FUNDAMENTALS OF

Infrared Detector Materials

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
Downloaded From: http://ebooks.spiedigitallibrary.org/ on 09/18/2013 Terms of Use: http://spiedl.org/terms

Tutorial Texts Series

- *Radiation Thermometry: Fundamentals and Applications in the Petrochemical Industry*, Peter Saunders, Vol. TT78
- *Matrix Methods for Optical Layout*, Gerhard Kloos, Vol. TT77
- *Fundamentals of Infrared Detector Materials*, Michael A. Kinch, Vol. TT76
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Third Edition*, Herbert Kaplan, Vol. TT75
- *Bioluminescence for Food and Environmental Microbiological Safety*, Lubov Y. Brovko, Vol. TT74
- *Introduction to Image Stabilization*, Scott W. Teare, Sergio R. Restaino, Vol. TT73
- *Logic-based Nonlinear Image Processing*, Stephen Marshall, Vol. TT72
- *The Physics and Engineering of Solid State Lasers*, Yehoshua Kalisky, Vol. TT71
- *Thermal Infrared Characterization of Ground Targets and Backgrounds, Second Edition*, Pieter A. Jacobs, Vol. TT70
- *Introduction to Confocal Fluorescence Microscopy*, Michiel Müller, Vol. TT69
- *Artificial Neural Networks: An Introduction*, Kevin L. Priddy and Paul E. Keller, Vol. TT68
- *Basics of Code Division Multiple Access (CDMA),* Raghuveer Rao and Sohail Dianat, Vol. TT67
- *Optical Imaging in Projection Microlithography*, Alfred Kwok-Kit Wong, Vol. TT66
- *Metrics for High-Quality Specular Surfaces*, Lionel R. Baker, Vol. TT65
- *Field Mathematics for Electromagnetics*, *Photonics*, *and Materials Science*, Bernard Maxum, Vol. TT64
- *High-Fidelity Medical Imaging Displays*, Aldo Badano, Michael J. Flynn, and Jerzy Kanicki, Vol. TT63
- *Diffractive Optics–Design*, *Fabrication*, *and Test*, Donald C. O'Shea, Thomas J. Suleski, Alan D. Kathman, and Dennis W. Prather, Vol. TT62
- *Fourier-Transform Spectroscopy Instrumentation Engineering*, Vidi Saptari, Vol. TT61
- *The Power- and Energy-Handling Capability of Optical Materials*, *Components*, *and Systems*, Roger M. Wood, Vol. TT60
- *Hands-on Morphological Image Processing*, Edward R. Dougherty, Roberto A. Lotufo, Vol. TT59
- *Integrated Optomechanical Analysis*, Keith B. Doyle, Victor L. Genberg, Gregory J. Michels, Vol. TT58
- *Thin-Film Design: Modulated Thickness and Other Stopband Design Methods*, Bruce Perilloux, Vol. TT57
- *Optische Grundlagen für Infrarotsysteme*, Max J. Riedl, Vol. TT56
- *An Engineering Introduction to Biotechnology*, J. Patrick Fitch, Vol. TT55
- *Image Performance in CRT Displays*, Kenneth Compton, Vol. TT54
- *Introduction to Laser Diode-Pumped Solid State Lasers*, Richard Scheps, Vol. TT53
- *Modulation Transfer Function in Optical and Electro-Optical Systems*, Glenn D. Boreman, Vol. TT52
- *Uncooled Thermal Imaging Arrays*, *Systems*, *and Applications*, Paul W. Kruse, Vol. TT51
- *Fundamentals of Antennas*, Christos G. Christodoulou and Parveen Wahid, Vol. TT50
- *Basics of Spectroscopy*, David W. Ball, Vol. TT49
- *Optical Design Fundamentals for Infrared Systems*, *Second Edition*, Max J. Riedl, Vol. TT48
- *Resolution Enhancement Techniques in Optical Lithography*, Alfred Kwok-Kit Wong, Vol. TT47
- *Copper Interconnect Technology*, Christoph Steinbrüchel and Barry L. Chin, Vol. TT46
- *Optical Design for Visual Systems*, Bruce H. Walker, Vol. TT45
- *Fundamentals of Contamination Control*, Alan C. Tribble, Vol. TT44
- *Evolutionary Computation: Principles and Practice for Signal Processing*, David Fogel, Vol. TT43
- *Infrared Optics and Zoom Lenses*, Allen Mann, Vol. TT42
- *Introduction to Adaptive Optics*, Robert K. Tyson, Vol. TT41
- *Fractal and Wavelet Image Compression Techniques*, Stephen Welstead, Vol. TT40
- *Analysis of Sampled Imaging Systems*, R. H. Vollmerhausen and R. G. Driggers, Vol. TT39
- *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*, Valery Tuchin, Vol. TT38
- *Fundamentos de Electro-Óptica para Ingenieros*, Glenn D. Boreman, translated by Javier Alda, Vol. TT37
- *Infrared Design Examples*, William L. Wolfe, Vol. TT36

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

FUNDAMENTALS OF

Infrared Detector Materials

Michael A. Kinch

Tutorial Texts in Optical Engineering
Volume TT76

SPIE
PRESS

Bellingham, Washington USA

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
Downloaded From: http://ebooks.spiedigitallibrary.org/ on 09/18/2013 Terms of Use: http://spiedl.org/terms

Library of Congress Cataloging-in-Publication Data

Kinch, M. A. (Michael A.)
Fundamentals of infrared detector materials / M.A. Kinch.
p. cm.
ISBN 978-0-8194-6731-7
1. Infrared detectors--Materials. I. Title.

TA1570.K575 2007
621.36'2--dc22

2007028531

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360 676 3290
Fax: +1 360 647 1445
Email: spie@spie.org
Web: SPIE.org

Copyright © 2007 The Society for Photo-optical Instrumentation Engineers

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

The content of this book reflects the work and thought of the author(s). Every effort has been made to publish reliable and accurate information herein, but the publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Printed in the United States of America.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Introduction to the Series

Since its conception in 1989, the Tutorial Texts series has grown to more than 70 titles covering many diverse fields of science and engineering. When the series was started, the goal of the series was to provide a way to make the material presented in SPIE short courses available to those who could not attend, and to provide a reference text for those who could. Many of the texts in this series are generated from notes that were presented during these short courses. But as stand-alone documents, short course notes do not generally serve the student or reader well. Short course notes typically are developed on the assumption that supporting material will be presented verbally to complement the notes, which are generally written in summary form to highlight key technical topics and therefore are not intended as stand-alone documents. Additionally, the figures, tables, and other graphically formatted information accompanying the notes require the further explanation given during the instructor's lecture. Thus, by adding the appropriate detail presented during the lecture, the course material can be read and used independently in a tutorial fashion.

What separates the books in this series from other technical monographs and textbooks is the way in which the material is presented. To keep in line with the tutorial nature of the series, many of the topics presented in these texts are followed by detailed examples that further explain the concepts presented. Many pictures and illustrations are included with each text and, where appropriate, tabular reference data are also included.

The topics within the series have grown from the initial areas of geometrical optics, optical detectors, and image processing to include the emerging fields of nanotechnology, biomedical optics, and micromachining. When a proposal for a text is received, each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that would serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the processes and technologies under discussion.

It is my goal to maintain the style and quality of books in the series, and to further expand the topic areas to include new emerging fields as they become of interest to our reading audience.

Arthur R. Weeks, Jr.
University of Central Florida

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

To

Rita and Betty

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Contents

1 Introduction **1**

2 IR Detector Performance Criteria **5**
2.1 Photon Detectors 5
2.1.1 IR detector operating temperature 5
2.1.2 IR detector sensitivity 7
2.2 Thermal Detectors 9

3 IR Detector Materials: A Technology Comparison **13**
3.1 Intrinsic Direct Bandgap Semiconductor 13
3.2 Extrinsic Semiconductor 16
3.3 Quantum Well IR Photodetectors (QWIPs) 18
3.4 Silicon Schottky Barrier Detectors 23
3.5 High-Temperature Superconductor 26
3.6 Conclusions 27

4 Intrinsic Direct Bandgap Semiconductors **31**
4.1 Minority Carrier Lifetime 32
4.1.1 Radiative recombination 32
4.1.2 Auger recombination 33
4.1.3 Shockley–Read recombination 34
4.2 Diode Dark Current Models 34
4.3 Binary Compounds 35
4.3.1 Indium antimonide: InSb 35
4.4 Ternary Alloys 37
4.4.1 Mercury cadmium telluride: $Hg_{1-x}Cd_xTe$ 37
4.5 $Pb_{1-x}Sn_xTe$ 42
4.5.1 Minority carrier lifetime 43
4.5.2 Dark currents 44
4.6 Type III Superlattices 45
4.6.1 Superlattice bandstructure 45
4.6.2 Band offsets and strain 47
4.6.3 Interdiffusion in HgTe/CdTe superlattices 48
4.6.4 Misfit dislocations 48
4.6.5 Absorption coefficient 49

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

4.6.6 Effective mass 51
4.6.7 Minority carrier lifetime 52
4.7 Type II Superlattices 53
4.7.1 Minority carrier lifetime 54
4.8 Direct Bandgap Materials: Conclusions 57
4.8.1 HgCdTe 57
4.8.2 InSb 57
4.8.3 PbSnTe 58
4.8.4 Type III superlattices 59
4.8.5 Type II superlattices 59
4.8.6 Final thoughts 59

5 HgCdTe: Material of Choice for Tactical Systems 61
5.1 HgCdTe Material Properties 61
5.1.1 Material growth 61
5.1.2 HgCdTe annealing 65
5.1.3 HgCdTe properties 67
5.2 HgCdTe Device Architectures 75
5.2.1 DLHJ architecture 76
5.2.2 Bump-bonded ion implant architecture 77
5.2.3 Vertically integrated photodiode (VIP and HDVIP) architectures 77
5.3 ROIC Requirements 81
5.3.1 Detector performance: Modeling 82
5.3.2 Dark current in HgCdTe diodes 82
5.3.3 $1/f$ noise 87
5.4 Detector Performance 89
5.5 HgCdTe: Conclusions 91

6 Uncooled Detection 93
6.1 Thermal Detection 93
6.2 Photon Detection 95
6.2.1 HOT detector theory 95
6.2.2 HOT detector data 101
6.2.3 HOT detector contacts 103
6.2.4 HOT detector options 103
6.3 Uncooled Photon vs. Thermal Detection Limits 105
6.4 Uncooled Detection: Conclusions 107

7 HgCdTe Electron Avalanche Photodiodes (EAPDs) 109
7.1 McIntyre's Avalanche Photodiode Model 110
7.2 Physics of HgCdTe EAPDs 112
7.2.1 High-energy scattering rates 113
7.2.2 Electron impact ionization rate in HgCdTe 115

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

7.3 Empirical Model for Electron Avalanche Gain in HgCdTe 121
7.4 Room-Temperature HgCdTe APD Performance 129
7.5 Monte Carlo Modeling 131
7.6 Conclusions 133

8 Future HgCdTe Developments 135
8.1 Dark Current Model 135
8.1.1 N-side 136
8.1.2 P-side 137
8.2 The Separate Absorption and Detection Diode Structure 139
8.3 Multicolor and Multispectral FPAs 141
8.4 High-Density FPAs 143
8.5 Low Background Operation 143
8.5.1 LWIR 14 μm at 40 K 143
8.5.2 Low background operation at a cutoff of 25 μm 144
8.6 Higher Operating Temperatures 145
8.6.1 High-gain APDs 147
8.7 Conclusion 148

Epilogue 149

Appendix A: Mathcad Program for HgCdTe Diode Dark Current Modeling 151

References 165

About the Author 169

Index 171

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 1

Introduction

The choice of available infrared (IR) detectors for insertion into modern IR systems is both large and confusing. The purpose of this volume is to provide a technical database from which rational IR detector selection criteria can evolve and thus clarify the options open to the modern IR system designer. Emphasis will be mainly on high-performance IR systems operating in a tactical environment, although there will be limited discussion of both strategic environments and low- to medium-performance system requirements.

Early IR imaging systems utilized extrinsically doped Ge as the detecting material and operated at 28 K. However, the development in the 1960s of the semiconductor alloys HgCdTe [1] and PbSnTe [2], with their tunable bandgaps that covered the complete IR spectrum from 1 to 20 μm, led to the birth of the first generation of modern high-performance IR systems in the early 1970s, with the advent of the so-called Common Module, first developed by Texas Instruments. The heart of this system was a simple 180-element linear parallel-scan HgCdTe photoconductive array mounted on a cold finger operating at approximately 77 K. The IR image formed at the focal plane of the system was scanned across the simple linear array of detectors by a rotating mirror, thus generating one line of the IR scene at a time, with an available integration time (or noise bandwidth) determined by the system line time, and the image was formatted by the subsequent off-focal plane electronics. This first-generation imaging system was thus mechanically complex, but geometrically simple from a detector, and hence array producibility, point of view. The array bias and amplifier electronics were mounted off the cold finger at an elevated temperature. HgCdTe had proved to be the material of choice at this time because of issues with PbSnTe associated with its high dielectric constant and inferior mechanical properties. The use of the direct bandgap alloy HgCdTe resulted in high absorption coefficients for the IR, and for the first time enabled the use of thin material, and hence the deployment of standard semiconductor photolithographic techniques in IR focal plane array fabrication. The Common Module has survived in one form or another for some 30 years and is still in production today.

The 1990s fostered the evolution of the second generation of IR systems, in which the detector bias and signal electronics were incorporated onto the cold finger itself. This was typically achieved by hybridizing the IR-sensitive material with the

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

silicon processor. Scanning arrays, with their simple linear geometry and relative mechanical complexity, still dominate the marketplace. The degree of detector complexity is increasing, as ever-higher performance is demanded of the scanning IR system. This is typically achieved by increasing the number of elements in the cross-scan direction of the focal plane array (FPA), and by the incorporation in the scan direction of time, delay, and integration (TDI) techniques, which essentially result in a linear array with an increased integration time equal to integral multiples of the available line time. HgCdTe is still the unambiguous material of choice and indeed has no rival in the scanning FPA marketplace. However, the ever-increasing complexity of second-generation scanning FPAs has in turn led to the development of staring FPAs of increasing size and density, particularly for applications in which mechanical complexity cannot be tolerated. The mechanical complexity of the scanning array is replaced in the staring array by a complexity of design in both the IR device and the associated hybridized silicon electronics. It is this aspect of second-generation FPAs that has led to the most concern regarding the possible deployment of alternative detector material technologies, other than HgCdTe, for various, specific staring-system applications. This concern stems from the perception, which is in many cases mistaken, that materials and processing issues prohibit the use of HgCdTe in the fabrication of production quantities of large-area FPAs in which the IR material is hybridized to the silicon readout integrated circuit (ROIC).

Third-generation high-performance IR FPAs are already a gleam in the eye of the IR system designer, and will almost certainly be staring architectures. They will demand capabilities such as multicolor (at least two, and maybe three or four different spectral bands) with possibly simultaneous detection in both space and time, ever larger array sizes, of say 2000×2000, and operation at higher temperatures, even to room temperature, for all cutoff wavelengths, to name but the obvious.

A successful IR materials technology must be capable of addressing any and all of these second- and third-generation IR focal plane issues. With this in mind, the current arsenal of IR materials technologies is examined from a strictly fundamental standpoint, in which the materials and device physics of each situation are allowed to determine the limitations of possible system implementation. This methodology allows for each of the IR technologies to be considered strictly on its fundamental merits, with no diversions provided by issues associated with a suitable reduction of the technology to practice. Problems, such as nonideal contacts or surfaces, obviously will exist and exacerbate the situation, but these problems will only serve to confuse a meaningful fundamental comparison of these technologies.

This book is a byproduct of the many years of experience in IR technology acquired by the author, primarily while employed at the original Texas Instruments Central Research Laboratories in Dallas, Texas. The author would like to acknowledge the expertise of all the scientists that worked there during his tenure, and the significant role that they played in the development of what has become a critical technology in the defense of the United States. Giants walked those corridors in the form of George Heilmeyer, Bob Stratton, Dennis Buss, Dick Chapman, Dick Reynolds, Al Tasch, Roland Haitz, Bob Bate, Jerry Hynecek, Seb Borrello, Bill Breazeale, Grady Roberts, Kent Carson, Herb Schaake, John Tregilgas, Francois Padovani, Larry Hornbeck, Don Shaw, Dean Collins, Bill Wisseman, and Turner

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Hasty, to name but a few. In many ways, this book is a tribute to the stimulating environment provided by their talent and innovation, in this and many other fields. The author feels privileged to have belonged to that community. Unfortunately, the Texas Instruments Central Research Laboratory, and many like it, no longer exist on the technology playing fields of the United States. The country is much the poorer for their demise.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 2

IR Detector Performance Criteria

IR detectors fall into two broad categories: photon and thermal. In the classical photon detector, photons are absorbed and generate free carriers, which are sensed by an electronic readout circuit. The thermal detector, on the other hand, is a power law detector. The incoming IR photons are absorbed by a thermally isolated detector element, resulting in an increase in the temperature of the element, which is sensed by monitoring a parameter such as resistivity or dielectric constant. The majority of this book is concerned with photon detectors, because of the somewhat more complex and varied materials physics associated with each of the various formats. However, comparisons are made in Chapter 6 between the ultimate forms of both photon and thermal detection.

2.1 Photon Detectors

2.1.1 IR detector operating temperature

There are two basic types of photon detector: the majority carrier and the minority carrier. If the dominant carrier is the majority carrier, then the sensing is photoconductive in nature. For minority carrier devices, both photoconductive and photovoltaic modes of detection can be utilized. Two parameters are typically used to compare the performance of different materials systems and device architectures, namely responsivity and noise. Responsivity is defined as the voltage, or current, signal response of the device to a stimulus of input signal power in the form of a known photon flux. Noise is the inherent fluctuation in voltage, or current, of the device measured on the same node as the responsivity. The primary criterion for responsivity is that it be large enough to elevate the detector noise floor above that of the signal processing system noise (i.e., preamplifier, integrating capacitor, etc.). If the detector node is considered to be a simple integrator, then the noise on that node is determined by the fluctuation of the integrated number of carriers. These carriers can be generated both thermally within the detector and by the incident background flux of IR radiation that is absorbed by the

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

detector material. The variance in the number of carriers on the detector node is given by

$$\left\langle \Delta N^2 \right\rangle = N_{\text{th}} + N_{\Phi}, \tag{2.1}$$

where N_{th} is the number of thermally generated carriers, and N_{Φ} is the number generated by the incident background. Equation (2.1) is self-evident for the majority carrier photoconductor or the minority carrier photodiode and is also true for the minority carrier photoconductor, provided that the integration of minority carriers alone is considered. The noise in the minority carrier photoconductor is due to minority carrier fluctuations, even though the resistivity is controlled by majority carriers, as was first pointed out by Long [3] using two-level fluctuation theory.

