the current modulation mode of operation. The kinetic relation between the thermal runaway and junction temperature rise is studied. The HBT power handling capabilities in relation to the device thermal resistance and ballasting schemes are discussed.

I. INTRODUCTION

Because of carrier confinement in the base, heterojunction bipolar transistors (HBTs) are capable of delivering high power at high frequencies. As the material and process techniques become mature, microwave power performance of AlGaAs/GaAs HBTs is being realized. The poor GaAs substrate thermal conductivity hinders the power performance of HBTs; as a result, they are unable to reach their ultimate limit as determined by the electronic properties of the device materials.

There are three thermal effects that cause AlGaAs/GaAs HBT to cease its transistor function. The first is the temperature dependency of current gain. The current gain of HBTs above room temperature is known to decrease monotonically with temperature. The HBT loses its current amplification function as the junction temperature reaches a "cut-off temperature," at which the current gain is unity. Depending on the device structure and fabrication, the cut-off temperature is varied and is usually larger than 500°K. The second thermal effect is second breakdown, a well known phenomenon in Si bipolar transistors under high power operation. It occurs when the device temperature reaches a semiconductor intrinsic temperature at which the electric properties are dominated by the thermally-induced intrinsic carriers. As a result, the reverse biased base-collector junction loses the rectification function, and the collector current exhibits a rapid increase while collector-emitter voltage shows a rapid decrease. Since the intrinsic temperature of AlGaAs/GaAs is high (greater than 600 °K in the collector, and much higher in the base), the second breakdown may not be a thermal problem. The third one is the thermal runaway. This thermal effect usually occurs at a temperature lower than the other two, and it is the primary reason for limiting the power performance of the devices.

Thermal runaway originates from the fact that the turn-on voltage of base-emitter junction has a negative temperature coefficient. In this work, we studied the thermal runaway phenomenon of multiple emitter AlGaAs/GaAs power HBTs using an analytical electro-thermal model. It will be

0-7803-0894-8/93 $3.00

shown that thermal runaway behavior is dependent on the bias scheme. For the voltage modulation mode, i.e., maintaining a constant base-emitter voltage bias while increasing the collector-emitter voltage, the thermal runaway destroys the device due to a large increase in the junction temperature. For the current modulation mode, i.e., maintaining a constant base current bias while increasing the collector-emitter voltage, the thermal runaway causes the thermally-induced current instability and formation of hot spots. This study also shows that thermal runaway is triggered when the junction temperature rise reaches a certain threshold value. This junction temperature rise threshold depends on the device electrical and thermal parameters such as the base current ideality factors, magnitude of the negative temperature coefficient of the base-emitter turn-on voltage, ballast resistance, and the self-induced and coupled thermal resistances. Finally, the methods to increase the power performance of HBTs through various thermal designs are summarized. These methods include the incorporation of ballast resistance (electrically and thermally) and reduction of the substrate thermal resistance.

II. THEORY

Based on the experimental [1,2] and calculated results of HBT base current density J_b, including the effect of emitter ballast resistance, a simplified analytical expression for J_b is given by:

$$J_b(T) = J_{b0}\, e^{\frac{V_{be}-(\beta+1)J_b\rho_e+\delta(T-T_0)}{n_b kT_0/q}} \qquad (1)$$

where β is the current gain, ρ_e is the specific emitter resistance in Ω-cm^2, k is Boltzmann's constant, T and T_0 are the junction and heat sink temperatures, respectively, n_b is the ideality factor for J_b at heat sink temperature, and J_{b0} is a pre-exponential factor. In this expression, the parameters J_{b0}, n_b and δ depend on the current density, structure parameters and material constants including those describing bulk and surface properties such as effective density of states in conduction and valence bands, minority carrier life time, density of the Shockley-Reed-Hall centers, surface recombination velocity, etc. All of these are involved in the detailed current calculation. For simplicity, we neglect the slight temperature dependence of n_b.

As the junction temperature increases, the charge carriers in the base are less confined by the valence band discontinuity at the emitter-base junction, and consequently, the emitter injection efficiency and current gain decrease [3]. The temperature dependence of current gain can be expressed by:

$$\beta(T) = \beta_0 e^{\frac{\Delta E}{kT}} \qquad (2)$$

where ΔE is the activation energy, also referred to as the "apparent" valence band discontinuity [4], and β_0 is a pre-exponential factor. The parameters of β_0 and ΔE depends on the device structure parameters and current density. The typical experimental value of ΔE is in the range of 0.04 to 0.07 eV [2]. The dissipation heat density of the device can be approximated by:

$$p(T) = \beta(T)J_b(T)V_{ce} \qquad (3)$$

where V_{ce} is the collector-emitter voltage.

A. One Emitter Element Case:

We first analyze a device with one emitter finger (one emitter element with unit

area) fabricated on a GaAs substrate. All heat is assumed to dissipate through the substrate, causing a uniform junction temperature rise of $\Delta T = r_{th}p$, where p is the dissipation heat density and r_{th} is the specific thermal resistance between the active device and the heat sink.

A.1. Voltage Modulation Mode: The dependence of p on V_{ce} for a given V_{be} can be obtained by solving the following nonlinear equation:

$$C_0 = \ln(p) + bc[1+\beta_0^{-1}e^{\frac{-a}{T_0+r_{th}p}}]\frac{p}{V_{ce}} - \frac{a}{T_0+r_{th}p} - br_{th}p \quad (4)$$

where

$$a = \frac{\Delta E}{k}, \quad b = \frac{q\delta}{n_b k T_0}, \quad c = \frac{\rho_e}{\delta}$$

$$C_0 = \ln(\beta_0 V_{ce} J_{b0} e^{\frac{qV_{be}}{n_b k T_0}})$$

The existence of an upper bound of V_{ce} for a real solution of p in (4) can be deduced. This upper bound, denoted as $V_{ce,max}$ corresponds to a bias point where the value of $d\Delta T/dV_{ce}$ goes to infinity, which is the onset of thermal runaway condition. The junction temperature rise at $V_{ce,max}$ is denoted as ΔT_{max}, and is an indicator of power handling capability. The larger the ΔT_{max} is, the more power a device can handle before the thermal runaway takes place.

With the parameters of $n_b = 2$, ΔE = 0.07 eV, δ = 1.57 meV/K, and $T_0 = 300\ ^oK$, ΔT as a function of V_{ce} with V_{be} and the ballast resistance as variable parameters is shown in Fig. 1. The three applied V_{be} values correspond to the collector current densities of 1, 2 and 3×10^4 A/cm^2 when the junction temperature is the same as that of the heat sink. There are four noted features in this figure: (i) The junction temperature rise goes to infinity as V_{ce} approaches $V_{ce,max}$. Forcing the device to operate beyond this bias will destroy the device. (ii) For a given V_{be}, as the ballast resistance increases so does $V_{ce,max}$. The common practice of implementing a ballast resistance to reduce the current mode thermal runaway problem was thus verified [5]. Assuming $\beta^{-1}\exp[-a/(T_0+\Delta T)] << 1$ in (4), a further analysis shows that the device can be made unconditionally thermally stable with a value of ρ_e larger than $\delta r_{th} V_{ce}$. (iii) For a zero ballast resistance, ΔT_{max} is a constant regardless of the value of V_{be} (though $V_{ce,max}$ increases as V_{be} increases). (iv) For a finite ballast resistance, both ΔT_{max} and $V_{ce,max}$ increases with V_{be}. This implies that operating a device with a nonzero ballast resistance at high current and low voltage has a larger safety operation range than those operating at low current and high voltage. Similar conclusions were reached in Si power transistors [6].

The maximum value of the real solution of p in (4) can be solved by analyzing the differential $d\Delta T/dV_{ce}$, and the corresponding ΔT_{max} can be obtained by solving the following equation numerically:

$$\frac{1}{\Delta T_{max}} - b + \frac{a}{(T_0+\Delta T_{max})^2} + bc\gamma = 0 \quad (5)$$

where $\gamma = \partial J_e/\partial(\Delta T)$ at $\Delta T = \Delta T_{max}$ and $J_e = (\beta+1)J_b$ is the emitter current density. With δ = 1.57 mV/K, $\beta(T_0) = 50$ and the applied V_{be} giving a zero-junction-temperature-rise collector current density of 2×10^4 A/cm^2, Fig. 2 shows the ΔT_{max} as a function of b (=

$q\delta/n_b kT_0$) with ΔE and ρ_e as parameters. ΔT_{max} is a characteristic quantity associated with the onset of thermal runaway and the device power handling capability under the voltage modulation mode. It can be increased by decreasing b (resulting from a combination of higher n_b, or lowering δ), and increasing ΔE and ρ_e.

A.2. Current Modulation Mode: When the device is under the current modulation mode of operation, the dependence of p on V_{ce} and the applied constant base current I_b, is given by the following expression:

$$C_0' = \ln(p) + bc[1+\beta_0^{-1} e^{\frac{-a}{T_0+r_{th}p}}]\frac{p}{V_{ce}} \quad (6)$$

$$- \frac{a}{T_0+r_{th}p}$$

$$C_0' = \ln(\beta_0 V_{ce} I_b)$$

V_{ce} in (6) can be proved to have no upper bound for a real solution of p. Therefore, a single emitter device is unconditionally thermally stable under the current modulation operation. This result is valid only when the electro-thermal model described by equations (1)-(3) is applicable. Apparently, the conclusion drawn here is no longer supported, if the temperature in the device exceeds the semiconductor intrinsic temperature, where the transistor action ceases.

B. Two Emitter Element Case:

In this case, the device is assumed to consist of two emitter fingers with each having an unit area (two emitter elements). The junction temperature is assumed to be constant for each emitter element. If p_1 and p_2 represent the heat dissipation by the two emitter elements, respectively, the corresponding temperature rises at the junctions are:

$$\begin{aligned} \Delta T_1 &= r_{11}p_1 + r_{12}p_2 \\ \Delta T_2 &= r_{21}p_1 + r_{22}p_2 \end{aligned} \quad (7)$$

where r_{11} and r_{22} are the thermal resistance due to self-induced temperature rise at element 1 and 2, respectively, r_{12} and r_{21} are the coupled thermal resistances between the two elements. In general, r_{ij} is the coupled thermal resistance characterizing the temperature increase at the i-th element due to the unit power density applied at the j-th element.

B.1. Voltage Modulation Mode: It can be shown that the thermal runaway of the device with two emitter elements behaves the same as those discussed in section A.1 for the device with single emitter element. Namely, $d\Delta T_{1,2}/dV_{ce} \rightarrow \infty$ as a result of thermal runaway.

B.2. Current Modulation Mode: When a device with two emitter elements is operated under the current modulation mode, the coupled nonlinear equations for p_1 and p_2 as functions of V_{ce} and the applied base current I_b are given by:

$$\ln(p_1) = C_0' + \frac{a}{T_0+r_{11}p_1+r_{12}p_2}$$
$$-\ln[1+e^{b(r_{22}p_2+r_{21}p_1-r_{11}p_1-r_{12}p_2)+bc(J_{e1}-J_{e2})}]$$

$$\ln(p_2) = C_0' + \frac{a}{T_0+r_{21}p_1+r_{22}p_2}$$
$$-\ln[1+e^{b(r_{11}p_1+r_{12}p_2-r_{21}p_1-r_{22}p_2)+bc(J_{e2}-J_{e1})}]$$

$$C_0' = \ln(\beta_0 V_{ce} I_b) \quad (8)$$

Using a symmetric case in which $r_{11} = r_{22} =$

$r_s = 2.2\times10^{-4}$ K-cm^2/W, $r_{12} = r_{21} = r_c$ and $\rho_e = 0$ for discussion, the trajectory of solutions on p_1-p_2 plan are shown in Fig. 3 for three different values of r_c. For the case of $r_c = 0$, the solution trajectory contains two branches. One is the straight line a, and the other is the arc c. For the case of $r_c = 1.1\times10^{-4}$ K-cm^2/W, the solution trajectory contains the same straight line a and the other quarter-circle-like branch b. For the case of $r_c = r_s$, only the straight line a is contained in the solution trajectory. It is not surprising to have the symmetric solutions for p_1 and p_2 for all the different r_c cases, since the two emitter elements are symmetric. The interesting feature is the quarter-circle-like branch which also appears as the solutions for the case of $r_c < r_s$. The physical situation of this feature is described as follows. When V_{ce} is small, p_1 and p_2 are small and equal to each other. As V_{ce} increases, p_1 and p_2 increase and remain equal to each other, until reaching a break point P (see Fig. 4), beyond which either p_1 continue to increase while p_2 decreases, or vice versa. This is the phenomenon of current instability, in which one element conducts most of the total current while the other becomes nearly inactive. It is one form of thermal runaway. It is different from the previous case of constant V_{be} operation, in which both elements conduct the same amount of current and show thermal runaway simultaneously.

Similar to ΔT_{max} in the voltage modulation case, the threshold temperature rise ΔT_{thrd} at the break point is the characteristic quantity associated with the thermal runaway of the device under the current modulation mode. Again, we use the same symmetric device example to illustrate the determination of ΔT_{thrd}. With a small perturbation of the power density away from the break point, the threshold power density p_{thrd} for the current instability occurrence can be calculated. The corresponding junction temperature rise, $\Delta T_{thrd} = (r_s+r_c)p_{thrd}$ can be calculated from the following equation:

$$\frac{1}{\Delta T_{thrd}} - b' + \frac{a'}{(T_0+\Delta T_{thrd})^2} + b'c\gamma = 0 \qquad (9)$$

where

$$b' = b\frac{r_s-r_c}{r_s+r_c}, \qquad a' = a\frac{r_s-r_c}{r_s+r_c}$$

and $\gamma = \partial J_{e1}/\partial(\Delta T_1) = \partial J_{e2}/\partial(\Delta T_2)$ at $\Delta T_1=\Delta T_2=\Delta T_{thrd}$, where J_{e1} and J_{e2} are the emitter current densities flowing into corresponding emitter elements. It is interesting to note that (10) has exactly the same form as (5) with the transformation of the coefficients a and b into a' and b'. Besides r_s and r_c, the discussion of the dependence of ΔT_{max} on the parameters b, ΔE, ρ_e, δ and the operating current density in Section A.1 is also applicable to ΔT_{thrd}. As long as r_s and r_c are known, the threshold junction temperature rise can be obtained using the graphical results shown in Fig. 3. It can be shown that r_c is always smaller than r_s, and r_s remains nearly constant while r_c decreases rapidly as the separation between the two emitter elements increases. For the case of two nearly thermally isolated elements (e.g., due to a large separation between elements), thermal runaway does occur in the current modulation mode. As the thermal coupling becomes larger, (e.g., through reducing the separation between elements, and r_c approaches r_s), ΔT_{thrd} becomes larger. When $r_c = r_s$, the emitter elements are completely thermally coupled, and the ΔT_{thrd} is infinity, which is essentially the same as the case of a single emitter element described in the previous section and the results shown in Fig. 3, where the solution trajectory for $r_c = r_s$ case contains no quarter-circle-like branch.

C. Multi-Emitter Case:

In this section, only the thermal runaway under the current modulation mode of operation is discussed. The emitter fingers on the conventional HBT chip were treated in the calculation as a composition of a number of small unit elements. In each unit element, the junction temperature is assumed to be constant. Assuming that the bottom surface is kept at the heatsink temperature T_0, the power density of the emitter is given by (3), and the remaining surface are adiabatic, the temperature distribution on the surface of the semiconductor chip is calculated by solving the 3-D heat transfer equation. An iterative numerical scheme used to obtain a self-consistent solution of the temperature and current distributions, and the common-emitter mode I-V characteristics in the linear region are described in detail in reference [7].

A schematic drawing of an HBT device with an array of 5×3 emitter elements is shown in Fig. 4. Each emitter element has an area of 2×2 μm^2. The separations between elements are 6 and 10 μm in the x- and y-directions, respectively. With the parameters of $n_b = 3$, $\delta = 1.57$ mV/K, $\beta(T_0) = 50$, $\Delta E = 0.07$ eV and $\rho_e = 0$, the I-V characteristics in the linear region are calculated and shown in Fig. 5. As a result of the thermally-induced current instability, a collector current crush phenomenon is observed beyond a certain power level. For the present case with the parameters used and $\rho_e = 0$, this power level is 110 mW, regardless of the base current values. Fig. 5 also shows the maximum temperature $\Delta T_{j,max}$ at the center emitter element and the average temperature $\Delta T_{j,ave}$ among all the emitter elements vs. V_{ce}. It is clearly shown that the current crush is coinciding with the rapid increase of $\Delta T_{j,max}$ and the decrease of $\Delta T_{j,ave}$. We define $\Delta T_{j,thrd}$ as the threshold of the junction temperature rise to be equal to $\Delta T_{j,ave}$ at the onset of current crush. In the present case, $\Delta T_{j,thrd}$ is 60 K, regardless of the base currents. This threshold value is fairly close to the graphical results shown in Fig. 2 after the correction factor due to the temperature-dependent thermal conductivity (this correction factor is discussed elsewhere). The temperature distributions on the device surface plan are displayed in Fig. 6(a), (b) and (c) for three different operating conditions. V_{ce} = 5, 8.5, and 9.5 V correspond to the operating points of below, at, and beyond the current crush point, respectively. Below the current crush, the temperature rise at the emitter element is rather uniform. At the current crush point, the temperature at the center element starts to increase more rapidly than the others. Beyond the current crush point, a localized hot spot is formed at the center element, and the temperature at the others decreases. The current crush is concurrent with the formation of hot spots through which most of the current is conducted.

The triggering of the current instability when the junction temperature rise reaches a threshold value can be qualitatively described as follows. In the emitter array case, the coupled thermal resistance between the two adjacent emitters is assumed small. As the power increases, the junction temperature increases almost uniformly at all the unit elements, although the center one has a slightly higher junction temperature due to all the coupled thermal resistances from the neighboring ones. The current instability occurs when the junction temperature rise at two adjacent elements reaches the threshold value. Beyond this threshold the one with a slightly higher junction temperature starts to take over the current from the other elements as the power dissipation is increased further. Once this

happens, the collector current starts to collapse.

Fig. 7 shows the I-V characteristics, $\Delta T_{j,max}$ and $\Delta T_{j,ave}$ vs. V_{ce} for the device with the same parameters as in Fig. 6, except now $\rho_e = 2\times10^{-6}$ Ω-cm^2 instead of 0. The $\Delta T_{j,thrd}$ are 80, 90 and 100 K, for I_b = 0.2, 0.4 and 0.6 mA, respectively. The threshold junction temperature rise in this case is larger than that of the previous $\rho_e = 0$ case. It is also higher when the device is operated at high current and low voltage rather than at low current and high voltage similar to the results obtained in section A.1. As the ballast resistance increases, the power handling capability increases, but the cut-off frequency decreases due to the extra charging time of the emitter junction [8]. There is a trade-off between power and frequency performance when ballast resistors are used.

III. DISCUSSION AND CONCLUSION

Since $\Delta T_{j,thrd} \approx r_{total} p_{thrd}$, improvement in the power handling capability is most directly obtainable through the reduction of r_{total}. Device fabrication techniques such as flip-chip configuration, peeled-film and AlGaAs/GaAs HBT built on Si are several means to achieve reduced total thermal resistance. The power threshold for the onset of the thermally-induced current instability is almost linearly proportional to the thermal conductivity of the substrate.

Another approach to increase the power handling capability of the device under current mode operation is to utilize the result obtained in the previous section, i.e. ΔT_{thrd} is infinity if $r_s = r_c$. One of the methods for obtaining a strong coupling between emitter elements (i.e., $r_c \rightarrow r_s$) is through reduced separation between emitter elements. However, this method increases the total thermal resistance (since $r_{total} \approx r_s + r_c$), and results in an inferior performance in both the dc and rf operations [7]. Another method to realize a strong coupling is through a thick metal air bridge, called thermal shunt, connecting all the emitter elements. The junction temperature difference between elements can be eliminated by the low thermal resistance path through the air bridge. The r_{total} can be made small through a careful layout design of the emitter elements. In addition, the air bridge can be posted to large pads which provide additional thermal paths to dissipate the generated heat during device operation which further reduces the total thermal resistance. The I-V characteristics of the device implemented thermal shunt as well as the conventional device with and without the ballast resistances are shown in Fig. 8. It is seen that the conventional device without ballast resistance has a current crush at lower power level, while the one with ballast resistance has a current hogging at a larger power level. The one with thermal shunt does not show sign of thermally-induced current instability.

REFERENCES

1. D. S. Whitefield, C. J. Wei, and J. C. M. Hwang, "Temperature-dependent large signal model of heterojunction bipolar transistors," Proceeding of IEEE GaAs IC Symposium, 1992.

2. J. R. Waldrop, K. C. Wang, and P. M. Asbeck, "Determination of Junction Temperature in AlGaAs/GaAs Heterojunction Bipolar Transistors by Electrical Measurement," IEEE Trans. on Electron Devices, vol. 39, pp. 1248-1250, 1992.

3. L. L. Liou, C. I. Huang and J. Ebel, "Numerical studies of thermal effects on

heterojunction bipolar transistor current-voltage characteristics using one-dimensional simulation," Solid-State Electronics Vol. 35, pp. 579-585, 1992.
4. G-B Gao, M-Z Wang, X. Gui and H. Morkoc, "Thermal design studies of high-power heterojunction bipolar transistors," IEEE Trans. on Electron Devices, vol. ED-36, pp. 854-862, 1989.
5. R. H. Winkler, "Thermal properties of high-power transistors," IEEE Trans. on Electron Devices, Vol. ED-14, pp. 260-264, 1967.
6. F. Bergmann and D. Gerstner, "Some New Aspects of Thermal Instability of the Current Distribution in Power Transistors," IEEE Trans. on Electron Devices, vol. ED-13, pp. 630, 1966.
7. L. L. Liou and B. Bayraktaroglu, "Thermal Stability Analysis of AlGaAs/GaAs Heterojunction Bipolar Transistors with Multiple Emitter Fingers," Submitted to IEEE Trans. on Electron Devices, 1993.
8. L. L. Liou, B. Bayraktaroglu and C. I. Huang, "Thermal Stability Analysis of Multiple Emitter Finger Microwave AlGaAs/GaAs Heterojunction Bipolar Transistors," IEEE MTT-S International Microwave Symposium Digest, pp. 281-284, 1993.

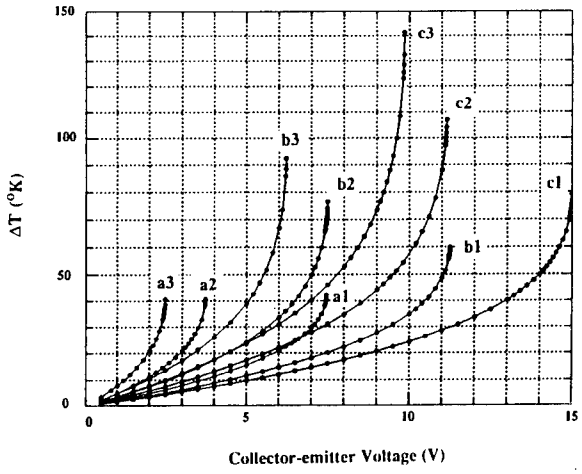


Fig. 1 The junction temperature rise ΔT as a function of V_{ce}, with V_{be} and ρ_e as the variable parameters. ρ_e = 0 (curves a1, a2 and a3), 1 (curves b1, b2 and b3) and 2 $\times 10^{-6}$ Ω-cm^2 (curves c1, c2 and c3), and the three different V_{be}'s, which give current densities of 1 (curves a1, b1 and c1), 2 (curves a2, b2 and c2) and 3 $\times 10^4$ (curves a3, b3 and c3) A/cm^2, respectively, when junction temperature is at the heat sink temperature. The other parameters used are n_b = 2, ΔE = 0.07 ev, δ = 1.57 mv/K, $\beta(T_0)$ = 50 and T_0 = 300 K.

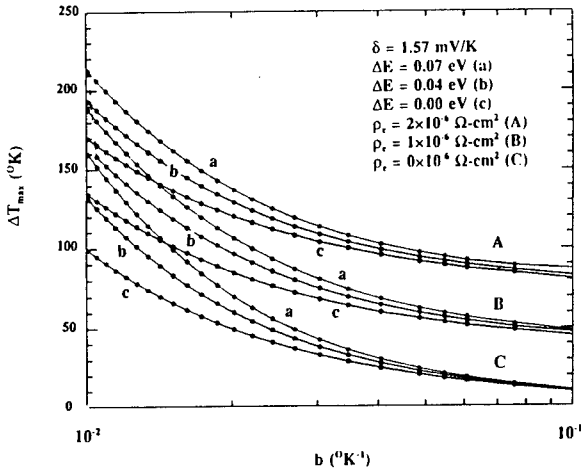


Fig. 2 ΔT_{max} as a function of b, with ΔE = 0.04, 0.07 and 0.1 eV, δ = 1.57 mV/K and ρ_e = 0, 1, and 2$\times 10^{-6}$ Ω-cm^2.

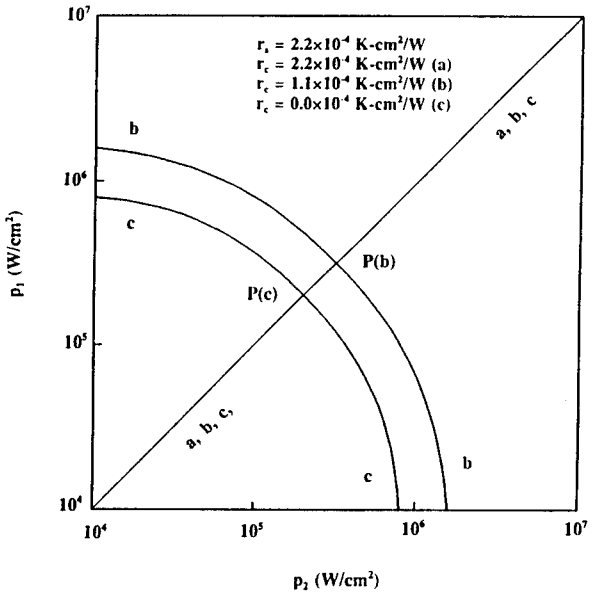


Fig. 3 The trajectory of the solutions of p_1 and p_2 for a two-emitter element with a symmetric thermal resistance. (See text for the explanation)

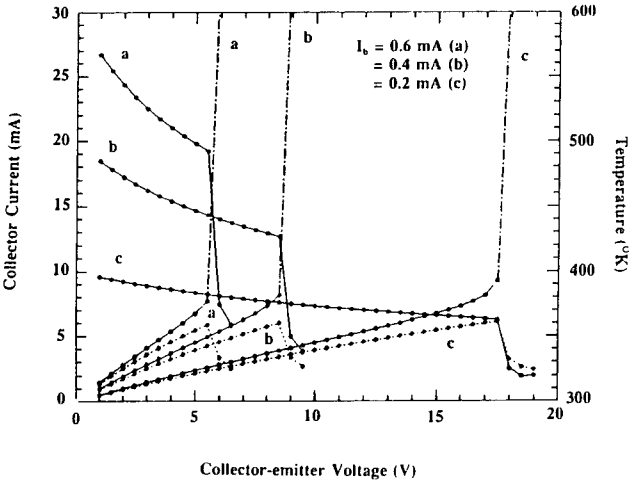


Fig. 5 The calculated I-V characteristics (solid line) in the linear region for I_b = 0.2, 0.4 and 0.6 mA. The parameters ΔE=0.07, ρ_e=0 and b=0.02/K are used. The maximum (dash-dot line) and the average temperature (dot line) among the emitter elements vs. V_{ce} are also shown.

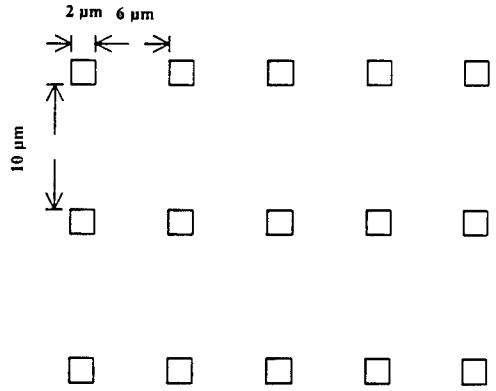
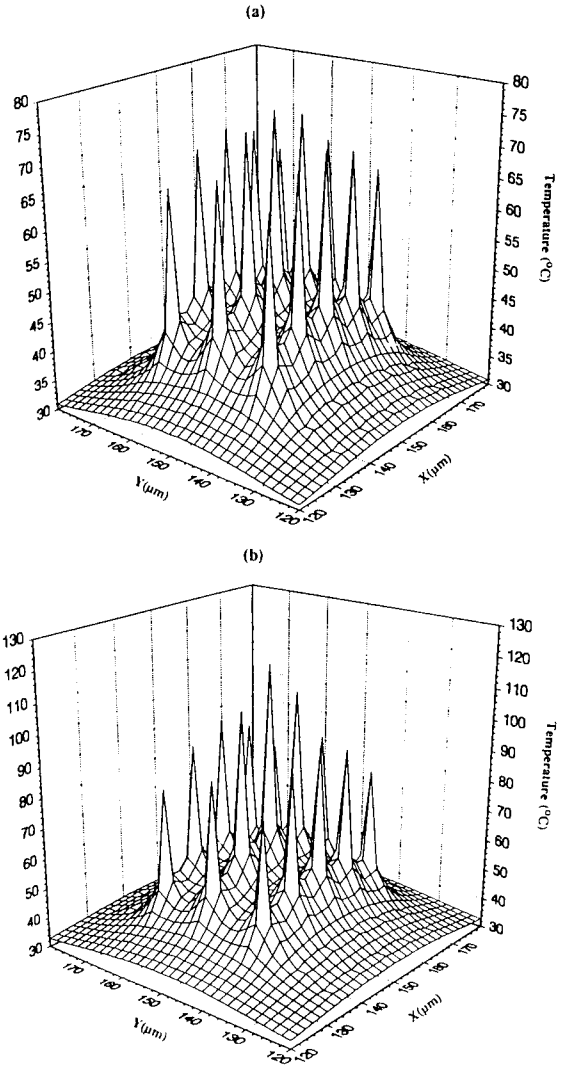


Fig. 4 The schematic diagram of the HBT with a 5×3 array of emitter elements on the top surface of the chip.



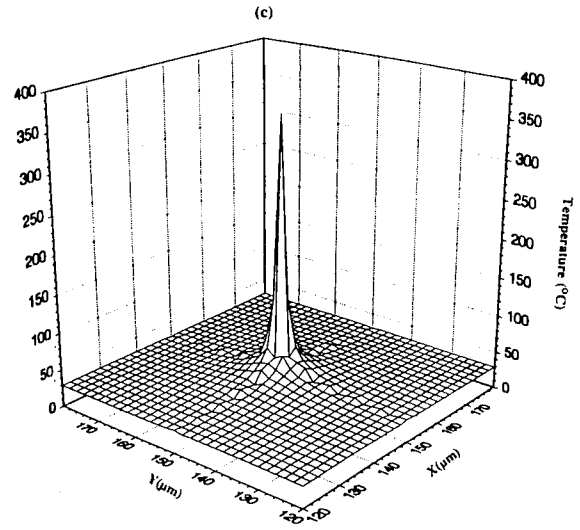


Fig. 6 The temperature distributions on the surface of the device operated at V_{ce} = 5 V (a), 8.5 V (b) and 9.5 V (c). Note the temperature scale is different.

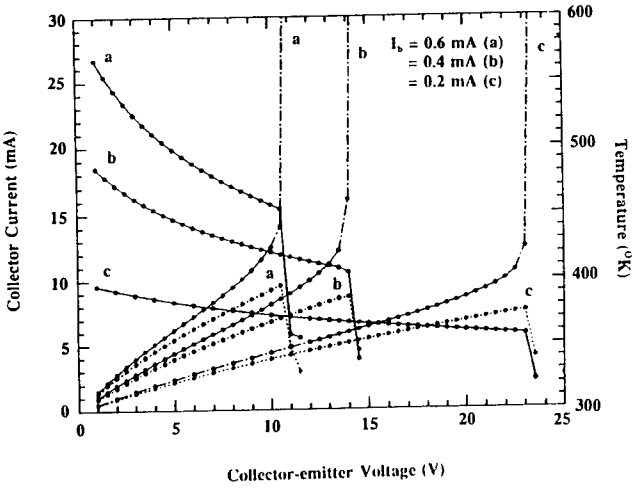
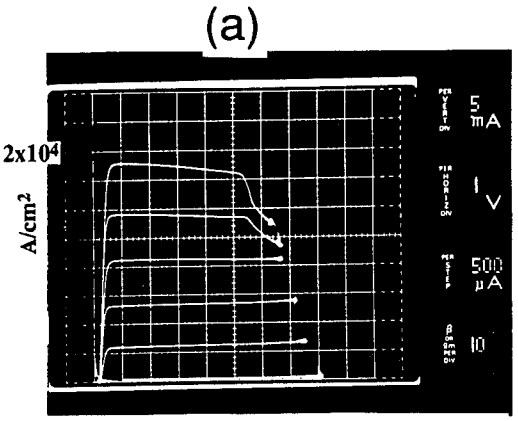
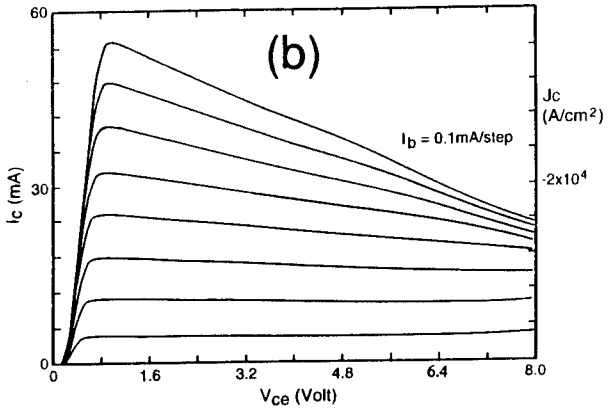


Fig. 7 The calculated I-V characteristics (solid line) in the linear region for I_b = 0.2, 0.4 and 0.6 mA. The parameters ΔE=0.07, ρ_e=2x10$^{-6}\Omega$-cm^2 and b=0.02/K are used. The maximum (dash-dot line) and the average temperature (dot line) among the emitter elements vs. V_{ce} are also shown.





