INFORMATION TO USERS

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

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

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

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

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

UMI

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

UNIVERSITY OF CALIFORNIA
Santa Barbara

Silicon-Indium-Gallium-Arsenide Avalanche Photodetectors

A dissertation submitted in partial satisfaction
of the requirements of the degree of

Doctor of Philosophy
in
Electrical and Computer Engineering

by
Aaron Roe Hawkins

Committee in charge:

Professor John E. Bowers, Chairperson

Professor Steven P. DenBaars

Professor Evelyn L. Hu

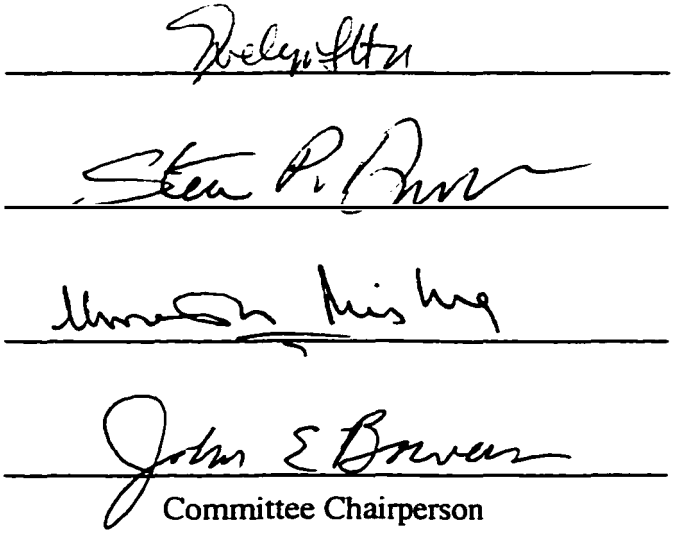
Professor Umesh K. Mishra

December 1998

UMI Number: 9921800

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

The dissertation of Aaron Roe Hawkins is approved



Committee Chairperson

December 1998

Silicon-Indium-Gallium-Arsenide Avalanche Photodetectors

December 1998

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

Acknowledgments

I would like to thank Prof. John Bowers for his support and faith in me. Many aspects of this project proved to be more difficult than initially expected, but he never let me give up and constantly pushed me forward. He also allowed me the freedom to pursue my own solutions to research problems and rarely second guessed me. He has taught me much more than just about solid state devices. I also thank the other members of my dissertation committee, Prof. Steve DenBaars, Prof. Umesh Mishra, and Prof. Evelyn Hu. They contributed not only to the direction of this work but to my development as a scientist.

I am also indebted to all of the graduate students I have learned from and with at UCSB. In particular I must thank Dubravko Babic for introducing me to wafer fusion and encouraging me from my first day. Dan Tauber also provided much needed support and advice. I will always have great memories of Near Margalit since we shared essentially the same undergraduate and graduate experiences. Holger Schmidt also made classes and research much more entertaining. Lance Rushing was so much fun I found a way to still work with him.

I am very grateful for all the help I received from others at UCSB as well. Undergraduates Derek England and James Reilley worked extremely diligently to help move this project along. Weishu Wu was always the brainpower behind the most difficult theoretical calculations. Jack Whaley,

Brian Carrelejo, Bob Hill, and Martin Vanderbrock also seemed to treat me especially well and were always willing to help with any equipment designs or problems. Gerry Robinson taught me more than anyone else about semiconductor processing - something many at UCSB can say. No one deserves more gratitude than Tom Reynolds. He helped me with every aspect of this project and helped maintain my sanity. I will always count him among my best and most respected friends.

It goes without saying that I could not have done much of anything without all those who grew semiconductor material for me: Mark Mondry, Archie Holmes, Patrick Abraham, and Klaus Streubel.

I was also very fortunate to be involved in several collaborations that allowed me to work with those outside of UCSB. I would like to thank Mike Jack from Raytheon, Richard Farrell from RMD, and Barry Levine from Lucent for all they taught me as well as their enthusiasm. Looking back, one of my most influential interactions was with the late Robert McIntyre. Our discussions were both humbling and inspiring.

I would also like to thank DARPA and Rome Labs for their financial support of the APD project and wafer fused devices in general.

Finally I owe my motivation to my wife Kellie. She provided the reason to keep trying.

Vita

February 26, 1970:	Born in Rehoboth, New Mexico
June 1994:B.S.:	Applied Physics, California Institute of Technology
March 1996:	M. S., Electrical and Computer Engineering University of California, Santa Barbara
1994-98:	Research Assistant, Department of Electrical and Computer Engineering,University of California, Santa Barbara

Publications

1. A. R. Hawkins, T. Reynolds, D. R. England, D. I. Babic, M. Mondry, and J. E. Bowers, "Silicon Hetero-Interface Photodetector," LEOS '95 Annual Meeting, San Francisco, Oct. 30-Nov. 2 (1995).

2. A. R. Hawkins and J. E. Bowers, "Silicon-InGaAs APDs - the best of both worlds," WOSCEMMAD '96, Santa Fe, New Mexico, February 19-21 (1996).

3. A. R. Hawkins, T. E. Reynolds, D. R. England, D. I. Babic, M. J. Mondry, K. Streubel, and J. E. Bowers, "Silicon Heterointerface Photodetector," *Applied Physics Letters*, **68** (26), 3692-3694, June 24 (1996).

4. A. R. Hawkins, T. E. Reynolds, D. England, D. I. Babic, K. Streubel, and J. E. Bowers, "81 GHz Gain-Bandwidth Product Silicon Hetero-Interface Photodetector," Device Research Conference, Santa Barbara, CA June 24-26 (1996).

5. A. R. Hawkins, W. Wu, P. Abraham, K. Streubel, and J. E. Bowers, "High Gain-Bandwidth Product Silicon Hetero-Interface Photodetector," Postdeadline Paper, LEOS '96, Boston, MA, Nov. 18-21 (1996).

6. W. Wu, A. R. Hawkins, and J. E. Bowers, "Frequency Response of Avalanche Photodetectors with Separate Absorption and Multiplication Layers," *Journal of Lightwave Technology*, **14** (12), 2778-2785, Dec. (1996).

7. A. R. Hawkins, W. Wu, P. Abraham, K. Streubel, and J. E. Bowers, "High Gain-Bandwidth-Product Silicon Heterointerface Photodetector," *Applied Physics Letters*, **70** (3), 303-305, January 20 (1997).

8. W. Wu, A. R. Hawkins, and J. E. Bowers, "Design of Si/InGaAs Avalanche Photodetectors for 200 GHz Gain-Bandwidth Product," invited paper, Optoelectronic Integrated Circuits-SPIE Photonics West '97, San Jose, CA, February 8-14 (1997).

9. A. R. Hawkins, W. Wu, and J. E. Bowers, "Improvement of Avalanche Photodetectors Through Integration of InGaAs and Si," invited paper, SPIE Optoelectronics '97, Conference Proceedings, 2999, 68-75, San Jose, CA, February 8-14 (1997).

10. W. Wu, A. R. Hawkins, and J. E. Bowers, "High Gain-Bandwidth Product Si/InGaAs Avalanche Photodetectors," OFC '97, Dallas, TX, February 16-21 (1997).

11. B. F. Levine, A. R. Hawkins, S. Hiu, B. J. Tseng, C. A. King, L. A. Gruezke, R. W. Johnson, D. R. Zolnowski, and J. E. Bowers, "20 GHz High Performance Si/InGaAs PIN Photodetector," *Applied Physics Letters*, **70** (18), 2449-2451, May 5 (1997).

12. P. Abraham, A. K. Black, N. M. Margalit, A. R. Hawkins, S. P. DenBaars, and J. E. Bowers, "Substrate Engineering of 1.55 μm Lasers," IPRM '97, Hyannis, MA, May 11-15 (1997).

13. W. Wu, A. R. Hawkins, and J. E. Bowers, "Effect of Electric Field Profile on Gain-Bandwidth Product of InGaAs/Si Avalanche Photodetectors," CLEO '97, Baltimore, MD, May 18-23 (1997).

14. J. E. Bowers, A. Hawkins, N. Margalit, P. Abraham, A. Black, A. Holmes, and E. Hu, "Wafer Fusion Technology for Optoelectronic Devices," Invited Paper, CLEO '97, Baltimore, MD, May 18-23 (1997).

15. A. Black, A. R. Hawkins, N. M. Margalit, D. I. Babic, A. L. Holmes, Jr., Y.-L. Chang, P. Abraham, J. E. Bowers, and E. L. Hu, "Wafer Fusion: Materials Issues and Device Results," *IEEE Journal of Selected Topics in Quantum Electronics*, **3** (3), 943-951, June (1997).

16. W. Wu, A. R. Hawkins, and J. E. Bowers, "Silicon Based Telecommunication Avalanche Photodetectors," Invited Paper, OECC '97, Seoul, Korea, July 8-11 (1997).

17. W. Wu, A. R. Hawkins, and J. E. Bowers, "Analysis of the Effect of an Electric-Field Profile on the Gain-Bandwidth Product of Avalanche Photodetectors," *Optics Letters*, **22** (5), 1183-1185, August (1997).

18. W. Wu, A. Hawkins and J. E. Bowers, "Design of Silicon Hetero-Interface Photodetectors," *Journal of Lightwave Technology*, **15** (8), 1608-1615, August (1997).

19. B. F. Levine, A. R. Hawkins, S. Hiu, J. P. Reilly, B. J. Tseng, C. A. King, L. A. Gruezke, R. W. Johnson, D. R. Zolnowski, and J. E. Bowers, "Characterization of Wafer Bonded Photodetectors Fabricated Using Various Annealing Temperatures and Ambients," *Applied Physics Letters*, **71** (11), September 15 (1997).

20. J. E. Bowers, A. R. Hawkins, and N. M. Margalit, "Novel Optoelectronic Devices Using Wafer Bonding," Invited Paper, 1997 Americal Vacuum Society Annual Meeting, San Jose, October 20 (1997).

21. A. R. Hawkins and J. E. Bowers, "High-Speed Wafer-Fused Photodetectors," Invited Paper, LEOS '97, San Francisco, CA, November 11-13 (1997).

22. K. A. Black, A. R. Hawkins, N. M. Margalit, D. I. Babic, A. L. Holmes Jr., Y.-L. Chang, P. Abraham, J. E. Bowers, and E. L. Hu, "Fusion Bonding: Hetero-interfacial Materials Analysis and Device Application," SPIE, San Jose, CA, January 26-29 (1998)

23. K. A. Black, A. R. Hawkins, N. M. Margalit, D. I. Babic, A. L. Holmes, Jr., Y.-L. Chang, P. Abraham, J. E. Bowers, and E. L. Hu, "Fusion Bonding: Hetero-interfacial Materials Analysis and Device Application," chapter in Heterogeneous Integration: Systems on a Chip, Vol. CR70, A. Husain, M. Fallahi, eds., SPIE, Bellingham, Washington, (1998).

ABSTRACT

Silicon-Indium-Gallium-Arsenide Avalanche Photodetectors

by

Aaron Roe Hawkins

Current optical communications systems rely heavily on avalanche photodetectors (APDs) to detect light exiting an optical fiber. APDs have two functions, the absorption and conversion of light to an electrical signal and the amplification of that signal through avalanche multiplication. With the advantage of internal amplification, the speed and sensitivity performance of APDs are unmatched by other types of detectors.

An ideal APD would use a material as an absorber with a high optical absorption coefficient to obtain high quantum efficiencies. A material used as a multiplier should have dissimilar electron and hole ionization coefficients. Very large or very small ratios (k) of these coefficients result in fast multiplication action with little signal noise. Given these criteria, the two materials of choice would be InGaAs, which has a very high optical absorption coefficient in the visible and near-infrared regimes, and silicon, which has the lowest k value reported for any semiconductor.

InGaAs and silicon are not a natural fit, although they would make a desirable combination. Joining dissimilar semiconductors like these through conventional epitaxy has proved in the past to be extremely difficult. A new

processing method known as wafer fusion has provided a path for their integration, however. This dissertation describes the fabrication of InGaAs-silicon APDs using the wafer fusion technique. Several generations of devices are described along with their electrical characteristics. InGaAs-silicon PIN detectors were also created, providing valuable information about the nature of the fused interface. Theoretical models and designs are also providing for describing the operation of the InGaAs-silicon APD. Such characteristics as voltage and temperature sensitivity, noise, and bandwidth are calculated. The theoretical results as well as measurements from fabricated detectors confirm the potential of InGaAs-silicon APDs to surpass the performance of present avalanche photodiodes made from other materials. This includes APDs operating at the important wavelengths of 1.3 and 1.55 μm used for optical fiber communications.

Table of Contents

1. Introduction and Motivation

1.0 A New Detector 1

1.1 Optical Fiber Communication Systems 3

1.2 Avalanche Photodiodes 11

1.3 Existing Avalanche Photodiodes 16

1.4 Optimum APD Using InGaAs and Silicon 19

1.5 Dissertation Layout 21

References 24

2. Design Considerations

2.0 Introduction 29

2.1 Requirements for a Useful APD 29

2.2 Design of Silicon APDs 31

2.3 Design of InGaAs/InP APDs 36

2.4 Designing an InGaAs-Silicon APD 40

2.5 Feasibility of InGaAs-Silicon APD Construction 43

2.6 Conclusions 48

References 49

3. Avalanche Photodiode Gain Equations

3.0 Introduction 52

3.1 The Gain Mechanism 53

3.2 Deriving the Gain Equations 57

3.3 Measuring Ionization Coefficients 63

3.4 Gain Versus Electric Field 67

3.5 Conclusions 68

References 70

4. Voltage and Temperature Sensitivity

4.0 Introduction 71

4.1 Experimental and Theoretical Gain Versus Voltage Curves 74

4.2 A Closer Look at Gain Versus Voltage 77

4.3 Temperature Sensitivity Derivation 86

4.4 Conclusions 94

References 95

5. Noise and Bandwidth

5.0 Introduction 97

5.1 Noise Equations 100

5.2 A Look at the Gain Relations 102

5.3 The Groundwork for Bandwidth Equations 108

5.4 Gain Bandwidth Equations –
Effective Multiplication Plane Method 114

5.5 Gain Bandwidth Equations - Hollenhorst Approach 120

5.6 Bandwidth Plots for InGaAs-Silicon APDs 124

5.7 Conclusion 131

References 133

6. Wafer Fusion

6.0 Introduction 137

6.1 The Process 139

6.2 Results - SEMs and TEMs 142

6.3 SIMS Profile 147

6.4 A Robust Bond 149

References 152

7. PIN Devices

7.0 Introduction 154

7.1 Fabricating an InGaAs-Silicon PIN 155

7.2 Current Versus Voltage and Quantum Efficiency 158

7.3 Capacitance and Bandwidth 162

7.4 Varying the Fusion Temperature and Atmosphere 165

7.5 Results of the Varying Fusion Conditions 168

7.6 An Interface Barrier 171

7.7 Conclusions 174

References 177

8. APD Devices - Fabrication and Measurements

8.0 Introduction 178

8.1 First Generation Device - Design and Fabrication 179

8.2 First Generation Device – Measurements 182

8.3 Second Generation Device - Design and Fabrication 189

8.4 Second Generation Device – Measurements 193

8.5 Third Generation Device - Design and Fabrication 202

8.6 Third Generation Device – Measurements 206

8.7 Conclusions 215

References 217

9. Conclusions and Outlook

9.0 Conclusions 219

9.1 Future Work 222

9.2 Outlook 225

Chapter 1

Introduction and Motivation

1.0 A New Detector

A silicon-indium-gallium-arsenide avalanche photodetector (InGaAs-silicon APD) is a semiconductor device made for the purpose of optical detection. This detector follows in a long line of man-made inventions and biological organs devoted to detecting the presence of light. The ability to sense light, or see, is a necessity of advanced life on our planet. A dividing line can be drawn between lesser organisms with no response to light and all other higher animal life that have at least primitive optical registration. Biological light detection is not just limited to mammals, birds, and fish, however. Even some single cell microbes react to light stimulus. Their reaction is the most basic of all light detection - they know simply whether light is present or not - on or off. There are many possibilities between this detector and the most sophisticated detector ever known - the human eye. The eye certainly detects the basic presence of light but also detects light levels with an incredible dynamic range. Optical wavelengths are also registered by the eye and show up as different colors in the brain. This is all done, of course, in a huge array that allows for the detection of images and objects in a dynamic "real time" display.

One use of optical detection that has advanced incredibly in the 20th century is optical communication. It is for this specific application that the InGaAs-silicon APD was invented. Although neither this nor any other detector can match the complexity of the human eye, it is not the best instrument for efficient optical communication, although throughout history eyes were used because no other detector was available. Optical communication boils down to detecting a light signal from a source transmitting information. This information could take the form of digital signals - on or off - or analog signals with variations in light intensity. We will never know what the first digital optical communication was. Perhaps it was a fire used to signal a warning or a victory. Early sailors were warned of impending danger by lighthouses and flashed Morse code to other ships using flashes of light. The American Revolution even saw its share of digital optical communication via the famous "one if by land, two if by sea" signal warning of British troop movements. A light sensitive microbe, however, could have picked up all of these signals, just as well as the human eye.

To do the simplest optical communication, all that is needed is a light source and a detector that produces a signal to indicate when light is present and when it is not. The way this is accomplished in many detectors is through the photoelectric effect - light strikes matter and its energy is converted to electron energy producing heat or electrical current. In the human eye light strikes cones

in the retina stimulating electrical signals to the brain. With the advancement of technology including the study and use of electricity, it became conceivable to try to duplicate this action in a man made device. One such device still in use today is the photomultiplier tube (PMT). The PMT utilizes the photoelectric effect by collecting electrons accelerated off of a metal surface when it is illuminated by light. With the advent of solid-state devices and the semiconductor era, advances in man-made detectors accelerated. Using ultra high purity semiconductor scientists have been able to create detectors with incredible sensitivities to light and mind boggling speeds. These detectors were designed around one purpose - optical communication and they excel at the simple task of sensing whether light is on or off and producing an electrical pulse to indicate this. Current state of the art detectors can perform this task over 10 billion times a second using a variety of light wavelengths.

Figure 1.1 is an elementary illustration of the elements required for optical communication. First a light source is needed to impart the information. This could be a light bulb, a light emitting diode, or even a laser. The light simply has to be modulated to transfer data. The light must then pass through some transport medium such as air, a vacuum, or an optical fiber. Whatever the medium, this usually results in a loss of light intensity as photons travel through it. At the end of a length of transport medium sits the optical detector. As stated above, this might be a human eye or a man made detector like a PMT. A

more likely detector these days, however, would be a solid-state PIN detector or avalanche photodiode (APD)[1]. This dissertation will concentrate on the latter of these devices, the APD. Here we will formalize the introduction of a new type of APD made by joining two very different semiconductors together, silicon and indium-gallium-arsenide, to take advantage of their material properties. Many of the mechanisms and designs for this detector have been built upon APDs made previously from other material combinations. The aim of the InGaAs-silicon APD is to improve on the performance of those previous devices and to increase the capacity of optical communication - to see faster and farther.

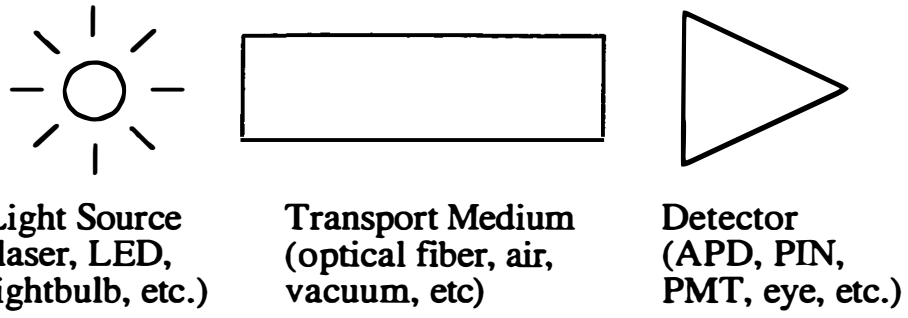


Figure 1.1 - The elements needed to make a basic optical communication system.

1.1 Optical Fiber Communications Systems

The real driving force behind detector development at the end of the 20th century has been the advent of optical fiber. These tiny strands of ultra-pure glass can now carry light signals further than anyone thought possible fifty years ago. Back then a good argument could not be made for the widespread use of optical communications systems. Radio waves and electrical signals could be transmitted much farther and with more information than light pulses that lacked a true medium besides the open air for transport. Optical fiber provided the medium that light needed to establish its preeminence in communication technology. Light traveling through fiber could be pulsed at seemingly limitless rates pushing bandwidths for information orders of magnitude higher than achievable with existing copper wire. Improvements in fiber manufacturing dropped the optical loss due to light absorption and led to longer transmission distances (tens of kilometers) before amplification was required.

Successful fiber-optic communication systems required very specialized components to serve as optical sources and optical detectors. The first necessity was for these devices to be able to couple light into the fiber and collect light out of the fiber efficiently. This is no small task given optical fiber's small dimensions - about 10 μm in core diameter for today's advanced fibers. This required the device dimensions for sources and detectors to be on the same order and also spawned a new research field - fiber optic packaging, the linking

of fiber with electrical components. Using semiconductor devices with small dimensions does have an intrinsic advantage, however. Decreasing device size decreases the capacitance of a semiconductor source or detector, thus increasing the achievable operation frequency. Semiconductor lasers and detectors are being put in today's optical systems with operating frequencies of 10 GHz - only achievable with extremely low capacitances.

Another requirement dictated by optical fiber is the wavelength of the light traveling through it. In selecting a wavelength for a fiber optic system, the preference would be the wavelength that could travel furthest down the fiber before losing its intensity. What has become standard optical fiber consists essentially of silicon dioxide. Light traveling through this medium will experience scattering losses as well as absorption as it propagates. All visible and near-infrared light can be used in optical fiber, but for silicon dioxide based fiber, two wavelengths, 1.3 μm and 1.55 μm, travel with the lowest loss. These wavelengths have become important targets for optoelectronic devices. Other wavelengths are usually reserved for short fiber length applications. The quest to create lasers that produce 1.3 and 1.55 μm light and detectors to detect that light, has driven heavily the compound semiconductor field. The earliest single element semiconductors, silicon and germanium, could not be used for lasers at this wavelength and the performance of germanium based detectors was soon eclipsed by PINs and APDs made from other materials. Ternary and quaternary

mixtures of indium, gallium, arsenic, aluminum, and phosphorous along with GaAs and InP substrates are now the standard building blocks for optoelectronics.

One of the great advantages of optical fiber is the long distance span possible between a source and a detector. Light traveling in a fiber between these components experiences more and more intensity loss the further it goes. In order to stretch these lengths and still have a sufficient signal to detect at the end of a fiber, optical sources need to put out large amounts of light and detectors need to be very sensitive. A large research effort has been devoted to semiconductor lasers to serve as sources for these systems. A wide variety of lasers are now available that produce light at the 1.3 and 1.55 μm wavelengths with high output powers. Power into the fiber at the source end is now 10 mW or more for commercial lasers. At the other end of the fiber, in which this dissertation is more concerned, sit semiconductor detectors and the supporting electronics for their operation. The detector and electronics are often packaged together as a "receiver"[2]. A good receiver is expected to take a very dim optical signal from a fiber and produce an electrical signal with the same pattern that can then used by other circuits. In essence, the less light necessary to produce a usable electrical pattern, the better the receiver.

Common to all receivers is an electrical amplifier that serves a dual role. First it converts the electric signal produced by a detector from a current swing

into a voltage swing. Solid state detectors produce signals that are essentially changes in total current flowing through them at a given voltage. The change in current is caused by light striking the detector and producing electrons, creating additional current flow. A detector's off state produces a low current flow and its on state a higher current, the level of which of course depends on the amount of light incident on the device. This current difference must be changed into a voltage difference before it can be sent out of the receiver to other circuits. The second function of an amplifier in a receiver is the amplification of the original signal. The amplification or gain available is limited however by the noise figure of the amplifier. This means a certain input power level into the amplifier is required or the output signal from the receiver will contain too many errors.

Current optical fiber receivers use one of two types of solid-state detectors, PINs or APDs. Both of these detectors are sensitive to light at 1.3 and 1.55 μm and are constructed from basically the same materials. A PIN detector or diode, refers simply to the doping structure of the device - P being a p-type doping region, I an intrinsic or undoped semiconductor region, and N an n-type doped region. The detector is operated under reverse bias so that light incident on the I region produces electrons and holes that get swept toward the diodes anode and cathode producing a photocurrent. An avalanche photodiode, or APD, performs the same function as a PIN but has a more complicated

structure that gives it the added capability of producing internal gain. The APD has an I region where electrons and holes are produced but also has a multiplication region where avalanche multiplication gain takes place. Figure 1.2 illustrates the action that takes place within an optical receiver using both a PIN and APD detector. Light exits an optical fiber hitting the detector. An electrical signal is produced by the detector that has variations in the current level. An amplifier then produces an amplified signal with voltage variations. The key difference between receivers built with the two detectors is that due to the APD internal gain, the APD receiver in general requires a lower level optical signal to produce the same electrical signal.

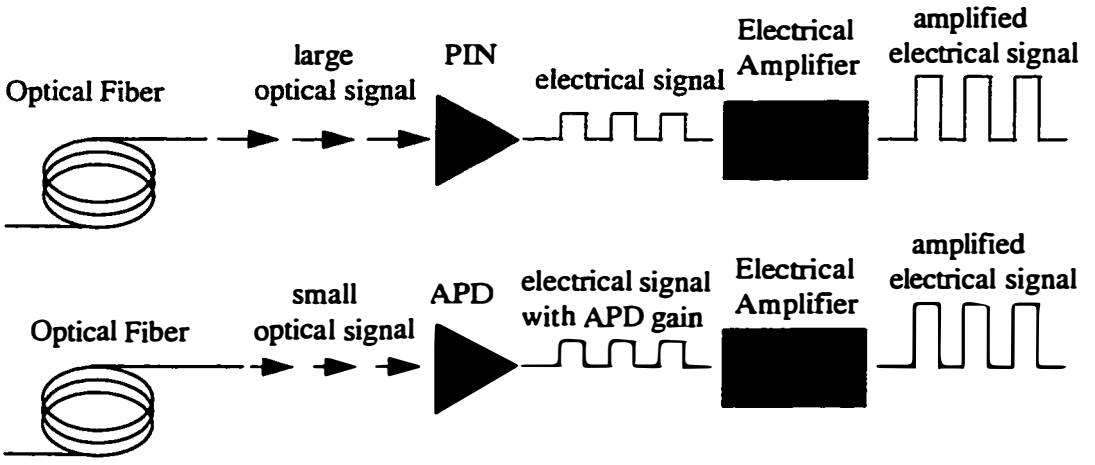


Figure 1.2 - The receiving end of an optical fiber communication system. The use of both a PIN and an APD detector are illustrated.

Optical system designers are often presented with the choice between using a PIN based receiver and an APD receiver. Due to its internal gain, the APD in general needs less incident light to produce the same electrical signal than does a PIN detector. There is added signal noise associated with the gain in the APD, however, so in some cases a PIN is preferred. The APD gain is only advantageous if the photon noise and avalanche multiplication noise levels are below the circuit noise for the receiver's amplifier.[3] For a given amplifier and data rate, there will be a given optical input strength below which the use an APD will increase the signal to noise ratio for the receiver. The term "sensitivity" is used to quantify the performance of an optical receiver. Sensitivity is the minimum optical input signal necessary to produce an electrical output signal with less than a required number of errors. In general, current commercial APD receivers operating at 2.5 GBit/s made from InP and InGaAs semiconductors need about ten times less light to produce the same electrical signal than do PIN receivers. Figure 1.3 shows the receiver sensitivity increase gained by switching from an InGaAs/InP PIN to an InGaAs/InP APD while using the same amplifier[4]. The sensitivity advantage increases with data rate as can be seen.

The decision to use an APD or a PIN in a system is often not only dictated by device performance but also by cost. APDs are more difficult to fabricate and require more sophisticated bias circuitry and so the price of using

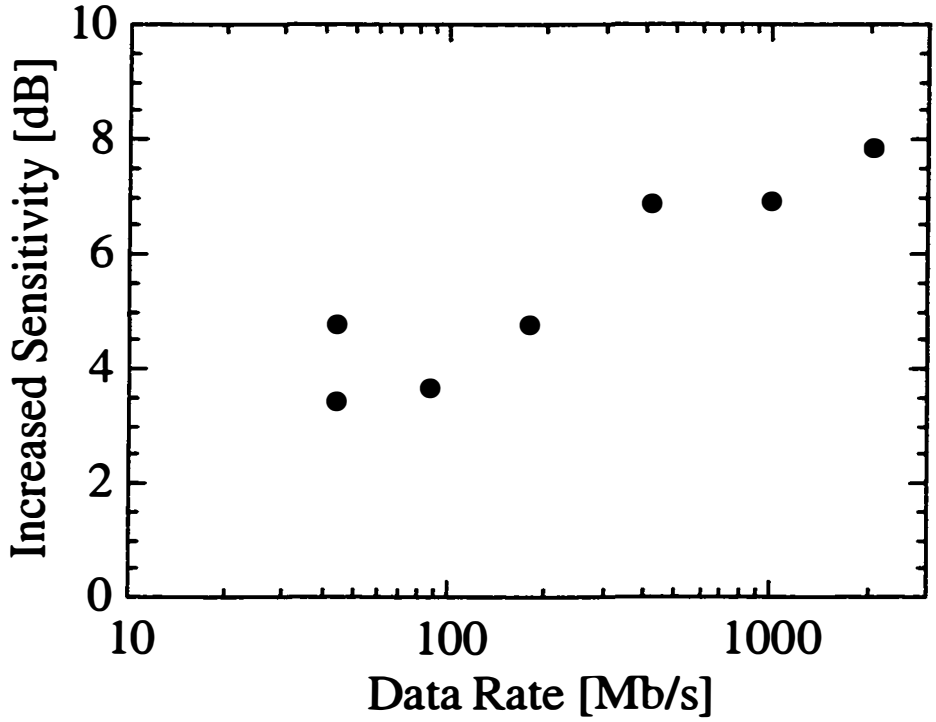


Figure 1.3 - Experimental results for the increase in sensitivity using an InP based APD versus using an InP based PIN in an optical receiver at different data rates[4].

an APD is higher than a PIN. When placing APD receivers in a system, designers will use them only when the sensitivity requirements are beyond what a PIN receiver can handle. Short stretches of fiber with plenty of output power are left to PINs while the long-haul stretches where signal power is a premium rely on APDs. As illustrated by Figure 1.3, the APD sensitivitity advantage is more pronounced at higher data rates and the use of the use of APDs is also

more prevelant in faster data rate systems. Another technology competing with APDs are erbium doped fiber amplifiers[5]. These can be used to increase the input power on an optical signal immediately before being detected by a receiver. The use of a fiber amplifier together with a PIN receiver can be used for lengths of fiber beyond the reach of APD receivers alone[6]. The cost of optical amplifiers, however, exceeds the cost of APD receivers. As a result, in the current state of optical communication systems, APD based receivers have found an important niche as a relatively low cost component for long-haul fiber applications - a niche that will remain into the foreseeable future.

1.2 Avalanche Photodiodes

An APD has two functions, the absorption and conversion of light to an electrical signal and the amplification of that signal through avalanche multiplication. The avalanche multiplication feature is what separates an APD from all other detectors. These devices are notoriously difficult to construct because they operate at very high voltages and electric fields. Their principle of operation is based on impact ionization in which a carrier, either an electron or a hole, is accelerated at high speeds in an electric field and then collides with an atom in a crystal lattice. The force of the collision ejects an electron from the valance band to the conduction band and leaves behind a hole. The original carrier as well as the newly created ones are then free to repeat this process

within a high field region. Although the principle is the same, the ionization process proceeds differently in different semiconductors. In some semiconductors, electrons are more likely than holes to experience ionization collisions. In other semiconductors the opposite is true and for some materials, electrons and holes are about equally as likely to ionize.

Two examples of the extreme differences possible for ionizing characteristics are displayed by silicon and indium phosphide (InP). Electrons in silicon are much more likely to experience ionizing collisions than holes. In InP, holes are more likely than electrons to ionize but only about twice as likely. Figure 1.4 illustrates the implications of these properties. The figure represents both silicon and InP in a high electric field region. In the case of silicon, an electron entering the region travels a certain distance and then has an ionizing collision and produces an electron-hole pair. The newly created hole is swept out of the high field region without ionizing. (Its probability of ionizing is extremely low). The original and ionized electrons then travel further in the field and both produce additional ionized electron-hole pairs. As before, the holes are swept out of the region without ionizing but the electrons continue on and ionize yet again. This process could of course go on many times and build on itself (avalanche). The created electron-hole pairs translate to additional current flowing out of the region producing what is known as avalanche gain. In general though, very few holes produce ionizing collisions. The case for InP

is much different where electrons and holes are both fairly likely to ionize. This means that an incident electron could create an electron-hole pair and the hole could then produce an ionized pair. From this pair, the electron could then produce an additional pair. This process could, of course, continue in any number of random ways creating additional carrier pairs and avalanche gain. Stated in a general way, the silicon ionization process is much more ordered while the InP process is more random.

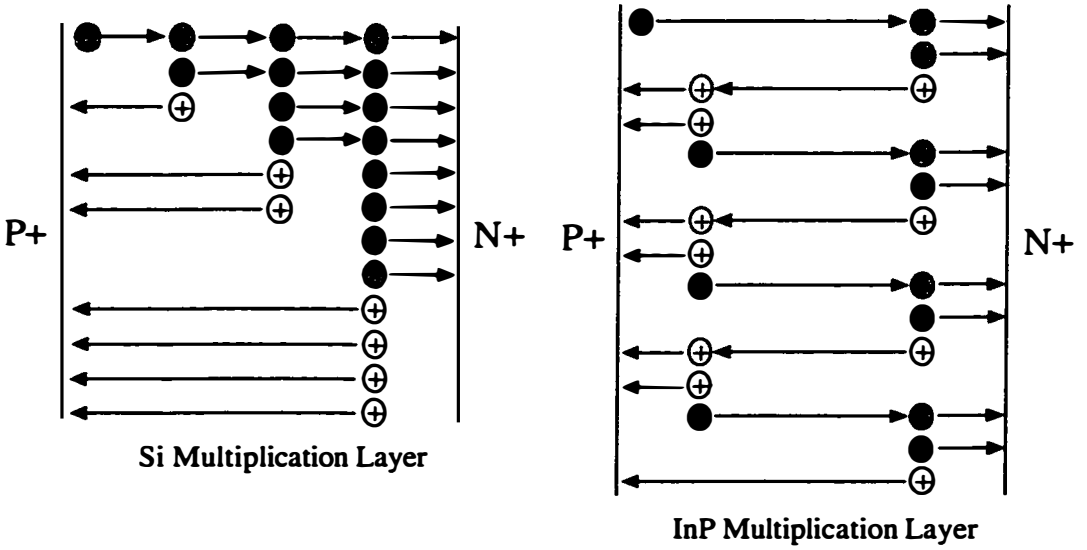


Figure 1.4 - The avalanche multiplication process in silicon and InP. The regions between the P+ and N+ dopings are assumed to be under an electric field so that positively charged holes travel to the left and electrons to the right. The creation of an electron-hole pair signifies an ionization event has occurred. In silicon holes are shown to create no ionization events, in reality they do produce ionizations but with a much

lower probability than electrons. For InP the electron and hole ionization lengths are drawn as equal, but are actually different by a factor of about two.

The more ordered silicon ionization process has many implications for the performance of APDs made from this material versus APDs made from semiconductors like InP. First of all, the multiplication process in silicon happens at a faster rate than in InP. In the example illustrated by Figure 1.4, electrons transit through the high field region in unison and the process is over when the last holes have been swept from the region. In the InP case, electrons and holes are produced at different points within the high field region and at different times after the first carriers are injected. The process is not complete until the last few carriers make it out of the region without ionizing. The whole action is something like a ball bouncing between two walls many times before losing its energy and falling to the ground. Because of this, the InP avalanche multiplication process takes longer to produce the same number of electron hole pairs. In terms of devices, this means APDs built from silicon can operate at higher speeds while maintaining the same gain as their InP counterparts.

Another device characteristic that stems from the ionization properties is the noise involved in the multiplication process. As shown in Figure 1.4, InP ionization is much more random than that of silicon. Some injected electrons could produce 10 ionization events and others no ionizations. This leads to

much greater signal noise associated with InP multiplication versus the more ordered silicon type. In addition, InP detectors are more sensitive to voltage and temperature changes than silicon as will be shown in later chapters.

Creating the APD with the best avalanche multiplication properties essentially means finding a semiconductor with ionization properties similar to that of silicon. That is, a semiconductor in which one of the carrier types, electrons or holes, ionizes much more readily than the other carrier type. The most common measure of the ionization probability of electrons and holes are the ionization coefficients, α for electrons, and β for holes. Figure 1.5 plots the ionization coefficients for four common semiconductors, GaAs, InGaAs, InP and silicon. The ionization coefficient is electric field dependent for all semiconductors as shown. The graph illustrates the large disparity between α and β for silicon, especially at low electric fields. For InP the coefficients have a ratio of about 2:1, also true for InGaAs. In GaAs, α and β are approximately equal. In fact, there is no semiconductor known in which the difference in coefficients is as large as that in silicon, making silicon the first choice for an avalanche multiplication material and the first choice when constructing an APD, no other factors considered.

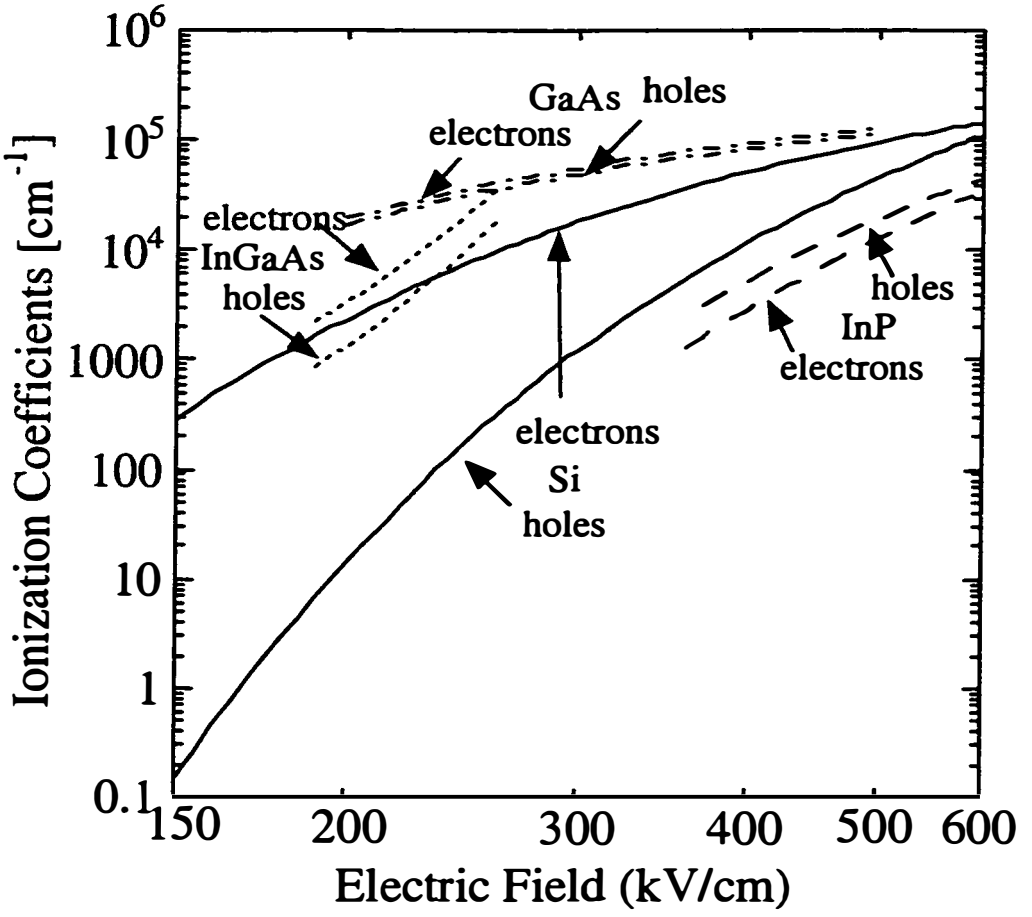


Figure 1.5 - Electrons and hole ionization coefficients versus electric field for the following semiconductors: silicon[7], InP[8], InGaAs[9], and GaAs[10].

1.3 Existing Avalanche Photodiodes

There is one important property that precludes silicon from being the only semiconductor from which APDs were made - the optical absorption

coefficient. This refers to how much light of a given wavelength is absorbed and converted to electron-hole pairs in a given distance. A graph of absorption coefficients for five different semiconductors is shown in Figure 1.6. As can be seen, silicon has fairly significant absorption in the UV and visible wavelengths

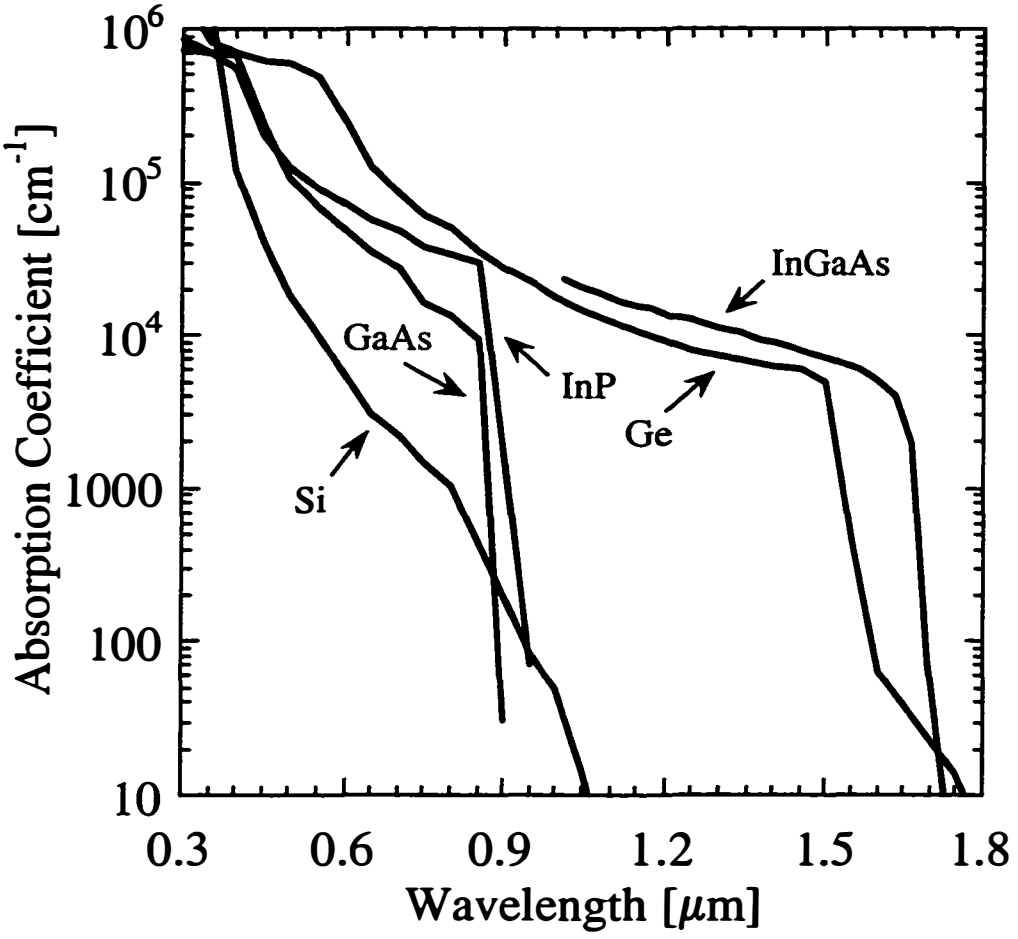


Figure 1.6 - Optical absorption coefficient[11] versus wavelenghth for the following semiconductors: silicon, InP, GaAs, InGaAs, and germanium.

(less than 0.7 μm) but its absorption drops off quickly in the near infrared. By 1.0 μm in wavelength, silicon absorption is almost non-existent. GaAs and InP cut off before 1.0 μm as well. Germanium and $In_{0.53}Ga_{0.47}As$ exhibit significant absorption in the near infrared with InGaAs displaying the highest absorption coefficient in this region. The significant absorption region for InGaAs also includes the wavelengths of 1.3 and 1.55 μm, so important when working with optical fiber.

Because silicon is practically invisible at 1.3 and 1.55 μm, building silicon only APDs to operate there is a bad idea. No matter how good the multiplication process, without any photocurrent generated by an absorption layer there is no optical detection. Very good silicon APDs have been[12] and are being built[13-14], and from the beginning the avalanche multiplication process was defined using silicon[15], but their operation is limited to wavelengths below 1.0 μm. Germanium APDs also developed very early[16] and were able to operate in the infrared regime out to 1.55 μm. As data rates increased and performance specifications became more demanding, $In_{0.53}Ga_{0.47}As$ and its host substrate InP began to emerge as the material system of choice for APDs[17] operating past 1.0 μm. Although InP is less than an ideal material in which to produce avalanche multiplication, InGaAs is readily grown on it through crystal epitaxy, and it does make a better multiplying region than InGaAs alone. Sophisticated designs emerged that separated the absorption and multiplication regions in the

APD (SAM)[18,19] and then provided a grading layer between the regions (SAGM)[20]. Due to their superior performance, InGaAs/InP APDs are the detector of choice in optical communication systems, for which receiver sensitivity is of primary concern[21]. Research on epitaxially grown APDs continues to advance, especially with designs intended for use at 10 GBit/s and beyond. A superlattice APD[22,23] has been proposed and designed that tailors the effective ionization coefficients in the multiplication layer in order to achieve higher speeds and lower noise. Resonant cavity APDs[24] have also been demonstrated that use very thin absorption layers between mirrors to decrease carrier transit times and maintain high quantum efficiency.

1.4 The Optimum APD Using InGaAs and Silicon

The superior avalanche multiplication properties of silicon, and the superb optical absorption properties of InGaAs in the near-infrared lead to an obvious question. What if they could be somehow combined? What about an InGaAs-silicon APD? Such a detector would combine the best of both worlds so to speak. One might speculate on the increased speed of the detector over existing InGaAs/InP APDs. Higher gains might also be possible due to the less noisy amplification in silicon versus InP. Figure 1.7 illustrates both existing APDs and a proposed InGaAs-silicon APD with an InGaAs absorption layer and a silicon multiplication layer. Such a device is certainly attractive, but

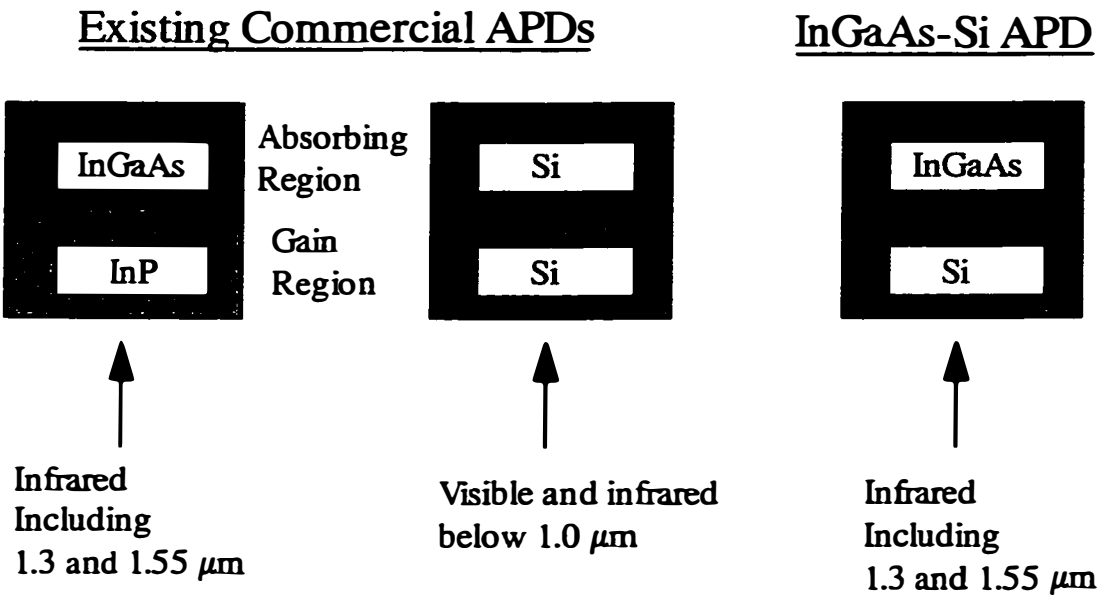


Figure 1.7 - Material composition of existing APDs and the proposed InGaAs-silicon APD. Existing APDs operating in the infrared out to 1.55 μm use an InGaAs absorption and InP gain region. Silicon APDs operating at visible wavelengths use silicon for both absorption and gain, while the proposed APD would use an InGaAs absorption region and silicon gain region.

seemingly impossible to fabricate. InGaAs, which grows readily on InP because of the lattice match between the two semiconductors, does not grow so readily on silicon. Scientists who have tried such a growth have ended up with InGaAs layers full of defects and strains - not tolerable in APD fabrication.

Attempting to create an InGaAs-silicon APD through epitaxial growth most likely would be impossible, no matter how much the growth process was

refined. In the early 1990's however, research was being done on a method that might just allow for the integration of InGaAs and silicon. Semiconductors of different composition and lattice constants were being bonded together successfully using a technique called wafer fusion. Electrically and mechanically stable bonds between semiconductors were reported and suddenly a path for the creation of an InGaAs-silicon APD was opened up. Although a possible method existed, the integration of InGaAs and silicon certainly did not happen overnight. Wafer fusion of individual material systems requires intricacies that only countless experiments can reveal. After about one year of attempts, it did happen however in 1995. Even when semiconductors were being successfully wafer fused, a lot of work was required in learning how to fabricate an APD from these materials. Silicon and InGaAs are not a "natural" fit and neither is the device processing used for them. Most silicon devices are fabricated using very high temperature doping and oxide growth steps, while InGaAs can withstand only relatively low temperature processing. Finding processes that worked for both materials was another barrier to creating working APDs.

1.5 Dissertation Layout

This dissertation was meant to give a broad look at many aspects of the InGaAs-silicon detector. In doing so, about the first half of the chapters are

devoted to the design and theoretical equations that explain the operation of the device. Much is borrowed from existing work on APDs which have been very well explored, but there are also new equations and calculations derived with specific implications for this particular detector. The main purpose of the first chapters is to optimize the design of InGaAs-silicon APDs, delineate inherent limitations, and quantify the advantages over existing detectors. The second part of the dissertation could be termed experimental, with fabrication and measurement results shown for real InGaAs-silicon devices. These results include material considerations like scattered ion mass spectroscopy (SIMS) and tunneling electron microscope (TEM) scans as well as electrical characterizations.