The ultimate in detector performance is achieved when the noise generated by the system background flux is larger than any thermally generated noise within the detector. The detector is then said to exhibit background limited performance (BLIP). For an incident background flux, Φ_{B} photons/cm^2/s, the photon generated density of carriers is $n_{\Phi} = \eta_{\text{a}}\Phi_{\text{B}}\tau/t$, where η_{a} is the absorption quantum efficiency, τ is the lifetime of the relevant carrier, and t is the detector thickness. Thus, BLIP is achieved when $\eta_{\text{a}}\Phi_{\text{B}}\tau/t > n_{\text{th}}$. Rearranging, we have for the BLIP requirement

$$\eta_{\text{a}}\Phi_{\text{B}} > \frac{n_{\text{th}}t}{\tau} = G_{\text{th}}; \tag{2.2}$$

that is, the photon generation rate per unit area needs to be greater than $n_{\text{th}}t/\tau$, the thermal generation rate per unit area. For $\eta_{\text{a}} \approx \alpha t$, where α is the material absorption coefficient, we have a requirement for BLIP of $\Phi_{\text{B}} > n_{\text{th}}/\alpha\tau$. The normalized thermal generation rate is thus defined as

$$G^*_{\text{th}} = \frac{n_{\text{th}}}{\alpha\tau}, \tag{2.3}$$

and it can be utilized to unambiguously predict the ultimate performance of any IR material and to compare the relative performance of different IR materials as a function of temperature. The only requirement is a knowledge of the dependence of n_{th}, α, and τ on temperature. The normalized thermal generation rate, G^*_{th}, given by Eq. (2.3), thus determines the BLIP operating temperature of any IR material, that is, the temperature below which the IR material has the capability to become background limited, and provides a precise vehicle for comparing different IR materials technologies. For $\alpha_{\lambda}t \gg 1$, $\eta_{\text{a}} = 1$ and maximum absorption efficiency, and hence detector sensitivity, is achieved at the expense of some penalty in increased dark current, and hence required lower operating temperature for BLIP.

The thermal generation rate, $G_{\text{th}} = n_{\text{th}}t/\tau$, represents the rate of return to equilibrium for any small departure of the number of carriers from their equilibrium value. Alternatively, it represents the rate at which carriers are generated under nonequilibrium conditions in the material that exist, for instance, in a reverse-biased diode within a diffusion length of the junction. Thus, $G_{\text{th}}q$, where q is the electronic charge, represents the ultimate limit of dark current flowing in a diode when excess

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

current sources such as tunnel and thermal generation within the diode depletion region are ignored. These phenomena are typically associated with defects in the bandgap, together with inappropriate semiconductor doping concentrations, and are considered avoidable in a suitably fabricated detector. They only serve to confuse any attempt at a fundamental comparison of different materials technologies and are ignored in that comparison.

2.1.2 IR detector sensitivity

The above discussion indicates the importance of absorption quantum efficiency in any comparison of IR materials. The lower the value of absorption quantum efficiency for a specific detector thickness, the lower the temperature required to achieve BLIP. However, the overall quantum efficiency of any IR device is composed of two factors: (1) the efficiency with which the incident radiation is absorbed, η_a, and (2) the efficiency with which the photogenerated carriers are detected, η_d, in the output circuit of the device. In a well-designed IR device, η_d is essentially unity. In a photodiode, for instance, good device design will ensure that photons will be absorbed within a diffusion length of the junction and hence detected in the external circuit. However, this is not always possible, as will be seen in the discussion of the internal quantum efficiency of silicon Schottky barrier diodes in Chapter 3. It is the absorption quantum efficiency that is relevant in Eq. (2.2), and that determines the BLIP temperature of operation of the IR material; however, it is the overall quantum efficiency that determines the maximum achievable device sensitivity.

The sensitivity of an IR detector can be expressed in a variety of ways. However, the limiting factor in all of these is the fluctuation in the relevant carrier concentration, as this determines the minimum observable signal of the detector. Again, this fluctuation is most easily envisioned by assuming that the detector current is integrated on a capacitive node, giving a variance on the node of

$$N^{1/2} = g[(I_d + I_\Phi)\tau/q]^{1/2}, \tag{2.4}$$

where g is the gain, I_d is the dark current, $I_\Phi = \eta\Phi_B A_d$ is the background flux current, and τ is the integration time (A_d is the detector area).

The minimum observable signal flux, $\Delta\Phi$, is thus given by $N^{1/2} = g\Delta\Phi\eta\tau A_d$, where in this case η represents the overall quantum efficiency of the detector, including the internal quantum efficiency. The optics transmission and cold shield efficiency are assumed to be unity. Thus, the noise equivalent flux is given by

$$\mathrm{NE}\Delta\Phi = g(1 + I_d/I_\Phi)\Phi_B/N^{1/2}. \tag{2.5}$$

This can be converted to a noise equivalent irradiance (NEI), which is defined as the minimum observable flux power incident on the system aperture, by renormalizing

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

the incident flux density on the detector to the system aperture area A_{opt}. The noise equivalent irradiance is given by

$$\text{NEI} = \text{NE}\Delta\Phi[A_{\text{d}}h\nu/A_{\text{opt}}], \tag{2.6}$$

where we have assumed monochromatic radiation of energy $h\nu$.

The sensitivity can also be expressed as a minimum temperature difference by considering that $\Delta\Phi = \Delta T[d\Phi_{\text{B}}/dT]$. Substituting in Eq. (2.5), we obtain for the noise equivalent temperature of the detector

$$\text{NE}\Delta\text{T} = \frac{g\left(1 + \dfrac{I_{\text{d}}}{I_{\Phi}}\right)}{N^{1/2}\left(\dfrac{1}{\Phi_{\text{B}}}\dfrac{d\Phi_{\text{B}}}{dT}\right)} = \frac{g}{N^{1/2}C}\left(1 + \frac{I_{\text{d}}}{I_{\Phi}}\right), \tag{2.7}$$

where $(d\Phi_{\text{B}}/dT)/\Phi_{\text{B}} = C$, the scene contrast. For the 8–12 μm spectral band, $C = 1.75 \times 10^{-2}$/K. For $I_{\Phi} > I_{\text{d}}$, and no ROIC noise, the required node capacity to achieve NE$\Delta T = 0.03$ K is $N = 4 \times 10^{6}$, which is well within the realm of available node capacities for current CMOS readout IC designs. Indeed, effective values for N are typically in the range 1×10^{7} to 2×10^{8} electrons for a 1 mil (=25 μm) pixel design. It is apparent from Eq. (2.7) that the larger the value of I_{d} relative to I_{Φ}, and/or the smaller the value of η, the more integration time will be required and the faster the speed of the optics, to achieve an NE$\Delta T = 0.3$. Thus, inefficient detectors can be utilized in faster optics and slower frame rate systems, particularly if temperature of operation and cooldown time are not an issue.

For unlimited well capacity, $N^{1/2}$ varies as g, and the parameters $\Delta\Phi$ and NEΔT are independent of the gain, g. However, for a limited maximum well capacity N, both $\Delta\Phi$ and NEΔT will vary as g. This effect can be thought of as a bandwidth issue.

All of the above parameters contain specific system and detector quantities such as area and bandwidth, where bandwidth is defined as $\Delta f = 1/2\tau$. These quantities can be eliminated by normalization, resulting in the detector parameter known as detectivity, D*, which is defined as [signal/flux]$A_{\text{d}}^{1/2}$/[noise/$\Delta f^{1/2}$]. Thus,

$$D^{*} = \frac{\eta}{2h\nu}\left(\frac{1}{G_{\text{th}} + \Phi_{\text{B}}}\right)^{1/2}. \tag{2.8}$$

If the background flux is negligible, then the D* expression in Eq. (2.8) depends entirely on the properties of the IR material. The maximum available D* at any temperature is directly related to the thermal generation rate per unit area of the material and is in fact inversely proportional to the square root of that parameter. It is also to be noted that the device sensitivity is linearly dependent on the device quantum efficiency, η ($=\eta_{\text{a}}\eta_{\text{d}}$). Thus, the IR material quantum efficiency plays a role, not only in determining the BLIP temperature of operation, but also in the ultimate sensitivity of the device. It should be pointed out here that device

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

sensitivity is perhaps more important in scanning IR FPAs, in which the available system integration time is limited to a horizontal line time, or some multiple thereof, by the use of TDI. In staring systems, the available integration time, particularly for tactical LWIR, is sufficiently long, even for high frame-rate systems, so as to provide exceedingly high performance at relatively modest device quantum efficiencies, as indicated by the discussion above on NEΔT.

2.2 Thermal Detectors

Thermal detectors are finding increasing use in IR imaging systems in the form of large-area staring FPAs. It is of interest to compare the potential performance of these arrays with photon detector-based architectures. For this purpose it is meaningful to consider the noise equivalent temperature (NEΔT) associated with the thermal detector.

The simple thermal detector consists of a detector element, of heat capacity C_{th}, thermally coupled to a heat sink by a thermal conductance G. In most modern versions of the thermal detector, the element is micromachined onto the surface of a silicon ROIC, and the resulting pedestal is connected to the ROIC by thin metal bus leads. The pedestal can be fashioned from any one of a number of temperature-sensitive materials that are process compatible with silicon ROIC processing. Current examples are vanadium oxide, α-silicon, and thin-film ferroelectrics. For the purposes of comparison with photon detectors, the exact nature of the detector element will be ignored here. We will assume that the temperature sensitivity of the element is sufficiently large that all noise sources other than temperature fluctuations can be ignored. This is not strictly true for the modern-day thermal detector, but it will give us an indication of the ultimate limit in thermal detector performance.

The power fluctuation of the element is given by

$$\Delta W_{\mathrm{n}} = [4kT^2G\Delta f]^{1/2}. \tag{2.9}$$

Thus, the detectivity, $D*$, of the detector is again defined as

$$\begin{aligned} D^* &= [\text{signal/flux power}][A_{\mathrm{d}}^{1/2}/(\text{noise}/\Delta f^{1/2})] \\ &= [\eta^2 A_{\mathrm{d}}/(4kT^2G)]^{1/2}. \end{aligned} \tag{2.10}$$

If G is radiatively limited, then $G_{\mathrm{rad}} = 4\sigma_{\mathrm{B}}T^3A_{\mathrm{d}} = 4PA_{\mathrm{d}}/T$, where $P = \sigma_{\mathrm{B}}T^4$, and σ_{B} is the Stefan–Boltzmann constant. Hence, $D^* \approx 2 \times 10^{10}\,\mathrm{cm\,Hz^{1/2}/W}$ and is independent of area.

This value of $D*$ is impressive but provides only a limited insight into the suitability of the thermal detector for specific IR systems. Signal and spectral bandwidth requirements are important, and to this end it is meaningful to consider the available noise equivalent temperature difference, NEΔT.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The limiting temperature fluctuation of the thermal detector associated with the power fluctuation ΔW_n is given by

$$\Delta T_n = [4kT^2 R_{th} \Delta f]^{1/2}, \tag{2.11}$$

where $R_{th} = 1/G$. Correlated double sampling [4] of this bandwidth-limited temperature fluctuation, with an integration time τ, gives

$$\Delta T_n = \{2kT^2[1 - \exp(-\tau/R_{th}C_{th})]/C_{th}\}^{1/2}, \tag{2.12}$$

where we have assumed $\Delta f = 1/4R_{th}C_{th}$.

Equating this to a signal temperature change for a given $F/\#$ (= optics focal length/diameter),

$$\Delta T_s = R_{th}(dP/dT)\Delta T[1 - \exp(-\tau/R_{th}C_{th})]/(1 + 4F^2), \tag{2.13}$$

gives

$$\text{NE}\Delta T = \{2kT^2/C_{th}/[1 - \exp(-\tau/R_{th}C_{th})]\}^{1/2}(1 + 4F^2)(dP/dT)/R_{th}, \tag{2.14}$$

where dP/dT is the differential change in radiated power per scene temperature change in the spectral region of interest. Image smearing considerations require $\tau \approx 2R_{th}C_{th}$, where the maximum value of R_{th} is given by radiative coupling. For a required system integration time, at the radiative limit, the only independent variable is the heat capacity, C_{th}. However, for micromachined resonant structures, there is a finite limit on thickness and specific heat that can be utilized. The available NEΔT for the perfect thermal detector, operating in a 300 K environment, as a function of integration time, is shown in Fig. 2.1, for both LWIR and MWIR spectral bands.

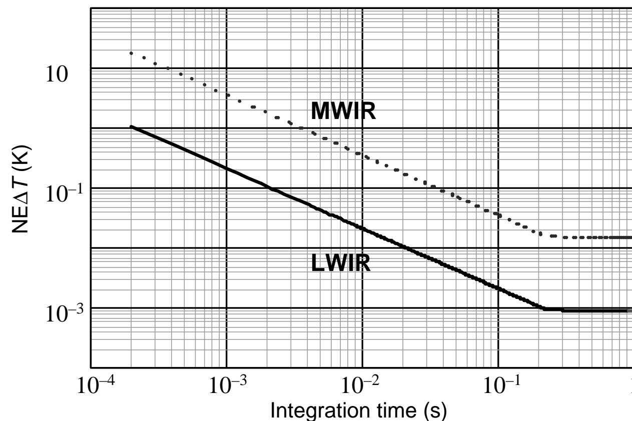


Figure 2.1 NEΔT limit vs. integration time for LWIR and MWIR thermal detection.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The assumed thickness, t, is 3000 Å, typical of today's FPAs, with $F/1$ optics, a 1 mil pixel size, and a volumetric specific heat of $2\,\mathrm{J/cm^3/K}$. The upper limit of R_{th}, due to radiative coupling, is 4×10^8 K/W for this pixel size. It is apparent that, for long integration times in the LWIR, excellent performance is possible, with NEΔT values $<$10 mK for frame rates of 30 Hz. However, for snapshot systems, with $\tau < 2$ ms, the available NEΔT is $>$100 mK, even for LWIR. For the MWIR the thermal detector has obvious performance limitations at any frame rate. For slower optics the situation is exacerbated, varying as $(F/\#)^2$. Thinner detector elements improve the NEΔT as $t^{1/2}$.

It should be emphasized again that the preceding arguments apply to the perfect thermal detector that is limited by thermal fluctuations of the detector pedestal and represent ultimate performance. The thermal detector technologies that most closely approach this desired performance are vanadium oxide, α-silicon, and various thin-film ferroelectrics. α-Silicon is attractive because of its processing compatibility with the Si ROIC and its processing versatility with regard to ultimate pedestal thickness. Thin-film ferroelectrics offer compatibility with first-generation thick-film ferroelectrics, together with relative immunity to $1/f$ noise and response nonuniformity issues, by virtue of the use of a mechanical chopper.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 3

IR Detector Materials: A Technology Comparison

In this chapter the following materials systems will be evaluated and compared as potential IR photon detectors, utilizing the normalized thermal generation rate parameter discussed in Chapter 2 as the performance differentiator:

1. Intrinsic direct bandgap semiconductors,
2. Extrinsic semiconductors,
3. IR quantum wells (QWIPs),
4. Silicon Schottky barriers, and
5. High-temperature superconductors.

This comparison will concentrate on the two spectral bands of prime importance to tactical IR systems, that is, the atmospheric transmission windows at 3–5 μm and 8–12 μm. The list is not meant to be exhaustive, but it does represent the primary materials candidates in vogue at the present time. It should be emphasized that any materials system proposed in the future for potential IR detection purposes, such as Type II superlattices, could also be easily compared to the current group by a simple consideration of its thermal generation rate, provided values of the relevant carrier concentration and lifetime as a function of temperature, together with the absorption quantum efficiency, are known. The high-temperature superconductor is a particularly interesting case, and it is included in the above list because it illustrates the enabling power of this modeling technique and provides a realistic estimate of potential detector performance without any expenditure of significant experimental resources.

3.1 Intrinsic Direct Bandgap Semiconductor

$Hg_{1-x}Cd_xTe$ is a classic example of the intrinsic direct bandgap semiconductor, and it has the added advantage of a bandgap that can be tuned to any cutoff wavelength in the IR spectrum by simply varying the x-value of the alloy composition. Following the arguments in Chapter 2, the relevant noise carrier is the minority carrier.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

For n-type materials, the minority carrier concentration is given by $p_o = 2n_i^2/\{n_d[1 + (1 + 4n_i^2/n_d^2)^{1/2}]\}$, and the value of n_i is known [5] both as a function of temperature and composition. The majority carrier concentration is given by $n_o = n_d[1 + (1 + 4n_i^2/n_d^2)^{1/2}]/2$. For the temperature range in which $n_d \gg n_i$, then $n_o = n_d$, and $p_o = n_i^2/n_d = n_i^2/n_o$. The minority carrier lifetime in good quality n-HgCdTe is limited by Auger recombination and is given by

$$\tau_A = 2\tau_{Ai} n_i^2/[n_o(n_o + p_o)],$$

and for $n_d \gg n_i$, we have

$$= 2\tau_{Ai}(n_i^2/n_o^2), \qquad (3.1)$$

where τ_{Ai} is the Auger lifetime for intrinsic material and is a known function [6] across all composition and temperature ranges. If we consider the n-type photoconductor in the range where $n_d \gg n_i$, then the thermal generation rate per unit area, defined by Eq. (2.2), is given by

$$\begin{aligned} G_{th} &= p_o t/\tau \\ &= n_i^2 t/n_o \tau_A \\ &= n_o t/2\tau_{Ai}. \end{aligned} \qquad (3.2)$$

All of the parameters in Eq. (3.2) are well known. t is determined by the absorption coefficient of the material, α_λ, which is a function of photon energy, and given by [7] $\sim 4 \times 10^4 [E_\lambda(E_\lambda - E_g)]^{1/2}$, where $E_\lambda = 1.24/\lambda$, E_λ is in eV when λ is in μm, and E_g is the bandgap of the material in eV. τ_{Ai} is given by $\tau_{Ai} \approx 6.3 \times 10^{-13}(E_g/k^3T^3)^{1/2}\exp(qE_g/kT)$. For $E_g = 0.124$ eV, t is typically chosen to be $\sim$10 μm, resulting in a cutoff wavelength (defined as the wavelength at which the spectral response is 0.5 of peak response) of 9.9 μm and an absorption quantum efficiency in the relevant 8–10μm system spectral band of 0.85. The thermal generation rate per unit area is thus given by

$$G_{th} = 6.3 \times 10^2\, n_o T^{3/2}/[E_g^{1/2}\exp(qE_g/kT)]. \qquad (3.3)$$

The normalized thermally generated minority carrier current is given by $G_{th}q/\eta_a$ and is plotted in Fig. 3.1 as a function of temperature, for fixed bandgap values of $E_g = 0.124$ eV and 0.25 eV, and a majority carrier concentration $n_o = 10^{15}$ cm^{-3}, assuming that Eq. (3.3) is valid, that is, that $n_d \gg n_i$. The value of G_{th}^* using exact expressions for n_o, p_o, and τ_A is also included, and Eq. (3.3) is seen to be valid over most of the appropriate temperature ranges for the two cutoff wavelengths. In fact, many present-day detectors are diodes operated under reverse bias, such that, within a diffusion length of the junction, the semiconductor is far from equilibrium. Under such conditions of operation, the majority carrier concentration is always n_d, regardless of the magnitude of n_i. This is discussed at length in Chapter 6, under high operating temperature (HOT) detectors. In Fig. 3.1, η_a is assumed to be 0.85 for $t = 10$ μm. Included in Fig. 3.1, on the right-hand axis,

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

is the equivalent incident background flux per unit area, Φ_B. For an $f/2.5$ system operating in the spectral bandwidth of 8–10 μm the incident background flux is 10^{16} photons/cm^2/s, and it is evident from Fig. 3.1 that BLIP performance (i.e., thermal current < background current) requires an operating temperature of <115 K. This temperature is referred to as the BLIP temperature of the material, and obviously the better the quality of the material, the higher this temperature will be. A similar system operating in the 3–5 μm spectral band will result in an incident background flux on the detector of 4×10^{14} photons/cm^2/s, with an indicated BLIP temperature of 180 K.