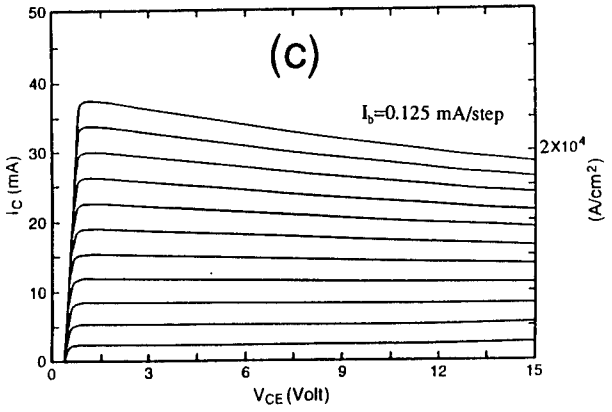


Fig. 8 The experimental I-V characteristics of a conventional AlGaAs/GaAs HBT devices (a) without any ballast resistor, (b) with an emitter ballast resistor and (c) with a thermal shunt.

Quasi-Optical Frequency Conversion Arrays

R. A. York and R. J. Ram
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106

ABSTRACT

A preliminary investigation of quasi-optical frequency conversion arrays is described. Using a diode-loaded dipole placed between two parallel conducting plates to simulate an array, we were able to demonstrate enhanced conversion efficiency into the second harmonic and obtain spontaneous parametric oscillation for certain array geometries. Sub-harmonic parametric oscillation was achieved by strongly driving the nonlinear antenna at a single pump frequency while the cavity length was adjusted for resonance at a subharmonic, and a well-defined threshold for this parametric down-conversion was observed.

I. INTRODUCTION

Quasi-optical power-combining is emerging as an attractive solution to the problem of limited power-handling capacity in millimeter-wave devices [1–3]. This paper proposes the application of these concepts to the realization of high-power frequency converters and parametric amplifiers. Some recent experiments at UCSB suggest that this can be accomplished using an array of coherently-pumped planar antennas, each containing a nonlinear parametric device. Impedance matching and frequency selectivity is achieved through antenna design and exploitation of the inherent mutual coupling between the antennas.

The preliminary experiments used a simple model consisting of a nonlinear dipole placed between parallel conducting plates (a Fabry-Perot cavity). According to image theory, the conducting plates simulate the response of the dipole in a large linear array (see figure 1). This not only simplifies the experiments, but also provides us with the following interpretation of mutual coupling in antenna arrays: the mutual coupling modifies the impedance of an antenna in a manner which reproduces the cavity mode structure, thus any antenna array is inherently frequency selective. If the antenna contains a nonlinear device which emits radiation at several frequencies, the filtering properties of the array are expected to have a profound influence on the output spectrum of the array.

II. INFLUENCE OF MUTUAL COUPLING ON EMISSION SPECTRA

First consider the behavior of the linear dipole driving-point impedance in the presence of the cavity. The mutual coupling that is introduced by the image elements can either enhance or inhibit emission at selected frequencies. For example, if the dipole is situated at the null of the standing cavity field, the radiation destructively interferes with image emission and the radiation resistance will decrease. Likewise, the radiation resistance is larger if the dipole is situated at the antinode of the cavity field, where there is constructive interference with image emission. The cavity length and dipole placement can therefore be tuned to selectively enhance emission at a certain frequency. This variation in radiation resistance can be quantitatively modelled using the induced EMF method from antenna theory [3,6]. It should be noted that the number of image dipoles in the equivalent

0-7803-0894-8/93 $3.00

image array is related to the cavity losses; an approximate relationship has been derived as [3]

$$N \approx \frac{Q}{2d/\lambda}$$

where N is the number of array elements, Q is the cavity Q-factor, and d is the element (cavity) spacing.

The influence of mutual coupling on the spectral emission of a nonlinear array element was explored experimentally by embedding a p-n diode (the base-emitter junction of an NE6448 microwave BJT, with $f_\tau \approx 10\,\mathrm{GHz}$) at the feedpoint of the dipole used in the previous experiments. This particular device was chosen for convenience and not because of any desirable properties for this application. The experimental setup is shown in figure 2. The antenna was driven by an HP8350 sweeper, typically at frequencies between 8 and 9 GHz. The output of the HP8350 was fed through a bandpass filter with a bandwidth of 4 GHz centered at 8 GHz, and out-of-band rejection of 60 dB. The filter removes any harmonic distortion in the sweeper output as well as isolating the sweeper from the harmonics generated in the antenna. The spectral response of the antenna was observed using a waveguide horn and an HP8563 spectrum analyzer, with the horn at the broadside position to the image array. A double-stub tuner was used to provide a small-signal impedance match to the nonlinear device.

For a single element in free space the dipole radiates at the pump frequency and harmonics. However, the relative power in the second harmonic was observed to change dramatically when the element was placed within the cavity [3]. For example, if the cavity is tuned so that its fundamental mode is at the second harmonic, the pump frequency is below cut-off and subsequently the power radiated into the second harmonic frequency increases and the dipole radiation at the pump frequency decreases.

The array can also be used to enhance emission at frequencies that are not emitted by the isolated element, if the device is pumped strongly. For certain values of the cavity spacing and pump power, we observed dipole emission at the *sub-harmonic* and corresponding mixing products, as shown in figure 3a. This is a result of spontaneous parametric down-conversion in our setup, which can be explained as follows. When a nonlinear reactance is modulated at a pump frequency ω_p, it can easily be shown [7] that a negative resistance is induced at the pump subharmonic, $\omega_p/2$. This is the basis for degenerate parametric amplification [5]. If the embedding circuit satisfies the oscillation conditions at this frequency, spontaneous down-conversion will occur. Interestingly, we consistently observed this phenomenon for cavity spacings below cutoff for both the fundamental and sub-harmonic frequencies. In some cases we were also able to observe non-degenerate parametric oscillation, as shown in figure 3b. In this case one of the lower frequency components can be thought of as the "idler" signal in parametric amplifier terminology [5]. At no time did we observe parametric oscillation without the cavity present.

The driven parametric device can be considered as providing gain, with positive feedback supplied by the dipole-cavity interaction. If the array elements are spaced to avoid such positive feedback, a low-noise negative resistance parametric amplifier or frequency up-converter (with gain) can be created. The magnitude of the induced negative resistance is proportional to the pump power, suggesting that there should be a threshold for parametric down-conversion. This was verified by measurement of the sub-harmonic power versus pump power, shown in figure 4. For pump powers below threshold the subharmonic component of the dipole response is indistinguishable from the

noise floor. At approximately 35 mW, a sharp increase in sub-harmonic output is observed. As pump power continues to increase, the slope of the curve changes due to the change in diode impedance. When the pump power is very large, the negative resistance saturates and sub-harmonic power levels off. The maximum power could be slightly increased by adjusting the diode bias and double-stub tuner for a better impedance match under large-signal conditions.

III. Conclusions

To the extent that mutual coupling can be reproducibly controlled in a practical array, it may be possible to realize high power frequency converters and parametric amplifiers using large arrays of nonlinear devices, by adjusting the mutual coupling to provide a good impedance match and the desired frequency response. Each array element would have its own pump source, with all the pump sources operating coherently. This could be achieved in several ways, such as through a traditional array feed network, injecting the pump source quasi-optically, or using a coupled-oscillator array as shown in figure 5. In the latter case, coupling between the oscillators allows them to synchronize through the phenomenon of injection-locking [9].

Although the experiments presented in this paper simulated an end-fire array, the results should extend to broadside arrays as well. Conceptually, this is equivalent to placing our dipole between two magnetic conducting planes, which has a similar mode structure to the electric conducting cavity. Either case can be realized in a planar array by controlling the relative phasing of the pump sources. In the case of a coupled-oscillator pump array (figure 5), the relative phasing could be adjusted using recently developed scanning techniques [10].

The observed spontaneous down-conversion could provide a useful technique for frequency division or generating sub-harmonic mixing products. However, we consider the observation of parametric oscillations to be significant only as an indication that other parametric effects could be exploited in arrays. These include negative resistance parametric amplification and parametric up-conversion [5], where the regime of parametric oscillations would necessarily be avoided. Such possibilities are intriguing—parametric amplifiers have not historically been used as power amplifiers, despite low-noise and high efficiency, because all the power in the system would converge on a single small-junction diode. Since the junction area of the device must be kept small for impedance matching and reduced parasitics, there are fundamental limitations on the power-handling capacity of a single device amplifier. By combining a large number of parameric amplifiers in a quasi-optical array, we can take advantage of this high efficiency and high output powers while still keeping the power requirements of a single device to a reasonable level. Since the signals are coupled to free-space, there is very little parasitic loss in the system, which is important for noise reduction.

Additionally, the degenerate negative-resistance parametric amplifier is known to possess a *phase-sensitive gain* [5]. When the pump signal is fully coherent with the input frequency, the maximum gain of the amplifier can be as much as 6 dB greater than for a non-coherent pump source. Therefore if the pump phase can be made to track the incoming signal phase, the signal-to-noise ratio can be improved by up to 6 dB since non-coherent noise would receive 6 dB less gain. This same principle has recently been demonstrated in parametric laser systems [8], and could be achieved by making the degenerate parametric amplifier part of a phase-locked loop.

IV. REFERENCES

[1] R. A. York and R. C. Compton, "Quasi-Optical Power-Combining using Mutually Synchronized Oscillator Arrays", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 1000-1009, June 1991.

[2] Z. B. Popovic, R. M. Weikle, M. Kim, and D. B. Rutledge, "A 100-MESFET Planar Grid Oscillator", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-39, pp. 193-200, Feb 1991.

[3] M. Kim, J. J. Rosenberg, R. P. Smith, R. M. Weikle, J. B. Hacker, M. P. DeLisio, and D. B. Rutledge, "A Grid Amplifier", *IEEE Microwave & Guided Wave Lett.*, vol. 1, pp. 322–324, Nov 1991.

[4] R. J. Ram and R. A. York, "Mutual Coupling in Nonlinear Dipole Arrays", submitted to *IEEE Trans. Antennas Prop.*, June 1993.

[5] L. A. Blackwell and K. L. Kotzebue, *Semiconductor-Diode Parametric Amplifiers*, Prentice-Hall: New Jersey, 1961.

[6] C. A. Balanis, *Antenna Theory: Analysis and Design*, Harper & Row, 1984.

[7] A. Yariv, *Quantum Electronics, 3rd ed.*, Wiley: New York, 1989, Chapter 17.

[8] L. Wu, H. J. Kimble, J. L. Hall, and H. Wu, "Generation of squeezed states by parametric down-conversion", *Phys. Rev. Lett*, vol. 57, p. 2520, 1986.

[9] K. Kurokawa, "Injection-Locking of Solid-State Microwave Oscillators", *Proc. IEEE*, vol. 61, pp. 1386-1409, Oct 1973.

[10] R. A. York and P. Liao, "A New Phase-Shifterless Beam-Scanning Technique using Arrays of Coupled Oscillators", to appear in *IEEE Trans. Microwave Theory Tech., Special Issue on Quasi-Optical Techniques*, October 1993.

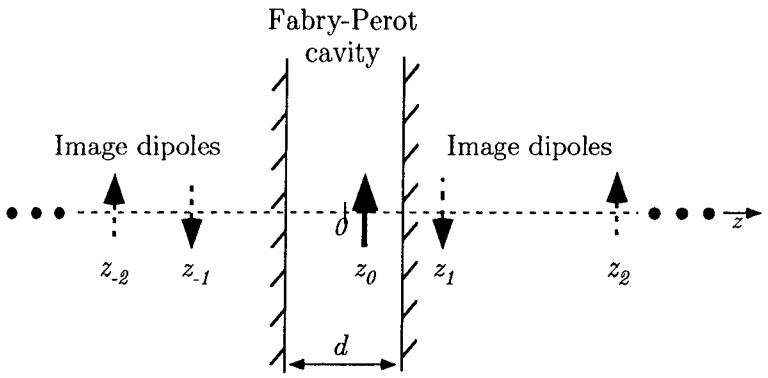


Figure 1 - Geometry for the dipole in a parallel-plate cavity, and the equivalent image array. Arrows denote the direction of current flow in the dipoles

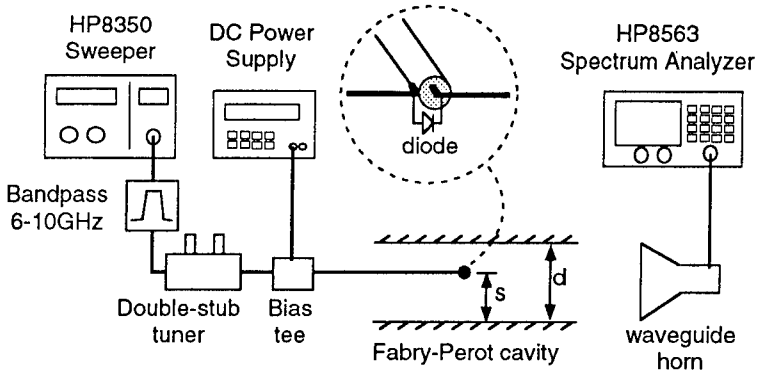


Figure 2 - Illustration of the experimental setup for observing the nonlinear dipole-cavity response. The HP8350 source plug-in was capable of delivering nearly 1 Watt at 8 GHz. The double-stub tuner was initially adjusted for a good match at small-signal drive levels using an HP8720 network analyzer.

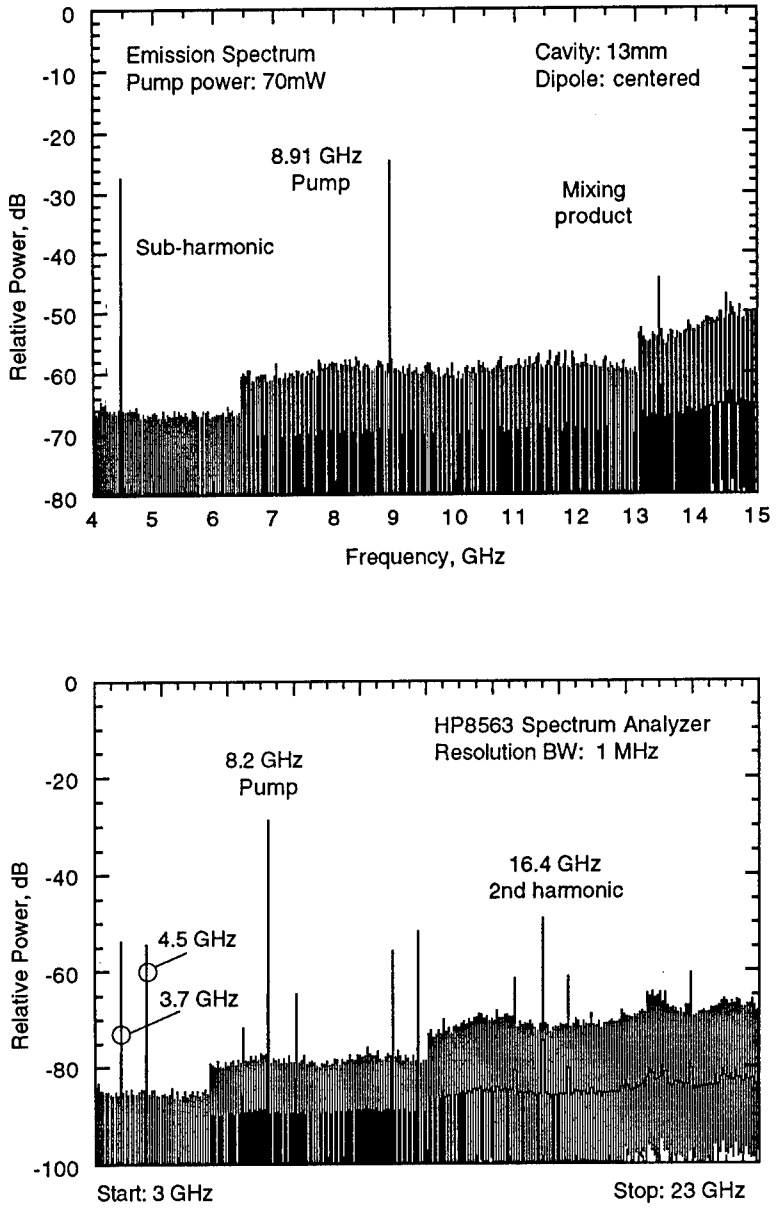


Figure 3 - Top: Dipole emission spectra under large-signal pump conditions, showing spontaneous parametric down-conversion (and mixing products) with cavity adjusted for resonance at sub-harmonic. Bottom: non-degenerate down-conversion was also observed with different cavity spacing and dipole placement.

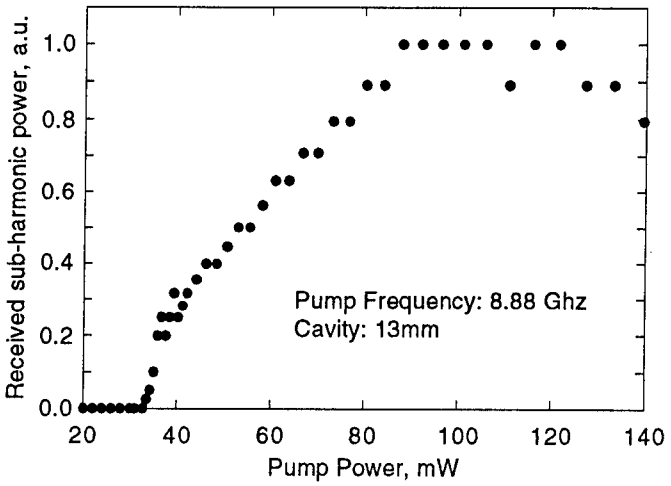


Figure 4 - Sub-harmonic power generated versus pump power. Above a threshold of approximately 35 mW, sub-harmonic power increases rapidly, then falls off as diode impedance changes under large-signal conditions.

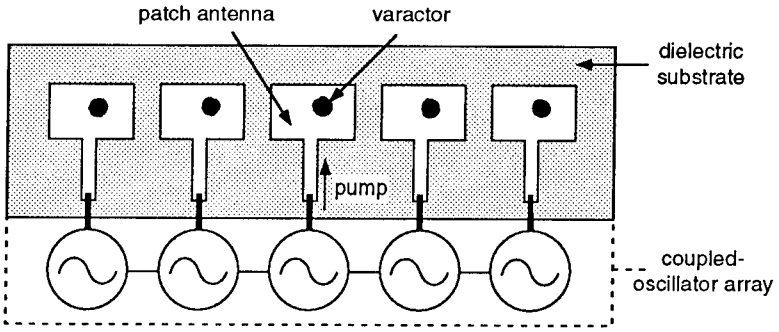


Figure 5 - Illustration of a quasi-optical frequency multiplier or parametric amplifier array. Varactor diodes embedded in patch antennas are driven by a local coupled-oscillator array. Antenna spacing, diode placement, and oscillator phasing determines the emission spectra.

Broadband Optoelectronic Wafer Probing

M. D. Feuer, S. C. Shunk, P. R. Smith[1], H. H. Law[1],
C. A. Burrus[2], and M. C. Nuss

AT&T Bell Labs,101 Crawfords Corner Rd., Holmdel, NJ 07733-3030
[1]AT&T Bell Labs,600 Mountain Ave., Murray Hill, NJ 07974-0636
[2]AT&T Bell Labs, 791 Holmdel-Keyport Rd., Holmdel, NJ 07733-0400

ABSTRACT

Optoelectronic S-parameter measurements offer the bandwidth needed to characterize today's state-of-the-art transistors, but have not yet achieved the throughput or accuracy provided by a vector network analyzer with microwave probes. In this paper, we discuss a new approach in which movable optoelectronic probes, calibrated by testing simple standard devices, are stepped around the wafer to provide accurate, high-throughput S-parameter measurements. Sub-picosecond laser pulses drive photoconductive switches on the probe tips, to generate electrical stimulus pulses and define sampling intervals, and signals are transferred to and from the wafer under test by coplanar waveguide transmission lines and plated contact bumps. The probes provide electrical pulses as short as 3 psec (FWHM), while maintaining a broadband 50 ohm termination to ensure stability of the device under test. Since probe flexure under contact significantly disturbs alignment of free-space beams, fiber-optic input is used to improve reproducibility. Analysis by vector error correction in the frequency domain removes systematic errors and separates the incident and reflected pulses without subjective time-window gating. We have demonstrated precise measurement of the complex reflection coefficient S_{11} at frequencies up to 175 GHz. Noise simulations have been performed to investigate the effect of various system parameters on the measurement uncertainty and useful bandwidth for Sparameter tests.

INTRODUCTION

Wafer probing at high speed has revolutionized research, development, and production of high-speed devices. Broadband, error-corrected S-parameter measurements up to 60 GHz are routinely obtained from vector network analyzers connected to microwave wafer probes based on transitions from coaxial lines to coplanar waveguide (CPW). However, the extrapolated performance of the best state-of-the-art transistors exceeds 300 GHz, presenting a large measurement gap. Waveguide-based probes are available to cover V-band (50-75GHz) and W-band (75-110 GHz), but the cost and complexity of multi-band waveguide systems are very high.

Optoelectronic(OE) testing driven by picosecond or femtosecond laser pulses offers a broad-bandwidth alternative to conventional, purely electronic S-parameter measurements. Typical OE setups use wirebonds or beam leads to connect external photoconductors to the device under test (D.U.T.)[1-4], or use integral photoconductors and transmission lines fabricated as part of the test wafer [5]. With the former approach, wafers cannot be measured whole, throughput is very limited, and vector error correction (accuracy enhancement) is not practical. The latter approach requires additional processing steps, may not be compatible with all substrates, consumes space on the test wafer, and relies heavily on the uniformity of the photoconductive gaps and calibration standards fabricated on each wafer.

To overcome these limitations, we have put the OE conversions onto a probe tip which can be moved to contact various sites on a wafer, as sketched in Figure 1a. Besides the obvious benefit to

0-7803-0894-8/93 $3.00

throughput, this approach can be used on whole or part wafers of any material, and permits vector error correction. Recalibration of the system is not needed each time the probe is moved to a new device. Our apparatus is similar to that of Scheuermann et al [6], but we have added fiber optic input, frequency domain analysis, error correction and independent bias of the D.U.T. and the stimulus gap. We have also used CPW-based probe layouts, for compatibility with devices designed for conventional microwave probing.

Error correction, based on the measurement of known standard devices, is critical to accurate high-speed vector measurements. Microwave network analyzers depend on it to control the reflections, loss, dispersion and delay due to connectors, discontinuities, aging, and simply the non-ideality of transmission lines, directional couplers and samplers. Although some of these sources of measurement error can be reduced by the simplicity and compact size of an OE test set, they are not eliminated. Also, error correction holds an added advantage for OE systems: it can provide directionality for the test set, eliminating the need for time gating or directional couplers, as discussed below.

FABRICATION and MEASUREMENTS

The probe tips were fabricated from commercial silicon-on-sapphire wafers which were 250 μm thick and had 0.6 μm of undoped epitaxial Si. CPW lines were defined in 0.5 μm gold films by photolithography, and the Si epi was removed from the open spaces by reactive ion etching. The epilayer was left in the photoconductive gaps, which measured 5 by 30 μm, and its carrier lifetimes were reduced into the sub-picosecond range by damage from an implant of $2x10^{14}$ Si^+ ions at an energy of 300 keV.

The central CPW line comprised a 78 μm center conductor and 39 μm spaces, and stimulus and sampling gaps were placed symmetrically, approximately 1mm from the tip, as shown in Figure 1b). Bias to the gaps was provided by CPW lines similar to the signal line, and both bias and signal lines were terminated in 50 ohm coaxial connectors to minimize back reflections. The three contact bumps were lithographically defined, with center-to-center spacing of 150 μm and typical sizes of 70 μm wide by 320 μm long by 20 μm thick, comprising 15 μm of nickel with a 5 μm gold overlayer for corrosion protection and low contact resistance. The contact bumps had a height deviation of <3 μm. For some measurements, ground-bridging straps were added to the basic probes to suppress the antisymmetric mode of the CPW line. The straps were sections of gold-mesh tape, typically 100 μm wide, attached with conductive epoxy.

For the fiber-input probe, single-mode fibers with cleaved ends were bonded to the top (back) surface of the sapphire probe. Since the fibers were not lensed, optical coupling to the photoconductive gaps was not optimally efficient, but ample power was available from the mode-locked Ti-sapphire laser source, and the gap separation of 80 μm provided reasonably low crosstalk.

Optical pulses at a wavelength of 760-800 nm were generated by a commercial Ar-pumped, mode-locked Ti-sapphire laser. Pulses from this system are typically 100 femtoseconds long, with a repetition rate of 82 MHz and an average power of ~1 Watt. For the free-beam-input experiments, the optical power was reduced to 25 mW per gap, while the power coupled into each fiber for the fiber-input probe was ~10 mW. A programmable translation stage was used to sweep the relative delay of the sampling pulses, and the stimulus beam was mechanically chopped to enable lock-in detection, eliminating signal due to dark current leakage in the stimulus gap.

The general methods of photoconductive stimulus/sampling have been discussed by Auston[7]. In brief, we applied a 10 Volt dc bias to the stimulus gap and monitored the average current from the sampling gap under zero bias. Optical pulses striking the stimulus gap launch electrical pulses along the line, which propagate to the D.U.T. and are partly reflected. When the delayed optical pulse strikes the sampling gap, current flow is proportional to the voltage present at that point on the signal line at that

instant, whether due to the incident or reflected pulse. The average sampled current is monitored as a function of delay time to obtain V(t) curves.

ANALYSIS & ERROR CORRECTION

Although time-domain responses can be useful in their own right, many design applications require accurate S-parameter data in the frequency domain. In typical optoelectronic measurements [2-4], the incident and reflected pulses are separated by time-window gating, then Fourier transformed and ratioed to obtain the complex reflection coefficient S_{11}. A simple propagation factor is used to displace the reference plane to the D.U.T. This approach has some limitations. Reflections at wirebonds or probe contact points cannot be removed. The time gate window must be much longer than the input pulse, so the sampling point must be distant from the D.U.T., which increases the delay, dispersion and loss. Reflections from the back end of the test set must be excluded by a separate time gate, limiting the maximum sampling time. Frankel et. al. [4] have noted that this limited sampling time can interfere with measurements of active devices which have long time constants.

Vector error correction techniques can deal with all of these errors, and more, while eliminating the need for time gating. By measuring an open circuit, short circuit, and load device with accurately known properties, one can establish three reference points in the S_{11} plane, to uniquely define the conformal transformation associated with the errors and non-idealities of the test set [8,9]. The calibration removes all linear errors which are local in frequency and repeatable from run to run, even a directivity error of unity (no directional sensitivity at all). This is why time gating is not needed: the incident signal has been extracted from the measurement of the standard devices. The overall effect of error correction is to place the burden of accuracy on the standard devices, demanding only reproducibility from the test equipment.

The actual procedure we used for calculation of accurate S-parameters from the time-domain data was:

1) Normalize each V(t) trace to the peak of the incident pulse, to eliminate scaling errors due to slow drift or fluctuations in laser power or fiber launch efficiency.

2) Apply the Fourier transform to generate frequency-domain curves.

3) Divide by the Fourier transform of an assumed Gaussian incident pulse to generate raw (uncorrected) S-parameters. This is just a convenience for looking at intermediate results: the final, corrected result is entirely insensitive to the assumed pulse shape.

4) Use raw S-parameters of short circuit, open circuit, and resistive load standard devices to calculate the error correction terms.

5) Apply the error correction terms to the raw S-parameters of the Device Under Test (D.U.T.) to obtain the corrected S-parameters.

The calibration standard devices used for this work, and their definitions, were the same as those used for conventional microwave probing: the short circuit and load standards were on a commercial calibration tile, while the open standard was the probe suspended in air. The calibration tile is supplied by Cascade Microtech for operation up to 40 GHz, but the small size of the standard devices suggests that they should be useful to much higher frequencies.

EXPERIMENTAL RESULTS

Figure 2 shows the time-domain response of the free-beam-input probe with various terminations. The incident pulse, a cross-correlation of the electrical stimulus and sampling intervals, has roughly symmetric rise and fall times, with a width (FWHM) of 2.8 ps. The average stimulus current was

~2.1 μA, implying a pulse voltage of about 450 mV. Figs. 2a-c) show the responses for an open circuit (tip suspended in air), short circuit (tip contacting a bar-shaped gold metallization on the wafer under test), and a 50 ohm load standard (tip contacting the metallized pads of a thin-film resistor), respectively. The latter two terminations are on a calibration tile made for conventional microwave probes by Cascade Microtech. As expected, the open circuit shows a large positive reflection, the short circuit shows a large negative reflection, and the load device shows a much smaller residual reflection due to impedance and mode matching errors at the contact point. In each case, the main reflection signal has an asymmetric shape due to the summation of many components from the finite-length contact bumps and D.U.T., and small features due to multiple reflections and back reflections are noticeable out to ~120 ps.

To demonstrate the effectiveness of error correction in OE probing, we have measured the reflection from a 12.6-ohm thin-film resistor, and calculated S_{11} over the frequency range from dc to 100GHz. Since it is small and simple, the resistor should be nearly ideal, appearing at the point -0.60 + j0.0 on the Smith chart for all frequencies tested. Figure 3a) shows the S_{11} of the test resistor obtained by the time-gated method outlined above, with the time window for the incident pulse set from -32 ps to +4 ps, and the reference plane displacement set at 13.6 ps. The points are scattered loosely about the ideal value, with $S_{11} = -0.57\pm0.24 + j(-0.04\pm0.18)$ over the frequency range. The dc value is actually the wrong sign, clearly demonstrating the importance of systematic errors. Neither adjustment of the time window, nor inclusion of loss or monotonic dispersion in the propagation correction, can improve the result.

Vector error correction, however, makes a substantial improvement, as shown in Fig. 3b). The short, open, and load terminations shown in Figs. 2a-c) were used. Specifications of the short and load standards were those supplied with the calibration tile, while the open-circuit capacitance, which varies slightly with probe geometry, was set to -18 fF by inspection of the calibrated curves. The error-corrected S_{11} data show less than half the scatter observed when using the time-gating method, with $S_{11} = -0.60\pm0.08 + j(0.01\pm0.11)$. The remaining discrepancies are due to run-to-run variations which cannot be corrected out.

Fiber-optic input to the probe improves reproducibility, yielding the error-corrected measurements shown in Fig. 3c). A tight distribution of values is observed: $S_{11} = -0.63\pm0.04 + j(0.01\pm0.04)$ over the 100 GHz frequency range. The dispersion from 0.6 m of fiber in the optical path is not a problem, since the FWHM of the incident pulse remains as short as 3.4 ps.

The importance of mode conversion effects in CPW structures becomes evident when S_{11} is computed for the open stub device, as shown in Figure 4a). The magnitude of S_{11} for the open stub should begin at unity and decrease smoothly as losses increase with frequency, while the phase should begin at zero and decrease nearly linearly. Although the measured S_{11} shows qualitatively correct phase winding, the magnitude varies widely, including excursions above unity which imply gain, impossible for this passive device. The explanation is that the open stub actually reflects back more power to the sampling gap than does the probe open in air, because of mode conversion. When the laser pulse strikes the stimulus gap, the wave launched in the CPW line on the probe tip is concentrated on one side of the line, in a combination of symmetric and asymmetric modes which resembles a slotline mode. When the probe is open in air, both symmetric and asymmetric components are reflected with a positive polarity, sending the pulse back along the line on the same side as it arrived, and presenting a rather weak signal to the sampling gap on the other side. The open stub, however, has a ring-shaped ground plane which connects the two outer contact points of the probe and presents a short circuit to the asymmetric mode. This short circuit reflects the asymmetric mode with a negative polarity, causing the combined mode to shift to the opposite side of the line from that on which it arrived and presenting a strong signal to the sampling gap. Thus the open stub device with the ring ground has a stronger reflection than the open standard, yielding the erratic S_{11} curve shown in Fig. 4a).