Chapter 2 introduces APD design and reviews existing silicon and InGaAs/InP APDs. Key parameters for the design of an InGaAs-silicon APD are also explored. Chapter 3 lays out the groundwork for the avalanche multiplication process in an APD including the classical APD gain equations. These will be used throughout the chapters that follow to quantify APD gain. Chapter 4 utilizes the gain equations to derive a relationship for the voltage sensitivity of the gain. Temperature sensitivity is also explored and comparisons made for silicon versus other semiconductors. Chapter 5 explores the bandwidth and noise characteristics of APDs and specifically an InGaAs-silicon APD. Calculations show that for these parameters, using silicon for

avalanche multiplication has its biggest advantages. Chapter 6 focuses on wafer fusion, the enabling technology for this new detector and includes characterization of the fused interface between InGaAs and silicon. Chapter 7 reports on the fabrication and measurement of PIN detectors made by fusing InGaAs and silicon. These devices provide an easy way to qualify the fused interface for use as a detector and also help in optimizing the fusion process. Chapter 8 finally reports on three generations of InGaAs-silicon APDs. This includes device design, fabrication, and electrical characterization. Many parameters are reported including a record setting gain-bandwidth number - a benchmark for APD performance.

References

[1] H. Melchior, "Sensitive High Speed Photodetectors for the Demodulation of Visible and Near Infrared Light," *Journal of Luminescence*, **7**, pp.390-414, (1973).

[2] B. L. Kasper and J. C. Campbell, "Multigigabit-per-Second Avalanche Photodiode Lightwave Receivers," *Journal of Lightwave Technology*," **LT-5**, No. 10, pp. 1351-1354, (1987).

[3] B. E. A. Saleh and M. C. Teich, Fudamentals of Photonics, pp. 666-695, John Wiley & Sons, New York, (1991).

[4] B. L. Kasper, J. C. Campbell, A. H. Gnauck, A. G. Dentai, and J. R. Talman, "SAGM avalanche photodiode receiver for 2 and 4 Gbit/s," *Electronic Letters*, **21**, pp. 982-984 (1985).

[5] E. Desurvire, C. R. Giles, J. R. Simpson, and J. L. Zyskind, "Efficient erbium-doped fiber amplifier at a 1.53-μm wavelength with a high output saturation power," *Optics Letters*, **14**, No. 22, pp. 1266-1268, (1989).

[6] S. Saito, T. Imai, T. Sugie, N. Ohkawa, Y. Ichihashi and T. Ito, "Coherent transmission experiment over 2,223 km at 2.5 Gbit/s using erbium-doped fibre amplifiers," *Electronics Letters*, **26**, no. 10, pp. 669, (1990).

[7] P.P. Webb, "Measurements of Ionization Coefficients in Silicon at Low Electric Fields," GE Canada Inc., Electro Optics Operations, unpublished

[8] L. W. Cook, G. E. Bulman, and G. E. Stillman, "Electron and hole impact ionization coefficients in InP determined by photomultiplication measurements," *Applied Physics Letters*, **40**, pp. 589-591, (1982).

[9] T. P. Pearsall, Applied Phys. Lett., **36**, (1980).

[10] H. D. Law and C. A. Lee, "Interband Scattering Effects on Secondary Ionization Coefficients in GaAs," *Solid-State Electronics*, **21**, pp. 331-340, (1978).

[11] Handbook of Optical Constants of Solids, edited by Edward D. Palik, (1985), Academic Press, NY.

[12] R. L. Batdorf, A. G. Chynoweth, G. C. Dacey, and P. W. Foy, *Journal of Applied Physics*, **31**, pp. 1153, (1960).

[13] R. J. McIntyre, "Recent developments in silicon avalanche photodiodes," *Measurement*, **3**, No. 4, pp 146-152, (1985).

[14] K. M. James, M. J. Masterson, and R. Farrell, "Performance evaluation of new large-area avalanche photodiodes for scintillation spectroscopy," *Nuclear Instruments and Methods in Physics Research*, **A313**, pp. 196-202, (1992).

[15] K. G. McKay and K. B. McAfee, *Physical Review*, **91**, pp. 1079, (1953).

[16] H. Melchoir and W. T. Lynch, "Signal and noise response of high speed germanium avalanche photodiodes," *IEEE Transactions on Electronic Devices*, **ED-13**, pp. 829-838, (1966).

[17] N. Susa, H. Nakagome, O. Mikami, H. Ando, and H. Kanbe, "New InGaAs/InP avalanche photodiode structure for the 1-1.6 μm wavelength region," *IEEE Journal of Quantum Electronics*, **QE-16**, pp. 8064-8069, (1980).

[18] N. Susa, H. Nakagome, H. Ando, and H. Kanbe, "Characteristics in InGaAs/InP avalanche photodiodes with separated absorption and multiplication regions," *IEEE Journal of Quantum Electronics*, **QE-17**, pp. 243-250, (1981).

[19] O. K. Kim, S. R. Forrest, W. A. Bonner, and R. G. Smith, "A high gain $In_{0.53}Ga_{0.47}As$/InP avalanche photodiode with no tunneling leakage current," *Applied Physics Letters*, **39**, pp. 402-404, (1981).

[20] J. C. Campbell, A. G. Dentai, W. S. Holden, and B. L. Kasper, "High-Performance Avalanche Photodiode with Separate Absorption 'Grading' and Multiplication Regions," *Electronics Letters*, **19**, No. 20, pp.818-820, (1983).

[21] F. Boudreau, "Optimizing APD's for High-Speed Communications," *Fiberoptic Product News*, **12**, pp. 29-30, (1995).

[22] F. Capasso, W. T. Tsang, A. L. Hutchinson, and G. F. Williams, "Enhancement of electron impact ionization in a superlattice: A new avalanche

photodiode with a large ionization rate ratio," *Applied Physics Letters*, **40**, No. 1, pp. 38-40, (1982).

[23] T. Kagawa, H. Iwamura, and O. Mikami, *Applied Physics Letters*, **54**, pp. 33-35, (1989).

[24] R. Kuchibhotla, A. Srimivasan, J. C. Campbell, C. Lei, D. G. Deppe, Y. S. He, and B. G. Streetman, *IEEE Photonics Technology Letters*, **3**, pp. 354-356, (1991).

Chapter 2

Design Considerations

2.0 Introduction

An avalanche photodetector is certainly one of the most difficult solid state devices to construct. By its nature, it operates in the avalanche multiplication region that designers of other devices do their best to avoid. What others refer to as avalanche breakdown, the APD utilizes as gain. To even reach such regimes, an APD requires high applied biases and consequently high electric fields while in operation. The necessity of very high electric fields often makes it very difficult to maintain low leakage currents through a diode. An added difficulty is that leakage currents required for a good APD are very low compared to other devices, usually in the nanoamp range, even at hundreds of volts of applied bias. Low dark current is of course just one of the many requirements in creating a useful APD. In this chapter these requirements will first be outlined, then a description will be made of existing silicon and InGaAs/InP APD designs, and finally the initial design concepts of an InGaAs-silicon APD will be examined.

2.1 Requirements for a Useful APD

As was mentioned above, one of the first requirements for an APD is a low leakage, or dark current. APDs are usually employed in very low light level applications. The dark current of the detector needs to be at least near the level of the photocurrent generated by incident photons. This light level of course depends on the application, but typical pre-gain dark currents in good commercial detectors are below 10 nA. Achieving such low dark currents requires very careful design and device fabrication as well as excellent semiconductor material. Electric fields also need to be carefully controlled especially in small bandgap materials, and device perimeters given special attention.

A second requirement in any successful detector design is a high quantum efficiency. This parameter quantifies the detector's ability to absorb light and convert it into electrical current. The semiconductor serving as the absorber-converter will absorb a percentage of incident light based on its thickness and absorption coefficient for a given wavelength. Obviously, the thicker the absorbing material, the more conversion of light to electrons. A very thick absorption layer also has its downsides, however. First of all it increases the total voltage needed to bias the detector correctly. In addition, thicker layers make the detector slower due to increased carrier transit times. A detector designer must balance the need for high quantum efficiency and the

necessary operating frequency. Of course the material with the largest optical absorption coefficient available should also be used as the absorber-converter.

Another aspect of the device speed that must be accounted for is the thickness of any multiplication layer in the detector. In a few APD designs, the absorbing semiconductor layer also serves as the avalanche multiplication layer. Most advanced detectors, however, utilize a separate layer for absorption and multiplication. In either case, the multiplication process also has an effect on the maximum operation frequency. Thinner layers lead to higher speeds but can be more difficult to fabricate and have more signal noise. Both absorption layer thickness and multiplication layer thickness must be chosen to fit within the total voltage limitations, speed requirements, and necessary quantum efficiency.

An aspect of device design often overlooked in initial photonic device design is their ease of manufacturing and packaging. An idea may look great in simulation, but unless it can be readily made it will remain forever only on paper. A key part of creating an APD is correctly doping layers of materials, one on top of the other. Achieving the correct doping levels can be very challenging. Slight variations in doping or layer thicknesses can make APD substrates unusable. A reproducible process must allow for reasonable error in the fabrication technologies available. A good APD structure also requires good ohmic contacts to both n and p contact regions to keep the circuit

resistance low. High resistance devices limit the speed of a detector and in APDs can limit the achievable current gains. Finally, all devices must be packaged to operate in the real world. Contact pads need to be large enough to bond to and detector area needs to be large enough to couple light or optical fiber correctly. Device geometry must also be such as to lend itself to easy integration with other electronics.

2.2 Design of Silicon APDs

Silicon avalanche photodiodes have been built for many years and have certainly undergone many refinements[1]. Typical applications for these APDs are at visible wavelengths with very low light levels. To maximize the performance of these diodes, most commercial APDs utilize a separate layer for absorption and one for multiplication. To absorb enough light, most silicon APDs have to be quite thick, on the order of 10 μm or more, with most of the total thickness due to the absorption region. The absorption and multiplication regions are distinguished by the electric field level in each during operation. The absorption region can operate at low fields, usually below 100 kV/cm, while the multiplication region requires higher fields that can range between 200 kV/cm and 500 kV/cm depending on the layer thickness. The electric field is manipulated in silicon by placing a p doped layer between the absorbing and multiplying layers. Under reverse bias, the p doping forces the field higher in

the latter region. Because the device layers are very thick and the required electric fields large, most silicon APDs operate at very high voltages. It is difficult to find this type of detector operating at less than 150 V and many require as much as 1500 V. Despite this high voltage, silicon APDs made of high quality material still maintain very low dark currents and detect very faint signals.

Figure 2.1 illustrates two common silicon APD structures. The first is a planar structure[2,3,4] that is constructed using dopant diffusion or a combination of diffusion and epitaxial growth. In the case of all diffusion, fabrication begins with an undoped or very low doped silicon substrate. The substrate is thinned to the required thickness of the device and boron is diffused into the backside to form a p+ region that serves as the anode for the diode. Once the anode and undoped silicon regions have been formed, a thermal oxide diffusion mask is deposited on the front side of the wafer. Windows are opened and boron is diffused into the top silicon surface, forming a p type layer to a thickness determined by the diffusion time and temperature. A larger window is then opened in the diffusion mask and phosphorous is diffused into the surface. The phosphorous diffusion depth is shallower than the boron but at a higher doping level. This creates a top n+ layer in the silicon that serves as the cathode for the device. The oxide diffusion mask also serves as a passivating layer for the silicon surface and prevents any edge leakage current. Additional passivation

layers and ring contacts are often added to further reduce leakage currents. Aluminum is added as the metallic contact to both the cathode and anode with a window in the anode for illumination.

When operating, the planar diode is reversed biased. This creates a large triangular shaped field profile in the p-doped region between the cathode and undoped silicon. When the bias is high enough, all of the p doping will be depleted and the field will quickly deplete through to the p+ silicon anode. Further bias increases the field in both the p region and the undoped region. The undoped region serves as the light absorbing region, and the p-doped region is the avalanche multiplier. Photons are incident through the thin p+ anode on the absorbing region and converted to electrons. Due to reverse bias, these electrons are swept towards the p doped multiplication region. Here the fields are much higher than in the absorption region and ionization occurs leading to avalanche gain. One of the main advantages of this structure is that the electric fields are kept isolated away from any surfaces due to the planar structure. In effect the device is encapsulated by the high resistivity, undoped layer. Of course one of the drawbacks of the structure is that the absorbing layer must be very wide to allow for any wafer handling. Variations can of course be made to the basic planar design, including top illuminated detectors constructed using epitaxial growth on n+ substrates.

Another class of silicon structures, the beveled APD[5], uses etching and epitaxy to create a detector as illustrated in Figure 2.1. Fabrication begins with an n+ substrate. A p-doped layer is grown first which will serve as the multiplication layer. A thick layer of undoped material is then grown which serves as the detector's absorption layer. Finally a p+ layer is added to the top of the structure to serve as an anode. Individual devices are made by first depositing a masking layer on the back n+ silicon. Mesas are etched into the mask layer and then the silicon is wet etched around the mesa mask. The silicon etchant naturally undercuts the mask to form the desired bevel shape around the structure. After etching, the bevel edge is encapsulated to protect it from moisture and aluminum contacts are deposited on the cathode and anode layers. A window is maintained in the aluminum contacting the p+ region so that light can penetrate to the absorber. In operation the beveled APD is reversed biased like the planar structure. Incident photons strike the silicon creating electron-hole pairs. The electrons migrate toward the p silicon multiplication region under a high electric field. The APD is designed to be broader at the anode and smaller at the cathode so as to prevent any high field lines at the etched surface. The electric field is effectively isolated away from any edges as in the planar APD case. An advantage of this structure is that it does not rely on any hard to control diffusion steps. Its disadvantage is that it does not lend itself well to wafer level processing. Once the bevel has been

created in the diode, it has separated itself from the rest of the wafer. These types of APDs must usually be large enough to handle individually. The planar and beveled structures outlined here are of course only a sample of the large array of available structures. The important p+ feature that most of them have in common however is the doping layer sequence: to undoped to p then to n+.

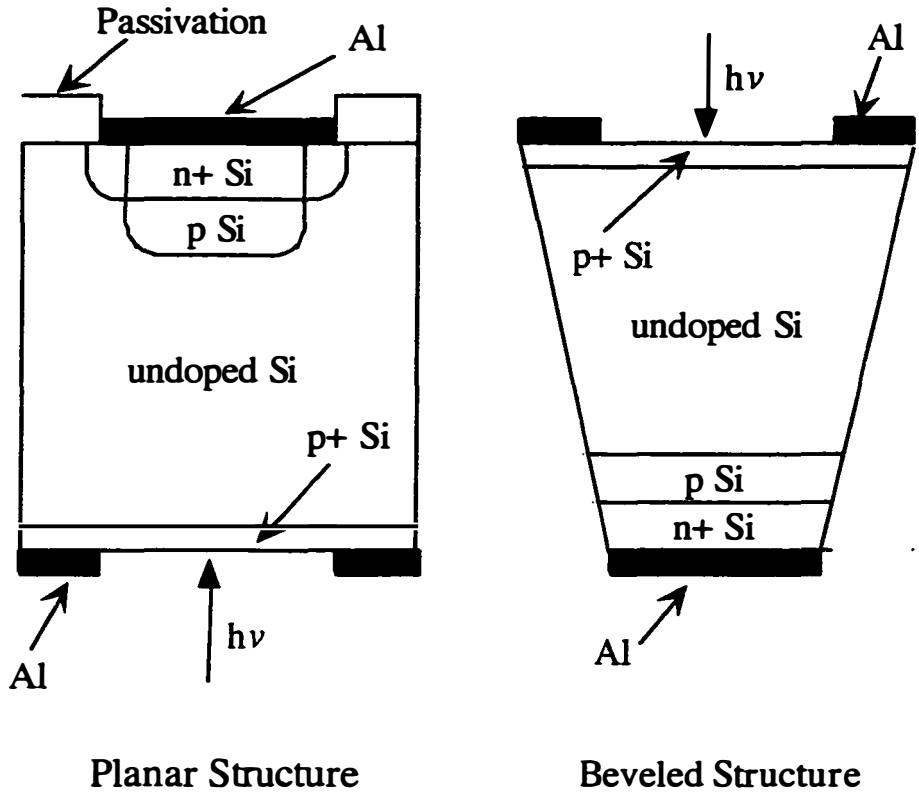


Figure 2.1 - Cross-sections of two types of silicon APDs. The planar structure provides device isolation though dopant diffusion while the beveled structure relies upon chemical etching.

2.3 Design of InGaAs/InP APDs

InGaAs/InP layer APDs were created to do the job silicon APDs could not - detect light in the near infrared. With the rise of optical fiber communications has come the rise of the InGaAs/InP APD to detect the critical wavelengths of 1.3 and 1.55 μm. Had the low dispersion and loss wavelengths in fiber been in the visible, it would be safe to say that InGaAs/InP APD would be a rather uninteresting, neglected device. The laws of nature make the rules, however, and the demand for higher performance optical systems has fueled the demand for better and better near-infrared APDs. Like silicon APDs, the III-V detectors utilize separate regions for absorption and multiplication (SAM APD). The heterojunction between InGaAs and InP has also led to a third specialized layer, an InGaAsP grading layer to grade the valance band between the two. Modern commercial devices are usually referred to as SAGM (Separate Avalanche Grading and Multiplication layer) APDs[6,7]. These layers are not only distinguished by their doping as in silicon, but also their material composition. Using a separate absorption layer allows for a very abrupt injection point of photogenerated carriers into the multiplication layer. The layers are also optimized for their functions - InGaAs as an absorber, InGaAsP for grading, and InP as the best multiplier lattice matched to InGaAs. In real world applications these detectors are run with gains of around ten, which is optimal for the performance of the device. Gains higher than ten create too much signal

noise. Typical operational voltages are from 50 to 75 Volts and the active layers are much thinner than in most silicon APDs. Layers are thin to obtain the necessary operation frequencies that can reach as high as 10 GHz in some InGaAs/InP structures. The design of a good InGaAs-InP APD is in many ways opposite to the design of a silicon APD. In silicon, the desired avalanching carriers are electrons because the ionization coefficient is larger for electrons than holes. InP represents the opposite case, where holes do most of the ionizing. InP based APDs are designed to force holes into a high electric field.

Figure 2.2 shows two designs for an InGaAs/InP APD. The first is a planar structure most common in commercial devices, especially those operating at 2.5 GHz and below. The most difficult aspect of building this type of detector is precise epitaxial growth required to get the layer dopings and thicknesses correct. The growth starts with an n+ InP substrate on which a very low doped n- InP buffer layer is grown. This is followed by a low doped n- $In_{0.53}Ga_{0.47}As$ layer that serves as the absorption region. This is followed by an n type InGaAsP grading layer which maintains the InP lattice constant but grades the valance band between InP and InGaAs. Above this in an n doped InP layer that eventually serves as the multiplication layer for the APD. The thickness and doping of this layer is the most critical. Dopings that are too high yield detectors that break down before depleting all the way through the device.

Dopings that are too low create fields in the absorption layer that are too high when the detector is biased to desired gain levels.

Upon the successful epitaxial growth of an APD layer structure, individual detectors are created by first depositing a dielectric layer on the top surface. The dielectric serves both as a surface passivation and as a diffusion mask. Windows are etched in the dielectric and the p dopant zinc is diffused into the top InP layer. The temperature and time of the process can tailor the total amount of zinc diffused. With the proper timing, a p+ layer is created in the InP leaving just the right thickness of n type InP below it from the original layer. The p+ InP serves as the anode for the detector and n-type InP layer sandwiched between the anode and the grading layer forms the multiplication region of the APD. A p metal contact such as Au:Zn is deposited onto the anode region, and an n metal contact like Au:Ge is deposited on the n+ substrate which serves as the cathode. Windows are maintained on the backside of the wafer to allow infrared light to pass through the InP substrate, which is invisible to the near-IR, and be absorbed in the InGaAs region. The detector is operated in reverse bias. As voltage is applied, the low doped n- InP and InGaAs regions are quickly depleted through. The electric field profile then rises linearly as the InGaAsP layer is depleted through. The electric field rises linearly again as the n doped InP multiplication layer is biased to depletion. Upon depletion of the n charge between the anode and cathode, the electric field

has the desired shape with the highest field present in the avalanche multiplication layer and the lowest field in the absorbing layer. The field is confined laterally by the p-n junction formed by the zinc diffusion as well as the low resistivity n- InGaAs and InP.

Figure 2.2 also illustrates a mesa type InGaAs/InP APD. This detector also utilizes sophisticated epitaxial growth. The layers are much the same as those for the planar device except that the growth is done on a p+ substrate. The n InP multiplication layer is grown first followed by the n InGaAsP grading layer, the n- InGaAs absorption layer, and finally an n- InP buffer layer and n+ InP cathode. Individual detectors are made by depositing a mask layer on the wafer surface then etching mesas in the mask. A mesa structure is then etched through the III-V layers down to the p+ InP substrate. The semiconductor etchants naturally undercut the top mask layer to give the device a desirable shape. Much like the beveled silicon APD, when this mesa detector is reversed biased, high field lines will remain away from the etched surfaces due to its geometry. After etching, the mesa sidewalls are encapsulated by a dielectric layer and metal contacts applied to the cathode and anode. Like the planar APD, the mesa APD is backside illuminated through the InP substrate. The main advantage of the mesa structure is that it eliminates the need for any dopant diffusion. This allows for more precise control of the layer thicknesses

and dopings - very critical for the highest speed APDs where layers are made very thin.

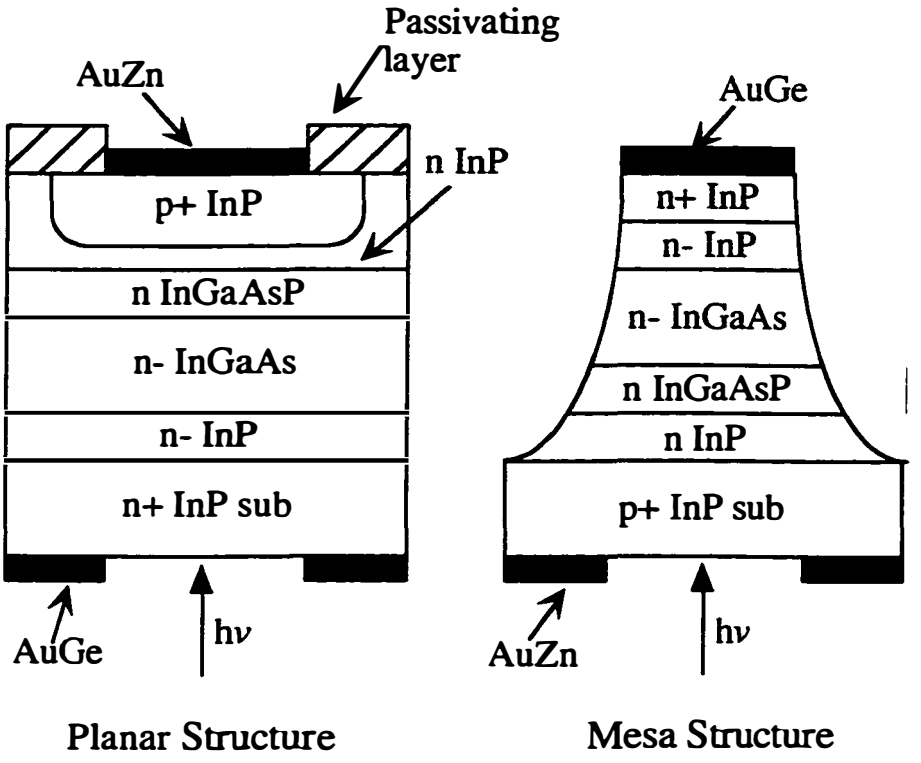


Figure 2.2 – Cross-section of two types of InGaAs/InP APDs. Both devices rely on epitaxy to create the layer structure. The planar structure provides device isolation through dopant diffusion while the mesa structure is created through chemical etching.

2.4 Designing an InGaAs-Silicon APD

The design for the fused InGaAs-silicon APD borrows from both the silicon and InGaAs/InP structures. The fused detector will use electrons as the

main ionizing carrier so in this respect the layer structure resembles a silicon APD with electrons being forced from the absorbing layer into the multiplication layer. Like the InGaAs/InP APD, the fused APD uses absorption and multiplication layers of completely different semiconductors and can utilize

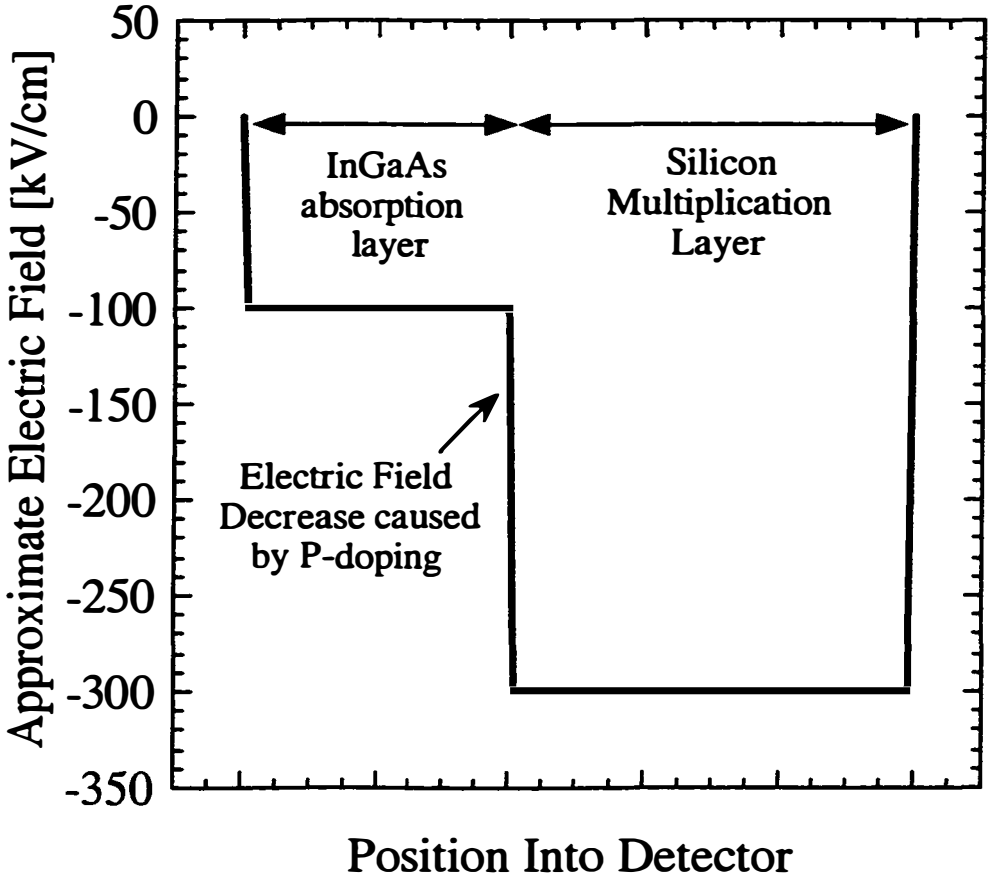


Figure 2.3 - Desired electric field profile in an InGaAs-silicon APD. The field in the InGaAs should be kept below 100 kV/cm. The field in the silicon multiplication layer will vary according to the layer thickness but in most cases should be over 250 kV/cm.

sharp material interfaces created by epitaxial growth. Without going into great detail about the specific parameters of InGaAs-silicon APD fabrication which is covered in Chapter 8, the remainder of this chapter will focus on what could be termed the "doping knob" parameter in the APD. Like its silicon and InGaAs/InP counterparts, the InGaAs-silicon APD requires a high electric field in the silicon multiplication layer, while maintaining a low field in the InGaAs absorption layer. This will be accomplished by placing a p doping between the InGaAs and the silicon layers. The total dose of the doping will essentially determine the difference in the electric field profiles of the two regions. The desired field profile is illustrated in Figure 2.3. To avoid tunneling breakdown in the InGaAs, the field needs to be kept below 100 kV/cm, while the field in the silicon is made high enough to reach the desired avalanche gains. Although the necessary field is thickness dependent, fields of around 300 kV/cm are required for thicknesses around 3.0 μm.

There are several options in applying the "doping knob" or p doping layer between the silicon and InGaAs layers. The first possibility is the growth of a p type layer on the top of the silicon multiplication layer. When the InGaAs is later integrated onto the wafer, the p layer will be sandwiched between. The drawback to this approach is that epitaxial growth of the desired silicon p layer is very hard to control. Most silicon epitaxial processes occur at temperatures around 1000°C and have been designed for rapid deposition of

thick layers. To have an effective "doping knob" the amount of doping in the p layer has to be very precisely controlled in a very thin layer, requirements not in harmony with most silicon growth reactors. Another possible solution that is more viable is the implantation of the desired p type doping. Implantation doses can be specified and measured very accurately. Their doping profiles are also well understood and routinely modeled. This method also allows itself more easily to "tuning" of the correct doping where identical wafers can be implanted with different doses to hone in on just the right electric field profile.

2.5 Feasibility of InGaAs-Silicon APD Construction

Using ion implantation as a "doping knob" in the InGaAs-silicon APD, a calculation can be done to examine the feasibility of creating such a detector given the current state of growth and processing technology. To make this calculation, the following assumptions are made. First the detector structure consists of an InGaAs layer on a silicon epitaxial layer targeted at 2.5 μm thick. The total field in the InGaAs must remain below 100 kV/cm. The operating gain in the silicon is assumed to be 30. The parameters that will have some process variation are the epitaxial thickness and the implant dose. The doping transition from the low-doped epitaxial layer to the highly doped substrate will not necessarily be sharp. There is typically a 500 nm transition region between doping levels of 10^{18} cm^{-3} and 10^{14} cm^{-3}. Any variation in this transition region

is assumed in the thickness variation of the epitaxial layer. Evaluating how much the layer thickness and implant can vary will consist of first determining the necessary electric field to achieve a gain of 30 for each thickness between 2.3 and 2.7 microns, an 8% thickness variation from the target. This is done using known ionization coefficients and the gain equation derived in Chapter 3[8]. Once the necessary electric field E_T is calculated, an implant dose can be calculated to create such a field. For the device to operate correctly, the implant dose must create a voltage difference between the InGaAs and the silicon between E_T and E_T - 100 kV/cm. This will ensure that a gain of 30 can be achieved in the detector and the full implant dose will be depleted under operational bias, allowing carriers to flow freely between the InGaAs and silicon layers. The values for E_T and E_T - 100 kV/cm establish a window of operation for the implant dose at a given thickness.

The results of the feasibility calculation are shown in Figure 2.4. The top solid line illustrates the maximum implant dose allowed - that which would create a field of E_T in the silicon layer. In this case there is very little field across the silicon but the implant dose has been completely depleted. The bottom line illustrates the minimum implant dose allowed, which would create a field of E_T - 100 kV/cm. This latter case would allow for an additional 100 kV/cm of electric field to be placed on both the InGaAs layer and the silicon layer so as to reach the desired avalanche gain. As can be seen from the graph,

the 100 kV/cm window allows for a relatively loose accuracy for the implant doping. The dashed line indicates the variation of implant allowed so as to allow for desired operation for any thickness between 2.3 and 2.7 μm. This 8%

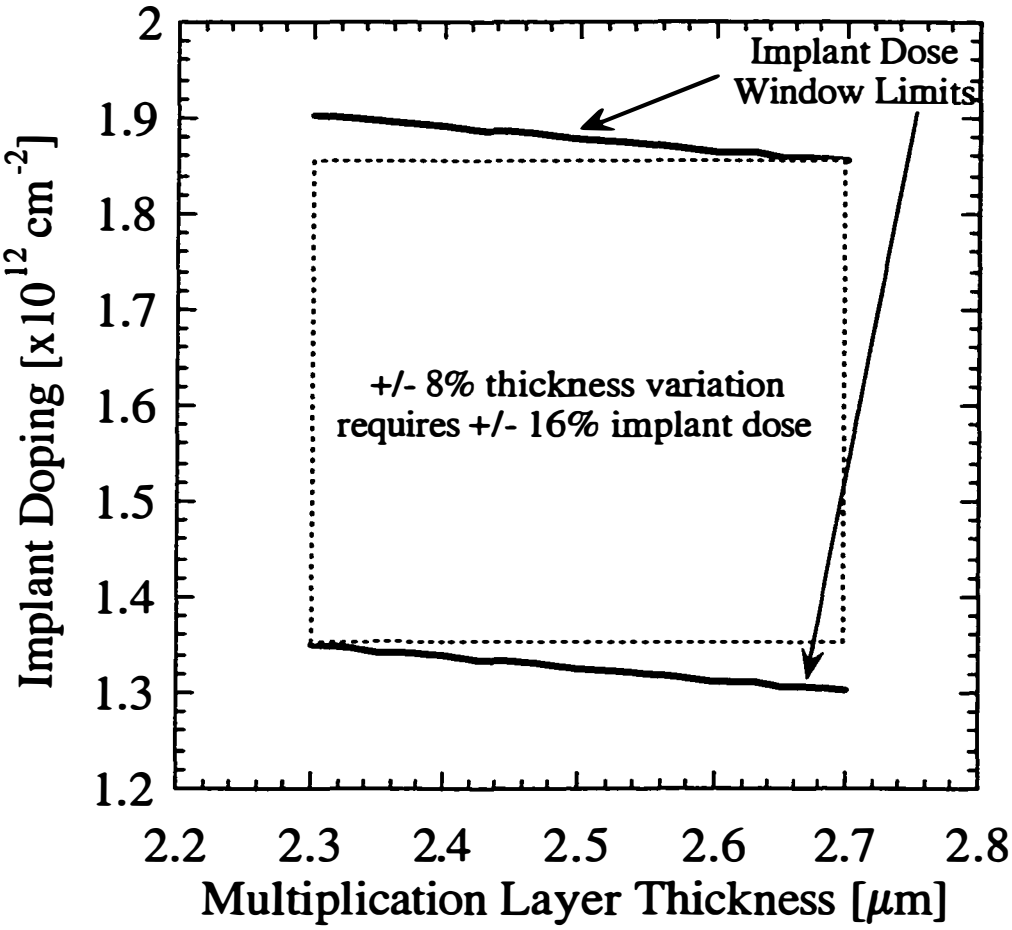


Figure 2.4 - Construction feasibility calculation for InGaAs-silicon APDs created through ion implantation. The two solid lines represent the maximum and minimum implant dosage allowed to make a correctly operating detector for a given multiplication layer thickness.

thickness variation requires only better than a 16% implant dose accuracy. Both of these variations are well within the specified ranges for commercial epitaxial growers and implantation houses. Typically growers will commit to a +/- 3-4% accuracy and implanters to +/- 2%. It appears that the InGaAs-silicon structure is well within the standard technology limits.

It is interesting to compare the tolerances necessary to create an InGaAs-silicon APD with those of creating an InGaAs/InP APD. The difficulty in creating the InP APD is getting the InP multiplication layer thickness and doping just right, as illustrated in Figure 2.2 above. To determine how accurate these need to be, calculations were made for the field necessary to obtain a multiplication gain of 10 for InP at different thicknesses around a target thickness of 0.7 μm. This thickness is close to where commercial APDs are grown, and as shown in Chapter 5, will have gain-bandwidths similar to the thicker silicon layer represented above. The same 100 kV/cm electric field window is assumed as for the InGaAs-silicon APD. This leads to a doping range required for a given thickness. The target is for all the doping in the n doped region to be depleted when the necessary electric field falls within the 100 kV/cm window. Again known InP ionization coefficients[9] were used in the calculation along with the gain equation from Chapter 3. Figure 2.5 shows the required doping versus thickness. As can be seen there is not much room for variation with the epitaxial growth. Even a 5% thickness variation requires a

better than 2% doping level accuracy. These are both very challenging specifications on a reactor, and make the implantation scheme for the InGaAs-silicon detector very attractive.

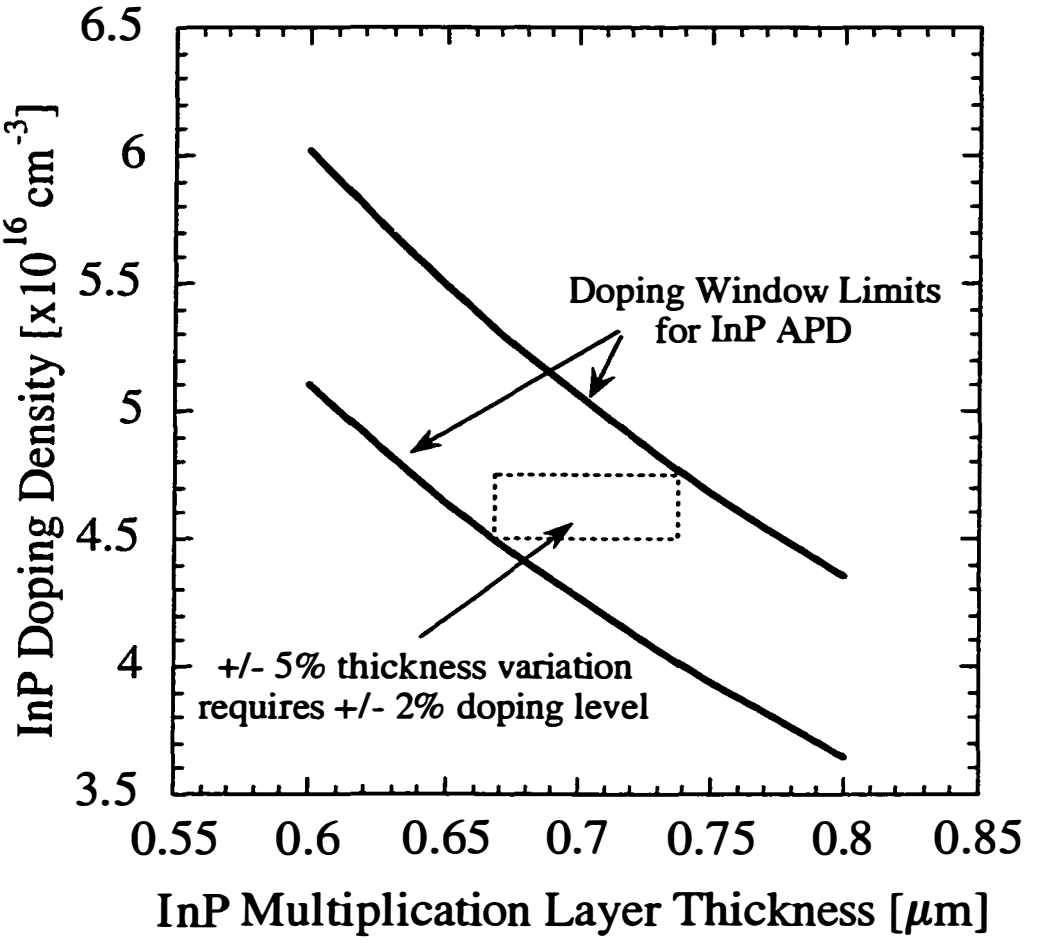


Figure 2.5 - Required doping accuracy in the InP multiplication layer of an InGaAs/InP APD. The top and bottom solid lines represent maximum and minimum dopings allowed to create an operational detector at a given multiplication layer thickness. The multiplication layer was assumed to have a constant n-type doping.

2.6 Conclusions

In this chapter designs for existing silicon and InGaAs/InP APDs were reviewed. With these devices as a springboard, a very important parameter of the InGaAs-silicon APD was examined - the implanted p dose which serves as a doping knob for the detector. Using different implant doses the electric field can be manipulated and optimized. Results of calculations for the feasibility of this approach were illustrated and discussed. It appears that using an implantation will allow the InGaAs-silicon hybrid to be fabricated with much greater control than currently exists with InGaAs/InP devices.

References

[1] R. J. McIntyre, "Recent developments in silicon avalanche photodiodes," *Measurement*, **3**, No. 4, pp 146-152, (1985).

[2] P. P. Webb and R. J. McIntyre, "Multi-element Reachthrough Avalanche Photodiodes," *IEEE Transactions on Electron Devices*, **ED-31**, No. 9, pp. 1206-1212, (1984).

[3] A. Ataman and J. Muller, "Double-mesa reach-through avalanche photodiodes with a large gain-bandwidth product made from thin silicon films," *Journal of Applied Physics*, **49**, No. 10, pp. 5324-5331, (1978).

[4] K. Berchtold, O. Krumpholz, and J. Suri, "Avalanche photodiodes with a gain-bandwidth product of more than 200 GHz, "*Applied Physics Letters*, **26**, No. 10, pp. 585-587, (1975).

[5] K. M. James, M. J. Masterson, and R. Farrell, "Performance evaluation of new large-area avalanche photodiodes for scintillation spectroscopy," *Nuclear Instruments and Methods in Physics Research*, **A313**, pp. 196-202, (1992).

[6] J. C. Campbell, A. G. Dentai, W. S. Holden, and B. L. Kasper, "High-performance Avalanche Photodiode with Separate Absorption 'Grading' and Multiplication Regions," *Electronics Letters*, **19**, No. 20, pp. 818-820, (1983).

[7] R. D. Dupuis, J. R. Velebir, J. C. Campbell, and G. J. Qua, "InP/InGaAsP/InGaAs Avalanche Photodiodes with Separate Absorption, Grading, and Multiplication Regions Grown by Metalorganic Chemical Vapour Deposition," *Electronics Letters*, **22**, No. 5, pp. 235-236, (1986).

[8] Quick Reference Manual for Silicon Integrated Circuit Technology, edited by W. E. Beadle, J. C. C. Tsai, and R. D. Plummer, 1985, John Wiley & Sons, Inc., NY.

[9] C. A. Armiento and S. H. Groves, "Impact ionization in (100)-, (110)-, and (111)-oriented InP avalanche photodiodes", *Applied Physics Letters*, **43**, No. 2, pp. 198-200, (1983).

Chapter 3

Avalanche Photodiode Gain Equations

3.0 Introduction

Although it is a diode, with only two contacts, an avalanche photodetector is still a very sophisticated semiconductor device. It distinguishes itself from most detectors because of its internal gain arising from the avalanche effect. It is different from devices like light sensitive phototransistors because it is has only two contacts. Due to this diode structure it is possible to inject very small currents into APD multiplication layers and still see gain. Attempting to simulate an APDs operation by using a PIN detector and an amplifier will not produce the same results. If low light levels are input, the small current created in the detector will be lost in the noise and leakage current of the amplifier. The gain produced in an APD can have very little noise and can occur very fast, making APDs applicable both to low signal as well as high frequency applications.

The method for producing gain in an APD is of course avalanche multiplication. This effect occurs in all known semiconductors when they are subjected to high enough electric fields. Designers of most devices try hard to avoid avalanching. In transistors for example a thin region reverse biased too

high can create huge leakage currents and the device is said to have broken down. The same effect then takes on two different names. If it is desired, as in an APD, it is referred to as avalanche multiplication, if it is avoided, which is most of the time, it is called avalanche breakdown. Whatever the name, the effect showed up very early in the history of semiconductor devices and so has been well studied. In this chapter, a brief description of the avalanche mechanism will be explained as it relates to impact ionization. Ionization coefficients will then be presented and their role in the avalanche gain equations derived. Finally the method of measuring ionization coefficients for different semiconductors will be outlined.

3.1 The Gain Mechanism

Impact ionization is the basic mechanism whereby avalanche gain may be created. This effect occurs in a semiconductor when a free carrier, electron or hole, is accelerated to high velocities in a high electric field. Such high-energy carriers may collide with atoms within a crystal lattice. If their energy is high enough, the impact can push an electron from a valance site on the impacted atom into the conduction band. This jump creates a vacancy or hole in the valence band. There is a minimum amount of energy needed by an impacting electron in order to create this electron-hole pair. The minimum energy, E_{min}, depends on the bandgap of the material between the valence and

conduction bands. It is no surprise then that small band gap semiconductors like InAs require very low electric fields to initiate ionizations, whereas wide-gap semiconductors like GaP require much higher fields.

The minimum ionization energy, E_{min}, is different for electrons and holes in most semiconductors. In Chapter 1, silicon was alluded too as the most desirable ionizing semiconductor based on the fact that electrons ionized in silicon much more readily than holes. There is no simple explanation for why this is and no one criterion that makes some semiconductors desirable ionization materials and others undesirable. Ionization energies are very complex functions of band structure, densities of state, phonon interaction, and effective masses. Many methods have been presented to calculate these energies and have been compiled and summarized by Capasso[1]. The one feature in silicon that may distinguish it from other semiconductors is its second conduction band located very close to the first, effectively increasing the electron state density. Because there a difference in E_{min} for electrons and holes in most semiconductors, ionization rates or coefficients are often used in describing the process for the two carriers. α is used to describe the electron ionization coefficient and β the ionization coefficient for holes. These quantities refer to the reciprocal of the average distance a carrier must travel in an electric field to acquire enough energy to produce an ionizing collision. α and β have units of cm^{-1} and are strongly dependent on the strength of the

electric field. Important to note too is that they are really a statistical average of the necessary distance for ionization collisions. Impact of an electron with an atom in a lattice is a random event and does not necessarily occur after any given length of travel. The only requirement is that the electron has traveled a sufficient distance to acquire the minimum energy E_{min}. Only by averaging the distance traveled by many of electrons can accurate ionization coefficients be arrived at. Computing ionization coefficients based on models of a semiconductor's band structure can be very difficult. Experimental measurements based on methods described in Section 3.3 are more desirable when constructing real devices.

The previous paragraph describes what could be termed a classical view of impact ionization in which a semiconductor has an intrinsic set of ionization coefficients for a given electric field. It is worth noting, however, that recent work[2,3] has shown that this may not be an accurate approach in some situations. Very thin multiplication layers under high electric fields are not necessarily well described by classic coefficients. Such layers might require considering non-localized effects that account for a carrier's path of travel to accurately calculate the avalanche gain mechanism. In general, layer thicknesses below 200 nm would fall into this category. The calculations done in this dissertation rely on classical ionization coefficients due to the relatively thick multiplication

layers of interest here. As layer thicknesses decrease though, a non-localized approach will become more and more important.

Based on the ionization impact mechanism, the concept of avalanche gain quickly follows. Carriers in a high field region travel an average distance $1/\alpha$ or $1/\beta$ and build up enough energy to cause an ionization event. If an electron initiated that event, afterwards two electrons and a hole exist. The electrons are then accelerated in the opposite direction of the electric field while the hole is accelerated with the electric field. After another average distance $1/\alpha$ for the electrons and $1/\beta$ for the holes, three new ionization events take place creating three new electron hole pairs yielding a total of five electrons and four holes. This process can then continue until the carriers have run out of room in the high field region and can then create no more ionization events. When the electric field region is long enough to allow for at least one ionization to occur, multiplication is said to have taken place because what began with only one electron or hole has turned into the original carrier plus an electron-hole pair. Because the process is exponential in its mechanism, if the electric field region is long enough to allow for multiple ionizations to occur, the number of electrons and holes in the region can multiply very rapidly, or avalanche. Computing the total number of electrons and holes created can quickly become overwhelming. This is especially complicated by the electrons and holes flowing in opposite directions. Holes created by an original

electron's ionization event can produce electrons in a subsequent ionization that lag behind the original electron and must be accounted for. Trying to account for each and every electron and hole is possible in a Monte-Carlo simulation, but there is an easier approach. Assuming that many electrons and holes are first injected into a high field multiplication region and many more are created there, the average current injected and then produced can be used to compute and define a total current gain based on the ionization mechanism. This approach will make use of the ionization coefficients with their built in statistical averages and is derived in Section 3.2.

3.2 Deriving the Gain Equations

In order to derive an equation to describe the gain produced by the avalanche multiplication process, we follow the analysis of Webb et. al.[1-4]. This is an averaged approach with conventional ionization coefficients, non-local theories[2,3] are more complicated and are not necessary in the long avalanche regions described here. To begin with, the currents entering and leaving a multiplication region will be closely studied. In this derivation, the multiplication layer is assumed to be of high quality (few defects) and the current densities through the layer are assumed to be small, reducing any effects of electric field screening by carriers. Figure 3.1 represents a semiconductor region with an electric field across it. In this particular diagram, the electric

field is in the direction opposite that of the directional variable x. Electrons in the field will drift in the direction opposite the electric field, while holes will drift with the electric field. All currents into and out of the region will be in the direction of the electric field, with $J_p(d)$ and $J_p(0)$ the current densities for holes into and out of the field region and $J_n(d)$ and $J_n(0)$ the current densities for electrons into and out of the region. The total width of the region under an electric field is d and the electric field strength is taken as arbitrary and can vary as a function of x.

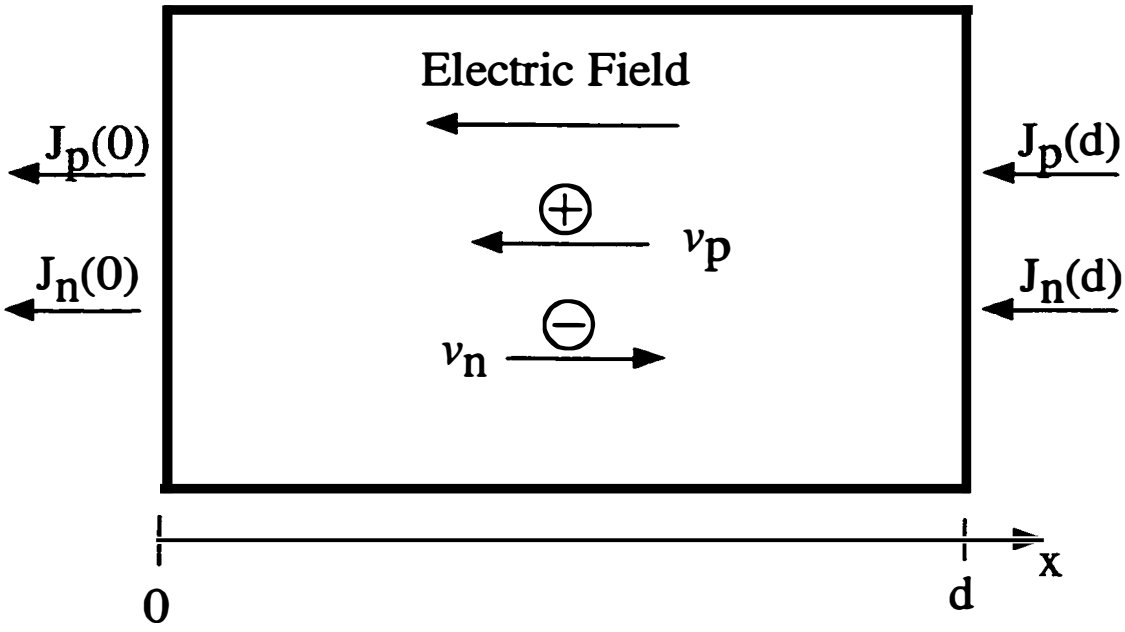


Figure 3.1 - A semiconductor region under an electric field such that holes travel in the -x direction and electrons in the +x direction. The width of the region under the electric

field is d and $J_p(0)$, $J_n(0)$, $J_p(d)$, and $J_n(0)$ represent the current densities for electrons and holes into and out of the region.

The first step in the gain equation derivation is to write down the differential equations that describe what is happening within the electric field region. The hole current and electron currents change within the region but the magnitude of their rate of change must remain the same for every position x. This is because for every conduction band electron that is created, there must be a hole created. For the electron current, the change in the current at a position x can be written as

$$dJ_n/dx = \alpha(x)J_n(x) + \beta(x)J_p(x). \qquad (3.1)$$

The first term in the equation represents the change created by ionizing electrons. The $\alpha(x)$ term is the ionization coefficient for electrons described earlier and in this case is x dependent because the electric field may vary with x. This term is multiplied by the electron current $J_n(x)$. The $\beta(x)$ term is the ionization coefficient for holes which is also x dependent and is multiplied by the hole current at x, $J_p(x)$. Terms for electrons and hole initiated ionization must be included because in both cases an electron is produced. The differential equation for holes is very similar to that of electrons except for a sign change indicating the opposite direction of travel for these carriers. This differential equation is written as

$$-dJ_p/dx = \alpha(x)J_n(x) + \beta(x)J_p(x) \quad (3.2)$$

Again there are terms representing electron and hole initiated ionization because both cases lead to the production of a hole.