The calculated thermal generation rate can be used to predict the specific peak detectivity D^*_λ associated with the material by using Eq. (2.8). This calculated parameter as a function of temperature is shown for the 0.124 eV HgCdTe photoconductor in Fig. 3.2, with an assumed overall quantum efficiency $\eta = 0.67$,

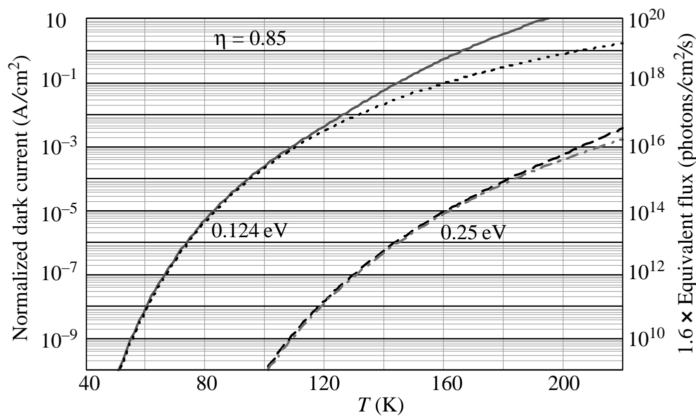


Figure 3.1 Normalized dark current vs. temperature for 0.124 and 0.25 eV HgCdTe.

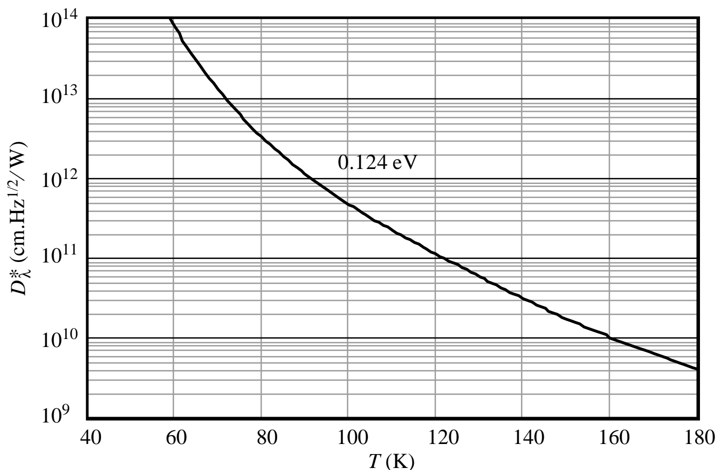


Figure 3.2 D^*_λ vs. temperature for 0.124 eV HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

and is in agreement with experimentally reported values [8]. Similar arguments can be made for HgCdTe photodiodes and will be discussed further in Chapter 4.

3.2 Extrinsic Semiconductor

The relevant noise carrier for the extrinsic semiconductor is the majority carrier, and the majority carrier concentration in a partially compensated singly ionized extrinsic semiconductor is given by [9]

$$n_o = [(N_d - N_a)/2N_a]N_c \exp(-qE_d/kT), \tag{3.4}$$

where N_d is the majority donor impurity concentration, N_a is the compensating acceptor impurity concentration, N_c is the density of states in the conduction band, and E_d is the energy level of the photon absorbing donor impurity relative to the conduction band.

The majority carrier lifetime is determined by the density of empty (ionized) donor levels, and for low temperatures, such that $n_o < N_a < N_d$, is given by $\tau = [\sigma v_{th} N_a]^{-1}$, where σ is the capture cross section for electrons into the donor level, and v_{th} is the majority carrier thermal velocity. In the relevant temperature range, σ is reported to be both mildly temperature- and electric-field-dependent. For Ge:Hg, at 30 K, a value of $\sigma = 10^{-13}\,\text{cm}^2$ has been reported [10], with an acceptor level activation energy $E_a = 0.089\,\text{eV}$. Larger values of σ in the range 2×10^{-13} to $1 \times 10^{-11}\,\text{cm}^2$ are reported for various impurities in Si [11]. $v_{th} = [8\,kT/\pi m^*]^{1/2}$, where m^* is the relevant effective mass. The σv_{th} product will be assumed a constant value at $5 \times 10^{-7}\,\text{cm}^3/\text{s}$ for Ge:Hg and $2 \times 10^{-5}\,\text{cm}^3/\text{s}$ for extrinsic silicon.

The required thickness of the IR material is determined by its absorption coefficient, which for the many extrinsic semiconductors [12] is given by $\alpha_\lambda = (10^{-15} N_{imp})$, where N_{imp} is the density of the majority carrier impurity. Thus, $t = 10^{15}/N_{imp}$, and the absorption quantum efficiency is $\eta_a = 0.67$.

The thermal generation rate per unit area for high-quality extrinsic n-type germanium, with $N_a \ll N_d$, is thus given by

$$\begin{aligned} G_{th} &= n_o t/\tau \\ &= 3 \times 10^{23} T^{3/2} \exp(-qE_d/kT), \end{aligned} \tag{3.5}$$

where we have assumed $m^* = 0.4\,m_o$. G_{th} is seen to be independent of N_a.

The dark current associated with the normalized generation rate, $G^*_{th}q$, is shown in Fig. 3.3(a) for a hypothetical impurity level $E_d = 0.124$ in germanium, compared to HgCdTe with a similar bandgap. The thermal generation current of the extrinsic semiconductor is considerably higher than the intrinsic semiconductor case by approximately five orders of magnitude at any specific operating temperature. The right-hand axis of Fig. 3.3(a) again indicates an equivalent background flux, and from this it is apparent that BLIP operation in a typical tactical environment of 10^{16} photons/cm^2/s is possible at a 50 K higher temperature for HgCdTe than

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

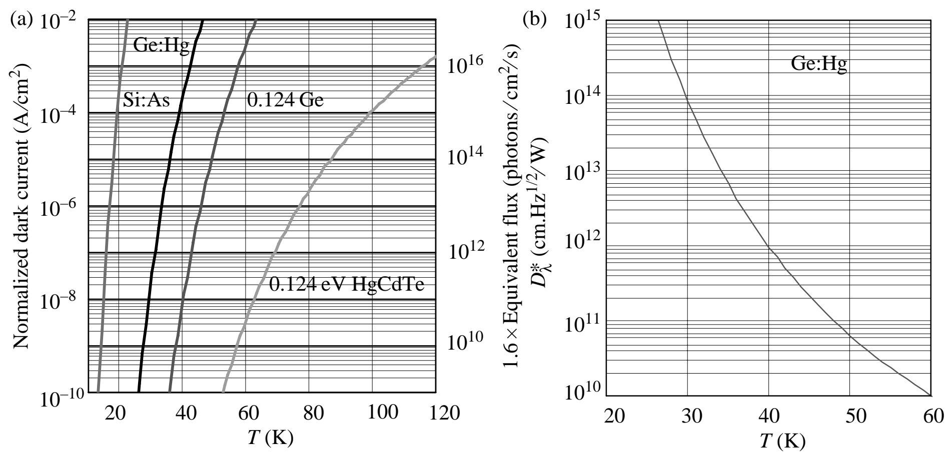


Figure 3.3 (a) Normalized dark current and (b) D^*_λ vs. temperature for Ge:Hg.

for the hypothetical extrinsic semiconductor. This problem is compounded by the fact that there is no identified impurity in germanium, or silicon, with an activation energy $E_d = 0.124\,\text{eV}$. Shallower levels are typically employed with an even larger temperature penalty relative to 0.124 eV HgCdTe. Included in Fig. 3.3(a) is an estimate of the normalized dark current associated with the mercury level in germanium, with an activation energy of 0.089 eV. The theoretical D^*_λ for Ge:Hg is shown in Fig. 3.3(b), which agrees both with published models [10] and unpublished experimental data taken at Texas Instruments.

Extrinsic silicon has been utilized for detection in the VLWIR and MWIR spectral bands, using the impurities arsenic, for a 26-μm cutoff, and indium and sulfur, for MWIR. The predicted dark current for Si:As is included in Fig. 3.3(a) and is seen to require very low temperatures of operation. As in the LWIR case, these impurity levels are not optimized for the atmospheric window at 5 μm; for instance, indium has an activation energy $E_a = 0.155\,\text{eV}$, which is equivalent to a cutoff wavelength of 7.4 μm. Sulfur is somewhat deeper, with $E_d = 0.187\,\text{eV}$ and a measured cutoff of 6.8 μm, but it exhibits lower quantum efficiencies, presumably because of lower solubilities in silicon. The normalized dark current as a function of temperature for a hypothetical 0.25 eV extrinsic Si detector is shown in Fig. 3.4(a), compared to HgCdTe of the same bandgap. Also included is a model for Si:In. The calculated D^*_λ for Si:In, as a function of temperature, is shown in Fig. 3.4(b), together with published data [11]. Again, for the purposes of IR detection, it is readily apparent that the extrinsic semiconductor is vastly inferior to the direct bandgap semiconductor.

It should be pointed out that, despite the BLIP temperature differential between extrinsic Si and HgCdTe, extrinsic detectors appear to have found a niche in very long cutoff wavelength strategic systems, due to issues associated with excess currents over and above thermal generation in HgCdTe IR devices operating at low background flux levels. This point will be discussed further in Chapter 8.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

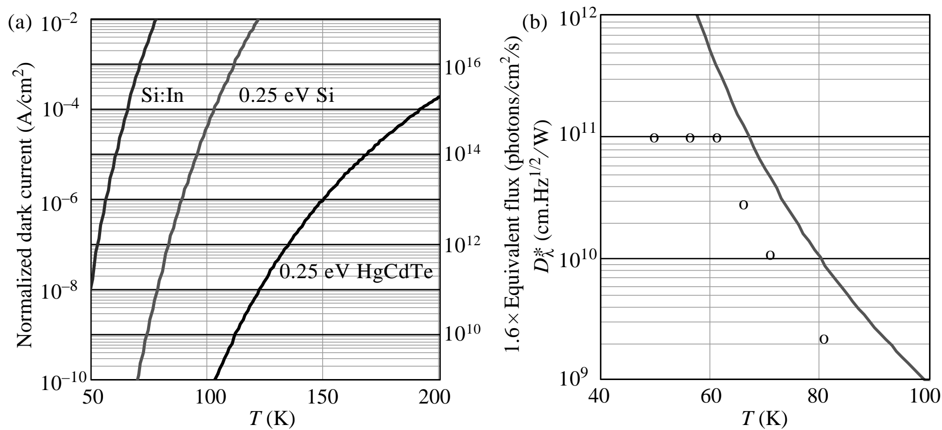


Figure 3.4 (a) Normalized dark current density vs. temperature for 0.25 and 0.155 eV extrinsic silicon, and (b) D^*_λ vs. temperature for Si:In.

3.3 Quantum Well IR Photodetectors (QWIPs)

Various types of quantum well superlattice have been proposed to detect IR radiation. The three simple basic types of superlattice are shown in Fig. 3.5. All have common features in that they consist of alternating thin layers (typically 20–100 Å) of semiconductors in which one-dimensional (1D) quantization occurs. The bands of the semiconductors in the individual layers are quantized into energy levels for motion in the direction of the superlattice growth, and bandgaps are generated that are determined to first order by the macroscopic dimensions of the layers. These levels are further modified by the interaction of adjacent layers, in much the same way as the motion of an electron is modified by a periodic lattice. This phenomenon

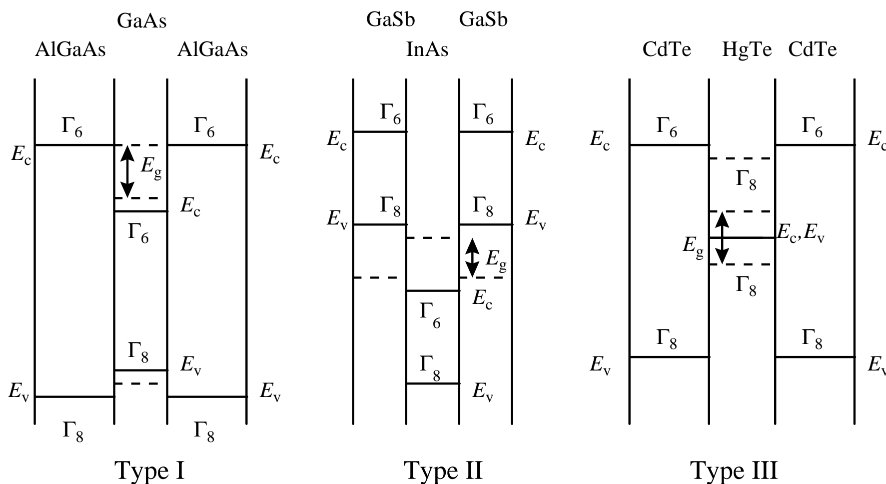


Figure 3.5 The three superlattice types.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

has been described in the literature as bandgap engineering. The different types of superlattice are defined by the relative position and symmetry of the valence and conduction bands in adjacent semiconductor layers. The relevant transitions defining the IR bandgap are designated by E_g in Fig. 3.5.

In a Type I superlattice, the alternating layers are composed of semiconductors in which the bandgaps are approximately aligned (i.e., the valence band of one does not overlap the conduction band of the other), and the conduction bands are of Γ_6 symmetry and the valence bands are Γ_8. A classical example is the GaAs/AlGaAs system.

A Type II superlattice is similar to Type I with the exception of overlapping conduction and valence bands in adjacent layers. An example is InAs/GaSb.

In a Type III superlattice the alternating layers are of different conduction and valence band symmetry. This is normally achieved by the use of a semimetal, in which the symmetry of the conduction and valence bands are flipped, as one layer. An example of a Type III superlattice is HgTe/CdTe.

All three types of superlattice have been proposed for use as IR materials. Type II and Type III superlattices are essentially minority carrier dominated intrinsic semiconductor materials when considered as IR detectors and will be discussed further in the next chapter. The Type I superlattice, better known as the QWIP (quantum well IR photodetector), is a majority carrier device and is viewed by many as a possible alternative to HgCdTe, mainly because of the perceived maturity of GaAs material technology and device processing. It is this materials system that will be considered here.

One version of the QWIP is shown in Fig. 3.6. IR absorption is achieved by transitions between the energy levels induced in the majority carrier conduction band by dimensional quantization. The superlattice shown in Fig. 3.6 involves GaAs/AlGaAs in which the transitions are between the two lowest level bound states in the GaAs conduction band. The AlGaAs layers are very thick and act as barrier layers that serve the dual purpose of providing the quantization and

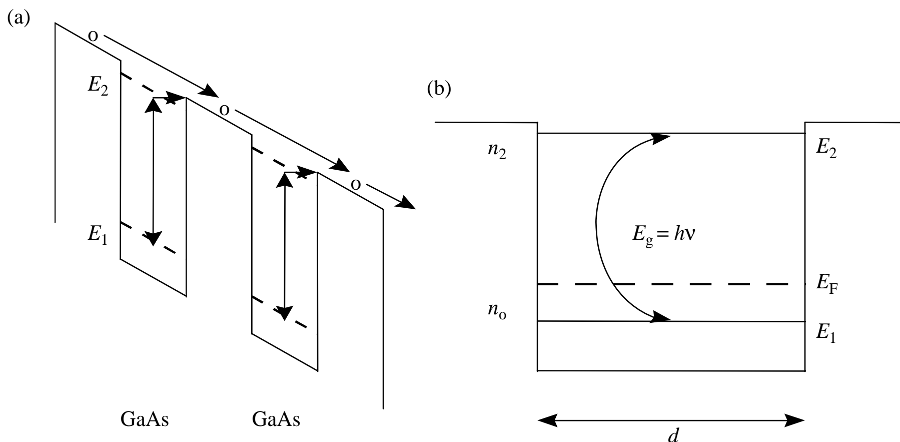


Figure 3.6 (a) QWIP structure under bias and (b) single quantum well.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

inhibiting any excess current, such as tunneling, through the superlattice. The GaAs layers thus, to first order, behave as isolated quantum wells, with energy levels determined by the well width. For detection in the LWIR region, well widths of 40–50 Å are used, and the number of wells is typically 50 to 100. The IR signal is detected as a photoconductive current through the superlattice by applying a bias so that photoexcited electrons in the upper level bound state can tunnel into the adjacent AlGaAs conduction band. This mode of operation is referred to as a bound-to-bound QWIP, and although not the most commonly used QWIP architecture, it does serve to illustrate most simply both the capabilities and limitations of this detector concept.

The thermally generated carrier density for the multiquantum well structure is obtained from a consideration of the density of states in the associated band structure of a single quantum well, shown in Fig. 3.6(b). In thermal equilibrium the first level is filled with the n_o electrons that are necessary to provide efficient free carrier absorption of the incident IR radiation. The Fermi level for these electrons is determined by the density of states in the first sub-band, namely

$$E_F = (n_o h^2 d)/(4\pi m^*), \tag{3.6}$$

where d is the width of the well, and h is Planck's constant.

The carrier concentration in the second sub-band is determined by E_F and the density of states in the second sub-band, such that

$$n_2 = n_o(kT/E_F)\exp q(E_F - E_2)/kT. \tag{3.7}$$

The lifetime of carriers in the E_2 sub-band is determined by hot electron–phonon interactions [13, 14] and is modified by the time the hot carrier spends in the barrier region relative to the quantum well. For typical geometries used in QWIPs in which the barrier layers are $\sim$5 to 10 times thicker than the active quantum wells, and assuming a saturation velocity of $\sim 10^7$ cm/s for electrons in the superlattice, then analysis of signal responsivity versus applied bias [15] indicates carrier lifetime values of $\sim 10^{-11}$ s, which is approximately in agreement with theory.

The thickness of the simple QWIP is determined by the absorption coefficient for incident radiation, which is complicated for the n-type version of this detector in that optical transition selection rules mandate an electric field polarization perpendicular to the surface of the superlattice for photon absorption. Various geometrical artifacts are employed to satisfy this requirement, such as etched gratings on detector surfaces, or effective angles of incidence other than normal. P-type material can also be used, as the degeneracy of the valence band of GaAs relaxes these selection rules. For geometrically optimized n-type layers, the unpolarized absorption coefficient for the active GaAs wells alone is given by [15] $\alpha_\lambda \approx 5 \times 10^{-15} n_o \text{ cm}^{-1}$. If it is normalized to the thickness of the whole detector, then this value is reduced by the ratio of the well to barrier thicknesses. The maximum value for the absorption quantum efficiency in this mode of operation is 0.5.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Thus, the thermal generation rate of the simple QWIP, with the total thickness of the active GaAs layers equal to $(\alpha_\lambda)^{-1}$, is given by

$$G_{\text{th}} = n_2 t/\tau$$
$$= 2 \times 10^{25} (kT/E_{\text{F}}) \exp q(E_{\text{F}} - E_2)/kT,$$

which minimizes at $E_{\text{F}} = kT$, giving

$$G_{\text{th}} = 5.5 \times 10^{25} \exp(-qE_{\text{g}}/kT), \tag{3.8}$$

where $E_{\text{g}} = E_2 - E_1$.