The best way to avoid this problem is to put the ground-bridging connections on the probe tip itself, so that the asymmetric mode will be suppressed as the pulse is launched, and the signal at the contact points will be predominantly the desired symmetric CPW mode. The resulting probe tip can be used successfully with either ring-shaped or separate grounds, as shown in Fig. 4b). The addition of ground-bridging straps increases the incident pulsewidth somewhat, to a typical value of 5 psec, but the useful bandwidth of the corrected system remains high, as discussed below.

NOISE EFFECTS

Noise and irreproducibilities limit the useful frequency range as well as the precision of S-parameters measured with the optoelectronic probe. Since the error correction adjusts for the frequency rolloff of the stimulus and sampling gaps, accurate measurements can be made well above the the 3-db rolloff of the time-domain pulse, up to the point where signal power and noise power are comparable and the precision of the measurement becomes unacceptable. This is illustrated in Fig. 5, which shows the simulated effect of measurement noise on the precision of the real part of S_{11}, for an ideal 12.5-ohm resistor. The simulations assume that the V(t) curve comprises 512 sample points ranging from -50 to 150 psec, and that the incident pulse is an ideal Gaussian with FWHM = 4.8 psec (3-db rolloff = 94 GHz). Random white noise was superimposed on the ideal V(t) curves for each standard and for the D.U.T. measurement. After forty runs with independent noise, the standard deviation of $Re(S_{11})$, which we will call σ, was obtained. As shown in Fig. 5, σ rises with increasing frequency, and we define useful bandwidth by the arbitrary condition $\sigma = 0.05$. With the very low noise assumed for Fig. 5a), the useful bandwidth is 210 GHz, but the more realistic noise levels assumed for Fig. 5b) yield a useful bandwidth of 155 GHz, still well above the 3-db rolloff of the incident pulse.

Changes in the sampling conditions also affect the useful bandwidth in the presence of noise. Fig. 6a) shows the simulated results for exactly the same noise conditions assumed in Fig. 5b), except for a time span from -100 to +400 ps. The sampling interval, though doubled, is still 0.98 ps, suggesting that the sampling limits would not be observed up to 500 GHz. However, in the presence of noise, the useful bandwidth is reduced by 47 GHz This occurs because there are only half as many measurement points which fall within the incident and reflected pulse peaks (about five points per peak), while the number of measurement points which fall in the noise-dominated background regions is slightly increased. Fig. 6b) shows the effect of shorter input pulses, with the original time span and sampling interval of Fig. 5b). Uncertainty in S_{11} is increased at low frequencies, because of the statistical effect noted in 6a), but at higher frequencies, σ is decreased because there is more signal power present in the shorter pulse.

The use of two-port (12-term) error correction routines, instead of 1-port routines, also exacerbates the effect of noise. Simulation of a two-port device consisting of two separate 12.5 ohm resistors ($S_{12} = S_{21} = 0$), under the same operating conditions as for Fig. 5b), shows a low-frequency value of σ more than twice as large as for the one-port correction. Although *systematic* errors are removed by the error correction routines, *uncorrelated* noise in the standard measurements tends to add, and the two-port calibration involves four times as many standard measurements as the one-port routine.

Unfortunately, signal levels cannot be raised arbitrarily to improve signal/noise ratio. Since the object is to measure active devices in their linear (small-signal) regime, we must keep the peak level of the incident pulse at about -10 dbmW (electrical). In addition, not all sources of noise will simply add to V(t). Laser amplitude noise, in particular, will multiply both the incident pulse height and the sensitivity of the photoconductive sampler and will not necessarily be improved by larger signal levels. Still, the overall conclusion drawn from our simulations is that the presence of random noise at realistic levels does not prohibit measurements of useful precision over broad operating bandwidth.

Several factors have not been addressed in the present simulations. First, the actual noise is not white, but 1/f in character, so noise measurements should ideally be made over a time equal to that of a

full calibration and measurement cycle. Second, there may be run-to-run variations above and beyond noise, such as variations in the electrical contact resistance to the D.U.T. Finally, active devices may generate significant 1/f noise on their own. As these factors are clarified and addressed,we expect to achieve a level of precision which will more clearly reveal fine structure of the D.U.T. characteristics.

Figure 7 shows the measured magnitude of S_{11} for a 12.5-ohm thin-film resistor, similar to Fig. 3c). The only change has been careful adjustments of the Ti-sapphire laser and the measurement electronics to minimize noise. The resulting S-parameter data shows a scatter of just ±0.02 over a frequency range from 0 to 175 GHz.

SUMMARY

Moveable optoelectronic wafer probes can provide the bandwidth, throughput and precision needed to characterize state-of-the-art devices. Fiber-optic input and CPW-based electrical transmission enable reproducible measurements on any wafer suitable for microwave probing, while presenting a broadband 50 ohm impedance for stable operation of the device under test. We have constructed such probes, based on photoconductive switches for stimulus and sampling, and implemented vector error correction routines to ensure accuracy. When driven by a mode-locked Ti-sapphire laser, the probes demonstrated calibrated S_{11} measurements with a scatter of ±0.02 over a bandwidth of 175 GHz. Ground-bridging straps were shown to be effective in suppressing mode-conversion errors.

Noise simulations display the complex interaction between noise, sampling interval, and incident pulsewidth in determining the measurement uncertainty and useful bandwidth of the S-parameter system.

The authors wish to thank A. Borges, K.F. Brown-Goebeler, B. Tell, and J.A. Zembas for experimental contributions, and M. Pollack for continuing support.

REFERENCES

[1] P.R. Smith, D.H. Auston, and W.M. Augustyniak, "Measurement of GaAs Field-Effect Transistor Electronic Impulse Response by Picosecond Optical Electronics", **Appl. Phys. Lett.**, vol. 39, pp.739-741, 1981.

[2] D.E. Cooper and S.C. Moss, "Picosecond Optoelectronic Measurement of the High-Frequency Scattering Parameters of a GaAs FET," **IEEE J. Quantum Electronics**, vol. QE-22, pp. 94-100, 1986.

[3] M. Matloubian, S.E. Rosenbaum, H.R. Fetterman, and P.T. Greiling, "Wide-Band Millimeter Wave Characterization of Sub-0.2 Micrometer Gate-Length AlInAs/GaInAs HEMTs," **IEEE Microwave and Guided Wave Lett.**, vol. 1, pp. 32-34, 1991.

[4] M.Y. Frankel, J.F. Whitaker, and G.A. Mourou, "Optoelectronic Transient Characterization of Ultrafast Devices," **IEEE J. Quantum Electron.,** vol. QE-28, pp. 2313-2324, 1992.

[5] S.L. Huang, E.A. Chauchard, C.H. Lee, T.T. Lee, H-L. A. Hung, and T. Joseph, "On-Wafer Testing of MMIC with Monolithically Integrated Photoconductive Switches," in **IEEE MTT-S Int. Microwave Symp Dig.**, pp 661-664, 1992.

[6] M. Scheuermann, R. Sprik, J.-M. Halbout, P.A. Moskowitz, and M. Ketchen, "Ultra-High Bandwidth Detachable Optoelectronic Probes," in **OSA Proceedings on Picosecond Electronics and Optoelectronics,** vol. 4, ed. by T.C.L.G. Sollner and D.M. Bloom, (Optical Society of America, Washington, D.C. 1989), pp. 22-26.

[7] D.H. Auston, "Ultrafast Optoelectronics," in **Ultrashort Laser Pulses and Applications,** ed. by W. Kaiser, (Springer-Verlag, Berlin 1988).

[8] J. Fitzpatrick, "Error Models for Systems Measurement", **Microwave J.**, pp.63-66, May 1978.

[9] I. Kasa, "Closed-Form Mathematical Solutions to Some Network Analyzer Calibration Equations," **IEEE Trans. on Instr. and Meas.**, vol. IM-23, pp.399-402, 1974.

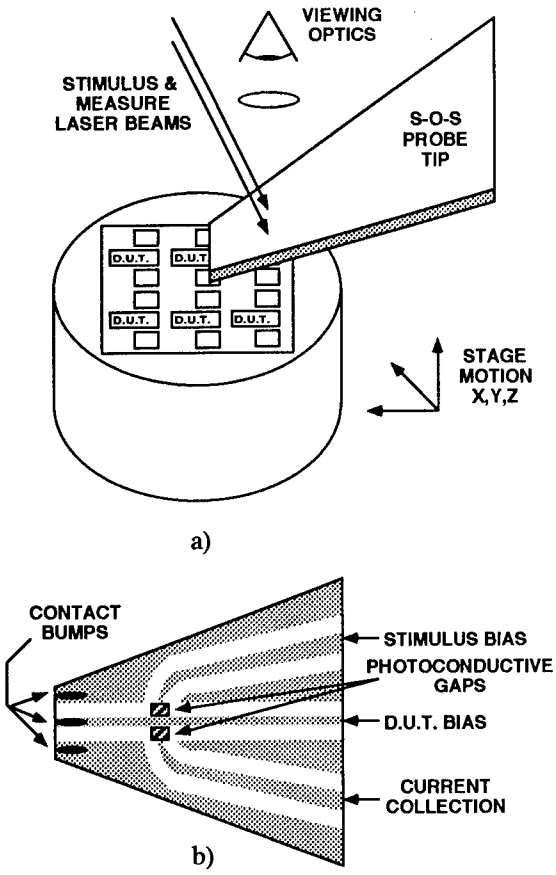


Figure 1. Schematic layout of optoelectronic wafer probe setup. The system provides high bandwidth, throughput and accuracy.

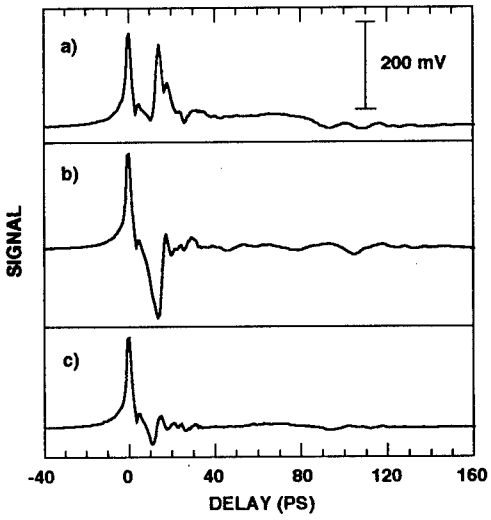


Figure 2. Time response of optoelectronic wafer probe with free-beam input, contacting different terminations. (a) Probe open in air; (b) Probe shorted on gold thin-film bar; and (c) Probe on 50-ohm calibration resistor.

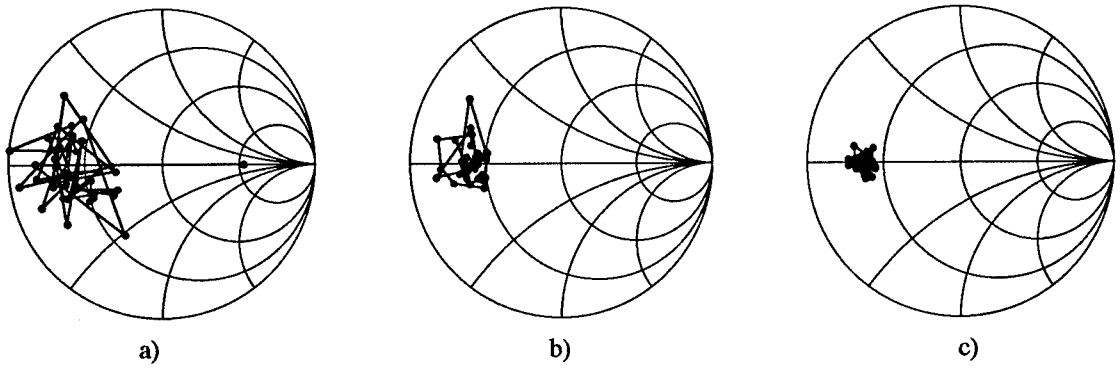


Figure 3. S11 of a thin-film resistor with a dc resistance of 12.6 ohm, from dc to 100 GHz at 2.5 GHz intervals. (a) Data from free-beam input probe, analyzed by conventional time-gating method, shows both systematic and irreproducible errors. (b) Data from free-beam input probe, analyzed by vector error correction, contains only the irreproducible errors. (c) Data from the fiber-input probe, analyzed by vector error correction, shows the best accuracy.

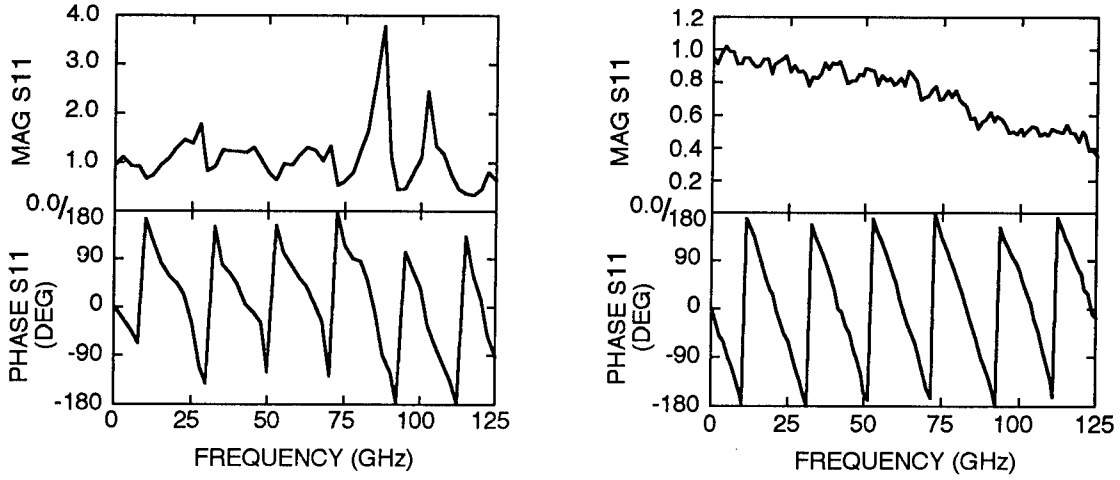


Figure 4. Magnitude and phase of S11 of 3.2-mm open stub transmission line. (a) Basic fiber probe without ground-bridging straps. (b) Fiber probe with gold mesh ground-bridging straps. Calibration standards and definitions are the same as for conventional microwave probing.

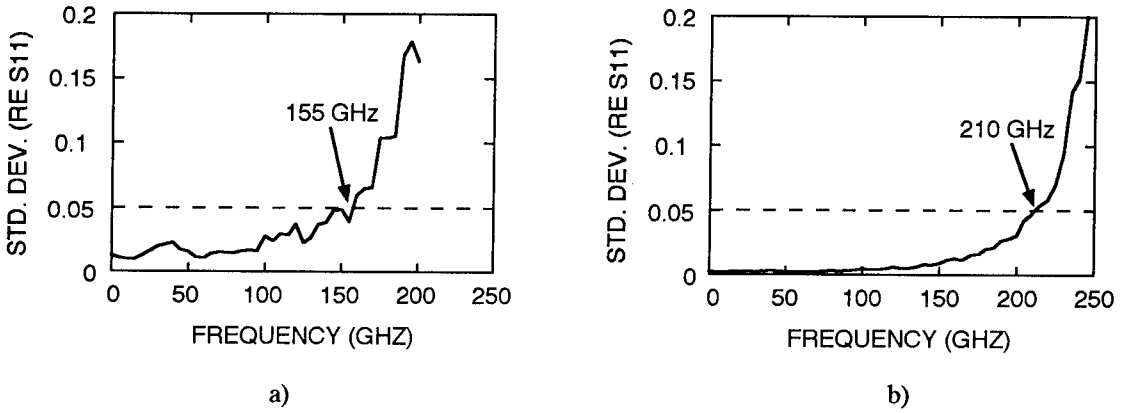


Figure 5. Simulated effects of random noise on S-parameter measurement. Standard deviation of the real part of S11 vs. frequency, for a 12.5-ohm resistor. The number of time samples is 512, the incident pulsewidth is 4.8 ps (FWHM), time sample interval is 0.39 ps. (a) Additive and multiplicative noise amplitudes are each 0.1% of incident pulse peak, resulting in a useful bandwidth of 210 GHz. (b) Additive noise amplitude is 0.5% of incident pulse peak, multiplicative noise amplitude is 1.0%, and the useful bandwidth is reduced to 155 GHz.

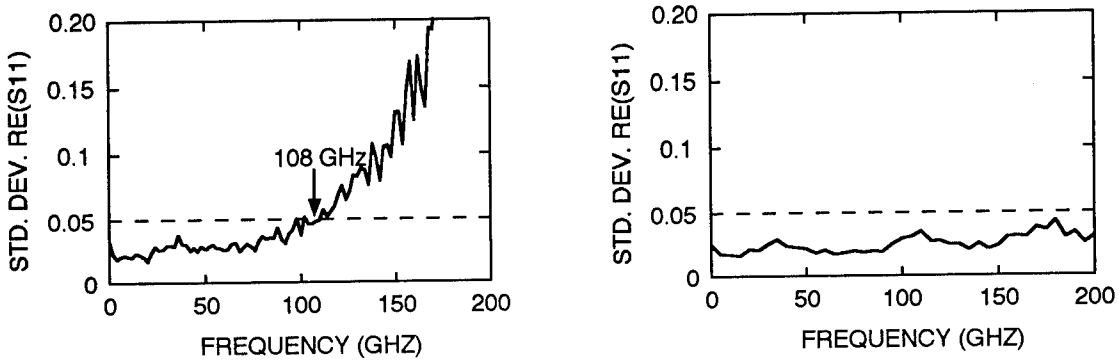


Figure 6. Simulated effects of random noise on S-parameter measurement. Parameters are the same as for Fig. 5b), except: a) time sample interval is increased to 0.98 ps; and b) incident pulsewidth is reduced to 2.4 ps (FWHM).

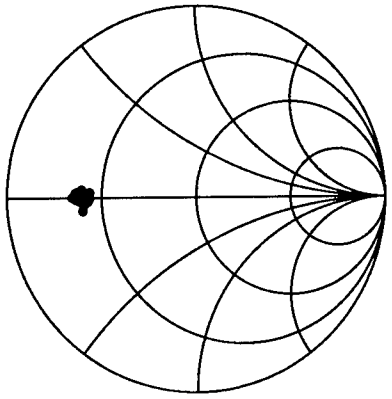


Figure 7. Measured S11 of a thin-film resistor with a dc resistance of 12.5 ohm, from dc to 175 GHz at 5 GHz intervals. Careful attention to noise sources has reduced to scatter in the data, while increasing useful bandwidth.

Two Layer Stepped-QW Channel HFETs on InP-Substrate

S. Strähle, B. Henle, E. Mittermeier, U. Erben, P.K. Rees°, E. Kohn

University of Ulm;
Department of Electron Devices and Circuits;
Albert-Einstein-Allee 45; D-89069 Ulm; Germany

°Epitaxial Products International Ltd. Cypress Drive, St. Mellons,
Cardiff, CF3 OEG, U.K.

Abstract

A novel HFET structure on InP containing an asymmetric two layer stepped QW channel sandwiched between two InAlAs barriers is proposed. A small bandgap InGaAs channel hosts the 2DEG-density and provides a low sheet resistance at low drain field. A higher bandgap sub-channel of (a) InP and (b) InGaAs/InAlAs superlattice is designed to act as high field drift layer. This approach allows to obtain high f_{max}/f_t-ratios and high output power at high speed. First results are given.

1. Introduction

InP based HFET devices have many advantages concerning their RF-performance compared to GaAs based HFET devices. Very high output power has been achieved with InP-channel FETs [1]. For InAlAs/InGaAs/InAlAs-HFET's grown by MBE as well as MOCVD very high cut-off frequencies and low noise performance have been demonstrated [2]. Therefore, it seems attractive to combine both features in a single structure and here a novel structure with a stepped Quantum Well (QW) channel of InGaAs and InP between InAlAs barriers is proposed. The idea of this structure design shall be illustrated considering the field-velocity curves for the materials used in the composed channel and barrier materials [3] (Fig. 1).

The InGaAs is the preferred channel material for the low field range since the electrons gain their maximum velocity already at low fields. For the transistor operation this means a low sheet resistance, low knee voltage and high f_t at low bias. However the velocity decreases sharply at high fields below $1*10^7$ cm/s. On the other hand, for the InP and AlInAs material the electrons reach their maximum velocity at high electric field. This material is therefore preferred to obtain a short transit time even at high drain bias. However the low mobility at low fields implies a large knee voltage. Combining both layers in one device represents an optimum condition to obtain simultaneously high f_{max} and f_t values.

In such a structure at low drain-source bias the 2-DEG-density introduced by modulation doping from the AlInAs supply layer is located in the InGaAs channel. Increasing the drain bias the hot electrons transfer into the second part of the channel, which is in this case InP. As the conduction band edge of InP is below that of the InAlAs barrier, the electrons are expected to be still partially confined to the stepped QW-channel and only at very high field

0-7803-0894-8/93 $3.00

they will spill out. Although to our knowledge the exact intervalley and real space transfer mechanisms have not been determined for such a structure. It is expected that the electrons transfer from the InGaAs Γ-center-valley into the InP Γ-center-valley and partially into the Γ-valley of the AlInAs. From there they might be transferred into the InP or InAlAs upper-valleys. To prevent noticable transfer into the InGaAs L-valley this layer is designed as a QW-channel as thin as possible. In other words, considering the velocity-field curves of Fig. 1 it is desired that the electrons transfer from the InGaAs velocity-field-curve onto the InP velocity-field-curve so that a large high-velocity, high-field-envelope-curve results.

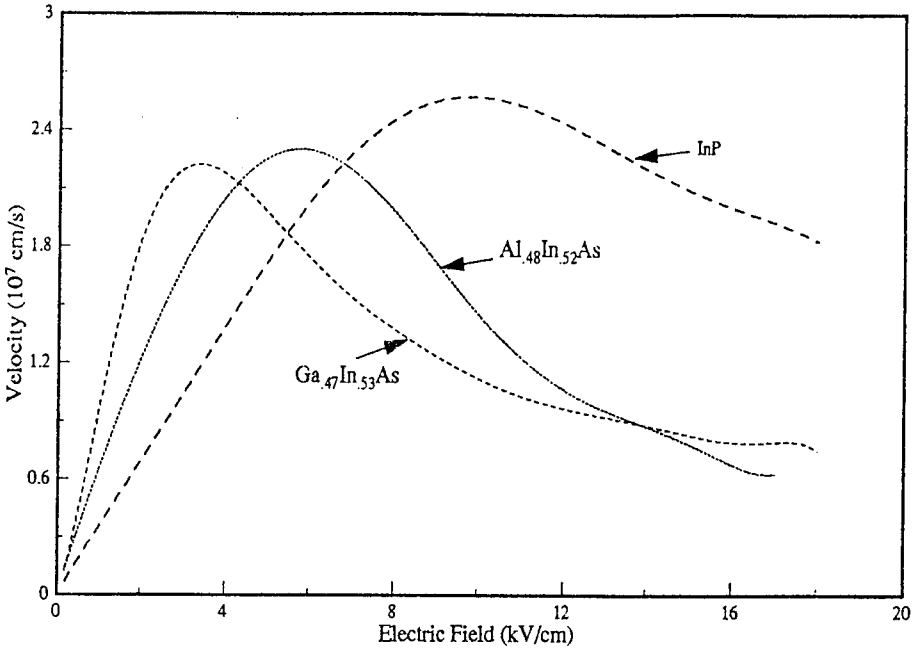


Fig. 1 Velocity Field Curves of the Materials used [3]

Usually, in a standard HFET structure on InP an InAlAs/InGaAs superlattice (SL) is inserted to smooth out the surface between the InAlAs-buffer layer and the InGaAs channel. In the light of the novel concept presented here this interfacial SL can also be considered and designed to be an active channel part. Therefore the alternative structure design uses an InGaAs/InAlAs superlattice instead of the InP-sub-channel with a similar effective conduction band edge and a effective thickness to that of InP. This phoshorous free structure design is very similar to the standard InGaAs-HEMT structure and allows to grow the samples with the well established Solid Source MBE technique and to use standard InP-HEMT processing technology.

2. Layer Design and Characterization

Several specific structures have been investigated for each type of design: (A1) and (A2) for the InP and (B) for SL-sub-channel. The InP containing channel structures (A1) and (A2) have been grown by MOCVD. For the AlInAs/InGaAs SL structures standard Solid Source MBE was used. All layers are grown on semi-insulating [100] substrate. Structure (A1) is lattice-matched to InP. The configuration is shown in Fig. 2. First an undoped buffer of 95 nm is

grown with a 2 nm p-δ-doped spike (p = 1-2*10^{12} cm^{-2}), located in the center of the AlInAs buffer to compensate the unintentional n-doping in the AlInAs. Then the InP sub-channel with a thickness of 40 nm is grown followed by the 15 nm thick InGaAs-channel. To separate the electrons from their donor atoms a 5 nm thick spacer was introduced. The 30 nm InAlAs-supply layer is uniformly doped with a doping concentration of 1.7*10^{18} cm^{-3}. Finally this is followed by a 5 nm thick highly doped InGaAs cap contact layer.

It is well known that the mobility in the InGaAs channel part can be further improved by increasing the In-content introducing this as a strained layer [7]. To enhance the Schottky barrier height an Al-rich InAlAs strained layer may also be used [8]. However these pseudomorphic structures have up to now only been realized using MBE techniques.

Structure (A2) is pseudomorphic on InP. Thus, here it has been attempted to transfer the advantage of the strained layers to the MOCVD technique. The structure is shown in Fig. 3. The InP sub-channel is unchanged. The InGaAs channel part is subdivided into a 15 nm lattice-matched part and a 10 nm strained part with an In-content of 60% at the interface to the AlInAs spacer layer. The AlInAs supply layer is pulse doped with a doping concentration of 8*$10^{18}$$cm^{-3}$. To improve the barrier height for the Schottky contact a strained $Al_{.59}In_{.41}As$ layer of 10 nm was inserted into the AlInAs buffer located as seen from Fig. 3.

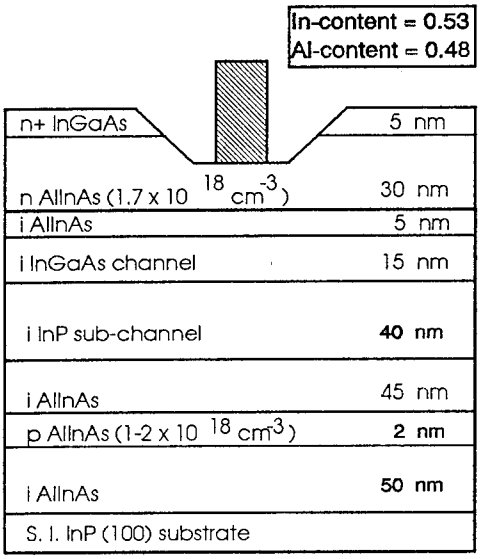


Fig. 2
Structure design A1 with stepped channel

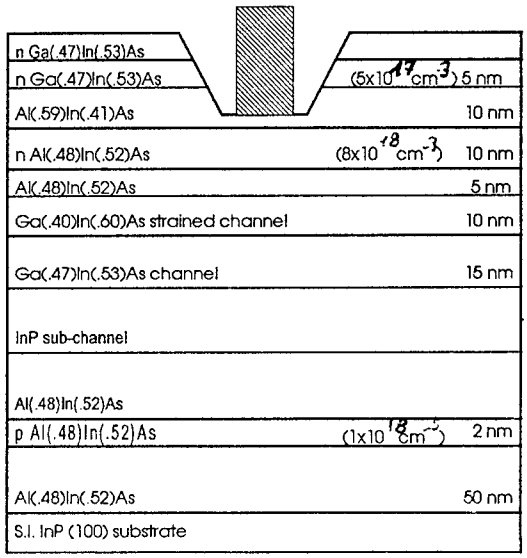


Fig. 3
Structure design A2 with strained channel and strained AlInAs part

The alternative structure design (B) is shown in Fig. 4. A 0.5 µm thick AlInAs buffer was grown underneath the InGaAs/AlInAs superlattice. Several SL configurations have been grown with different SL spacings resulting in various effective conduction band levels. The different superlattices contain 7 periods of 4nm/4nm, 4nm/3nm and 4nm/2nm AlInAs/InGaAs well and barrier thicknesses. The InGaAs channel is 15 nm thick in both cases. The 15 nm thick AlInAs supply layer is doped to 6*$10^{18}$$cm^{-3}$. In the AlInAs supply layer a 20 nm thick undoped layer is inserted at the expected gate recess depth to improve the Schottky barrier characteristics. A highly doped GaInAs cap layer concludes the sequence.

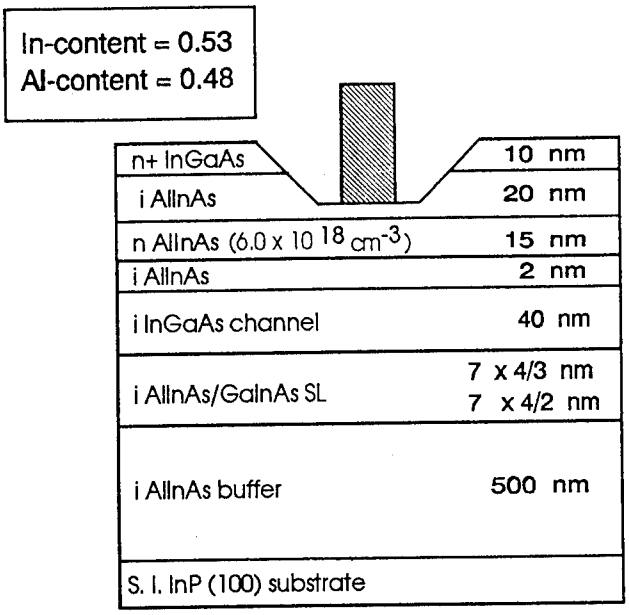


Fig. 4 Structure design with SL configuration. For sample B the thicknesses are AlInAs/InGaAs 4/3nm. The period is 7.

The structures were optimized using the one-dimensional charge control programm CBAND [4], which solves Schrödinger and Poisson equation selfconsistently. The calculated band diagrams for the two stepped-QW (A1, A2) and the SL (B) structures with a SL period of 3nm/4nm are shown in Fig. 5, 6 and 7.

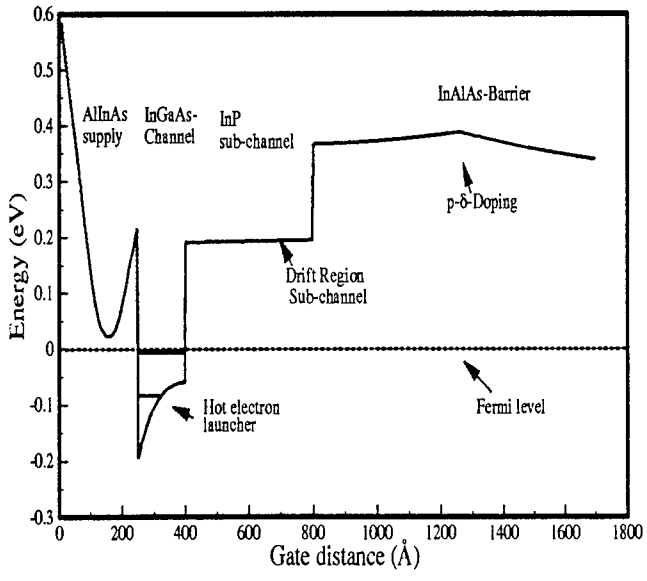


Fig 5 Band diagram for sample A1

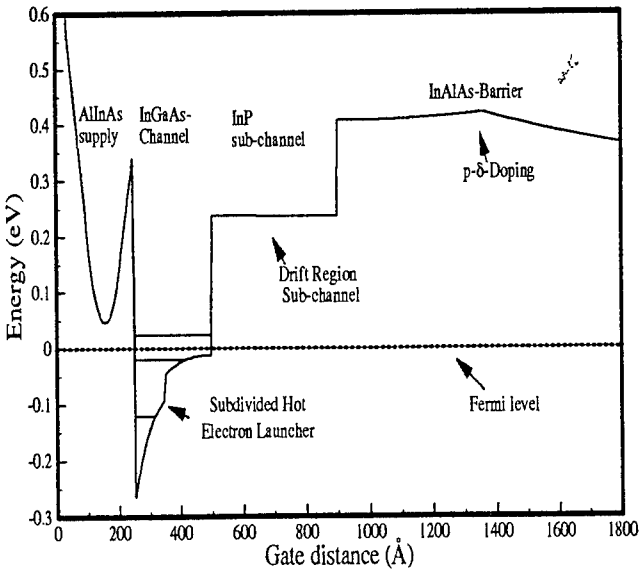


Fig 6. Band diagram for sample A2

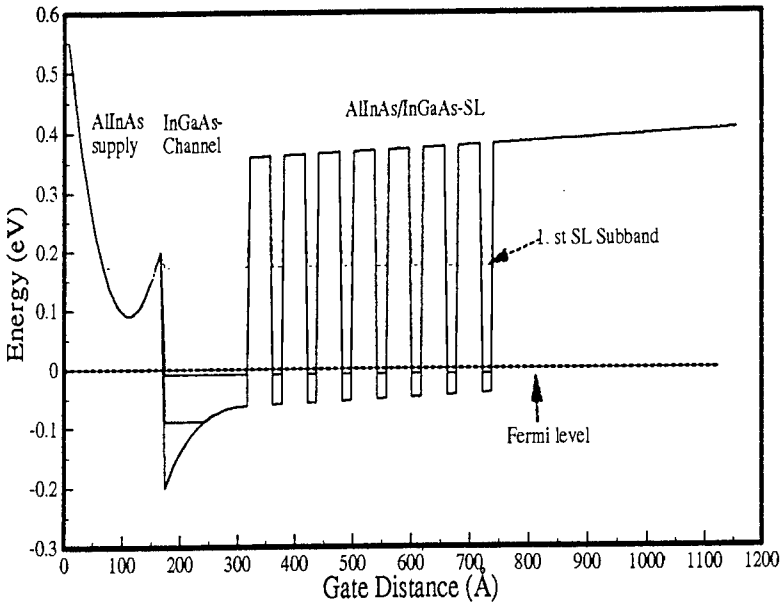


Fig. 7 Band diagram for the SL configuration

In Fig. 5 the stepped channel configuration can clearly be seen (shown for the case V_g=0). The entire 2DEG-density is confined in the InGaAs channel part from which the hot electrons are launched into the InP sub-channel. The band discontinuity between the InP sub-channel and the InGaAs well is 250 meV. In comparison to the unstrained case, the strained layer part of sample (A2) (Fig. 6) leads to a shift of the subband levels to smaller energies and the interface discontinuity to both the InAlAs spacer layer and InP sub-channel is increased as expected.