In solving these differential equations two cases will be important for APDs, pure electron and pure hole injection into the high field region. Most SAM APD structures simulate this situation well using one layer to absorb light and convert it into electrons and holes, and another layer to multiply either electrons or holes that are injected into it. The first case to be examined is pure electron injection. Referring to Figure 3.1, the electron current is injected at x=0 and increases with increasing x, while at d their is no hole injection so $J_p(d)=0$. To maintain charge neutrality under dc conditions the total current must remain constant at any point x so that

$$J = J_n(x) + J_p(x) = J_n(d) = constant. \quad (3.3)$$

The gain created by ionization for electron injection within the region from 0 to d is defined by the amount of charge created and exiting the region at x=d divided by the electrons injected at x=0, or

$$M_n = J_n(d) / J_n(0). \quad (3.4)$$

To solve the differential equation for electron injection, Equation 3.3, is substituted into Equation 3.1 to become

$$\frac{dJ_n(x)}{dx} - (\alpha(x) - \beta(x))J_n(x) = \beta(x)J_n(d). \quad (3.5)$$

This is an ordinary differential equation with a solution given by the following:

$$J_n(x) = \frac{\int_0^x \beta J_n(d) \exp\left[-\int_0^y (\alpha(x') - \beta(x'))dx'\right] dy + J_n(0)}{\exp\left[-\int_0^x (\alpha(x') - \beta(x'))dx'\right]} \tag{3.6}$$

Given the definition of gain from Equation 3.4, Equation 3.6 can be written for the case x=d as

$$M_n = \frac{J_n(d)}{J_n(0)} = \frac{1}{\exp\left[-\int_0^d (\alpha(x) - \beta(x))dx\right] - \int_0^d \beta(x) \exp\left[-\int_0^x (\alpha(x') - \beta(x'))dx'\right] dx}. \tag{3.7}$$

This equation can be written in a more simple form utilizing integration by parts. The second term in the denominator can be rewritten as follows:

$$\int_0^d \beta(x) \exp\left[-\int_0^x (\alpha(x') - \beta(x'))dx'\right] dx = \int_0^d \beta(x) \exp\left[\int_0^x \beta(x')dx'\right] \exp\left[-\int_0^x \alpha(x')dx'\right] dx \tag{3.8}$$

Integrating by parts this becomes

$$\int_0^d \beta \exp\left[\int_0^x \beta dx'\right] \exp\left[-\int_0^x \alpha dx'\right] dx = \left[\exp\left[\int_0^x \beta dx'\right] \exp\left[-\int_0^x \alpha dx'\right]\right]_0^d$$

$$-\int_0^d \alpha \exp\left[\int_0^x \beta dx'\right] \exp\left[-\int_0^x \alpha dx'\right] dx$$

$$= \exp\left[-\int_0^d (\alpha-\beta)dx\right] - 1 + \int_0^d \alpha \exp\left[-\int_0^d (\alpha-\beta)dx'\right]dx. \qquad (3.9)$$

Substituting this relation into Equation 3.7 yields the final generalized formula for injected electron gain

$$M_n = \frac{1}{1 - \int_0^d \alpha \exp\left[-\int_0^x (\alpha-\beta)dx'\right]dx}. \qquad (3.10)$$

A similar equation can be derived for hole initiated gain that is given by

$$M_p = \frac{1}{1 - \int_0^d \beta \exp\left[\int_x^d (\alpha-\beta)dx'\right]dx}. \qquad (3.11)$$

Both of these equations are very general for injected carriers entering a multiplication region and allow for a varying electric field. A special case that can also be derived is when the electric field throughout the multiplication layer is constant. This case is considered by using constant values for the ionization coefficients α and β. Constant field multiplication regions are a good approximation in many APDs, especially those made from silicon because it is possible to create layers with low background impurity densities. For constant coefficients, the gain equation 3.10 and 3.11 become:

$$M_n = \frac{1}{1 + \frac{\alpha}{\alpha-\beta}\left[\exp(-x(\alpha-\beta))\right]_0^d} = \frac{(\alpha-\beta)\exp(d(\alpha-\beta))}{\alpha-\beta\exp(d(\alpha-\beta))} \qquad (3.12)$$

$$M_p = \frac{1}{1+\frac{\beta}{\beta-\alpha}\left[\exp(-x(\beta-\alpha))\right]_0^d} = \frac{(\beta-\alpha)\exp(d(\beta-\alpha))}{\beta-\alpha\exp(d(\beta-\alpha))} \qquad (3.13)$$

These equations will be utilized in this chapter as well as in succeeding chapters to quantify APD gain. The two assumptions that must hold true for 3.12 and 3.13 to be accurate are the injection of only one type of carrier into a multiplication layer and an electric field within the layer that is relatively constant.

3.3 Measuring Ionization Coefficients

The gain equations derived in Section 3.2 are highly dependent on the ionization coefficients α and β. As stated above, there is not a good model for calculating these coefficients and they must be experimentally measured. To do so, the gain equations will be utilized along with specially designed diode structures. These diodes must meet the following criteria in order to provide accurate coefficients based on the gain equations. First pure electron and hole injection must be allowed for in the diode. This is often accomplished by illuminating a diode from either the p or n doped side. This is illustrated in Figure 8 for the case of a PIN structure, where for electron injection high energy light is illuminated from the P+ side. This creates electron-hole pairs very close to the surface and electrons are injected into the undoped region while holes are swept to the anode. To create hole injection, high-energy light is incident on

the surface of the N+ side of the diode. Electron-hole pairs are created, with the holes injected into the undoped region and the electrons being swept out to the cathode. Using very high energy light with a short optical absorption length, close to pure carrier injection can be created. The total photocurrent created in the diode must also be known very accurately in order to determine a number for the avalanche gain.

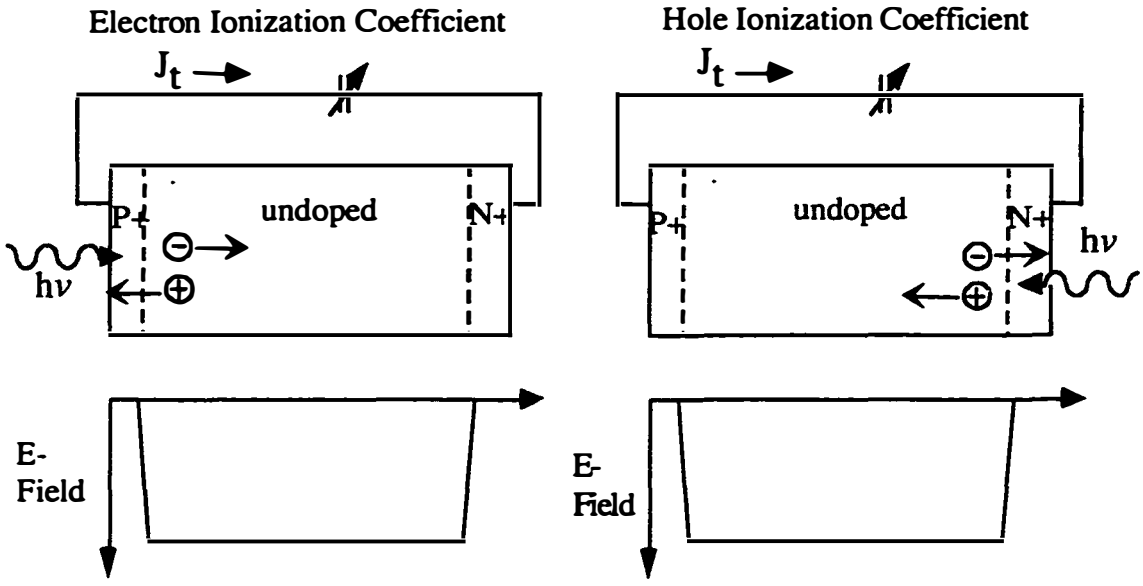


Figure 3.2 - Typical schemes used for measuring the electron and hole ionization coefficients. A semiconductor structure with an undoped region between very thin P+ and N+ layer is used. To measure electron ionization, light shines from the P+ side,

injecting electrons into the undoped layer. Hole ionization is measured by shining light from the N+ side, injecting holes into the undoped layer.

For the PIN structures in Figure 3.2, another criteria that must be met is an almost constant electric field profile in the undoped region. For this field to be constant, there must be minimal residual doping in this region. Without any doping in this region, when the diode is reverse biased, this part of the detector is quickly depleted through and the electric field remains constant versus position as additional voltage is applied. As epitaxial growth techniques continue to improve, PIN structures come closer and closer to this ideal. Silicon PINs can now be grown with undoped layers with greater than 10,000 Ω-cm resistivities. In addition to having a very low doped layer, the thickness of the layer must also be very accurately known so that the electric field in the device can be determined. Ionization coefficients are direct functions of electric field and without precise knowledge of the field, values for α and β are of little value. Figure 3.2, illustrates the type of diode necessary for these measurements as well as the preferred electric field profile under reverse bias.

For the case of electron injection, the electron gain M_n can be calculated simply by dividing the total current measured for a given electric field J_t by the photocurrent created by incident light. M_n versus E (electric field) can then be plotted for the semiconductor. The case for hole injection is much the same with M_h equal to the total current J_t at a given field divided by the incident

photocurrent. Once the quantities M_n and M_h are known, the ionization coefficients can be computed using the following relations derived from the gain equations. From equations 3.12 and 3.13, M_n, M_h, α, and β can be related by

$$M_n = M_p \exp(d(\alpha - \beta)). \qquad (3.14)$$

Using this relation and substituting into 3.12, the following can be determined

$$M_n = \frac{(\alpha - \beta)\frac{M_n}{M_p}}{\alpha - \beta\frac{M_n}{M_p}}. \qquad (3.15)$$

Further substitutions yield

$$\alpha = \frac{1}{d}\ln\left(\frac{M_n}{M_p}\right)\left[\frac{M_n - 1}{M_n - M_p}\right]. \qquad (3.16)$$

Solving for β gives

$$\beta = \frac{1}{d}\ln\left(\frac{M_n}{M_p}\right)\left[\frac{M_p - 1}{M_n - M_p}\right] = \left(\frac{M_p - 1}{M_n - 1}\right)\alpha. \qquad (3.17)$$

Much work has been devoted to the accurate measurement of ionization coefficients for various semiconductors. Although this is often a difficult task, precise knowledge of coefficients gives the APD designer a powerful tool. Semiconductor layer thicknesses and dopings can be designed around required gains and voltages, and as will be shown in Chapter 5 even bandwidths and noise figures can be predicted by these coefficients.

3.4 Gain Versus Electric Field

The gain equations from Section 3.2 can be utilized along with previously measured ionization coefficients[5,6] to create plots of gain versus electric field. Figure 3.3 demonstrates what these curves look like for silicon and indium phosphide for various multiplication layer thicknesses. Multiplication layers refer to the constant high electric field regions created by placing an undoped semiconductor layer between n and p doped layers. These curves are specific to silicon and InP but do have characteristics that are demonstrated by all semiconductors. First of all, avalanche gain increases monotonically with electric field strength. The rate of rise of the avalanche gain also increases with e-field. Thicker multiplication layers require lower electric fields to achieve the same gains than do thinner layers. This is easily explained by considering the minimum ionization energy E_{min}. Thicker layers allow for more carrier travel distance to achieve this energy so a lower net field is required. Figure 3.2 also illustrates how two different semiconductors can have different gain characteristics. Compared to silicon, InP requires a higher electric field to achieve equal gains for given multiplication layers and the gain curves are much "sharper". These sharper curves show up because the gain in InP is more electric field sensitive than that in silicon - something explained in greater depth in Chapter 4.

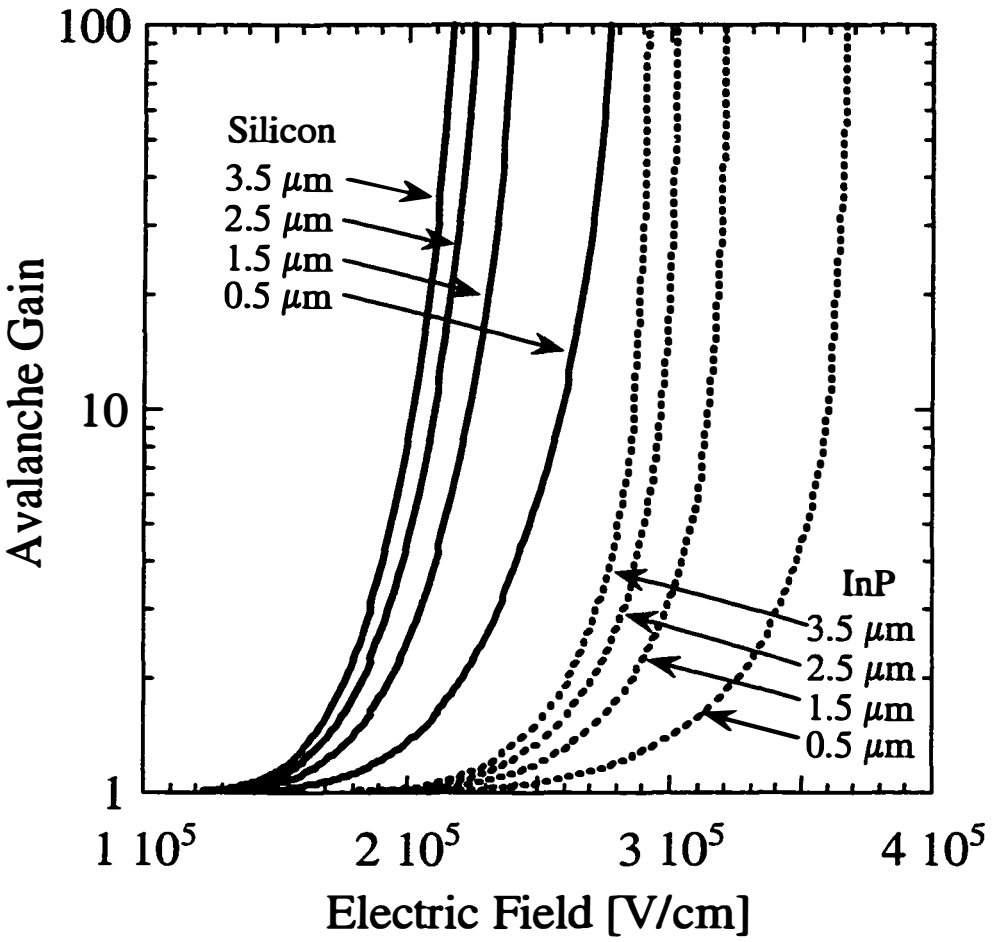


Figure 3.3 - Avalanche gain versus electric field strength for multiplication layers made from silicon and InP. For both types of semiconductors, four different multiplication layer thicknesses are plotted.

3.5 Conclusions

This chapter reviewed the classical derivation of the gain mechanism for impact ionization. Equations were given for gain as a function of ionization

coefficients and multiplication layer thickness. The ionization coefficients are highly dependent on electric field and an accurate determination of these coefficients must be done experimentally. A method for their measurement was also described which is based on the gain produced by photocurrents created on either side of a PIN structure. Based on previously published ionization coefficients, a plot of avalanche gain versus electric field was made for InP and silicon. In the chapters that follow, further aspects of the gain mechanism will be explored - with derivations starting at the gain equation from this chapter. The same ionization coefficients used here will also be utilized to quantify gain sensitivity, signal noise, and frequency bandwidth.

References

[1] Semiconductors and Semimetals, Volume 22 - Lightwave Communications Technology, edited by W. T. Tsang, "Physics of Avalanche Photodiodes," by F. Capasso, pp. 1-172, (1985), Academic Press, NY.

[2] R. J. McIntyre, "A New Look at Impact Ionization - Part I: A Theory of Gain, Noise, Breakdown Probability, and Frequency Response," to be published, (1998).

[3] P. Yuan, K. A. Anselm, C. Hu, H. Nie, C. Lenox, B. G. Streetman, J. C. Campbell, and R. J. McIntyre, "A New Look at Impact Ionization - Part II: Gain and Noise in Short Avalanche Photodiodes," to be published, (1998).

[4] P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Review*, **35**, pp 235-278, (1974).

[5] Quick Reference Manual for Silicon Integrated Circuit Technology, edited by W. E. Beadle, J. C. C. Tsai, and R. D. Plummer, (1985), John Wiley & Sons, Inc., NY.

[6] C. A. Armiento and S. H. Groves, "Impact ionization in (100)-, (110)-, and (111)-oriented InP avalanche photodiodes", *Applied Physics Letters*, **43**, No. 2, pp. 198-200, (1983).

Chapter 4

Voltage and Temperature Sensitivity

4.0 Introduction

Two of the most basic and important characteristics of any avalanche photodetector are its voltage and temperature sensitivity as a function of gain. The significance of voltage sensitivity becomes very evident the first time one looks at a plot of current versus bias for an APD. As shown in Chapter 3, the avalanche gain in a semiconductor is dependent on the electric field present. In the real world, varying the electric field, and in turn the avalanche gain, is done by applying different voltages across a given thickness of a semiconductor. At low voltages, the electric fields are low enough that no avalanche gain is produced. Photocurrent produced by incident light on a semiconductor with a low electric field is said to have a gain of one. When more voltage is applied, increasing the electric field, "avalanching" begins to occur. Photocurrents are magnified by the amount of gain provided for by the increased voltage. A photocurrent versus voltage curve has some characteristics that are the same for all APDs. There is initially a range of voltages starting from zero volts where the curve looks flat because the gain is essentially equal to one. As voltages increase, the photocurrent curve begins to "roll over" and increase

monotonically with increasing voltage. The rate of increase becomes larger and larger as bias increases. The exact nature of this curve is different for every semiconductor. Some materials have extremely sharp gain versus voltage curves where the gain seems to go from one to nearly infinity in a very small voltage span. Other semiconductors have much more gradual curves and the transition to very high gains occurs in a large voltage span. In building solid state devices, the latter characteristic is usually preferred. A larger voltage span translates into greater control of the gain in a device and makes operating at very high gains conceivable. Given the option of a detector that required a constant voltage of 60 V +/- .01 V and a detector that required 60 V +/- 1 V to achieve the same gain, very few circuit designers would accept the challenge of the more sensitive device.

The understanding of the nature of the gain versus voltage curve for a semiconductor is vital in the design of an APD. One would like to be able to understand and predict how voltage sensitive a given material will be when used in a detector. Any APD can achieve a gain as large as 100, but only a voltage insensitive detector could maintain this gain over long periods of time. Voltage sensitivity will in effect help to determine the maximum gains the device can be used at. In this chapter, a derivation will be made that will show why different semiconductors are more voltage sensitive than others. The derivation begins with the APD gain equations explained in Chapter 3 and

yields a comparison of different materials based on k, the ratio of hole ionization to electron ionization coefficients. Comparisons are made for silicon and indium phosphide which represent the most common semiconductors used for APD avalanche gain.

Another important characteristic that determines how well APDs operate in real world environments is the temperature sensitivity as a function of gain. In all semiconductors, the ionization process is temperature dependent. In general, for the same electric field a higher temperature produces less avalanche gain and higher gains are much more sensitive to temperature shifts. Again, just how the gain changes versus temperature is different for each semiconductor. Some have gains that change very little as the temperature varies while others have wild gain swings as temperature varies. It is easy to realize how this can effect where an APD can be deployed. Even for the most temperature insensitive APDs, some type of temperature control is usually needed. A fiber optic repeater station in the middle of the desert with 40°C temperature swings could kill any highly biased detector. There is a big difference, however, in maintaining a +/- 1°C temperature variation that a robust detector might need and a +/- .01°C variation that a more delicate detector requires.

A derivation for temperature sensitivity, similar to that made for voltage sensitivity will also be made in this chapter. The starting point will again be the

APD gain equations from Chapter 3. The ionization coefficient ratio k will play an important role in comparing the differences between semiconductors and which will be more temperature sensitive. Silicon and indium phosphide will be specifically examined because of their prominent roles in the APD world.

4.1 Experimental and Theoretical Gain versus Voltage Curves

The best way to illustrate the differences between the gain curves of two semiconductors is simply to plot both on the same graph. Experimental gain versus voltage curves are shown in Figure 4.1 for silicon[1] and InP[2]. These curves are for a silicon device with a 30 μm absorption layer and a 2 μm multiplication layer. The InP detector has a 3 μm absorption layer and a 2 μm multiplication layer. Both curves show the expected profile for gain versus voltage, a flat region with gain equal to one at low electric fields, and a "rolling over" region where the gain increases faster and faster as voltage increases. The first temptation is to look at these curves and conclude that the silicon diode has vastly superior gain sensitivity characteristics. As we can see, the increase from 10 to 20 in the InP diode occurs in about a 1 volt span, but for the silicon this change occurs in a 45 volt span. This is not a fair comparison, however, because the silicon diode has a different structure than that of the InP detector. In order to achieve high quantum efficiencies, the silicon diode requires a much thicker absorption region (30 μm versus 3 μm for the

InP/InGaAs diode). The voltage across the entire device includes the voltage across the absorption region as well as the multiplication region. In the case of the silicon diode, most of the voltage is dropped across the absorption region

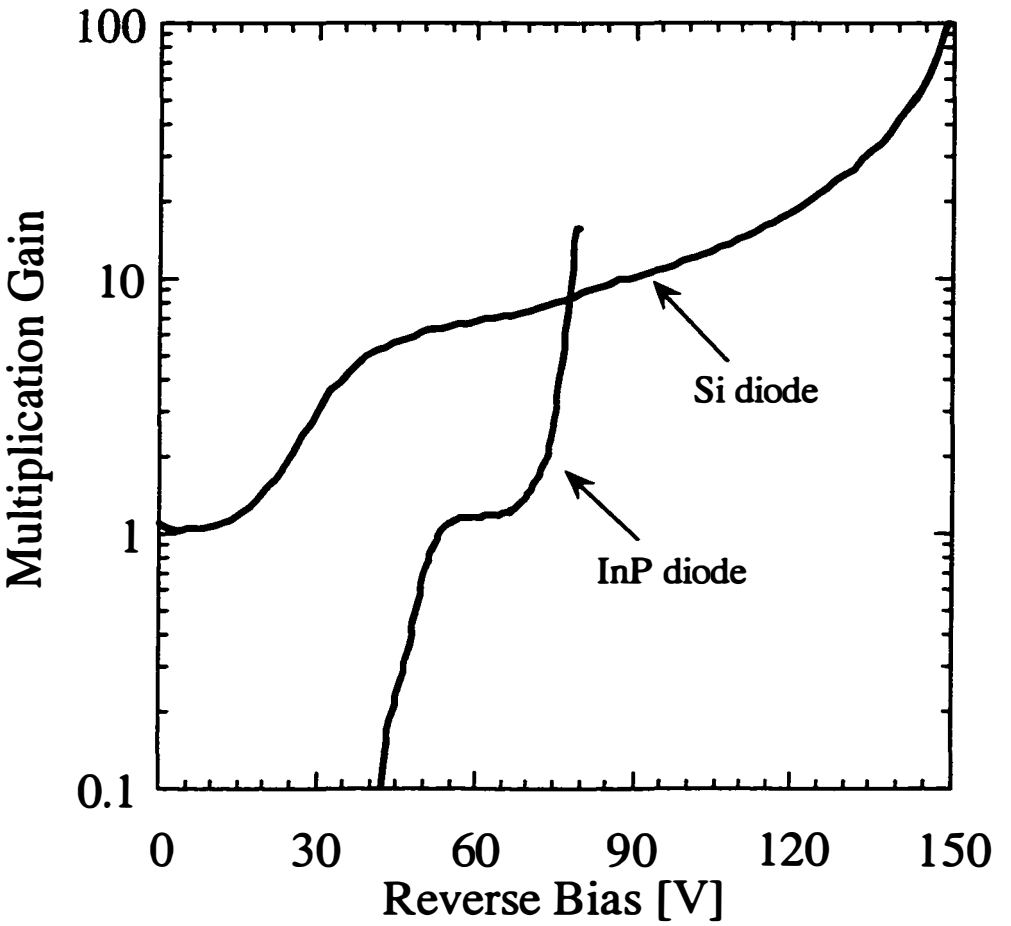


Figure 4.1 - Measured avalanche multiplication gain versus reverse bias curves for silicon and InP diodes. The silicon diode has a 30 μm absorption layer and a 2.0 μm multiplication layer. The InP diode has a 3.0 μm absorption layer and a 2.0 μm multiplication layer.

and does not contribute to the avalanche gain mechanism. To do a true comparison of the voltage sensitivity would require a silicon and InP APD of the same layer thicknesses. This of course is difficult to find. A silicon APD with a thin absorption layer is not very useful due to its low quantum efficiency and so is rarely fabricated.

The lack of comparable real silicon and InP APDs is motivation for a theoretical model of their gain versus voltage characteristics. This of course is done easily enough using the APD gain equations from Chapter 3. By using a given multiplication layer thickness we can easily convert an applied bias to an electric field and then plot gain versus bias for that thickness. This is done for silicon and InP in Figure 4.2. Measured ionization coefficients from previous literature[3,4] were used for the plots. The accuracy of the plots are dependent on the accuracy of the ionization coefficients, but the coefficients used are reasonably similar to those reported by other researchers. The gain versus bias plots illustrate many things about the nature of the gain curve as well as provide a fair comparison of voltage sensitivity for InP and Si - voltage sensitivity only across the multiplication region.

The first thing to notice about the curves in Figure 4.2, are the differences in the voltages needed to achieve comparable gains for multiplication layers of different widths. Thinner layers require less voltage to achieve equal gains although the required electric fields are higher. The second

obvious feature of the graph is how the gain curves are more "gradual" or less voltage sensitive for thicker layers. One reason for this is that the thicker regions require more voltage to cause an equal change in electric field. A second effect is simply that thinner layers operating under higher electric fields

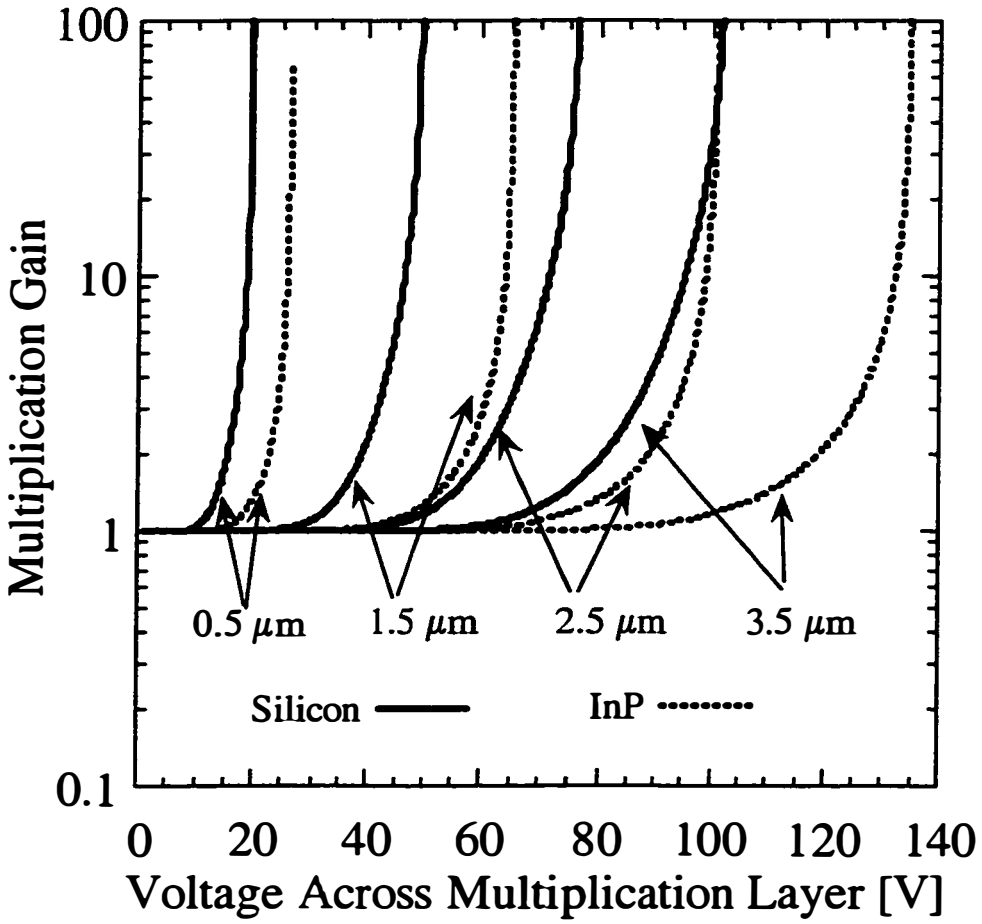


Figure 4.2 - Avalanche multiplication gain versus applied voltage across silicon and InP multiplication layers. Four multiplication layer widths are shown for the two semiconductors.

are more sensitive to equivalent changes in the electric field. In addition to the differences in multiplication layer thicknesses, it is important to notice the differences in the curves for silicon and InP. At the same thicknesses, InP diodes require about 30% more voltage to achieve the same gain as silicon devices. It is usually desirable when designing systems, to use as low a voltage as possible on individual devices. In this category then, silicon is the clear winner over InP. Another distinct advantage to the silicon multiplication layer is also clearly illustrated - the voltage sensitivity. For equal multiplication layer thicknesses there is a more gradual gain curve for silicon devices, especially at gains over 10. For example, a 2.5 μm silicon layer requires over 5 volts to move from a gain of 10 to 100, whereas a 2.5 μm InP layer needs only about 1 volts to make the same jump.

4.2 A Closer Look at Gain Versus Voltage

Figure 4.2 makes clear what was suspected from the experimental APD curves regarding voltage sensitivity. The figure also inspires new questions. Silicon is less voltage sensitive than InP, but by how much? What is the nature of the gain versus voltage curve? Does gain increase as voltage squared, an exponent, or maybe some other power? Is there a general parameter that determines which semiconductors have better gain-voltage characteristics. All

of these questions can be answered with some mathematics, which draw out the fruits of the original APD gain equations[5].

A voltage sensitivity derivation begins by assuming a simple model for an avalanche photodetector with separate absorption and multiplication (SAM) layers. In such a model any photogenerated electron-hole pairs are produced in the absorption region. One of the carrier types, electrons or holes, is then injected into the multiplication region where impact ionization takes place. Also assumed is that the electric field is constant in the region, a good approximation for minimally doped semiconductors. The electric field across this multiplication region is then simply V/w, where V is the applied voltage. The multiplication gain M for injected carriers is given by the gain equation from Chapter 3:

$$M = \frac{(\alpha - \beta)e^{w(\alpha-\beta)}}{\alpha - \beta e^{w(\alpha-\beta)}} \qquad (4.1)$$

This gain equation is valid for electrons injected in a multiplication region. The gain for injected holes is obtained simply by exchanging α's and β's. To really understand the dependence of gain on voltage first requires differentiating to obtain dM/dV, the rate of change of voltage versus gain. This will reveal just how steep the curve is and how it changes with voltage.

$$\frac{dM}{dV} = M^2\left[\frac{\alpha\frac{\partial\beta}{\partial V} - \beta\frac{\partial\alpha}{\partial V}}{(\alpha-\beta)^2} + \frac{e^{-w(\alpha-\beta)}}{(\alpha-\beta)^2}\left(\beta\frac{\partial\alpha}{\partial V} - \alpha\frac{\partial\beta}{\partial V}\right) + \frac{e^{-w(\alpha-\beta)}w\alpha}{(\alpha-\beta)}\left(\frac{\partial\alpha}{\partial V} - \frac{\partial\beta}{\partial V}\right)\right]$$

(4.2)

This expression is exact and completely general. Unfortunately it does not reveal anything obvious about the dM/dV relation. Making the equation more workable requires the following approximation for high gains, the accuracy of which increases with increasing M,

$$\alpha - \beta e^{w(\alpha-\beta)} \to 0. \qquad (4.3)$$

Using this approximation and k = ß/ α , it can be shown that

$$\frac{dM}{dV} \approx \frac{M^2}{\left(\frac{1}{k}-1\right)}\left[\frac{\partial\alpha}{\partial V}\left(w-\frac{1}{\alpha}\right) - \frac{\partial\beta}{\partial V}\left(w-\frac{1}{\beta}\right)\right]. \qquad (4.4)$$

Equation 4.4 is valid for electron initiated gain. For hole initiated gain, the equivalent equation is

$$\frac{dM}{dV} \approx \frac{M^2}{(k-1)}\left[\frac{\partial\beta}{\partial V}\left(w-\frac{1}{\beta}\right) - \frac{\partial\alpha}{\partial V}\left(w-\frac{1}{\alpha}\right)\right]. \qquad (4.5).$$

These equations can be further simplified by recognizing that the ionization coefficients α and ß are usually described by the empirical formulas $\alpha = a_1 e^{-a_2/E}$ and $\beta = b_1 e^{-b_2/E}$, where E is the electric field in the

multiplication region. Using these relations as well as Equation 4.3, results in the following for electron initiated gain:

$$\frac{dM}{dV} \approx \frac{M^2}{V^2}\left(\frac{-k\ln k}{(k-1)^2}\right)\left[(a_2 - b_2 k)w - \frac{1}{\alpha}(a_2 - b_2)\right]. \qquad (4.6)$$

A further approximation can be made for electron initiated gain where w >> $1/\alpha$ and hole initiated gain equation where w >> 1/ß. These relations lead to the following:

$$\frac{dM}{dV} \approx \frac{M^2}{V^2} w\left(\frac{-k\ln k}{(k-1)^2}\right)(a_2 - b_2 k), \qquad (4.7)$$

$$\frac{dM}{dV} \approx \frac{M^2}{V^2} w\left(\frac{k\ln k}{(k-1)^2}\right)\left(b_2 - \frac{a_2}{k}\right), \qquad (4.8)$$

for electron and hole initiated gain, respectively. Due to their simplicity these equations are useful in explaining the gain sensitivity dependence for APDs. To examine how accurate they are after all the approximations that were made, Figures 4.3 and 4.4 show a comparison of dM/dV for an exact solution and then the approximations (Equations 4.4, 4.5, 4.7 and 4.8). Plots are shown for InP and silicon at two different thicknesses. Ionization coefficient parameters were taken from existing literature as in Figure 4.2. As can be seen from the graphs, Equations 4.4 and 4.5 agree very well with an exact solution. Equations 4.7 and 4.8 are less accurate but are still good approximations.

The simplicity of equations 4.7 and 4.8 reveal first that dM/dV increases as gain squared, not surprising after viewing the fast rise of a typical gain curve.

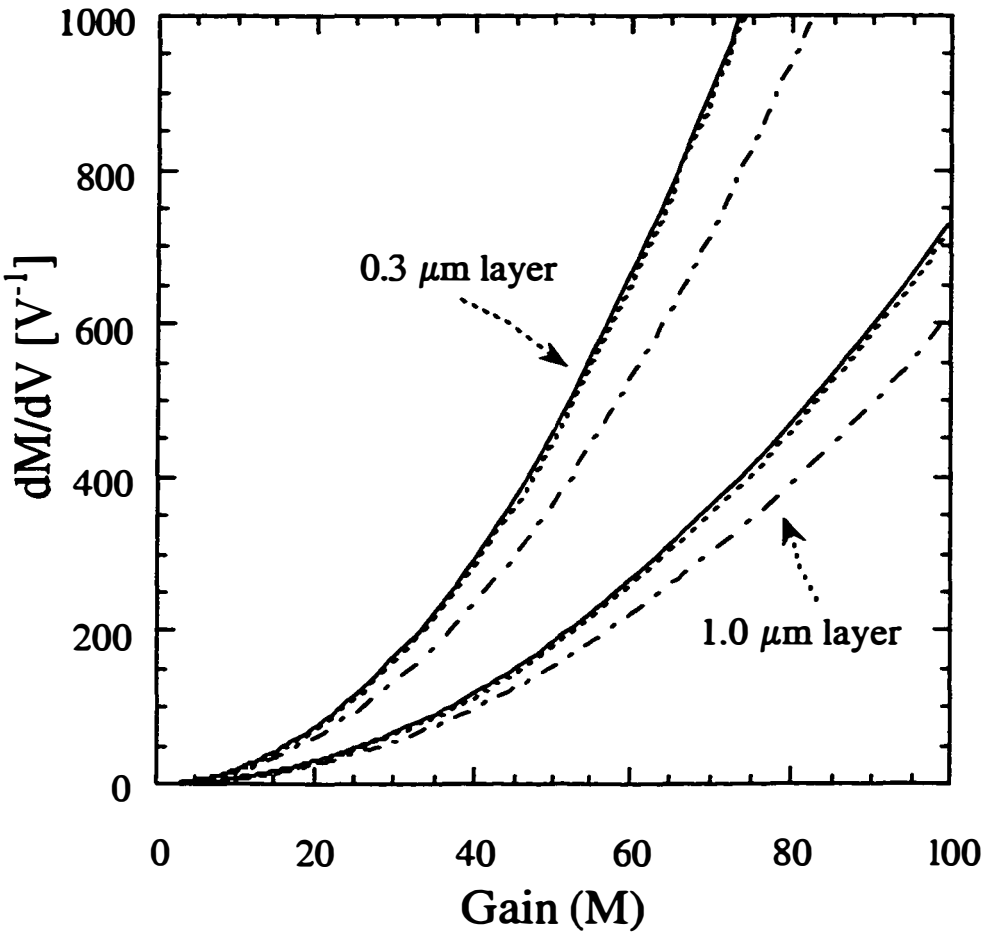


Figure 4.3 - dM/dV as a function of gain for a silicon avalanche multiplication layer. Plots are shown for two different layer thicknesses. The solid line represents the exact solution, the dotted line is the approximation given by Equation 4.5, and the dotted and dashed line is the approximation given by Equation 4.8

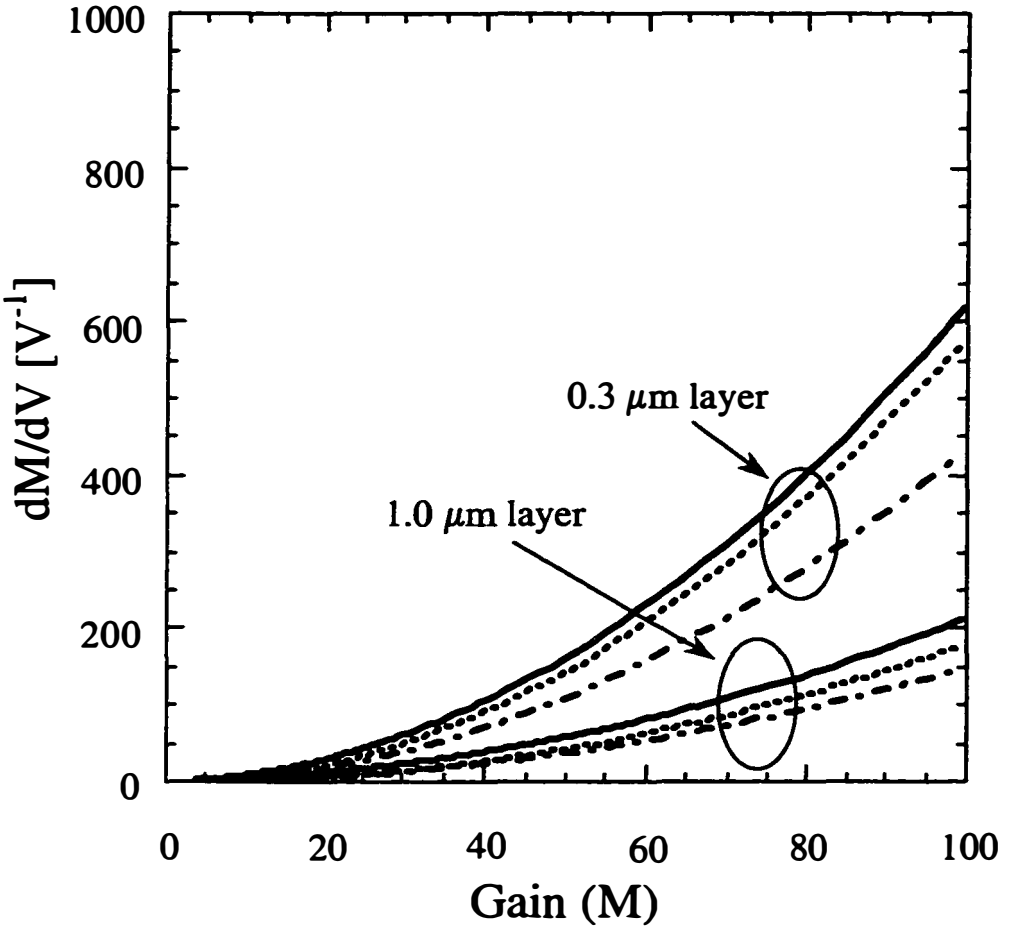


Figure 4.4 - dM/dV as a function of gain for an InP avalanche multiplication layer. Curves for two multiplication layer thicknesses are shown. The solid line represents the exact solution, the dotted line is the approximation given by Equation 4.4, and the dotted and dashed line is the approximation given by Equation 4.7.

These expressions also expose the dependence of dM/dV on the material parameter k. This combined with the following facts about the constants in equations 4.7 and 4.8, allows one to make some sweeping conclusions about gain sensitivity. The constants a_2 and b_2 are both on the order of 10^6 V/cm for many materials, in fact the quantity $a_2 - b_2k$ or $b_2 - a_2/k$ is ~$1x10^6$ V/cm at the electric fields most APDs operate. Table I shows experimental values of k, a_2, b_2 and either $a_2 - b_2k$ or $b_2 - a_2/k$ (depending on whether $k>1$ or $k<1$) for several materials measured between a range of electric field values. The value of k is for an electric field in the middle of the measured range.

TABLE I. k, a_2,b_2 and either $a_2 - b_2k$ or $b_2 - a_2/k$ for materials listed.

Material	k	a_2	b_2	$a_2 - b_2k$ or $b_2 - a_2/k$
Si[6]	.002	$1.24x10^6$	$2.71x10^6$	$1.2x10^6$
Ge[7]	1.6	$1.4x10^6$	$1.27x10^6$	$.43x10^6$
InP[4]	1.9	$3.10x10^6$	$2.56x10^6$	$.93x10^6$
$In_{0.53}Ga_{0.47}As$[8]	.42	$1.86x10^6$	$2.06x10^6$	$1.0x10^6$
$In_{0.14}Ga_{0.86}As$[9]	2.9	$3.6x10^6$	$2.7x10^6$	$1.5x10^6$
$In_{0.89}Ga_{0.11}As_{0.74}P_{0.26}$[10]	.13	$3.2x10^6$	$3.07x10^6$	$2.8x10^6$
GaAs/AlGaAs Superlattice[11]	.14	$1.7x10^6$	$2.17x10^6$	$1.4x10^6$

Figure 4.5 shows a plot of $dM/dV(V^2/M^2)$ versus k for various thicknesses at a gain of 100 assuming the $1x10^6$ V/cm value for the constants in the expression outlined in Table 1. As the graph shows, as k approaches 1, the

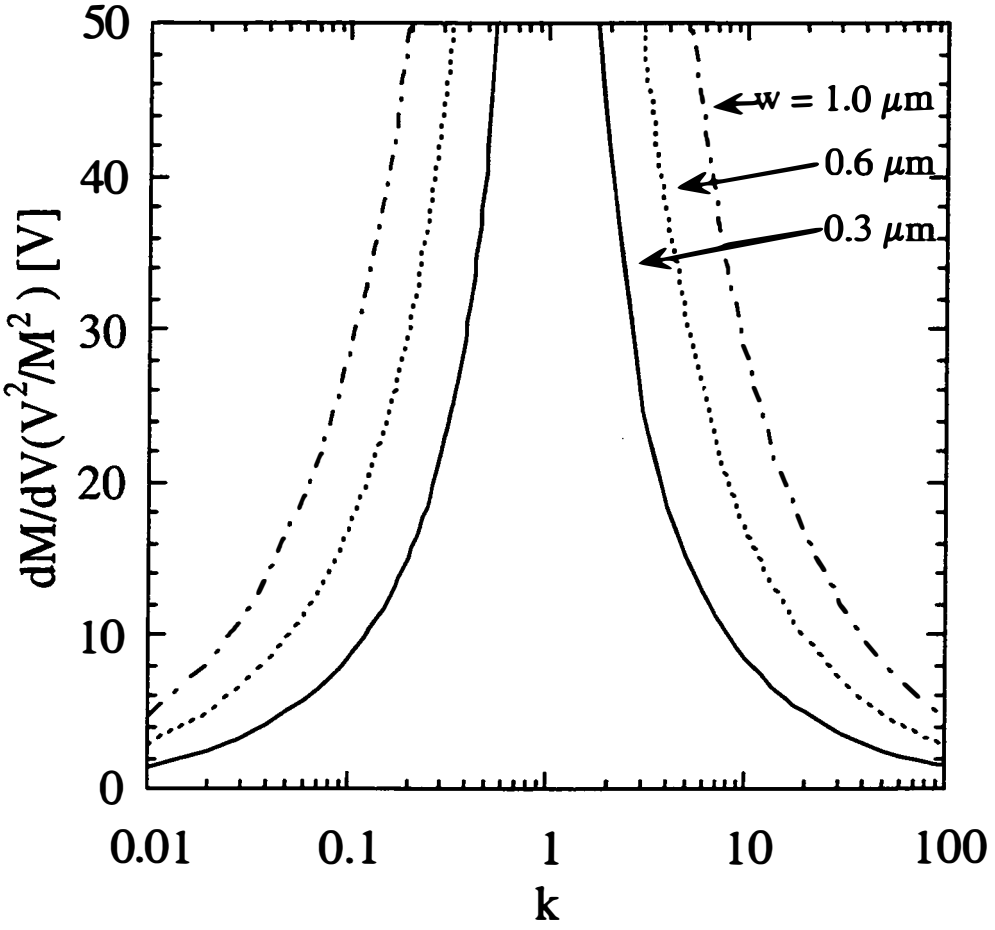


Figure 4.5 - $dM/dVx(V^2/M^2)$ versus k plotted for three different thicknesses calculated for a gain of 100. A value of $1x10^6$ V/cm is assumed for the quantity $a_2 - b_2k$ or $b_2 - a_2k$ from Equations 4.7 and 4.8.

gain sensitivity increases rapidly. This explains why gain versus voltage curves in silicon have a more gradual curvature than those for InP, as illustrated in Figure 4.2, given that the k value for InP is much closer to 1 than the k value for silicon. If the above assumption for the constant values a_2 and b_2 are valid for any material. Equations 4.7 and 4.8 can also be used to determine k based on a plot of $dM/dV(V^2/M^2)$.

4.3 Temperature Sensitivity Derivation

We now turn our attention towards an in depth look at the temperature sensitivity of the gain in an avalanche photodetector. As was stated earlier, changes in temperature cause changes in the impact ionizations of carriers in a semiconductor. Just how much these ionization events are effected will determine the temperature sensitivity of a device. Temperature effects are not only important in extreme environments where optical systems are deployed. Even devices operating in a "room temperature" laboratory environment can experience temperature changes due to self heating when large amounts of electrical current flow through them. This will show up as a drift in avalanche gain over time as higher temperatures produce states with less gain. It is obviously important then to have active temperature control of an operating APD and very desirable to have a detector with the lowest temperature sensitivity possible.

The treatment of temperature sensitivity will begin with the gain equations for APDs from Chapter 3. Embedded within these equations are the ionization coefficients that provide specific parameters for individual semiconductors. To obtain a temperature dependence these ionization coefficients must include some temperature information. The most widely studied semiconductor for APD applications is silicon and a large amount of data is available on its ionization coefficients. Included in this data is how the ionization coefficients vary with temperature. Conradi[12] determined that the coefficients could be written as $\alpha = a_1 e^{\frac{-(a_2 + a_3 T)}{E}}$ and $\beta = b_1 e^{\frac{-(b_2 + b_3 T)}{E}}$. This reflects the fact that the ionization coefficients become smaller at higher temperatures. Figure 4.6 illustrates how the gain changes with temperature based on the Conradi coefficients. As expected higher temperatures shift the gain curve, in effect requiring more applied voltage to produce the same gain. This is due to the decrease in the ionization coefficients. Another aspect of the graph to notice is that higher gains are more sensitive to temperature changes. For example, an APD operating at 23°C at a gain of 100 will see its gain decrease to 20 when the temperature is raised to 50°C. The same APD operating at 23°C and a gain of 10 will only see a gain decrease to 7 when the temperature is raised to 50°C.

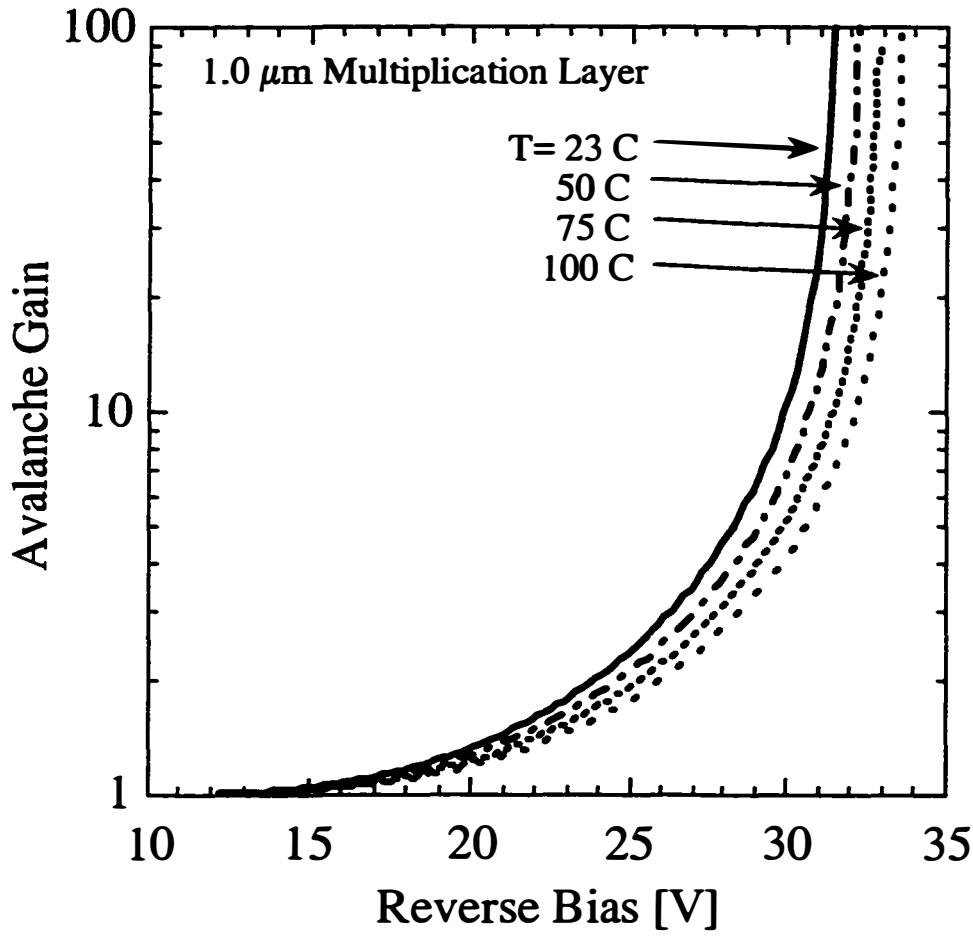


Figure 4.6 - Avalanche multiplication gain versus reverse bias for a silicon diode. Curves are plotted for four different operational temperatures and were generated using the Condradi coefficients.