The normalized thermally generated dark current values as a function of temperature for $E_{\text{g}} = 0.124$ and 0.25 eV, as given by Eq. (3.8), are shown in Fig. 3.7 and compared with the corresponding values for HgCdTe with the same bandgaps. The dark current values on the left-hand axis are again normalized to the absorption quantum efficiency of the detectors for the assumed thicknesses, namely $t = 1/\alpha_\lambda$. Thus, for the QWIP, $\eta_{\text{a}} = 0.33$, and the value assumed for HgCdTe is $\eta = 0.67$.

It is apparent from Fig. 3.7 that the basic thermally generated dark current of the QWIP is approximately four to five orders of magnitude higher than for HgCdTe with the same bandgap, operating at the same temperature. For tactical background fluxes, the QWIP requires a temperature of operation some 50° below that of HgCdTe to achieve BLIP. It should be stressed here that this dark current is not due to any excess current mechanism, but it is merely due to the unavoidable thermal generation mechanisms inherent in the physics of the material system. The above simple reasoning has been applied to the specific QWIP that utilizes bound-to-bound subband transitions. Similar arguments apply to the more popular bound-to-continuum QWIP, with similar results [15]. The corresponding D^*_λ values

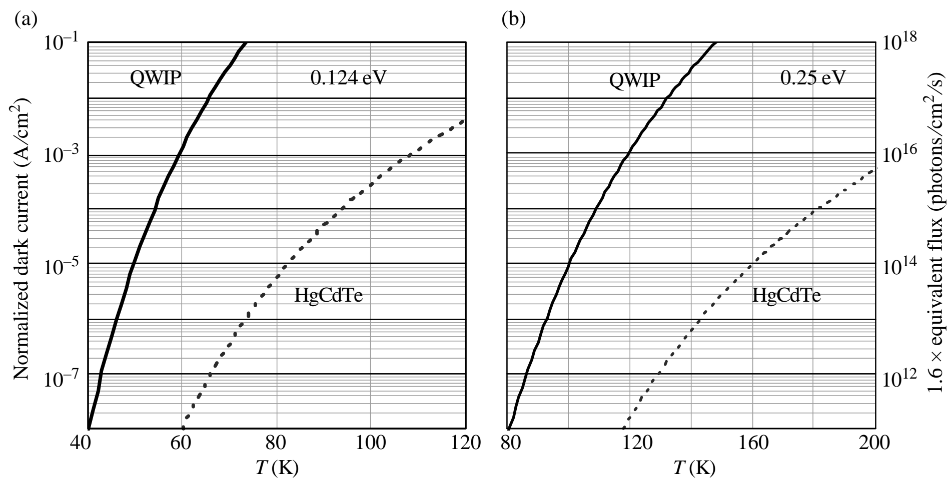


Figure 3.7 Normalized dark current vs. temperature for (a) 0.124 and (b) 0.25 eV GaAs QWIPs. The dark current axes are common.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

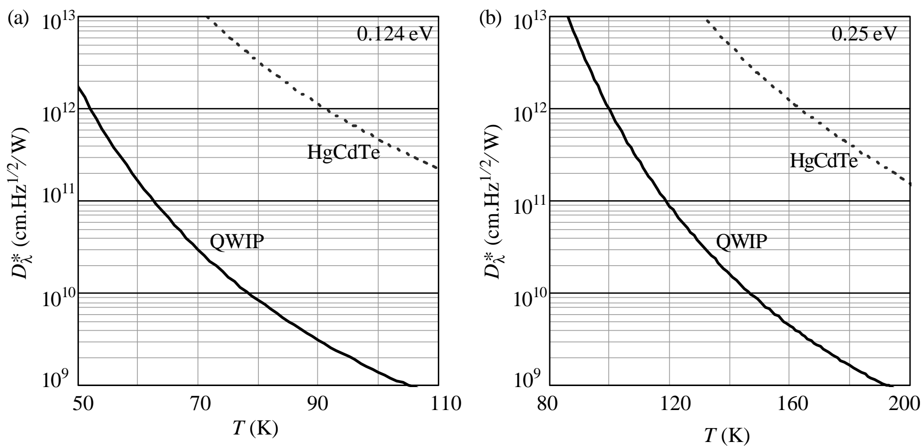


Figure 3.8 D^*_λ vs. temperature for 0.124 and 0.25 eV QWIPs relative to HgCdTe.

for 0.124 and 0.25 eV QWIPs as a function of temperature are shown in Fig. 3.8 and compared to HgCdTe.

Given that the thermal generation rate is an unavoidable consequence of the physics of QWIPs, the only options available for improving the situation, vis-à-vis the background generation rate, involve techniques to (a) reduce the dark current volume of the detector while maintaining the same optical absorption and (b) increase the effective lifetime of the detector.

Two potential solutions have been proposed to reduce the dark volume: (1) microlenses and (2) resonant or antenna structures. The microlens is a lens fabricated typically in silicon that is used to focus an IR pixel down in size by a factor of 10, so that the dark volume of the IR-sensitive material can be reduced by a similar amount, with no loss of IR signal. This concept has its limitations for $\lambda_c = 10\,\mu$m detectors, on a 25-μm pitch, due to diffraction effects for usable F/#s values. The resonant structure approach utilizes an optically thin IR detector that is backed by a resonant cavity of optical thickness $\lambda_c/4$. Antenna theory indicates an increased absorption for the resonant IR structure of $\sim 4n$ relative to the simple single layer of IR material with an antireflection coating, where n is the refractive index of the IR material. This factor is ~ 16 for GaAs and most other semiconducting materials. Thus, the dark-current-generating volume can be reduced by this factor while still generating the same IR signal. Theoretically, this same technique can be used for any IR materials system, but the thicknesses involved are small ($\sim$ a few thousand Angstroms), and surface effects could dominate a low carrier density minority carrier device, for instance, and thus may limit the use of this concept to majority carrier IR materials systems with large carrier densities, such as QWIPs. The dark-current-generating volume of the QWIP can thus theoretically be reduced by approximately an order of magnitude below that used for the calculations of Fig. 3.7, with an improvement in required operating temperature for BLIP. However, some penalty will be paid by the resonant nature of the spectral response of the detector, and its subsequent effect on black-body responsivity in a broad spectral bandpass system, and hence system NEΔT.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The effective lifetime of the detector can theoretically be increased by the introduction of extra dimensionality into the equation, as for instance in the form of quantum dots (QDIPs), which may or may not be embedded in a QWIP structure, forming a dot-in-the well (DWELL). The additional geometrical quantization associated with the dot dimensions is minimal; however, the size can be such that the dimensions are less than the mean free path due to optical phonon scattering (typically $\sim$1000 Å), which is the main vehicle limiting carrier lifetime in the QWIP, resulting in longer effective carrier lifetimes. The lifetime enhancement associated with this so-called phonon bottleneck has been discussed by Phillips [16].

As stated earlier, the attraction of QWIPs is related primarily to the common perception of the advanced state of GaAs materials and device processing technologies, and the availability of a number of GaAs/AlGaAs MBE foundries to meet the materials needs of FPA manufacturers. This is particularly attractive to IR system houses with no in-house HgCdTe capability. This aspect of QWIPs has, to a certain extent, been overemphasized, and it is indeed not immediately obvious that the required multilayer technology is reproducible enough for large-area FPA production purposes, even if one were prepared to accept the significant decrease in performance associated with the QWIP relative to HgCdTe.

3.4 Silicon Schottky Barrier Detectors

Silicon Schottky barriers (SBs) are majority carrier devices in which the IR absorption results in carrier transport over a metal–silicon SB. The IR absorption can be considered a three-stage process [17].

1. The incident radiation is absorbed by the carriers in the metal, but only a fraction of the resulting excited carriers will have sufficient energy to overcome the SB.
2. These carriers must be transported to the barrier while still retaining sufficient energy to escape.
3. The carriers escape over the barrier, where they are detected in the external circuit.

The first two stages constitute the absorption quantum yield in that they define the photon-generated carrier concentration in the neighborhood of the barrier relative to the thermally generated carrier concentration. The third stage represents an internal quantum yield, which is important in any calculation of device sensitivity such as D^*_λ or NEΔT, but it is not relevant to defining the BLIP operating temperature of the material. A simple SB incorporating a resonant structure is shown in Fig. 3.9(a), together with its associated energy band diagram in Fig. 3.9(b).

The thermal density of available carriers at the top of the SB, obtained by a simple integration over the density of states, is given by

$$\begin{aligned} n_o &= [(8\pi m^{3/2} kT(2(E_F + E_g))^{1/2})/h^3] \exp(-qE_g/kT) \\ &= 2 \times 10^{18} T \exp(-qE_g/kT), \end{aligned} \tag{3.9}$$

where we have assumed that $E_F + E_g \approx 8$eV. E_g is the barrier height.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

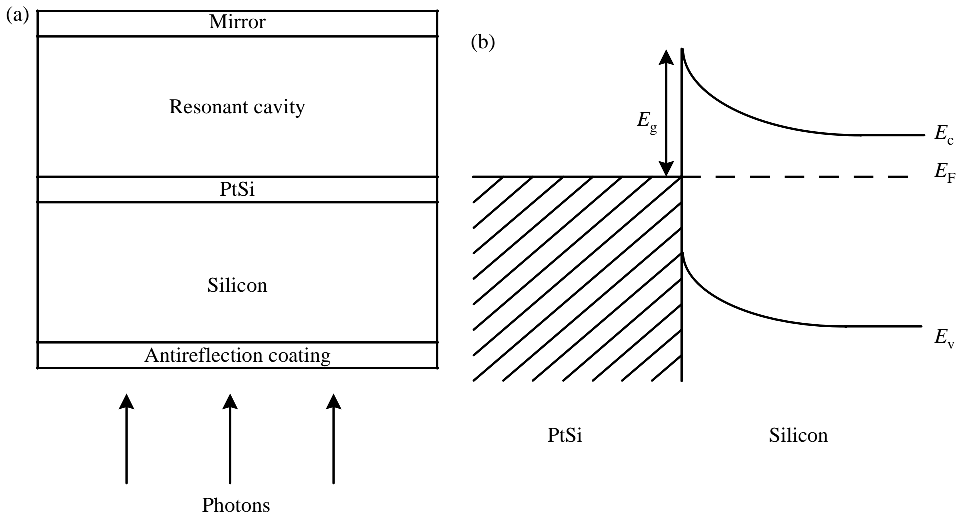


Figure 3.9 (a) Resonant SB structure and (b) associated band structure.

The lifetime for hot carriers in a metal is determined primarily by carrier–carrier scattering, and values obtained from calculations and measurements of the mean free path of hot electrons in metals [18] indicate $\tau \approx 1.5 \times 10^{-14}/E_g^2$ s, where E_g is in eV.

The optimum thickness of the IR absorbing layer in a PtSi resonant structure has been found to be ~10–20 Å, yielding an absorption efficiency of ~0.3, with an overall quantum efficiency due to the low internal quantum yield of 0.007, for the specific case of a 4.07 μm detection wavelength [17]. This indicates that, although the absorption depth for this material is ~300 Å [19], an almost theoretical absorption efficiency (equal to $(h\nu - E_g)/h\nu$, where $h\nu$ is the photon energy) can be achieved with a resonant structure a mere 20 Å thick. We will assume that a similar geometrical enhancement can be achieved at a 10-μm cutoff wavelength. The thermal generation rate for the silicon SB is thus given by

$$G_{th} = n_o t/\tau = 2 \times 10^{25} T E_g^2 \exp(-qE_g/kT), \quad (3.10)$$

where we have assumed a resonant thickness of 15 Å. The normalized dark current, $G_{th}^* q$, is shown in Fig. 3.10 as a function of temperature, for hypothetical 5- and 10-μm cutoff wavelengths, for $\eta_a = 0.3$. Included in Fig. 3.10 are the values for HgCdTe with the same cutoff wavelengths. The higher BLIP temperature for HgCdTe is readily apparent. This advantage for HgCdTe is further compounded by the low value of the internal quantum yield of the SB detector relative to the HgCdTe device, and it shows up in the overall quantum efficiency in the expression for the thermally limited D_λ^* given by Eq. (2.8). A comparison of the thermally limited D_λ^* values for HgCdTe, and Schottky barriers operating at wavelengths of

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

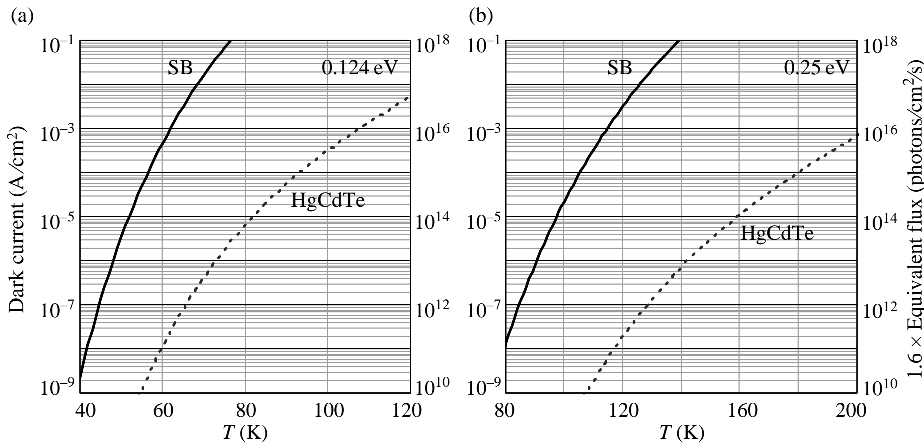


Figure 3.10 Normalized dark current vs. temperature for (a) 0.124 eV and (b) 0.25 eV SBs compared to HgCdTe.

5 μm (PtSi) and 10 μm (IrSi) are shown in Fig. 3.11, utilizing an overall quantum efficiency of 0.007 for the SB detectors.

Proponents of the Si SB detector technology would argue that the inherent dark current and quantum yield disadvantages of Si SBs relative to HgCdTe can be overcome by operating with longer system integration times (slower frame rates) and lower temperatures, with a considerable advantage to be gained in the area of array uniformity and manufacturing cost due to the use of standard silicon processing in array fabrication.

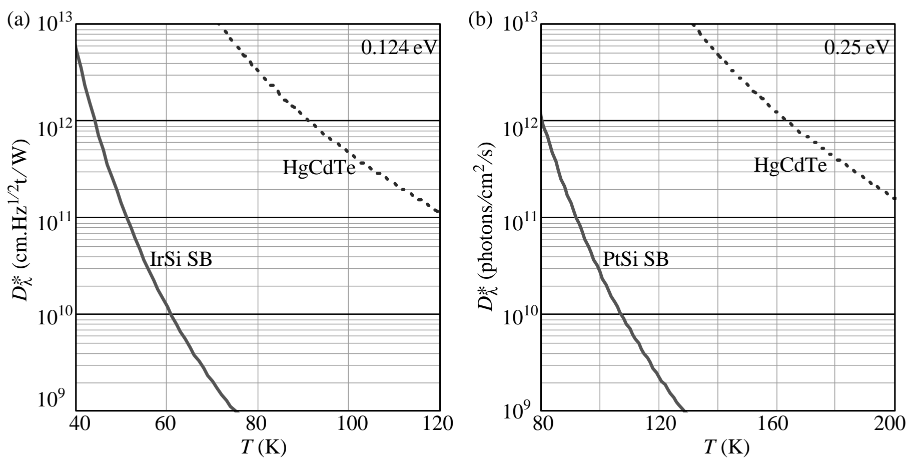


Figure 3.11 D^*_λ vs. temperature for (a) IrSi and (b) PtSi SB detectors compared to LWIR and MWIR HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

3.5 High-Temperature Superconductor

The advent of high-temperature (HTC) superconductivity in the late 1980s quickly led to proposals to detect IR radiation by excitation across the superconducting energy gap in much the same way as in direct gap semiconductors. We will consider the performance limitations of hypothetical HTC superconductors as photon detectors, using well-documented theory and experimental findings associated with conventional superconductors, assuming the existence of an HTC superconductor with an energy gap $2\Delta = 0.124\,\text{eV}$. Because a material with this large a value for 2Δ has not yet been reported, the discussion in this section will be limited to cutoff wavelengths of $10\,\mu\text{m}$.

The noise fluctuations will be determined in this case by the density of quasiparticles (or normal electrons) in the superconductor, which is given by [20]

$$n_{\text{t}} = 4N_{\text{o}}[\pi\Delta kT/2q]^{1/2}\exp(-q\Delta/kT), \tag{3.11}$$

where N_{o} is the single spin density of states at the Fermi level at $T = 0$.

Assuming the recombination of excess quasiparticles into Cooper pairs is dominated by electron–phonon interactions, as it appears to be for low-temperature superconductors, then the effective quasiparticle lifetime is given by [21]

$$\tau_{\text{eff}} = \tau + (n_t/4N_{\omega t})\tau_{\text{ph}}, \tag{3.12}$$

where τ represents the inherent lifetime associated with the phonon emission process, $N_{\omega t}$ is the equilibrium density of phonons with energy $h\omega/2\pi > 2\Delta$, and τ_{ph}^{-1} is the net transition probability for phonons to be lost from the energy range $h\omega/2\pi > 2\Delta$ by processes other than pair excitation. The second term in Eq. (3.12) essentially represents the modification to the quasiparticle lifetime associated with the departure of the phonon distribution from thermal equilibrium; that is, it represents the lifetime enhancement effect due to excess phonons generated by the recombining excess quasiparticles. The smallest value for τ_{ph} occurs when the quasiparticles are dissipated by simply leaving the device and is given by $t/2s$, where s is the velocity of sound. For device thicknesses of 200–2000 Å, then $\tau_{\text{ph}} \approx$ 10–100 ps. In actuality, acoustic mismatch will result in a phonon loss efficiency at the device boundaries of $\beta < 1$, and τ_{ph} will be $t/2s\beta$.

The analysis by Rothwarth and Taylor indicates that for low-temperature superconductors, the phonon-limited term of Eq. (3.12) will be dominant at all temperatures (the temperature dependence of the two terms is identical). An application of this analysis to the HTC superconductor $YBa_2Cu_3O_7$ yields a similar conclusion. Thus, we have $\tau_{\text{eff}} = n_t t/(8N_{\omega t}s\beta)$. The device thickness t will be approximately equal to the absorption thickness, in the range 200–1000 Å, and we will assume an absorption efficiency $\eta_{\text{a}} = 0.67$.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

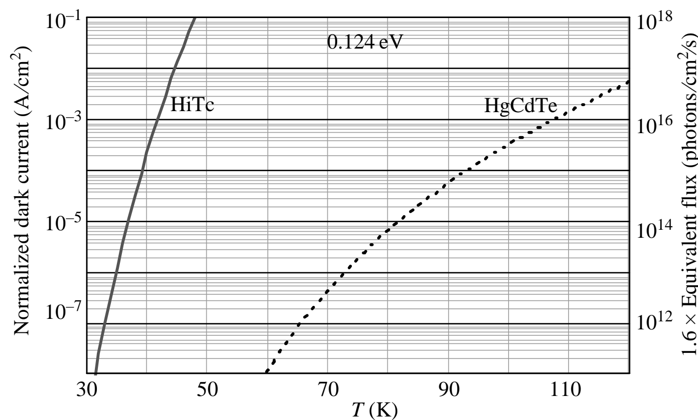


Figure 3.12 Normalized dark current vs. temperature for a hypothetical high-temperature superconductor with a quasiparticle bandgap. $2\Delta = 0.124$ eV. $\beta = 1$ is assumed.