In Fig. 7 the band diagram for the InGaAs/SL-HFET is plotted. The lowest energy level of the SL represents an effective discontinuity to the conduction band of the InGaAs well of about 210 meV which is close to that of the InGaAs/InP case.

The PL spectra (Fig. 8 and Fig. 9) are used to determine the energy gaps of the InGaAs channels and additionally the 2DEG-density is extracted. Sample (A1) shows two clear transitions and the Fermi edge can be seen, which means that two subbands are occupied as expected from the calculation. For sample (A2), the strained QW layer, a small red shift occurs and the energy difference between the two transitions becomes smaller. Also both subbands are occupied. In both cases the entire 2DEG-density is confined to the InGaAs-QW part. Therefore, a strained layer as commonly used in MBE technology can also be introduced in MOCVD structures without degradation.

In Fig. 9 the two PL spectra of sample (B) with different SL thicknesses of 4nm/3nm and 4nm/2nm InAlAs/InGaAs are compared. The two peaks on the low energy side are transitions from the QW, where the first two subbands are occupied. The peaks on the high energy side originate from the SL. For the SL with the 4nm/2nm-period the subband level shifts to higher energy compared to the 4nm/3nm-period SL and therefore the PL transition should appear at higher energy. This shift could be seen in the PL spectrum. Again the 2DEG-densiy is well confined to the InGaAs channel part.

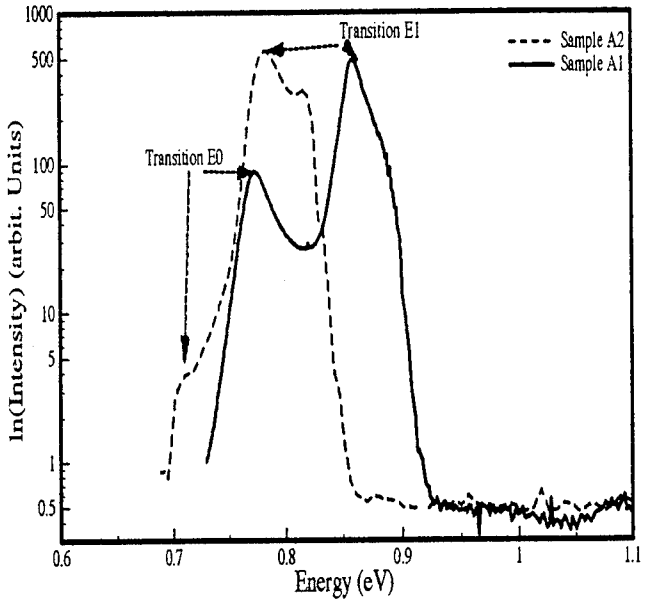


Fig. 8 PL spectra of sample A1 and A2

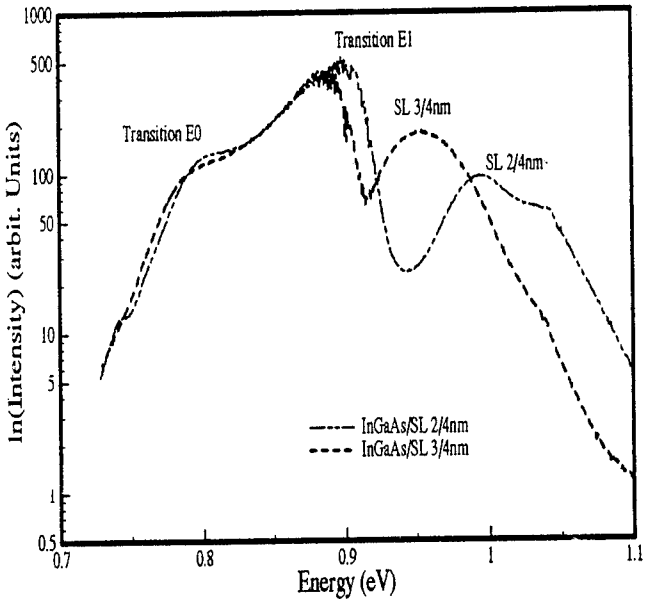


Fig. 9 PL spectra of sample B

From the differences between the Fermi levels and the occupied subbands the sheet carrier concentrations are determined. For sample (A1) a sheet carrier concentration of $2.2*10^{12}cm^{-2}$, for (A2) $2,5*10^{12}cm^{-2}$ and for (B) $2,3*10^{12}$ cm^{-2} is extracted. This is in reasonable agreement with the theoretical data.

3. Technology

All structures are isolated by wet chemical mesa etching. In the case of sample (A1) and (A2) the mesa etching was performed selectively using preferential etching of InGaAs, InAlAs and the InP sub-channel. First the InGaAs and AlInAs was etched down to the InP sub-channel. Next the InP sub-channel was etched selectively and finally the buffer was etched. The InGaAs and InP layers have to be undercut slightly to prevent a short circuit between these layers and the gate metallization crossing the mesa slope. The present technology allows 4 to 5 V gate-drain bias.

Ohmic contacts were patterned by lift-off using a two layer resist process [5]. The GeNiAu contacts were E-Beam evaporated and annealed by RTP at 410° C.

The gates were defined by optical lithography with the same two layer resist process as used for the ohmic contacts resulting in gate lengths between 0,5 and 1,2 µm. The gate was recessed wet chemically. The gate metallization is TiPtAu.

4. DC and RF Results

The output characeristics for the strained InGaAs/InP stepped channel sample and the InGaAs/SL sample are shown in Fig. 10 and Fig. 11. Very similar characteristics have been obtained for the MOCVD grown and the MBE grown structures. No kink effect appears and the small output conductance indicates a well compensated buffer layer by the p-δ-puls. The maximum drain current (I_{dmax}) is 500 mA/mm to 400 mA/mm.

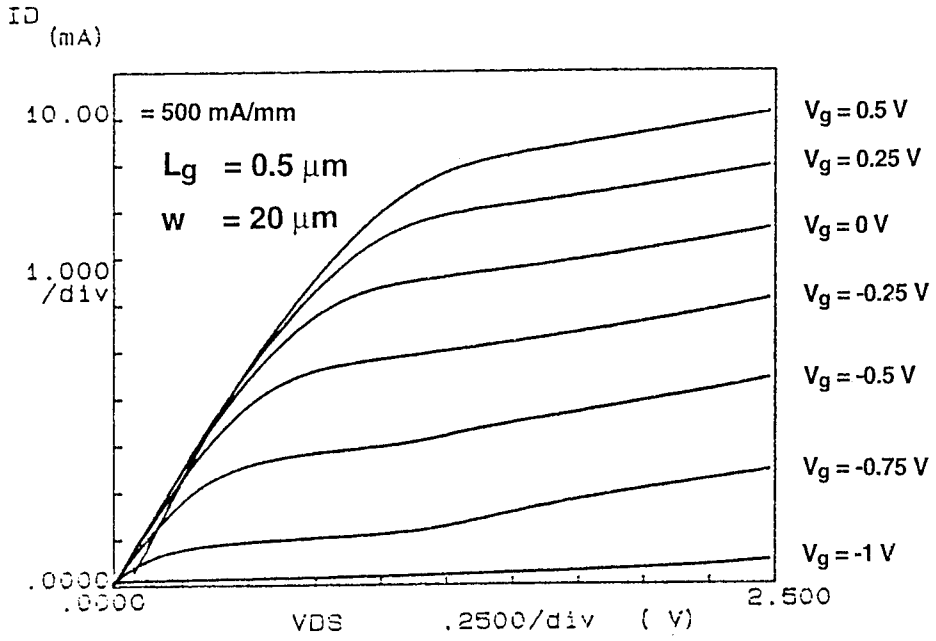


Fig. 10 Output characteristic for the stepped InGaAs/InP QW-Channel HFET

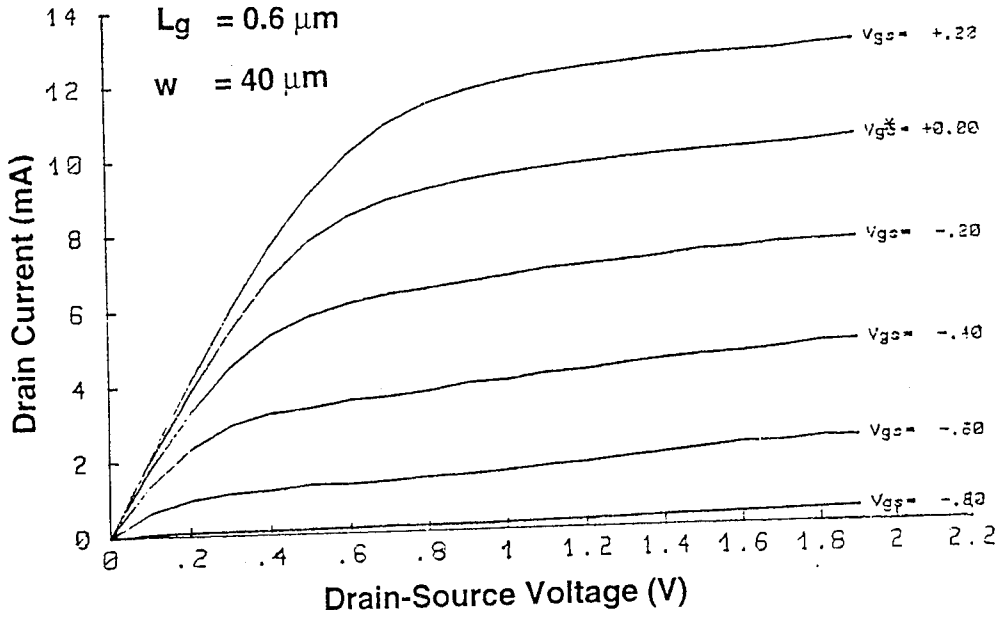


Fig. 11 Output characteristic for the InGaAs/SL HFET

The peak transconductances are between 450 mS/mm (sample A1) (Fig. 9) and 400 mS/mm (sample B).

S-parameter measurements were carried out between 50 MHz and 40 GHz. To extract the small signal circuit diagram parameters the commercially available NP5 software from ATN was used. The main device parameters obtained are listed in Table I.

For the 0.6 µm gate length stepped QW-HFET a f_t of 40 GHz and a f_{max} of 85 GHz has been achieved (Fig. 12) at a drain bias condition of 2.5V which results in a f_{max}/f_t ratio of 2.1. In the case of the InGaAs/SL HFET a f_t of 35 GHz and a f_{max} of 85 GHz are extracted (Fig. 13), which results in a f_{max}/f_t ratio of 2.4. These data are very comparable since high f_{max}/f_t ratios are normally achieved at high drain bias. The differences in f_t could be associated with parasitic elements.

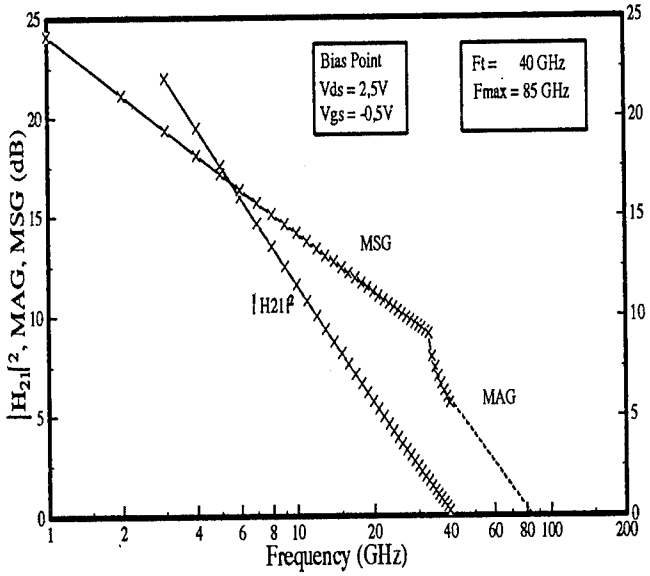


Fig. 12 $|H21|^2$, MAG, MSG over Frequency for the stepped channel sample

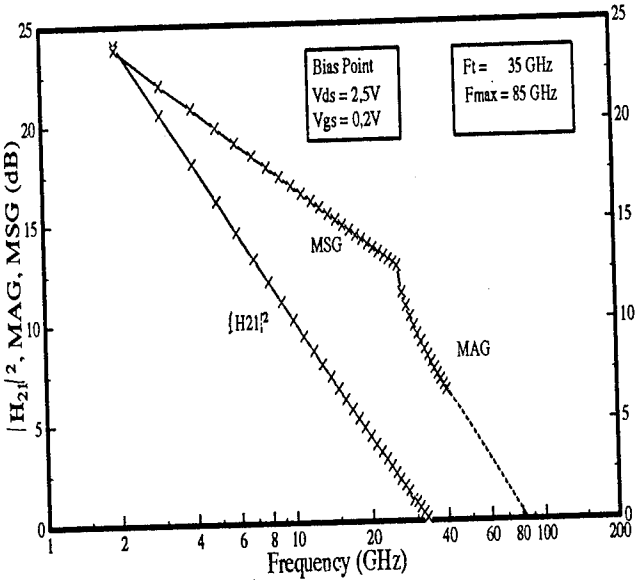


Fig. 13 $|H21|^2$, MAG, MSG over Frequency for the InGaAs/SL sample

Table I

Sample	InGaAs/InP stepped Channel	strained InGaAs/InP stepped Channel	InGaAs/SL
Gate Lenth	Lg = 0.6 µm	Lg = 1.2 µm	Lg = 0.6 µm
I_{dmax} (mA/mm)	500	350	460
g_{mmax} (mS/mm)	400	450	400
f_t (GHz)	40	20	35
f_{max} (GHz)	85	37	85
C_{gs}/C_{gd}	5,3	8	12
f_{max}/f_t	2,1	1,85	2,4

5 Discussion and Conclusion

A novel concept has been presented to combine and improve high speed and high power gain performance of InP based HFETs by introducing a composed channel. In the first part (InGaAs) the electrons are confined in the deep channel and are launched into the second sub-channel of InP, which features a broad velocity-field characteristic. FETs with this stepped channel concept and with 0.6 µm gate length have been realized and tested up to 3V drain bias. At the optimum bias point a f_t of 40 GHz was obtained with an associated f_{max} of 80 to 85 GHz. For a drain bias of 2.5V this results in a favorable f_{max}/f_t ratio of 2.1 to 2.5. High f_{max}/f_t-ratios are normally only achieved at high drain bias. A further important improvement is expected operating at higher drain bias, however the presented devices are still plagued with premature breakdown at the mesa edge.

To quantitatively assess the predicted advantage of a large high-velocity high-field drift region an analysis as suggested in [6] was performed. Here the S-parameters are scanned across the entire operating field of V_g and V_d. The analysis of the small signal parameters C_{gs}, C_{gd}, g_m and g_{ds} allows to extract the intrinsic length of the drift region which is a figure of merit for f_{max} optimization. The drift region length L_{Deff} is given by Eq. 1.

$$L_{Deff} = L_g (\varepsilon W/C_{gs})^2 * ((C_{gs}/C_{gd})-1) \qquad \text{Eq. 1}$$

It can be plotted as a function of the intrinsic channel voltage drop across this region given by Eq. 2.

$$V_{Drift} = V_D - (R_S+R_D)*I_D - E_C L_g + 0.1(V_g - R_S I_D) \qquad \text{Eq. 2}$$

The deduction of Eq. 1 and 2 and the meaning of the symbols are given in [6] found in these proceedings.

This procedure has been used to compare both structure type (A) and (B). In Fig. 14 the development of L_{Deff} with the intrinsic drift voltage (V_{Drift}) is shown. The common base for the comparison is the gate bias condition for optimum f_t. One can see, that L_{Deff} develops essentially linear. In the interpretation of the high-field drift region as space charge layer the L_{Deff}-V_D-characteristics will depend on the lateral space charge distribution. A linear relationship can be explained by a hyperbolic space charge profile with a high electron density at the gate end and a small electron density travelling at high overshoot velocity at the drain end. If a high velocity is reached at high drain field, the linear L_{Deff}-V_{Drift} relationship is maintained. This is the case for the stepped InP channel device up to the maximum voltage applied. If the overshoot velocity decreases to the high field saturated velocity, the slope should decrease. This tendency is indeed observed for the InGaAs channel devices. Shown are two devices: structure (B), which comes very close to the standard InGaAs HEMT structure and a standard InGaAs-HEMT with 0.3 µm gate length, which has been analysed in [6].

In conclusion, it has been possible to demonstrate a novel stepped channel concept for InP-based HEMTs. The heterostructure grown by MOCVD included pseudomorphic layers. First RF-results with InP-sub-channel indicate indeed an improved high-field drift region. However, devices with a InGaAs/InAlAs sub-channel show very similar behavior to standard InGaAs HEMTs. Although the technology for this novel InGaAs/InP channel device has essentially been established, gate to channel breakdown has still to be improved to demonstrate the full potential of this structure.

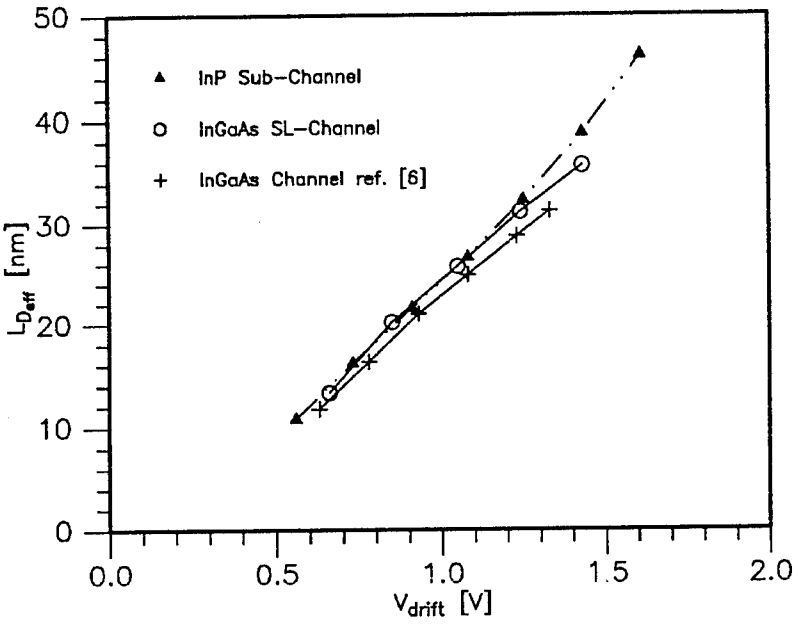


Fig. 14 Drift region as function of the effective drain voltage plotted for the two types of configuration

References:

[1] O. Aina, M. Burgess, M. Mattigly, A. Meerschaert, J. M. O´Connor; "A 1.45-W/mm, 30-GHz InP-Channel Power HEMT"; IEEE Electron Device Letters, Vol. 13, 1992; 300-302

[2] G. I. Ng, D. Pavlidis, Y. Kwon, T.Brock, J. I. Davies, G. Clarke, P. K. Rees; "0.1 µm MOVPE Grown InAlAs/InGaAs HEMT´s with Above 150 GHz Operation Capability" Fourth International Conference on InP and Related Materials; April 20-24,1992; Newport, Rhode Island, USA

[3] J. M. Golio, R. J. Trew; "Optimum Semiconductors for High-Frequency and Low-Noise MESFET Applications"; IEEE Transaction on Electron Devices, Vol. ED-30, NO. 10, Oct. 1983;

4. Marc Foisy Ph. D May 1990 Cornell University; "A Physical Model for the Bias dependence of Modulation Doped Field Effect Transistors High Frequency Performance"; Progamm CBAND

[5] X. Zhang, S. Strähle, E. Mittermeier, and E. Kohn; "Novel Optical high Resolution Three Layer Resist Process"; SOTAPOCS, Honolulu Hawai, May 16-21, 1993 Proceedings

[6] E. Kohn, S. Strähle, D. Geiger, U. Erben; "High Field Drift Domains in GaAs and InP Based Heterostructure Field Effect Devices"; This IEEE/Cornell Conference Paper 6.2

[7] Loi D. Ngygen, April S. Brown, Mark A. Thomson, Linda M. Jelloian, Larry E. Larson and Mehran Matloubian; "650-A Self-Aligned-Gate Pseudomorphic $Al_{0.48}In_{0.52}As$/-$Ga_{0.20}In_{0.80}As$ High Electron Mobility Transistros";IEEE Electron Device Letters, Vol. 13, 1992; 143-145

[8] Sandeep R. Bahl, Brian R. Bennett, and Jesús A. del Alamo; "Doubly Strained $In_{0.41}Al_{0.59}As/n^+$ -$In_{0.65}Ga_{0.35}As$ HFET with High Breakdown Voltage"; IEEE Electron Device Vol. 14, January 1993, 22-24;

This work is supported by the Deutsche Forschungsgemeinschaft under the Contract Nr. Ko 1201/2-1

Strained Layer Device Epitaxy on Patterned Substrates

W. J. Schaff, K. Y. Hur, L. F. Eastman, and R. C. Compton
School of Electrical Engineering
Cornell University
Ithaca, New York 14853, USA

P. Mandeville
Bell-Northern Research
Ottawa, Ontario K1Y 4H7, Canada

Abstract

Growth of pseudomorphic MODFET structures on patterned substrate was investigated as means of increasing the critical layer thickness of strained layers. Prior to growth, semi-insulating GaAs substrates were patterned and etched using chemically assisted ion beam etching to define a series of mesas. Double-doped pseudomorphic MODFET layers were then grown on the substrates by molecular beam epitaxy. To fabricate MODFETs on the resulting non-planar wafer, a new fabrication technique has been developed.

Introduction

The epitaxial thickness of a strained layer is limited by the critical layer thickness due to the formation of misfit dislocations[1]. However, it has been previously demonstrated that the density of misfit dislocations can be reduced by limiting the growth areas[2,3]. Localized growth areas, defined by patterning the initial substrate, have been previously used to fabricate InGaAs/GaAs photodetectors with enhanced quantum efficiency[4].

Thicker strained layers with higher InAs mole fractions are desired in MODFET channels for higher sheet densities and higher output currents. In this work, the patterned substrate technique was explored for pseudomorphic InGaAs/AlGaAs MODFETs. Each device spans a group of mesas that are defined in the initial patterning step. A trench planarization technique and a vacuum-passivated gate technique is then used. To assess the quality of the device layers after the growth, an open channel current measurements scheme was utilized on test structures described in the following section.

0-7803-0894-8/93 $3.00

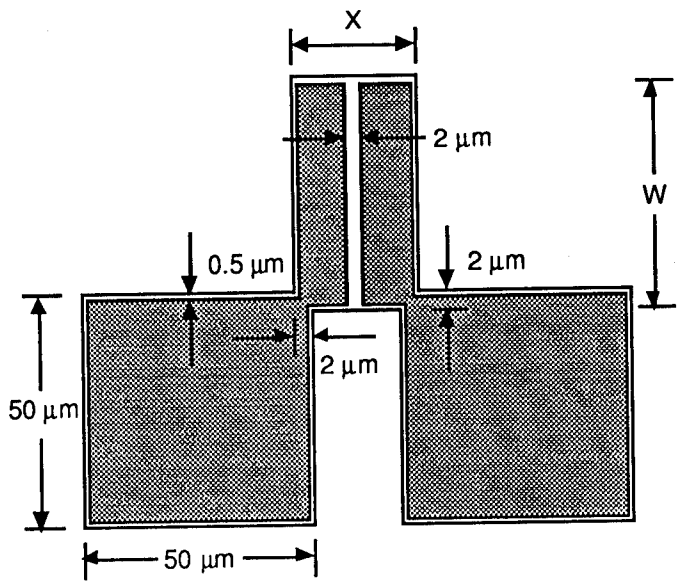


Figure 1. Test structure for open channel current study. Active areas range from $25\,\mu m^2$ to $10{,}000\,\mu m^2$.

Study of Open Channel Currents

A series of isolated mesa with a range of active areas were first patterned and etched in semi-insulating GaAs using Cl_2-based chemically assisted ion beam etching (CAIBE) (Figure 1). On these samples, a variety of double-doped pseudomorphic MODFET structures with $In_xGa_{1-x}As$ channels (Figure 2) were then grown by molecular beam epitaxy (MBE). The channel thickness was 120Å, while the InAs mole fraction ranged from 0.2 to 0.3. Two ohmic contacts were deposited on each mesa by evaporation of 100Å Ni, 300Å Ge, 600Å Au, 1000Å Ag, and 1000Å Au, followed by rapid thermal annealing at 450° C for 15 seconds in Ar/H_2 ambient.

DC Measurements were performed on the test structures to investigate the effects of channel currents on InAs mole fractions and mesa areas. As illustrated in Figure 3, the currents decrease as the InAs mole fraction was increased. Furthermore, on the 30% InAs sample, currents also decrease gradually as the mesa area is increased. These observations indicate that the density of the misfit dislocations is considerably higher on the 30% indium sample, indicating perhaps that the growth conditions are not yet optimal.

MODFET Fabrication and Results

For device integration on the relatively mesas ($10\,\mu m$ by $10\,\mu m$), a dielectric planarization and vacuum-passivated gate technique have been developed to fabricate pseudomorphic

Cap	n^+ GaAs	400 Å
Barrier	undoped $Al_{0.3}Ga_{0.7}As$	250 Å
APD	Si: $6x10^{12}$ /cm^2	
Spacer	undoped $Al_{0.3}Ga_{0.7}As$	30 Å
	undoped GaAs	20 Å
Channel	undoped $In_xGa_{1-x}As$	120 Å
Spacer	undoped GaAs	20 Å
	undoped $Al_{0.3}Ga_{0.7}As$	50 Å
APD	Si: $2x10^{12}$ /cm^2	
Barrier	undoped $Al_{0.3}Ga_{0.7}As$/GaAs 90Å/10Å SL	600 Å
Buffer	undoped GaAs	5000 Å
SI GaAs Substrate		

Figure 2. Layer structure of $In_xGa_{1-x}As$ MODFETs. InAs mole fractions range from 20% to 30% for the open channel current study (no gate present).

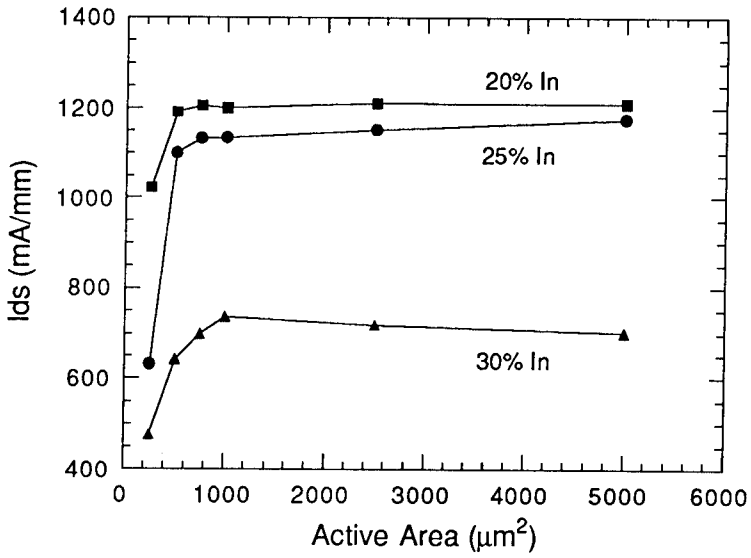


Figure 3. Variation of open channel currents as a function of InAs mole fraction and mesa area. The test structure shown in Figure 1 was utilized for these measurements.

MODFETs that span many mesas. The device structure is embedded in a coplanar waveg-

uide configuration for on-wafer measurements (Figure 4). $In_{0.2}Ga_{0.8}As$ MODFETs with 0.3 μm gate length and 200 μm gate width were fabricated to demonstrate the patterned substrate growth and fabrication techniques. Figure 5 is an SEM of the cross-section of the MODFET, showing the planarized trenches on the source and drain sides. Unetched Hall mobility and sheet carrier density were 5113 cm^2/V-sec and 5.87×10^{12} cm^{-2} at 300 K. The current-voltage characteristics are shown in Figure 6.

A linearly graded, 120Å-wide, $In_xGa_{1-x}As$ MODFET structure ,where x is graded from 0.16 up to 0.27 and then down to 0.19 (triangular), were also grown and processed. The gate length for the linearly graded MODFET is approximately 0.4 μm. The IV measurements, shown in Figure 7, illustrate similar characteristics as the uniform channel MODFET above. However, the graded MODFETs showed smaller knee voltages with lower on-resistance. This is an indication of better transport properties in the graded channel structure as a result of the uniform and symmetric electron distribution in the center of the quantum well[5]. In both MODFET structures, maximum open channel currents of approximately 750 mA/mm and output currents of 360 mA/mm at V_{gs}=0V were measured. For the two pseudomorphic MODFET structures fabricated on patterned GaAs substrates, unity current gain frequency, f_t of 50 GHz and maximum frequency of oscillation, f_{max} in excess of 100 GHz, with peak extrinsic transconductances of 380 mS/mm were obtained. The gate lengths of the MODFETs can be further reduced down to a 0.1 μm-scale to obtain higher frequency response.

Summary

In this paper, a new fabrication sequence for pseudomorphic MODFETs on patterned GaAs substrates was described. An open channel measurement technique was developed was used to gauge the effects of misfit dislocation densities in the strained layer channel. For a MODFET consisting of a 120Å-wide strained layer channel with a 30% InAs mole fraction, the currents were still considerably lower, believed to be as a result of higher density of misfit dislocations. Therefore, the patterned substrate technique did not appear to reduce the density of misfit dislocations on the 30% InAs sample.

Acknowledgments

This work is supported by the National Science Foundation, the Defense Advanced Research Projects Agency, and the U.S. Army Research Office. Device fabrication was performed at the National Nanofabrication Facility (NNF) at Cornell, which is supported by the National Science Foundation (ESC-8619049), Cornell University, and Industrial Affiliates. The authors would like to thank M. J. Skvarla and R. J. Bojko of the NNF.

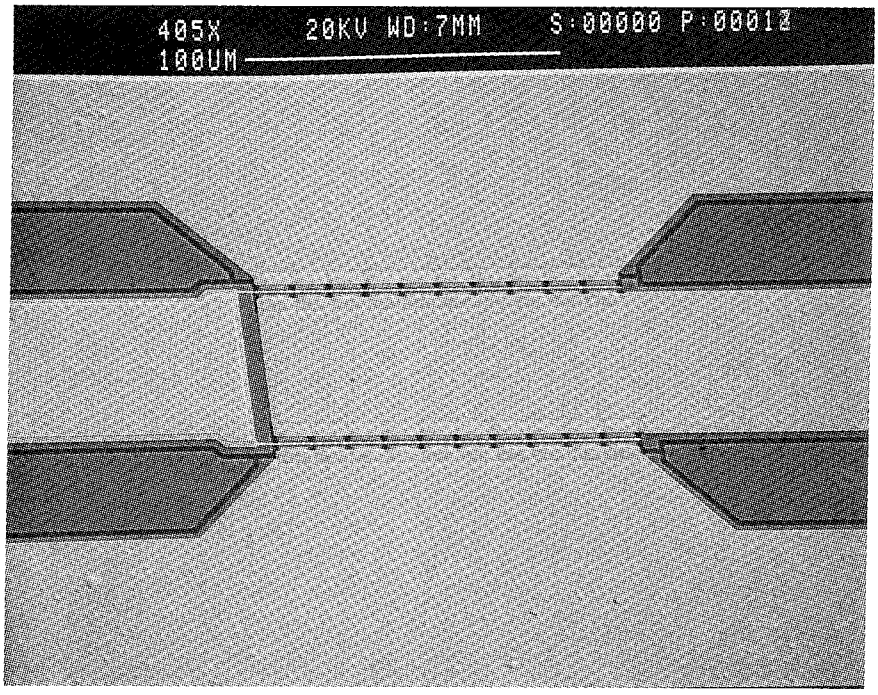


Figure 4. SEM of a pseudomorphic MODFET fabricated on the patterned GaAs substrate. The vacuum-passivated gate sits on a group of mesas that are isolated by a planarized layer of SiO_2.