Now we will develop a formalism for temperature sensitivity similar to that for voltage sensitivity. We will base the derivation on Conradi's temperature dependence. Again we assume a simple model for an avalanche

photodetector with separate absorption and multiplication (SAM) layers. Photogenerated electron-hole pairs are produced in the absorption region, and one type of carrier is then injected into the multiplication region where impact ionization takes place. Beginning with equation 4.1 for gain, the change of gain with temperature dM/dT can be written as the following:

$$\frac{dM}{dT} = M^2 \left[\frac{\alpha \frac{\partial \beta}{\partial T} - \beta \frac{\partial \alpha}{\partial T}}{(\alpha - \beta)^2} + \frac{e^{-w(\alpha-\beta)}}{(\alpha-\beta)^2}\left(\beta \frac{\partial \alpha}{\partial T} - \alpha \frac{\partial \beta}{\partial T}\right) + \frac{e^{-w(\alpha-\beta)} w\alpha}{(\alpha-\beta)}\left(\frac{\partial \alpha}{\partial T} - \frac{\partial \beta}{\partial T}\right) \right]$$

(4.9)

This expression is exact and completely general. Unfortunately it reveals little about the dM/dT relation. To make the equation more workable, we can use the approximation for high gains found in equation 4.3. Using this and the relation $k = \beta/\alpha$, it can be shown that

$$\frac{dM}{dT} \approx \frac{M^2}{\left(\frac{1}{k} - 1\right)} \left[\frac{\partial \alpha}{\partial T}\left(w - \frac{1}{\alpha}\right) - \frac{\partial \beta}{\partial T}\left(w - \frac{1}{\beta}\right) \right]. \qquad (4.10).$$

Equation 4.10 is valid for electron initiated gain. For hole initiated gain the equivalent expression is the following

$$\frac{dM}{dT} \approx \frac{M^2}{(k-1)} \left[\frac{\partial \beta}{\partial T}\left(w - \frac{1}{\beta}\right) - \frac{\partial \alpha}{\partial T}\left(w - \frac{1}{\alpha}\right) \right]. \qquad (4.11).$$

To simplify these equations further, we use the empirical formulas for ionization coefficients $\alpha = a_1 e^{\frac{-(a_2 + a_3 T)}{E}}$ and $\beta = b_1 e^{\frac{-(b_2 + b_3 T)}{E}}$, where E is the

electric field in the multiplication region and T is the temperature of the device. Using these relations as well as Equation 4.3 yields the following for electron initiated gain

$$\frac{dM}{dT} \approx \frac{M^2}{E}\left(\frac{-k\alpha}{(k-1)}\right)\left[(b_3 k - a_3)w + \frac{1}{\alpha}(a_3 - b_3)\right] \qquad (4.12).$$

We can make a further approximation for electron initiated gain where w >> $1/\alpha$ and hole initiated gain where w >> 1/ß. These relations lead to the following

$$\frac{dM}{dT} \approx \frac{M^2}{E}\left(\frac{-k\ln k}{(k-1)^2}\right)(b_3 k - a_3), \qquad (4.13)$$

$$\frac{dM}{dT} \approx \frac{M^2}{E}\left(\frac{k\ln k}{(k-1)^2}\right)\left(\frac{a_3}{k} - b_3\right), \qquad (4.14)$$

for electron and hole initiated gain, respectively. These equations are very revealing in what they teach us about temperature sensitivity but we still must examine how accurate they are. As one might notice, neither 4.13 nor 4.14 are dependent on temperature whereas the exact solution 4.9 is. Any check of accuracy must include different temperatures to ensure that the temperature dependence of dM/dT is minimal. Figure 4.7 plots dM/dT versus gain for the exact solution from equation 4.9 at temperatures of 23°C and 200°C and the approximation from equation 4.13 using a 1.0 μm silicon multiplication layer. Conradi's ionization coefficients are used for both curves. For the approximation, k=1/7 and E= 300kV/cm were used. As Figure 4.7 illustrates,

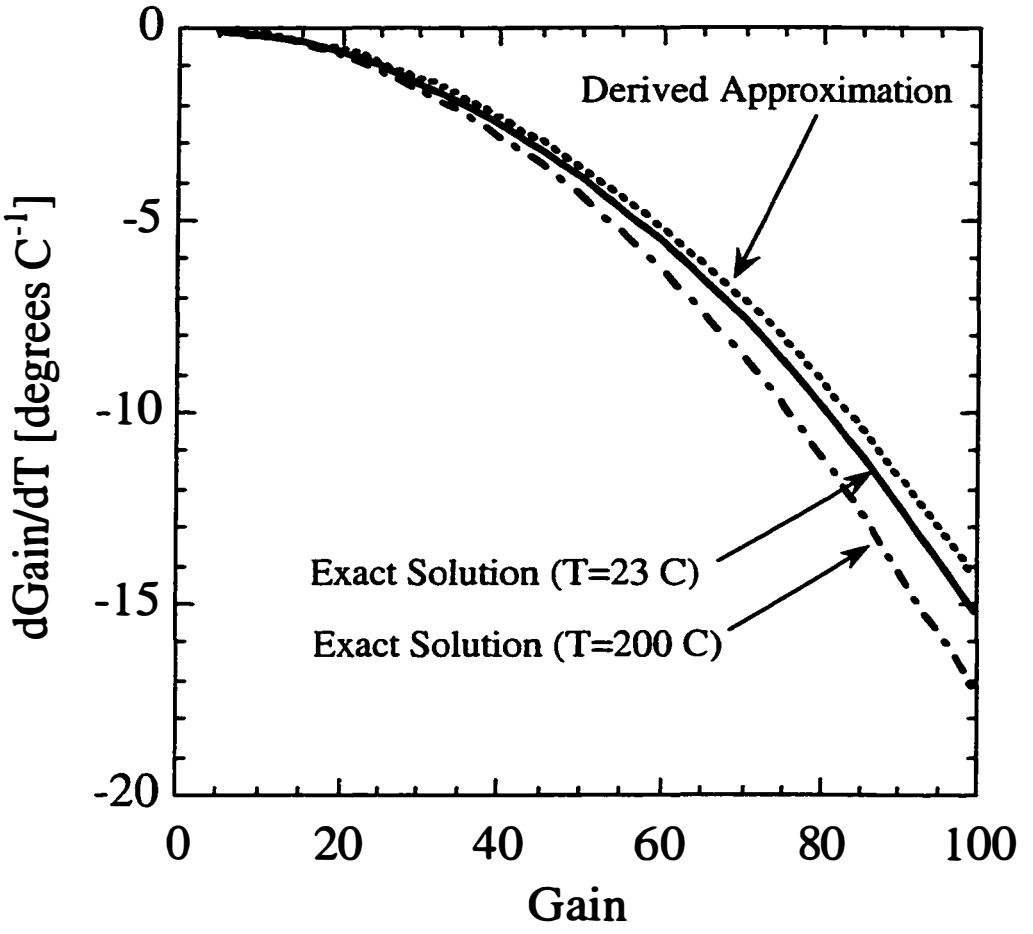


Figure 4.7 - dM/dT versus avalanche multiplication gain for a silicon multiplication layer. The exact solution from Equation 4.9 at temperatures of 23°C and 200°C and are plotted along with the approximation from equation 4.13. In generating the curves, a 1.0 μm silicon multiplication layer was assumed and Conradi's ionization coefficients were used. For the approximation, k=1/7, E= 300kV/cm, b_3=1.1x10^{-3}, and a_3=1.3x10^{-3} were used.

the approximation is remarkably accurate even for a large temperature range. The figure also quantifies what was stated earlier about the sensitivity versus gain - the higher the gain the more temperature sensitive.

Given that approximations 4.13 and 4.14 have been shown to be very accurate in describing temperature sensitivity, we would now like to make some generalizations for all semiconductors based on the parameter k. The key to generalizing 4.13 and 4.14 will be in assuming that the constants $b_3k - a_3$ and $a_3/k - b_3$ will in general be very similar for all semiconductors, to same type of assumption that was done in Section 4.2. This is difficult to prove absolutely, however, because of the lack of temperature dependent data for ionization coefficients of semiconductors other than silicon and InP. For now, it can be shown that for these two materials the constants are similar, allowing for direct comparison. Using the data from Taguchi[13] at several temperatures, we can compute values for a_3 and b_3 for InP of $a_3=1.9\text{x}10^3$ $Vcm^{-1}C^{-1}$ and $b_3=1.8\text{x}10^3$ $Vcm^{-1}C^{-1}$. Conradi's values for silicon are $a_3=1.3\text{x}10^3$ $Vcm^{-1}C^{-1}$ and $b_3= 1.1\text{x}10^3$ $Vcm^{-1}C^{-1}$. This yields a constant of $b_3k - a_3 \approx -1.3\text{x}10^3$ $Vcm^{-1}C^{-1}$ for silicon with electron injection and $a_3 / k - b_3 \approx -0.85\text{x}10^3$ $Vcm^{-1}C^{-1}$ for InP with hole injection. Based on these numbers, we will make the assumption then that in general the constant $b_3k - a_3$ for electron injection and $a_3/k - b_3$ for hole injection is approximately $-1.0\text{x}10^3$ $Vcm^{-1}C^{-1}$. Figure 4.8 shows a plot of $\frac{dM}{dT}\frac{E}{M^2}$ versus

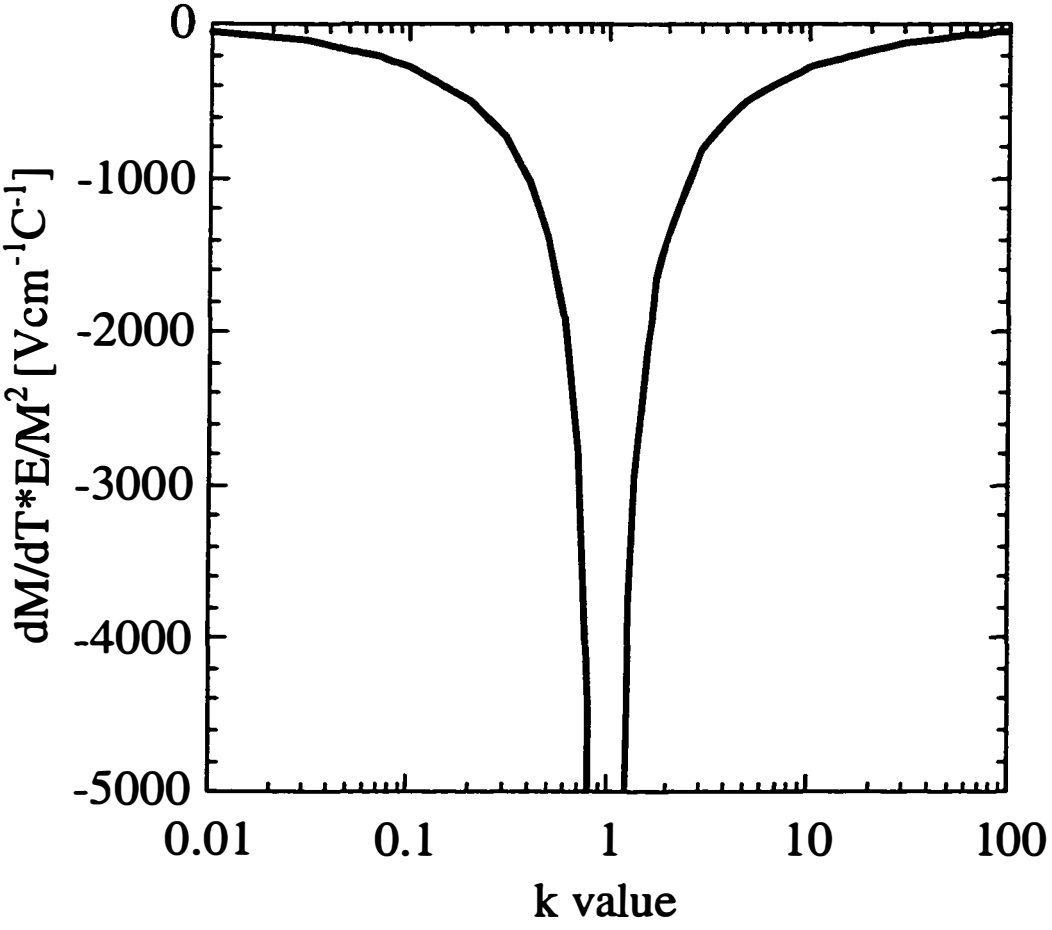


Figure 4.8 - $\frac{dM}{dT}\frac{E}{M^2}$ versus k plotted using Equations 4.13 and 4.14. The assumed value of -1.0×10^{3} $Vcm^{-1}C^{-1}$ was used for the constant $b_3k - a_3$ for electron injection and $a_3/k - b_3$ for hole injection.

k using the assumed value of the constant and equations 4.13 and 4.14. As the plot shows, the closer the k value is to 1, the more temperature sensitive the gain in the device will be. A semiconductor like silicon with a low value for k

will have much less sensitivity to temperature variations than InP for example with a k value of around two. Other semiconductors like GaAs whose k ratio approaches one should be very sensitive to temperature.

4.4 Conclusions

In conclusion, we have derived several important relations to explain APD voltage and temperature sensitivity. Although these equations are based on a simplified APD structure, the trends should be valid for more complex devices. We have shown that dM/dV and dM/dT vary as a function of M^2 . The other quantities in our equations remain relatively constant at significant gains. We have also derived the relation of gain sensitivity to k and can conclude that APDs made from materials with k ratios much larger or smaller than one will result in devices less sensitive to voltage and temperature variations. These conclusions have important ramifications for Silicon/InGaAs APDs. Silicon, whose k ratio is lower than any know semiconductor will be the best possible avalanche multiplication layer in regards to these first important properties. We can expect then a clear advantage of silicon/InGaAs APDs over InP/InGaAs APDs in voltage and temperature sensitivity.

References

[1] Semiconductors and Semimetals, Volume 22 - Lightwave Communications Technology, edited by W. T. Tsang, "Silicon and Germanium Avalanche Photodiodes," by T. Kaneda, pp. 247-328, (1985), Academic Press, NY.

[2] Semiconductors and Semimetals, Volume 22 - Lightwave Communications Technology, edited by W. T. Tsang, "Compound Semiconductor Photodiodes," by T. P. Pearsall and M. A. Pollack, pp. 173-245, (1985), Academic Press, NY.

[3] Quick Reference Manual for Silicon Integrated Circuit Technology, edited by W. E. Beadle, J. C. C. Tsai, and R. D. Plummer, (1985), John Wiley & Sons, Inc., NY.

[4] C. A. Armiento and S. H. Groves, "Impact ionization in (100)-, (110)-, and (111)-oriented InP avalanche photodiodes", *Applied Physics Letters*, **43**, No. 2, pp. 198-200, (1983).

[5] A. R. Hawkins, W. Wu, and J. E. Bowers, "Improvement of Avalanche Photodetectors Through Integration of InGaAs and Si," invited paper, SPIE Optoelectronics '97, Conference Proceedings, 2999, 68-75, San Jose, CA, February 8-14, (1997).

[6] P.P. Webb, "Measurements of Ionization Coefficients in Silicon at Low Electric Fields," GE Canada Inc., Electro Optics Operations, unpublished.

[7] T. Mikawa, S. Kagawa, T. Kaneda, Y. Toyama, and O. Mikami, "Crystal orientation dependence of ionization rates in germanium," *Applied Physics Letters*, **37**, No. 4, pp. 387-389, (1980).

[8] T.P. Pearsall, "Impact ionization rates for electrons and holes in $Ga_{0.47}In_{0.53}As$," *Applied Physics Letters*, **36**, No. 3, pp. 218-220, (1980).

[9] T.P. Pearsall, "Impact ionization rates for electrons and holes in $Ga_{0.14}In_{0.86}As$," *Applied Physics Letters*, **27**, No. 6, pp. 330-332, (1975).

[10] Y. Takanashi and Y. Horikoshi, "Ionization Coefficient of InGaAsP/InP APD," *Japanese Journal of Applied Physics*, **18**, No. 11, pp. 2173-2174, (1979).

[11] F. Capasso, W.T. Tsang, A. L. Hutchinson, and G. F. Williams, "Enhancement of electron impact ionization in a superlattice: A new avalanche photodiode with a large ionization rate ratio," *Applied Physics Letters*, **40**, No. 1, pp. 38-40, (1982).

[12] J. Conradi, "Temperature Effects in Silicon Avalanche Diodes," *Solid-State Electronics*, **17**, pp. 99-106, (1973).

[13] K. Taguchi, T. Torikai, Y. Sugimoto, K. Makita, and H. Ishihara,"Temperature dependence of impact ionization coefficients in InP," *Journal of Applied Physics*, **59**, pp. 476-481, (1986).

Chapter 5

Noise and Bandwidth

5.0 Introduction

This chapter will focus on probably the two factors that give the InGaAs-silicon APD the largest advantage over existing near infrared avalanche photodiodes. In comparing APDs, they cannot be judged simply by comparing the highest achievable gain a detector is capable of. It could be said that all APDs could reach extremely high gains given tight enough bias controls. The more important measure is whether or not APD gain is usable. This is determined by two things, the noise produced by the detector and the speed of operation or frequency bandwidth. These two properties are intimately linked to the ionization process within a semiconductor and can be directly related to the ionization coefficients α and β.

Examining the ionization process as described in Section 1.2, it becomes clear how avalanche multiplication can lead to increased signal noise. Ionization is a very random event for an individual carrier as it travels through a high electric field region. An ionizing collision can occur whenever a carrier has obtained a minimum ionization energy E_{min}, but such a collision does not necessary occur exactly after an energy E_{min} has been achieved. As a result,

different carriers will experience different numbers of collisions and effectively different gains. The random nature of the process is accentuated in semiconductors in which both electrons and holes are likely to ionize. Forward moving electrons and backward moving holes can create ionization events throughout the width of a multiplication layer. This results in a very random process and in effect a noisy one if a signal is being amplified by the APD. The noise added to a signal by the avalanche multiplication process is always greater than the simple shot noise associated with transit across a semiconductor layer. This added or excess noise in an APD is denoted as F and is a function of k, the ratio of the hole and electron ionization coefficients. Excess noise increases monotonically with the amount of gain produced in a detector and the amount of noise created is a very important consideration in communications systems. The signal noise level out of the detector needs to be less than the noise level in the electrical amplifiers that follow the APD. In existing systems utilizing InGaAs/InP APDs, noise factors limit the operation gain on detectors to around 10. Operating at lower gains requires a higher input light level and reduces the device's sensitivity. Gains higher than 10 have noise levels so high that larger optical signals are again required, limiting the device's sensitivity. Section 5.1 will summarize the derived relationship for excess noise versus k, the ionization ratio. A clear advantage will be shown for silicon multiplication

layers. Section 5.2 will show plots for the noise equations and illustrate their dependence on k, gain, and multiplication layer width.

Like the noise in an APD, the limits on operating speed or bandwidth can also be explained by examining the multiplication process described in Section 1.2. The electron entering the multiplication regions can represent an input signal. Two extreme cases are illustrated, one in which only electrons have ionizing collisions and another in which electrons and holes ionize at equal rates. In the first case, all electrons are swept to the right and holes to the left. Due to the fact that holes do not ionize, even holes produced from ionization events on the far edge of the multiplication region travel back across the region without producing ionization events of their own. In effect then the time required for the multiplication process to be completed is the total time for an electron to transit the region plus the time for a hole to transit a region. The second case, with equal ionization coefficients, is much more complicated. As illustrated in Section 1.2, with electrons and holes both ionizing, the initial signal appears to ricochet back and forth within the region as ionized electrons and holes are created on both ends. A first look at this case might lead to the conclusion that the process would never end and the multiplication time would be infinite. In the real world however, solid-state events are rarely eternal and this one too would come to an end. Individual electrons and holes would eventually travel out of the region without ionizing due to the random nature of

that event and multiplication would cease. The time for this multiplication would obviously be much longer than the earlier case however. No semiconductor fits exactly into these two extremes - all lie somewhere in between. Silicon's electrons have much higher ionization probabilities than do its holes, but some holes do ionize. InP's electrons and hole probabilities aren't equal, the holes are more than twice as likely to ionize as the electrons. The larger the difference between hole and electron coefficients, the faster the multiplication time but for these semiconductors it lies between the transit time limit and the very, very long "ricocheting" signal limit. The speed of the process has a direct bearing on the signal frequency for which the detector can be used. As will be shown in later sections, the frequency limit is dependent on many factors including the operation gain, the multiplication layer thickness, and the RC limits of a given detector. These sections will also illustrate the relationship between the frequency limit and k, the ionization coefficient ratio. Device performance for APDs made from different materials can then be compared by plotting their frequency dependence on gain.

5.1 Noise Equations

The following equations for the noise produced in an avalanche multiplication process follow the derivation by McIntyre[1,2] and are accurate under the following assumptions. First, as in Section 3.2, a multiplication

region is assumed to exist which consists of semiconductor layer of distance d under a constant electric field. Electrons representing a signal are injected from the left of the region at x=0 while any injected holes enter at x = d. This situation represents closely what occurs in an APD with separate light absorption and multiplication layers. It is also assumed that only one type of carrier is injected into the region for a given semiconductor. When using silicon, electrons would be injected and for InP holes injected. Given this situation, ϕ, the noise spectral density, can be derived by starting with $d\phi$ the differential noise density. In a differential distance dx, $d\phi$ can be written as

$$d\phi(x) = 2qM^2(x)dI_p(x), \qquad (5.1)$$

where q is the electron charge, M(x) is the gain at a point x, and $I_p(x)$ is the hole current at a point x. Integrating this equation from 0 to d yields

$$\phi = 2q\left[I_n(0)M^2(0) + I_p(d)M^2(d) + \int_0^d \frac{dI_p}{dx} M^2(x)dx \right] \qquad (5.2)$$

where $I_n(0)$ and $I_p(d)$ are integration constants representing injected currents into the multiplication region. Equation 5.2 can be further reduced using integration by parts on the integral, first recognizing the relationship for the differential current (Equation 3.1) and using the following relationship for differential gain:

$$\frac{dM(x)}{dx} = (\alpha - \beta)M(x). \qquad (5.3)$$

Upon substitution and integration it can be shown for injected electrons and k < 1

$$\phi = 2qI_{in}M^3\left[1-(1-k)\left(\frac{M-1}{M}\right)^2\right], \qquad (5.4)$$

or for injected holes and k > 1

$$\phi = 2qI_{in}M^3\left[1+\frac{(1-k)}{k}\left(\frac{M-1}{M}\right)^2\right]. \qquad (5.5)$$

Here I_{in} is the current injected into the multiplication region, M is the multiplication of the injected current, and k is the ionization coefficient ratio β/α. Equations 5.4 and 5.5 can both be written as

$$\phi = 2qI_{in}M^2F \qquad (5.6)$$

where the first part of the term is due to multiplied input shot noise and F is defined as the excess noise due to the avalanche multiplication process. F could also be defined as the ratio of total noise to shot noise in an APD. For electron initiated gain, the excess noise factor F_n is given by

$$F_n = M\left[1-(1-k)\left(\frac{M-1}{M}\right)^2\right] \qquad (5.7)$$

and for hole initiated gain F_p is given by

$$F_p = M\left[1+\frac{(1-k)}{k}\left(\frac{M-1}{M}\right)^2\right] \qquad (5.8)$$

5.2 A Look at the Gain Relations

With the equations derived in 5.1 for excess noise, plots can be made to illustrate the dependence on gain, multiplication layer thickness, and material. Figure 5.1 shows a plot of the excess noise factor F_n for injected electrons into a silicon multiplication layer of various thicknesses. This plot was made by using

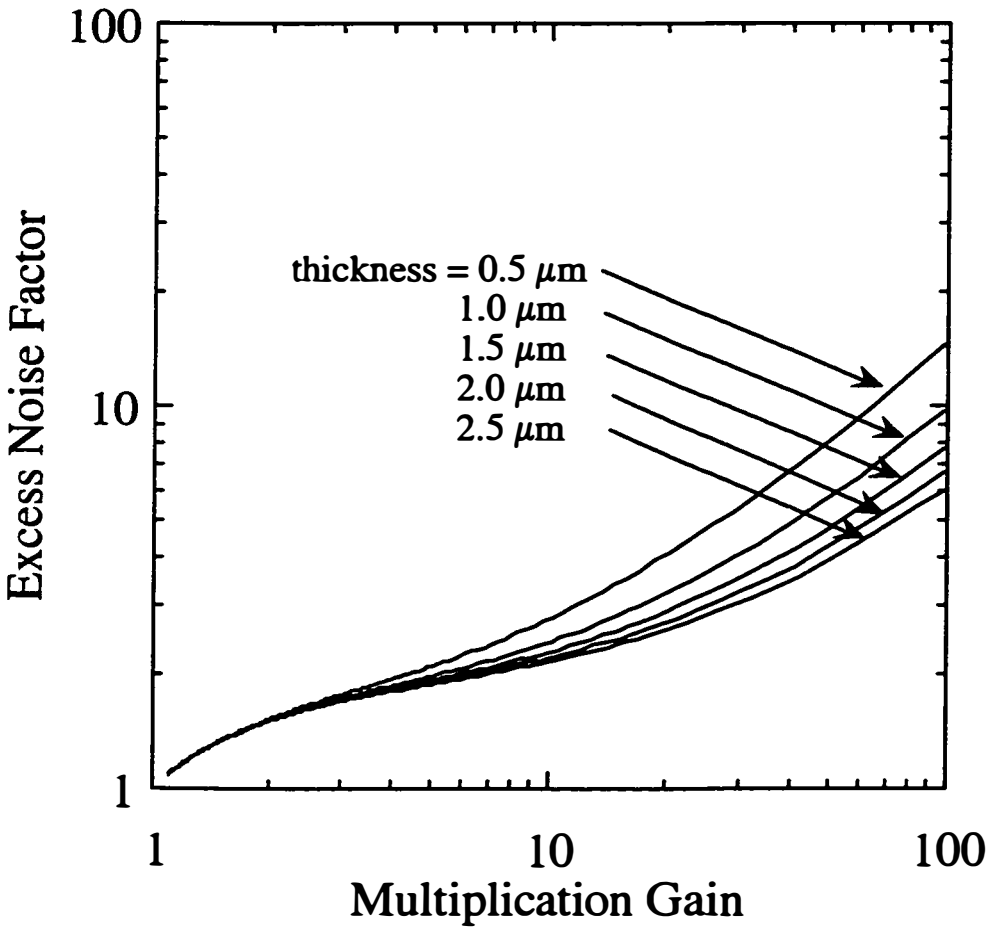


Figure 5.1 - Excess noise factor versus avalanche multiplication gain for a silicon multiplication layer. Five different layer thicknesses are shown and the curves were generated using Equation 5.7 and published ionization coefficients.[3]

reported values[3] for the ionization coefficients of silicon to obtain values of k at a given gain. As shown by the figure, increases in the gain causes increases in the noise. For all thicknesses, F_n remains fairly low, around two for gains up to ten. Above gains of ten, the curves begin to separate and the excess noise begins to change more rapidly. In the worst case, for a layer thickness of 0.5 μm and a gain of 100, the excess noise factor is still less than 15. Thicknesses above 1.0 μm have noise factors less than ten for gains up to 100. Figures that follow will show that this is quite good compared to InP.

For large values of gain, the excess noise factor approaches $F_n \approx kM$, the ionization coefficient ratio times the gain. The value for k can often be determined by using this relationship for excess noise at high gains. The curves indicate that the thicker multiplication layers have lower noise factors, and in effect lower k values. This can be understood by remembering that ionization coefficients are very electric field dependent and in silicon the k ratio gets larger at higher electric fields. The thinner multiplication layers require higher fields than thicker layer to produce equivalent gains. Thus as a function of gain, thicker regions have lower k values and lower excess noise values. This is true for all semiconductors based on the noise model set forth here. If noise was the only consideration, multiplication regions would be made very thick to take advantage of the lowest noise operation possible. There are many other factors that come into play however that dictate the multiplication region

thickness such as operating voltage and bandwidth. In both of these cases a thinner region gives more desirable results. The correct choice for thickness is in general the thickest layer that still operates in the desired bandwidth and voltage range.

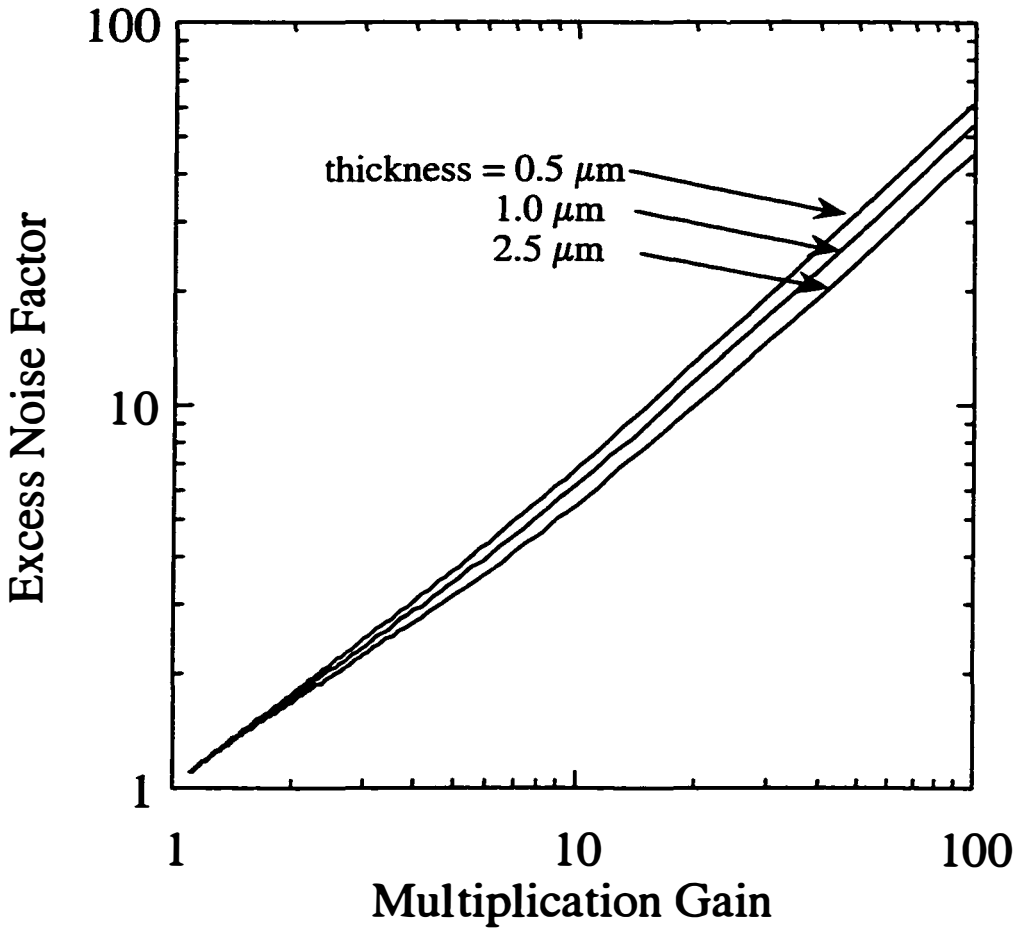


Figure 5.2 - Excess noise factor versus avalanche multiplication gain for an InP multiplication layer. Three different layer thicknesses are shown and the curves were generated using Equation 5.8 and published ionization coefficients[4].

Figure 5.2 shows a plot of F_p for InP as a function of gain for various multiplication layer thicknesses. This graph was made by using reported values[6] for the ionization coefficients at a given gain to obtain the correct k value. As with the plot for silicon shown in Figure 5.1, the excess noise factor increases with gain. This case shows a much faster rise in the excess noise at lower gains however. The excess noise values at a gain of 10 are near those of silicon at a gain of 100. The InP gain curves in this case approach the limit where $F_p \approx M/k$ as M becomes large and this limit is quite accurate for InP even at gains as low as ten. Another difference between the silicon curves in Figure 5.1 and those in Figure 5.2 is the relative insensitivity of the excess noise factor to multiplication layer thickness for InP. The 0.5 μm and 2.5 μm thickness cases are not dramatically different. This indicates that the k value does not vary much from the thick to thin layer cases in InP.

The curves in Figures 5.1 and 5.2 are based on the theoretical model for excess noise factor. Another comparison can be made using experimental curves for InP[5] and silicon APDs. Figure 5.3 shows F_p versus gain measured on an APD with an InP multiplication layer approximately 2.0 μm thick. F_a is shown for a silicon multiplication layer of around the same thickness. These plots are shown on a linear scale in the x and y axis as opposed to a log-log scale as before. As the graph shows, the silicon device has a much lower excess

noise factor. At a gain of 100, its noise is comparable to an InP APD at a gain of 10.

The noise models as well as experimental results show a clear advantage for using silicon as a multiplication layer over InP. In essence, a silicon APD

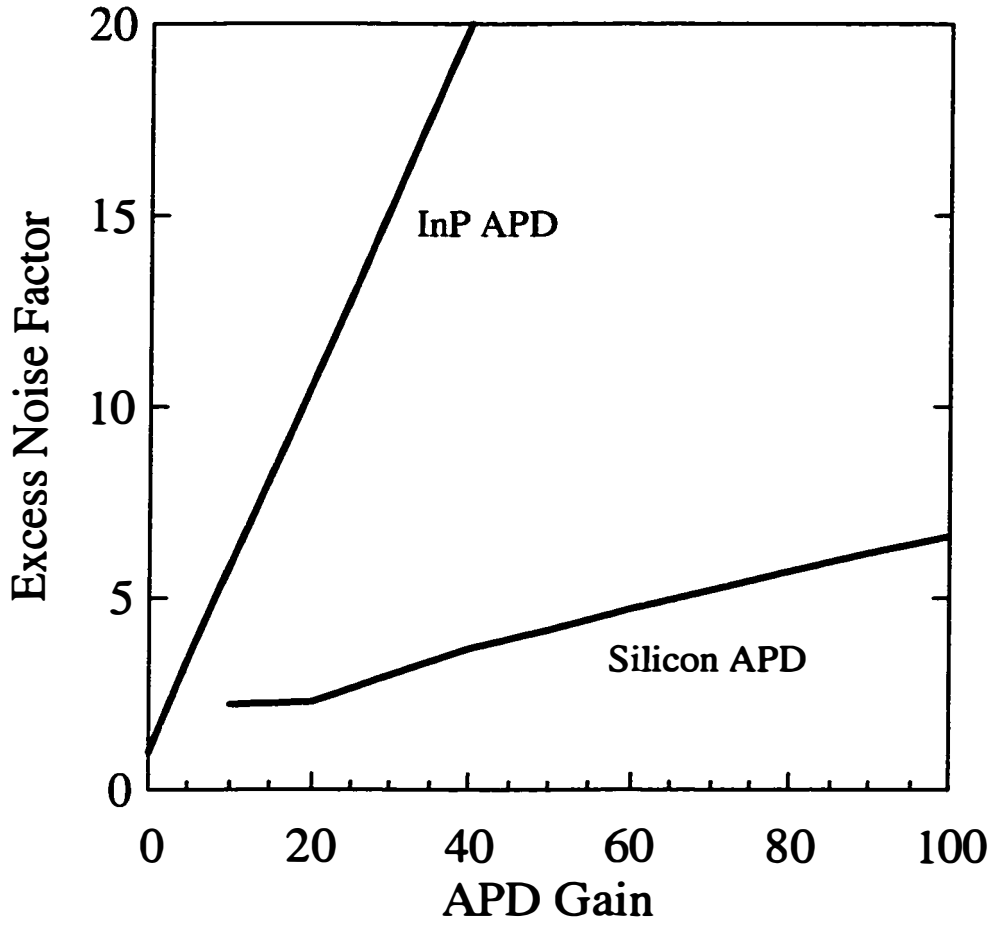


Figure 5.3 - Measured excess noise factor versus gain for one APD with an InP[6] multiplication layer and another with a silicon[7] multiplication layer. The multiplication layers were both approximately 2.0 μm thick.

can run at 10 times the gain of an InP APD and still maintain the same noise level. This could well translate into an increase in the sensitivity[8,9,10] of an APD detector used for optical communication systems since less light would be needed to produce an equivalent electrical signal. This would mean longer lines of optical fiber could be stretched between an optical transmitter and receiver with an equivalent performance.

5.3 The Groundwork for Bandwidth Equations

The frequency response for an APD will be shown in later sections to be a very complex function involving the ionization coefficients for a given semiconductor. In order to calculate an accurate bandwidth, accurate values for these coefficients must be known. As stated previously, these ionization coefficients vary with electric field and experiments to determine these coefficients by measuring carrier multiplication (Section 3.3) have yielded widely different results[11,12,13] for silicon. This is especially true at high electric fields. One reason for the discrepancies is that the APDs used to make these measurements utilized the same region for multiplication and absorption. A very exact knowledge of electric field profiles and device structures are required to make an accurate determination of ionization coefficients. Obtaining exact numbers for electric fields and thicknesses can be difficult for the thin semiconductor layers required to measure gain at high fields. Figure 5.4 shows

the differences in the ionization coefficients reported by three experimentalists. Theoretical models for the silicon ionization coefficients have also been reported[14,15,16,17], but their accuracy depends on the experimental values one believes.

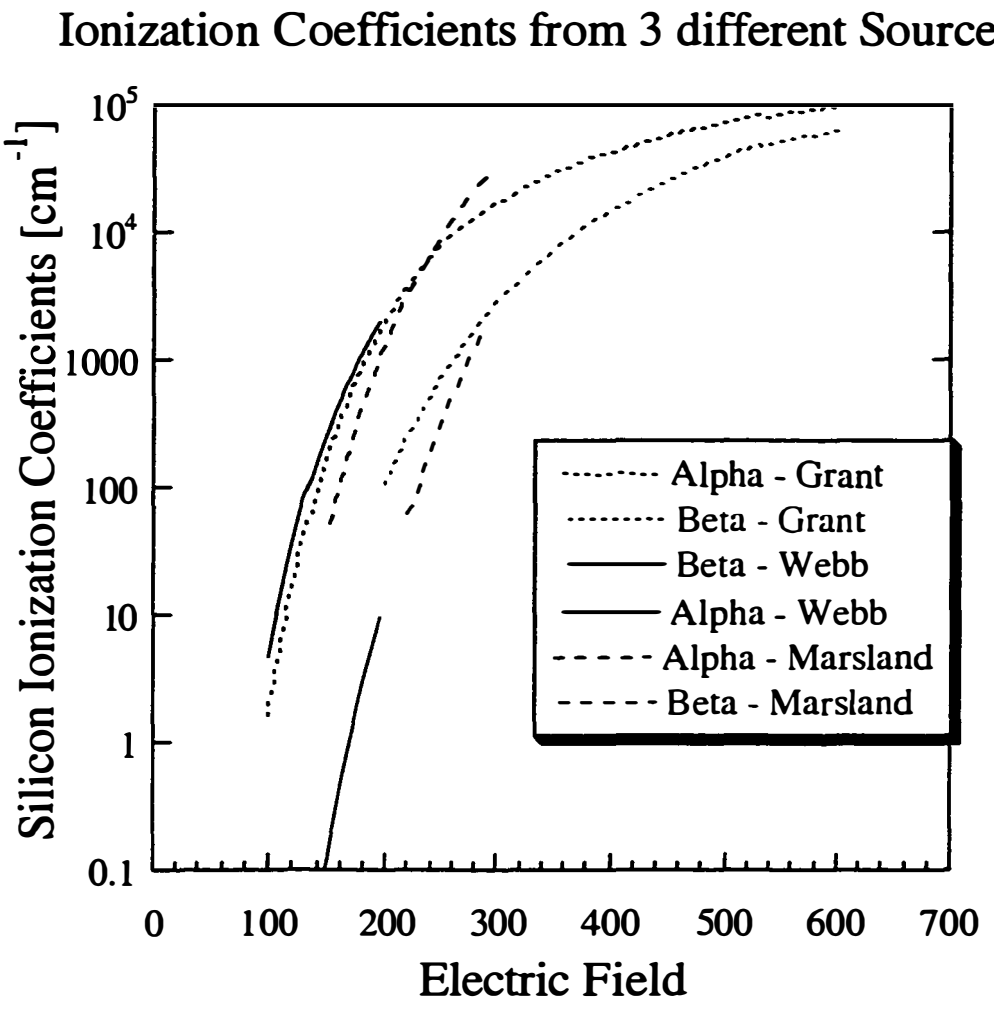


Figure 5.4 - Experimental silicon ionization coefficients versus electric field reported by three different authors, Grant, Webb, and Marsland.

An alternative method for determining the ionization coefficients is to measure the electron to hole coefficient ratio, k, using noise versus gain measurements as was alluded to in Section 3.2. The individual coefficients at

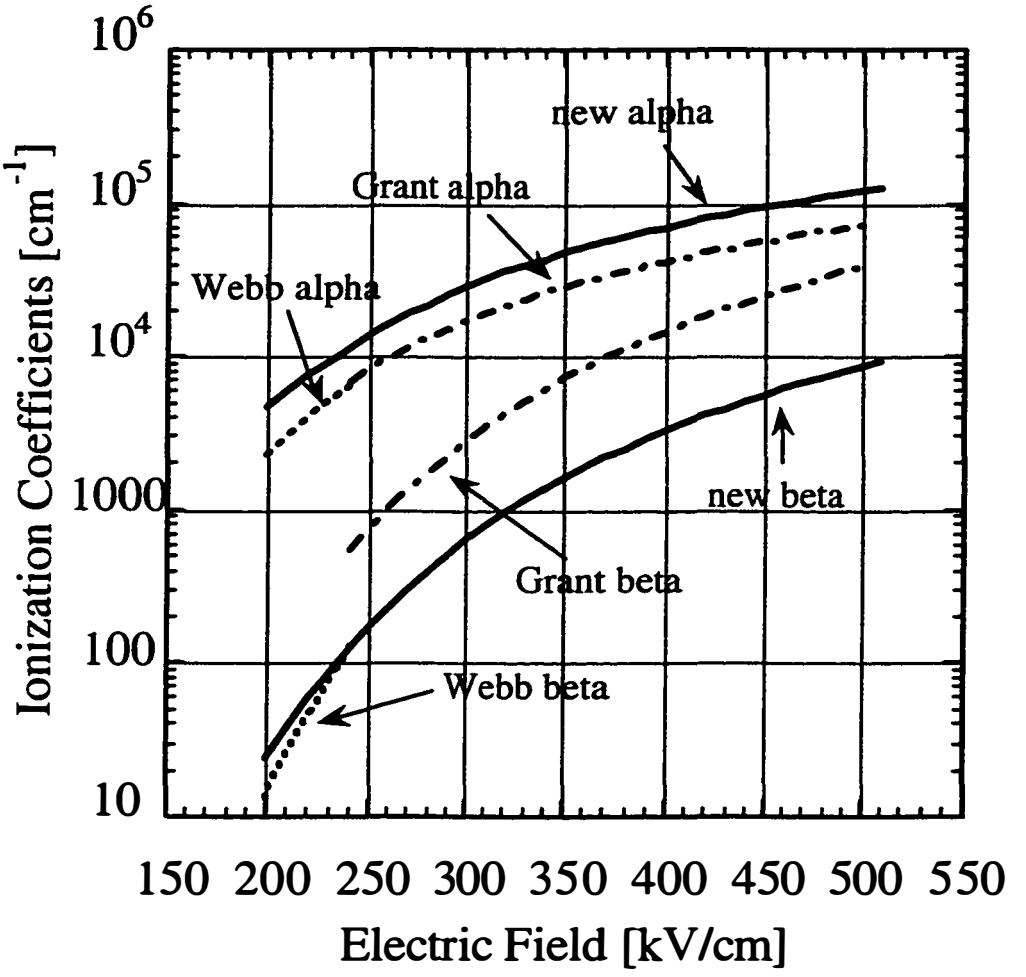


Figure 5.5 - New silicon ionization coefficients derived from Kaneda's experimental data, plotted versus electric field (Equations 5.9 and 5.10). Coefficients previously reported by Webb and Grant are also plotted.

different electric fields can then be determined based on the width of the multiplication region. This method should be much more accurate at high electric fields. Keneda[1] did measurements of the ratio, *k*, in silicon using a "reach-through" structure similar to a SAM APD. The multiplication layers were relatively thin for these reach-through structures allowing for measurements of a number of detectors at high electric fields. Based on the measurements by Keneda, new electron and hole ionization coefficients have been computed for silicon[18]. The values for these coefficients are given by the following equations

$$\alpha = 1.04 \times 10^6 \exp(-1.08 \times 10^6 / E) \qquad (5.9)$$

$$\beta = 0.45 \times 10^6 \exp(-1.97 \times 10^6 / E) \qquad (5.10)$$

where E is the electric field in units of V/cm. These coefficients are shown along with those measured by Grant and Webb in Figure 5.5.

An important measure of the accuracy of ionization coefficients is how well they predict the k ratio at a given electric field. Figure 5.6 plots k ratio versus electric field for all of the coefficients shown in Figure 5.5. Experimental k values are also plotted, taken from Kaneda's noise measurements on reach-through silicon diodes. As can be seen, in the region from 300 kV/cm to 500 kV/cm the new coefficients from Equations 5.9 and 5.10 most accurately reflect the measured k values. Grant's coefficients predict k values much too high in this range and Webb's coefficients were not intended

for fields above 250 kV/cm. The coefficients derived from the Keneda data should be more applicable in computing the frequency response of InGaAs-silicon APDs for several reasons. First of all the values will be accurate at high electric fields

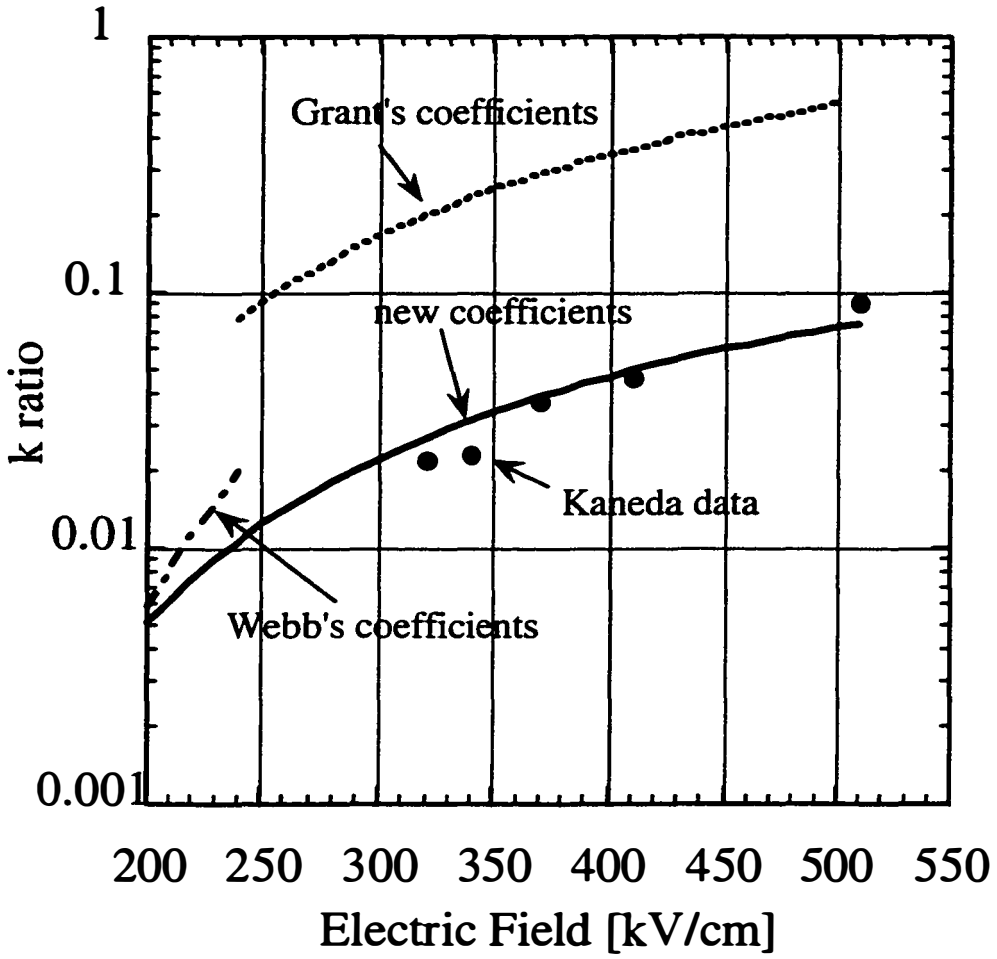


Figure 5.6 - k ratio versus electric field for three sets of ionization coefficients. The solid line represents the new coefficients calculated from Kaneda's data (Equations 5.9

and 5.10), the dotted line comes from Grant's coefficients, and the dotted and dashed line is calculated from Webb's coefficients. Kaneda's experimental data is shown by the solid circles.

where very fast APDs will be required to operate. In order to calculate ultimate bandwidth in the InGaAs-silicon APD, accuracy at these fields is crucial. A second argument for using the Keneda data stems from the similarities in the structures between the devices he measured and the InGaAs-silicon design (Chapter 2). Both utilize thin multiplication layers and almost pure electron injection. Silicon APDs used in other experiments had thicker multiplication layers, mainly because their absorption layers were very thick and thinning up the multiplication layer did not improve the speed significantly - high frequency operation was not the end goal. For the InGaAs-silicon APD to see its ultimate potential, it must operate in the GHz frequency domain, a region foreign to most silicon-only APDs.

The calculations done in Sections 5.4 and 5.5 use the new coefficients derived from Kaneda's data. All of the other silicon APD curves generated up to this point used previously published ionization coefficients[3]. There are several reasons for this. First of all, the previous results would not change significantly with different coefficients. Using accurate coefficients for thin multiplication layers is more important for frequency response calculations,

especially when comparing them with experimental data. Examining the different coefficients in Figure 5.5 and 5.6, the new coefficients give the most "optimistic" view of silicon - the lowest k ratio. This brings up a second reason why previously reported coefficients were used - they represent a conservative measure. Showing silicon's superiority with a "worst case" scenario is more credible than presenting calculations one might argue are superior only because they are based on optimistic coefficients tailored for that purpose. Using the new coefficients to determine frequency response seemed appropriate, however, since there was experimental data available to verify the calculations (Chapter 8).

5.4 Gain-bandwidth Equations - Effective Multiplication Plane Method

The first approach used for calculating bandwidth for InGaAs-silicon APDs was the effective multiplication plane method put forth by Wu[19]. In this approach carriers created through the multiplication process were assumed to have all been created at a distance Δd from the edge of the transition between the absorption layer and the multiplication layer in a SAM APD. This approach was more sophisticated than those presented previously which either ignored the contribution of secondary carriers[6] (created through multiplication action) or assumed that contributions from transit time effects, multiplication buildup time, and any carrier trapping were independent of each other[20]. While the latter

approaches were applicable to APDs with thin multiplication layers and triangular electric field profiles, they are not adequate for InGaAs-silicon APDs with comparatively thick multiplication layers. The equations that follow represent analytical expressions for frequency response for any APD with a separate absorption and multiplication layer. In this treatment, electrons are assumed to be the predominant ionizing carriers. Additional assumptions are that incident light strikes the absorption region through a transparent multiplication layer. Absorption layer thickness is given by w_a and multiplication layer thickness by d. These layers are assumed to be fully depleted so that the full depletion width is given by $w = w_a + d$.