The thermal generation rate for the HTC superconductor thus becomes

$$\begin{aligned} G_{\text{th}} &= 8s\beta N_{\omega t} \\ &= (3 \times 10^{3}\beta\Delta^{2}kTq^{2}/\pi^{2}s^{2}h^{3})\exp(-2q\Delta/kT) \\ &= [4 \times 10^{41}T\beta E_{g}^{2}/s^{2}]\exp(-qE_{g}/kT), \end{aligned} \tag{3.13}$$

where E_g is in eV, and s is in cm/s.

The normalized thermal generation rate per unit area for the hypothetical HTC superconductor with $2\Delta = 0.124$ eV is shown in Fig. 3.12 and compared to 0.124 eV HgCdTe. It is obvious that the BLIP temperature for this material, even if it were found to exist, is much lower than for 0.124 eV HgCdTe, and it indicates that resource expenditure to develop a HTC superconducting IR detector would not be a profitable exercise.

3.6 Conclusions

It is apparent from the above discussions that arguments based solely on thermally generated dark current, and overall device quantum efficiency, indicate that the intrinsic direct bandgap semiconductor, or its bandgap engineered equivalent, has no equal in the detection of IR radiation. This can be emphasized by a quantitative comparison of the normalized dark current for a specific bandgap as a function of temperature for the five materials systems considered in this chapter. This comparison is shown in Fig. 3.13 for a nominal bandgap of 0.124 eV, together with the assumed values for absorption quantum efficiency. At standard tactical backgrounds of 10^{16}photons/cm^2/s the temperature advantage for HgCdTe, which is representative of the intrinsic variable bandgap semiconductor, is approximately 50 K over the group of three IR materials systems consisting of the silicon SB,

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

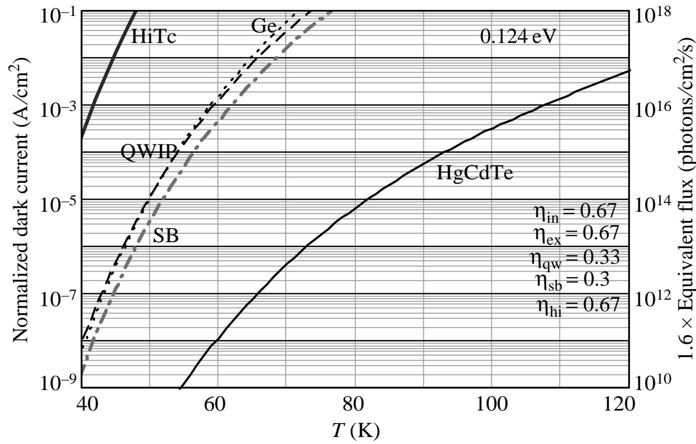


Figure 3.13 Comparison of normalized thermally generated dark currents vs. temperature for five different materials systems with a bandgap of 0.124 eV.

extrinsic germanium, and the GaAs/AlGaAs QWIP. The high-temperature superconductor suffers a further 20 K disadvantage. The second-place grouping is thinned out somewhat by the impact of the inherent low quantum yield of the silicon SB device on its overall quantum efficiency, and this is reflected in the comparison of peak D^*_λ values for the same five materials systems, shown in Fig. 3.14. Further differentiation between the five materials types results from other materials considerations, such as the availability of the required energy levels with $E_g = 0.124$ eV in the materials system in question. In this regard, the only truly variable bandgap systems are the intrinsic direct bandgap alloy and the III-V QWIP. It should be pointed out here that the performance of the QWIP relative to HgCdTe will be improved somewhat by the use of resonant structures, which was not considered in

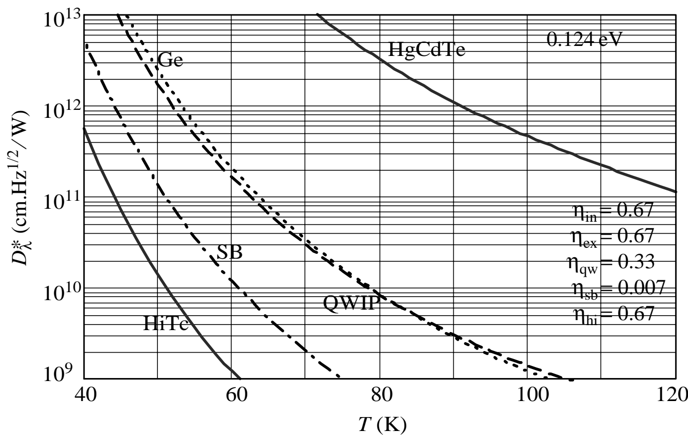


Figure 3.14 D^*_λ vs. temperature for five materials systems with a bandgap of 0.124 eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Figs. 3.13 and 3.14, resulting in a decrease in dark current of approximately an order of magnitude, with a subsequent increase in D^*_λ by a factor of approximately three, at the expense of some reduction of the overall broadness of the spectral response.

It can be argued that parameters other than thermally generated dark current and quantum efficiency should be considered in any comparison of IR materials technologies. Only the simple physics of the materials systems have been considered here. Reduction of the IR materials system to the form of a practical device may not be a trivial matter, and the development of surface passivation, contacts, and device processing technologies will also play a vital role in the potential application of a given technology. Other factors, some of which have been alluded to above, include

- Maturity, and hence cost, of the manufacturing technology;
- Material availability;
- Suitability of the device architecture for optimally interfacing to a Si readout IC;
- FPA uniformity;
- Suitability of the materials technology for advanced concepts.

In the author's opinion, these factors are no longer an issue in today's tactical IRFPA marketplace. The direct bandgap semiconductor in the form of HgCdTe, and to a lesser extent, InSb, satisfy all the above requirements, and for this reason the overriding consideration will be the physics of the materials system, and one is justified in assuming that the only effective performance differentiator that is at all relevant is the normalized thermal generation rate, G^*_{th}. It is for this reason, and this reason alone, that Texas Instruments, after the advent of the common module, invested its IR resources solely in the development of direct, tunable bandgap alloy materials systems and their bandgap engineered equivalents. The remainder of this book will thus be confined to a discussion of the complexities of these materials systems.

Note that the limitation assumed above for lifetime in HgCdTe is Auger1 recombination in n-type material. In fact, if p-HgCdTe is considered, then the limiting mechanism is Auger7, which is $\sim$10 times longer, resulting in a further lowering of dark current by a factor of $\sim$10 and a corresponding increase in D^* by a factor $\sim[10]^{1/2}$. Furthermore, if the detector is designed to contain no majority carrier regions, such as in a P-I-N diode, then all Auger generation is completely eliminated, and dark current is determined by generation through Shockley–Read (S-R) centers in the HgCdTe bandgap. Depending on the purity of the material, and the absence of native defects, there is theoretically no limit to the S-R lifetime in HgCdTe, resulting in no lower limit on the dark current in such a device. These thoughts are discussed more fully in Sec. 6.2.1.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 4

Intrinsic Direct Bandgap Semiconductors

Direct bandgap semiconductors are mainly compound semiconductors derived from combinations of elements from the II, III, IV, V, and VI groups of the periodic table. The earliest direct bandgap IR detectors involved simple binary compounds, such as InSb (III-V) and PbS (IV-VI), and were either photoconductive or photovoltaic. In the early 1970s, these binary compounds were superceded in many instances, but not all, by the variable bandgap alloys $Hg_{1-x}Cd_xTe$ and $Pb_{1-x}Sn_xTe$, whose compositions could be tuned to cover the important regions of the IR spectrum. HgCdTe, in its photoconductive form, ultimately became the material of choice for first-generation FPAs, and its photovoltaic counterpart is now the de facto standard for today's FPA technology. This alloy can be tuned to cover the complete IR spectrum from 1 to 20 μm for x compositions in the range $0.18 < x < 1.0$. PbSnTe is somewhat less versatile, covering the spectral range from 7 to >20 μm, but this can be extended to include the important MWIR region by the use of another member of the Pb salt family, PbGeTe. The demise of the Pb salt alloys for detection purposes was in part due to the large dielectric constants associated with this materials system, leading to high capacitance values and response time issues in first-generation scanning systems. It is interesting to speculate how this alloy might perform in today's proposed third-generation staring FPAs, although the question of its mechanical strength might still be a cause for concern. The advent of quantum wells has provided bandgap-engineered equivalents to the ternary alloys in the form of the Type II superlattice, such as InAs/GaSb, and the Type III superlattice, such as HgTe/CdTe. These may prove of value in certain applications where HgCdTe faces significant materials physics issues due to tunneling, such as in the detection of very long wavelength IR photons.

The thermal generation rate considerations of the last chapter indicated the inherent advantage of the direct bandgap semiconductor, as typified by $Hg_{1-x}Cd_xTe$, for IR detection. The normalized thermal generation rate for the direct gap semiconductor is given by $G^*_{th} = n_i^2/[n_{maj}\alpha\tau_{min}]$, where n_{maj} is the majority carrier concentration. For an ideal band-to-band Auger-limited semiconductor, such that $\tau_{min} = 2(n_i/n_{maj})^2\tau_{Ai}$, then $G^*_{th} = n_{maj}/(2\tau_{Ai}\alpha)$. For materials systems with the same

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

direct bandgap, α will be relatively constant, and G^*_{th} will be determined primarily by how closely the material's minority carrier lifetime approaches band-to-band Auger-limited behavior and by the minimum available doping concentration.

For this reason, the important contributions to minority carrier lifetime, τ_{min}, will first be defined, so as to better appreciate the subsequent discussions, which concern the limitations of the various forms of the direct bandgap semiconductor—namely binary compounds, ternary alloys, and superlattices—relevant to their use in photovoltaic devices.

In IR photodiodes, it should be stressed that G^*_{th} gives a direct indication of the magnitude of diffusion current from both sides of the device junction, at any particular temperature. Other sources of current may well be evident, to a greater or lesser degree, depending on the presence of defects in the material system, leading to a further degradation of device performance. These currents can be thermally generated or tunnel related.

4.1 Minority Carrier Lifetime

Minority carrier lifetimes in direct bandgap semiconductors are limited primarily by three recombination mechanisms: (1) radiative, (2) Auger, and (3) Shockley–Read recombination centers located in the bandgap of the bulk or surface of the material. The first two mechanisms are determined primarily by the band structure and doping concentration of the materials, whereas Shockley–Read is determined by defects or impurities in the semiconductor that are not necessarily associated with the doping and hence are to a certain extent avoidable. These three mechanisms are considered below.

4.1.1 Radiative recombination

The standard literature treatment of radiative recombination by van Roosbroeck and Shockley [22], used for many years, yields a value for the radiative lifetime of

$$\tau_R = 1/(B(n+p)) = 2\tau_{Ri}[n_i/(n+p)], \tag{4.1}$$

where

$$B = 5.8 \times 10^{-13}\varepsilon_{hf}^{1/2}(m_o/(m_e+m_h))^{3/2}(1+m_o/m_e+m_o/m_h)(300/T)^{3/2}E_g^2. \tag{4.2}$$

$\varepsilon_{hf}^{1/2}$ refers to the refractive index, and m_e and m_h represent the electron and hole effective masses.

This expression has been shown by Humphreys [23] to represent a gross underestimate of the radiative lifetime in IR semiconductors, due to noiseless photon reabsorption effects. The simplest way to consider the problem is to envision, say, a

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

p-type semiconductor of thickness t, in thermal equilibrium with its surroundings at temperature T. In equilibrium, the generation rate of minority carriers per unit volume within the semiconductor due to radiative transitions is $n_i^2/p\tau_r$, where τ_r is the radiative lifetime, n_i is the intrinsic carrier concentration, and p is the majority carrier concentration. This must equal the number of photons absorbed per unit volume from the surrounding enclosure, namely $2\eta\Phi_B/t$, where η is the quantum efficiency of absorption, and Φ_B is the total incident background flux density for wavelengths shorter than the semiconductor bandgap. Thus, the radiative lifetime for the semiconductor in thermal equilibrium is $\tau_r = n_i^2 t/2p\eta\Phi_B$, which is considerably longer than the value given by Eq. (4.2), which can be referred to as the internal radiative lifetime. The treatment of van Roosbroeck and Shockley is correct with regard to single photon emission, but it ignores the subsequent interaction of the emitted photons within the semiconductor. This phenomenon is often referred to as photon recycling and in equilibrium is responsible for the much longer radiative lifetime defined by Humphreys. However, in many experimental determinations of minority carrier lifetime, the situation is far from equilibrium. The decay of carriers generated by large fluxes of high-energy photons is measured, and if recycled photons can leak out of the semiconductor faster than they generate fresh carriers, then it is possible for the measurement to be limited by the internal radiative lifetime of Eq. (4.2). This phenomenon is in many aspects analogous to the phonon recycling encountered in lifetime measurements of quasiparticles in superconductors discussed in Sec. 3.5. However, from the point of view of mechanisms that limit device performance, internal radiative recombination can be ignored in a well-designed unit cell. The ultimate in detector performance is associated with background flux limited operation, which for Humphreys' lifetime is automatically satisfied.

4.1.2 Auger recombination

Auger recombination in direct gap semiconductors is a well-understood phenomenon involving the interaction of three carriers. The Auger1 mechanism in n-type material involves the recombination of a minority carrier hole and an electron, with the resulting energy and momentum conserved by a second hot electron in the conduction band, giving

$$\tau_{A1} = 2\tau_{Ai1} n_i^2/n(n + p), \tag{4.3}$$

where

$$\begin{aligned}\tau_{Ai1} = \{&7.6 \times 10^{-18}\varepsilon^2(1 + \mu)^{1/2}(1 + 2\mu)/[(m_e/m_o)|F_1F_2|^2(kT/E_g)^{3/2}]\}\\ &\times [\exp((1 + 2\mu)qE_g/(1 + \mu)kT)],\end{aligned} \tag{4.4}$$

ε is the dielectric constant, and $\mu = (m_e/m_h)$. τ_{Ai1} is defined as the intrinsic Auger1 lifetime. The largest uncertainty in the Auger model is in the calculation of the overlap integral F_1F_2, and empirical values, provided by a comparison to experimental

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

data, are typically used. However, they are invariably within a factor of 2 of theoretical values.

In p-type material, the primary Auger mechanism involves the recombination of a minority carrier electron and a heavy hole, with excess energy and momentum conserved by the production of a hot light hole, and is referred to as Auger7. Calculations by Casselman and Krishnamurthy [24], and Beattie [25], indicate that $\tau_{Ai7} \approx 10\tau_{Ai1}$.

4.1.3 Shockley–Read recombination

Simple equilibrium generation–recombination kinetics through a concentration of S-R centers, N_r, located at an energy E_r above the valence band, give the following expressions for the associated electron and hole lifetimes:

$$\tau_n = \left\{\tau_{po}(n + n_1) + \tau_{no}\left[p + p_1 + N_r p_1/(p + p_1)\right]\right\}/\left[n + p + N_r p p_1/(p + p_1)^2\right], \tag{4.5}$$

$$\tau_p = \left\{\tau_{no}(p + p_1) + \tau_{po}\left[n + n_1 + N_r p/(p + p_1)\right]\right\}/\left[n + p + N_r p p_1/(p + p_1)^2\right], \tag{4.6}$$

with

$$n_1 = N_c \exp\left[-(E_g - E_r)q/kT\right] \tag{4.7}$$

$$p_1 = N_v \exp(-E_r q/kT), \tag{4.8}$$

where $\tau_{no} = 1/\gamma_n N_r$, $\tau_{po} = 1/\gamma_p N_r$, γ_n and γ_p are the recombination coefficients for electrons and holes into the N_r centers, N_c and N_v are the densities of states in the conduction band and valence band respectively, and E_g is the bandgap. N_r is considered independent of p. The optimum thermal stepping stone occurs when E_r is at the intrinsic energy level E_i.

A consideration of the lifetime components above indicates that radiative recombination is not an issue, relative to Auger and S-R recombination. Furthermore, it is apparent that the Auger limit for a material containing majority carriers will be achieved provided the density of S-R centers is under control.

4.2 Diode Dark Current Models

Diode dark current models are discussed fully in Sec. 5.6 with reference to HgCdTe, but the relevant expressions for diffusion current, depletion current, surface generated currents, direct tunneling across the bandgap, and trap-assisted tunneling are reproduced here to allow for a more meaningful comparison of the direct bandgap materials systems discussed in this chapter.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Diffusion current is given by

$$J_{\text{dif}} = qn_i^2 t[(1/n\tau_n) + (1/p\tau_p)][\exp(qV/kT) - 1], \tag{4.9}$$

where n and p are the doping levels on either side of the junction, and τ_n and τ_p are the limiting minority carrier lifetimes.

Thermal generation in the depletion region through an S-R center at E_r is given by

$$J_{\text{dep}} = qn_i^2 W/[\tau_{no} p_r + \tau_{po} n_r], \tag{4.10}$$

where $\tau_{no} = 1/\gamma_n N_r$, $\tau_{po} = 1/\gamma_p N_r$, $n_r = N_c \exp(-qE_r/kT)$, $p_r = N_v \exp[q(E_r - E_g)/kT]$, and W is the width of the depletion region given by $W \approx [2\varepsilon\varepsilon_o(E_g - V)/qn]^{1/2}$. N_c and N_v are the conduction and heavy hole valence band density of states respectively, γ_n and γ_p are recombination coefficients for electrons and holes, and N_r is the density of S-R centers at E_r. For the optimum levels at E_i, the intrinsic energy level, $n_r = p_r = n_i$, and $J_{\text{dep}} = qn_i W/(\tau_{no} + \tau_{po})$, the familiar expression for diode depletion current. The bias dependence of J_{dep} is contained in the depletion width, W, which varies as $\sim V^{1/2}$ for an abrupt junction with an unlimited dimension to accommodate junction spread.

Surface generation currents are given by

$$J_s = qn_i s/2, \tag{4.11}$$

where s is the surface recombination velocity with no explicit bias dependence.

Direct tunneling between the conduction and valence bands, for a parabolic barrier, in a uniform electric field, with doping n, gives a bias dependent current

$$J_{\text{tdir}} = 3.93\varepsilon^{-3/2} V[m_e^*/E_g m_o][n(E_g + V)]^{1/2}$$
$$\exp[-2.1 \times 10^{10} \varepsilon^{1/2} m_e^{*\,1/2} E_g^{3/2}/n^{1/2}/(E_g + V)^{1/2}]. \tag{4.12}$$

Tunneling via N_t bandgap states at E_t, for a parabolic barrier, in a uniform electric field, with doping n, gives a bias voltage dependent current

$$J_{\text{tsr}} = [2 \times 10^{-12} N_t V/\varepsilon E_g] \exp[-3.35 \times 10^9 \pi \varepsilon^{1/2} m_e^{*\,1/2} E_g^{3/2}/n^{1/2}/(E_g + V)^{1/2}]. \tag{4.13}$$

These expressions for tunneling based on a uniform field should give an overestimate of tunnel current for standard abrupt junction diodes. Allowances for this effect can be made by the inclusion of a factor, f, in the exponents in Eqs. (4.12) and (4.13). Modeling suggests that $1 < f < 2$.