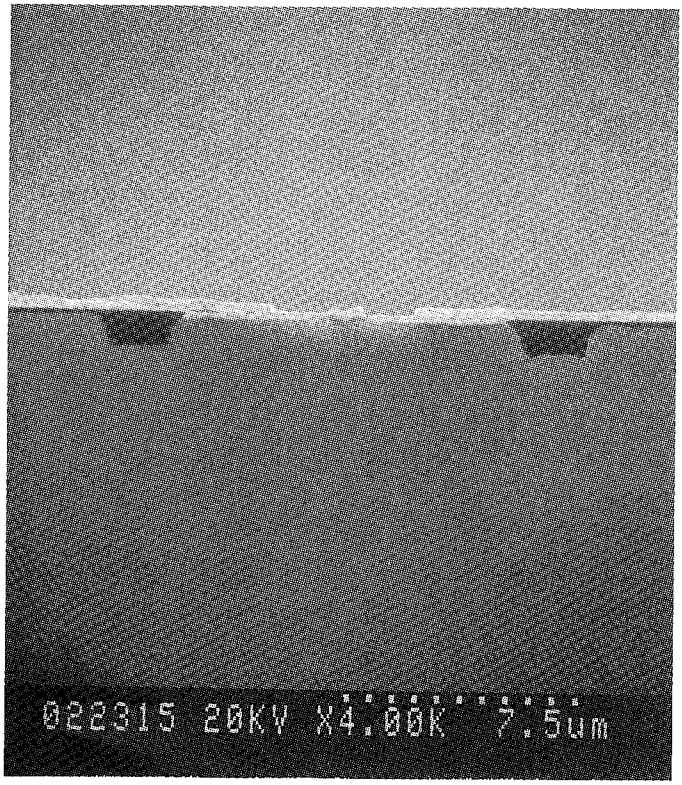


Figure 5. SEM of a pseudomorphic MODFET cross-section. The vacuum-passivated T-gate sits in the middle with the planarized trenches after the source and drain.

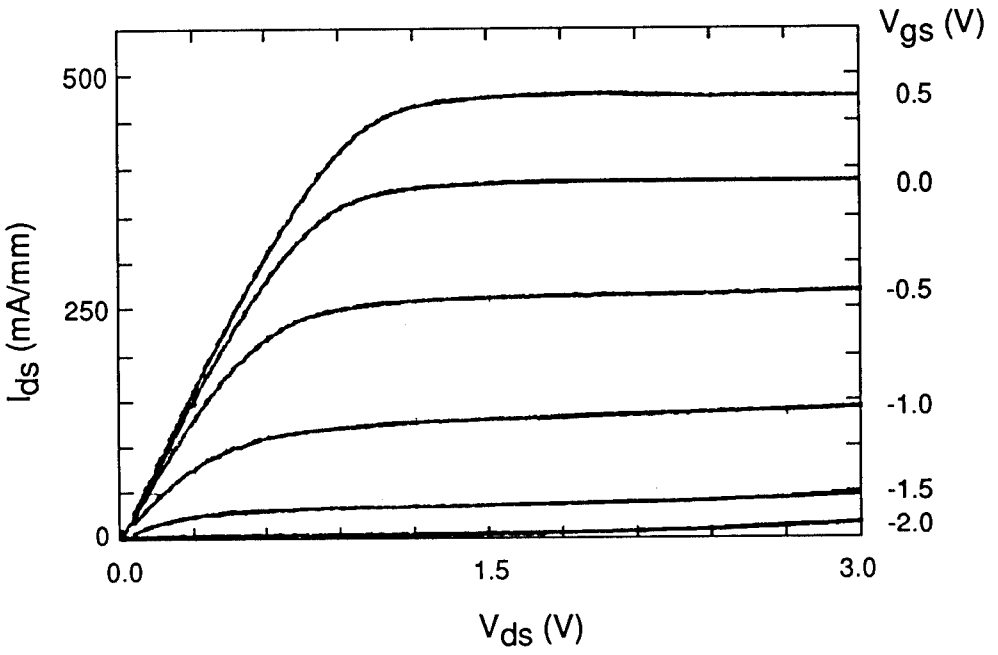


Figure 6. Current-voltage characteristics for a 0.3 μm-long, 200 μm-wide, uniform $In_{0.2}Ga_{0.8}As$ MODFET fabricated on a group of 20 mesas.

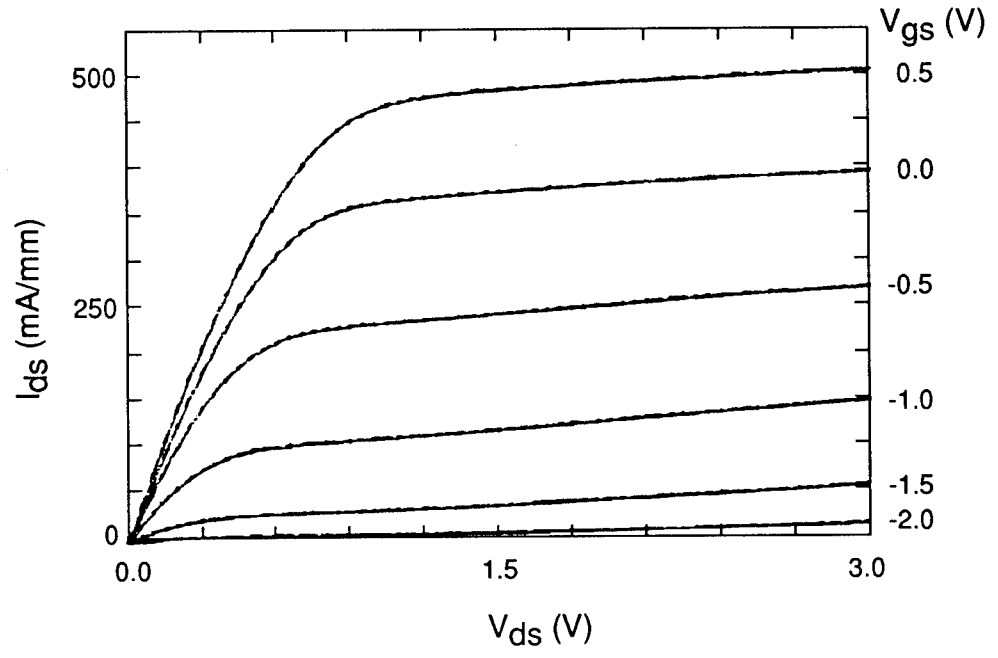


Figure 7. Current-voltage characteristics for a 0.4 μm-long, linearly-graded $In_xGa_{1-x}As$ MODFET, where x is graded from 0.16 to 0.27 to 0.19.

References

[1] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multi-layers I. Misfit dislocations," *Journal of Crystal Growth*, Vol. 27, 118-125 (1974).

[2] E. A. Fitzgerald, G. P. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall, "Nucleation mechanisms and the elimination of misfit dislocations at mismatched interfaces by reduction in growth area," *Journal of Applied Physics*, 65 (6), 2220-2237 (1989).

[3] G. P. Watson, Ph.D. Thesis, Cornell University (1992).

[4] W. Q. Li, P. K. Bhattacharya, and R. L. Tober, "Improved performance of strained InGaAs/GaAs photodiodes grown on patterned GaAs substrates by molecular beam epitaxy," *Applied Physics Letters*, 58 (19), 1931-1933 (1991).

[5] T. K. Yoo, P. Mandeville, H. C. Park, W. J. Schaff, and L. F. Eastman, "Double modulation-doped AlGaAs/InGaAs heterostructure with a graded composition in the quantum well," *Applied Physics Letters*, 61 (16), 1942-1944 (1992).

METAMORPHIC $In_{0.3}Ga_{0.7}As/In_{0.29}Al_{0.71}As$ LAYER ON GaAs: A NEW STRUCTURE FOR MILLIMETER WAVE ICs.

P. Win, Y. Druelle and A. Cappy

Institut d'Electronique et de Microélectronique du Nord, Centre Hyperfréquences et Semiconducteurs, UMR CNRS 9929. Bâtiment P4. Université des Sciences et Technologies de LILLE. 59655 VILLENEUVE D'ASCQ CEDEX, FRANCE.

Y. Cordier, D Adam and J. Favre

Laboratoire Central de Recherche, THOMSON-LCR, Domaine de Corbeville, 91401 ORSAY CEDEX, FRANCE.

ABSTRACT

A new High Electron Mobility Transistor (HEMT) using InAlAs/InGaAs grown on GaAs has been successfully realized. This device, with an In content close to 30%, presents several advantages over conventional pseudomorphic HEMT on GaAs and over lattice matched HEMT on InP. To accomodate the mismatch between the active layer and the GaAs substrate, a metamorphic buffer has been grown at low temperature. High electron mobility with high two dimensional electron gas density (25000 $cm^2/V.s$ with $3x10^{12}$ cm^{-2} at 77K) as well as high Schottky barrier quality (Vb=0.68V with η=1.1) have been obtained. A device with a 0.4x150 μm^2 gate geometry has shown a transconductance as high as 700 mS/mm at a current density of 230 mA/mm. The measured f_T was 45 GHz and f_{max} was 115 GHz. These values, showing the great potentiality of this structure, are, to the authors knowledge, the first reported for submicrometer gate metamorphic InAlAs/InGaAs/GaAs HEMTs with an Indium content of 30%.

I. INTRODUCTION

For several years, the work on InGaAs based High Electron Mobility Transistors (HEMTs) has shown the superiority of these structures over GaAs MESFETs and over conventional AlGaAs/GaAs HEMTs. Pseudomorphic HEMT on GaAs (PM-HEMT) is now commercially available, and has shown excellent performance in the field of low noise and power amplifiers [1],[2], while lattice matched HEMT on InP (LM-HEMT) has shown world record performance in the field of low noise amplifiers in the millimeter wave range [3],[4]. For these devices, the main limitations are the following:

In PM-HEMT, Indium content is limited to 25-30% to preserve high layer quality. Therefore, the conduction band discontinuity ΔEc is limited to about 0.35 eV and the sheet carrier density is limited to about $2.0x10^{12}cm^{-2}$ for one quantum well. In addition, the electron transport properties (mobility, peak velocity.....) in strained InGaAs are lower than in relaxed InGaAs with the same indium content. In fact, it was shown, using Monte Carlo

0-7803-0894-8/93 $3.00

calculation including strain effects, that the electron transport in the channel of a PM-HEMT is similar to the electron transport in a conventional GaAlAs/GaAs HEMT [5].

In LM-HEMT, the power performance is limited by the low value of the InGaAs bandgap (0.75 eV) which induces impact ionization in the channel for drain voltages higher than 1.5V [6]. In addition the built-in potential on $In_{0.52}Al_{0.48}As$ (Vb ~ 0.55V) is rather low and usually the Schottky barrier quality is rather poor on this material [3].

To overcome these difficulties, a new structure showing higher electron velocity than in PM-HEMTs and larger breakdown voltage than in LM-HEMTs is proposed. The design of this new structure is based on the following considerations.

Let us first consider the bandgap variation of InAlAs and InGaAs ternary alloys as well as the conduction band discontinuity ΔEc at the InAlAs/InGaAs heterojunction interface as a function of the Indium content (figure 1). To plot ΔEc, the In content in InAlAs has been chosen to ensure lattice matching and the rule $\Delta Ec/\Delta Eg=0.65$ has been used. Figure 1 shows that the bandgap energy of the InGaAs and InAlAs alloys is decreasing with increasing In content, while the conduction band discontinuity ΔEc shows a maximum for an In content of about 30% as the InAlAs bandgap becomes direct. So, it seems interesting to use unstrained InGaAs/InAlAs heterojunction with an Indium content close to 30% for HEMT realization [7] Such an active layer has the following advantages:

(i) the high conduction band discontinuity (ΔEc ~ 0.65-0.7 eV) leads to high sheet carrier density and good electron confinement,

(ii) unstrained material is better to produce high electron velocity [5],

(iii) the high bandgap of InAlAs involves better Schottky barrier quality,

(iv) the high InGaAs bandgap allows higher drain voltage operation and reduces impact ionization.

It should be noted that similar considerations have been given in reference [8]. Unfortunately, no commercial substrate allows to grow the lattice matched $In_xGa_{1-x}As$-$In_yAl_{1-y}As$ system for x and y close to 0.30. So a relaxed $In_xGa_{1-x}As$ buffer , referred to herein as metamorphic (MM) buffer [9] grown on the GaAs substrate is needed. This buffer has two functions:

(i) To accomodate the large lattice mismatch (~ 2%) between the active layer and the GaAs substrate by the formation of misfit dislocations.

(ii) To trap these misfit dislocations and to prevent their propagation in the active layer which is grown on this buffer.

II MATERIAL GROWTH AND CHARACTERIZATION

Several InGaAs buffers were grown on GaAs using a RIBER 32P MBE machine. According to previous works performed in the field of the growth of relaxed-mismatched buffer [10], a low growth temperature (350°C) was chosen to avoid island mode growth and a V/III flux ratio of 25 was used. The Indium content measured by photoluminescence at 77K and by double crystal X-Ray diffraction was close to 30%.

Satisfactory results have been obtained with two different types of buffer. Buffer #1 is a 1-μm linear Indium composition graded layer of InGaAs, with x_{In} varying from 10 to 30%, grown on a thin GaAs buffer (figure 2(a)) while Buffer #2 is more complex: 50 nm of GaAs, 450 nm of $In_{0.15}Ga_{0.85}As$, a linear composition graded layer of 100 nm to x_{In}=30% and 450 nm of $In_{0.3}Ga_{0.7}As$ (figure 2(b)).

In these two cases, the dislocation density, estimated by Cross Sectional Transmission Electron Microscopy, is less than 10^7 dislocations/cm^2. Nomarski interference observations have shown cross-hatching in both buffers, indicating strain relaxation in the structure [11].

According to these satisfactory results, the growth of HEMT active layers on Metamorphic buffer has been carried out.

The MM-HEMT active layer was grown on the MM-buffer at higher substrate temperature (500°C). It is constituted by 200 nm of $In_{0.29}Al_{0.71}As$, 100 nm of undoped $In_{0.3}Ga_{0.7}As$, 3 nm of $In_{0.29}Al_{0.71}As$ spacer, a 30 nm of undoped $In_{0.29}Al_{0.71}As$ electron supplying layer and a $In_{0.3}Ga_{0.7}As$ cap layer (figure 3). For electron supplying layer, both bulk doping (20 nm doped at $2x10^{24}$ m^{-3}) and planar doping ($\delta = 5x10^{16}$ m^{-2}) were used. The influence of the doping technique on the device performance was found to be similar to those observed in pseudomorphic HEMTs. Concerning the cap layer, both undoped cap and doped cap were used. However, even in the case of the undoped cap layer, fairly good access resistances (0.6 to 1.0 to Ω.mm) were obtained due to the high sheet carrier density in the channel.

Van der Pauw-Hall patterns were fabricated on as grown layers. Figure 4 shows the measured Hall mobility μ_H as a function of the measured sheet carrier density n_H. At room temperature , hall mobilities as high as 8500 $cm^2/V.s$ have been obtained for sheet carrier densities ranging from $5x10^{11}$ to $3x10^{12}$ cm^{-2}. At 77K, the highest mobility was 41000 $cm^2/V.s$ with a sheet carrier $n_H=3x10^{11}$ cm^{-2} for a layer with a 200 Å spacer. As the spacer thickness decreases, the mobility decreases while the sheet carrier density increases. At $n_H=3x10^{12}cm^{-2}$, the hall mobility is close to 25000 $cm^2/V.s$. These values compare favourably with pseudomorphic layers, confirming the potential superiority of MM-HEMT over PM-HEMT.

III DEVICE PROCESSING AND CHARACTERIZATION

Long gate devices (Lg = 3 to 10 μm) have been realized by using conventional photolithography and wet chemical etching. Mesa isolation is realized by chemical etching with H_3PO_4-based solution. For ohmic contact realization, Au/Ge/Ni metals were evaporated and annealed at 420°C for 1 minute. After gate recess etching, Ti/Al metals have been evaporated to realize gate contact. I-V, C-V , magneto-resistance and magneto-transconductance measurements have been realized on these long gate devices .

The Ids(Vds,Vgs) characteristics of a 3x140 μm^2 device are presented in figure 5. Extrinsic transconductance of 390 mS/mm is obtained at Vgs=+0.42V. After series resistance correction, this corresponds to an intrinsic transconducatnce of 530 mS/mm and it should be emphasized that this value is extremely high for a 3 μm gate length device. Another interesting result of the device characterization is the high Schottky barrier quality deposited on undoped $In_{0.29}Al_{0.71}As$. Built-in potential as high as 0.68V has been obtained with an excellent ideality factor (1.1), while the reverse breakdown voltage is larger than -10V (figure 6). These

experimental results clearly show that the Schottky barrier realized on metamorphic material is better than that usually obtained on InAlAs lattice matched on InP. The capacitance- voltage and differential mobility-voltage characteristics, deduced from magneto-transconductance measurements, are shown in figure 7.

The same technology was used to realize submicrometer gate length MM-HEMTs. The source-to-drain spacing is nominally 1.5 μm. The 0.4x150 μm^2 Ti/Pt/Au triangular gate was evaporated through a recess window defined by e-beam lithography. The measured gate length was in the range 0.4-0.45 μm across the wafer. High quality Schottky contact has been achieved (Vb = 0.64 V with $\eta = 1.17$), due to the high bandgap value of $In_{0.29}Al_{0.71}As$. Due to the undoped cap layer, the maximum reverse breakdown voltage is high (Ig = 5 mA at -10 V) while the access resistance is rather high (0.95 Ω.mm).

DC current voltage characteristics of the 0.4x150 μm^2 MM-HEMT are shown in figure 8. Good pinch-off behaviour can be observed. 400 mA/mm maximum saturation current Ids_{max} is measured at Vgs = +0.4 V. S-parameters measurement was performed in the 1-40GHz frequency range using on-wafer probing. The small signal equivalent circuit was directly determined from S-parameter according to the method described in [11]. Figure 9 shows the intrinsic cutoff frequency $f_c = \frac{gm_0}{2\pi Cgs}$ and the intrinsic transconductance as a function of the gate voltage at Vds = +2 V. For a current density of 230 mA/mm, a maximum of intrinsic cutoff frequency of 65 GHz is obtained with an associated transconductance of 700 mS/mm (420 mS/mm extrinsic). The average electron velocity <v> can be deduced from the intrinsic cutoff frequency according to $\langle v \rangle = 2\pi Lg\, f_c$. For the 0.4 μm gate metamorphic device, the calculated value is $\langle v \rangle = 1.6 \times 10^7$ cm/s showing the high electron transport properties in metamorphic layers confirming thus the good material quality.
The current gain ($|H_{21}|^2$), the maximum stable gain (MSG), the maximum available gain (MAG) and the unilateral gain (U) are plotted versus frequency in figure 10. The extrapolation of the current gain to unity using a 6 dB/octave slope yields an f_T of 45 GHz, and the extrapolation of the unilateral gain U gives a maximum frequency of oscillations f_{max} of 115 GHz. This high value of f_{max} is due to the low value of the gate to drain capacitance Cgd (about 53 fF/mm) resulting from the quasi-planar topology of the device.

IV CONCLUSION

Simple theoretical considerations have shown that $In_{0.29}Al_{0.71}As/In_{0.3}Ga_{0.7}As/GaAs$ HEMTs can be a good candidate for applications in the microwave and millimeter-wave range. To accomodate the mismatch between the active layer and the GaAs substrate and to trap the dislocations, a metamorphic buffer was grown at low temperature between theactive layer and the GaAs substrate. Hall mobility as high as 41000 $cm^2/V.s$ has been obtained at 77K. Both long and submicrometer gate metamorphic HEMTs have been successfully realized. High microwave performance was obtained (gm_{max} = 700 mS/mm at Vds = +2V and Ids = 230 mA/mm, f_c = 65 GHz and f_{max} = 115 GHz). These performances are better than those obtained in our laboratory for a pseudomorphic HEMT with similar gate geometry. These results show the great interest of metamorphic approach for HEMT realization.

REFERENCES

[1] M.Y. Kao *et al.*, *IEEE Elect. Dev. Lett.*, **10**, 580 (1989).

[2] K.L. Tan *et al.*, *IEEE Elect. Dev. Lett.*, **11**, 585 (1990).

[3] U.K. Mishra *et al.*, *IEEE Elect. Dev. Lett.*, **9**, 647 (1988).

[4] P.C. Chao *et al.*, *IEEE Elect. Dev. Lett.*, **11**, 59 (1990).

[5] J.L. Thobel, *Appl. Phys. Lett.*, **56**, 346 (1990).

[6] K.H.G. Duh *et al.*, *IEEE MTT-S Digest*, 805 (1989).

[7] P. Win *et al.*, *Proc. of S.O.T.A.P.O.C.S. XV*, Phoenix, AZ, 100 (1992).

[8] P. Chu *et al.*, *Thin Solid Films*, **166**, 155 (1988).

[9] D.E. Grider *et al.*, *J. Vac. Sci. Technol.*, **B8**, 301 (1990).

[10] J.C. Harmand *et al.*, *Jpn Jal Appl. Phys.*, **28**, L1101 (1989).

[11] G.H. Olsen, *J. Cryst. Growth*, **31**, 223 (1975).

[12] G. Dambrine *et al.*, *IEEE - MTT*, **36**, 1151 (1988).

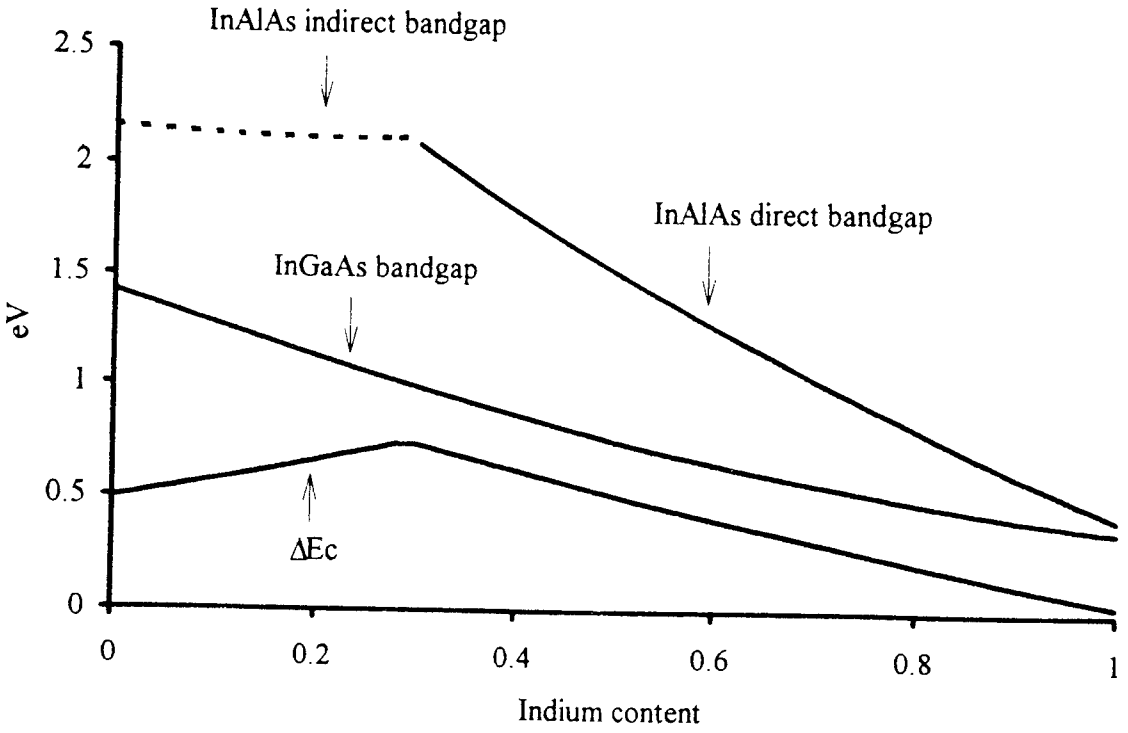


Figure 1: Bandgaps of InAlAs and InGaAs and related conduction band discontinuity ΔEc versus Indium content.

1 μm	Graded $In_xGa_{1-x}As$	undoped
	x: 0.1 → 0.3	

(001) S.I. GaAs substrate

(a)

450 nm	$In_{0.3}Ga_{0.7}As$	undoped
100 nm	Graded InGaAs	undoped
	x: 0.15 → 0.3	
450 nm	$In_{0.15}Ga_{0.85}As$	undoped
50 nm	GaAs	undoped

(001) S.I. GaAs substrate

(b)

Figure 2: Cross sectional view of metamorphic buffers #1 (a) and #2 (b).

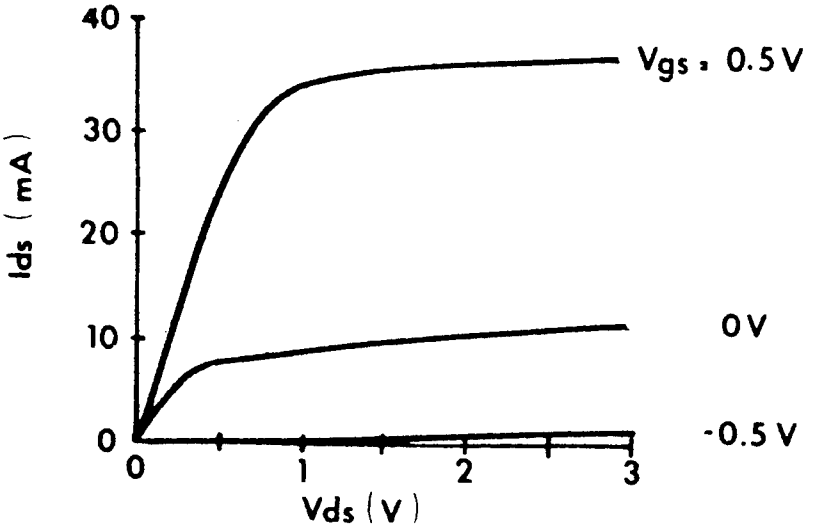


Figure 5: DC characteristic of a 3 μm gate MM-HEMT.

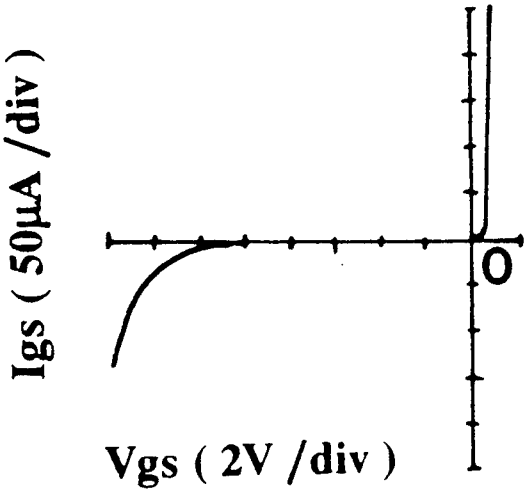


Figure 6: Schottky contact characteristic.

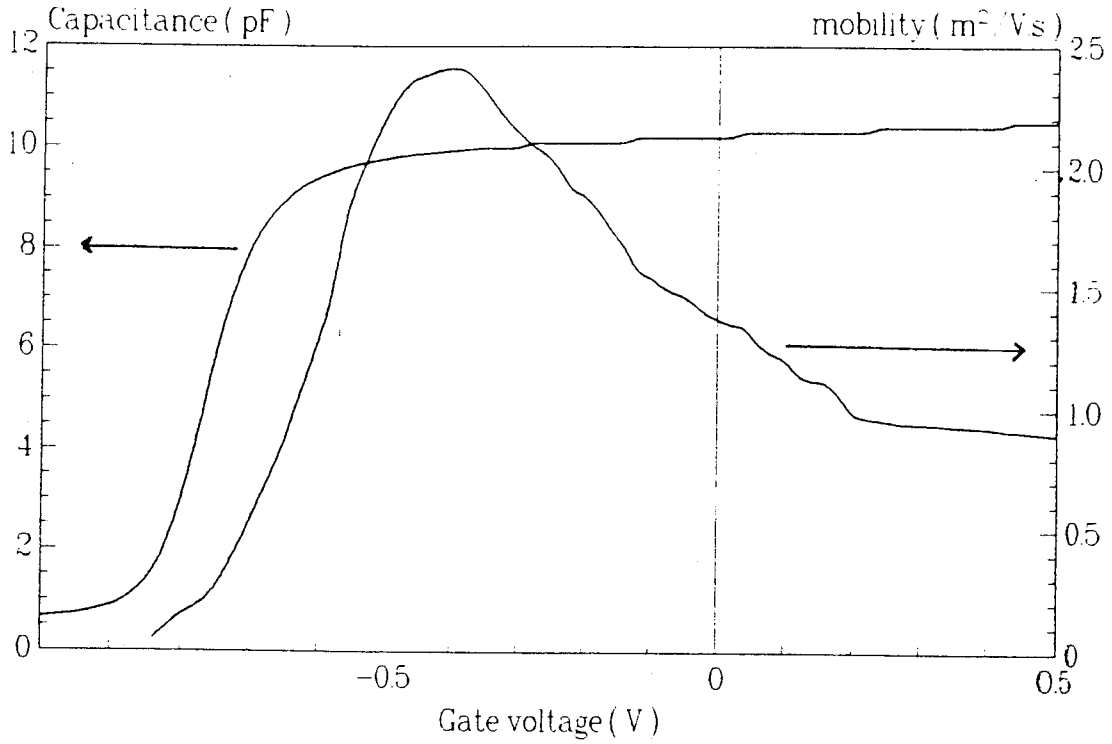


Figure 7: Gate capacitance (pF) and differential mobility ($m^2/V.s$) at 77K versus gate voltage (V).

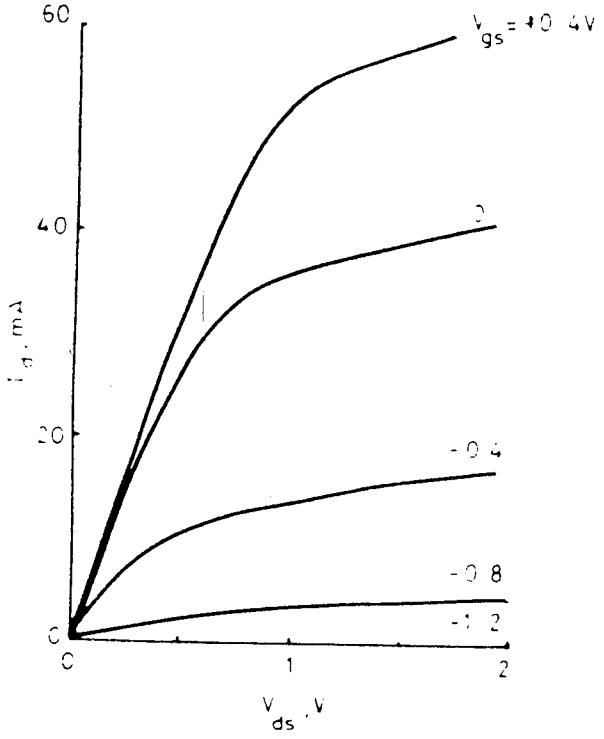


Figure 8: DC characteristic of a 0.4 μm gate MM-HEMT.

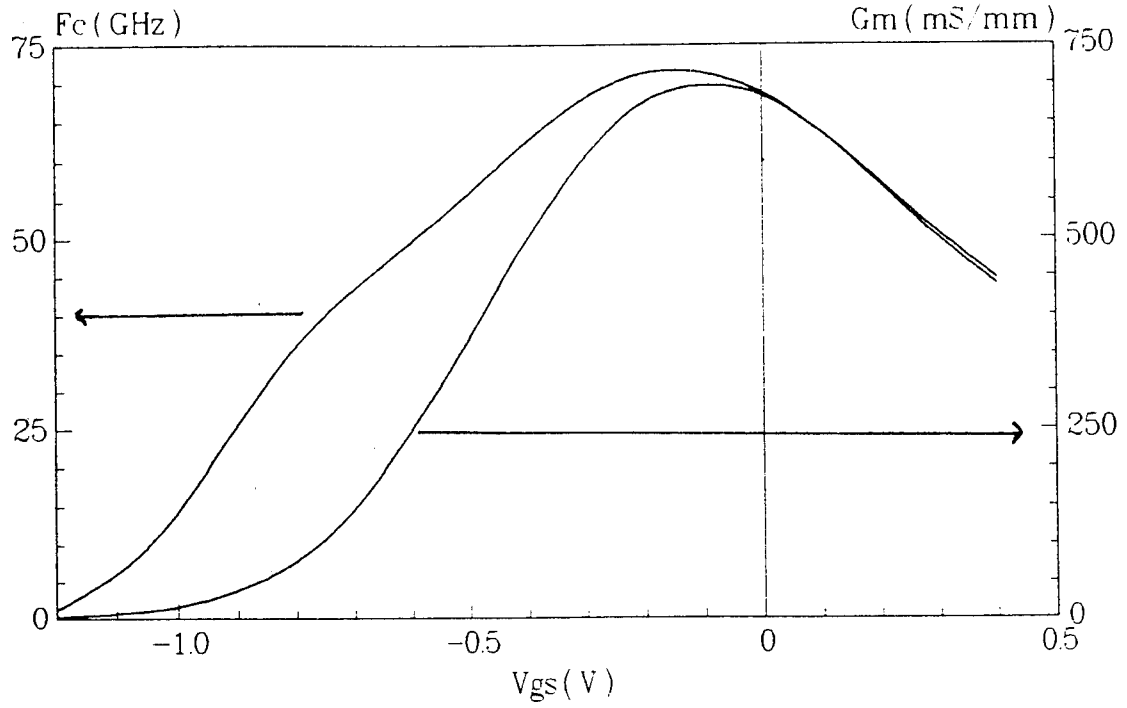


Figure 9: Intrinsic cutoff frequency and transconductance versus gate voltage.