The frequency response derivation begins by writing the photocurrent as

$$i_{ph}(t) = \frac{q}{w}[v_n N(t) + v_p P(t) + v_n N_s(t) + v_p P_s(t)] \qquad (5.11)$$

where v_n and v_p are the electron and hole saturated drift velocities, and N(t), P(t), $N_s(t)$, and $P_s(t)$, are the total number of uncollected photogenerated primary electrons, primary holes, secondary electrons, and secondary holes in the depletion region, respectively. In the frequency domain, the signal current can be expressed as

$$i_s(\omega) = \frac{q}{w} \frac{v_n \tilde{N}(\omega) + v_n \tilde{N}_s(\omega) + v_p \tilde{P}(\omega) + v_p \tilde{P}_s(\omega)}{1 - \omega^2 LC + j\omega C(R_s + R_l)} \qquad (5.12)$$

where L is the parasitic inductance, C is the total device capacitance, R_s the series resistance, R_l the load resistance, and $\tilde{N}(\omega)$, $\tilde{P}(\omega)$, $\tilde{N}_s(\omega)$, and $\tilde{P}_s(\omega)$

are the Fourier transforms of N(t), P(t), $N_s(t)$, and $P_s(t)$ respectively. The normalized frequency response is given by the signal current divided by the dc photocurrent or $i_s(\omega)/i_s(0)$. And the dc photocurrent can be written as

$$i_s(0) = \frac{q\eta P_o}{h\nu} M\left[1 - \exp(-\alpha_o w_a)\right] \qquad (5.13)$$

where η is the quantum efficiency of the absorption layer, P_o is the input signal power, M the dc gain in the APD, and α_o the optical absorption coefficient in the absorption layer. The frequency response can then be written as

$$\frac{i_s(\omega)}{i_s(0)} = \frac{h\nu}{\eta P_o M\left[1 - \exp(-\alpha_o w_a)\right]} \times \frac{1}{1 - \omega^2 LC + j\omega RC} \times \frac{v_n \tilde{N}(\omega) + v_n \tilde{N}_s(\omega) + v_p \tilde{P}(\omega) + v_p \tilde{P}_s(\omega)}{w} \qquad (5.14)$$

To evaluate the frequency response, Fourier transforms must be performed on each of the expressions describing carriers generated by an input optical signal. The impulse response for an APD illuminated from the multiplication layer side is given by

$$N(t) = \frac{\eta P_o}{h\nu}\left[1 - \exp(-\alpha_o w_a)\right]\left[u(t) - u(t - t2)\right] + \left[\exp(-\alpha_o v_n (t - t_2)) - \exp(-\alpha_o w_a)\right]\left[u(t - t_2) - u(t - t_2 - t_3)\right] \qquad (5.15)$$

$$P(t) = \frac{\eta P_o}{h\nu}\left[1 - \exp(\alpha_o v_p (t - t_4))\right]\left[u(t) - u(t - t_4)\right] \qquad (5.16)$$

$$\frac{P_s(t)}{M_o - 1} = \frac{\eta P_o}{h\nu}\left\{\begin{array}{l}\left[1-\exp(-\alpha_o \nu_n (t-t_1))\right]\left[u(t-t_1)-u(t-t_1-t_3)\right] \\ +\left[1-\exp(-\alpha_o w_a)\right]\left[u(t-t_1-t_3)-u(t-t_5)\right] \\ +\left[\exp(-\alpha_o \nu_n (t-t_5))-\exp(-\alpha_o w_a)\right]\times \\ \left[u(t-t_5)-u(t-t_3-t_5)\right]\end{array}\right\} \otimes \frac{\exp\left[\frac{-t}{(M-1)\tau_m}\right]}{(M-1)\tau_m}$$

(5.17)

$$\frac{N_s(t)}{M-1} = \frac{\eta P_o}{h\nu}\left\{\begin{array}{l}\left[1-\exp(-\alpha_o \nu_n (t-t_1))\right]\left[u(t-t_1)-u(t-t_2)\right] \\ +\left[\exp(-\alpha_o \nu_n (t-t_2))-\exp(-\alpha \nu_n (t-t_1))\right] \\ \times\left[u(t-t_2)-u(t-t_1-t_3)\right]\end{array}\right\}$$

$$+\frac{\eta P_o}{h\nu}\left\{\begin{array}{l}\left[\exp(-\alpha_o \nu_n (t-t_2))-\exp(-\alpha_0 w_a)\right] \\ \times\left[u(t-t_1-t_3)-u(t-t_2-t_3)\right]\end{array}\right\} \otimes \frac{\exp\left[\frac{-t}{(M-1)\tau_m}\right]}{(M-1)\tau_m}$$

(5.18)

$$(d < w_a + 2\Delta d)$$

or

$$\frac{N_s(t)}{M-1} = \frac{\eta P_o}{h\nu}\left\{\begin{array}{l}\left[1-\exp(-\alpha_o \nu_n (t-t_1))\right]\left[u(t-t_1)-u(t-t_1-t_3)\right] \\ +\left[1-\exp(-\alpha_o w_a)\right]\left[u(t-t_1-t_3)-u(t-t_2)\right]\end{array}\right\}$$

$$+\frac{\eta P_o}{h\nu}\left\{\begin{array}{l}\left[\exp(-\alpha_o \nu_n (t-t_2))-\exp(-\alpha_o w_a)\right] \\ \times\left[u(t-t_2)-u(t-t_2-t_3)\right]\end{array}\right\} \otimes \frac{\exp\left[\frac{-t}{(M-1)\tau_m}\right]}{(M-1)\tau_m}$$

(5.19)

$$(d \geq w_a + 2\Delta d)$$

where $t_1 = \Delta d/v_n$, $t_2 = d/v_n$, $t_3 = w_a/v_n$, $t_4 = w_a/v_p$, $t_5 = w_a/v_p + \Delta d/v_n + \Delta d/v_p$, and τ_m is the avalanche buildup time constant of the APD given by

$$\tau_m = \frac{1}{v_n + v_p}\int_0^d \kappa \exp\left[-\int_0^x [\alpha(x') - \beta(x')]dx'\right]dx. \qquad (5.20).$$

κ is a correction factor that varies with β/α and the ratio of the saturation velocities.[21] The τ_m[22] term contains the information about the multiplication properties of the semiconductor through the ionization coefficients. The Fourier transforms of the impulse responses are given by

$$\tilde{N}(\omega) = \frac{\eta P_o}{h\nu}\left[\begin{array}{l}\left[\dfrac{1-\exp(-\alpha_o w_a)}{j\omega} + \exp(-j\omega d/v_n)\right] \\ \times\left[1-\exp(-\alpha_o w_a - j\omega w_a/vn)\right]\left\{\dfrac{1}{j\omega + \alpha_o v_n} - \dfrac{1}{j\omega}\right\}\end{array}\right], \qquad (5.21)$$

$$\tilde{P}(\omega) = \frac{\eta P_o}{h\nu}\left[\frac{1-\exp(-j\omega w_a/v_p)}{j\omega} + \frac{\exp(-j\omega w_a/v_p) - \exp(-\alpha_o w_a)}{j\omega - \alpha_o v_p}\right],$$

(5.22)

$$\frac{\tilde{P}_s(\omega)}{M-1} = \frac{\eta P_o}{h\nu}\exp(-j\omega\Delta d)\left[1-\exp(-\alpha_o w_a - j\omega w_a/v_n)\right] \times\left[1-\exp(-j\omega(w_a+\Delta d)/v_p)\right]\left[\frac{1}{j\omega} - \frac{1}{j\omega + \alpha_o v_n}\right]\frac{1}{1+j\omega(M-1)\tau_m},$$

(5.23)

$$\frac{\tilde{N}_s(\omega)}{M-1} = \frac{\eta P_o}{h\nu}\exp(-j\omega\Delta d)\left[1-\exp(-\alpha_o w_a - j\omega w_a / v_n)\right]$$
$$\times\left[1-\exp(-j\omega(d+\Delta d)/v_n)\right]\left[\frac{1}{j\omega}-\frac{1}{j\omega+\alpha_o v_n}\right]\frac{1}{1+j\omega(M-1)\tau_m}$$

(5.24)

These expressions provide analytical expressions for the frequency response, but they are not the perfect solution. Their accuracy depends to a large degree on the selection of the Δd parameter that places the multiplication plane within the multiplication region. There is no way to know exactly what this quantity should be for a given APD. In practice the equations would be fit to experimental data, shifting the Δd parameter until the theoretical curves matched experimental ones. To predict bandwidths for unknown structures really requires first constructing and measuring a device typifying the structure in question. A measure for Δd can then be made and projections made for similar structures. This can be very useful for some situations and can provide additional insight into the multiplication properties of the detector. A plot of 3-dB bandwidth versus gain created using the effective-plane method is shown in Figure 5.8. The 3-dB bandwidth was determined at a given gain for an InGaAs-silicon APD with a 0.5 μm multiplication layer and a 1.0 μm absorption layer. Δd was assumed to be 0.3 μm in this case. These results are plotted with those created from a more exact approach described in Section 5.5 to show their similarities. Both types of curves are discussed and compared in Section 5.6.

5.5 Gain-bandwidth Equations - Hollenhorst approach

The effective-plane equations provide a good first look at frequency response in simple APD structures. The equations can also be modified to take into account more and more complex layer structures, but this necessarily complicates the math. This solution will always be lacking however because of the assumption that multiplication all occurs at one place and because that place is not really well defined. Fortunately a more exact method was developed by Hollenhorst[23] and independently by Kahraman[24]. This approach though not great for modeling very complex devices, is very applicable for InGaAs-silicon APDs with only two layers and relatively uniform electric field profiles[18]. The Hollenhorst approach is more mathematically difficult than the effective-plane method, but it eliminates the nagging assumption of multiplication at only one plane by allowing for ionization events throughout the multiplication layer. In the equations that follow, the intent is not to rigorously describe the Hollenhorst method, but to provide a brief outline, pointing towards references to render more exact description of some parameters. Perhaps more importantly, in Section 5.6 that follows curves generated by this method are shown and discussed.

Hollenhorst's method uses vector and matrix notation to keep track of current density flowing into and out of a semiconductor region. Using his

notation, the current density components at a given frequency can be written as the following equation

$$\begin{bmatrix} J_{p1} \\ J_{n1} \end{bmatrix} = \begin{bmatrix} T_{pp} & T_{pn} \\ T_{np} & T_{nn} \end{bmatrix} \begin{bmatrix} J_{p0} \\ J_{n0} \end{bmatrix} + \begin{bmatrix} S_p \\ S_n \end{bmatrix}, \qquad (5.25)$$

or more simply

$$\vec{J}_1 = \mathbf{T}\vec{J}_0 + \vec{S}. \qquad (5.26)$$

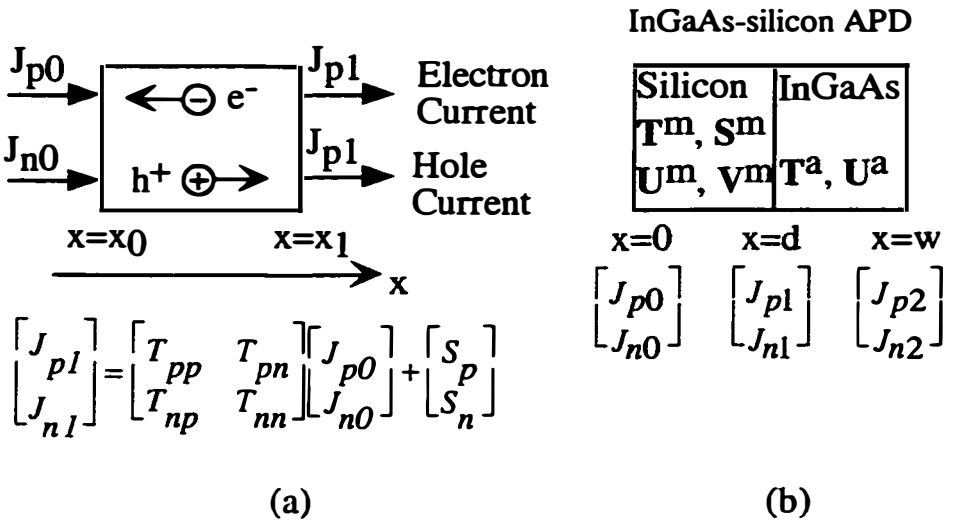
Here p and n are used to designate hole and electron components. **T** is a current transfer matrix that varies with different layers depending on whether they are multiplying layers or absorption layers. $\vec{S}$ is a source current vector which represents current created from absorption of optical light. J_{p1}, J_{n1}, J_{p0}, and J_{n0} represent the electron and hole currents into and out of a region as illustrated in Figure 5.7(a). The electrode current density due to $\vec{J}_0$ is given by the equation

$$I(\omega) = \frac{1}{w}(\vec{U} \cdot \vec{J}_0 + V) \qquad (5.27)$$

where w is the depletion layer thickness of the APD and $\vec{U} = \begin{bmatrix} U_p \\ U_n \end{bmatrix}$ is a vector that relates the proportionality of the left-hand current densities (J_{p0} and J_{n0}) and the electrode current. This vector takes into account mechanisms such as avalanche gain that add to or diminish the current densities in transitioning across the region. V represents the contribution to the current caused by optical absorption. Detailed expressions have been derived by Hollenhorst[23] for the

quantities **T**, $\vec{S}$, $\vec{U}$, and V applicable for layers most commonly used in the construction of APDs. Hollenhorst assumes that holes propagate in the +x direction and electrons in the -x direction as in Figure 5.7(a).

A two-layer device such as the InGaAs-silicon APD shown in Figure 5.7(b) fits quite naturally into the Hollenhorst derivations. The important quantities relating current densities from x=0 to x = w are given by the



$$\begin{bmatrix} J_{p1} \\ J_{n1} \end{bmatrix} = \begin{bmatrix} T_{pp} & T_{pn} \\ T_{np} & T_{nn} \end{bmatrix} \begin{bmatrix} J_{p0} \\ J_{n0} \end{bmatrix} + \begin{bmatrix} S_p \\ S_n \end{bmatrix}$$

Figure 5.7 (a) Electron and hole current densities in a uniform semiconductor layer of length x_1. Current transport through this layer is described by the transmission matrix **T**, and the source vector $\vec{S}$. (b) A two layer InGaAs-silicon APD with a silicon multiplication layer of width d and an InGaAs absorption layer of width w-d. Transmission through the multiplication layer is described by the terms with the m superscript, transmission through the absorption layer by the a superscript.

following, according to the composition rules of Equation 8 in Reference 23.

$$\mathbf{T} = \mathbf{T}^a\mathbf{T}^m \qquad (5.28)$$

$$\vec{S} = \vec{S^a} \quad (\vec{S^m} = 0) \qquad (5.29)$$

$$\vec{U^T} = (\vec{U^m})^T + (\vec{U^a})^T\mathbf{T}^m \qquad (5.30)$$

$$V = V^a \quad (V^m = 0) \qquad (5.31)$$

The superscripts a and m represent absorption and multiplication terms respectively. The transfer matrix $\mathbf{T}$ and proportionality vector $\vec{U}$ have elements taking into account the optical absorption as well as avalanche multiplication. The source vector $\vec{S}$ and quantity V have only absorption layer components since it is assumed no optical absorption occurs in the silicon multiplication layer. A remaining relation that must be considered is that for injected currents from the left and right side of the detector structure defined as $\vec{J_{in}} = \begin{bmatrix} J_{p0} \\ J_{n2} \end{bmatrix}$. Using this quantity the electrode current can be rewritten as

$$I(\omega) = \frac{1}{w}(\vec{\rho} \cdot \vec{J_{in}} + \delta). \qquad (5.32)$$

This designates the electrode current as a combination of currents from injected sources and optical sources. For the InGaAs-silicon APD, the assumption is made that the contribution from injected currents is small (low dark currents).

Given Equations 5.26 through 5.32, the following equation can be written that express the electrode current that results from photon absorption (δ/w):

$$I(\omega)=\frac{1}{w}\left[\left(V^{a}-\frac{U_{n}^{a}S_{n}^{a}}{T_{nn}^{a}}\right)-\frac{S_{n}^{a}}{T_{nn}^{a}}\left(U_{n}^{m}+U_{p}^{a}T_{pn}^{m}\right)\right]. \qquad (5.33)$$

The quantities within the brackets are all given in Reference 23. The dc electrode current is given by Equation 5.11. Including parasitic effects, the frequency response of the detector illustrated in Figure 5.6(b) can be written as

$$\text{Frequency Response}\equiv\frac{1}{1-\omega^{2}LC+j\omega RC}\frac{I(\omega)}{I(0)} \qquad (5.34)$$

where L is the inductance, C the total capacitance, and R the sum of the series and load resistances. A combination of these equations was used to create the graphs for various aspects of bandwidth plotted in the next section.

5.6 Bandwidth Plots for InGaAs-silicon APDs

With the methods for determining bandwidth outlined in Sections 5.4 and 5.5, plots can be made to determine the performance to be expected from the InGaAs-silicon APD along with comparing this to known InP APDs. The first plot shown in Figure 5.8 is a typical measure for avalanche photodiodes. The graph plots 3-dB bandwidth (the bandwidth for which the current response has dropped 3-dB from the DC value) versus gain. Four curves are shown. The

first was generated using the effective-plane method based on an InGaAs-silicon APD with a 0.5 μm multiplication layer and a 1.0 μm absorption layer. The next curves were generated using the Hollenhorst method for InGaAs-silicon APDs with 0.4 μm and 1.0 μm multiplication layers, and 1.0 μm absorption layers. The final curve is an experimental one for an InGaAs/InP APD with a 0.5 μm multiplication layer and 2.0 μm absorption layer[25] (experimental plots for InGaAs-silicon APDs will be shown in Chapter 8). The silicon plots in the figure illustrate the classic gain-bandwidth curves typical for APDs. There is an initial portion of the curve that remains relatively flat out to a given gain and the bandwidth is limited by parasitic or transit time effects. In the case of the silicon based APDs this region ends around a gain of 20. There is then a gradual roll-off of the bandwidth that asymptotes to an ultimate gain-bandwidth product. This roll off occurs as avalanche multiplication build-up time becomes longer than either the RC time constants or simple transit times. The gain-bandwidth product is simply the bandwidth at a given gain times that gain, and all APDs have an ultimate gain-bandwidth product. The 0.4 μm and 1.0 μm silicon based diodes in Figure 24 have ultimate gain-bandwidths of around 500 GHz and 320 GHz respectively, while the InP based APD has an ultimate gain-bandwidth of 70 GHz. Interesting to note are the differences between the curves in the figure. APDs with more narrow multiplication

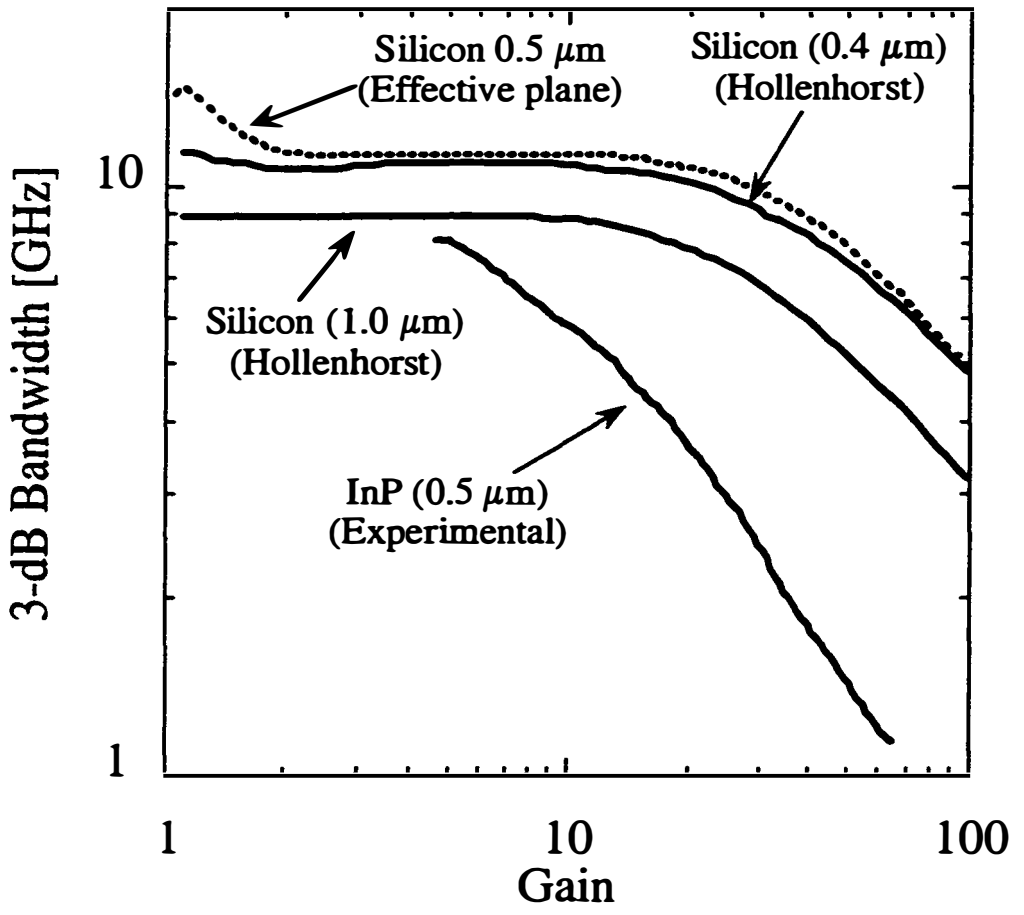


Figure 5.8 - 3-dB bandwidth versus gain for InGaAs-silicon and InGaAs/InP APDs. The InGaAs-silicon curves are theoretical, generated using the effective-plane and Hollenhorst methods. The silicon diodes were assumed to have a 1.0 μm InGaAs layer and silicon layer thicknesses as indicated on the plot. The InGaAs/InP curve was measured using an APD with a 2.0 InGaAs layer and a 0.5 μm InP multiplication layer.

regions should exhibit higher bandwidths and higher gain-bandwidth products, as in the differences between the 0.4 μm and 1.0 μm curves generated using Hollenhorst's method. The curve from the effective-plane method actually indicates higher bandwidth than the thinner diode from the Hollenhorst method. This illustrates the shortcoming of the effective-plane theory, where not all of the multiplication effects are accounted for. The curve also illustrates the essential differences between the silicon based multiplication layer and the InP one. Even at a gain of four, the InP bandwidth is beginning to decrease. This decrease is once again dictated by the ionization coefficient ratio k. The ultimate gain-bandwidth product for APDs with predominantly electron ionization can be shown to reach the value given by

$$GB = \frac{1}{\pi\kappa k(w / v_p + w / v_n)} \qquad (5.35).$$

Silicon with its very disparate ionization coefficients can thus achieve very high gain-bandwidths, much higher than possible in InP. This has important implications in commercial fiber optic systems operating at 2.5 and 10 GHz. For 2.5 GHz systems, silicon based APDs could operate at higher gains than InP diodes while maintaining the required 2.5 GHz bandwidth. This is true for 10 GHz systems as well, for which it is difficult to build an InP APD with any gain at 10 GHz.

Speed is not the only important parameter for an APD. If it were, APDs would be made extremely thin to shorten any transit times. Speed is always balanced by quantum efficiency when selecting device thickness however. Very thin diodes have almost no optical absorption. Thick diodes have lots of light absorption, but can have limited bandwidths. Figure 5.9 examines the balance between speed and quantum efficiency for an InGaAs-silicon APD. Curves were generated using Hollenhorst's approach, which show quantum efficiency versus ultimate gain-bandwidth for incident 1.3 μm light. Four modeled devices are plotted with different thicknesses represented by their 3-dB bandwidth at a gain of ten. The 3 GHz curve represents the case in which either the InGaAs absorption layer or the silicon layer is fairly thick. If the InGaAs layer is thick, the quantum efficiency is very high and the silicon must be very thin, pushing the gain-bandwidth very high. If the InGaAs layer is thin, the silicon layer must be quite thick to drop the bandwidth to 3 GHz and so the gain-bandwidth drops fairly low. The 10 GHz case is very different. There is a limit to the amount of InGaAs one may use and still operate at 10 GHz. For quantum efficiencies above 0.6, the silicon must be quite thin leading to high gain-bandwidths. Even with very little InGaAs and low quantum efficiencies, the gain-bandwidth must be quite high (thin silicon layer) to operate at 10 GHz. This curve demonstrates that high quantum efficiency InGaAs-silicon APDs are certainly possible however, even at a 10 GHz operating frequency.

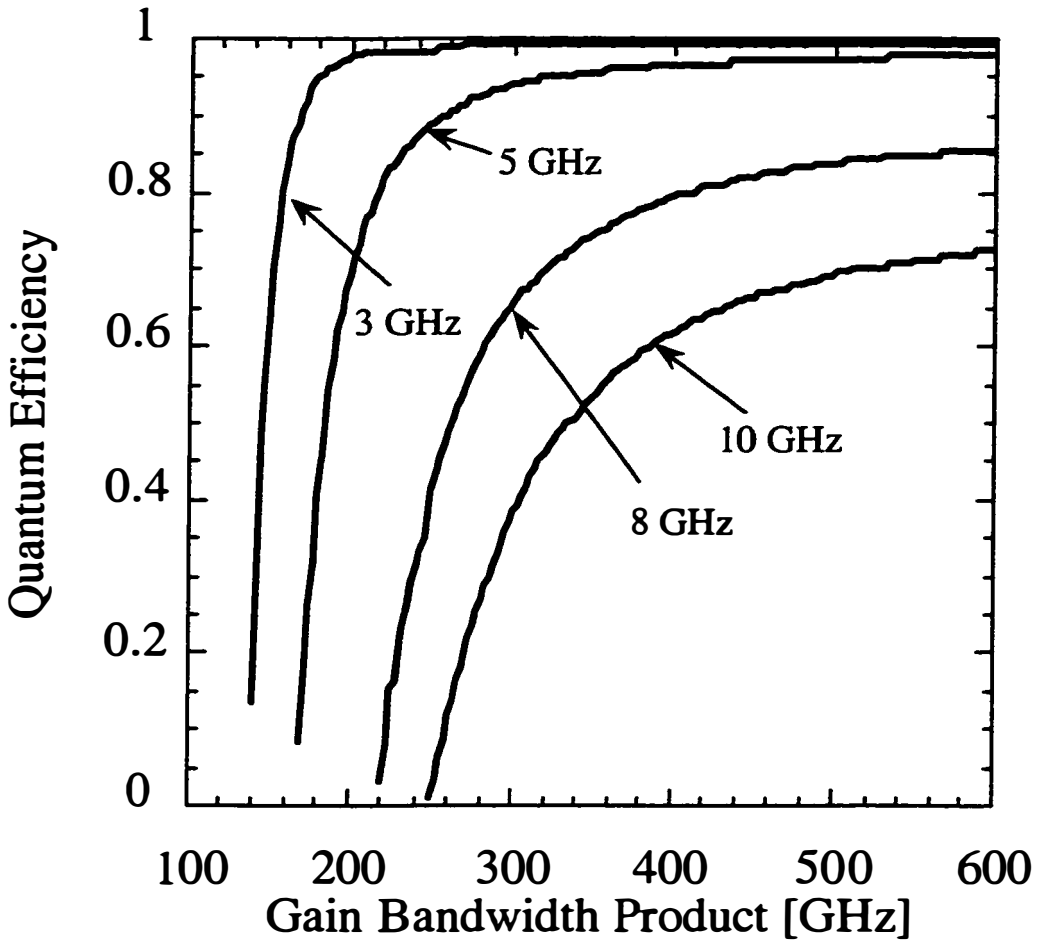


Figure 5.9 - Quantum efficiency versus gain-bandwidth product for InGaAs-silicon APDs illuminated with 1.3 μm light. Curves were calculated using the Hollenhorst approach. Four modeled devices are plotted. The thickness ratio between InGaAs and silicon was calculated to maintain a constant 3-dB bandwidth at a gain of ten.

Another way to examine the ultimate performance of an InGaAs-silicon. APD is found in Figure 5.10. This is a plot of ultimate gain-bandwidth product versus multiplication layer thickness, assuming a 1.0 μm absorption layer, a

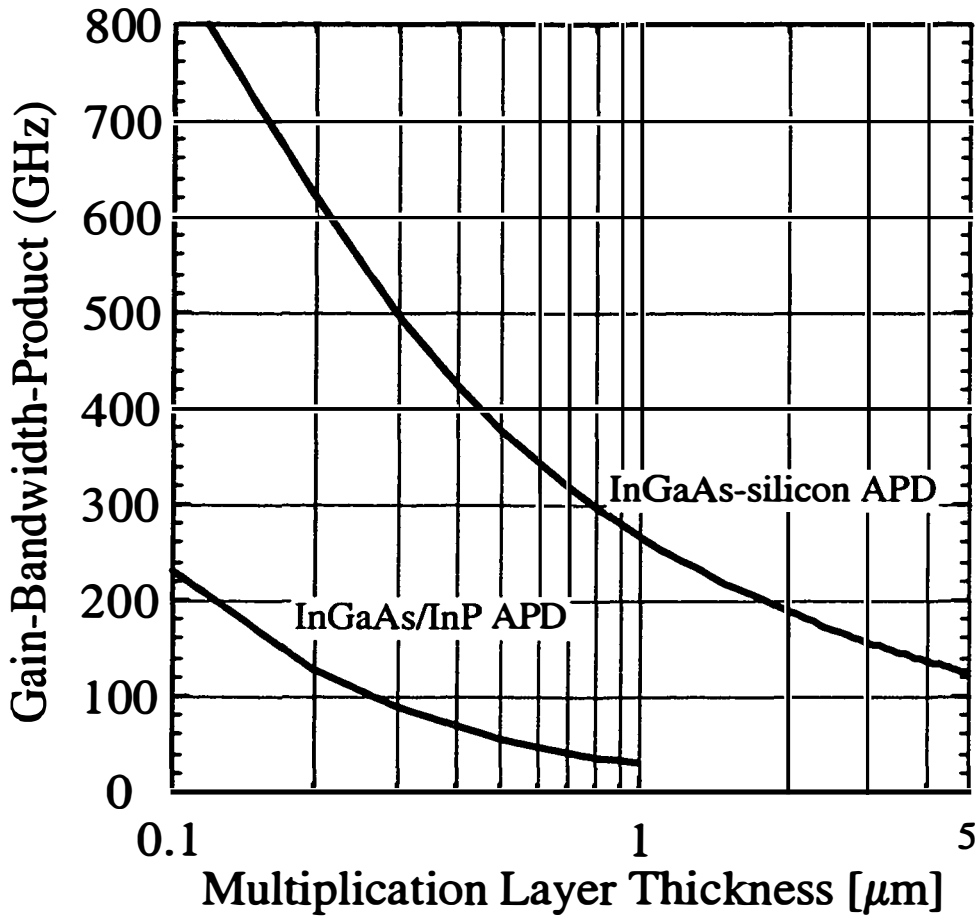


Figure 5.10 - Gain-bandwidth product for InGaAs-silicon and InGaAs/InP APDs versus multiplication layer thickness. The Hollenhorst approach was used for calculating gain-bandwidth assuming a 1.0 μm InGaAs absorption layer.

constant electric field profile in the multiplication layer, and using Hollenhorst's theory to generate the curve. Also plotted is the expected performance of a similar InP diode. The graph indicates the clear superiority of silicon. For example, with a 0.5 μm multiplication layer, the gain-bandwidth product achievable in silicon is five times higher than in InP. If the multiplication layer in InGaAs-silicon APDs can be made thin enough, the gain bandwidth product could conceivably be over 600 GHz.

5.7 Conclusion

This chapter has shown the expected noise and bandwidth characteristics of the InGaAs-silicon APD and demonstrated what should be its superiority over existing InGaAs/InP diodes. The equations for excess noise caused by multiplication were derived and silicon had effectively an order of magnitude less noise than InP at the same operating gains. The bandwidth and gain-bandwidth of these two types of diodes were also examined. The importance of accurate ionization coefficients was first discussed and new coefficients applicable to the InGaAs-silicon APD were presented. Two methods for calculating bandwidth were then presented, the effective-plane method, and the Hollenhorst method based on a current transfer matrix. The Hollenhorst technique, though more complicated, yields more accurate results for simple InGaAs-silicon structures. Both approaches, however, predict very high

bandwidths and gain-bandwidths for these detectors. Ultimate gain-bandwidth in silicon based APDs was shown to be nearly five times higher than for InP based diodes.

References

[1] R. J. McIntyre, "Multiplication Noise in Uniform Avalanche Diodes," IEEE Transactions on Electronic Devices, **ED-13**, pp. 164-168, (1966).

[2] Semiconductors and Semimetals, Volume 22 - Lightwave Communications Technology, edited by W. T. Tsang, "Physics of Avalanche Photodiodes," by F. Capasso, pp. 1-172, (1985), Academic Press, NY.

[3] Quick Reference Manual for Silicon Integrated Circuit Technology, edited by W. E. Beadle, J. C. C. Tsai, and R. D. Plummer, (1985), John Wiley & Sons, Inc., NY.

[4] C. A. Armiento and S. H. Groves, "Impact ionization in (100)-, (110)-, and (111)-oriented InP avalanche photodiodes", *Applied Physics Letters*, **43**, No. 2, pp. 198-200, (1983).

[5] J. C. Campbell, S. Chandrasekhar, W. T. Tsang, G. J. Qua, and B. C. Johnson, "Multiplication Noise of Wide-Bandwidth InP/InGaAsP/InGaAs Avalanche Photodiodes," *Journal of Lightwave Technology*, **7**, No. 3, pp. 473-478, (1989).

[6] J. C. Campbell, B. C. Johnson, G. J. Qua, and W. T. Tsang, "Frequency Response of InP/InGaAsP/InGaAs Avalanche Photodiodes," *Journal of Lightwave Technology*, **7**, No. 5, pp. 778-784, (1989).

[7] Semiconductors and Semimetals, Volume 22 - Lightwave Communications Technology, edited by W. T. Tsang, "Silicon and Germanium Avalanche Photodiodes," by T. Kaneda, pp. 247-328, (1985), Academic Press, NY.

[8] R. G. Smith and S. R. Forrest, "Sensitivity of Avalanche Photodetector Receivers for Long-Wavelength Optical Communications," *The Bell System Technical Journal*, **61**, No. 10, pp. 2929-2946, (1982).

[9] B. L. Kasper and J. C. Campbell, "Multgigabit-per-Second Avalanche Photodiode Lightwave Receivers," *Journal of Lightwave Technology*," **LT-5**, No. 10, pp. 1351-1364, (1987).

[10] F. Boudreau, "Optimizing APD's for High-Speed Communications," *Fiberoptic Product News*, **12**, pp. 29-30, (1995).

[11] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, **16**, pp. 1189-1203, (1973).

[12] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Hole and Electrons in Silicon," *Physical Review*, **134**, pp. A761-773, (1964).

[13] P.P. Webb, "Measurements of Ionization Coefficients in Silicon at Low Electric Fields," GE Canada Inc., Electro Optics Operations, unpublished.

[14] N. Goldsman, Y. -J. Wu, and J. Frey, "Efficient calculation of ionization coefficients in silicon from the energy destribution function," *Journal of Applied Physics*, **68**, No. 3, pp. 1075-1081, (1990).

[15] Y. -J. Wu and N. Goldsman, "Deterministic modeling of impact ionization with a random-k approximation and the multiband Bolzmann equation," *Journal of Applied Physics*, **78**, No. 8, pp. 5174-5176, (1995).

[16] V. M. Robbins, T. Wang, K. F. Bennan, K. Hess, and G. E. Stillman, "Electron and hole impact ionization coefficients in (100) and in (111) Si," *Journal of Applied Physics*, **58**, No. 12, pp. 4614-4617, (1985).

[17] T. Kunikiyo, M. Takenaka, Y. Kamakura, M. Yamaji, H. Mizuno, M. Murifuji, K. Taniguchi, and C. Hamaguchi, "A Monte Carlo simulation of anisotropic electron transport in silicon including full band structure and anisotropic impact-ionization model," *Journal of Applied Physics*, **75**, No. 1, pp. 297-312, (1994).

[18] W. Wu, A. Hawkins, and J. E. Bowers, "Design of Silicon Hetero-Interface Photodetectors," *Journal of Lightwave Technology*, **15**, No. 8, pp. 1608-1615, (1997).

[19] W. Wu, A. Hawkins, and J. Bowers, "Frequency Response of Avalanche Photodetectors with Separate Absorption and Multiplication Layers," *Journal of Lightwave Technology*, **14**, No. 12, pp. 2778-2785, (1996).

[20] J. C. Campbell, W. S. Holden, G. J. Qua, and A. G. Dentai, "Frequency Response of InP/InGaAsP/InGaAs Avalanche Photodiodes with Separate Absorption "Grading" and Multiplication Regions," *IEEE Journal of Quantum Electronics*, **QE-21**, pp. 1743-1746, (1985).

[21] I. M. Naqvi, "Effects of time dependence of multiplication process on avalanche noise," *Solid-State Electron.*ics, **16**, pp. 19-28, (1973).

[22] R. B. Emmons, "Avalanche-Photodiode Frequency Response," *Journal of Applied Physics*, **38**, No. 9, pp. 3705-3714, (1967)

[23] J. N. Hollenhorst, "Frequency Response Theory for Multilayer Photodiodes," *Journal of Lightwave Technology*, **8**, pp. 531-537, (1990).

[24] G. Kahraman, B. E. A. Saleh, W. L. Sargeant, and M. C. Teich, "Time and frequency response of avalanche photodiodes with arbitrary structure," *IEEE Transactions on Electronic Devices*, **39**, pp. 553-560, (1992).

[25] J. C. Campbell, W. T. Tsang, F. J. Qua, and J. E. Bowers, "InP/InGaAsP/InGaAs avalanche photodiodes with 70 GHz gain-bandwidth product," *Applied Physics Letters*, **51**, No. 18, pp. 1454-1456, (1987).

Chapter 6

Wafer Fusion

6.0 Introduction

Up to this point, we have explained in depth the advantages of using silicon and InGaAs together in an avalanche photodiode, but have paid little attention to how such a diode could be made. We now turn our attention to this rather challenging problem. A detector made from these materials will require some stringent characteristics. First, current must readily pass through the two materials and through any interface between them. If this were not necessary, glue or epoxy could be used to integrate the materials together. A second requirement is that we must be able to apply an electric field across both the InGaAs and the silicon. The electric field dictates carrier movement when bias is applied. We would also like to control the electric field profile and maintain low fields in the InGaAs layer and high fields in the silicon layer. Without this requirement, we might conceive of simply connecting the two materials with a metallic solder layer.

Yet another requirement, especially for an avalanche photodiode, is the need for semiconductor layers without defects. High electric fields present in these diodes break down along any defects and cause increases in dark current.

With too many defects the dark currents become so high that the detectors are unusable. This low defect requirement excludes epitaxial growth as a possibility for combining InGaAs with silicon, because of the lattice mismatch between the two semiconductors. Crystal growers who have tried to grow compound semiconductors on silicon have found that beyond a "critical thickness" the strain created by the two crystals attempting to "fit" together creates threading dislocations that propagate through an epitaxial layer. The critical thickness varies for each pair of semiconductors[1], but for InGaAs and silicon should be less than 10 nm. Forming threading dislocations relieves the strain caused by the growth, but create just the type of electrical defect that destroys photodetector performance. Despite the problems associated with such epitaxial growth, some groups have tried to create PIN detectors using an InGaAs layer grown on silicon.[2] The dark currents in these devices are very high, however, even at very low electric fields.

A final requirement for effectively combining InGaAs and silicon is that any interface between the two must be of high quality. This means it must be robust enough to stand up to further device fabrication steps such as wet etching, plasma etching, and dielectric deposition. The interface must also freely pass current between the materials and not add to dark current through defects that exist only at the interface. The interface must also have no voids or regions where there is no contact between the two semiconductors. To achieve

such an interface would seem to require a chemical bond between the semiconductors, strong enough to bind them together physically and to allow the sharing of electrons for current transfer. Simple Van der Waals bonding, another integration technique, would in this respect fall short.

Until recently, it might have seemed impossible to effectively combine two such disparate materials as silicon and InGaAs while maintaining high quality crystal layers. That has changed, however, with the development of a technique known as wafer fusion[3]. The process, which will be described below, yields high quality material layers and interfaces. Previously wafer fusion has allowed the integration of materials such as InP and GaAs[4,5], InP and GaN[6], AlGaInP and GaP[7], and InP and silicon[8,9]. This technique has also now been used in the direct integration of InGaAs and silicon, paving the way for the creation of an InGaAs-silicon APD. In this chapter, some of the results of InGaAs to silicon fusion will be shown as well as some first looks at the InGaAs-silicon interface.

6.1 The Process

The wafer fusion process can be summarized in four stages. The first of these stages is the epitaxial growth of bonding layers. For example, in InGaAs to silicon fusion, lattice matching InGaAs layers are grown on InP substrates. Successful fusions have been done using both MBE and MOCVD grown III-V

layers. The growth substrate, in this case InP, is usually removed after the fusion process leaving only the grown layer attached to another wafer. The silicon substrate can also have an epitaxial layer grown on it as well, although this is not necessary. Other examples of epitaxial layers grown for fusion are AlGaAs/GaAs layers grown on a GaAs substrate later fused to an InP substrate. Sophisticated wafer growths may be used to provide features such as etch stops or waveguides.

After the correct epitaxial growths are completed and the host substrates chosen, the next step towards fusion is the careful preparation of the wafer surfaces. As one might imagine, any organic residues or particles on the wafers will prevent the semiconductor atoms from coming into contact with each other. These and other particles must be carefully removed either through wet etches or plasma cleans. Only after wafers are free from particles are they ready to be placed in contact. This is done as one polished surface is placed atop another. If the surfaces have been properly prepared, Van der Waals bonding will occur joining the two substrates. This direct bond must now be converted to a chemical bond through the fusion process.

Wafer fusion of InGaAs to silicon is accomplished by placing the two directly bonded wafers (InP with an InGaAs epitaxial layer, and silicon) under high pressures of several hundred MPa. The wafers are then placed in a hydrogen rich environment and the temperature is raised to 650°C, near the

epitaxial growth temperature of the InGaAs. The wafers are kept at this elevated temperature for 20 minutes and then returned to room temperature. The interface between the InGaAs and silicon has then become chemically bonded and the crystal lattices have attempted to align themselves as illustrated by TEM (tunneling electron microscope) scans. Effects of fusing at different temperatures and in different atmospheres are discussed in Sections 7.4-7.6.

The completion of the fusion process has created a chemical bond between the InGaAs and silicon that is very robust and can withstand severe

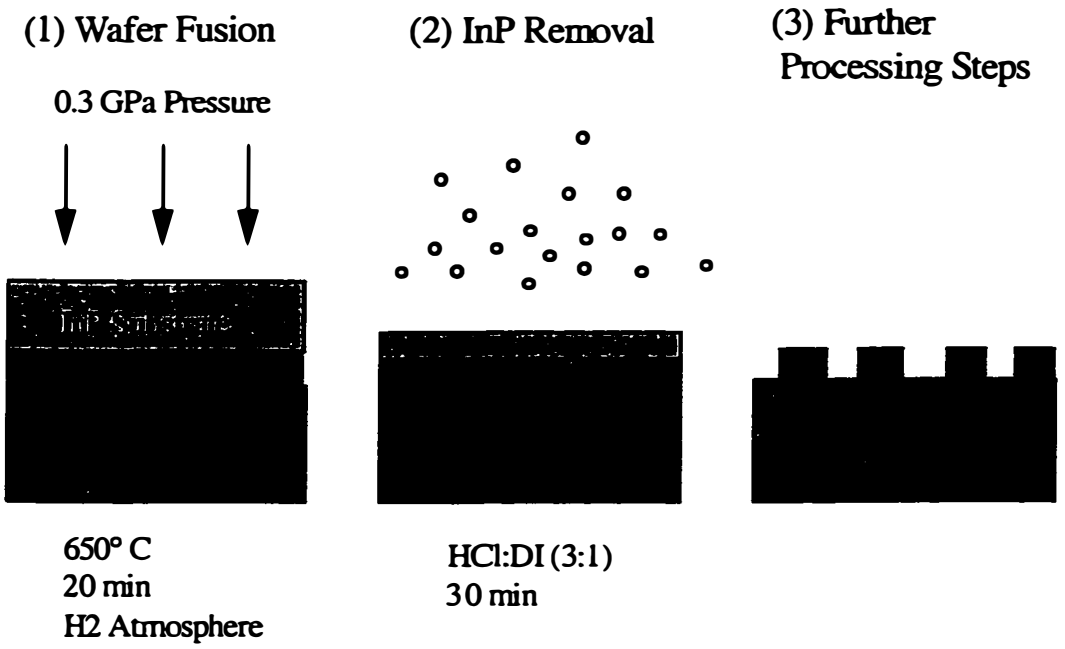


Figure 6.1 - Elementary look at the use of wafer fusion to create an InGaAs-silicon hybrid substrate for the creation of InGaAs-silicon APDs. Fusion occurs at high

temperatures and pressures in a H_2 atmosphere. The InP substrate is then removed from the InGaAs layer initially grown on it, leaving material layers for creating an APD.

processing steps. In order to make the new structure useful, the first of these process steps is the removal of the InP substrate on which the InGaAs was grown. This is done using hydrochloric acid that vigorously etches InP but leaves InGaAs untouched. When the InP is gone, we are left with what appears to be a smooth epitaxial layer of InGaAs on a silicon substrate. Further device processing can proceed from this point, essentially treating the InGaAs as if it had been grown on the silicon. Figure 6.1 illustrates the steps involved going from the fusion process to a silicon wafer with a new material layer attached to it.

6.2 Results - SEMs and TEMs

The development of a wafer fusion process for a specific material system often requires a great deal of trial and error. It is important to have a measure of the extent of the bonding between two materials, or if there is any bonding at all. The first method for evaluating bonds is the SEM (scanning electron microscope) scan. Such scans are made when the wafer fusion process is complete and either before or after a substrate has been etched away. SEM scans can even be used to evaluate the bonding of wafers without an epitaxial

layer, for example an InP wafer fused to a silicon wafer. The first step in making a scan is to cleave the substrate along a crystallographic axis to produce a flat surface on which to perform a side-view scan. The bonded wafers can be placed in the SEM at this point, but it is often difficult to distinguish one material from another, especially for two well bonded layers. Usually a better approach is a type of "stain etch" that will distinguish one material from another. The wafer is placed in a chemical solution that etches one of the semiconductors very slightly, producing an indention at the interface that is much easier to see in the SEM. This method is the most effective for very thin layers sandwiched in between layers that will not etch in the "stain etch"

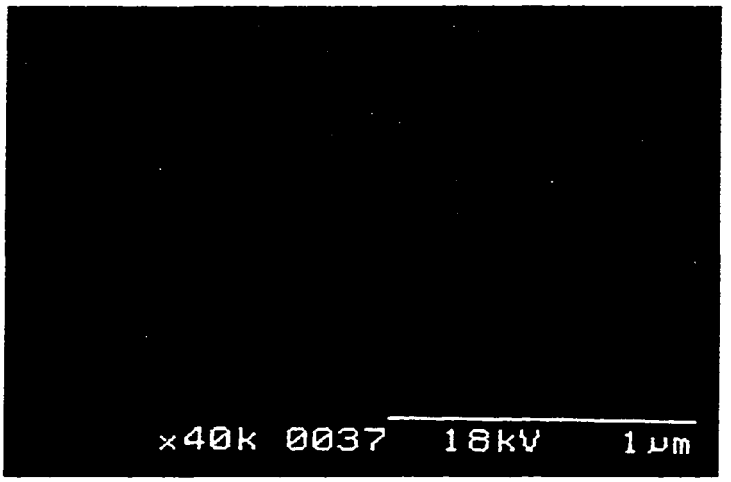


Figure 6.2 - SEM micrograph of the junction for fused InGaAs and silicon. The InGaAs was subjected to a slight stain etch and the material interface is visible as a brightness change in the middle of the photo.

solution. This "stain etch" method also helps to accentuate any fusion gaps in the wafers. Any weakly or partially bonded interfaces will quickly be attacked by the etchant and show up as black voids. Figure 6.2 shows an SEM of an InGaAs and silicon fused interface. The InGaAs has been subjected to a slight stain etch. The interface between the materials can be faintly made out, the InGaAs appearing lighter and the silicon darker in the scan. This is of course an example of a successful fusion run without any visible voids or irregularities at the interface.

A more difficult, but probably more important evaluation tool for wafer fusion is the TEM. These pictures allow views of the crystallographic interfaces on an atomic scale to provide a close up look at the interactions taking place. The TEMs shown in this chapter are cross sections of interfaces made by slicing across an interface and then thinning the cross section. Figure 6.3 shows a TEM view of a wafer created by attempting to epitaxially grow GaAs on silicon[10]. This picture was included to illustrate the type of interface unacceptable for the creation of a device. The first feature to notice in the figure is the white amorphous looking layer right at the interface. Crystal structure breaks down as GaAs tries and fails to conform to the silicon lattice constant. Above this amorphous layer are threading dislocations that propagate diagonally toward the GaAs surface. These results are typical for III-V layers grown on silicon and the defects created would be disastrous for an APD.

In contrast to the TEM produced by the epitaxial growth of GaAs on silicon, Figure 6.4 shows a scan of an InGaAs layer fused to silicon (TEM scan made at UCSB by Ryan Naomi). The first important difference is the lack of an amorphous layer at the material interface. Instead there is a continuation of the crystal lattice right through the interface suggesting chemical bonding between the materials. There is also the absence of any threading dislocations that were so prominent in Figure 6.3. What are present are edge dislocations sitting at the interface. These occur in regularly spaced intervals and accommodate the

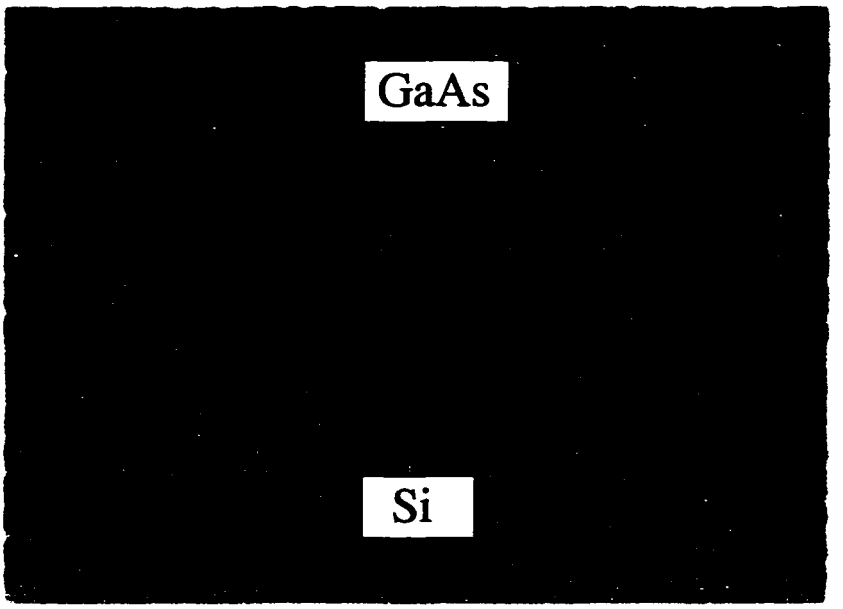


Figure 6.3 - TEM view of a wafer created by attempting to epitaxially grow GaAs on silicon[10]. An amorphous interface is visible along with threading dislocations in the GaAs layer.

lattice mismatch between the materials. Figure 6.5 more closely illustrates the fused interface and shows how, at the interface, a silicon atom is periodically "squeezed into" the lattice because of the larger lattice spacing in the InGaAs, leading to the creation of edge dislocations. The lattice constant for silicon is 0.543 nm and 0.587 nm for InGaAs. Based on the lattice constant difference, we expect an edge dislocation periodicity of one per 13.4 lattice planes in InGaAs. The TEMs produced from fused layers are much more promising in the creation of detectors and other devices from a defect standpoint. Edge dislocations isolated at the interface will not lead to the breakdown problems created by threading dislocations.