4.3 Binary Compounds

4.3.1 Indium antimonide: InSb

The only relevant binary compound utilized in current FPAs is indium antimonide (InSb), which is readily available in a highly pure n-type form with doping levels $\sim 10^{14}\ \text{cm}^{-3}$. The alloy has a bandgap of 0.224 eV at 77 K, decreasing to a value

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

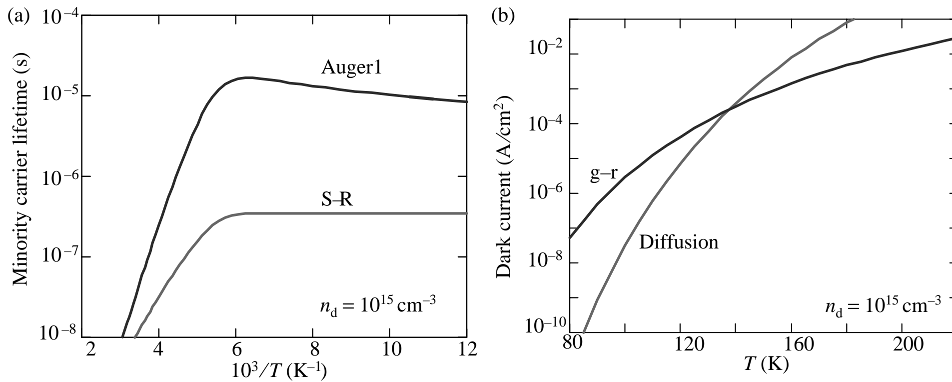


Figure 4.1 (a) Minority carrier lifetime vs. $10^3/T$ for InSb Auger1 and S-R, and (b) dark current vs. temperature for InSb with $N_d = 10^{15}$ cm^{-3}.

of 0.17 eV at 300 K. Photodiodes are typically diffused or implanted p^+/n mesa-defined structures fabricated on thinned bulk or LPE grown, low dislocation density ($<100/cm^2$), n-type material for backside illumination. The dark current in these diodes is determined at best by diffusion current from the n-region, with additional components due to generation–recombination from Shockley–Read (S-R) centers located within the depletion region of the device and diffusion from the p-region.

Despite the availability of high purity n-type material since 1960, band-to-band Auger recombination has not been observed in InSb at low temperatures. Minority carrier lifetimes in bulk n-InSb have been reported by Pines and Stafsudd [26]. The data are modeled using an Auger-like three carrier interaction via S-R centers located at approximately mid-gap, as opposed to the normal three-carrier band-to-band recombination mechanism described earlier. The dependence of minority carrier lifetime on inverse temperature for this S-R mechanism is shown in Fig. 4.1(a) for a donor center located 0.124 eV from the valence band, with a concentration $N_r = 8 \times 10^{13}$ cm^{-3}. The capture cross section of the center was normalized to yield a value for $\tau_{SR} = 7 \times 10^{-6}$ s, for a carrier concentration of 5×10^{13} cm^{-3}.

Also included in Fig. 4.1(a) is a calculation of the band-to-band Auger1-limited lifetime for InSb for the same donor concentration, with allowances for the known variation of bandgap parameters with temperature, and a value for the overlap integral of 0.235. More recent data [27] have compared lifetimes of bulk and LPE grown n-InSb and again found that S-R mechanisms dominate both material types, with better values for LPE than bulk, although both appear inferior to Pines' original bulk data.

This lifetime model can be utilized to predict the dependence of diffusion current on temperature in p^+/n diodes, assuming that the current is dominated by the contribution from the n-region, and that the minority carrier diffusion length is greater than the diode dimensions. The diffusion current generated in a reverse-biased diode is $qn_i^2t/n\tau$, where t is the n-region thickness, τ the hole lifetime in the n-region, n_i the intrinsic carrier concentration, and n the electron concentration. This current is shown in Fig. 4.1(b) for $N_d = 10^{15}$ cm^{-3}, assuming a thickness of

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

10 μm, which is entirely appropriate for intrinsic IR materials. Shockley–Read dominated material requires that generation–recombination within the depletion region of the diode be considered. Because of the Auger-like nature of Pines' recombination mechanism, it is unlikely to be important in the depletion region of the reverse-biased diode where carrier concentrations are essentially zero. An estimate of depletion current in the p^+/n InSb diode is made assuming a temperature and doping concentration independent S-R lifetime of 3×10^{-7} s, as found by Jost [27] in high-quality LPE material. This depletion current, $J_{dep} = qn_i W/2\tau_{SR}$, for a width W that is consistent with $N_d = 10^{15}\,cm^{-3}$, is shown in Fig. 4.1(b), and agrees with published diode data [28, 29] at 77 K. The diode is depletion current limited up to temperatures of 125 K, and BLIP for an $F/2$ system only at temperatures <110 K.

The absorption coefficient in direct gap intrinsic materials for photon energies of $h\nu$ is given, to a good approximation, by $\alpha = 4 \times 10^4 (h\nu(h\nu - E_g))^{1/2}\,cm^{-1}$. For InSb, the mobility of electrons and holes is sufficiently large over the range of temperatures utilized for IR detection that minority carrier diffusion lengths are $\gg 1/\alpha$, and quantum efficiencies approaching unity are readily achieved for device thicknesses of $5-10\,\mu m$.

4.4 Ternary Alloys

4.4.1 Mercury cadmium telluride: $Hg_{1-x}Cd_xTe$

$Hg_{1-x}Cd_xTe$ is a ternary alloy with a bandgap varying from the semimetallic HgTe ($x = 0$), with a negative bandgap of -0.3 eV at low temperatures, to the wide-bandgap semiconductor CdTe ($x = 1$), with a value of 1.5 eV. The essential properties of this alloy are described in the review article by Long and Schmit [30], and some aspects that are relevant to diode fabrication and performance are considered further in Chapter 5. Background impurity concentrations in LPE are typically donor dominated, at the level of 1 to $2 \times 10^{14}\,cm^{-3}$. MBE growth on the $\langle 211B \rangle$ orientation is also donor dominated to approximately the same degree, but reported values for MOCVD are somewhat higher. The alloy composition is continuously tunable over the whole range of x, but the main regions of interest coincide with the spectral windows of the earth's atmosphere, namely $x = 0.2$ ($8-14\,\mu m$), $x = 0.3$ ($3-5\,\mu m$), and $0.4 < x < 0.7$ (for various near-IR bands in the $1.5-3\,\mu m$ range). Of particular importance for tactical IR systems are the $3-5\,\mu m$ (MWIR) and $8-14\,\mu m$ (LWIR) spectral bands.

4.4.1.1 n-Type

The minority carrier lifetime in high-quality n-HgCdTe is determined by band-to-band Auger1 recombination for donor concentrations $>5 \times 10^{14}\,cm^{-3}$. The dependence of lifetime on temperature and carrier concentration is given by $\tau_{A1} = 2\tau_{Ai1} n_i^2/n(n + p)$, where τ_{Ai1} represents the Auger1 limited lifetime for intrinsic material at any temperature T, and it is shown graphically in Fig. 4.2(a) for

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

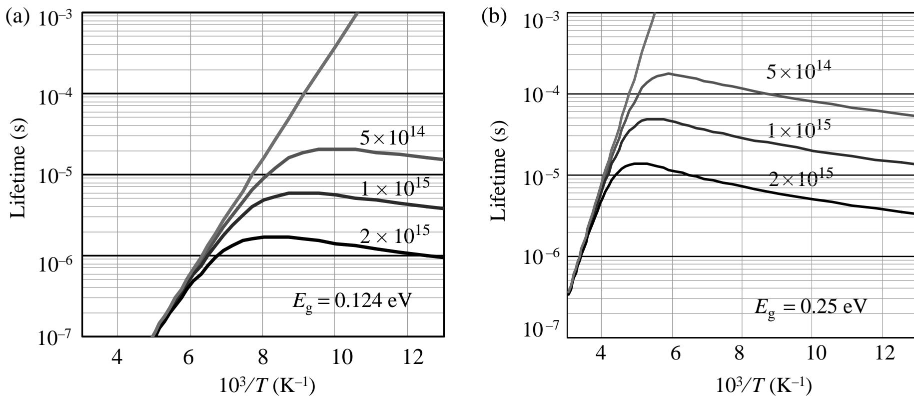


Figure 4.2 Modeled Auger1 limited lifetime in n-type HgCdTe vs. $10^3/T$ for various carrier concentrations, with (a) constant bandgap 0.124 eV and (b) constant bandgap 0.25 eV.

a constant bandgap of $E_g = 0.124$ eV and in Fig. 4.2(b) for $E_g = 0.25$ eV for various donor concentrations. τ_{Ai1} has been modeled from first principles by a number of authors and agrees with experimental data if an overlap integral of $\sim$0.2 is assumed.

Shockley–Read limited lifetime behavior [6] is observed in n-HgCdTe, at doping concentrations $<5 \times 10^{14}$ cm^{-3}. For high-quality n-HgCdTe, measured values of $\tau_{SR} > 5$ μs are typical for $x = 0.225$, and $\tau_{SR} > 30$ μs for $x = 0.3$.

For n-type HgCdTe, $L_d > 1/\alpha$, and simple planar diode geometries can be used. The corresponding Auger1 and depletion dark current densities as a function of temperature are shown in Fig. 4.3 for 10-μm-thick p$^+$/n planar LWIR and MWIR double layer heterojunction (DLHJ) diodes with $N_d = 2 \times 10^{15}$ cm^{-3}. The lower dark current for the 0.25 eV HgCdTe DLHJ relative to InSb shown in Fig. 4.1(b) is due partly to the tuning ability of the alloy system relative to InSb, whose bandgap is decreasing with increasing temperature, and partly due to the presence of the

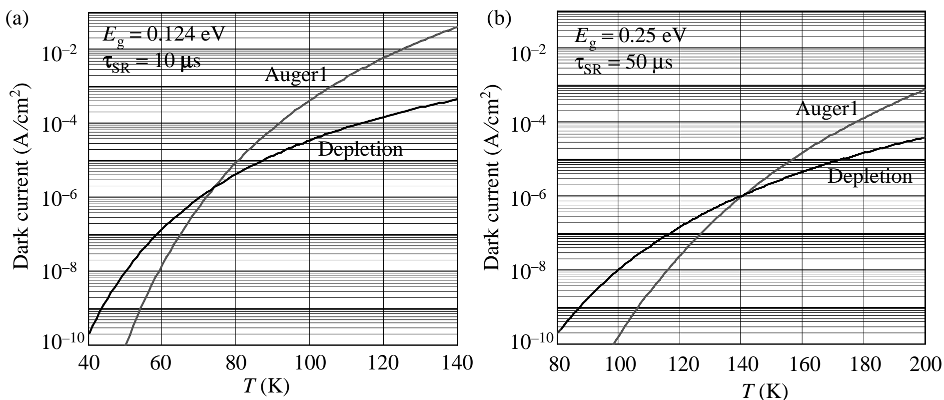


Figure 4.3 Auger1 and depletion dark current density vs. temperature in 10-μm-thick p$^+$/n HgCdTe DLHJs doped at 2×10^{15} cm^{-3}, with constant bandgap (a) $E_g = 0.124$ eV and (b) $E_g = 0.25$ eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

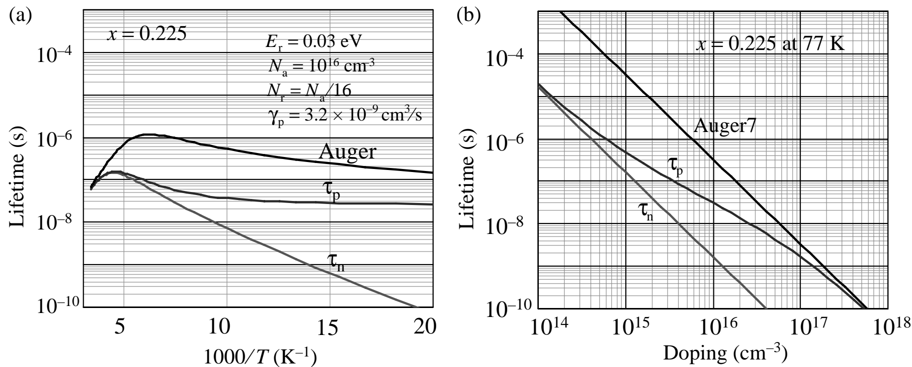


Figure 4.4 (a) $\tau_n, \tau_p, \tau_{A7}$ vs. $10^3/T$ for a donor-like S-R center in $x = 0.225$ HgCdTe located 0.030 eV from the conduction band and a hole concentration due to vacancies of 10^{16} cm^{-3} and (b) the dependence of these lifetimes on the hole concentration associated with vacancies at 77 K.

large S-R generated depletion current component in InSb photodiodes, associated with its relatively short lifetime. The assumption of Auger1 limited performance for HgCdTe is thus valid at higher temperatures.

4.4.1.2 p-Type

For minority carriers in p-HgCdTe, a careful distinction must be drawn between extrinsically and vacancy-doped material and any combination of these p-type dopants. For the sake of clarity, the two types of p-type material will be considered separately here.

(a) Vacancy-doped p-type HgCdTe

The minority carrier lifetime in purely vacancy-doped material is limited by a Shockley–Read (S-R) mechanism, with no evidence of Auger recombination at low temperatures. The expressions for the electron and hole lifetimes in p-HgCdTe are given by Eqs. (4.5) to (4.8). The published lifetime data for vacancy-doped material indicates that the S-R center is donor-like, even though the metal vacancy is an acceptor and located some 30 mV from the conduction band for both $x = 0.225$ and $x = 0.3$. An S-R lifetime calculation for such a center is shown in Fig. 4.4, as a function of (a) inverse temperature and (b) hole concentration, for a constant HgCdTe composition of $x = 0.225$.

Because of the donor-like nature of the S-R center, the recombination coefficient for electrons is assumed sufficiently large that τ_{no} is negligible relative to τ_{po}. The values shown in Fig. 4.4 are in good agreement with published data, and we have assumed $\tau_{Ai7} = 12\tau_{Ai1}$. For large N_r values, the carrier lifetimes are not equal, and $\tau_p \gg \tau_n$, resulting in photoconductive response times that can vary as τ_p, and diode dark currents that vary as $1/\tau_n$, where $\tau_n = n_1\tau_{po}/p$. For the doping

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

concentration range where p is sufficiently large that $N_r > n_1$, but not so large that Auger7 dominates, then $\tau_p = 1/\gamma_p p$, providing an unambiguous determination of $\gamma_p = 3.2 \times 10^{-9}\,\text{cm}^3/\text{s}$. The required concentration of donor-like S-R centers to fit the data is $\sim N_a/16$. The source of these donors is presently unknown, but it is tempting to think that they are in some way related to a compensating native donor defect that is strongly correlated to the concentration of metal vacancies, which is typically obtained by annealing at an appropriate temperature on the Hg-rich side of the phase field.

Care must be exercised in calculating the diffusion-limited dark current from the purely vacancy-doped p-region of a planar n^+/p diode due to the magnitude of the diffusion length L_d, which can easily be $< 1/\alpha$ at high doping concentrations and low temperatures. Creative diode geometries must be used in these circumstances to obtain high quantum efficiencies. It is obvious from the above discussion that diffusion-limited dark currents in vacancy-doped HgCdTe at low temperatures, unlike the case for n-HgCdTe, will not achieve Auger-limited performance.

(b) Extrinsically doped p-HgCdTe

Extrinsically (Cu, Au, Ag, As) doped p-HgCdTe containing no vacancies exhibits considerably longer minority carrier lifetimes than purely vacancy-doped p-type material. Shockley–Read and internal radiative recombination can be used to model the measured lifetime at lower carrier concentrations, and there is evidence of Auger7 recombination limiting the lifetime at higher hole concentrations, at values indicating $\tau_{Ai7} \approx 10\tau_{Ai1}$.

In the case of extrinsically doped p-type material, the S-R center appears to be acceptor-like or neutral, located approximately at the intrinsic level. The density of N_r centers is independent of p and is not large, so $\tau_n \approx \tau_p \approx \tau_{no}(p + n_i)/p$. The dependence of τ_p, τ_n on inverse temperature is shown in Fig. 4.5(a) for $p = 2.5 \times 10^{15}\,\text{cm}^{-3}$, $\tau_{no} = 5\,\mu\text{s}$.

The modeled dependence of τ_p, τ_n on doping concentration for $x = 0.225$ material is shown in Fig. 4.5(b), including the Auger7 contribution as calculated by Beattie, and internal radiative and S-R components. For $p < 10^{16}\,\text{cm}^{-3}$ the measured lifetime is dominated by S-R and is internal radiative, and for $p > 10^{16}\,\text{cm}^{-3}$ Auger dominates and the lifetime varies as $1/p^2$.

Similar arguments apply to MWIR p-HgCdTe, and the modeled dependence of minority carrier lifetime vs. $10^3/T$, and doping concentration is shown in Fig. 4.6, for an assumed $\tau_{no} = 20\,\mu\text{s}$.

(c) General p-type doping

Invariably, p-HgCdTe contains both vacancies and extrinsic dopants. For this case, the S-R lifetimes are controlled by a combination of the above mechanisms. For these modes of operation, the electron lifetime $\tau_n = \left[\sum(\tau_{no} + \tau_{po} n_1/p)^{-1}\right]^{-1}$. It is thus possible for n_1, N_r, γ_n, and γ_p to be determined by the metal vacancies, and the

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

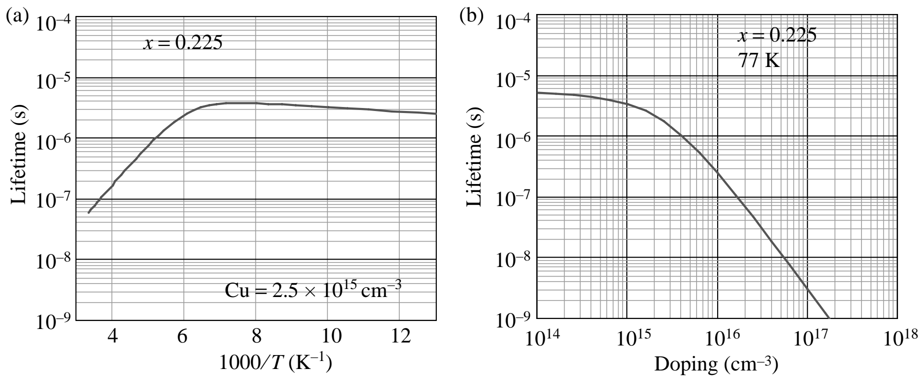


Figure 4.5 Lifetime of $x = 0.225$ HgCdTe (a) vs. $10^3/T$ for 2.5×10^{15} cm^{-3} Cu and (b) vs. Cu doping concentration at 77 K, for $\tau_{no} = 5\,\mu$s.