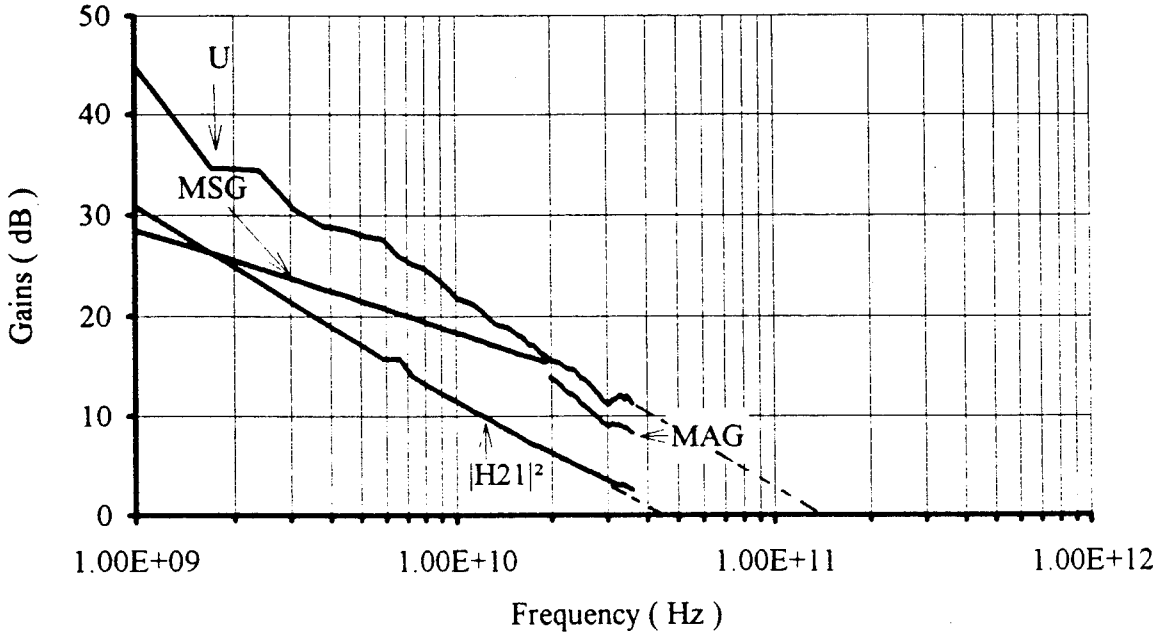


Figure 10: Extrinsic current gain ($|H_{21}|^2$), maximum stable gain (MSG), maximum available gain (MAG) and unilateral gain (U) plotted versus frequency.

InP-BASED HEMTS FOR THE REALIZATION OF ULTRA-HIGH EFFICIENCY MILLIMETERWAVE POWER AMPLIFIERS

M. Matloubian, L. Larson, A. Brown, L. Jelloian, L. Nguyen, M. Lui, T. Liu, J. Brown, and M. Thompson

Hughes Research Laboratories
Malibu, CA 90265

W. Lam, and A. Kurdoghlian

Hughes Aircraft Company - Gallium Arsenide Operations
Torrance, CA 90509

R. Rhodes, M. Delaney, and J. Pence

Hughes Aircraft Company - Space and Communications Sector
El Segundo, CA 90245

I. INTRODUCTION

Although they are acknowledged as the premiere devices for the realization of millimeterwave low-noise amplifiers [1], InP-based GaInAs channel HEMTs have not until recently demonstrated high performance for high power amplifiers in the millimeterwave frequency range. This is due to their historically lower gate-to-drain and drain-to-source breakdown voltages compared with GaAs-based pseudomorphic HEMTs (PHEMTs) [2]. These low breakdown voltages could result in a requirement for lower drain voltage operation, and hence lower drain and power-added efficiency. However, if the breakdown voltage limitations can be overcome, these devices are excellent candidates for high-efficiency millimeterwave power amplifier applications due to their high f_T and high power gain.

We have conducted a systematic effort to improve the breakdown voltage of these devices without compromising their high frequency performance, and have demonstrated millimeterwave circuit results that are comparable to or exceed those of the best GaAs-based PHEMTs in the critical area of power-added efficiency and output power. This improvement was accomplished by a combination of developments in material growth and device design and fabrication.

0-7803-0894-8/93 $3.00

II. InP-BASED HEMT MATERIAL AND DEVICE DESIGN

A materials cross-section of a typical InP-based millimeterwave power HEMT appears in Fig. 1. An ideal power HEMT would exhibit a high breakdown voltage and saturated channel current (for high power density) combined with a high electron mobility and velocity (for high millimeterwave power gain). The material profile for these devices is designed to simultaneously optimize these sometimes conflicting goals.

An improved breakdown voltage is accomplished by increasing the Al percentage in the Schottky layer beyond the lattice-matched 48% to between 60 and 70%. The resulting increase in the Schottky barrier height (from ≈ 0.5 eV to ≈ 0.8 eV) results in a typical increase in BVgds from 4 V to over 7 V, as shown in Fig. 2 [3]. The growth of pseudomorphic strained AlInAs layers for power HEMT applications is detailed in [4]. The breakdown voltage is further improved by reducing the thickness of the narrow bandgap GaInAs channel layer as far as possible before channel mobility and sheet charge density begin to degrade. δ-doping is used within the channel to supply additional charge and in order to minimize the density of ionized donors near the gate and increase the current density at a given breakdown voltage [5].

InP has recently been employed in the channels of these devices as a replacement for GaInAs, in the hope of improving their breakdown voltages without compromising their gain at millimeterwave frequencies [6]. Although the use of InP alone for the channel may result in a drop in f_T because of the lower electron velocity — at low drain bias voltages — the *combination* of GaInAs and InP in the channel has been shown to increase the breakdown voltage without degrading the rf performance of the device [7].

The material layers grown for our InP-based millimeterwave power HEMTs have sheet charge densities between 3.5×10^{12} cm^{-2} and 4.5×10^{12} cm^{-2}, and room temperature Hall mobilities between 8500 and 10,000 $cm^2/Vsec$. The resulting 0.15 μm devices exhibit peak f_T's of approximately 150 GHz, with g_m's of 650 mS/mm, and drain-source breakdown voltages (measured at 1 mA/mm current) of 6-9 V. This combination of characteristics is nearly ideal for millimeterwave power amplifier applications.

The devices are fabricated with a planar process. Source and drain ohmic contacts are fabricated with a AuGe/Ni/Au alloy with drain-to-source

spacing of 2 μm. Boron ion implantation was used for device isolation. The gates have a T-shaped cross-section and a footprint of 0.15 μm. The wafers are lapped and polished, and through substrate vias are wet chemically etched.

III. InP-BASED HEMT MILLIMETERWAVE POWER AMPLIFIERS — DISCRETE DEVICE AND MMIC RESULTS

Millimeterwave power amplifiers are being actively developed in V-band frequency range for a variety of applications. We have recently demonstrated extremely high power-added efficiencies from our InP-based GaInAs channel HEMTs at 60 GHz. We have obtained an output power of over 155 mW from a single device, with power-added efficiency of 30% and linear gain of 8 dB. When power combined in hybrid form, the devices produced 288 mW at a power-added efficiency of 20.4%. By using a GaInAs/InP composite channel structure we have obtained an output power of 174 mW with power-added efficiency of 30%. A summary of our InP-based HEMT results at 60 GHz appears in Fig. 3, along with a comparison to competing results based on GaAs PHEMT and HBT technologies [8-10]. The improved performance of the InP-based HEMT devices is attributed to the higher gain that these devices exhibit compared to a comparable GaAs-based PHEMT. This significantly improves the power-added efficiency of the amplifier, despite the fact that the breakdown voltage may be somewhat lower than a GaAs-based PHEMT of comparable gate length.

We have also developed a family of 44 GHz InP-based HEMT MMIC power amplifiers for satellite communications applications. A single-stage design, shown in Fig. 4, produces 250 mW of output power at 44 GHz, with 33% power-added efficiency and 8.5 dB small-signal gain. The HEMT is a single 450 μm x 0.15 μm device, producing 550 mW/mm of gate periphery. The output power and power-added efficiency results of this amplifier is shown in Figure 5. By comparison, the best GaAs MESFET MMIC results at this frequency produced 15% power-added efficiency and 22.5 dBm output power [11]. Also by comparison, a PHEMT power amplifier recently demonstrated 32% power-added efficiency and 27 dBm of output power at 35 GHz [12]. So the performance of these MMICs is comparable to or exceeds that of the best GaAs results reported to date.

IV. CONCLUSIONS

InP HEMT power amplifier technology is still relatively immature compared with GaAs PHEMT technology. Despite this, the results presented here demonstrate the impressive potential of the technology for future millimeterwave systems that require high efficiency power amplification. Future work on the technology will involve further improvements in breakdown voltage, device yield, and reliability.

V. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of Dr. Paul Greiling and Dr. Richard Reynolds of the Hughes Research Laboratories.

VI. REFERENCES

[1] K. Duh, P.C. Chao, P. Ho, M. Kao, P. Smith, J. Ballingall, and A. Jabra, "High-Performance InP-based millimeter-wave low-noise amplifiers," 1989 IEEE MTT-S Symposium Digest, pp. 805-808.

[2] U.K. Mishra and A.S. Brown, "GaInAs/AlInAs HEMT Technology for millimeter-wave applications," IEEE GaAs IC Symposium Digest, pp. 97-99, 1988.

[3] J. Brown, A. Brown, S. Rosenbaum, A. Schmitz, M. Matloubian, L. Larson, M. Melendes, M. Thompson, J. Henige, D. Cramton, and J. Pusl, "Study of the dependence of GaInAs/AlInAs power HEMT breakdown voltage on Schottky layer design and device layout," presented at 1993 IEEE DRC, Santa Barbara, CA.

[4] A. Brown, J. Brown, J. Henige, R. Wilson, R. Metzger, M. Melendes, M. Matloubian, L. Jelloian, and L. Larson, "The growth of high performance GaInAs-AlInAs power HEMT devices by MBE: Si segregation and activation and strained layer structural quality," June, 1993, EM, Santa Barbara, CA .

[5] M. Matloubian, L. Jelloian, A. Brown, L. Nguyen, L. Larson, M. Delaney, M. Thompson, R. Rhodes, and J. Pence, "V-band high-efficiency high-power AlInAs/GaInAs/InP HEMTs," 1993 IEEE MTT-S Symposium Digest, pp. 535-537.

[6] L. Aina, M. Burgess, M. Mattingly, J. M. O'Connor, A. Meerschaert, M. Tong, A. Ketterson, and I. Adesida, "0.33-μm gate-length millimeter-wave InP-channel HEMT's with high f_T and f_{max}," IEEE Elect. Dev. Lett., vol. 12, no. 9, pp. 483-485 Sept. 1991.

[7] M. Matloubian, L. Jelloian, M. Lui, T. Liu, L. Larson, L. Nguyen, and M. Le, "GaInAs/InP composite channel HEMTs," 1993 IEEE Device Research Conference, Santa Barbara, CA.

[8] K.L. Tan, D.C. Streit, R.M. Dia, S.K. Wang, A.C. Han, P.-M.D. Chow, T.Q. Trinh, P.H. Liu, J.R. Velebir, and H.C. Yen, "High-power V-band pseudomorphic InGaAs HEMT," IEEE Electron Device Lett., vol. 12, no. 5, pp. 213-214, May 1991.

[9] P.M. Smith, M.Y. Kao, P. Ho, P.C. Chao, K.H.G. Duh, A.A. Jabra, R.P. Smith, and J.M. Ballingall, "A 0.15 μm gate-length pseudomorphic HEMT," 1989 IEEE MTT-S Symposium Digest, pp. 983-986.

[10] J.A. Higgins, "Heterojunction bipolar transistors for high efficiency power amplifiers," 1988 GaAs IC Symposium Digest, pp. 33-36.

[11] G. Hegazi, H-L. A. Hung, J.L. Singer, F.R. Phelleps, A.B. Cornfeld, T. Smith, J.F. Bass, H.E. Carlson, and H.C. Huang, "GaAs molecular beam epitaxy monolithic power amplifiers at U-band," IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium Digest, pp. 121-125, 1989.

[12] A. Kurdoghlian, C.S. Wu, W. Yau, J. Chen, M. Hu, C. Pao, and D. Bosch, "The demonstration of Ka-band multi-functional MMIC circuits fabricated on the same PHEMT wafer with superior performance," IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium, pp. 97-98, 1993.

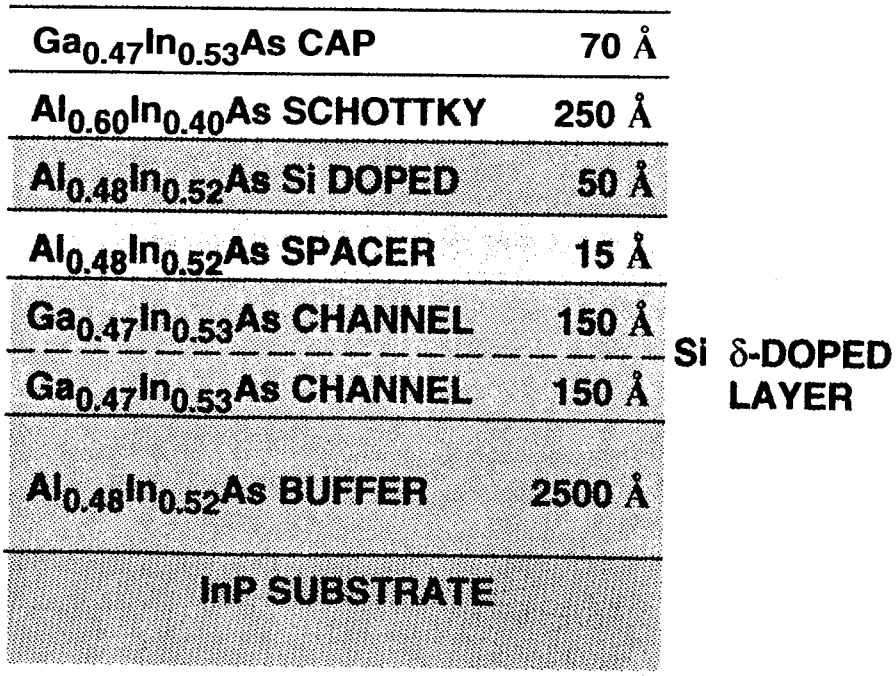


Fig. 1 - Material Profile of the InP-based millimeterwave power HEMTs.

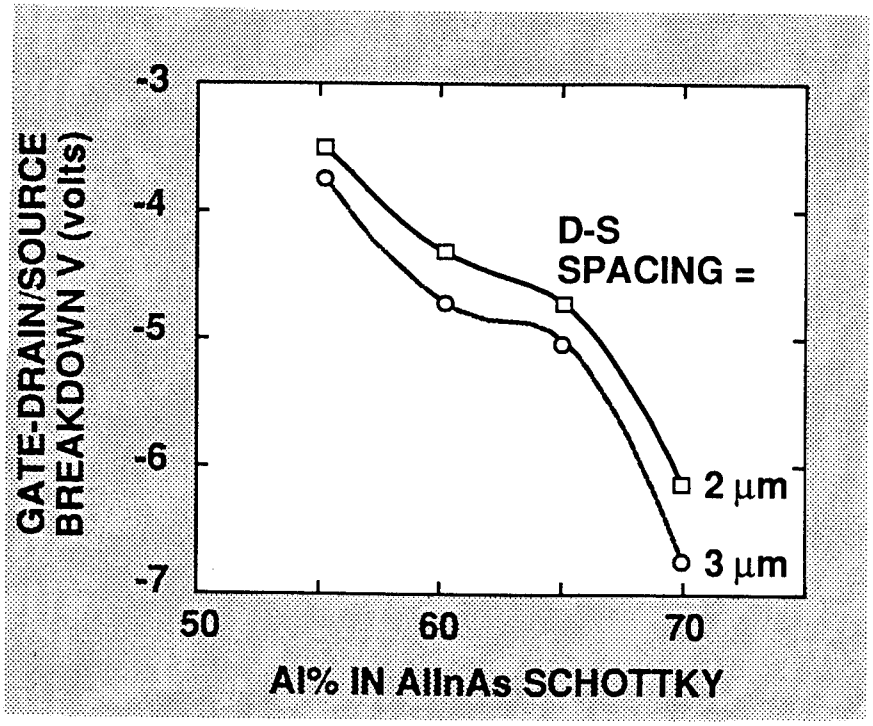


Fig. 2 - Improvement in gate-to-drain-source breakdown (BV_{gds}) as a function of Al% in the Schottky layer [3].

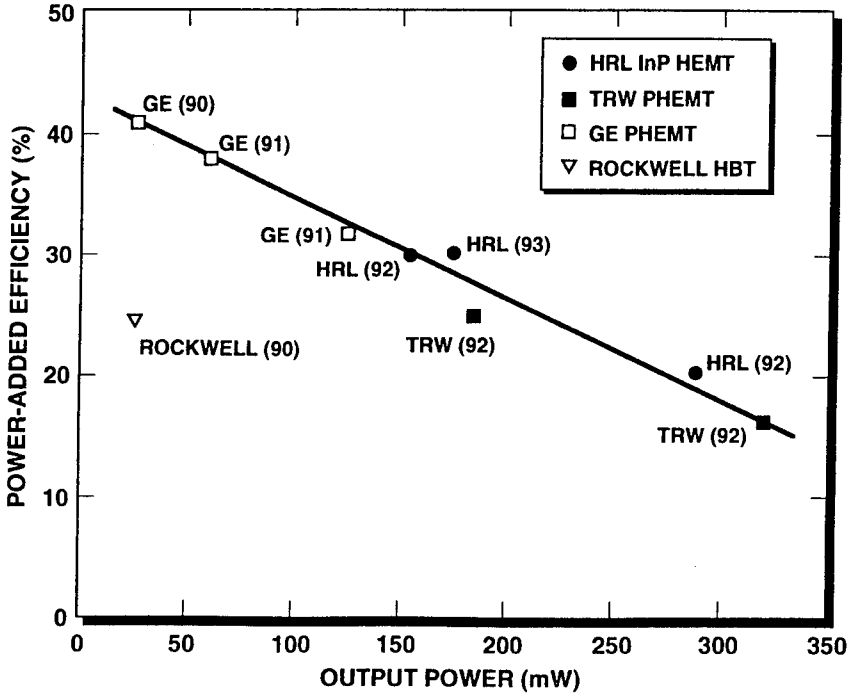


Fig. 3 - Comparison of the state-of-the-art power performance of 60 GHz amplifiers.

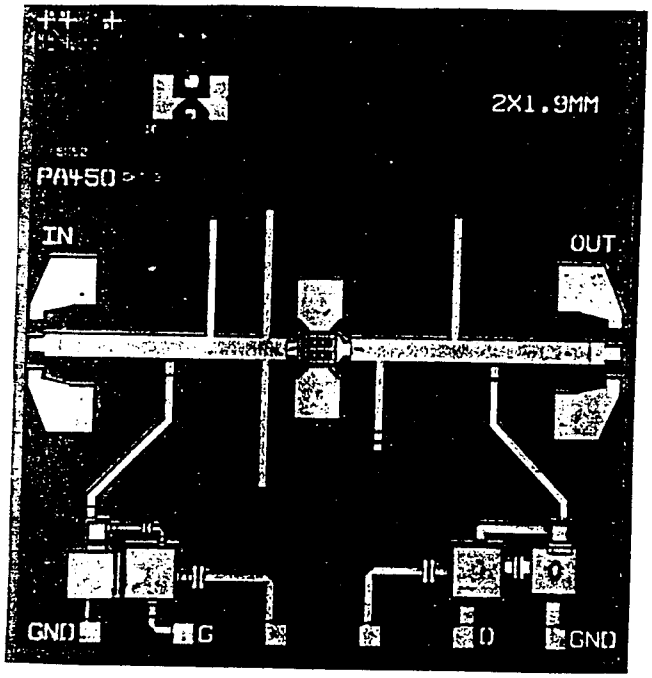


Fig. 4 - Microphotograph of single-stage 44 GHz InP-based HEMT MMIC amplifier.

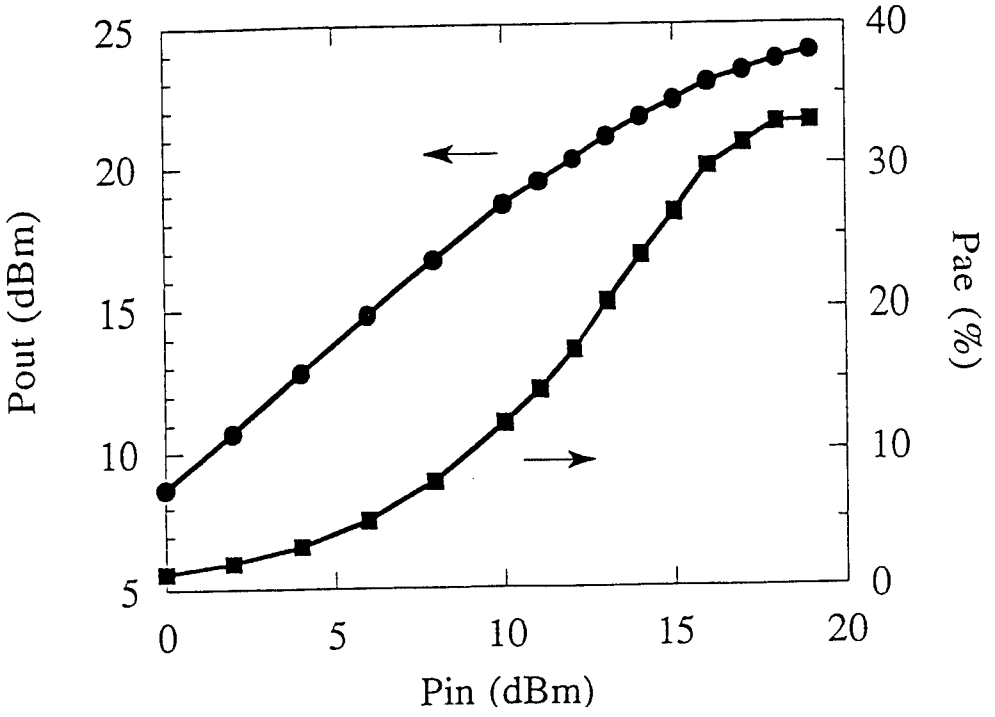


Fig. 5 - Output power and power-added efficiency of the circuit of Fig. 4.

p^+- Thin Surface Layer Schottky-Barrier Enhanced High Speed Pseudomorphic $Al_{0.25}Ga_{0.75}As/In_{0.15}Ga_{0.85}As$ and $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$ MODFETs

J. Dickmann[1], M. Berg[2], Th. Hackbarth[1], R. Deufel[1], H. Daembkes[1], F. Scholz[3], M. Moser[3]

[1] Daimler-Benz Research Center Ulm, D-89081 Ulm, Germany

[2] Technical University Darmstadt, Institute for High Frequency Technology, D-64283 Darmstadt, Germany

[3] University Stuttgart, 4 th Institute of Physics, D-70569 Stuttgart, Germany

Abstract

In this paper we report about our investigation of the introduction of a highly p-doped very thin surface layer in combination with GaInP as the wide bandgap material in order to improve the breakdown voltage of GaAs based pseudomorphic MODFETs without deteriorating other device performances. In order to proof the quality of this approach we compare the device performance of the new device with that of a conventional AlGaAs/InGaAs MODFET also having a p^+-surface layer.The device performances of AlGaAs/InGaAs and GaInP/InGaAs devices with 1.8µm gate length are, g_{mmax}=224mS/mm, 200mS/mm, I_{DSmax}=300mA/mm, 400mA/mm, V_{BrDG} (I_G=1mA/mm) = 10V, 14V, f_T=15GHz, 12GHz, f_{max}=59GHz, 42GHz, respectively.

I. Introduction

GaAs based HFETs are key elements in transmitter and receiver components for millimeter wave sensor and communication systems. In the past pseudomorphic AlGaAs/InGaAs MODFETs have demonstrated their great potential as a microwave and millimeter-wave device for high gain and low noise applications. To further optimize the transmitter modules, the output power should be increased, which necessitates the improvement of the power performance of the present MODFET devices. In order to improve the power performance of a MODFET both, the current driveability and the breakdown voltage have to be increased. The well known approaches to improve the breakdown voltage in

0-7803-0894-8/93 $3.00

conventional AlGaAs/InGaAs MODFETs without deteriorating the high speed performance was to use the double-recess approach [1], or to use high bandgap insulator AlGaAs layers [2].
In this paper we report about the combination of two approaches to further improve the breakdown voltage of GaAs based MODFETs using an InGaAs channel layer. One approach follows the basic idea, that the breakdown voltage of a MODFET device can be improved by putting a thin layer of material with a high breakdown field under the gate. According to this idea we use very thin highly p-doped surface layers under the gate as first proposed in [3,4]. The second approach is to substitute AlGaAs by a different suitable wide bandgap material. Considering the bandgap, lattice constant and cross over point from a direct to indirect semiconductor, $Ga_XIn_{1-X}P$ seems to be a well suitable material. In order to compare the effectiveness of a combined application of both approaches on a MODFET, in this paper we compare the device performance of a conventional AlGaAs/InGaAs MODFET with that of a GaInP/InGaAs MODFET both having a very thin p^+-surface layer.

II. Material growth and device layer structure

The $Al_{0.25}Ga_{0.75}As/In_{0.15}Ga_{0.85}As$ layer structure was grown by conventional molecular beam epitaxy (MBE) in a varian Gen II system on a semi insulating (100) GaAs substrate. Silicon is used for n-doping and berillium is used for p-doping. Starting from the substrate the layer structure is as follows (see Fig.1): 100nm GaAs buffer, a 10 period superlattice (2nm/2nm) AlAs/GaAs as an impurity filter, 500nm GaAs buffer in order to smoothen the growth front, 15nm $In_{0.15}Ga_{0.85}As$ channel layer, 3nm GaAs pre-spacer in order to improve the interface quality between the two ternary alloys InGaAs and AlGaAs, a 2nm $Al_{0.25}Ga_{0.75}As$ spacer, a 15nm thick homogeniously doped $Al_{0.25}Ga_{0.75}As$ layer doped to $N_D=5x10^{18}cm^{-3}$ followed by a 15nm thick homogeniously doped $Al_XGa_{1-X}As$ layer doped to $N_D=5x10^{18}cm^{-3}$ (x=0 at the cap layer). The layer structure is completed with a 10nm thick GaAs p^+-layer which is doped to $N_A=5x10^{18}cm^{-3}$. The grading in the barrier layer was introduced in order to enable the fabrication of low resistive ohmic contacts and to suppress the heterojunction in the access regions, which contributes to lower source- and drain-resistances. The p-doping concentration was used so that the surface layer just modifies the surface potential distribution around the gate electrode and not to change dramatically the Schottky-Barrier height. The basic idea that lies behind is to generate a quasi surface depleted field distribution around the gate electrode as utilized with the surface depleted cap layer concept in InAlAs/InGaAs MODFETs [5].

GaAs	$N_A=5x10^{18}cm^{-3}$	10nm
GaAs	$N_D=3x10^{18}cm^{-3}$	10nm
$Ga_{0.5}In_{0.5}P$	$N_D=3x10^{18}cm^{-3}$	20nm
$Ga_{0.5}In_{0.5}P$	Spacer	2nm
$In_{0.15}Ga_{0.85}As$	Channel	12nm
GaAs	undoped Buffer	
s.i. (100) GaAs substrate		

a)

GaAs	$N_A=5x10^{18}cm^{-3}$	10nm
$Al_xGa_{1-x}As$	$N_D=5x10^{18}cm^{-3}$	10nm
$Al_{0.25}Ga_{0.75}As$	$N_D=5x10^{18}cm^{-3}$	15nm
$Al_{0.25}Ga_{0.75}As$	Spacer	3nm
Ga As	Spacer	2nm
$In_{0.15}Ga_{0.85}As$	Channel	15nm
GaAs Buffer + AlAs/GaAs SL		
s.i. (100) GaAs substrate		

b)

Fig.1: Layer structure of the **a)** $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$, **b)** $Al_{0.25}Ga_{0.75}As/In_{0.15}Ga_{0.85}As$ MODFET

From Shubnikov de Haas (SdH) measurements a 2DEG sheet carrier concentration of $n_S= 1.5x10^{12}cm^{-2}$ was determined on comparable test structures.
The $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$ MODFET layer structure was grown by low pressure metal organic chemical vapour deposition (MOCVD) on a semi-insulating (100) GaAs substrate. More details about the growth procedure is given in [6].
From Shubnikov de Haas (SdH) measurements a 2DEG sheet carrier concentration of $n_S= 1.25x10^{12}cm^{-2}$ was determined for this layer structure, proofing the high quality of the sample and the fact that a reasonable 2DEG concentration can be generated in such a layer structure.

III. Device Fabrication

The sequence of the fabrication process is shown in Fig.2 and is as follows :
The first step in course of the process sequence is the fabrication of the T-gate. For the fabrication of the T-gate structure as well as for the ohmic contacts, we use multi layer sequences consisting of Az-resists and germanium intermediate layers. For the T-gate structure we use a five layer sequence where for this work the gate length is defined into the top Az-resist via optical contact lithography. The generation of submicron gates using this process is in principle also possible. This generated pattern is then transferred into the subsequent layers by exploiting

the RIE capability of etching highly material selective Az-resists over germanium layers as described in [7]. The total process yields a final gate length of L_G=1.8µm. In case of the ohmic contact fabrication we use a three layer sequence.