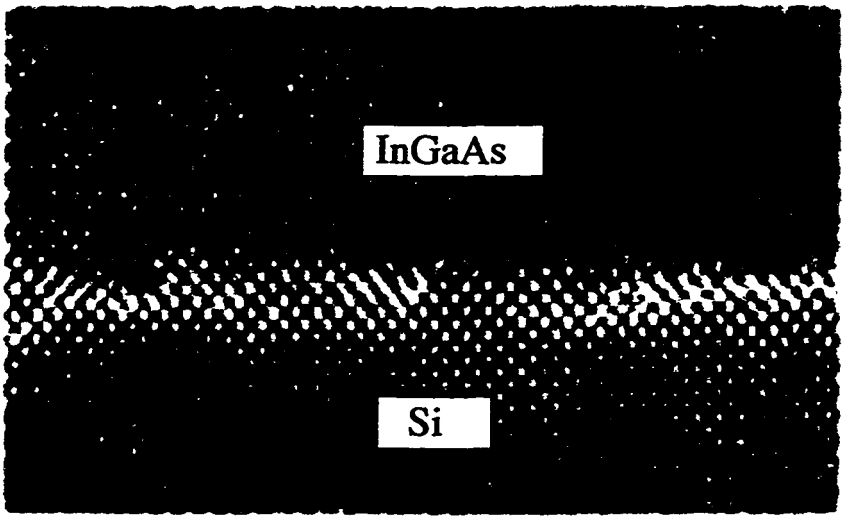


Figure 6.4 - TEM scan of the fused interface between InGaAs and silicon. The transition between the two semiconductors occurs within the light shaded band in the middle of the photo.

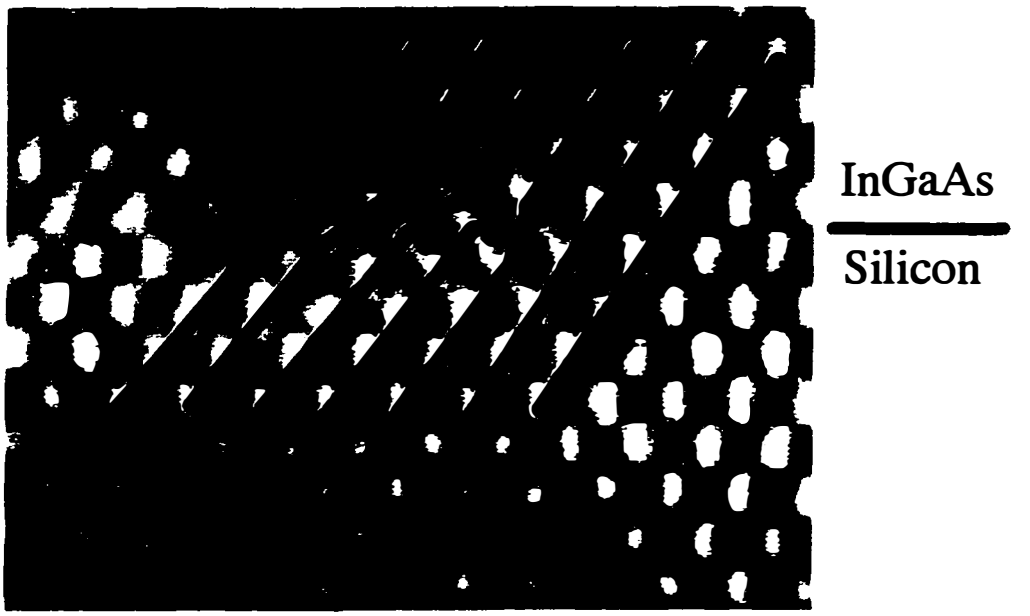


Figure 6.5 - Close-up view of the TEM scan shown in Figure 6.4. Dark lines were superimposed on the photo to indicate semiconductor lattice planes. A misfit dislocation at the InGaAs-silicon transition is illustrated by the termination of one of the lattice planes in the silicon.

6.3 SIMS profile

Secondary ion mass spectroscopy or SIMS is another way to evaluate the state of the fused interface. In this technique, ions are sputtered from a sample through different material layers. Ions of different atomic masses can then be counted to determine what species are present in a sample at a given depth. Figure 6.6 shows a SIMS scan of a fused structure consisting of a 500 nm InGaAs layer attached to a silicon substrate. In this scan we are mostly

interested in what residual species exist at the interface. The species shown in the plot are oxygen, hydrogen, and carbon, chosen because these three would make up any organic residues present on the surface even after a cleaning procedure. The important part of the graph then is at the 500 nm point, the

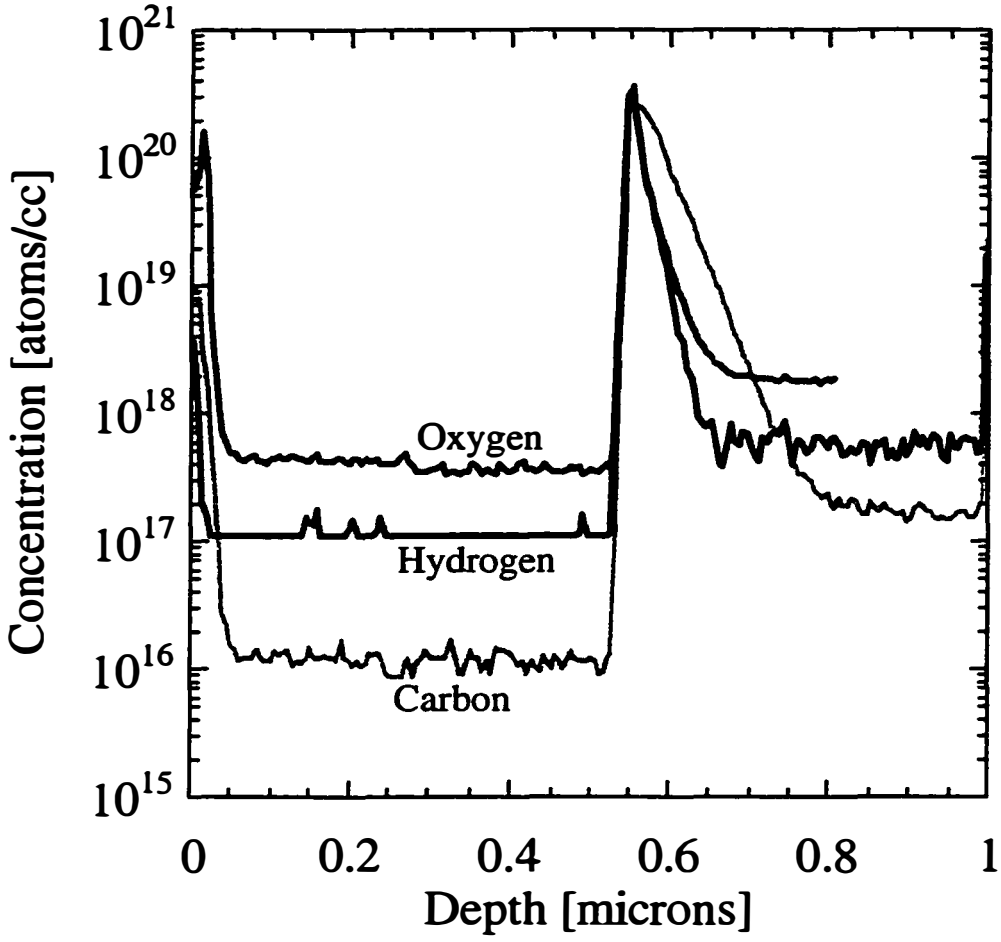


Figure 6.6 - SIMS scan through an InGaAs-silicon fused interface. The interface occurs in the graph at about 0.53 μm, where the large jump in the oxygen, hydrogen, and carbon concentration occurs.

transition from InGaAs to silicon. The three species in question are below the resolvable limit until this transition and then all of them rise dramatically to a peak right at the fusion interface. This peak then tails off as we sputter deeper and deeper into the silicon. The remarkable thing about this plot is that despite the high hydrocarbon count at the interface, there was still obtain good adhesion of the InGaAs and silicon. The bond between the materials is not only strong, but as will be shown in later chapters, even with the high carbon count the junction is very electrically conductive.

6.4 A Robust Bond

One of the important requirements for the wafer fusion process was the production of a bond between InGaAs and silicon that was strong enough to withstand device processing. Some of the processing steps required for the production of a detector are chemical etching, high temperature film depositions, plasma etching, and contact lithography. Of these processes, the most difficult for a weak bond to withstand is chemical etching. In the particular case of InGaAs to silicon bonding, one of the more violent processes is the InP etch used to remove the InP substrate. This etch is very fast (about 25 μm per minute) and vigorous, producing phosphine gas bubbles that float to the surface of the etch solution. Any areas in which fusion did not take place

between two substrates is quickly revealed in this etch because the unfused regions simply peel off under the stress.

For areas where InGaAs-silicon fusion did occur, one of the best demonstrations of good bonding is the etching of either the InGaAs or the silicon in a mesa structure. Poorly bonded regions are quickly undercut and peeled away by such an etch. Figure 6.7 shows a mesa structure created from a chemical etch to demonstrate the fusion bond strength. The figure is an SEM scan of a silicon wafer that had been fused to an InGaAs and InP layer. After the removal of a much thicker InP substrate, the mesa structure seen was created through wet chemical etching and lithography. The top InP layer was etched in an $HCl:H_2O$ solution. The InGaAs layer was etched in an $H_3PO_4:H_2O_2:H_2O$ solution, while the silicon was etched in a hot $KOH:H_2O$ solution. If the bond between the silicon and the InGaAs were anything but a robust chemical bond, the $H_3PO_4:H_2O_2:H_2O$ and $KOH:H_2O$ would be especially damaging to it. The even etching shown in Figure 6.7 would not be possible, and areas of weak bonding would show tremendous undercutting of the mesa structure.

The structure tests and scans made of the InGaAs and silicon fused structure all indicate a very promising interface. The bonds between the materials are very robust and lattice defects are limited to edge dislocations at the interface. Fusion is also possible over large areas even with less than

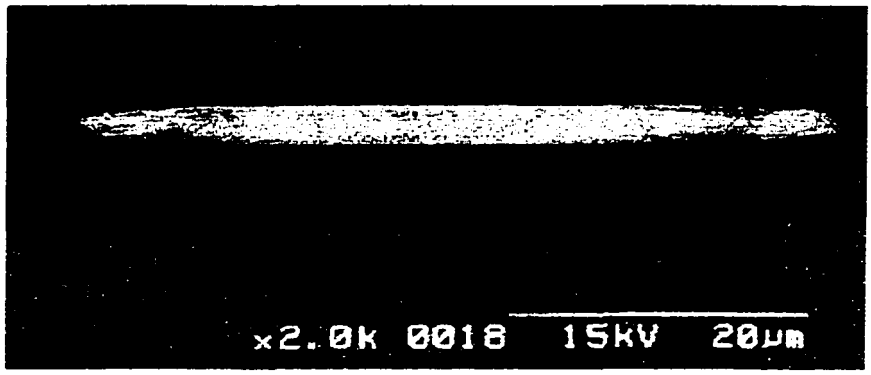


Figure 6.7 - Micrograph of a mesa structure created from InGaAs and InP epitaxial layers fused to a silicon substrate. The mesa structure was created first by wet etching the InP layer with a hydrochloric acid solution, then etching the InGaAs layer with a phosphoric acid/hydrogen peroxide solution, and finally etching into the silicon substrate with a hot potassium hydroxide solution.

perfect surfaces as SIMS scans show. The real test of the utility of the fusion process, however, comes only with the creation of detector structures. The electrical characteristics of the fusion interface remain the deciding factors in the feasibility of creating a fused APD.

References

[1] J. W. Matthews, "Coherent interfaces and misfit dislocations," in Epitaxial Growth (edited by J. W. Matthews) Part B, New York, Academic, (1975).

[2] A. M. Joshi, R. Brown, "Monolithic InGaAs-on-silicon short wave infrared detector arrays," SPIE Photonics West '97, Conference 2999, San Jose, CA.

[3] A. Black, A. R. Hawkins, N. M. Margalit, D. I. Babic, A. L. Holmes, Jr., Y. -L. Chang, P. Abraham, J. E. Bowers, and E. L. Hu, "Wafer Fusion: Materials Issues and Device Results," *IEEE Journal of Selected Topics in Quantum Electronics*, **3**, No. 3, pp. 943-951, (1997).

[4] N. M. Margalit, D. I. Babic, K. Streubel, R. P. Mirin, R. L. Naone, J. E. Bowers, and E. L. Hu, "Submilliamp Long-Wavelength Vertical Cavity Lasers," *Electronics Letters*, **32** (18), 1675-1677, August 29 (1996).

[5] R. J. Ram, J. J. Dudley, and J. E. Bowers, "GaAs to InP Wafer Fusion,"*Journal of Applied Physics*, **78**, No. 6, pp. 4227-4237, (1995).

[6] R. K. Sink, S. Keller, B. P. Keller, D. I. Babic, A. L. Holmes, D. Kapolnek, S. P. DenBaars, J. E. Bowers, X. H. Wu, and J. S. Speck, "Cleaved Facets in GaN by Wafer Fusion of GaN to InP," *Applied Physics Letters*, **68** (15), pp. 2147-2149, (1996).

[7] F. A. Kish, D. A. DeFevere, D. A. Vanderwater, G. R. Trott, R. J. Weiss, and J. S. Major, Jr., "High luminous flux semiconductor wafer-bonded AlGaInP/GaP large-area emitters," *Electronics Letters*, **30** (21), 1790-1792, (1994).

[8] H. Wada, T. Takamori, and T. Kamijoh, "Room-Temperature Photo-Pumped Operation of 1.58-μm Vertical-Cavity Lasers Fabricated on Si Substrates Using Wafer Bonding," *IEEE Photonics Technology Letters*, **8** (11), 1426-1428, (1996).

[9] Y. H. Lo, R. Bhat, D. M. Hwang, C. Chua, and C. -H. Lin, "Semiconductor lasers on Si substrates using the technology of bonding by atomic rearrangement," *Applied Physics Letters*, **62**, No. 10, pp. 1038-1040, (1993).

[10] M. Akiyama, et al., Journal of Crystal Growth, **68**, 21 (1984).

Chapter 7

PIN Devices

7.0 Introduction

The complexity in the design of an avalanche photodetector makes it a difficult device to study. It is often very hard to determine what part of the structure is causing a given characteristic. For example, there are many factors that can lead to high dark currents in a detector, such as material defects, poor device design, or unpassivated surfaces. The InGaAs-silicon APD presents a particularly challenging device to study due to the unknowns about the fused interface. This interface can have dramatic effects on device performance, but it is difficult to isolate its impact in a complicated structure. A much better detector structure with which to study fused InGaAs and silicon is a PIN. PIN refers to the P type, I or intrinsic, and N type dopings in such a detector. With only three device layers, compared to an APD, fabricating a PIN detector is very straightforward.

This chapter will first outline the device design and fabrication steps used in the creation of PIN detectors[1]. Detector measurements are then shown and conclusions made about the quality of the fused interface. A more detailed study of the different fusion conditions is then described which utilizes the PIN

structure and repeats the same measurements, allowing for even more conclusions about the fused interface and how to optimize its properties. The work described in this chapter was done in collaboration with Lucent Technologies

7.1 Fabricating an InGaAs-Silicon PIN

The PIN design was chosen to isolate the fused InGaAs-silicon interface as much as possible, using known fabrication techniques. Mesa etching was avoided to eliminate any uncertainty caused by the damage such etching can create, most often leading to an increase in dark current and premature voltage breakdowns. A planar device is a better alternative. With this in mind, fabrication proceeded as follows. First a 0.5 μm layer of undoped silicon was epitaxially grown on an N+ silicon substrate. III-V layers were then grown on an InP N+ substrate through MOCVD (metal organic chemical vapor deposition). The layers were first a 0.3 μm $In_{0.53}Ga_{0.47}As$ stop etch layer, then a 0.6 μm InP layer, followed by a 1.0 μm $In_{0.53}Ga_{0.47}As$ layer. All of the layers were grown without doping, but were slightly n type due to the non-intentional doping of the growth chamber. The silicon and III-V wafers were then carefully cleaned in preparation for the fusion process described in Chapter 6. The two wafers were then placed in contact and heated under pressure in a hydrogen environment. A temperature of 650°C, near the InGaAs growth temperature,

was maintained for 20 minutes. The wafers were then cooled and removed from the hydrogen atmosphere. The InP substrate was removed through wet etching using an $HCl:H_2O$ solution. Etching cleanly stopped on the 0.3 μm InGaAs layer. This etch stop layer was subsequently removed with an $H_3PO_4:H_2O_2:H_2O$ solution, stopping on the 0.6 μm InP layer.

With the necessary epitaxial layers in place, PIN detector fabrication proceeded. The first step was the deposition of a 150 nm SiO_2 layer on top of the remaining InP layer. This SiO_2 layer served as a diffusion mask for the next important step - zinc diffusion. In order to create a p-type layer in the InGaAs while maintaining a planar structure, the dopant had to be diffused in after the fusion process. This of course allows for the isolation of p- type regions, which in effect define individual PIN detectors. To accomplish the zinc diffusion, 22 μm diameter windows were first etched into the SiO_2 layer to the InP layer beneath. Zinc was then diffused through the 0.6 μm InP layer and into the InGaAs layer at a temperature of 550°C for 45 minutes. The total zinc diffusion depth was 1.0 μm, or 0.6 μm in the InP and 0.4 μm into the InGaAs. This left an undoped InGaAs layer 0.6 μm thick beneath the P+ InP and InGaAs regions.

Upon completion of the zinc diffusion, another SiO_2 layer was deposited on the top of the structure. 12 μm diameter holes were opened in the SiO_2 over the top of the original 22 μm diameter zinc diffusion windows. These holes were then etched deeper through the InP layer and stopping on the heavily

doped InGaAs layer beneath. This etch allowed for the deposition of a metal contact to the P+ layer. A 30 μm diameter Au:Be metal layer was then deposited using a metal liftoff technique, the gold extending from the P+ InGaAs layer to on top of the SiO_2 layer and served as the p contact metal. Contact then had to be made to the N+ silicon substrate. To do this a 70 μm diameter mesa was formed centered over the original zinc diffused window. The top SiO_2 layer was first etched followed by the InP and InGaAs layers below. Finally the silicon was etched using an RIE (reactive ion etcher) to a depth greater than the 0.5 μm undoped silicon thickness. Aluminum was then deposited over the surface of the etched N+ silicon to form the n contact to the device. The structure is illustrated in Figure 7.1. As can be seen from the

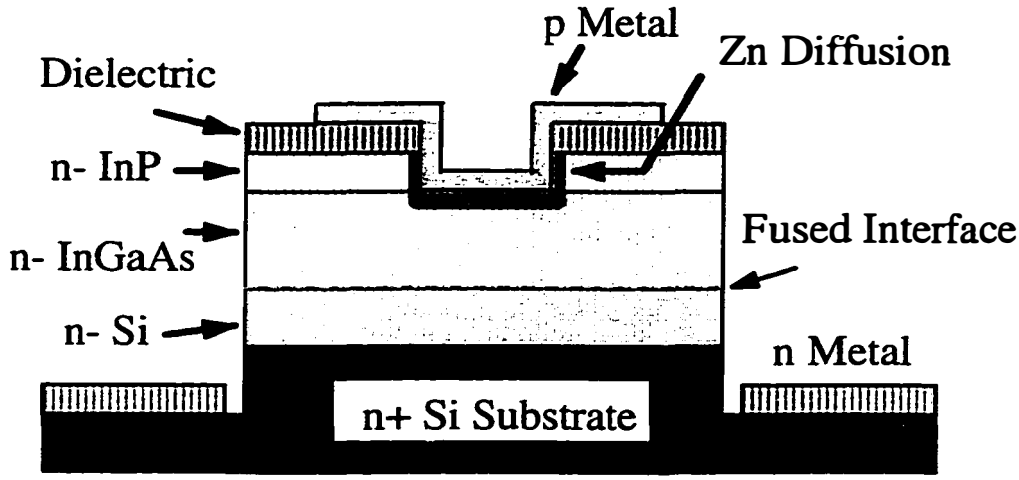


Figure 7.1 - Cross section of an InGaAs-silicon PIN detector.

figure, bias between the n and p metals will be applied through the fused interface. The interface is isolated from the mesa etched used to contact the N+ silicon by the high resistivity n- InGaAs and n- silicon layers.

For comparison, PIN devices of a similar structure to Figure 7.1 were made simply using InGaAs and InP. A 1.0 μm InGaAs layer was grown on an N+ InP substrate followed by a 0.6 μm InP layer. Device fabrication was then the same as that used for the InGaAs-silicon PINs except that gold-germanium instead of aluminum was used as the contact for the N+ InP layer.

7.2 Current Versus Voltage and Quantum Efficiency

The first measurements made on the fused InGaAs-silicon PINs were the dark current and photocurrent versus reverse bias. Photocurrent was measured by polishing the backside of the silicon substrate and illuminating the detector through the silicon. The absorption of silicon is not significant at wavelengths above 1.0 μm so either 1.3 or 1.55 μm lasers can be used as the light source. For this measurement, the longer wavelength was used. Figure 7.2 plots the photocurrent and dark current versus bias from 0 to -10V. The first significant part of the graph to notice is the constant photocurrent versus bias, even at zero volts. This points towards little or no charge trapping occurring at the fusion interface. The dark current is also very low for these devices. Below -4V, a dark current of less than 100 pA was measured, very similar to that measured

for the InGaAs/InP PINs made in parallel with the fused PINs. This too indicates a high quality fused interface with no significant dark current increases coming from generation and recombination at the InGaAs-silicon junction. The carrier lifetime can be computed using the dark current

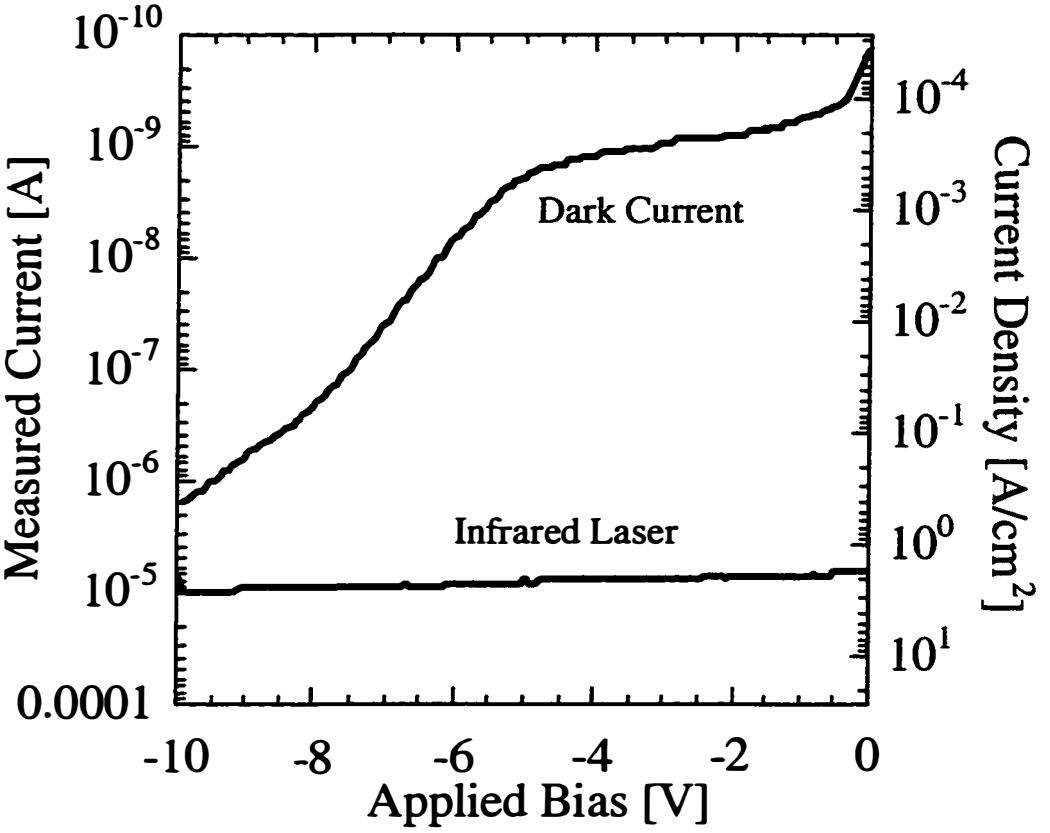


Figure 7.2 - Measured current versus reverse bias for an InGaAs-silicon PIN. The curve labeled "infrared laser" is the photocurrent produced when the PIN was illuminated with a 1.55 μm laser.

measurements and the equation[2]

$$I_D = Aqn_i w / \tau, \qquad 7.1$$

where I_D is the dark current, A the device area, q electric charge, n_i the intrinsic carrier concentration, w the depletion width, and τ the carrier lifetime. A carrier lifetime of 400 ns was calculated based on the fused device parameters. This lifetime is long enough to indicate not only a good fused junction, but also the bulk of the InGaAs depletion layer is practically free of defects.

The forward bias characteristics of the InGaAs-silicon PINs were also measured. In the range between 0.1 and 0.4 V, an ideality factor of n=1.1 to 1.2 was determined using the equation

$$I = I_o \exp(qV / nkT)^2, \qquad 7.2$$

where I was the measured current, I_o a normalizing current, q electric charge, V voltage, n the ideality factor, k the Boltzmann constant, and T the temperature. An ideality factor so close to one is another indicator of an interface free of defects and recombination sights.

Two extremely critical properties of any detector are closely related - the responsivity R and quantum efficiency η. These quantities are so important because they express how well a detector converts optical power into electrical power. The responsivity and its relation to the quantum efficiency can be written as

$$R = \frac{I_p}{P_{in}} = \frac{q\eta}{h\nu} \qquad 7.3$$

where I_p is the measured photocurrent, P_{in} the total power of the light incident on the detector, h Planck's constant, and ν the photon frequency. To express responsivity in terms of wavelength, λ, we may use the relation hν=1.24/λ eV to write

$$R = \frac{\eta\lambda}{1.24}. \qquad 7.4$$

An expression can also be written for the quantum efficiency which expresses it in terms of absorption efficiency η_a and an internal quantum efficiency η_i which represents how well carriers are conveyed through a device. Especially interesting is the internal quantum efficiency because any interface trapping of carriers will result in an η_a of less than 100%. These quantities can be expressed as

$$\eta = \eta_a \eta_i \qquad 7.5$$

where $n_a = 1 - e^{-\alpha L}$, α is the optical absorption coefficient, and L is the length of the absorption region. Using a 1.55 μm laser source, calibrated using the InGaAs/InP diodes and a power meter, and correcting for the reflection off the polished silicon surface, the responsivity was measured for the fused PIN at a reverse bias of 4 V. A value of R=0.38 A/W was obtained, translating to η=0.30, and η_a=100%. The accuracy of the experiment was approximately +/-

5%. These results are very significant and indicate no measurable loss of carriers across the fused interface. The high quantum efficiency bodes very well for any InGaAs-silicon detector. Any benefits derived from an InGaAs-silicon combination in an APD would have been undermined if a high quantum efficiency were not possible.

7.3 Capacitance and Bandwidth

The capacitance of the InGaAs-silicon PINs was also measured to determine the depletion layer doping and whether any excess charge is present in the structure. The measurement was done for voltages between 0 and -10 V using an HP Lightwave Component Analyzer at a frequency of 1 GHz. The measured capacitance decreased with increasing reverse bias without any slope changes, which are indicative of excess charges in the structure. If the fused interface had any significant charge buildup larger than the background doping, the capacitance curve would have a kink in it. In fact the structure behaves very much like a standard PIN diode and the capacitance curve is well described by the relation[2]

$$\frac{C(V)}{A} = \left[\frac{q\varepsilon N}{2(V+V_{bi})}\right]^{1/2} + C_o \qquad 7.6$$

where ε is the dielectric constant, N the charge density, V_{bi} the built in voltage, and C_o the stray and pad capacitance, measured to be 50 fF. Figure 7.3 shows a plot of $C(V)^{-2}$ versus V, which should have a slope given by Equation 7.6 of

$\frac{2}{q\varepsilon NA^2}$ and an intercept of $\frac{2V_{bi}}{q\varepsilon NA^2}$. As can be seen from the figure, the slope is very straight and yields a value for N = 1.4×10^{16} cm^{-3} and V_{bi} = -0.6V. The doping value is consistent with resistivity and SIMS measurements made on the silicon and InGaAs layers. The total capacitance of 100 fF at -10 V is also consistent with a calculated value given the device area and depletion depths.

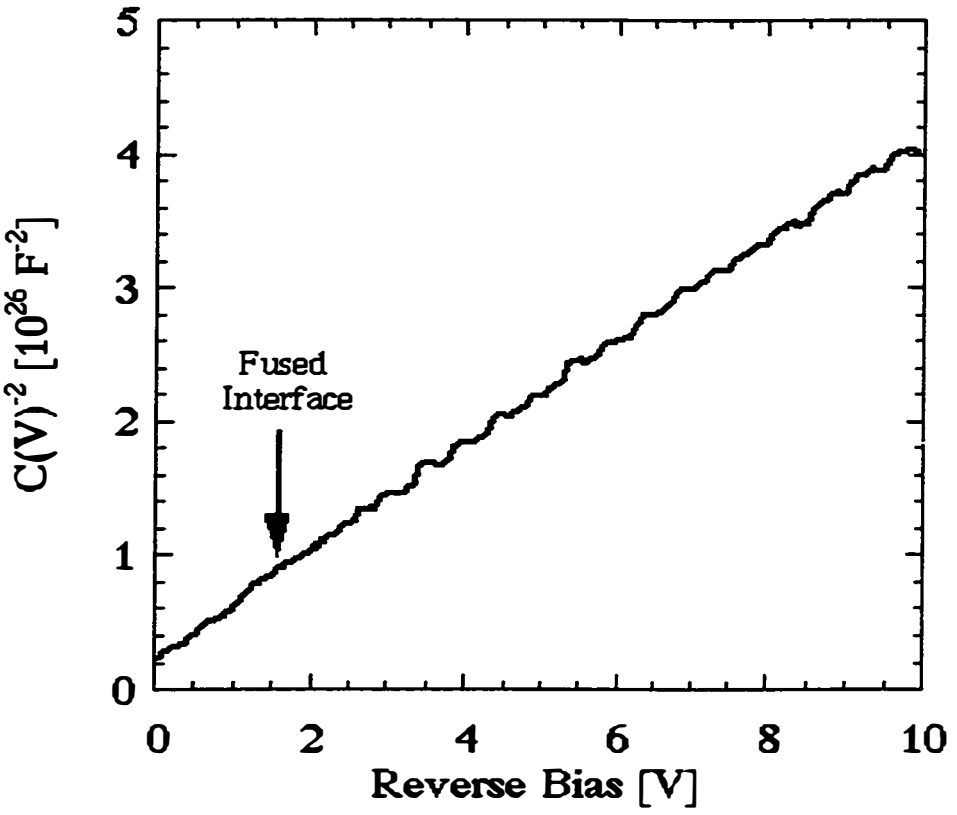


Figure 7.3 - Inverse square of the measured capacitance versus reverse bias for the InGaAs-silicon PIN. The fused interface should be depleted through by 1.7 V based on doping density.

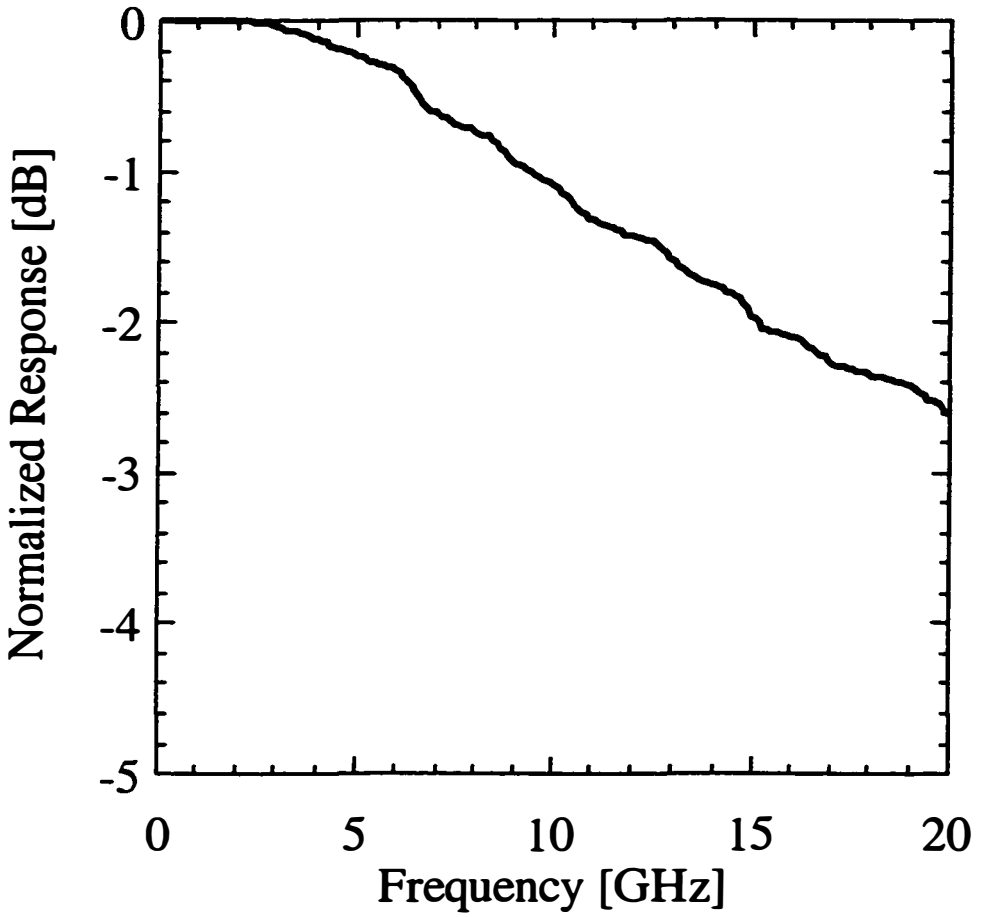


Figure 7.4 - Frequency response of the InGaAs-silicon PIN. The measured 3-dB down frequency was 21 GHz.

Frequency response and 3dB bandwidth were measured for the InGaAs-silicon PINs also using an HP Lightwave Component Analyzer with a range from 130 MHz to 20 GHz. A high-speed coplanar probe was used to contact the detectors and voltage was applied through a high speed bias tee. Figure 7.4

shows the frequency response of the detector at a reverse bias of 10 V. The 3 dB down point is about 22 GHz. RC effects should limit the bandwidth. The contact resistance measured 40 Ω and combined with the 50 Ω resistance of the line gives a total circuit resistance of 90 Ω. The capacitance was 100fF so the frequency response limit is $1/(2\pi RC)$ = 18 GHz, corresponding very closely to the measured 3 dB frequency. Even at zero bias the experimental frequency response corresponded well with calculated RC effects. The smooth roll-off of the response curve is another indicator of an interface relatively free of carrier traps, unless such traps had extremely short lifetimes and were beyond the 20 GHz instrument limit. Fast carrier transport is another important ingredient in the making of a successful APD for use at high speeds, and it appears the fused InGaAs-silicon interface will allow for operation at frequencies even above the 10 GHz bandwidths currently being planned for.

7.4 Varying the Fusion Temperature and Atmosphere

The results described in the previous section for the InGaAs-silicon PIN indicate that the wafer fusion process used to construct the detector produced a nearly ideal interface between the III-V layer and silicon. The interface was so invisible to carriers, the InGaAs-silicon PINs were very similar in their characteristics to InGaAs/InP PINs without a fused interface. The fusion process as described in Chapter 6, included a 650°C heating step in hydrogen.

Using the data compiled on the PINs from the first fabrication as a baseline, a study was done to investigate the effects of different fusion parameters on device characteristics[3], especially the dark and photocurrents in forward and reverse bias. Changes were made to both fusion temperature and the fusion atmosphere. The motivation for this study was to hopefully learn something new about the fusion process as well as explore possible process limits. It is usually desirable to do any processing at as low a temperature as possible to avoid any unwanted diffusion of dopants or contaminants, so lower temperature fusion is an attractive possibility. If the temperature could be reduced low enough, applications such as fusion to already fabricated VLSI become an interesting proposition. A change in the fusion atmosphere is probably most desirable from a safety standpoint. If hydrogen could be replaced

Table 2 - Varied Parameters for five different fusion experiments.

Fusion Run	Temperature	Atmosphere
1	650	H_2
2	550	H_2
3	475	H_2
4	650	10% H_2, 90% N_2
5	650	N_2

by a more inert gas, the fusion process would be much easier to make manufacturable. Table 2 outlines five fusion process runs completed in this study, including the temperature and atmosphere used.

All of the runs used the same silicon substrate that had been cleaved from a 6" wafer into smaller pieces. The wafer consisted of a 0.5 μm undoped layer epitaxially grown on an N+ substrate. Run 1 was simply the PIN structure already described which used the III-V epitaxial growth outlined in Section 7.1. Runs 2 through 5 used a III-V wafer that began with a N+ InP substrate and then using MOCVD growth consisted of the following layers: first a 0.3 μm $In_{0.53}Ga_{0.47}As$ stop etch layer, followed by a 0.3 μm InP layer, and a top layer of 1.0 μm $In_{0.53}Ga_{0.47}As$. All of the epitaxial layers were grown undoped, but were unintentionally doped slightly n-type due to the background doping in the growth chamber. Subsequent to the wafer fusion process, the InP substrate, and InGaAs stop etch layers were removed from all of the bonded wafers. All of the fusion runs 1-5 displayed very good mechanical adhesion of the III-V epitaxial layers to the silicon substrate after the vigorous chemical-etching step.

Device fabrication for Run 1 has been described and the process was very similar for the latter runs. First a 0.15 μm SiO_2 layer was deposited, then 30 μm diameter windows were opened down to the top InP layer. Zinc diffusion was then done through the 0.3 μm InP layer and 0.5 μm into the InGaAs layer. SiO_2 was again deposited and a 20 μm diameter hole aligned to

the first window was opened in the oxide layer to the InP. A 50 μm diameter Au:Be metal contact pad was then deposited over the SiO_2, making contact to the zinc doped InP layer. Contact to the N+ substrate was made by etching a 70 μm diameter mesa structure, centered at the original window, through the III-V layers and below the undoped silicon epitaxial layer. Aluminum was then deposited onto the silicon surface as an n contact metal. The differences between this structure and that described in Section 7.1 is the larger p contact window and zinc diffusion window and the p contact metal was deposited on the p doped InP instead of the InGaAs. The change to the larger window was to decrease contact resistance and the change to the InP contact layer was made to reduce the dark current.

7.5 Results of the Varying Fusion Conditions

The first measurements made on all of the fused PIN devices were the dark current and photocurrent versus voltage measurements. These currents are shown in Figure 7.5 for reverse biases between zero and -10 V. For the samples from Runs 1, 4, and 5 that were fused at the highest temperature of 650°C, the dark current remains very low and is similar for all three runs. The hydrogen atmosphere does seem to be the optimal condition, followed by the hydrogen containing forming gas, and then the nitrogen atmosphere. The dark current curves at different temperatures show a much greater deviation, indicating the

quality of the fusion interface is more dependent on temperature than on atmosphere. The 650°C sample exhibits the lowest dark current, followed by the 550°C then the 475°C sample. The PINs fused at the lowest temperature exhibited substantial dark current increases, nearly two orders of magnitude larger than the optimally fused structure. This result points to defects retained in the interface for the lower temperature processes[3].

The photocurrent versus bias curve also shows an interesting result as shown in Figure 7.5. The optimally fused Run 1 sample has a constant photocurrent value even down to zero volts. The Runs 2-5 at the non-optimal conditions show a slight turn-on voltage for the photocurrent, indicative of a voltage barrier at the InGaAs-silicon interface. The value of the photocurrent was very similar for all of the detectors, however, once the turn-on voltage was overcome. A calculation of quantum efficiency was done using the photocurrent measured when the diodes were illuminated with a 1.55 μm laser source calibrated using InGaAs/InP devices of a similar layer structure. For all the devices measured at reverse biases over .5 V, the internal quantum efficiency, as defined in section 7.2, was measured to be 100%. At zero bias, the optimally fused structure still had an η_i = 100%, while the non-optimally fused Runs 2-5 had an η_i=95%. The 5% efficiency loss indicates that there is some type of barrier at the interface, though it is rather small and does not block many carriers.

The capacitance of the InGaAs-silicon PINs was measured as described in Section 7.3. The capacitance versus reverse voltage curve was similar for all

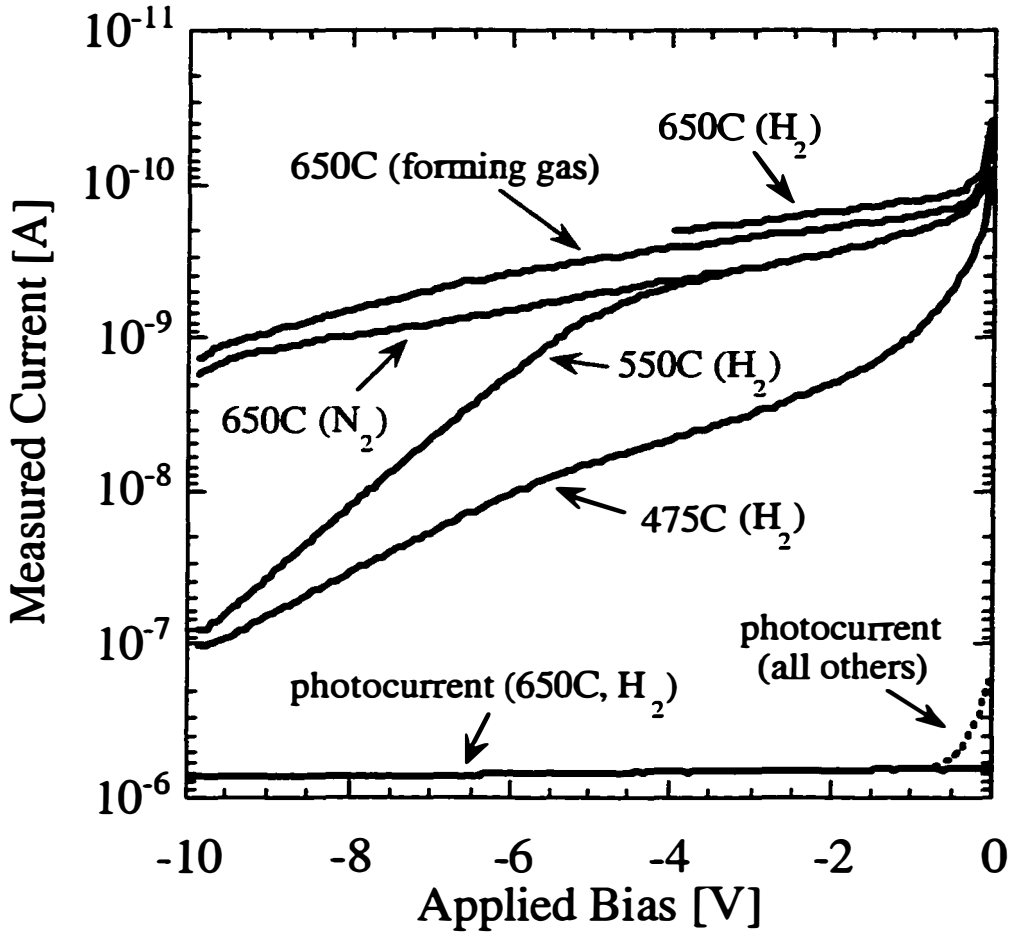


Figure 7.5 - Dark and photocurrents for InGaAs-silicon PINs fabricated under different fusion conditions. The different fusion conditions are indicated on the graph first with a temperature then the atmosphere present in the fusion chamber. The photocurrent for all devices was measured by illuminating with a broadband source.

the detectors from Runs 1-5 with a smooth decrease from 0 to -10 V. The total capacitance was slightly larger for the larger area devices, with a junction capacitance of 82 fF measured at -10V, and a stray and pad capacitance of 80 fF. Using these values and the slope of the C^{-2} versus V curve as described in Section 7.3, the doping density of the material was calculated to be $N=1.5x10^{14}$ cm^{-3}, in excellent agreement with the previous calculation as well as measured doping levels.

The frequency response was expected to be limited mostly by the RC constraints of the PINs. The 3-dB down frequency for fusion Run 1 was already shown to be 22 GHz. The PINs from Runs 2-5 should have a 3-dB limit of $1/(2\pi RC)$ 13.3 GHz (total circuit resistance being 66 Ω). Measurements using a 20 GHz Lightwave Component Analyzer showed 3-dB down frequencies of 13.4 +/- 0.8 GHz for all of the non-optimized detectors when measured at 10 volts reverse bias. At zero bias, there is a radical difference compared to the optimized PINs. The measured bandwidth is only around 100 MHz, while the first PINs had a 6 GHz bandwidth. Again this suggests an interface barrier caused by the non-optimized fusion, trapping carriers at very low biases. The bandwidth at high biases shows that this barrier can be overcome, however, with high-energy carriers.

7.6 An Interface Barrier

The measurements of photocurrent, quantum efficiency, and frequency response indicate there is a barrier present at the InGaAs to silicon interface in non-optimally fused samples. Further evidence for such a barrier comes from

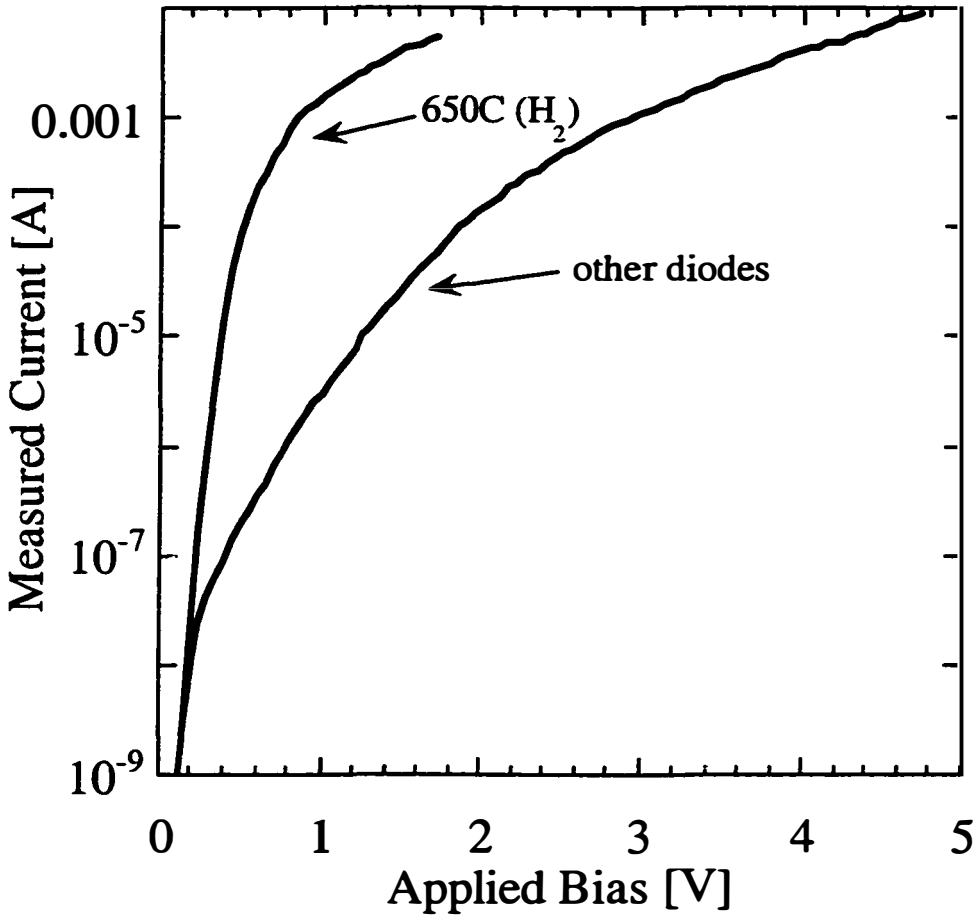


Figure 7.6 - Measured current for forward biased InGaAs-silicon PINs. The diodes fused at 650°C in hydrogen had different forward bias curves than the devices fused under any other conditions - which all had very similar curves.

forward bias versus current information. The sample fused at 650°C in hydrogen had ideality factors in the forward bias near one, indicating the lack of recombination centers. The detectors formed from samples at all other temperatures and atmospheres had much different forward bias curves. As shown in Figure 7.6, the measured current increases much more slowly for all the non-ideal diodes and the curve was essentially the same for Runs 2-5.

Levine[3] conjectured that the interface barrier could be described by a very thin high bandgap material or dielectric layer which required high field tunneling for penetration. This effect would only show up at low energy, low bias conditions as are seen for the non-ideal PINs. If a thin layer of silicon dioxide with a Gaussian thickness distribution is assumed for the barrier, with a bandgap of 9 eV, we can use the equation for tunneling current to fit a layer thickness to measured current in forward bias. The equation for the oxide thickness probability can be written as $p(t)=\exp[-(t-t_{av})^2/\sigma^2]$, where σ is the halfwidth. The tunneling current as a function of thickness t and bias V is described by the equation[2]

$$I(t,V) = AV^2 \exp\left[\frac{-8\pi(2m^*)^{1/2} E_{gb}^{3/2} t}{3qhV}\right] \qquad 7.7$$

where m* is the effective mass, E_{gb} is the bandgap barrier, and A is a constant. The total tunneling current is then express by

$$I_T = \int p(t) I(t,V)\, dt \qquad 7.8.$$

When the measured forward current is fit to these equations, tunneling parameters equal to t_{av}=0.39 nm and σ=0.135 nm are found, with limits of integration from t=0 to t=0.43 nm. This theoretical model gives a good fit to the data over eight orders of magnitude although the exact tunneling parameters are approximations given the effective mass was assumed to be one and E_{gb}= 9 eV. The easy fit of the tunneling model to the data does indicate, however, that a tunneling barrier is likely present. The calculated parameters of a barrier less than 0.5 nm wide are also very believable for this device.

7.7 Conclusions

The study of fused InGaAs-silicon PINs has led to many conclusions that can be perhaps be summed up by the following sentence. Under the proper conditions a fused interface can be created between InGaAs and silicon that is in all respects ideal for the creation of a photodetector. Although a conduction band offset has not been explicitly measured, given the electrical characteristics we can speculate about the nature of the offset between fused InGaAs and silicon. The absence of any type of barrier under ideal fusion conditions indicates there is either no offset in the conduction band or a very small offset. Figure 7.7 illustrates what this offset would look like when combining n-type InGaAs with n-type silicon at an unbiased junction. The figure shows all of the bandgap energy difference accounted for by an offset in the valence band. This

structure would allow for the free transfer of electrons from InGaAs to silicon and holes from silicon to InGaAs - a necessity for an InGaAs-silicon APD.