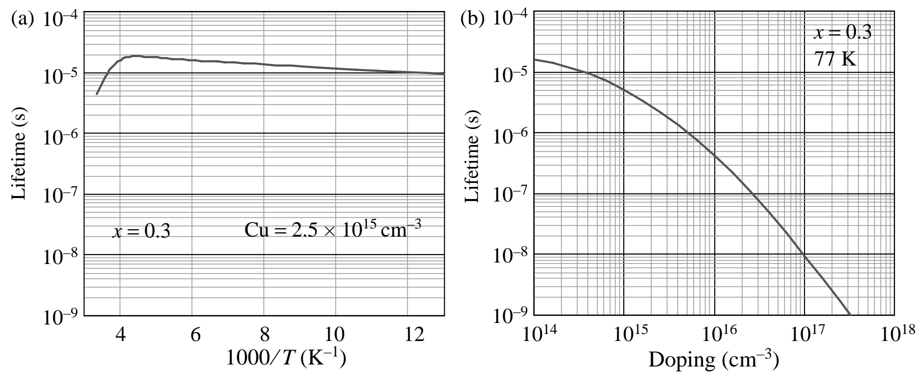


Figure 4.6 Lifetime of $x = 0.3$ HgCdTe (a) vs. $10^3/T$ for 10^{15} cm^{-3} Cu and (b) vs. Cu doping concentration at 77 K, for $\tau_{no} = 20\,\mu$s.

carrier concentration, p, to be determined by the extrinsic dopant. This behavior is in qualitative agreement with data found in the literature [31–33].

In summary, we can say that ultimate HgCdTe performance at low temperatures is achieved in the absence of metal vacancies. The arguments concerning internal radiative lifetime also indicate that this recombination mechanism can be ignored in a well-designed unit cell. Under these circumstances, the state-of-the-art limiting lifetimes in HgCdTe are associated with Auger and S-R, as shown in Fig. 4.7, for MWIR and LWIR HgCdTe at 77 K, as a function of doping concentration, for both n- and p-type material. It is apparent that p-HgCdTe exhibits longer lifetimes than Auger1 for n-type material, for all doping concentrations $>10^{15}$ cm^{-3}. The premise of Auger-limited performance assumed in Chapter 3 for HgCdTe is thus seen to be valid for both n-type and p-type HgCdTe, fully justifying the dark current arguments made there relative to other materials systems.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

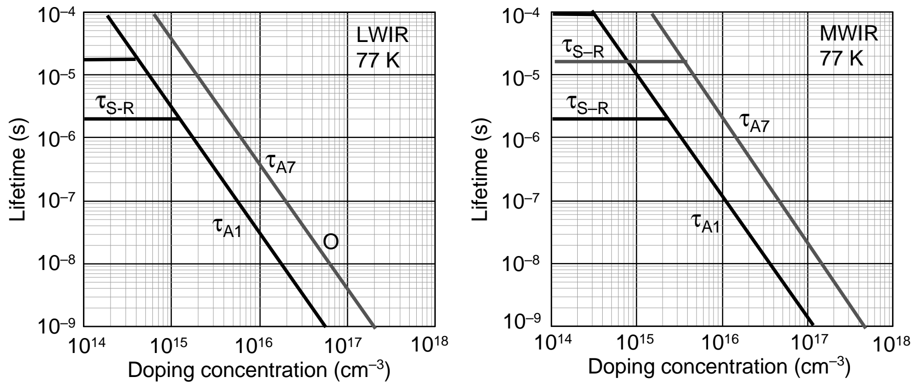


Figure 4.7 State-of-the-art measured minority carrier lifetime values for n- and p-type MWIR and LWIR HgCdTe at 77 K.

4.5 $Pb_{1-x}Sn_xTe$

Before leaving the realm of ternary alloys, it is worth revisiting the materials technology that lost out in the Common Module race in the early 1970s, namely PbSnTe. From a purely physics point of view the direct bandgap alloy PbSnTe is a fascinating materials system.

- There is evidence of para-electric behavior [34], in which the static dielectric constant obeys a Curie law above a ferroelectric transition temperature which is less than absolute zero.
- The static dielectric constant can exhibit very large values, ~1000, and is electric field dependent.
- The direct bandgap is at the L-point of the Brillouin zone.
- The band extrema are prolate ellipsoids in the ⟨111⟩ directions, with relatively large effective mass values for the bandgaps in question.
- The effective masses are equal in the conduction and valence bands.
- The phase fields of the compounds are well defined at 50% stoichiometry.
- Crystalline quality is extremely high.

The consequences of the above properties for advanced IR detection systems are potentially significant, particularly for the very long wavelength (VLWIR) spectral region. The high dielectric constant combined with the larger effective mass implies that tunneling currents, for any bandgap, will be greatly reduced compared to HgCdTe for similar doping concentrations. The equal valence and conduction band structures mandate that Auger recombination will be negligible for PbSnTe, because the volume of phase space available for the simultaneous satisfaction of energy and momentum conservation will be extremely limited. This was first pointed out by Junga et al. [35] as the principle reason for the ease of laser action in Pb salt diodes relative to comparable II-VI and III-V devices, particularly for long wavelengths. The optical transitions are also stronger in PbSnTe because of the four equivalent valleys and the large density of states in the conduction band. However, the values of n_i are similar to HgCdTe with the same bandgap, despite

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

the presence of the multiple band minima and the larger conduction band mass, due to the relatively low valence band effective mass in PbSnTe.

The general IR properties of the Pb-Sn chalcogenides are covered in the excellent review article of Melngailis and Harman [36] and will not be repeated here. The parameters of PbSnTe of primary interest here, for comparison purposes, are minority carrier lifetime and achievable doping concentrations, together with an estimate of excess sources of dark current over and above thermally generated diffusion currents.

4.5.1 Minority carrier lifetime

For comparison purposes, it is of interest to examine the generic expression utilized earlier for Auger1 recombination, namely

$$\tau_{Ai1} = 7.6 \times 10^{-18}\varepsilon^2(1+\mu)^{1/2}(1+2\mu)/\left[(m_e/m_o)|F_1F_2|^2(kT/E_g)^{3/2}\right] \times \left\{\exp\left[(1+2\mu)E_g/(1+\mu)kT\right]\right\}, \quad (4.14)$$

and apply it to $x = 0.21$ PbSnTe, with $E_g = 0.113$ eV at 77 K. Estimates for ε are ~500, with density of states effective masses of $0.0728\,m_o$ for the electron, and $0.082m_o$ for the hole [37]. Assuming $|F_1F_2|$ is ~0.2, as for HgCdTe, then Eq. (4.14) gives $\tau_{Ai1} \approx 8.2 \times 10^3$ s for 11 μm cutoff PbSnTe at 77 K, as opposed to $\sim 6.9 \times 10^{-3}$ s for 11 μm HgCdTe at the same temperature. The large increase is due in part to the threshold energy, as given by the exponent in Eq. (4.14), increasing to $\sim 3E_g/2$, and partly to the large increase in dielectric constant, which in turn reduces the Coulomb interaction between interacting electron wavefunctions that are responsible for the Auger effect and modifies the pre-exponent factor. It is thus apparent that Auger recombination is indeed not relevant for minority carriers in PbSnTe.

The earlier comments regarding radiative recombination in HgCdTe will apply equally to PbSnTe, and the only remaining potential issue is S-R recombination. The early work [38] on minority carrier lifetime in ~0.1 eV photoconductive n- and p-PbSnTe at 77 K indicated a lifetime $\tau_{SR} \approx 2 \times 10^{-8}$ s that was relatively independent of temperature from 77 K to 300 K, with doping concentrations in the 10^{15} to 10^{16} cm^{-3} range. Similar values for τ_{SR} were estimated from depletion current modeling [35] of *I–V*s at 77 K utilizing material with doping concentrations of 10^{17} to 10^{18} cm^{-3}. The lack of a temperature dependence of τ_{SR} for both n- and p-type is to be expected in the absence of Auger recombination, provided the S-R center is neutral, which is highly likely in view of the large values of dielectric constant in the PbSnTe alloy. Typical values of capture cross-section for a neutral S-R center are $\sigma \approx 10^{-16}$ cm^2, which will give a capture coefficient for the center, $\gamma = \sigma v_{th} \approx 2 \times 10^{-9}$ cm^3/ s at 77 K. Thus, a value for $\tau_{SR} = 2 \times 10^{-8}$ s will require a density of neutral recombination centers, $N_r \approx 2.5 \times 10^{16}$ cm^{-3}.

No identification of the S-R center has ever been proposed, and it is not clear whether it is due to residual impurities or a native defect in the material. However, it is obvious that this S-R center will exert the dominating influence over both

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

thermally generated and tunnel-generated dark currents in diodes fabricated in this material.

4.5.2 Dark currents

A model of the thermal- and tunnel-generated dark currents for 11-μm cutoff PbSnTe at 77 K, utilizing the above parameters for the alloy, is shown in Fig. 4.8. No distinction is made between the n- and p-regions of the diode, and the trap-assisted tunneling is assumed to occur through the same S-R centers that are responsible for the thermal recombination times. This is a perfectly reasonable assumption for PbSnTe, due to the similarity of the conduction and valence band effective masses, resulting in an intrinsic energy level that is at mid-gap. Given that τ_{SR} is independent of n- and p-type doping concentration, Eqs. (5.17) and (5.18) indicate that optimum thermally generated dark currents will vary as $1/n$ and $1/p$, provided that band-filling with its associated Burstein–Moss shift does not occur. The doping concentration assumed in Fig. 4.8 is n-type at 10^{17} cm^{-3}, and the junction model is one-sided abrupt.

A number of features are apparent. Depletion current dominates even at 77 K, unlike the case of HgCdTe with the same bandgap, requiring a lower temperature of operation for BLIP performance than for HgCdTe. The tunnel currents are low, even for the high doping concentration of 10^{17} cm^{-3}, and should not be the limiting current issue even at very low temperatures and background flux levels. This will even be true for very long cutoff wavelengths, particularly if the doping concentration is reduced to $\sim 10^{15}$ cm^{-3}. However, the overriding factor in this materials system, as it exists at the present time, is the magnitude of the depletion current, which is in turn driven by the S-R lifetime.

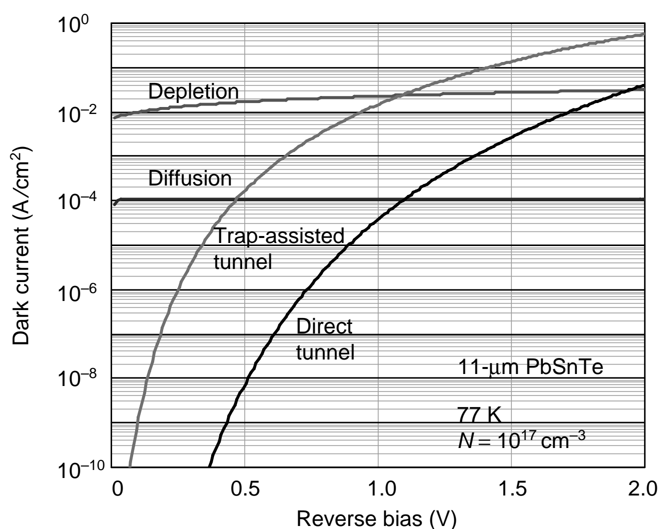


Figure 4.8 Modeled dark current components at 77 K for 11-μm cutoff PbSnTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

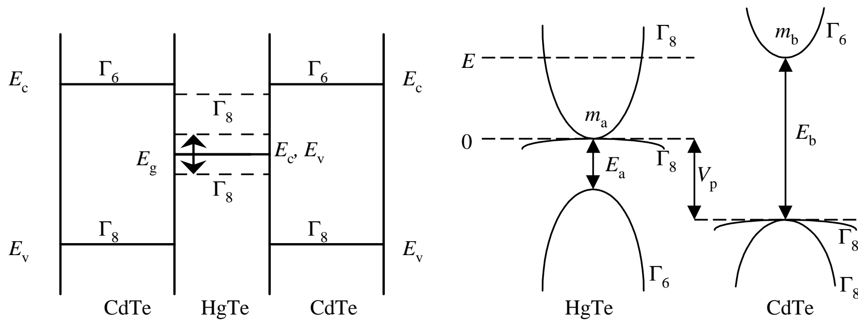


Figure 4.9 HgTe/CdTe type III superlattice and associated bandstructure.

4.6 Type III Superlattices

The Type III superlattice is described in Fig. 4.9, and in its simplest form consists of alternating layers of a wide bandgap semiconductor together with a semiconductor of the inverted gray tin-like band structure. The structure was first proposed as a potential IR detector by Schulman and McGill [39], utilizing HgTe/CdTe, and provides a large effective mass in the superlattice direction that is relatively independent of the engineered bandgap, which is a particularly attractive property for tunnel current limited, narrow bandgap device situations. Other candidate Type III superlattice materials systems would be HgSe/CdSe, βHgS/CdS, and αSn/Ge.

4.6.1 Superlattice bandstructure

Theoretical techniques of varying degrees of complexity have been employed to predict superlattice properties. The envelope function approximation (EFA) of Bastard [40, 41] is a particularly useful concept for II-VI Type III superlattice systems, and it is more than adequate to illustrate the relevant physics for IR device purposes. Considering the bandstructure in Fig. 4.9, and ignoring the heavy hole valence band, utilizing a simple two-band Kane model, the EFA gives energy bands that are solutions of the simple dispersion relation

$$\cos qd = \cos k_a l_a \cdot \cos k_b l_b - 1/2(\zeta + \zeta^{-1}) \sin k_a l_a \cdot \sin k_b l_b, \tag{4.15}$$

where q is the wavevector in the superlattice direction, and

$$\begin{aligned} \zeta &= (k_a/k_b)[E - E_b - V_p]/[E + |E_a|], \\ k_a^2 &= 8\pi^2 m_a E(E + |E_a|)/h^2|E_a| - k_{perp}^2, \\ k_b^2 &= 8\pi^2 (E - V_p)(E - V_p - E_b)/h^2 E_b - k_{perp}^2. \end{aligned}$$

The subscript a refers to HgTe, and b to CdTe; V_p is the valence band offset between HgTe and CdTe, k_a and k_b represent the k values along z, the superlattice

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

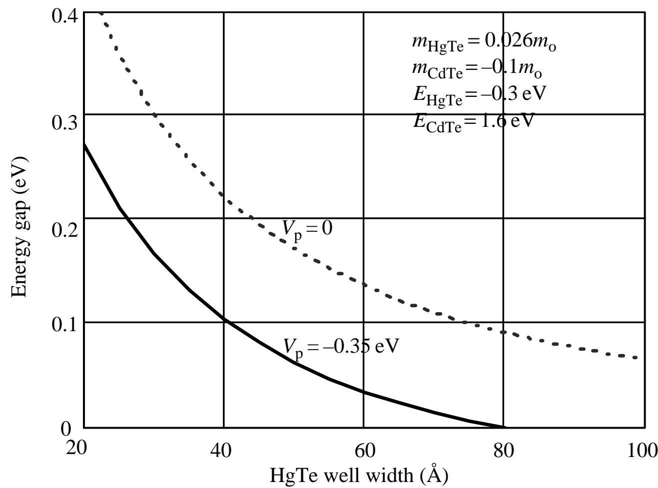


Figure 4.10 The predicted dependence of the quantum well energy gap, $E_1 - E_0$, on well width for HgTe/CdTe, for valence band offset values $V_p = 0$ and –0.35 eV.

direction, $d = l_a + l_b$, the sum of the individual layer thicknesses, and k_{perp} is the k vector perpendicular to z. The real solutions of Eq. (4.15) lie between $+1$ and -1.

For sufficiently thick layers of CdTe, the superlattice can be viewed as a series of isolated HgTe quantum wells, whose energy levels are determined by the solution of Eq. (4.15) with l_b very large, namely $\tan(k_a l_a/2) + |\zeta| = 0$. The solution involving the lowest level as a function of HgTe layer thickness, assuming $|E_a| = 0.3$ eV, $E_b = 1.6$ eV, $m_a = 0.026$, $m_b = -0.1$, $V_p = 0$, is shown in Fig. 4.10. The heavy hole band has been ignored but can be included as a separate non-interacting band located as shown in Fig. 4.9 with the same effective mass for both HgTe and CdTe. The effect of any valence band offset can be accommodated by a simple one-band analysis. The density of states for the heavy hole is very large relative to the light hole band and will dominate optical transitions to the conduction band. Hence, the effective IR bandgap is the value for the heavy hole to conduction band transition. The data in Fig. 4.10 are in agreement with the published theoretical data of Smith et al. [42] for HgTe/CdTe superlattices with $l_a = l_b$, indicating that to first order, the energy levels of the superlattice (and hence the IR bandgap) depend only on the HgTe layer thickness and the barrier function. Recent modeling [43] indicates that the assumption of $V_p = 0$ (the common anion rule) is not correct, and that $V_p = -0.35$ eV.

The modification to the HgTe conduction band energy levels caused by this change in barrier function is shown in Fig. 4.10.

As the thickness of the CdTe layers in the superlattice is reduced, the coupling between the HgTe quantum wells increases, resulting from the increased probability of electrons tunneling through the barrier. The discrete energy levels associated with the HgTe quantum well become bands of finite width and can support transport of electrons in an applied electric field along z, the superlattice direction.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

This transport is described in terms of an effective mass in the z-direction, which is essentially independent of the energy level in question (and hence the IR bandgap). This effective mass is determined primarily by the CdTe layer thickness l_b and can be calculated from the appropriate $E-q$ curves given by Eq. (4.15) by performing numerical differentiation.

4.6.2 Band offsets and strain

Valence band offsets significantly affect the band structure and associated cutoff wavelength of the superlattice, as shown in Fig. 4.10. The largest reported offset value is -0.35 eV, which is enough to generate a degree of quantization even in the heavy hole band. This effect must be taken into account in tailoring the superlattice bandgap for a particular cutoff wavelength.

Considerations of lattice constant mismatch between HgTe and CdTe ($\Delta a/a = 3 \times 10^{-3}$) indicate that in any IR superlattice, whether lattice matched to the substrate or otherwise, a degree of strain will exist between alternating layers. For a CdTe substrate supporting a superlattice that is sufficiently thick, which misfit dislocations have formed at the superlattice–substrate interface, the major portion of the superlattice will be free-floating, with an average lattice constant determined by the equivalent composition of the layers. The HgTe layers will be subjected to a biaxial tensile stress, and the CdTe layers to a biaxial compression. These stresses can be reduced to equivalent hydrostatic and uniaxial stress components. The hydrostatic stress effectively changes the direct bandgap values without removing any band degeneracies. However, the uniaxial stress components lift the valence band degeneracy, resulting in the modified band structure shown in Fig. 4.11.

In the direction of the uniaxial stress (the superlattice direction) a splitting is observed in the Γ_8 light and heavy mass bands, and at the same time a mixing of the light and heavy mass bands occurs in the plane of the layers. Calculations

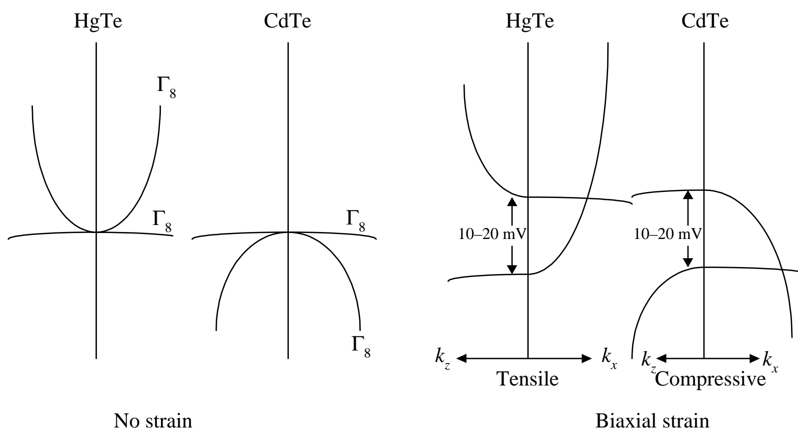


Figure 4.11 Strain effects in HgTe/CdTe superlattices.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

indicate that the splitting of these bands, although somewhat orientation dependent, is $\sim$10 mV.