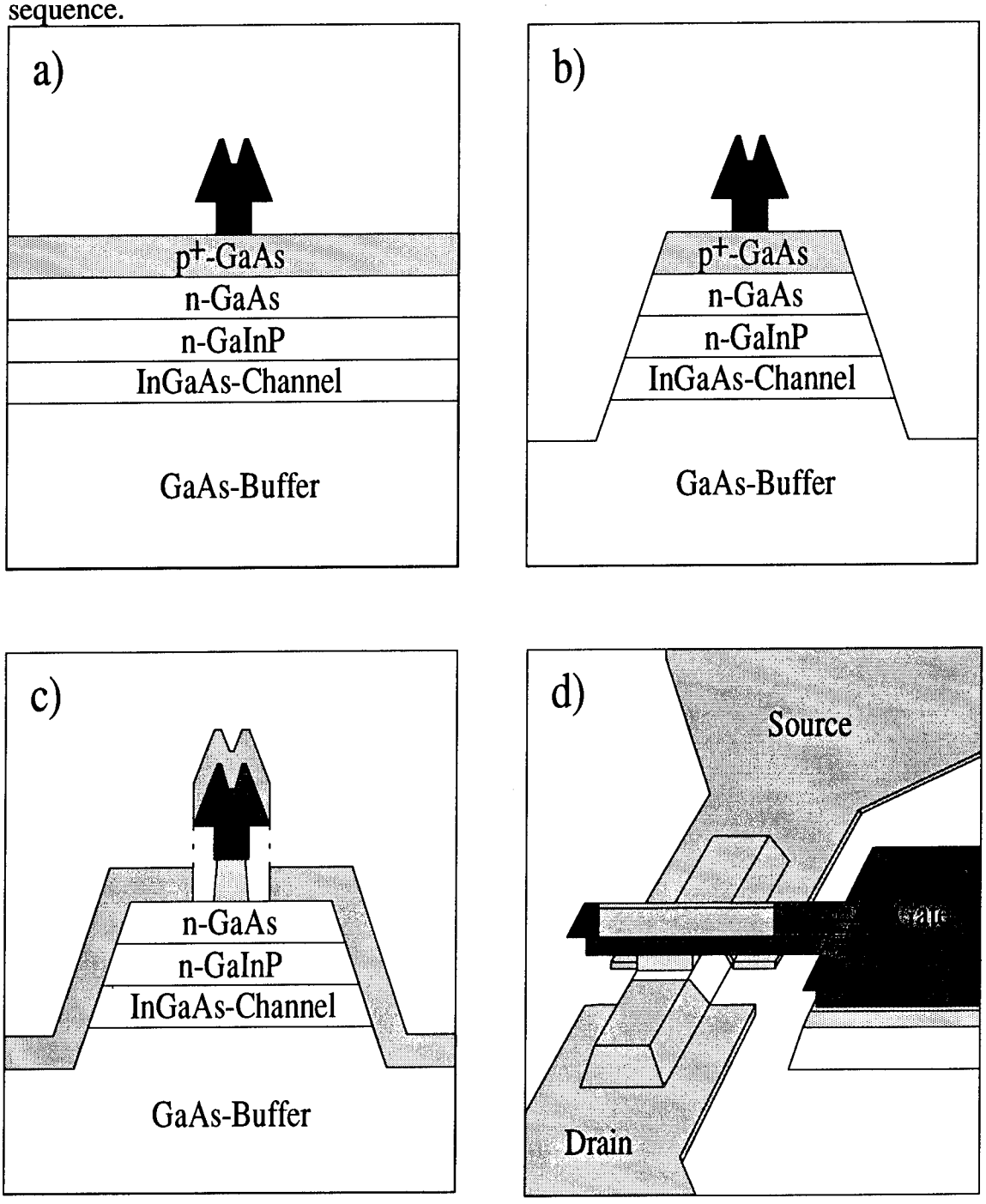


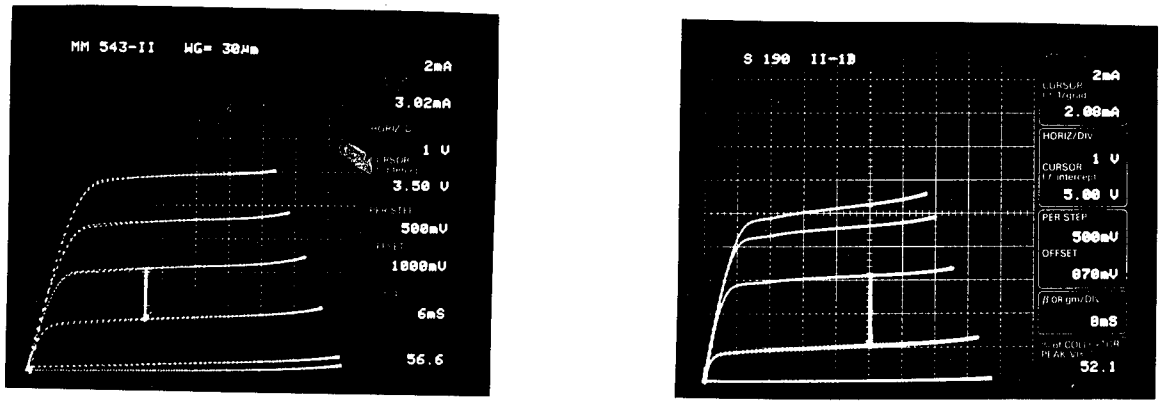
Fig.2: Process sequence of the p^+-surface layer MODFETs

The gate metalization is Ti/Pt/Au (15nm/30nm/600nm). One important advantage of the p^+-layer approach is, that there is no need for a gate recessing, since the effective distance between the gate and the center of the 2DEG concentration (d_{eff}) is adjusted by epitaxy during the growth making device processing more comfortable. The next step is device isolation. In order to isolate the gate bondpad from the active FET area in order to minimize the parasitic contribution of the gate pad capacitance to the total gate-source capacitance and to avoid a cross contacting of the p-n junction which would cause a large gate leakage current, a double-mesa approach is used as first described in [8]. With this approach an air gap is created under the link between the bondpad and the active gate. For the etching of GaAs a standard H_2SO_4:H_2O_2:H_2O etchant and for the GaInP layer an HCL:H_3PO_4 etchant is used.
Subsequent to this, ohmic contact fabrication is done. The inherent disadvantage of the p^+-approach is that it gives only minor space to take steps to reduce the parasitic source and drain resistances. Hence, in order to minimize the total access resistance the T-bar of the gate is used to self-align the ohmic contacts. Prior the deposition of the Ge/Ni/Au ohmic metalization the entire p^+-GaAs layer is removed. Doing so allows the deposition of the metal directly on highly n-doped GaAs. The undercut of the T-bar with respect to the gate foot print represents the gate-source/drain distance and is adjusted to be 0.3µm. The total source-drain distance is L_{SD}=2.4µm.

IV. Device Performance

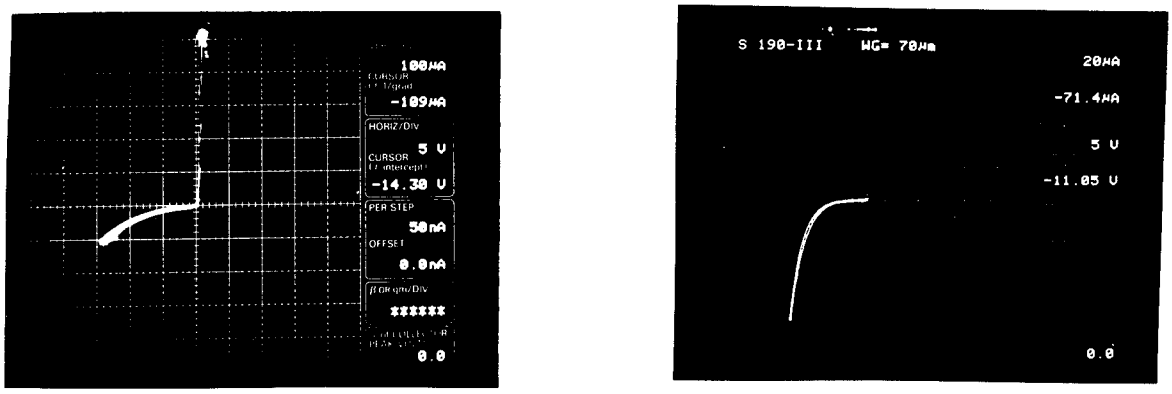
DC-Performance:
The typical IV-characteristic of both types of MODFETs is shown in Fig.3. The maximum positive gate voltage is defined by the onset of gate leakage current and is V_{GStop}=0.87V, 1V for the MODFET with the AlGaAs and GaInP barrier layers, respectively. In both cases the devices show good pinch-off and saturation behaviour which is advantageous for power gain and output-power operation. With values of g_d<5mS/mm in both cases the output conductance is extremely small and is mainly caused by the large aspect ratio (L_G/d_{eff}) in these devices. The device performances of AlGaAs/InGaAs and GaInP/InGaAs devices with 1.8µm gate length are, g_{mmax}=224mS/mm, 200mS/mm, I_{DSmax}=315mA/mm, 400mA/mm and V_{BrDG} (I_G=1mA/mm) = 10V, 14V, respectively. Comparing the DC-results of the two different devices it becomes obvious, that the transconductances are about the same while the saturation current is slightly higher in the case of the GaInP device. However, the most distinct difference can be observed for the breakdown voltage. The gate-drain breakdown voltage is

about 4V higher in the case of the GaInP device supporting the assumption of GaInP being an attractive substitute for AlGaAs for GaAs based pseudomorphic power MODFETs (compare Fig.4). The most dramatic difference is observed in the drain-source burn-out voltage. The maximum values for the AlGaAs devices were measured to be around 12 V while for the GaInP devices burn-out voltages in excess of 20V were measured, [6].



a) b)

Fig.3: Typical IV-characteristic of a **a**) $Ga_{0.5}In_{0.5}P/ In_{0.15}Ga_{0.85}As$ MODFET. $L_G x W_G$=1.8µmx30µm. V_{GStop}=1.0V, **b**)$Al_{0.25}Ga_{0.75}As/ In_{0.15}Ga_{0.85}As$ MODFET. $L_G x W_G$=1.8µmx35µm. V_{GStop}=0.87V



a) b)

Fig.4: Typical gate-drain IV-characteristic of a **a**)$Ga_{0.5}In_{0.5}P/ In_{0.15}Ga_{0.85}As$ MODFET, **b**)$Al_{0.25}Ga_{0.75}As/ In_{0.15}Ga_{0.85}As$ MODFET.

RF-Performance:

Microwave characterization was performed from 0.45-26.5 GHz using an on wafer Cascade Microtech probe station and an HP8510 network analyzer. The cut-off frequencies f_{max}, f_T were determined by extrapolating the maximum available gain (MAG) and the current gain ($|h_{21}|^2$) with a -20dB/decade roll off. In order to minimize the amount of parasitc charge that is modulated and in order to reduce the gate resistance, we use in this work the concept of a very thin p^+-surface layer with a low resistive Ti/Pt/Au T-gate electrode instead of using a gate electrode being formed by a thick p^+-layer as commonly used for JFET´s, [8].

For devices with a gate geometry of $L_G x W_G$=1.8µmx80µm (2x40µm) the cut-off frequencies were determined to be f_T=15GHz, 12GHz and f_{max}=59GHz, 42GHz for the AlGaAs and GaInP device, respectively. The high ratios of f_{max}/f_T >3 can be explained by relatively high ratios of C_{GS}/C_{GD}>10 and g_m/g_d>20 (best values). These unusally high ratios for single-heterojunction MODFETs can be explained by the combination of two effects. First, the large gate-length leads to naturally high aspect ratios resulting in excellent small output conductance. Second, the layer structure used for the p^+-surface layer approach is designed to reduce the peak electric field at the drain side edge of the gate-electrode, thus allowing the depletion region to extend towards the drain with increased drain bias. This results in a decrease of C_{GD} with increased drain bias, [5]. For a more detailed device characterization, the S-parameters were measured as a function of gate-source and drain-source bias, while the drain-source current was measured. From bias dependent measurements we found, that f_{max} started to saturate at around V_{DS}= 3.5V, 1.6V and remained nearly unchanged up to 7V for the GaInP device and 5V for the AlGaAs device, reflecting the higher breakdown voltages of the GaInP device. Fig.5 and Fig.6 show the current gain cut-off frequency as a function of the measured drain-source saturation current. As can be seen, the GaInP device shows a clear distinct maximum of f_T=12GHz for a value of I_{DS}=140mA/mm, while the AlGaAs device shows a broad plateau over I_{DS} with a maximum value of f_T=15GHz for a value of I_{DS}=100mA/mm. This behaviour reflects the measured DC results, where the GaInP device showed the slightly higher saturation current but lower transconductance. The wider plateau for the AlGaAs device can be explained with the different modulation efficencies of the different devices [9]. The slightly higher 2DEG concentration of the AlGaAs compared to that of the GaInP device allows it to operate more efficiently at medium and high current levels. This finally yields to the observed higher value of f_T and the broader plateau over I_{DS}.

A comparison of the performance of both devices shows, that the GaInP device results in similar high speed performance, but shows a much higher breakdown

voltage and thus a better behaviour at high voltage operation as the AlGaAs device, which makes the GaInP device an attractive candidate for power application up into the millimeter-wave range.

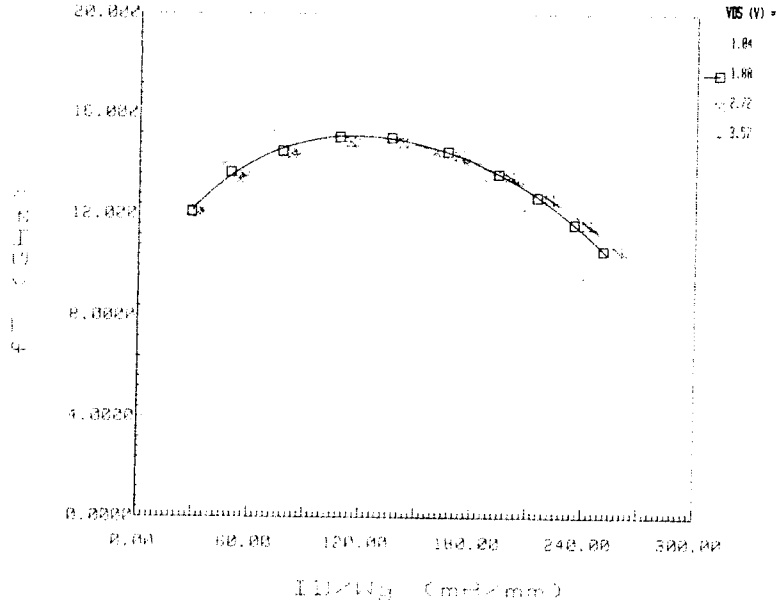


Fig.5: Current gain cut-off frequency of the $Al_{0.25}Ga_{0.75}As$/ $In_{0.15}Ga_{0.85}As$ MODFET as a function of V_{DS}. $L_G x W_G$=1.8μmx80μm.

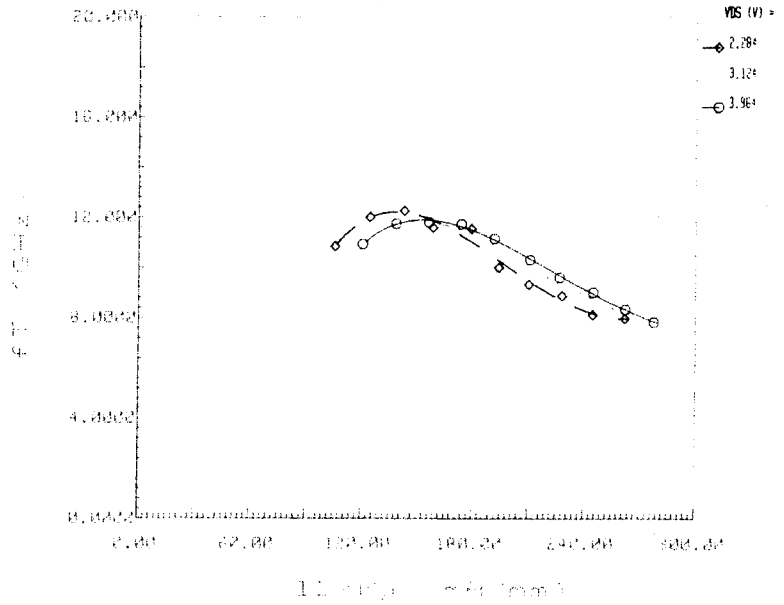


Fig.6: Current gain cut-off frequency of the $Ga_{0.5}In_{0.5}P$/ $In_{0.15}Ga_{0.85}As$ MODFET as a function of V_{DS}. $L_G x W_G$=1.8μmx80μm.

Conclusion

In conclusion we have compared the device performance of standard $Al_{0.25}Ga_{0.75}As/In_{0.15}Ga_{0.85}As$ and $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$ MODFETs with a p^+-surface layer. In order to reduce parasitic gate resistance and charge modulation, we used surface depleted p^+-layers. From our investigation we found that the GaInP devices showed similar high speed performance but better breakdown voltages. This distinct advantage over AlGaAs devices makes the $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$ MODFETs an attractive candidate for power application.

Acknowledgment: We would like to thank Joachim Hugo for his expert work in Shubnikov de Haas measurements and Holger Koser for Hall measurements.

References

[1] L.F.Lester, P.M.Smith, P.Ho, P.C.Chao, R.C.Tiberio, K.H.Duh, E.D.Wolf, "0.15 μm gate-length double recess pseudomorphic HEMT with f_{max} of 350GHz", IEEE IEDM Tech.Dig., pp.172-175, 1988

[2] H.Hida, A.Okllamato, H.Toyoshima, K.Ohata, "A high current driveability i-AlGaAs/n-GaAs doped channel MIS-like FET (DMT)", IEEE Electron Device Letters, vol.EDL-7, No.11, pp.625-627, 1986

[3] J.M.Shannon, " Control of Schottky Barrier Heigth Using Highly Doped Surface Layers", Solid-State Electronics, vol.19, pp.537-543, 1976

[4] K.L.Priddy, D.R.Kitchen, J.A.Grzyb, C.W.Litton, T.S. Henderson, C.K.Peng, W.F.Kopp, H.Morko`c, "Design of enhanced Schottky-barrier AlGaAs/GaAsMODFET's using highly doped p^+ surface layers", IEEE Trans. on Electr. Devices, vol.ED-34, No.2, pp.175-179, 1987

[5] J. Dickmann, H.Dämbkes, H.Nickel, R. Lösch, J. Böttcher, H. Künzel, Influence of Surface Layers on the RF-Performance of AlInAs-GaInAs FET´s", IEEE Microwave and Guided Wave Letters, vol.2, No.11, pp.472-74, 1992

[6] J. Dickmann, M. Berg, A. Geyer, H. Daembkes, F. Scholz, M. Moser, " High Speed $Ga_{0.5}In_{0.5}P/In_{0.15}Ga_{0.85}As$ Pseudomorphic p+- Schottky-Enhanced Barrier Heterojunction-Fieldeffect Transistors (HFETs)", to be published in Tech. Dig. GaAs & Rel. Compounds, Freiburg, Germany, August 1993

[7] J.Dickmann, A.Geyer, H.Daembkes, H.Nickel, R.Lösch, W.Schlapp, Fabrication of Low Resistance Submicron Gates in Pseudomorphic MODFETs Using Optical Contactlithography", J.Electrochem.Soc., vol.138, No.2, pp.491-493, 1991

[8] D. Wake, A.W. Livingstone, D.A. Andrews, G.J. Davis, " A self-aligned $In_{0.53}Ga_{0.47}As$ Junction Field Effect Transistor Grown by Molecular Beam Epitaxy", IEEE Electron Device Letters, vol. ED-1 5, No.7, pp.285-287, 1984

[9] M.C. Foisy, P.J. Tasker, B. Hughes, L.F. Eastman, "The role of Inefficient Charge Modulation in Limiting the Current Gain Cut Off Frequency of the MODFET", IEEE Trans. Electron Devices, vol. ED-35, No.7, pp.871-877, 1988

ENCAPSULATED GaAs POWER MESFET

N. X. Nguyen, J. P. Ibbetson, J. C. Yen, M. H. Hashemi*, U. K. Mishra

Department of Electrical and Computer Engineering, University of California
Santa Barbara, CA 93106

*Motorola Incorporate, Phoenix Corporate Research Laboratories
Tempe, Arizona 85284

ABSTRACT

Utilizing a combination of Low-Temperature-Grown $Al_{0.3}Ga_{0.7}As$ passivation, overlapping gate, MOCVD regrown of the source/drain contacts, and isolation by boron ion implantation, we have fabricated a GaAs MESFET with completely encapsulated channel. Electrical characterization of the device shows that the early catastrophic on-state breakdown is suppressed. In contrast to the usually observed characteristics in GaAs MESFET, the breakdown locus of the device also has a positive slope. These improvements should allow us to bias the device at a higher operating point, thereby increasing the obtainable maximum output power in the device.

I. INTRODUCTION

The surface of the device has been recognized as a major source of many detrimental effects on the performance of GaAs MESFET. Various theoretical studies over the years have led to the conclusion that the surface played a significant role in the device characteristics[1-3]. In particular, the breakdown voltage and instability in the current-voltage family of curves are two surface related parameters that directly effect the power performance of a device. In recent years, the surface of the device in the vicinity of the gate has been indirectly addressed by the advent of the low-temperature-grown (LTG) passivation and overlapping-gate technology [4],[5]. Devices which utilized these technology have demonstrated dramatic improvement in the gate-drain breakdown voltage, which in turn drastically enhanced the power output by the devices [4-6]. However, little attention has been paid to the exposed surfaces near the source and drain regions. In this work, we have attempted to address this issue by developing a technology to fabricate MESFETs with an encapsulated channel; and from the study of these fabricated devices, we have gained a better understanding in the role of the surface in GaAs MESFETs.

II. DEVICE DESIGN AND FABRICATION

Figure 1 shows a schematic representation of the epitaxial structure of the samples used in this experiment. All the samples were grown by a Varian Gen II MBE system on

0-7803-0894-8/93 $3.00

semi-insulating LEC GaAs substrate. Standard growith procedure started with a smoothing buffer layer of 5000 Å of undoped GaAs grown at 600 °C, followed by the channel; then an arsenic-diffusion-barrier of 200 Å of AlAs. The substrate temperature was then lowered to 200 °C, where 2000 Å of LTG-$Al_{0.3}Ga_{0.7}As$ was grown. The samples were then annealed in-situ at 600 °C for 10 minutes.

Preparation for selective regrowth of the source/drain regions began with a blanket deposition of 1500 Å of SiO_2 by PECVD at 250 °C. The samples were then patterned with a source/drain mask by photolithography. The SiO_2 in the open areas were removed using reactive ion etching with plasma CF_4; followed by a selective wet etching of the LTG-$Al_{0.3}Ga_{0.7}As$ by citric acid. Then the AlAs was etched in diluted HF, and finally citric acid was used to etch about 300 Å into the GaAs channel layer. n++ GaAs was regrown in the open areas using MOCVD at 550 °C. After the MOCVD regrowth of the source/drain regions, the samples were patterned for isolation by multiple boron ion implantation. Subsequent standard processing steps to define the ohmic and gate metals have been reported elsewhere [6].

III. RESULT AND DISCUSSION

Figure 2 shows the characteristics of a device without complete encapsulation of the channel, the device has a gate-drain breakdown voltage of about 29 V. However, as shown in the drain current-voltage family of curves, the device has a catastrophic breakdown at V_{DS} = 17.5 V and V_{GS} = 0.5 V; therefore this is an inherent limitation of the voltage swing of the device for power performance. Continuous wave on-wafer measurement of the power of the device at 4 GHz is shown in Figure 3; the device delivered 620 mW/mm with a power-added-efficiency of 30 %.

In comparison, the preliminary device characteristics of a GaAs MESFET with an encapsulated channel is shown in Figure 4. Although the gate-drain breakdown voltage of this device is only 11 V; we were able to bias the device up to V_{DS} = 7 V and V_{GS} = -2 V for power measurement. From the drain current-voltage characteristics, we also observed that the device breakdown locus shows a definite positive slope. The device outputs 520 mW/mm with a PAE of 18 % at 4 GHz (Figure 6). Although the maximum output power of this device is lower than that of the above device, the limitation is due to the lower gate-drain breakdown voltage not the catastrophic breakdown. Therefore, we believed that we should be able to improve the power performance of the device by just improving the gate-drain breakdown voltage.

On another sample with higher gate-drain breakdown voltage (20 V), Figure 5; we again noticed that the catastrophic on-state breakdown is suppressed. The device also displays a positive slope of the breakdown locus. The regrowth contact resistance obtained from the TLM measurement is 0.7-0.8 Ω-mm. However, it should be noted that this quoted resistance included the resistance of the regrown layer along with the interfacial resistance.

IV. SUMMARY

We have proposed a technology for the fabrication of a GaAs MESFET with a completely encapsulated channel. Electrical characterization of these devices show that the catastrophic on-state breakdown voltage is suppressed, and the breakdown locus of the devices displays a positive slope. These DC characteristics allow the device to be biased at a higher operating point, thereby increasing the power obtainable from the device.

ACKNOWLEDGMENT

The authors would like to thank Bob Wilson at Hughes Research Laboratories, Malibu, California for the ion implantation. This work is supported by the Air Force Office of Scientific Research.

REFERENCES

[1] H. Itoh, K. Ohata, G. Hasegawa "Influence of the surfaceand the episubstrate interface on the drain current drift of GaAs MESFET's," IEEE Tran. Electron Devices, Vol. ED-28, No. 7, pp. 878-882, 1981

[2] W. R. Frensley "Power-limiting Breakdown effects in GaAs MESFET's," IEEE Tran. Electron Devices, Vol. ED-28, No. 8, pp. 962-970, 1981

[3] T. M. Barton, P. H. Ladbrooke "The role of the device surface in the high voltage behavior ofthe GaAs MESFET," Solid-State Electronics Vol. 29 No. 8, pp. 807-813, 1986

[4] L.-W. Yin, Y. Hwang, J. H. Lee, R. M. Kolbas, R. J. Trew, U. K. Mishra, "Improved breakdown voltage in GaAs MESFETs utilizing surface layers of GaAs grown at a low temperature by MBE," IEEE Electron Device Letters, Vol. 11, No. 12, pp. 561-563, 1990

[5] C.-L. Chen, F. W. Smith, B. J. Clifton, L. J. Mahoney, M. J. Manfra, A. R. Calawa, "High-Power-Density GaAs MISFETs with a Low-Temperature-Grown Epitaxial Layer as the Insulator," IEEE Electron Device Letters, Vol. 12, No. 6, pp. 306-308, 1991

[6] L.-W. Yin, N. X. Nguyen, Y. Hwang, J. P. Ibbetson, R. M. Kolbas, A. C. Gossard, U. K. Mishra, "Investigation of gate-drain breakdown characteristics of power GaAs MESFET with LTG passivation as a function of temperature," presented at 1993 Spring MRS Meeting, San Francisco, California, April 12-16, 1993.

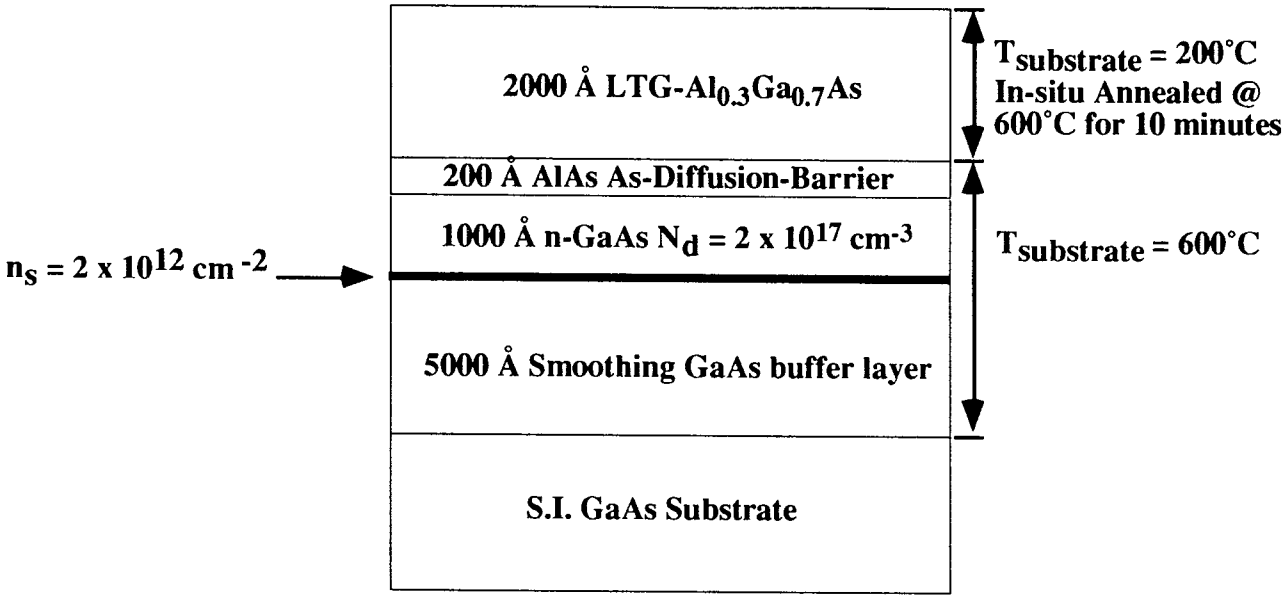


Figure 1 : MBE epitaxial structure and growth conditions

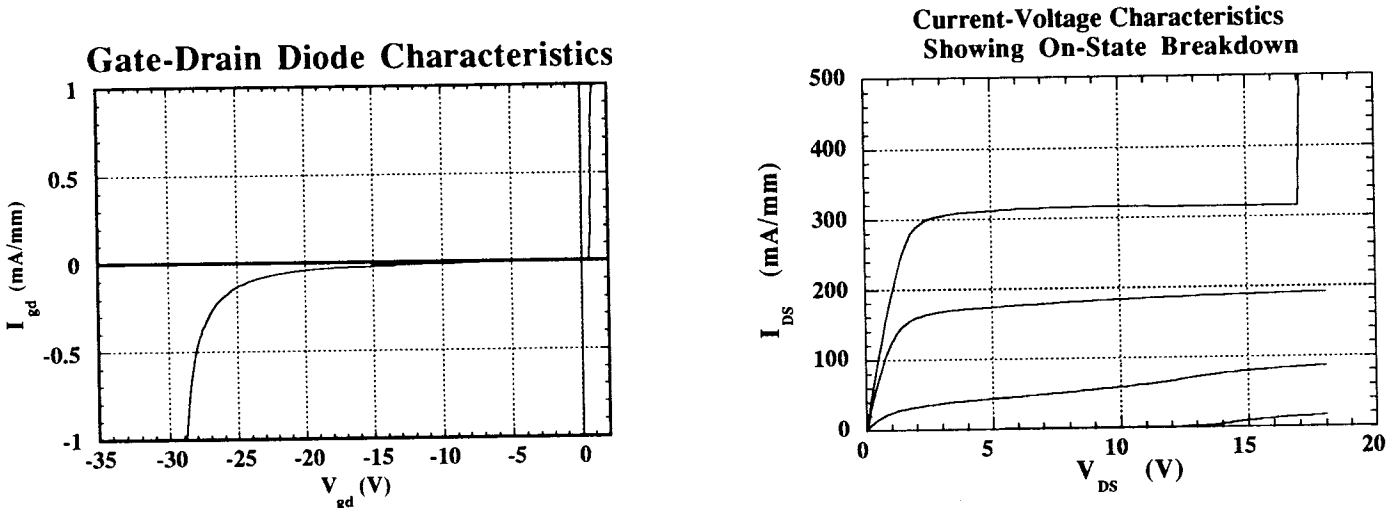


Figure 2 : DC characteristics of the device without complete encapsulation of the channel

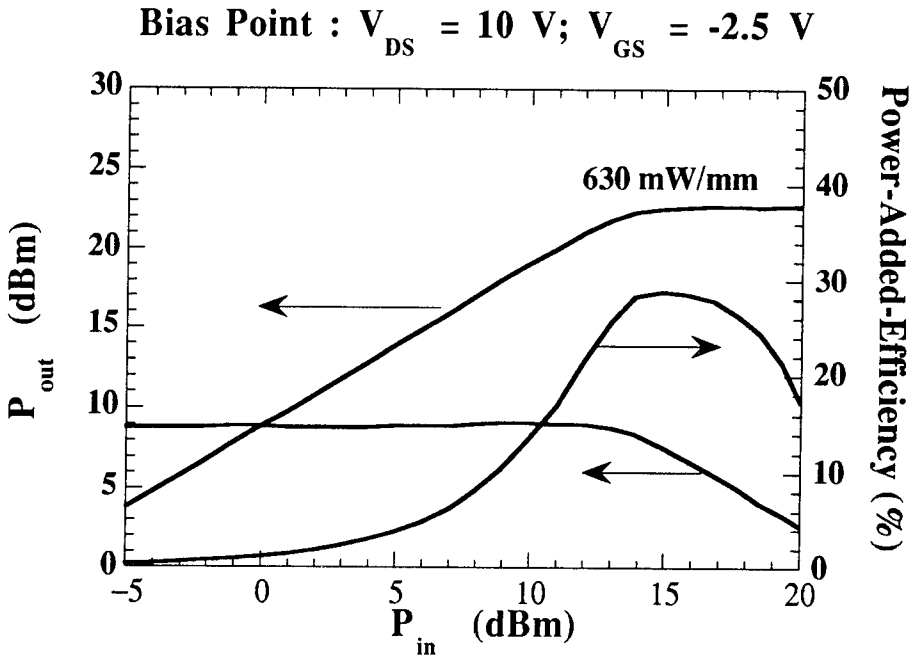


Figure 3 : Power performance of a GaAs MESFET without complete encapsulation of the channel at 4 GHz

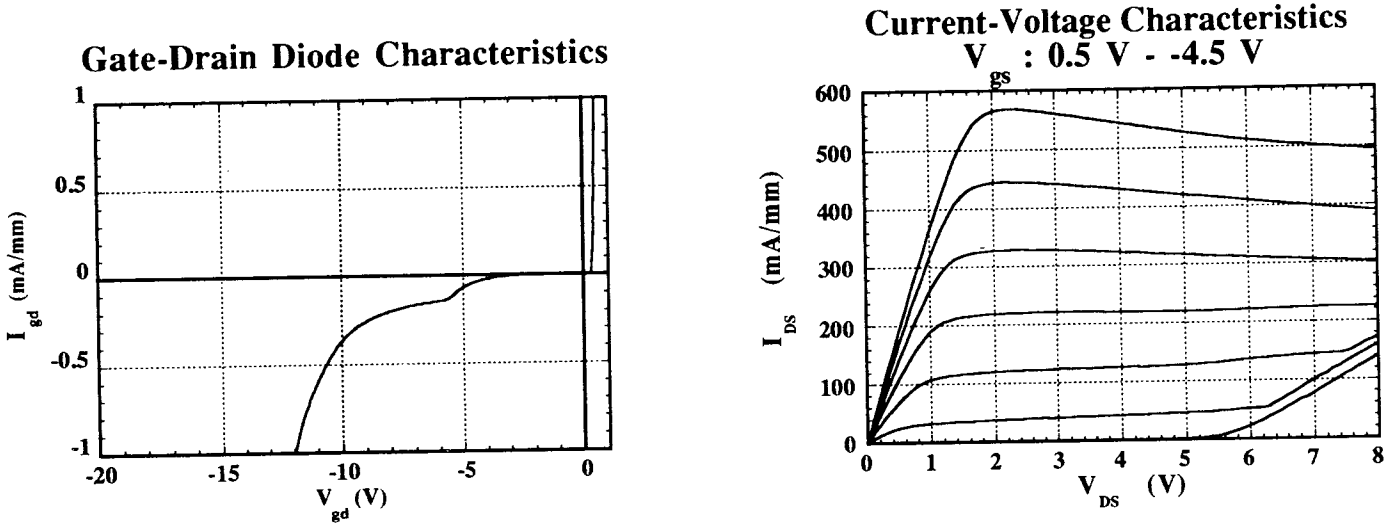


Figure 4 : DC characteristics of preliminary GaAs MESFETs with encapsulated channel