Measurements of dark currents, photocurrents, quantum efficiencies, capacitances, and bandwidths have shown that an InGaAs-silicon PIN performs just as well as a non-fused InGaAs-InP PIN with a comparable InGaAs layer. This bodes very well for more complicated APD devices, given that performance should not be limited by the fused interface. The ideal conditions for the creation of such an interface are a hydrogen atmosphere at 650°C. In other atmospheres and lower temperatures, the interface formed is less than

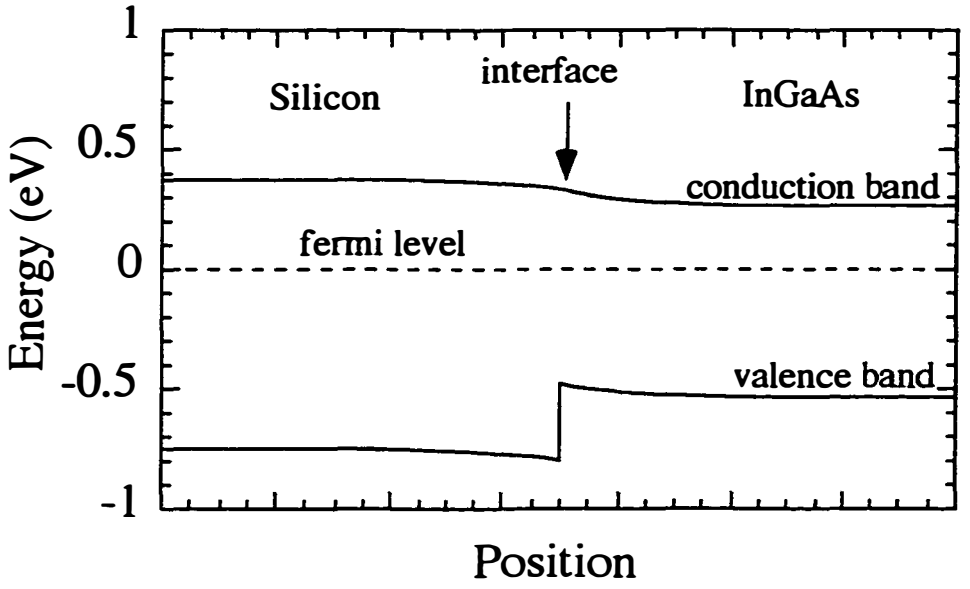


Figure 7.7 - Possible band structure for n-type InGaAs fused to n-type silicon. The diagram is a supposition based on the electrical characteristics of the fused interface.

ideal. This leads to increases in dark currents and a very thin interface barrier. A tunneling model supports the conclusion that carriers tunnel through this thin barrier and that its effects show up only at low carrier energies.

References

[1] B. F. Levine, A. R. Hawkins, S. Hiu, B. J. Tseng, C. A. King, L. A Gruezke, R. W. Johnson, d. R. Zolnowski, and J. E. Bowers, "20 GHz High Performance Si/InGaAs PIN Photodetector," *Applied Physics Letters*, **70**, No. 18, pp. 2449-2451, (1997).

[2] Physics of Semiconductor Devices, S. M. Sze, John Wiley & Sons, New York, (1981).

[3] B. F. Levine, A. R. Hawkins, S. Hiu, B. J. Tseng, J. P. Reilley, C. A. King, L. A Gruezke, R. W. Johnson, D. R. Zolnowski, and J. E. Bowers, "Characterization of Wafer Bonded Photodetectors Fabricated using Various Annealing Temperatures and Ambients," *Applied Physics Letters*, **71**, No. 17, pp. 2508-2510, (1997).

Chapter 8

APD Devices - Fabrication and Measurement

8.0 Introduction

All of the chapters up to this point have laid the groundwork for the construction of an InGaAs-silicon APD. The advantages of this device design over existing technology are clear. These advantages will remain theoretical, however, until an actual APD is created and measured. The technique of wafer fusion has opened up a means of construction, and as shown, can provide a robust integration method. The fused interface has been well measured using PIN structures and is excellent for use in a detector. A basic design for the detector grounded in existing silicon APDs has been introduced. Now all of these elements can be brought together.

This chapter will outline the design, fabrication, and measurements of three generations of APD detectors. The first detector presented is a very simple design and served as a proof of concept. This device was actually the first ever to use an InGaAs layer fused to a silicon layer. Fabrication was kept very simple with this first design and the main goal was to determine if any avalanche gain was possible in the detector and what the photocurrent was like. The first device could be designated a success in this respect and looked very

promising. This led to the second device design which was significantly more complex. More sophisticated fabrication methods were used in its creation and features like coplanar contact pads were added to allow for frequency response measurements. The second-generation device also proved successful and more promising and this of course led to a third generation design. This detector was created with even more advanced processing techniques and more careful design rules. Its goal was to achieve high bandwidths and make clear the gain-bandwidth-product advantage of this type of detector. This device also could be termed a success and remains the highest gain-bandwidth-product APD reported to date. All three generations of detectors were a challenge to create, and as with most research devices, created as many new questions about the ultimate performance of an InGaAs-silicon APD as they answered.

8.1 First Generation Device - Design and Fabrication

The device design and layer structure for the first-generation of InGaAs-silicon APDs[1] are shown in Figure 8.1. This is a back illuminated detector much like a typical InGaAs/InP APD. Metal contacts were attached to the cathode and anode of the diode. As Figure 45 shows the bottom contact is on the backside of the detector. The absence of two topside contacts allowed only dc measurements of the detector. Other notable features are the p-type layer in the silicon created through implantation. This is sandwiched between the 2.0

μm InGaAs layer that serves as the absorber for the detector and the 2.2 μm silicon layer that serves as the multiplier. Under bias, this scheme creates a high electric field in the silicon while maintaining a low field in the InGaAs.

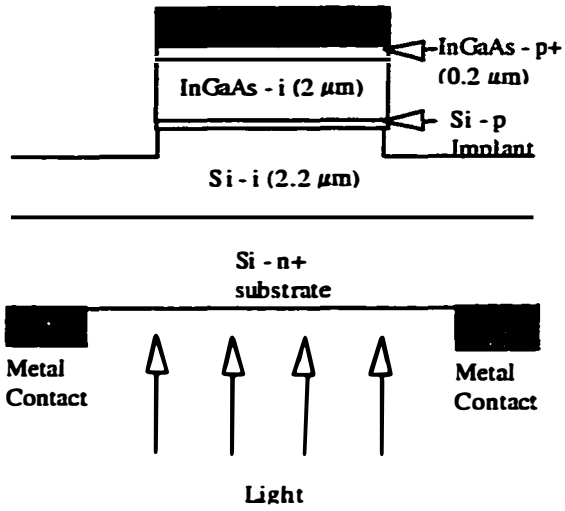


Figure 8.1 - Cross section of a first generation InGaAs-silicon APD with a 2.0 μm InGaAs absorption layer and a 2.2 μm silicon multiplication layer. The detector is back illuminated with infrared light through the silicon substrate.

The actual fabrication of this detector began with the epitaxial growth of two $In_{0.53}Ga_{0.47}As$ layers on an InP (100) substrate. The first region was a 0.2 μm thick p+ layer with a doping level of $2x10^{19}$ cm^{-3} used for ohmic metal contact. The second region was a 2.0 μm thick intrinsic layer unintentionally doped n-

type with a doping level of $4x10^{14}$ cm^{-3} used for photon absorption. The growth in this case was done through MBE (molecular beam epitaxy). At the same time, an undoped 2.2 μm epitaxial layer was grown on an N+ (100) silicon substrate. The unintentional doping of this layer was approximately $5x10^{14}$ cm^{-3} while the doping level of the substrate was $1x10^{18}$ cm^{-3}. The surface of the silicon wafer was then implanted with a shallow $1.3x10^{12}$ cm^{-2} dose of boron ions to create a p doped layer. The implantation dose was calculated to create an electric field layer difference between the InGaAs and silicon of 200 kV/cm. According to this design, at a reverse bias of 65 volts, the field in the InGaAs absorption region is kept between 50 kV/cm and 35 kV/cm. Field strengths of this magnitude allow for high carrier electron velocity through the region, but the field is well below that required for InGaAs avalanche multiplication, given the expected hole and electron ionization coefficients for such fields. The electric field in this silicon region is calculated to vary between 220 kV/cm and 240 kV/cm at a reverse bias of 65 volts, which should result in significant avalanche multiplication.

After the growth of the epitaxial layers and the implantation, fabrication proceeded by fusing the InGaAs layer to the top of the silicon substrate. The fusion conditions were a temperature of 650°C, maintained for 20 minutes in a hydrogen atmosphere. After fusion, the InP substrate was selectively removed with $HCl:H_2O$ leaving the InGaAs integrated onto the silicon. Individual

detectors were made by etching 50 μm x 50 μm square mesas into the structure. The mesas were etched using a methane-hydrogen-argon reactive ion etcher for the InGaAs and a CF_4 reactive ion etcher for the silicon. Total mesa height was 3.0 μm. A top metal layer of Au:Zn and nickel served as an etch mask for the devices as well as a p type contact. Gold was also evaporated on the back of the silicon substrate to make contact to the n+ layer. The devices were left unpassivated for testing. A top-view photograph of several SHIP detectors is shown in Figure 8.2.

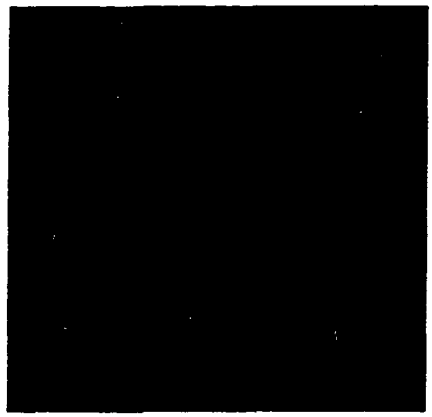


Figure 8.2 - Top view photograph of four first-generation InGaAs-silicon APDs. The detectors are square in shape with 50 μm sides.

8.2 First Generation Device - Measurements

The first measurements made on these detectors were dark current and photocurrent versus reverse bias curves as shown in Figure 8.3. As the figure shows, the total dark current remains below 1 μA for these detectors out to

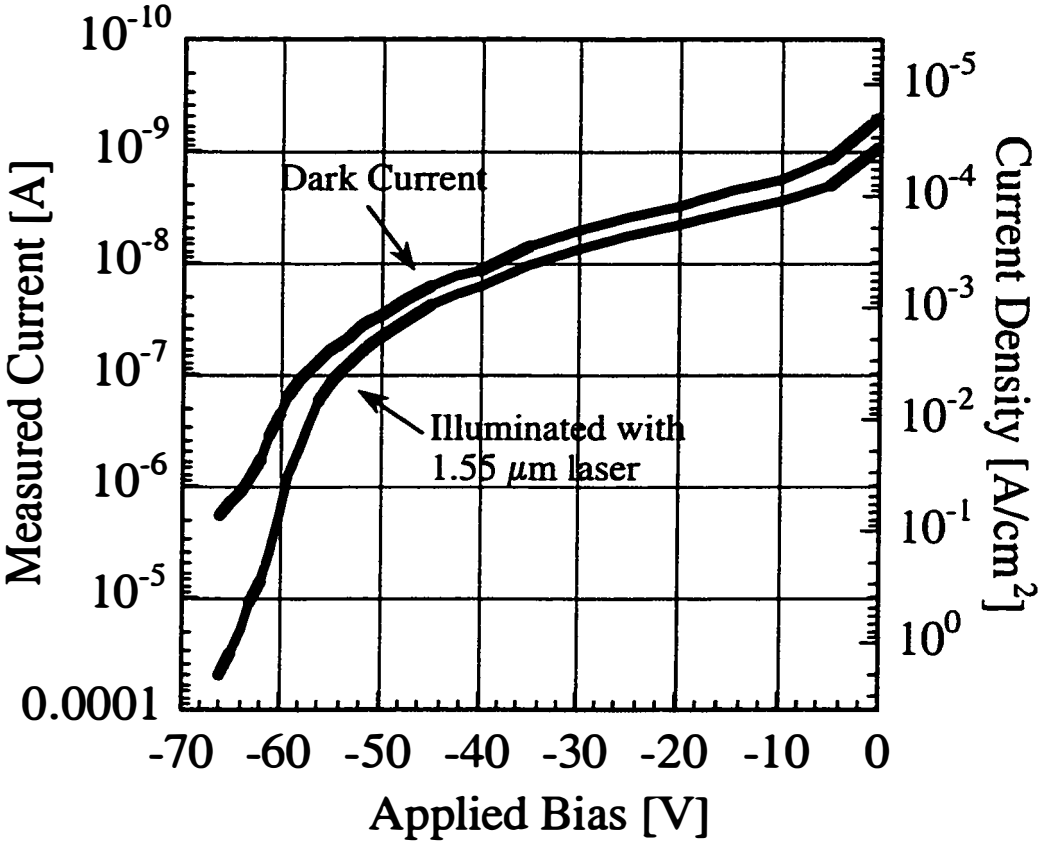


Figure 8.3 - Measured current versus reverse bias for first generation InGaAs-silicon APD. The dark current is shown along with the photocurrent curve produced from illumination with a 1.55 μm laser.

approximately 65 V reverse bias. This corresponds to a peak dark current density of 0.04 A/cm^2. As bias is increased beyond 50 V, dark current begins to increase rapidly as higher electric fields are placed across the device layers. The photocurrent shown in Figure 8.2 is the response to a 1.55 μm laser. This curve tracks proportionally to the dark current until around 55 V where it begins to increase more rapidly. This is due to the full depletion of the p implant layer near this voltage that allows for free carrier movement between the absorption and multiplication layers.

The responsivity of these devices was calculated by measuring the total light incident on the absorbing region using a calibrated detector. Assuming that the devices quantum efficiency was 100%, Figure 8.4 shows the gain versus voltage based on these responsivity calculations. Responsivities over 4 A/W were measured when the detectors were illuminated with a 1.55 μm laser and operated in the avalanche regime near a reverse bias of 65 V, indicating avalanche gains of at least five. Based on the calculated field profile this gain is taking place in the silicon multiplication region.

A more comprehensive look at the electric field profile in the detector at different biases is shown in Figure 8.5. The different profiles were calculated based on a depletion approximation model given the layer thicknesses and dopings assumed from the fabricated structure. On the scale, 0.0 μm represents the beginning of the depleted regions. Between 0.0 and 2 μm is the InGaAs

minimally doped absorption region. At a distance of 2 μm into the detector, is the p doped region in the silicon created through implantation. 2.0 μm to 4.2 μm represents the minimally doped epitaxial silicon layer serving as the multiplier.

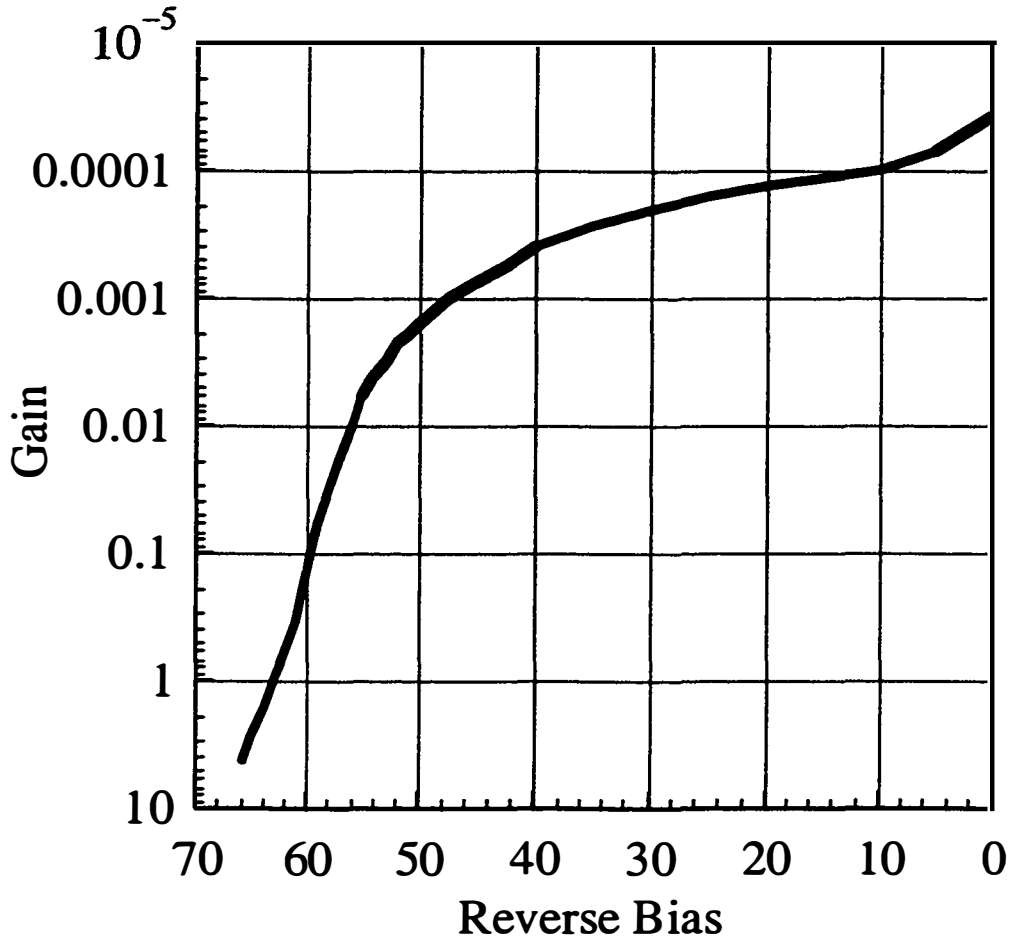


Figure 8.4 - Multiplication gain versus reverse bias for a first -generation InGaAs-silicon APD based on measured responsivity to known 1.55 μm light source.

As the profiles indicate, a 20 V reverse bias produces electric field only in the silicon and a barrier for electrons remains at the InGaAs-silicon interface. At 40 V reverse bias, the carrier barrier has been reduced so that electrons can begin to transition into the silicon, while the electric field has grown in the

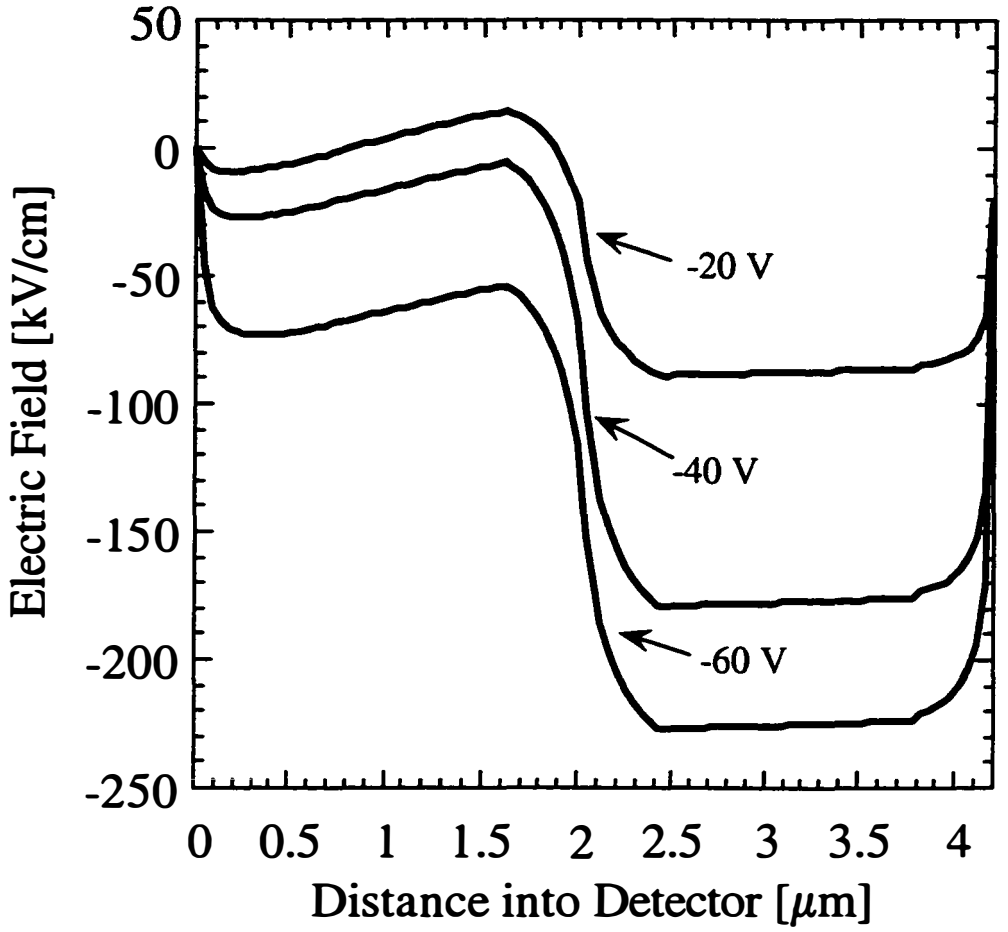


Figure 8.5 - Electric field profile in a first generation InGaAs-silicon APD. The calculated electric field is shown at three different biases based on the absorption and multiplication layer thicknesses.

silicon. An additional 20 V of reverse bias increases the field in both the InGaAs and silicon regions so that the necessary field strength for avalanche multiplication is reached. The characteristics of the photo and dark current curves correspond fairly well with that predicted by the simple model, with avalanche gains first seen at biases above 60 V and an increase in the photocurrent corresponding to an electron field barrier lowering.

Measurements of the optical response of the InGaAs-silicon detector to different light wavelengths were also made. Figure 8.6 shows the voltage dependence of the photocurrent normalized to the photocurrent at -35 V for three different wavelengths of incident light. Since the device is back illuminated through a silicon substrate, we expect an amplified response only for wavelengths between 1.0 μm and 1.65 μm. This represents the window between the cutoff wavelengths for InGaAs and silicon absorption. At 1300 nm and 1550 nm, photocurrent multiplication was observed indicating absorption in the InGaAs and electron injection into the silicon multiplication region. At a wavelength of 920 nm, a small photocurrent was observed that did not increase with increasing reverse bias, indicating it was due to light absorbed in the substrate and hole diffusion to the junction without amplification in the avalanche region.

The measurements on this first generation device indicate the hoped for properties for the APD. There was an observed avalanche gain in response to

infrared light which is a significant step for any APD design. The photocurrent followed expected behavior based on modeling of the electric field and the dark current was reasonably low. These factors provided the motivation for the creation of a more complete, second generation APD.

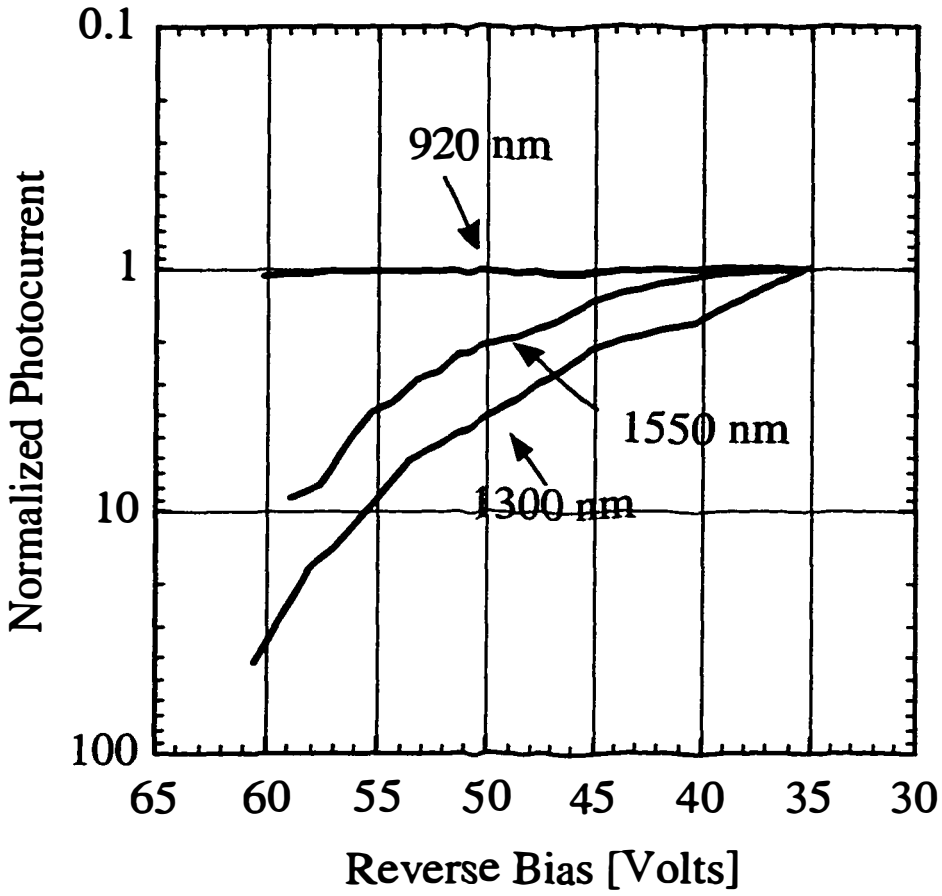


Figure 8.6 - Normalized photocurrent response versus reverse bias for first generation InGaAs-silicon APD, measured using light sources at three different wavelengths. The longer wavelengths exhibited a multiplication gain.

8.3 Second Generation Device - Design and Fabrication

The device design and layer structure for the second-generation of InGaAs-silicon APDs[2] are shown in Figure 8.7. This is again a back illuminated detector much like the first generation device. The first feature that distinguishes this detector from its predecessor is that both the p and n contact metals are on the top of the device in a coplanar fashion. This allows for frequency response measurements using a coplanar probe. Another advancement is the presence of passivating layers encapsulating the mesa that makes up the device. This prevents deterioration of etched surfaces from exposure over time and also provides a high dielectric constant layer for the contact pads to rest on. In addition, this device was made using two fusion steps, one to attach a 1.0 μm InGaAs absorption layer and another to attach a

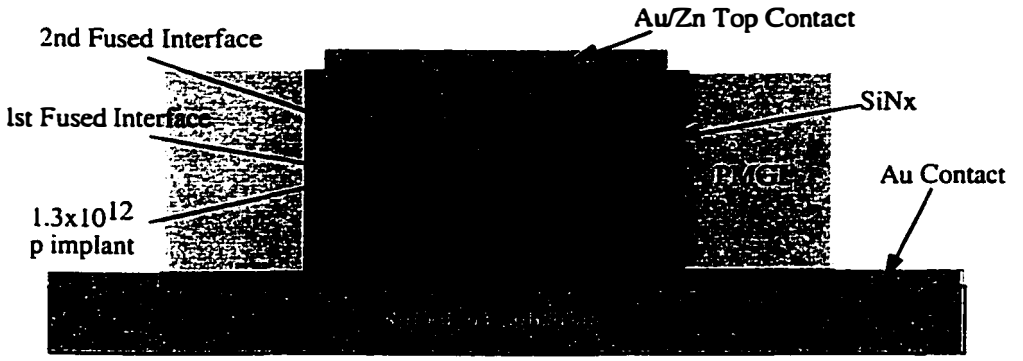


Figure 8.7 – Cross-section diagram of a second generation InGaAs-silicon APD. This device was created through RIE etching and encapsulated with dielectric layers.

0.25 μm contact layer. Using two separately fused layers prevents some of the diffusion of zinc caused when a p+ InGaAs epitaxial layer is grown under an n- epitaxial layer, as was the case with the first generation device.

Detector fabrication began first with the epitaxial growth of three different layers. First a 1.0 μm undoped layer of $In_{0.53}Ga_{0.47}As$ was grown on an n+ InP substrate through MOCVD (metal-organic chemical vapor deposition). A second 0.25 μm p+ $In_{0.53}Ga_{0.47}As$ layer was then grown on a p+ InP substrate. An intrinsic epitaxial silicon layer similar to that used for the first generation device was then grown on an n+ silicon substrate. A shallow ion implantation dose of $1.3x10^{12}$ cm^{-2} boron ions were placed in the silicon surface to provide a shallow p doping to control the electric field profile as before. Upon completion of the epitaxy and implantation, the InP wafer with the undoped InGaAs layer was fused to the silicon wafer at a temperature of 650°C for 20 minutes in a hydrogen atmosphere. After the first fusion step, the InP substrate was selectively removed leaving only the InGaAs epitaxial layer. The second MOCVD grown p+ doped InGaAs layer was then fused to the first InGaAs layer and its InP substrate subsequently removed. This second fusion step was also done at 650°C for 20 minutes in a hydrogen atmosphere. Again, without this double fusion scheme, some dopant diffusion is unavoidable when the intrinsic layer is epitaxially grown directly over the p+ layer. This is especially true when using zinc which has a tendency to migrate toward the grown surface.

The epitaxial layers of the finished device are as follows, starting from the topmost layer. First a 0.2 μm thick $In_{.53}Ga_{.47}As$ p+ layer with a doping level of $2x10^{19}$ cm^{-3} is used for ohmic metal contact. The second region is a 1.0 μm thick intrinsic InGaAs layer unintentionally doped n-type used for photon absorption. This layer was fused to a silicon surface implanted with a shallow, $1.3x10^{12}$ cm^{-2} dose of boron ions. Below this implant was a 2.2 μm intrinsic epitaxial silicon layer, unintentionally doped n-type with a doping level of approximately $5x10^{14}$ cm^{-3}. This layer serves as the multiplication region for the detector and sits on an n+ substrate with a doping level of $1x10^{18}$ cm^{-3}. Again, the implantation dose in the silicon was calculated to ensure that the electric field in the intrinsic silicon region is higher than that in the intrinsic InGaAs region when the device is biased at operating voltages. For significant avalanche gain in the multiplication region, electric fields of 240 kV/cm to 300 kV/cm are required, while the field in the InGaAs needs to remain below 100 kV/cm. Fields of this strength in the InGaAs layer allow for electron velocities of over $7x10^{6}$ cm/sec through the region but inhibit avalanche multiplication.

After the fusion and InP substrate removal steps, further fabrication steps proceeded as follows. First the epitaxial InGaAs and silicon layers were etched down to the n+ silicon substrate, leaving only circular mesas of variable diameters between 20 and 35 μm. This mesa etch provided device isolation and determined the active area of the detector. A reactive ion etcher (RIE) using a

mixture of methane-hydrogen-argon was used for etching InGaAs and an RIE using Cl_2 was used for etching Si. A top metal layer of AuZn/Ni served as an etch mask for the devices as well as a top p-type contact. A dielectric layer of silicon nitride was first applied followed by a spun- on layer of PMGI - a polyimide type passivant. Gold contacts were then added to provide a contact to the n+ substrate layer on the silicon surface to allow for coplanar probing.

A top-view photograph of a single completed device is shown in Figure 8.8. The light U-shaped figure in the photo is the metal contact to the silicon substrate. This shape was chosen to match a 200 μm pitch coplanar probe and to provide maximum surface contact. The p metal contact consists of the square

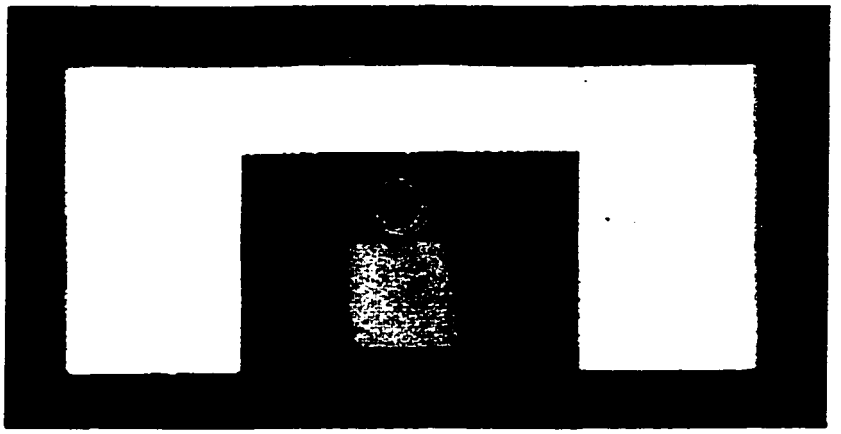


Figure 8.8 - Top view photograph of a second generation InGaAs-silicon APD. The U-shaped figure is the n metal contact and the p metal contact is the square figure attached to the circle within the U-shaped figure. The active device area had an approximately 25 μm diameter and is under the circle part of the p contact.

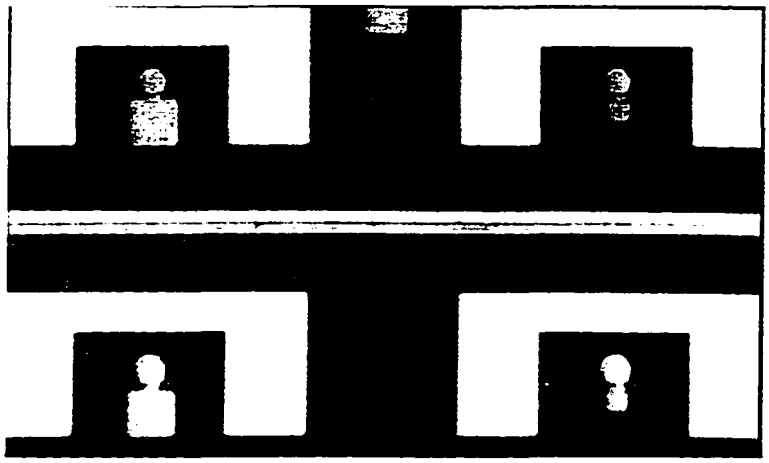


Figure 8.9 - Top view photograph of several second generation InGaAs-silicon APDs constructed on the same wafer. Two different p contact pad sizes are shown.

figure (50 μm x 50 μm) attached to the circle, inside the U-shaped figure. The active device area is below the circle part of the p contact and is approximately 25 μm in diameter. Figure 8.9 shows a group of detectors on the surface of the wafer. Some of the p-metal pads have areas of 25 μm x 25 μm, smaller than that for the detector in Figure 8.8. Of course the relatively small size of the detectors allows for the creation of many on a small substrate. For a 1 inch x 1 inch square substrate, 4000 APDs of this size can be constructed.

8.4 Second Generation Device - Measurements

The first measurements made on the second generation of InGaAs-silicon devices were dark and photocurrent versus bias. Infrared response of the detectors was measured by back illuminating them through the silicon substrate with a 1.3 μm laser. Figure 8.10 shows a 23 μm diameter device illuminated

with approximately 20 μW of laser light. The photocurrent curve is close to what would be expected. There is a large initial increase in the photocurrent and dark current for a small increase in the reverse bias, then a relatively flat region where first the p-type ion implant in the silicon and then the InGaAs absorption layer are being depleted and the gain is approximately one. There is then a visible kink in the response curve with the onset of avalanche gain.

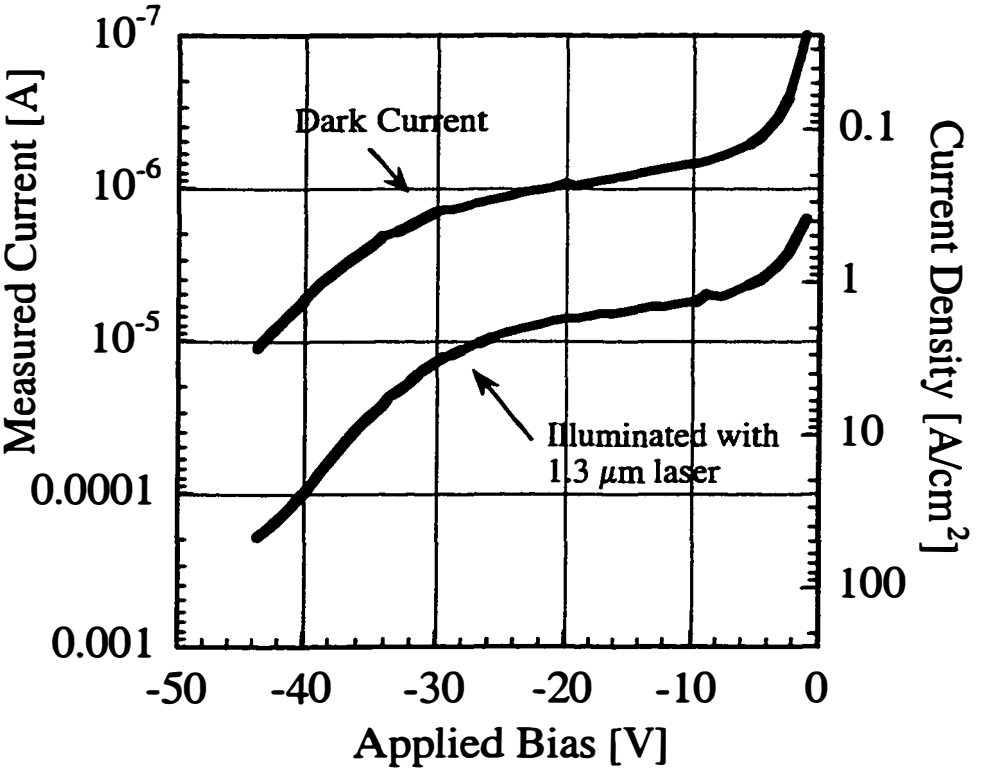


Figure 8.10 - Measured current versus reverse bias for a second generation InGaAs-silicon APD. Shown are the dark current curve and the photocurrent curve produced by illumination with a 1.3 μm laser.

Unfortunately, the dark current is higher than desired and even higher than the first generation device. Some of the reasons for this could be the RIE etch steps that were different for the detectors. The deposition of a silicon-nitride layer to serve as a surface passivant might also have created surface damage leading to higher leakage current.

Given the shape of the photocurrent curve shown in Figure 8.10, multiplication gain was determined for this device. This was done by assuming that the photocurrent in the flat region of the curve, at -25 V, represents a gain=1, before the onset of avalanching. Gains at other voltages were then determined by dividing their photocurrent by the photocurrent at -25 V. Gains of over 25 were measured for incident light levels of around 20 μW and gains of over 130 were measured for light levels of around 2 μW. The difference in the achievable gains at different light levels is due to the high series resistance of the detector. Large currents create voltage drops across the series resistance and decrease the available voltage across the detector, lowering the net gain. The gain for the 23 μm device from Figure 8.10 is shown in Figure 8.11. Again there is 20 μW of 1.3 μm laser light incident on the detector. The curve indicates a very voltage insensitive device as predicted in Chapter 4. As seen in the figure, at a gain of ten even a five volt increase in bias only increases the gain to 25. Measurements of devices of different diameters showed that larger diameter devices had higher avalanche onset voltages, indicating that the

lateral electric field profile varies, with higher gains near the perimeter. This is often the case with etched mesa APD structures and is one of their biggest problems. Uneven profiles create excess noise and dark current from the regions with higher than intended gains.

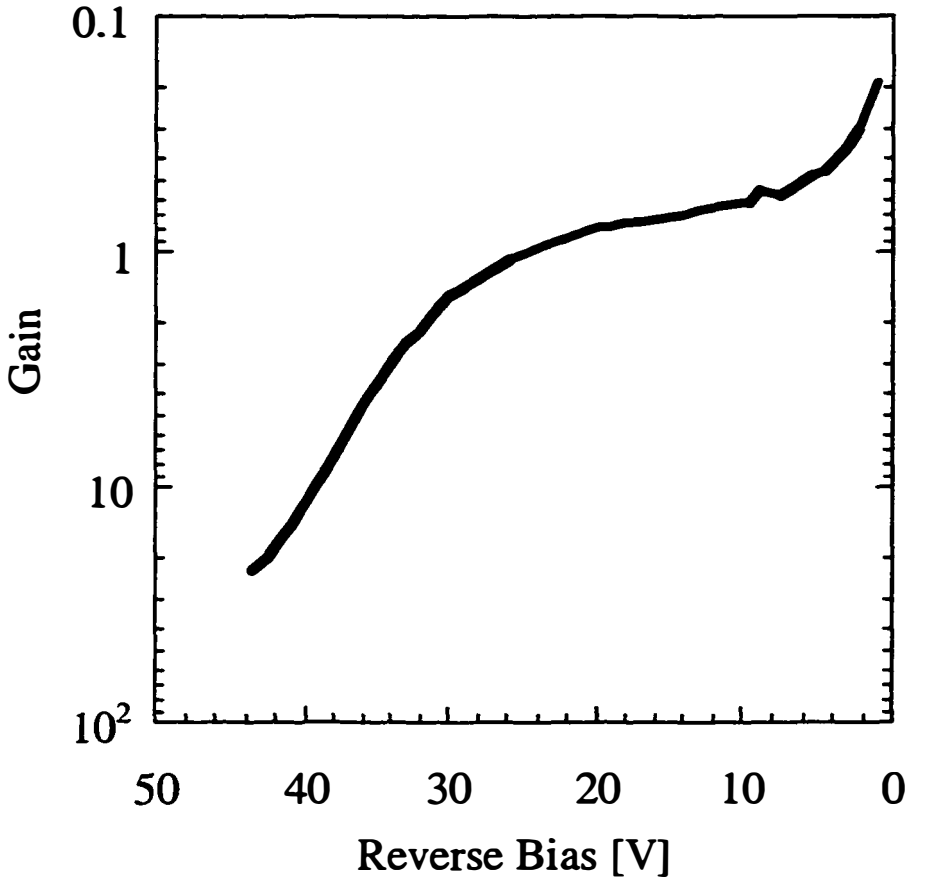


Figure 8.11 - Multiplication gain versus reverse bias for a second generation InGaAs-silicon APD. Gain is based on normalizing the photocurrent to the response at a reverse bias of 25 V.

The expected electric field profile based on measured device parameters can be calculated using the depletion approximation. Figure 8.12 shows what the fields should look like at different points in the detector for various biases. From 0.0 to 1.0 μm, is the InGaAs n- absorption region. The implanted p-type region sits at 1.0 μm on the scale, and from 1.0 μm to 3.2 μm is the silicon multiplication region. As with the first generation device, the p implant region was intended to create a high field in the silicon and maintain a low field in the InGaAs. Given the assumed layer thicknesses and dopings, at a negative bias of 20 V, most of the voltage drop occurs only in the silicon and an electric field barrier resides at the InGaAs-silicon interface. At -40 V, the implant has been depleted and the barrier lowered to allow for free carrier transfer between the materials. Added bias increases the electric field to a magnitude high enough to produce significant multiplication in the silicon by -60 V. While these features are seen qualitatively in Figures 8.10 and 8.11, the calculated avalanche gain occurs at voltages less than predicted by the simple model. This seems to indicate that the actual layer thicknesses of the silicon and InGaAs are thinner than expected. This is certainly possible in the InGaAs where zinc could have diffused from the p+ into the n- layer during fusion. The 2.2 μm silicon layer could also be much thinner than specified due to diffusion of substrate doping into the n- layer during epitaxial growth.

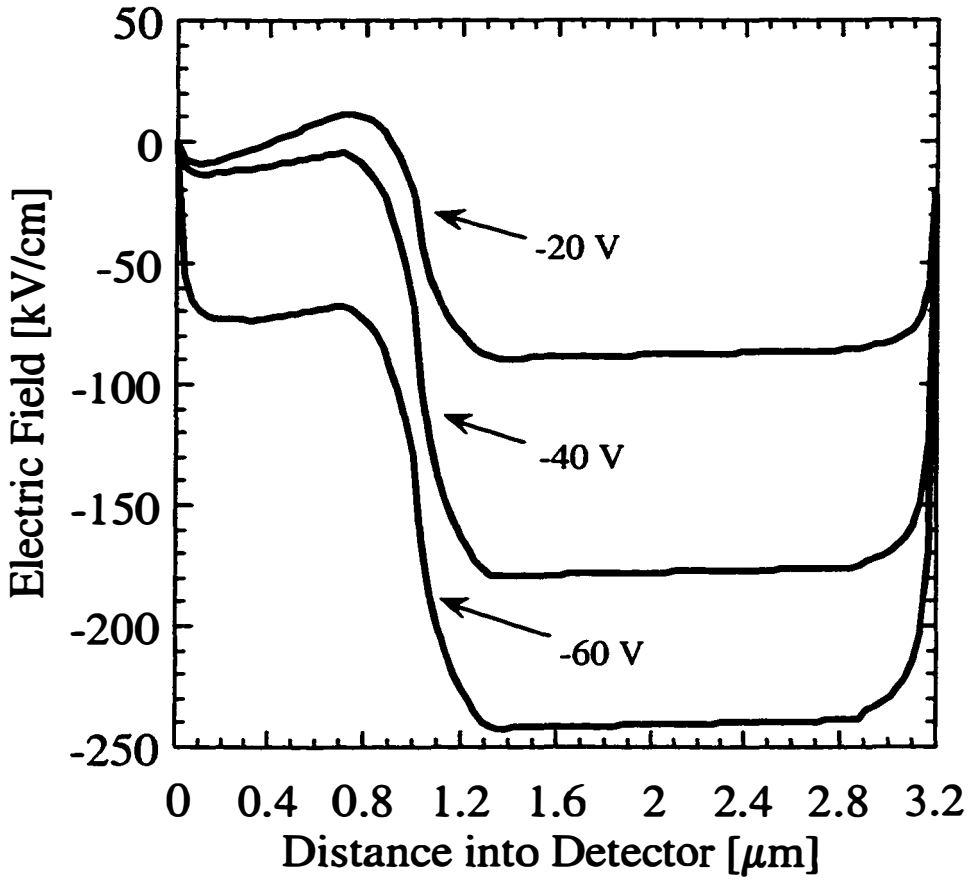


Figure 8.12 - Electric field profile in a second generation InGaAs-silicon APD. The calculated electric field is shown at three different biases based on the absorption and multiplication layer thicknesses.

Frequency response measurements were also made on the detectors using an HP 8703a Lightwave Component Analyzer and coplanar probes. The component analyzer has a measurement range from .13 GHz to 20 GHz. For a

23μm diameter device illuminated with a modulated 1.3 μm laser, at a gain of 10, a 3 dB bandwidth of 820 MHz was measured. At a gain of 135, a 3 dB bandwidth of 600 MHz was measured yielding a gain-bandwidth product of 81 GHz. The frequency response is shown in Figure 8.13. The maximum gain-

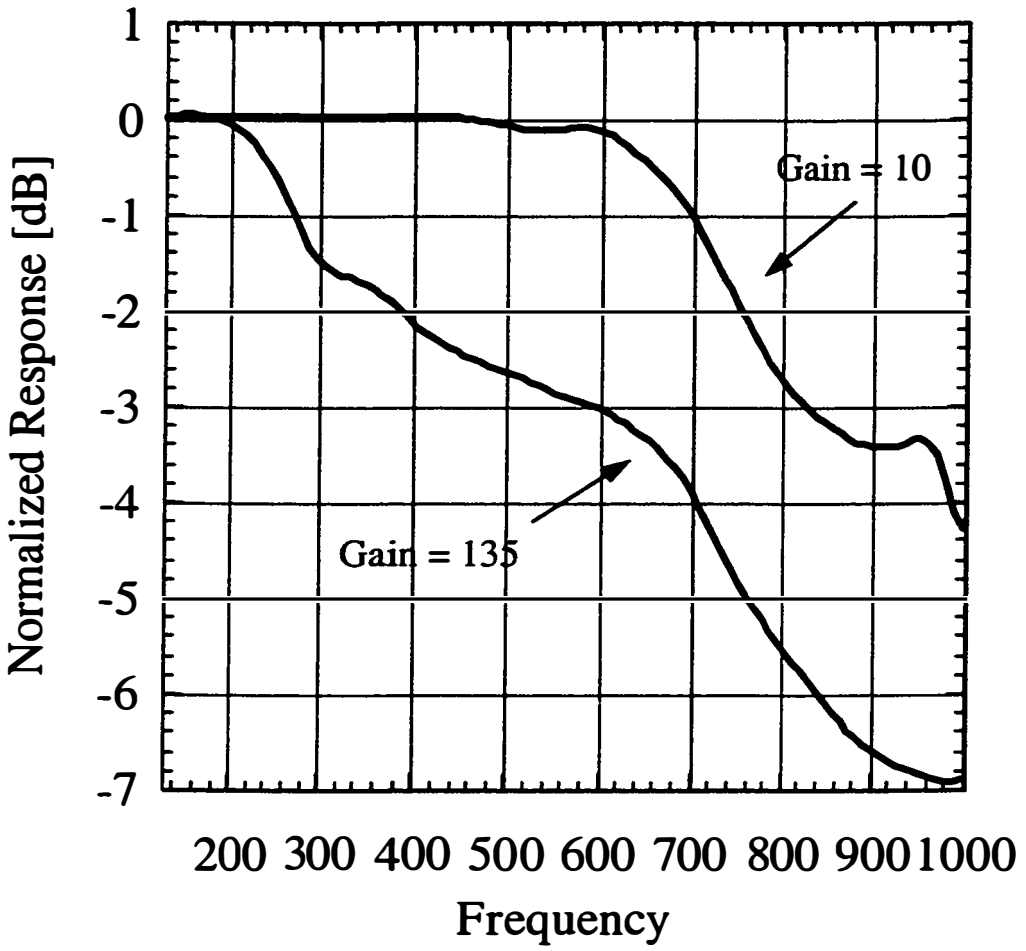


Figure 8.13 - Normalized frequency response for a second generation InGaAs-silicon APD. Curves for the same device biased at two different gain levels are shown.

bandwidth product is near the highest gain-bandwidth products reported for InGaAs/InP avalanche photodetectors. The main limiting factor for the speed of these second-generation devices was the large series resistance caused by a choice of bad ohmic contacts. A series resistance of over 2000 ohms was

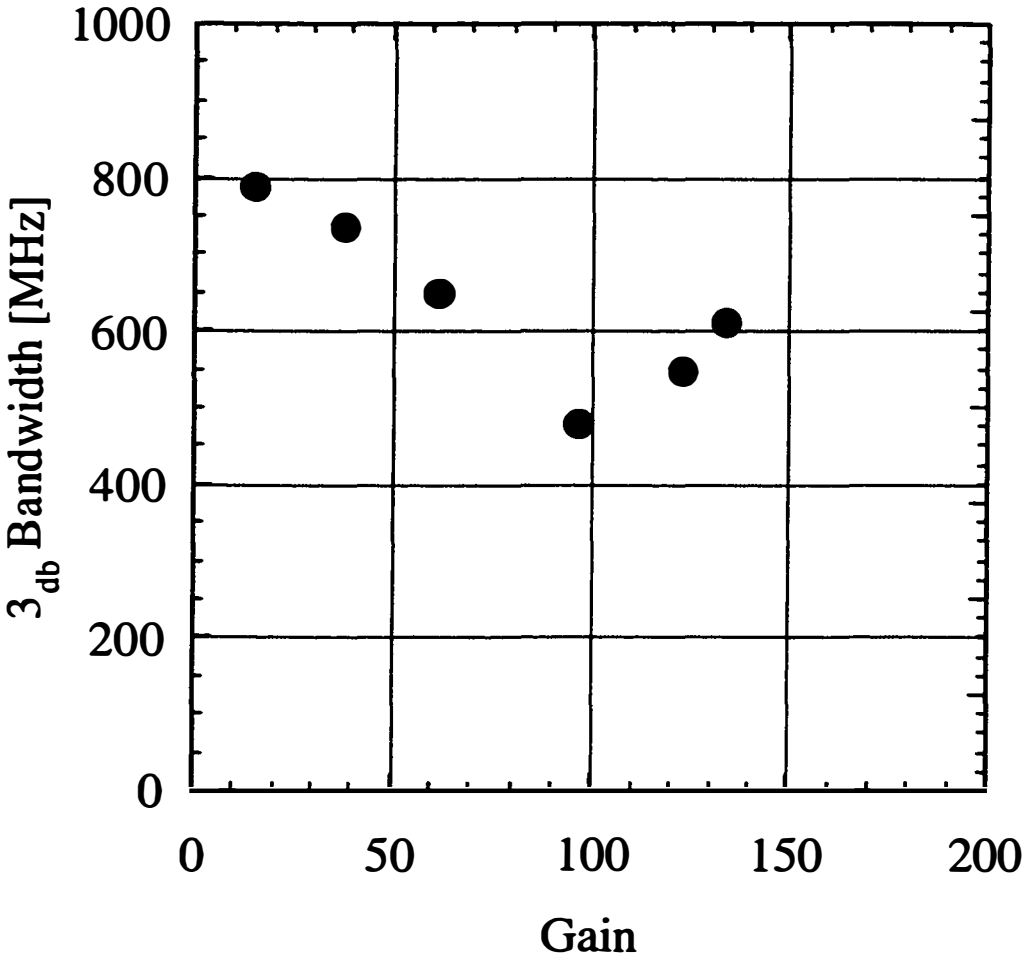


Figure 8.14 - 3-dB bandwidth versus multiplication gain for second generation InGaAs-silicon APD.

measured, combined with a capacitance of 0.1 pF. These factors combine to limit the frequency of the detector to ($1/2\pi RC$) = 800 MHz, which is exactly what was measured on the component analyzer at the relatively low gain of 10.