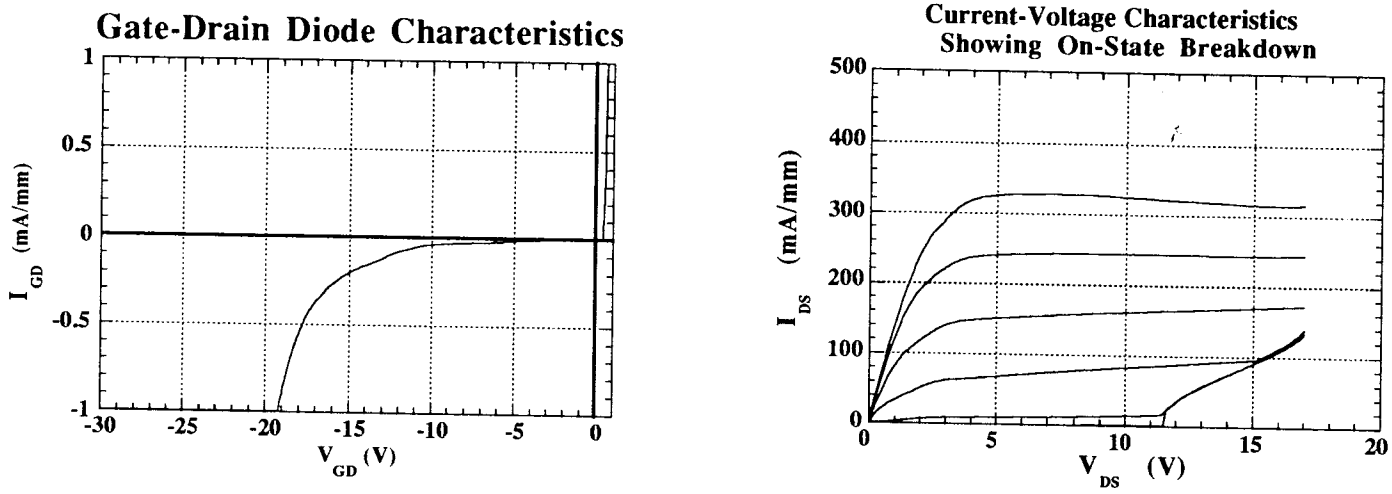


Figure 5 : Device characteristics of a GaAs MESSFET with an encapulated channel and high gate-drain breakdown

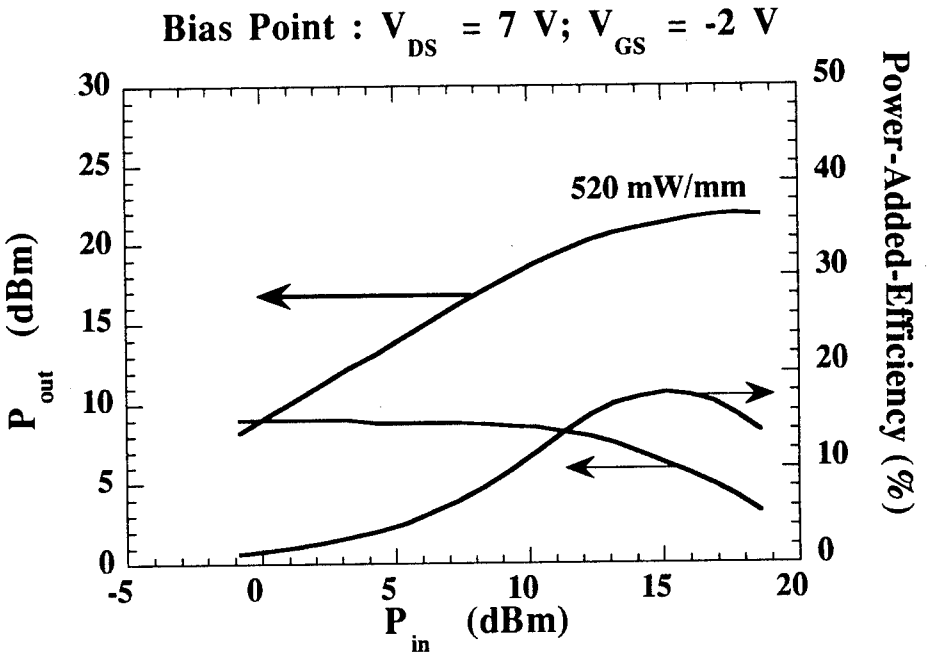


Figure 6 : Power performance of an encapsulated channel GaAs MESFET at 4 GHz

Breakdown Characterization of AlInAs/GaInAs Junction Modulated HEMTs (JHEMTs) with Regrown Ohmic Contacts by MOCVD

J.B. Shealy+, M.M. Hashemi*, S.P. DenBaars, and U.K. Mishra.

Department of Electrical and Computer Engineering
University of California, Santa Barbara CA 93106

+ Hughes Doctoral Fellow
*Motorola, Phoenix Corporate Research Laboratory
Tempe, AZ 85282

T.K. Liu, J.J. Brown, and M. Lui.

Hughes Research Laboratories
Malibu, CA 90265

Abstract

We present a technology to increase both the two-terminal gate-drain breakdown and subsequently the three-terminal-off-state breakdown of AlInAs-GaInAs HEMTs to record values without substantial impact on other parameters such as Idss and g_m. The breakdown in these structures is dependent on the injection of electrons from the source (channel current) and the gate (gate leakage) into the channel where multiplication occurs (due to high electric fields at the drain) producing holes which are swept back into the gate and source electrodes.

These phenomena can be suppressed by increasing the gate barrier height and alleviating the fields at the drain. In our approach we have achieved both by incorporating a p+ - 2DEG junction as the gate which modulates the 2DEG gas and by utilizing selective regrowth of the source and drain regions by MOCVD. The 1μm gate length devices fabricated show a full channel current of 350 mA/mm, transconductance of 240 mS/mm and **record high** two-terminal gate-drain and three-terminal-off-state breakdown voltages of **31V** and **28V**, respectively. The gate-to-drain spacing is 1μm and the breakdown is defined at 1mA/mm gate leakage. Further, temperature measurements were made to characterize both two-terminal and three-terminal-on-state breakdown. The gate current behavior is presented along with a calculated ionization rate which is compared with rates previously reported.

I. Introduction

AlInAs/GaInAs High Electron Mobility Transistors (HEMTs) have shown excellent high frequency and low noise performance [1]. However, the power performance of the AlInAs/GaInAs channel HEMT has yet to be exploited [2]. The reason attributed is a combination of the lower breakdown voltage in the InGaAs channel and the weak Schottky-barrier height on AlInAs. On the other hand, the high current densities (due to the large sheet charge density) available in the AlInAs/GaInAs system make it attractive for millimeter-wave power. Therefore, to optimize the power performance of the AlInAs/GaInAs HEMT, the gate barrier height and the breakdown voltage must be improved without sacrificing the current density.

0-7803-0894-8/93 $3.00 © 1993 IEEE

The breakdown in GaInAs channel HEMTS is determined by the multiplication of electrons injected into the channel from both the source and gate electrodes. Furthermore, holes are generated and swept back into the gate and source electrodes. The injection of electrons from the source is determined by the source current whereas the injection from the gate is determined by the gate barrier. The multiplication is determined by the electric field at the drain. A schematic of the breakdown in GaInAs channel HEMTs is shown in figure 1. Thus, for a given source current the breakdown should increase if the gate barrier is increased and the electric field at the drain is reduced.

We present a technology which improves the gate barrier height by incorporating a junction between a surface p+ layer and the 2DEG as the gate electrode. In addition, the drain fields are reduced by forming stable non-alloyed contacts to the 2DEG using selective regrowth of the source and drain regions by MOCVD. As a result, the two-terminal gate-drain and subsequently three-terminal-off-state breakdown voltage (for a 1μm gate-drain spacing) have increased to record values of 31V and 28.8 V, respectively. This is a substantial increase in breakdown over devices previously reported [3].

In addition, the three-terminal on-state breakdown of JHEMTs with regrown ohmic contacts has improved over similar JHEMTs with standard alloyed ohmic contacts. In an attempt to characterize three-terminal on-state breakdown, we have measured an ionization rate in a JHEMT with regrown ohmic contacts. This rate is compared with previous reported data.

II. MOCVD REGROWTH

Growth of the lattice matched n-InGaAs contact layer was achieved by low pressure metal-organic chemical vapor deposition (MOCVD). The sources were trimethylindium (TMI), trimethylgallium (TMG), tertiarybutylarsine (TBA) and disilane (150ppm in hydrogen) for the n-type source. The liquid organometallic group V source, TBA was purchased from Air Products and Chemicals, Inc. Bubbler temperatures were kept at 5°C, -10°C and 20°C for the TBA, TMGa, and TMIn sources, respectively. To eliminate organometallic vapor condensation all the source lines are resistively heated from the output of the bubblers to the injection manifold. The reactor employs two separate injection manifolds to keep group-V and group-III sources separate until they are mixed immediately upstream from the susceptor. This design was implemented to avoid any pre-reactions that might occur between the group-III alkyls and the organometallic group-V sources.

Before regrowing the contact layers, optimization of the bulk InGaAs doping characteristics with TBA and disilane was undertaken. A 150 ppm disilane in hydrogen mixture is used for n-doping of InGaAs. Linear doping behavior is observed for both InP and $In_{0.53}Ga_{0.47}As$ epilayers over a wide range of conditions. Figure 2 shows a plot of flow-rate of the disilane/hydrogen mixture versus carrier concentration at room temperature. Note that a relatively high n-doping saturation of $5(10)^{19}cm^{-3}$ in the InP is achieved when using TBP. For the n-InGaAs regrown contact layer a doping level of $2(10)^{18}cm^{-3}$ was employed. The best undoped $In_{0.53}Ga_{0.47}As$ epilayers had 77K mobilities of $59{,}000cm^2/V{\cdot}s$ and unintentional background impurity level of $5.5(10)^{15}$.

At a growth temperature of 600°C and reactor pressure of 100 Torr no deposition was observed on the oxide mask. The sample was initially heated up to 650°C for 3 minutes to remove the native oxide layer and produce specular morphology regrowth. Complete selectivity of the regrown InGaAs was obtained because of the increased diffusivity of the column III species at 100 Torr. The growth rate for regrown InGaAs epilayers in the device structure was approximately 11Å/sec. This results in a V/III ratio of 22 which ensures specular surface morphology at 100 Torr. A total gas flowrate of 5.5 slpm was found to yield a uniform film deposition in our system. After regrowth the sample was etched in concentrated BOE to remove the SiO_2 mask.

III. Fabrication

The AlInAs-GaInAs junction modulated HEMT (JHEMT) structures studied are shown in figure 3. The original epilayer structure was grown by Gas Source MBE. The electron mobility was 7800 and 38,000 cm^2/Vs at 300K and 77K, respectively. Next, the wafer was cleaved into two samples. Sample A was cleaned and coated with 1000Å of PECVD SiO_2, then patterned with the source-drain mask. Next, the patterned SiO_2 was etched back in CF_4 plasma. The sample was then wet chemically etched down to the channel, within 100Å of the buffer as shown in figure 3a. Before loading in the MOCVD reactor, the sample was dipped into a dilute BOE etch. InGaAs ($n=6x10^{18}\ cm^{-3}$) was regrown as discussed in the previous section.

Device processing of both samples started with gate metalization (Ta/Au, 200Å/2300Å) and lift-off to define a 1-μm-long gate. Next, the gate metal was used as the mask for dry etching (CH_4:H_2:Ar) through both the p+ InGaAs and the p+ AlInAs. The source-drain mask was then re-aligned and standard AuGe/Ni/Au (1000Å/200Å/1000Å) ohmic contacts were evaporated. Following the lift-off, the samples were patterned for mesa isolation. The isolation etch was done using RIE with chlorine. The standard JHEMT wafer, sample B, was alloyed on a strip heater at 370° C for 3 seconds (now called sample B-370). Sample A was cleaved and one piece alloyed at 350° C for 3 seconds (sample A-350) while the other piece was left non-alloyed (sample A-0).

IV. DC and RF Results

The I-V characteristics of 1-μm x 150-μm devices from both sample A-0 and sample B-370 were similar at low drain voltages. The devices produced good pinch-off characteristics and similar transconductance and current density. For sample A-0, with regrown ohmic contact regions, the specific contact resistance and channel sheet resistance were measured to be 0.43 Ω-mm and 300 ohms per square, respectively. Sample B-370 had a specific contact resistance of 0.53 Ω-mm with a channel sheet resistance of 334 ohms per square. Thus, the source resistance for sample A-0 and sample B-370 were similar and calculated to be 0.73 Ω-mm and 0.86 Ω-mm, respectively. In addition, the s-parameters were measured for sample A-0 yielding values of f_t and f_{max} of 22 GHz and 75 GHz, respectively, which are typical values of HEMTs with gate length of 1μm. A gate and drain bias dependence of both f_t and f_{max} was previously reported [4].

V. Breakdown

The main advantage of sample A-0 is the uniform non-alloyed ohmic contacts which we believe contributed to the higher breakdown voltages achieved. The two-terminal gate-drain diode characteristics of the devices discussed above are presented in figure 4. The two-terminal gate-drain breakdown voltage of sample A-0 is approximately 31V, which is 40% higher than the 22V breakdown measured for sample B-370. The two-terminal gate-drain breakdown of sample A-350 was 20% lower than sample A-0 but was still better than the gate-drain breakdown of any device from sample B-370. This illustrates the importance of (i) relieving the field at the drain by increasing the doping in drain region and (ii) creating smooth contacts.

The most important breakdown voltage for RF power amplification is the three terminal breakdown voltage. The three-terminal-off-state breakdown voltage is defined as the drain voltage, Vds, at which the gate current, Ig, is 1 mA/mm under pinch-off conditions. The three-terminal-off-state breakdown voltage of a typical device from sample A-0 was 28V and is shown in figure 5. The difference between the two-terminal gate-drain breakdown voltage, $V_{gd(BV)}$, and the three-terminal off-state breakdown voltage, $V_{ds(BV)}$, is the pinch-off voltage of the device. This indicates that the three-terminal off-state breakdown is dominated by the same mechanism as the two terminal gate-drain breakdown.

In addition, the three-terminal-on-state breakdown voltage characteristic for both sample A-0 and sample B-370 is shown in figure 6. The drain current was approximately Idss/2. The catastrophic three-terminal on-state breakdown voltage occurs at 4V for sample B-370 and increased to 7V for sample A-0. Further, we measured the ionization rate similar to recent measurements made on AlGaAs/GaAs HEMTs [5]. An effective ionization length, L , of 0.8 μm is chosen and the measured ionization rate is plotted in figure 7. Also plotted in figure 7 are calculated ionization rates derived from monte carlo simulations [6] and ionization rates measured on photodiodes [7].

VI. Conclusion

The breakdown of a new type of junction modulated AlInAs-GaInAs HEMT (JHEMT) which utilizes ohmic contact regrowth to increase both two-terminal breakdown and three-terminal breakdown has been characterized. Since the increase in breakdown voltage occurs without sacrificing the current density, this technology is attractive for the fabrication of power devices. These experimental results indicate that the source current is instrumental in determining the three-terminal on-state breakdown voltage whereas the three-terminal off-state breakdown voltage is dominated by the same mechanism observed in the two-terminal gate-drain breakdown characteristics. In conclusion, these results illustrate the important role of both the gate barrier and contact uniformity in determining the breakdown voltage of InGaAs channel HEMTs.

Acknowledgments

This work was supported by BMDO/ARO contract #DAAH04-93-G-0033 and NSF contract #DMR-9202290. U.K. Mishra is supported by an NSF PYI award # ECS-9144723.

References

[1] U. K. Mishra, A. S. Brown, M. J. Delaney, P.T. Greiling, and C. F. Krumm, IEEE's MTT, vol. 37, no. 9, September 1989.

[2] L. D. Nguyen, L. E. Larson, and U. K. Mishra, Proc. IEEE, vol 80, no. 4, April 1992.

[3] J. B. Boos and W. Kruppa, Electron. Lett., vol 27, no 21, pg 1909, October 1991.

[4] J.B. Shealy, M.M. Hashemi, K. Kiziloglu, S.P. DenBaars, and U.K. Mishra, T.K. Liu, J.J. Brown, and M. Lui. 51st Device Research Conference, 1993.

[5] C. Canali, A. Paccagnella, P. Pisoni, C. Tedesco, P. Telaroli, and E. Zanoni, IEEE's Transactions on Electron Devices, vol. 38, no. 11, pg 2571, November 1991.

[6] J. Bude and K. Hess, J. Appl. Phys. vol. 72, no. 8, pg 3554, October 1992.

[7] T. P. Pearsall, Appl. Phys. Lett. vol. 36, no. 3, pg 218, February 1980.

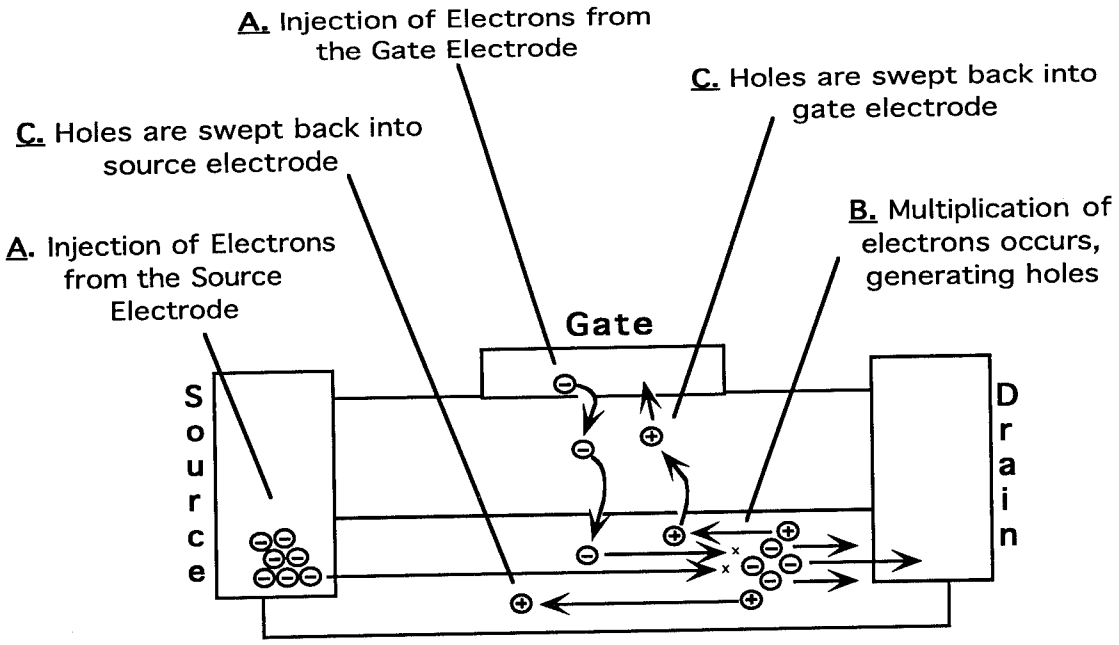


Figure 1. Schematic of Breakdown in GaInAs Channel HEMTs.

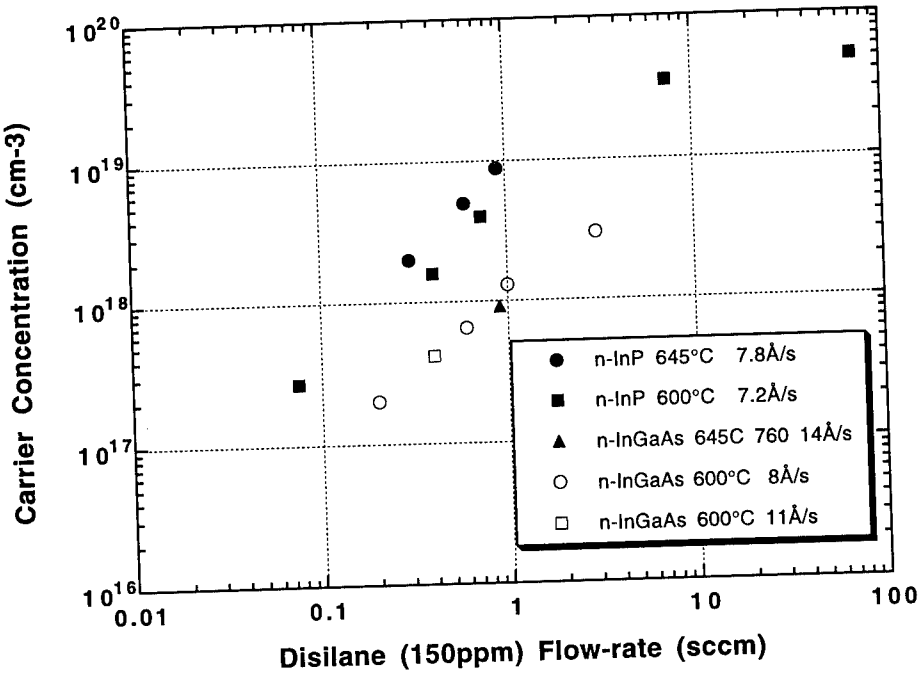


Figure 2. Free carrier concentration as a function of disilane effective flow-rate at various InP and InGaAs growth rates and growth temperatures.

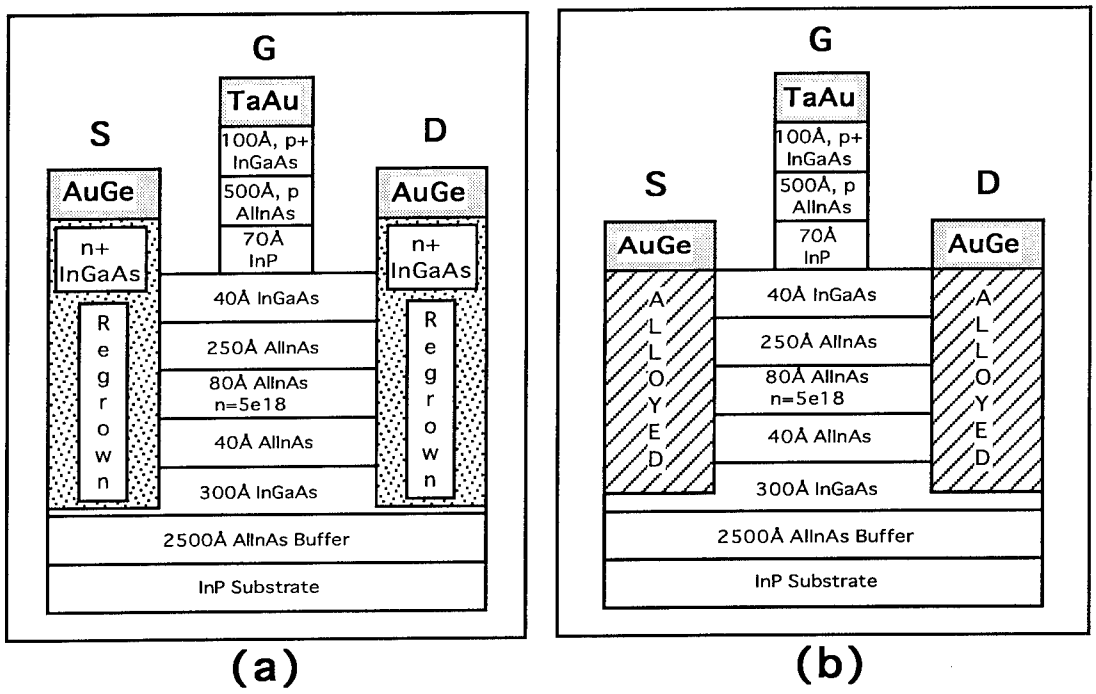


Figure 3. JHEMT Structures Fabricated: (a) Sample A-0 and (b) Sample B-370

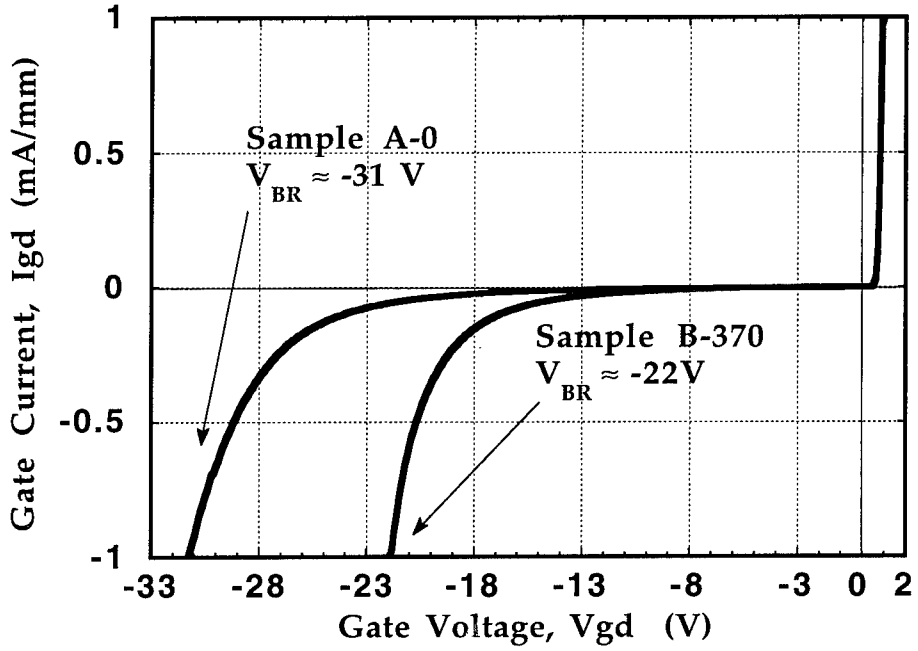


Figure 4. JHEMT Two-Terminal Gate-Drain Breakdown for both Sample A-0 and Sample B-370.

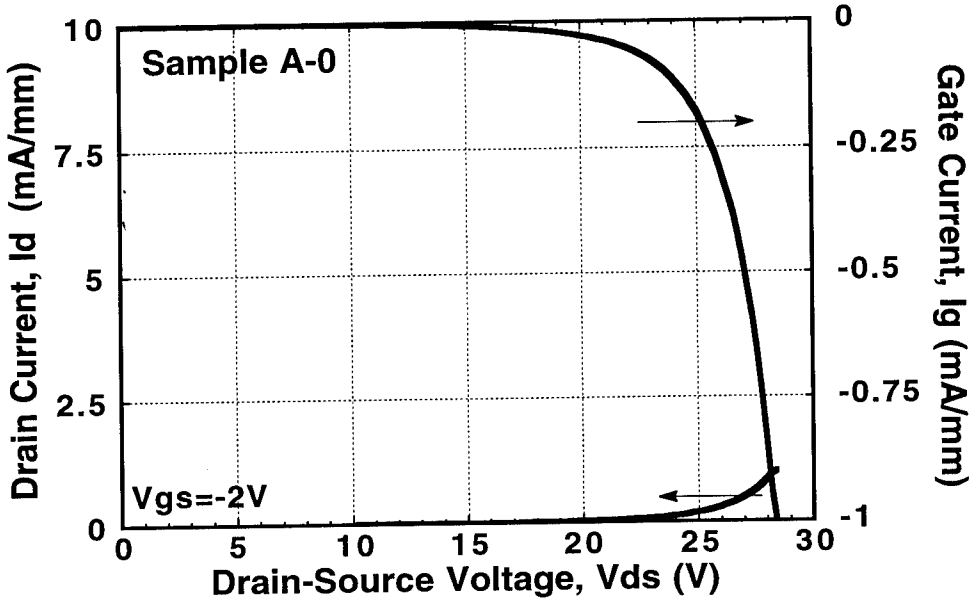


Figure 5. JHEMT Three-Terminal-Off-State Breakdown for Sample A-0.

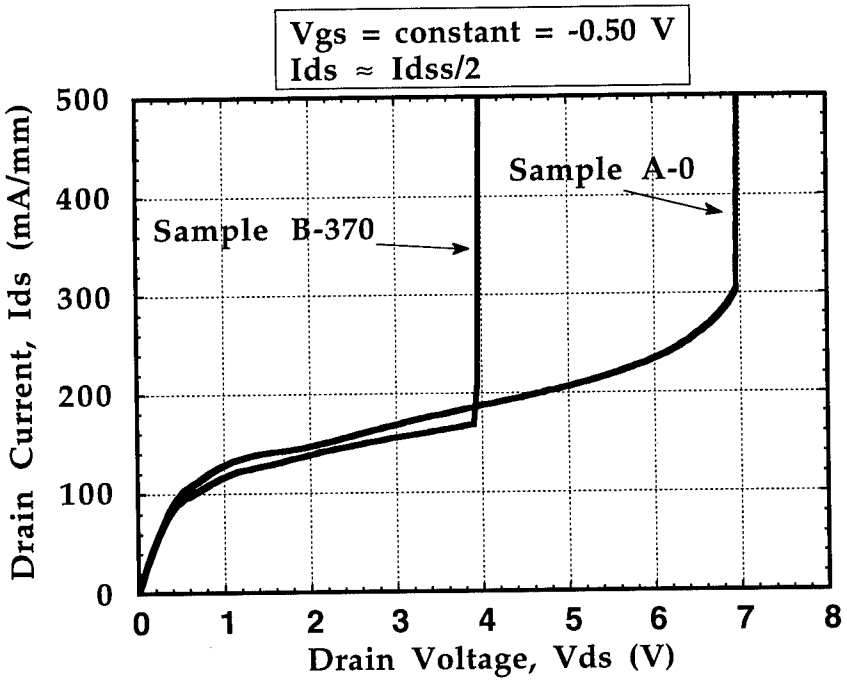


Figure 6. JHEMT Three-Terminal-On-State Breakdown for both Sample A-0 and Sample B-370.

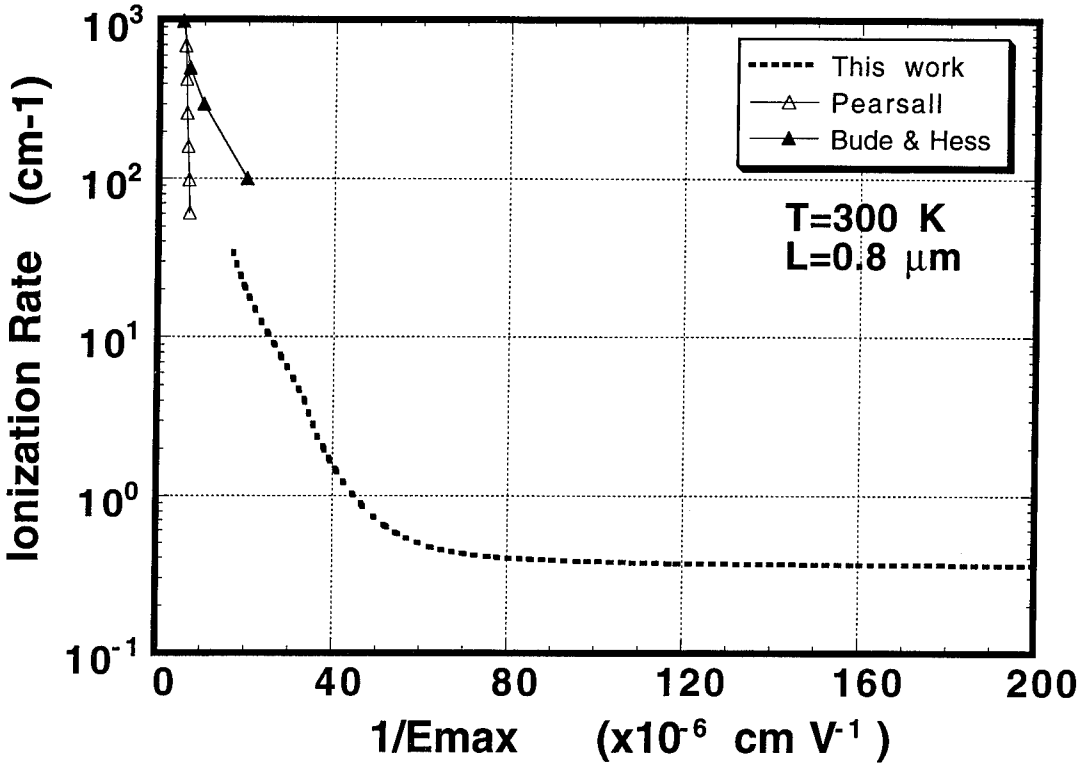


Figure 7. Calculated Ionization Rates measured from JHEMTs with regrown ohmic contacts (Sample A-0). Comparison is made to values published by references [6] and [7].