Figure 8.14 plots the 3 dB bandwidth versus the gain for the second generation detector. This type of curve is a typical performance measure for avalanche photodetectors. For all APDs, as the gain becomes higher, the frequency response drops. This is due to the increasing avalanche multiplication time required for higher gains. In this particular case, the 3 dB bandwidth begins at 800 MHz which is the RC limit. Over gains of 50 there is an obvious decrease with a dip at a gain of 95. This dip is not due to any multiplication effect but is simply due to microwave resonances in the measurement.

The second generation fused APD, much like the first was reasonably successful but also created many questions. High gains were measured, but these were accompanied by high dark currents. Was it possible to make a lower dark current device with high gains? The first frequency response measurements were made but the devices were RC limited. The first gain-bandwidth number was reported, but was a higher number possible with a device better designed for higher speeds? What could be done to improve the processing and better define the layers? All these questions of course led to another device generation.

8.5 Third Generation Device - Design and Fabrication

The device design and layer structure for the third-generation InGaAs-silicon APD[3] are shown in Figure 8.15. There are significant changes compared to the previous designs. The first significant change is the absence of a silicon epitaxial layer. This was used in the earlier versions together with an implant to define the multiplication layer in the detector. In this third generation, the multiplication layer was created only through implantation of a silicon substrate. Another big change was the use only of PMGI as a passivating dielectric instead of a combination of silicon nitride and PMGI. Also present are guard ring structures surrounding the active detector mesa, intended to reduce the total dark current by eliminating surface leakage currents. The guard ring had a separate contact pad and could be biased independently. The metal system was also changed in an effort to reduce the series resistance. Aluminum was substituted for gold for the ohmic contact to n+ silicon. The top p contact remained the same. The double fusion scheme was replaced by the single fusion of two InGaAs layers designed to have less zinc diffusion than could be expected in the first generation device. The detector remains a back-illuminated device and features topside metal contacts suitable for coplanar probing.

The fabrication of this third generation of InGaAs-silicon detectors began with the epitaxial growth of the two InGaAs regions. First a 1700 Å p+ $In_{0.53}Ga_{0.47}As$ layer zinc doped to $1x10^{19}$ cm^{-3} was MOCVD grown on an InP

substrate. A 300 Å $In_{0.53}Ga_{0.47}As$ layer was then grown with the zinc doping graded from $1x10^{19}$ to $1x10^{18}$ cm^{-3}. A 1.0 μm nominally undoped n- $In_{0.53}Ga_{0.47}As$ layer with doping less than $1x10^{15}$ cm^{-3} was then grown above the p doped layers. The 300 Å doping grading layer was used to reduce zinc diffusion into the n- layer during growth and subsequent wafer fusion. The

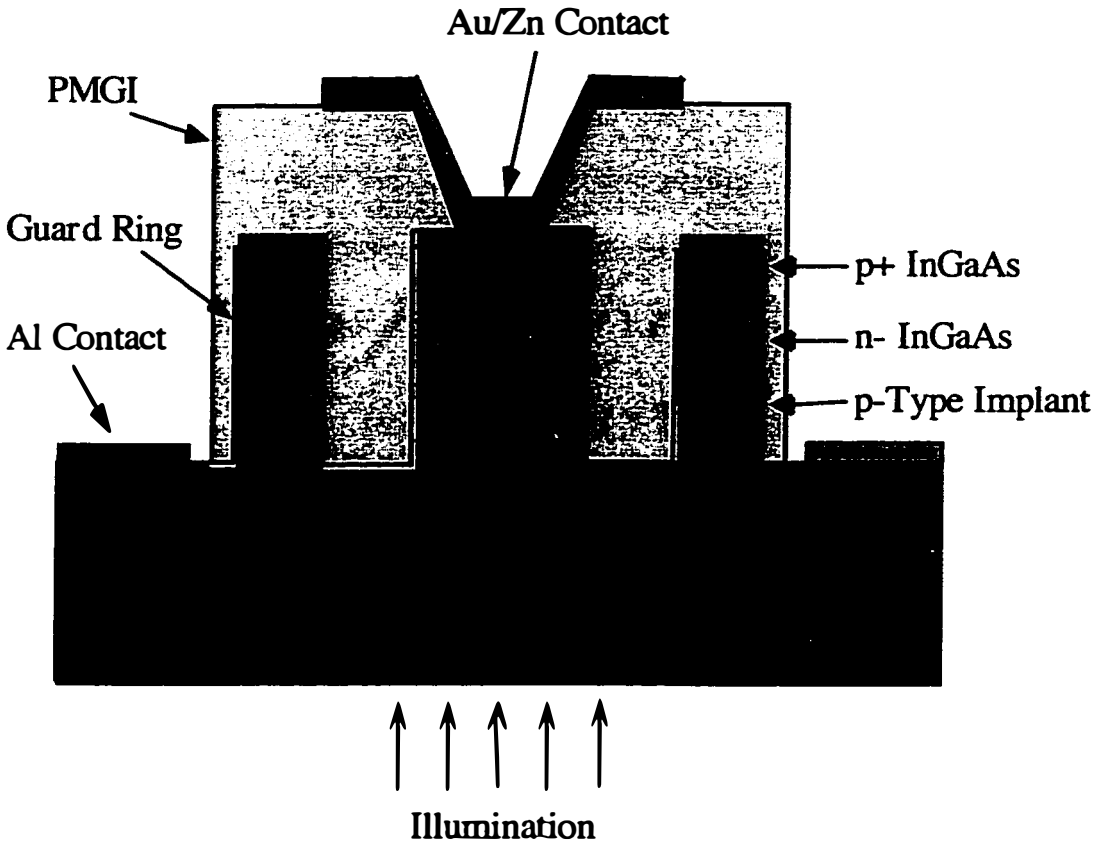


Figure 8.15 – Cross-section for a third generation InGaAs-silicon APD. This device was created with RIE etching and was encapsulated with the dielectric PMGI.

silicon portion of the detector was made up of an n-type silicon wafer with a resistivity of 0.142ohm-cm implanted with a $3.0x10^{12}$ cm^{-2} dose of 10 keV boron ions. This dose was calculated to compensate the n doping in the silicon and provide a p doped layer at the surface of the wafer.

The top n-InGaAs layer was then fused to the silicon surface by placing them in direct contact under pressure at a temperature of 650°C for 10 minutes in a hydrogen atmosphere. After fusion, the InP substrate was selectively removed from the InGaAs epitaxial layers. The resulting structure had the following layers, starting at the topmost layer (Figure 8.15). First, a p+ InGaAs layer was used for ohmic contact, followed by a 300 Å doping grading layer. Below this was the 1.0 μm n- layer used as the photon absorption layer followed by a shallow implanted p layer in silicon above a thick n type silicon substrate.

Fabrication of devices involved first applying a top Au/Zn metal that served as an ohmic contact to the p+ layer as well as a mask for reactive ion etching (RIE). Circular mesas and guard rings were defined by etching through the InGaAs layers and 1.0 μm into the silicon substrate with a Cl_2 RIE. A 3.0 μm thick dielectric coating of PMGI was then applied. Via holes on top of the circular mesas were opened using deep UV photolithography and gold probing pads evaporated on the PMGI surface. Openings in the PMGI were then made for contact to the n type silicon substrate, and aluminum was used as the contact metal. The final devices had metal contact pads suitable for coplanar probing

using a 200 μm pitch probe and a variety of mesa diameters between 20 and 30 μm.

Figure 8.16 shows a top-view photograph of a completed third generation InGaAs-silicon APD. The active 23 μm diameter mesa of the device is in the center of the picture within the light colored ring. The p-contact metal extends from this active region out to a larger square pad resting on the PMGI dielectric. The guard ring structure surrounds the active mesa and has its own contact pad. On either side of these structures are two rectangular n-metal contact pads which rest directly on the n+ silicon substrate. Figure 8.17 is a top view photograph of several detectors on the wafer surface along with several test structures and alignment marks. As with the second-generation detector, the die size per detector would allow for a total of 4000 devices in a 1 in^2 area.

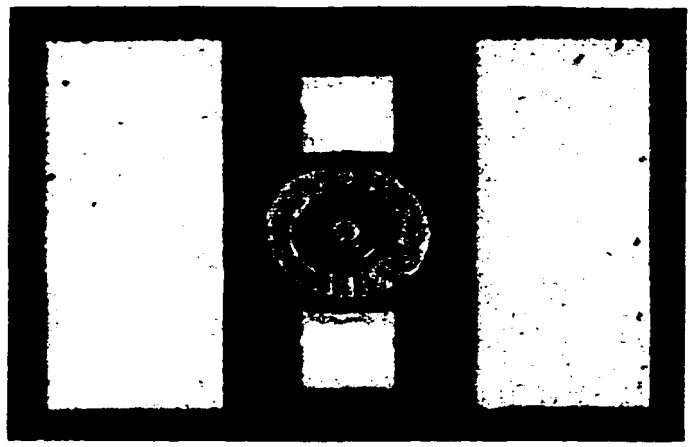


Figure 8.16 - Top view photograph of a third-generation InGaAs-silicon APD. The active area of the detector is the circular mesa between the ring in the middle of the

picture. N-metal contacts are on the outside of the ring and separate p-metal contacts for the active area and the guard ring are in between. The metal pad spacing was designed to allow for coplanar probing.

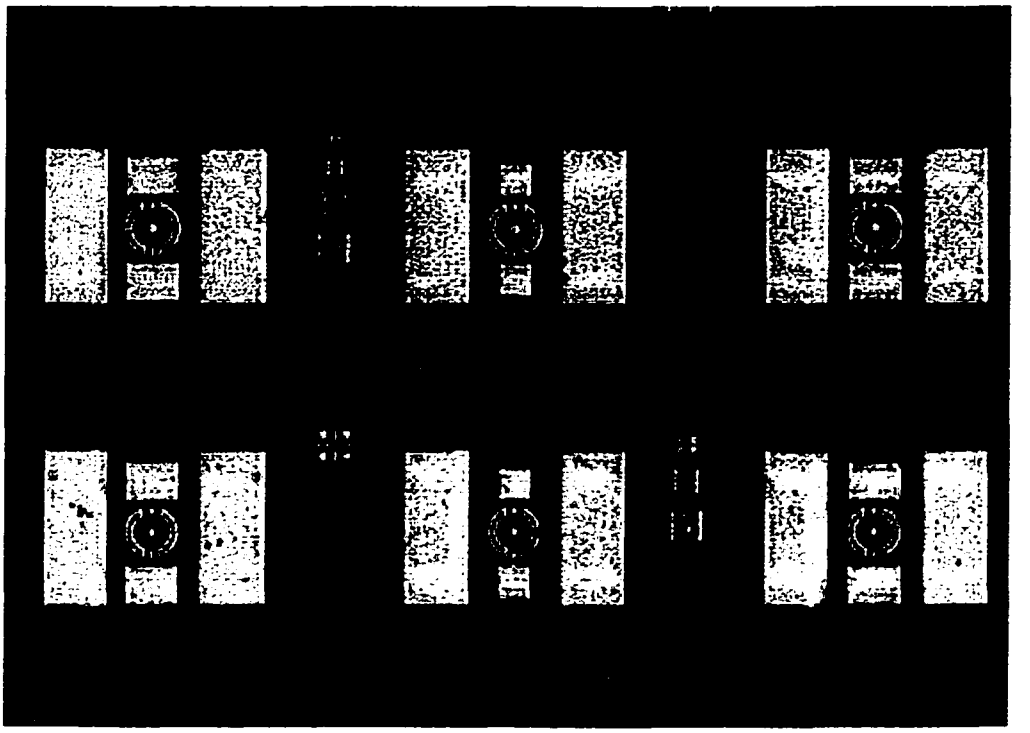


Figure 8.17 - Top view photograph of several third-generation InGaAs-silicon APDs constructed on the same wafer. Two different p contact pad sizes are shown.

8.6 Third Generation Device - Measurements

The first characteristics investigated for this third generation device were its dc leakage currents and response to infrared light. The quantum

efficiency of the detector was measured by back illuminating non-avalanching structures similar to Figure 8.14 but without the p type multiplication region. These devices exhibited no current gain as expected. Using a 1.3 μm laser, 150 μW of light was incident on the silicon backside with an air-silicon interface transmission of 0.70. Photocurrent was then measured at different reverse biases. After an initial photocurrent increase with increasing bias due to the depletion of the InGaAs n- layer, the photocurrent curve flattened and a photocurrent of 60 $\pm$ 5 μA was measured. Correcting for the reflection at the air-silicon interface, the responsivity R for the detector at full layer depletion was 0.57 A/W and the quantum efficiency η was 0.60 (these quantities are defined in the PIN device chapter). For a 1.0 μm absorption layer (w) of InGaAs with an absorption coefficient (α) of 1.16 μm^{-1} at a wavelength of 1.3 μm, we would expect a quantum efficiency of $1 - e^{(-\alpha \times w)}$, or 0.69. The 12% disparity in our measured and expected quantum efficiencies can likely be explained by a smaller than expected absorption layer due to zinc diffusion from the p+ layer during the high temperature fusion step. The measured quantum efficiency would predict an absorption layer approximately 0.7 μm thick, giving a 300 nm diffusion layer width from the p+ region into the n- absorption region. This diffusion distance is certainly within reason given the fusion parameters. Another explanation for the slightly low quantum efficiency could be the loss of carriers across the InGaAs-silicon interface, although the earlier data from the PIN devices would contradict this..

The shape of the photocurrent verses reverse bias curve (Figure 8.18) shows the two regions we would expect to see in this APD. First there is a

region where the photocurrent increases rapidly due to depletion through the absorption layer (between 0 V and -9 V). Then after full depletion, there is a region of gradual increase in the photocurrent gain as the electric field is increased (Figure 8.18) in both the InGaAs and silicon layers (between -9 V and -24 V). Note that the dependence of gain on voltage is much more gradual than for InP APDs, as expected for a silicon multiplication layer. The implant dose in the surface of the silicon was calculated to provide a triangular electric field with a peak height of over 500 kV/cm in the silicon while maintaining a lower electric field of approximately 100 kV/cm in the InGaAs. This profile allows for avalanche gain only in the silicon. The determination of gain is difficult in an APD for which the absorption layer does not deplete until the multiplication layer is biased high enough to yield significant gains. To calculate the gain, the device was modeled based on known parameters and the shape of the photocurrent versus bias curve. Electric field profiles were generated[4] for different biases based on the InGaAs-silicon structure shown in Figure 8.14 and avalanche gain was determined based on these field profiles and ionization coefficients used to calculate the bandwidth of the detectors[5]. Based on the absorption layer depleting at 9 V, and the ratio of measured photocurrent at voltages of -9 V and -24 V, the computed electric field profiles at various voltages are shown in Figure 8.19 and the gain curve is shown in Figure 8.18. The parameters used to obtain these curves are a silicon substrate doping of 6.7 x 10^{16} cm^{-3}, ion implant of 3.18 x 10^{12} cm^{-2}, and InGaAs absorption layer thickness of 0.66 μm. These doping levels fall within the range of measured values and the absorption thickness corresponds well with what we would

expect based on zinc diffusion and quantum efficiency measurements. Measured photocurrent gain is also plotted in Figure 8.17 normalized to the

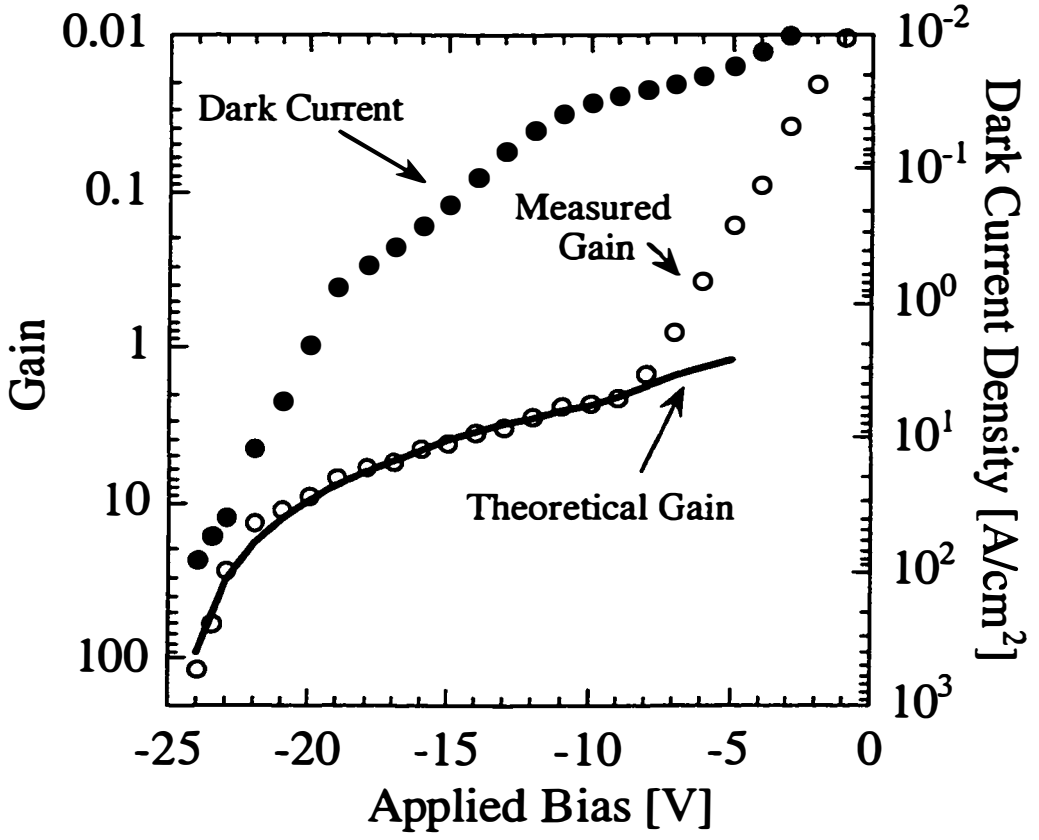


Figure 8.18 - Multiplication gain and dark current versus reverse bias for a 23 μm diameter third generation InGaAs-silicon APD. The detector was illuminated with 5 μW of 1.3 μm light. The solid line indicates gain based on modeling, the open circles represent measured photocurrent gain normalized to the gain at 9.5 V, which should be 2.2. The solid circles represent measured dark current density.

gain at 9.5 V, which based on our model is 2.2. Measured dark current versus reverse bias is also shown in Figure 8.17. As indicated by the figure, the dark current is initially in the nA range but increases rapidly as the detector is biased into the avalanche gain regime. The rapid increase in dark current made photocurrent measurements at high gains difficult and the current became unstable at gains of over 40, due to device heating. Much of the leakage current was probably due to edge breakdown.

Another feature of Figure 8.18 to notice is the relative insensitivity of the gain to bias changes. At a gain of ten, a ten percent voltage increase only increases the gain to about 20, while a ten percent voltage decrease only decreases the gain to about six. A comparable InGaAs/InP APD would be much more voltage sensitive and a ten percent voltage swing would lead to much larger swings in the avalanche gain. InGaAs-silicon devices are more insensitive due to the large difference in ionization coefficients in silicon, as explained in Chapter 4.

The frequency response of the third generation detector was improved over the second- generation detector which had a maximum bandwidth of 800 MHz. The 800 MHz design was limited primarily by RC roll off due to the large contact resistance (2000 •) with the n-metal pads and silicon. In addition the multiplication layer thickness in the earlier device was relatively thick (2.5 μm). The latter detector had a device resistance of less than 50 ohms and a

capacitance of less than 0.1 pF so the frequency response was limited by carrier transit time instead of RC roll off. In addition the multiplication layer for the silicon was calculated to be around 0.65 μm thick, much thinner than that of the previous device. The normalized response versus frequency for the APD is shown in Figure 8.20. The measurements were made using coplanar probes and

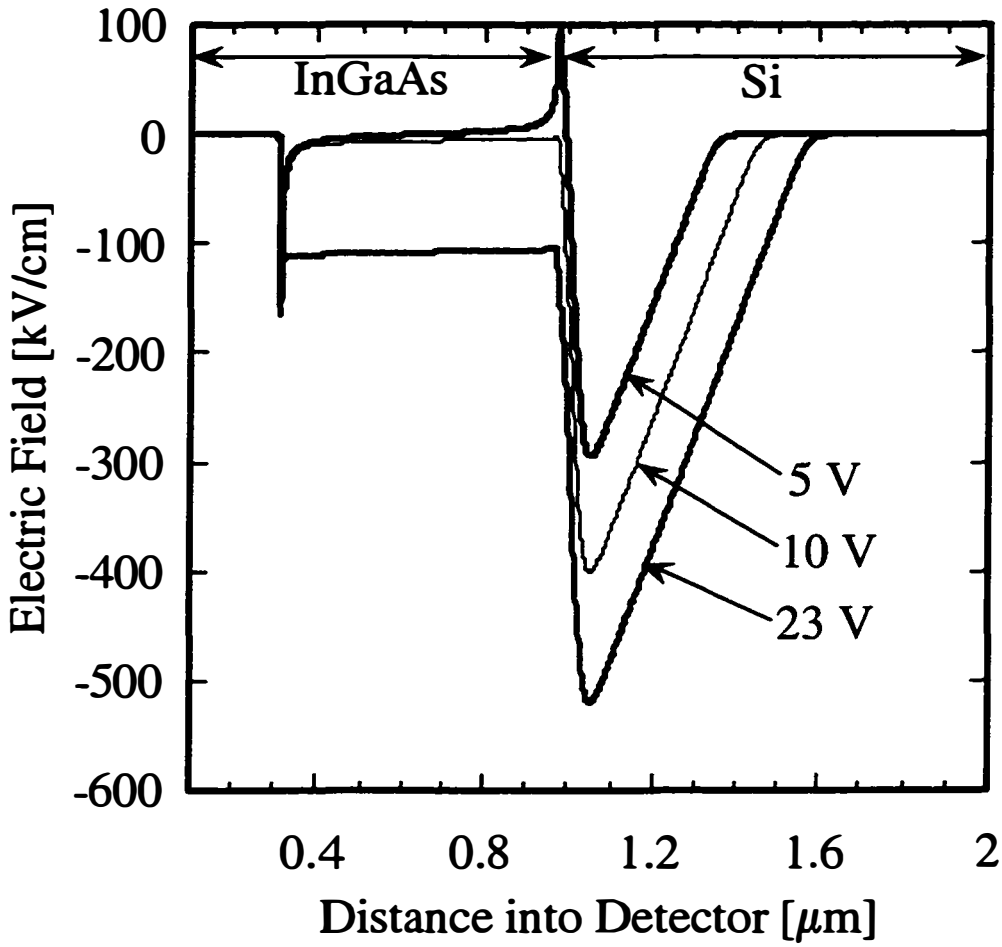


Figure 8.19 - Calculated electric field profiles in a third generation InGaAs-silicon APD at three different biases.

an HP Optical Component Analyzer with a range from 0.13 MHz to 20 GHz. The component analyzer used a 1.3 μm modulated laser as a source. The frequency response was relatively flat up to gains of around 20 where it began to drop slightly. At lower gains the 3-dB down frequency is about 13 GHz.

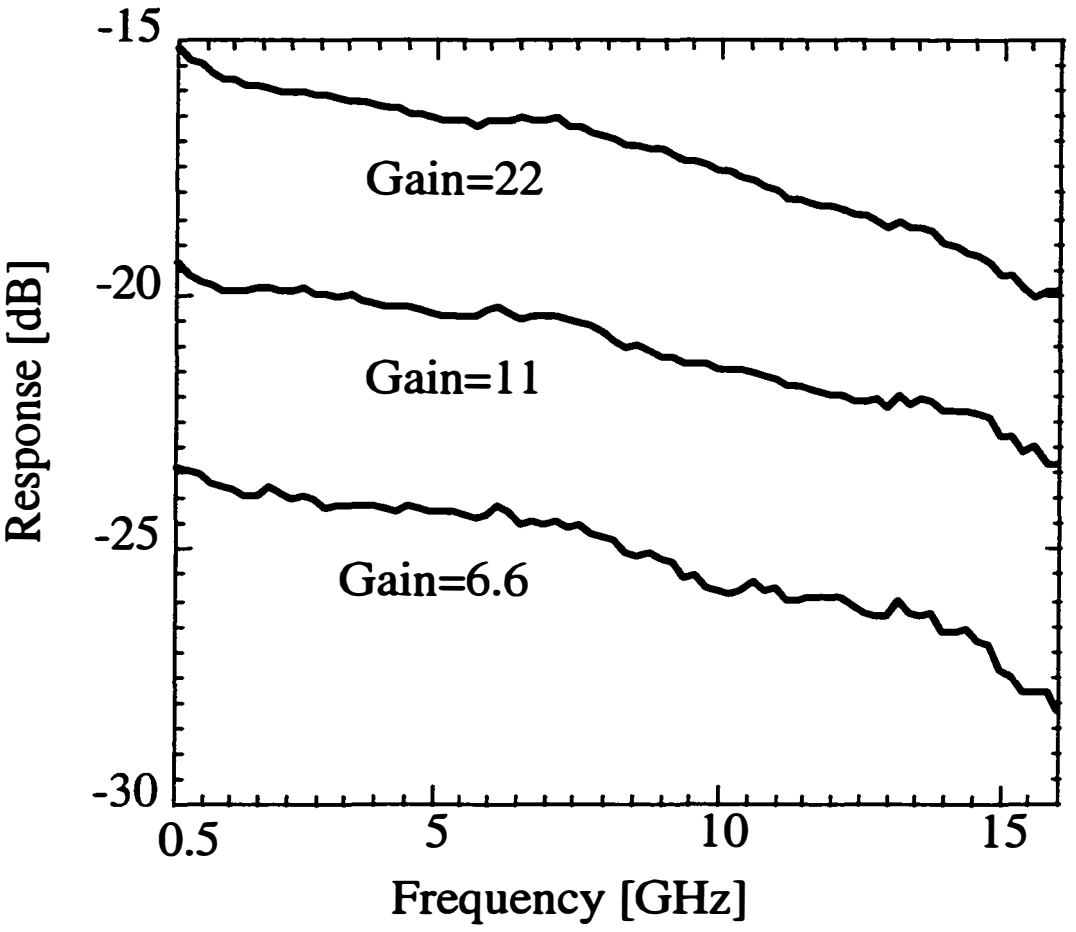


Figure 8.20 - Frequency response of a third generation InGaAs-silicon APD. Shown are the responses for the same detector biased at three different gain levels.

The 3 dB bandwidth as a function of gain is plotted in Figure 8.21. As can be seen, the bandwidth is approximately 13 GHz with gains up to 20 and decreases to 9 GHz at a gain of 35. Measurements of gains beyond this were

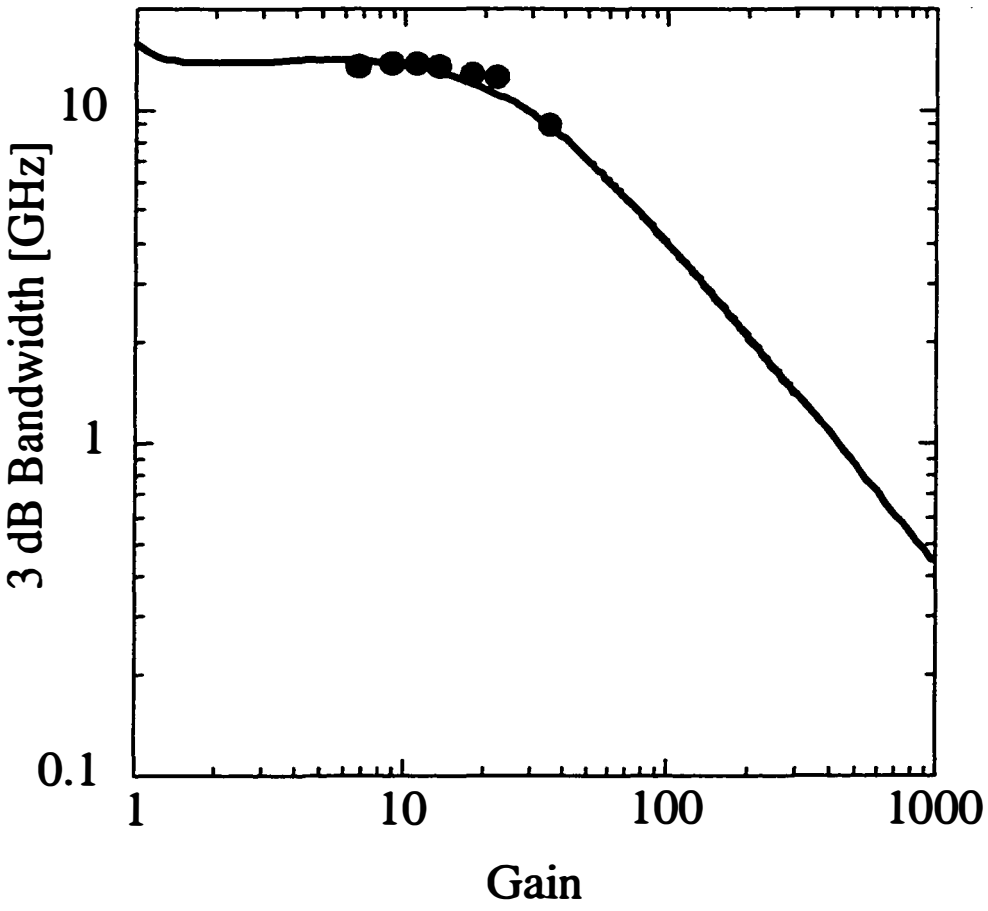


Figure 8.21 - 3-dB bandwidth versus multiplication gain for a 23 μm InGaAs-silicon APD with a 0.7 μm absorption layer and a 0.6 μm multiplication layer. Points indicate measurements, line indicates theoretical response based on Hollenhorst's method (Section 5.5).

inhibited by the rapid increase in dark current and thermal heating. The calculated dependence of bandwidth on gain[5] is also shown in Figure 8.21 and

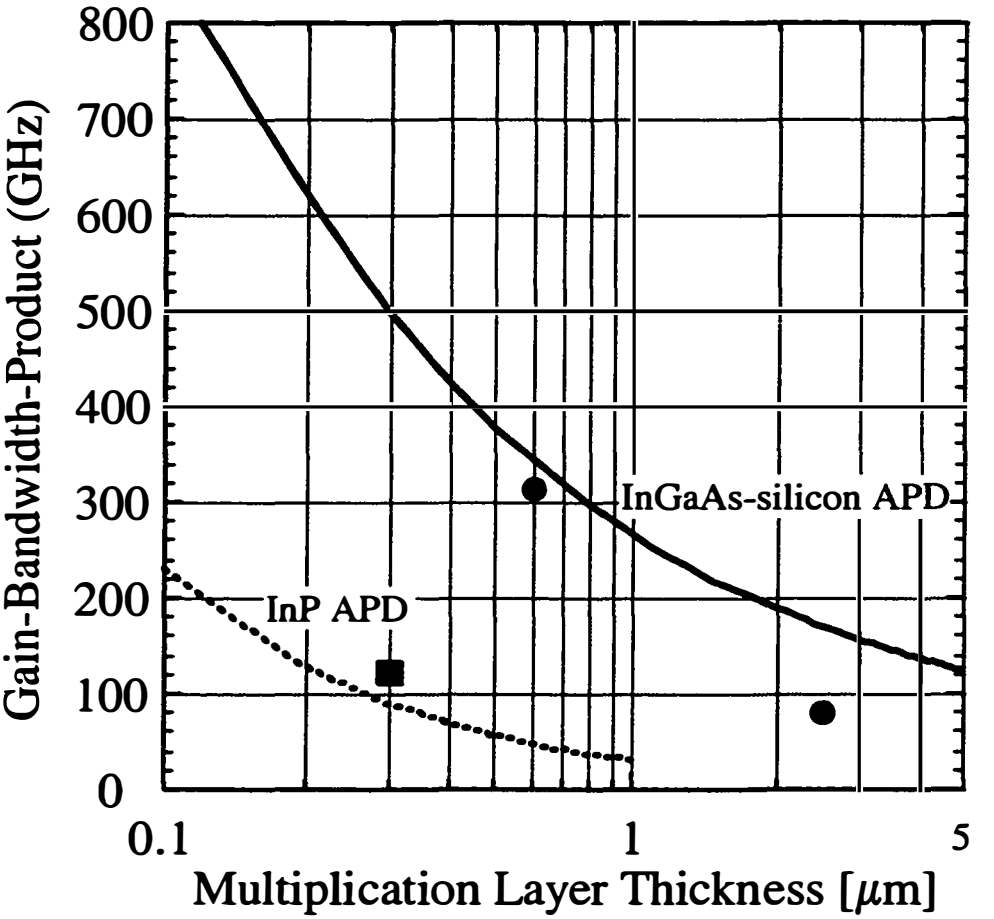


Figure 8.22 - Gain-bandwidth product versus multiplication thickness for two types of APDs. The square and dashed lines indicate highest measured gain bandwidth and theoretical performance, respectively, for InGaAs/InP. The circles and solid line represent measured values and theoretical prediction for InGaAs-silicon.

agrees well with the measurements. The highest measured gain-bandwidth product was 35 x 9 GHz, or 315 GHz. As illustrated by the figure, gain-bandwidths of around 350GHz should be possible when measurements at higher gains are achievable.

To compare the performance of the third generation detector to existing and predicted results for InGaAs/InP[6] APDs, Figure 8.22 plots the gain-bandwidth-product versus multiplication layer thickness. The highest reported gain-bandwidths for the two types of detector are indicated along with a curve predicting performance. As can be seen, the latest InGaAs-silicon device surpasses the gain-bandwidth of the best InGaAs/InP APD (120 GHz)[7] and indicates the potential for much greater performance as dark currents are decreased and multiplication layers are made thinner. Gain-bandwidth products of 600 GHz may be possible.

8.7 Conclusions

The construction of InGaAs-silicon APDs allowed for the first measurements of their electrical properties such as dark current, photo-response, and frequency response. The fabrication of these detectors was accomplished with standard semiconductor processing techniques except for the critical wafer fusion step used to bind the InGaAs and silicon layers together. The three generations of devices that were made all had measurable avalanche

gain and photoresponses that confirm one of the conclusions from the InGaAs-silicon PINs - the fused interface does not act as a barrier for electrons and holes. All of the APDs created had dark currents higher than desirable for integration into real system, but these seemed to be caused by the fabrication techniques rather than any intrinsic material properties. Frequency response measurements made on the most sophisticated APD design showed 3-dB down frequencies of 13 GHz - very fast for an APD. The measured gain-bandwidth-product of 315 GHz for this device was also the highest every reported in the near-infrared. All of the measurements answered questions about how well the device could perform, but also leave many remaining. What is the ultimate gain-bandwidth-product for this type of detector? How low can the dark current be made? The noise figure has been measured for an InGaAs-silicon PIN but what is it like for an APD? How do these detectors perform in actual transmission systems? Hopefully, these devices will serve to pave the way toward further research and the construction of devices that outperform any that exist today.

References

[1] A. R. Hawkins, T. Reynolds, D. R. England, D. I. Babic, M. Mondry, and J. E. Bowers, "Silicon Hetero-Interface Photodetector," LEOS '95 Annual Meeting, San Franscisco, Oct. 30-Nov. 2, (1995).

[2] A. R. Hawkins, T. E. Reynolds, D. R. England, D. I. Babic, M. J. Mondry, K. Streubel, and J. E. Bowers, "Silicon heterointerface photodetector," *Applied Physics Letters*, **68**, No. 26, pp. 3692-3694, (1996).

[3] A. R. Hawkins, W. Wu, P. Abraham, K. Streubel, and J. E. Bowers, "High gain-bandwidth-product silicon heterointerface photodetector," *Applied Physics Letters*, **70**, No. 3, pp. 303-305, (1997).

[4] Calculations done using computer codes VWF, ATLAS, and ATHENA (SILVACO International, Santa Clara, CA, 1995).

[5] W. Wu, A. R. Hawkins, and J.E. Bowers, "Frequency Response of Avalanche Photodetectors with Separate Absorption and Multiplication Layers," *Journal of Lightwave Technology*, **14**, No 12, pp. 2778-2785, (1996).

[6] L. E. Tarof, J. Yu, R. Bruce, D. G. Knight, T. Baird, and B. Oosterbrink, "High-Frequency Performance of Separate Absorption Grading, Charge, and Multiplication InP/InGaAs Avalanche Photodiodes," *IEEE Photonics Technology Letters*, **5**, pp. 672-674 (1993).

[7] L. E. Tarof, J. Yu, R. Bruce, D. G. Knight, T. Baird, and B. Oosterbrink, *IEEE Photonics Technology Letters*, **5**, pp. 672-674, (1993).

Chapter 9

Conclusions and Outlook

9.0 Conclusions

The first conclusion to be drawn from this dissertation is that silicon is the material of choice for the gain region in an avalanche photodiode. The reasons for this were explained qualitatively based on an elementary look at the avalanche multiplication process. Electrons in silicon are much more likely to ionize than holes leading to a very low ratio (k) between the hole and electron ionization coefficients. This creates a gain scenario that is faster and less random than that in semiconductors with k ratios closer to one. Four different aspects of the gain mechanism were linked to this k ratio. First the gain sensitivity to voltage and temperature were derived as a function of k. It was found that the more disparate the ionization coefficients the less sensitive the gain to voltage and temperature variations - very desirable when operating an APD in a system. The dependence of signal noise on the ratio k was also derived. Again semiconductors with k values much less or much greater than one produced less noise with increasing gain. Finally two methods of computing bandwidth were presented and calculations were made for APD structures based on the ionization coefficients of their multiplication layers. In

each of these areas, voltage sensitivity, temperature sensitivity, signal noise, and bandwidth, silicon was compared directly to InP and in each case silicon was shown to be clearly superior.

An obvious second conclusion is that InGaAs is the best material for the absorption region of an APD intended for use in fiber optic systems. Its high optical absorption coefficient at the 1.3 and 1.55 μm has made InGaAs the most desirable detector material at these wavelengths. The idea of combining InGaAs and silicon to take advantage of their properties follows naturally. One of the important requirements when working with APDs containing InGaAs is control of the electric field profile. Calculations made here show that if InGaAs and silicon can be successfully integrated, ion implantation can be used as a "doping knob" to tailor electric fields. The use of implantation makes an InGaAs-silicon APD very feasible to build in terms of the present processing technology.

A third conclusion is that wafer fusion provides a good interface between InGaAs and silicon. For years the lack of an effective integration method limited any attempts to combine III-V semiconductors and silicon to construct an APD. The junction created through wafer fusion is mechanically robust and provides excellent current transport properties. Studies of the interface found no signs of recombination centers or excess charge. High quantum efficiency PIN detectors made from fused InGaAs and silicon also had

large frequency bandwidths limited only by the RC time constants of the detector. It was also discovered that the proper fusion conditions are important to producing a high quality interface. Ideal conditions seem to be a pure hydrogen atmosphere at temperatures near the epitaxial growth temperatures of InGaAs. Lower temperatures and lower concentrations of hydrogen produced detectors with higher dark currents and what appeared to be a thin interface barrier.

Perhaps the most important conclusion to be drawn is that high performance APDs can be made from the InGaAs-silicon material system. Three generations of APDs were constructed and their characteristics measured. The first design was a very elementary one that served as a proof of concept. Photocurrent gains measured for this device showed that carrier injection from an InGaAs layer into a high field silicon layer was possible. The lessons learned from the first generation detector were applied to a second, more sophisticated design. High photocurrent gains were measured and the gain versus voltage curves showed the expected characteristics of a low k semiconductor. The first bandwidth measurements were also made on this device in spite of the fact that the design needed improvements in order to achieve very high-speed operation. This led to the third device generation, intended to push the bandwidth limits of the InGaAs-silicon system. This detector again exhibited high photocurrent gains and low voltage sensitivity as

expected for a silicon multiplication layer. Very high bandwidths of 13 GHz were also measured along with a record gain-bandwidth-product of 315 GHz.

9.1 Future Work

It could be said that for every research project the work is never finished. There is always some question that remains unanswered some property yet to be explored. The InGaAs-silicon APD is certainly one of those projects and the research done here has brought up as many new questions as answers it has provided. Some of the first questions that remain involve the nature of the fused InGaAs-silicon interface. From what has been measured it appears to be almost ideal for use in a detector with an undetectable offset in the conduction band. Is the conduction band all in the valance band? What exactly occurs as electrons transit from the direct bandgap semiconductor InGaAs into the indirect bandgap silicon? Several experiments could be done in an effort to quantify this interface. First of all a reversed PIN structure diode could be constructed with p-doped silicon and n-doped InGaAs. Reverse bias operation of this structure would demonstrate a valence band offset as holes would run into this barrier as they were swept towards the silicon layer. Further capacitance-voltage measurements could also be done, especially on this reverse PIN structure to try and pick up any valence band offset.

Another area that merits further effort is the dark current reduction in InGaAs-silicon APDs. Much of the dark current produced in the devices from Chapter 8 is due to the processing techniques used in their production. All of the APDs were made from dry etching, which produces material damage and excess current leakage. Less damaging processes need to be developed that are compatible with InGaAs as well as silicon. When the dark current is reduced sufficiently other detector characteristics can be measured more accurately, including noise and precise voltage and temperature sensitivities. Making these measurements on diodes with high dark currents is fairly difficult and can also produce misleading errors.

Another looming quantity that remains unknown for this APD is the device lifetime. This is of concern due to the fused interface so critical to the detector. Although the interface appears to be very good initially, it could very well deteriorate over time. APDs must operate under high electric fields and this could contribute to any interface degeneration. The only way to really know what happens to the interface is to do accelerated life testing at high temperatures and voltages. Again, having a low dark current device is important for this test as well.

One aspect of semiconductor device development that all new inventions must go through is the transition from being produced and measured in a research laboratory to being manufacturable. For InGaAs-silicon APDs to

be at all useful in the real world they must be produced reliably and in a state that is readily packaged. Many of the processes used to produce the APDs described here are low temperature in nature and do not lend themselves to either bump bonding packaging or even wire bond attachment. Higher temperature fabrication processes must be developed to allow integration of detector die with other electronics.

Creating APD die capable of being packaged also opens up a whole other set of measurements possible for the detector. They could then be integrated with a transimpedance amplifier in a receiver circuit. Receiver sensitivity could then be measured - the ultimate measure for an APD intended for optical communications. Operation of an APD receiver in a real communication system, using actual lengths of optical fiber, is also a future landmark.

Of course there is also plenty of work to be done on more basic aspects of the detector. Investigating the properties of different silicon multiplication layer thicknesses is one that comes to mind. An ultimate gain bandwidth product well over 800 GHz may be possible with enough effort concentrated on producing the "ultimate" thin avalanching layer in silicon. More useful work might actually come in investigating very thick silicon layers capable of producing very controllable gains with very low noise. For every data rate,

there is a silicon thickness that would optimize an InGaAs-silicon APD receiver's sensitivity, and a lot of work left to find each one.

9.2 Outlook

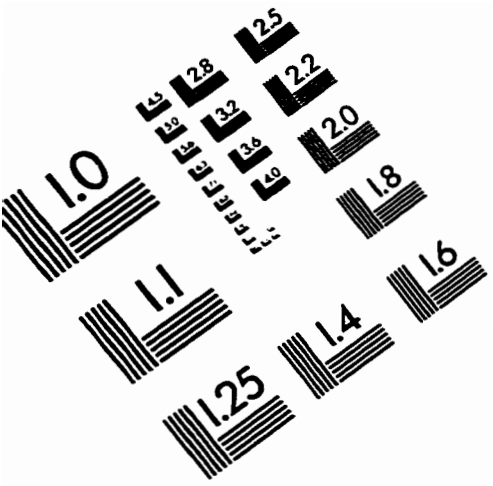
It is always difficult to predict what may happen with a new invention and if it will make any real impact. A few projections can be ventured upon, however, based upon an optimistic view of the InGaAs-silicon APD and assuming it can be manufactured and packaged successfully. The first area in which the device could have an impact is in existing communication systems, especially those operating at 2.5 GBit/s. These systems are presently using InGaAs/InP APD receivers at the end of long fiber runs. If the InGaAs-silicon combination can provide an increase in receiver sensitivity, as theoretically it should, longer lengths of fiber can be used before optical amplification is necessary. More important than this application may be access networks in which fiber rings connect nodes of transmitters and receivers. Higher sensitivity APDs translate into longer fiber runs between nodes and more nodes in a ring, without the huge cost of using optical amplifiers. This advantage could be applicable at 2.5 GBit/s as well as lower data rates.

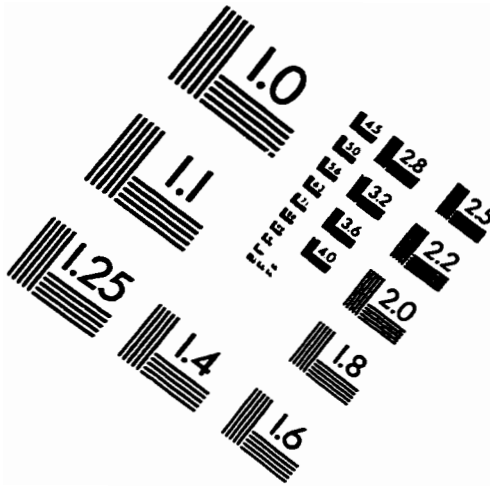
The application for which the InGaAs-silicon APD should really excel is in a receiver for 10 GBit/s. Systems operating at this speed are still in their prototyping stage and one of the real limitations designers are faced with is an

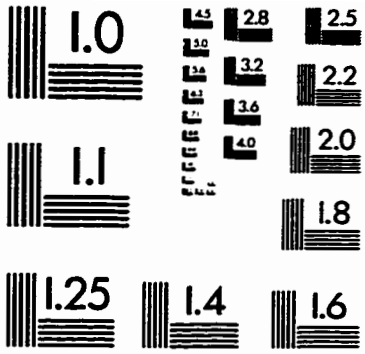
adequate receiver. It is extremely difficult to push APDs based on existing InGaAs/InP technology to work with any significant gains at 10 GBit/s. The avalanche build up times in these detectors is simply too great. Systems are now being put together without APDs, instead using PIN receivers and amplifiers - adding greatly to the cost. It has been demonstrated, however, that InGaAs-silicon can operate at the required high speed and maintain significant gains. If InGaAs-silicon APDs can be made commercially available, they would not only be a nice improvement, but an enabling technology for 10 GBit/s systems.

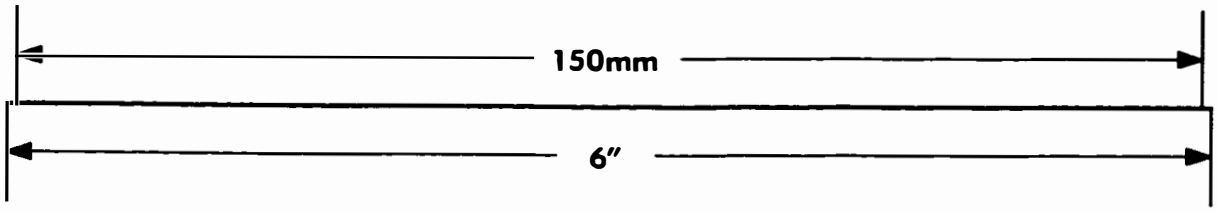
There are many other advantages that the use of the InGaAs-silicon system could provide. Infrared APDs could conceivably be made by fusing InGaAs to silicon wafers to pre-fabricated VLSI circuitry. This could lead to highly integrated APD chips with detector, amplifier, and even clock and data recovery electronics. Such a chip would be desirable because of its small size and also reduction in cost. APDs for applications other than optical communications are also a possibility with InGaAs-silicon. More sensitive detectors for things like LIDAR and medical imaging are constantly being sought. In these areas high speed is not as important as high photocurrent gains and low signal noise. Outside of the intended focus of this dissertation, there are also likely applications of the new InGaAs-silicon system that have nothing to do with light detection. Time and necessity will tell.

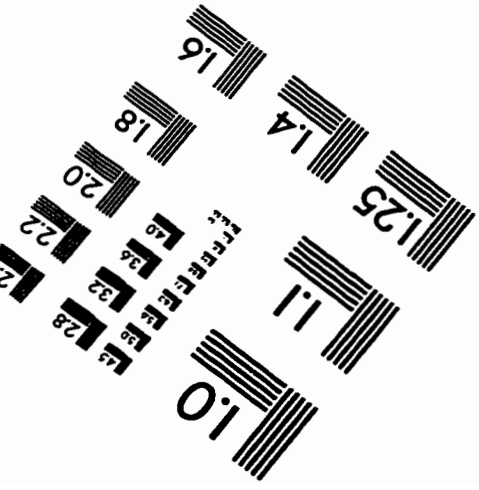
IMAGE EVALUATION
TEST TARGET (QA-3)

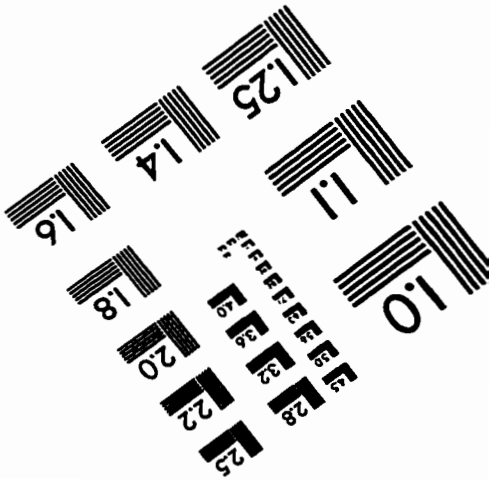












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989