The strain due to lattice mismatch between superlattice layers can also generate internal electric fields, associated with the piezoelectric nature of II-VI semiconductors, as was first pointed out by Smith [44, 45]. The effect is strongly orientation dependent, being zero for [001] and a maximum for the [111] growth direction. The magnitude of the effect is determined by the degree of lattice mismatch and the values of the elastic moduli and piezoelectric coefficients for HgTe and CdTe. The piezoelectric coefficient for $x = 0.22$ HgCdTe is 0.0136 C/m^2, and for CdTe is 0.031 C/m^2. Extrapolating from the published data on $x = 0.22$ HgCdTe, the piezoelectric-induced electric field in the HgTe layers of any superlattice will be $\sim 1.75 \times 10^3$ V/cm per 0.1% of lattice mismatch. This value should exert a negligible effect on the electronic wavefunctions of the superlattice, but it may need to be considered in the performance of diode devices fabricated in the material.

4.6.3 Interdiffusion in HgTe/CdTe superlattices

Measurements [45] of the interdiffusion of HgTe into CdTe indicate that the superlattice is relatively stable at the growth temperatures and times employed, but not at the temperatures typically used for certain aspects of HgCdTe device processing, such as impurity activation, native defect reduction, and surface passivation. We know that MOCVD growth of $Hg_{1-x}Cd_xTe$ employs layer interdiffusion of alternating HgTe and CdTe layers of the requisite thickness during the growth run to achieve a uniform alloy composition. Growth typically occurs at $\sim$360° C, and layer thicknesses are $\sim$1000 Å, indicating that the processing of superlattice thicknesses of tens of Angstroms at $\sim$250° C for any length of time will be an issue.

The interdiffusion coefficient at high temperatures for a composition x is given by an expression such as

$$D = A \exp[-Bx - (C/T)], \tag{4.16}$$

where A, B, and C are constants. The diffusion coefficient for Cd is much lower than for Hg and is the rate-limiting process for interdiffusion. The interface between HgTe and CdTe layers remains relatively abrupt during interdiffusion, and the CdTe is dissolved away layer by layer, resulting in a reasonably uniform HgCdTe layer to replace the HgTe. The CdTe barrier layers become thinner, and the effective mass of the electron becomes significantly less in the layers that were originally HgTe. The end result is a wider bandgap superlattice than the as-grown system.

4.6.4 Misfit dislocations

An epitaxial layer may be grown on a substrate with a different lattice parameter without the formation of misfit dislocations provided that the thickness is less than a critical value. The formation of an array of misfit dislocations is favored when the

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

energy of the misfit array becomes less than the energy of the elastically deformed coherent epitaxial layer. A similar criterion applies to HgTe/CdTe superlattices grown on CdZnTe substrates regarding the strain in the superlattice, its thickness, and the formation of a misfit array between the superlattice and the substrate. The degree of lattice mismatch between the superlattice and substrate depends on the average lattice constant of the superlattice. Typical values are $\sim$0.2%. The critical thickness given by the Matthew's criterion [47] for this mismatch is $\sim$800 Å. Larger thicknesses result in a partitioning of the total misfit energy between elastic strain and dislocations. The greater the thickness of the film, the more the misfit is accommodated by dislocations.

Misfit dislocations are undesirable in the device structure if they form in high field regions such as the depletion region of a diode. These macroscopic defects are thought to result in increased tunnel currents and $1/f$ noise in diodes fabricated on HgCdTe alloys.

4.6.5 Absorption coefficient

Early thoughts on the absorption coefficient of HgTe/CdTe superlattices were concerned with the fact that the inverted band structure of HgTe results in a primary IR absorption mechanism involving forbidden transitions from the Γ_8 heavy hole band to the Γ_8 conduction band. However, the strong admixing of the Γ_8 (p-like) and Γ_6 (s-like) wavefunctions in the conduction band away from $k = 0$ leads to significant absorption in the superlattice. Numerical calculations close to the band edge have been carried out to confirm this. To understand the absorption coefficient of the HgTe/CdTe superlattice it is instructive to consider photon absorption in HgTe [47a].

The absorption coefficient at an energy E is given by [7]

$$\alpha(E) = [hc\mu q^2/2Em_o^2]M^2\rho(E), \tag{4.17}$$

where n is the refractive index, M is the matrix element representing the electron–photon interaction between initial and final states, and $\rho(E)$ is the density of states. A two-band Kane [47b] model for the admixing of the $|S\rangle$ and $|P\rangle$ wavefunctions gives a matrix element of $M^2 = 3m_o^2|E_a|E/[8m_a(2E + |E_a|)]$ averaged over all polarizations in the plane of the surface, where m_a is the effective mass for the conduction band at $k = 0$, and E_a is the negative bandgap of HgTe.

The density of states term is given by

$$\rho(E) = 8\pi^2 m_a^{3/2}[1 + 2E/|E_a|][2E(1 + E/|E_a|)]^{1/2}/h^3. \tag{4.18}$$

The resulting absorption coefficient for HgTe is

$$\alpha(E) = 2.45 \times 10^4[E(1 + E/|E_a|)]^{1/2}\,\text{cm}^{-1}, \tag{4.19}$$

where E and E_a are in units of eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

A similar consideration for the HgCdTe alloy gives

$$\alpha(E) = 1.5 \times 10^5 (m_c/m_o)^{1/2}[E(E - E_g)/E_g]^{1/2}, \tag{4.20}$$

where m_c is the alloy conduction band effective mass at $k = 0$, and E_g is the alloy bandgap. Substituting $m_c \approx 7 \times 10^{-2}\, E_g m_o$ then

$$\alpha(E) = 4 \times 10^4 [E(E - E_g)]^{1/2}\,\text{cm}^{-1}. \tag{4.21}$$

The absorption coefficient for the HgTe/CdTe superlattice is given by Eq. (4.16) with appropriate values for M^2 and $\rho(E)$. For reasonably large values of l_b, the superlattice wavefunction will be localized primarily in the HgTe with k_a and k_b values given by the dispersion relation Eq. (4.15), and the energy bandwidth (and hence range of k_a values) will be small. Under these conditions, the matrix element M^2 is given approximately by the corresponding value in HgTe at the energy in question.

The energy–momentum relationship for motion through the superlattice for large l_b is given by

$$E = E_1 + E_0(1 + \cos qd) + E_{\text{perp}}, \tag{4.22}$$

where the first two terms represent the quantized motion along the superlattice direction, E_1 is the superlattice bandgap, $2E_0$ the bandwidth, and E_{perp} is the energy in the plane of the HgTe layers. Allowing for nonparabolicity in E_{perp}, the density of states is given by

$$\rho(E) = 4\pi m_a[1 + 2E/|E_a|]/h^2 d, \qquad \text{for } q > \pi/\text{d}, \tag{4.23}$$

and

$$\rho(E) = 4m_a[1 + 2E/|E_a|]q_o/h^2, \qquad \text{for } q < \pi/\text{d},$$

where q_o is given by

$$h\nu = E_1 + E_0(1 - \cos q_o d). \tag{4.24}$$

The absorption coefficient for the superlattice is

$$\alpha(h\nu) = 9.2 \times 10^5/d\,\text{cm}^{-1}, \qquad \text{for } q > \pi/d, \tag{4.25}$$

and

$$\alpha(h\nu) = 2.93 \times 10^5 q_o \text{cm}^{-1}, \quad \text{for } q < \pi/d,$$

where d is in Angstroms, and q_o is given by Eq. (4.24).

Figure 4.12 shows the predicted variation of absorption coefficient with photon energy for HgTe, a HgCdTe alloy with $E_g = 0.1$ eV, and a HgTe/CdTe superlattice with a bandgap of 0.1 eV. The superlattice absorption is significantly larger than that of the HgCdTe alloy with an identical bandgap, implying that superlattice thicknesses of $\sim$1 μm should be adequate for absorbing IR radiation. This is not a peculiarity of the superlattice, but it is due to the high absorption coefficient

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

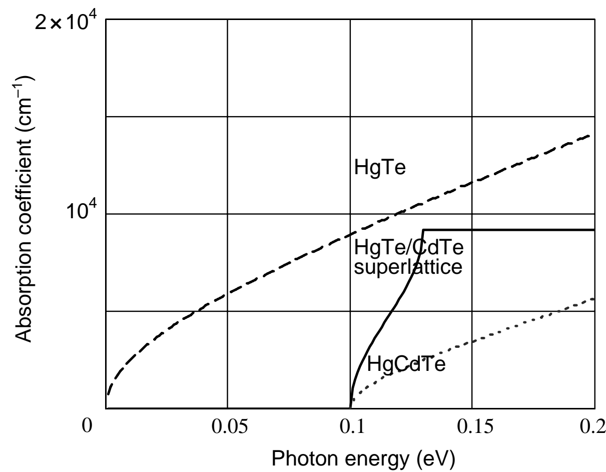


Figure 4.12 Model of absorption coefficient vs. photon energy for HgTe, HgCdTe with a bandgap of 0.1 eV, and a HgTe/CdTe superlattice with a bandgap of 0.1 eV.

associated with HgTe, which is in turn a direct result of the large effective mass (and hence density of states) of HgTe relative to the HgCdTe alloy. The matrix element for optical transitions in the photon energy range shown in Fig. 4.12 is actually smaller than in the HgCdTe alloy due to the smaller admixture of $|S>$ states in HgTe, but this is more than compensated for by the larger density-of-states factor.

4.6.6 Effective mass

The large effective mass of HgTe will have a significant effect on tunnel currents in junction devices fabricated in HgTe/CdTe superlattices. Theoretically, the effective mass in the superlattice direction can be made arbitrarily large by a suitable selection of CdTe layer thickness; however, the width of the associated energy levels will be very small. IR devices are required to support built-in electric fields of $\sim$1 V/ μm, and under such fields the offset of energy levels in adjacent HgTe quantum wells will be sufficient to strongly inhibit tunneling between wells and hence carrier transport. (This phenomenon is analogous to resonant tunneling in coupled quantum wells.) If minority carrier transport is required in the device, then the width of the relevant energy levels must be sufficiently large to allow tunneling between adjacent quantum wells in the electric field region. The CdTe layer width must be reduced to a value consistent with a level width of 5−10 mV for the HgTe quantum wells, to sustain carrier transport in fields of 1V/ μm. Calculations of the effective mass in the superlattice direction for such CdTe layer thicknesses yield values that are approximately equal to the effective mass of electrons in bulk HgTe. A further important consideration with regard to possible tunnel current reduction in the superlattice device concerns minority carrier transport in the plane of the HgTe layers. Tangential fields at the periphery of a junction device are limited by tunneling considerations in the plane of these layers, and they will be determined by the effective mass in HgTe at the energy level E_n, given by $m_a^* = m_{ao}^*[1 + 2E_n/|E_a|]$,

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

where m^*_{ao} is the effective mass at the conduction band minimum. These arguments suggest that tunnel currents in HgTe/CdTe superlattice junction devices will be limited to values consistent with an effective mass in the range $0.03m_o$–$0.04m_o$ (i.e., the approximate effective mass of HgTe).

Tunnel breakdown in HgCdTe diodes can be due to direct band-to-band tunneling, or to tunneling via Shockley–Read centers in the bandgap. Tunneling due to direct band-to-band transitions through a parabolic barrier in a uniform electric field E is given by

$$\begin{aligned} J_{td} &= [q^3(2m^*)^{1/2}EV/(\pi h^2 E_g^{1/2})]\exp -[\pi^2(m^*/2)^{1/2}E_g^{3/2}/qEh] \\ &= [5.7 \times 10^8(m^*/E_g)^{1/2}E^3/N]\exp[-3.5 \times 10^7 m^{*\,1/2}E_g^{3/2}/E], \end{aligned} \tag{4.26}$$

where V is the applied bias, and N the doping concentration. For the HgCdTe alloy system the relative effective mass (m^*/m_o) is empirically given by [6] $\sim(7 \times 10^{-2}E_g)$ with E_g in eV.

Tunneling through Shockley–Read centers located in a parabolic barrier and uniform electric field is given by

$$\begin{aligned} J_{tsr} &= [\pi^2 q^2 m^* M^2 N_r W E/h^3/(E_g - E_r)]\exp[-\pi(m^*/2)^{1/2}E_g^{3/2}F(a)/qEh] \\ &= [2.75 \times 10^{-7}N_r E^2/N(E_g - E_r)]\exp[-1.27 \times 10^7 m^{*\,1/2}E_g^{3/2}F(a)/E], \end{aligned} \tag{4.27}$$

where $F(a)$ is a geometrical factor depending on the position of the center in the bandgap. For mid-gap, $E_r = E_g/2$ and $F(a) = \pi/2$. M is a matrix element associated with the trap potential and value of $M^2(m^*/m_o) = 2 \times 10^{-24}$ eV^{2}cm^3 is assumed.

A calculation of the reduction in tunnel current for a diode fabricated on a superlattice with a cutoff wavelength of 14 μm, as compared to a HgCdTe alloy of the same bandgap is shown in Fig. 4.13, for both the direct and Shockley–Read assisted cases. The Shockley–Read center is assumed at mid-gap with a density $N_r = 2 \times 10^{13}$ cm^{-3}, the doping concentration is 3×10^{14} cm^{-3}, and $m^*_a = 0.03\,m_o$.

4.6.7 Minority carrier lifetime

Minority carrier lifetime data on HgTe/CdTe superlattices is sparse. Early values for both n-type and p-type superlattices with a variety of cutoff wavelengths were <20 ns. However, values by the GE group [48] are more in keeping with accepted Auger band-to-band recombination theory as applied to superlattices, at relatively high temperatures, for cutoff wavelengths in the 4–15 μm cutoff wavelength range. There is evidence of Shockley–Read recombination at temperatures below 125 K, but minority lifetime values in the range of 0.2–20 μs are obtained for n-type superlattices at 77 K. Smaller lifetime values are obtained for p-type superlattices at 77 K with stronger evidence of Shockley–Read effects.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

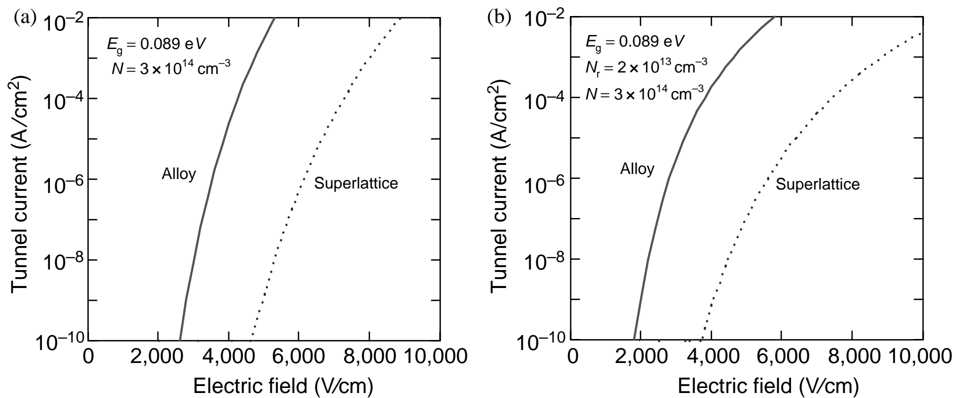


Figure 4.13 Predicted tunnel current vs. electric field for a HgCdTe alloy and a HgTe/CdTe superlattice with a 14 μm cutoff wavelength for (a) direct tunneling and (b) for tunneling via mid-gap states; ($n = 3 \times 10^{14}$ cm^{-3}, $N_r = 2 \times 10^{13}$ cm^{-3}).

These data suggest that minority carrier lifetimes in HgTe/CdTe superlattices are somewhat inferior to those found in the equivalent HgCdTe alloy. However, the published $I-V$ characteristics [49] for mesa-etched p^+/n homojunction superlattice diodes, with a base layer indium doping of 10^{16} cm^{-3}, indicate a larger breakdown voltage due to tunneling via bandgap states than would be predicted for the HgCdTe alloy with the same bandgap. The observed breakdown is in approximate agreement with a superlattice effective mass ~0.03 m_o. Similar observations [50] have been made on MWIR planar MIS devices fabricated on n-type HgTe/CdTe superlattices. These results indicate that the basic tenet that Type III superlattices offer a reduced tunneling capability relative to the alloy HgCdTe are indeed true, despite the obvious difficulty of achieving dark current values that are competitive with the corresponding alloy and the problems associated with the limited range of processing temperatures that can be employed.

4.7 Type II Superlattices

The Type II superlattice, typified by the InAs/GaSb system, is shown in Fig. 4.14(a). The conduction band edge of InAs lies below the valence band edge of GaSb; however, layer quantization is employed to produce a positive bandgap between the InAs and GaSb. The quantization effects for this materials system can be modeled using the envelope function approximation used above for Type III superlattices. Bandgaps of 0.1 eV are achieved with layer thicknesses in the 50–100 Å range. However, the electrons and holes are localized in different layers, and the optical matrix element for photon absorption depends directly on the degree of electron and hole wavefunction overlap in the layers. For layer thicknesses appropriate for LWIR detection in the InAs/GaSb system, this overlap is small, resulting in low absorption coefficients. This problem is eliminated by the proposal of Smith and Mailhiot [51] to replace GaSb with the alloy GaInSb, thus utilizing the effects of strain associated with the increasing lattice constant of $Ga_{1-x}In_xSb$ with x, relative to the lattice

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

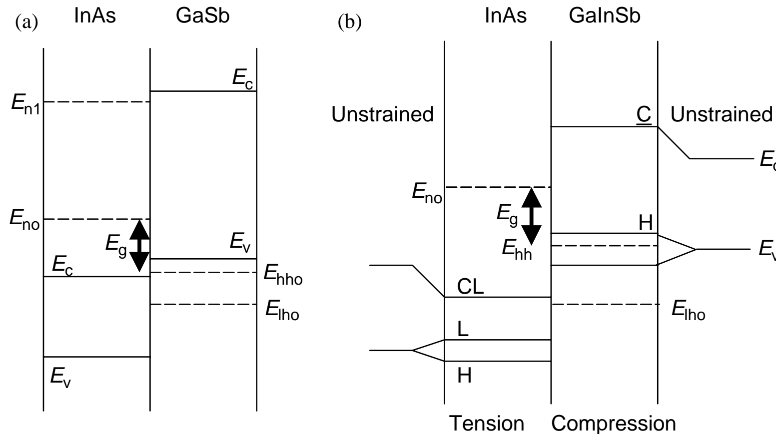


Figure 4.14 Type II superlattice (a) InAs/GaSb and (b) InAs/GaInSb with effects of strain on band structure.

constant of InAs. The strain places the InAs in biaxial tension and the GaInSb in biaxial compression. Again, these stresses can be reduced to equivalent hydrostatic and uniaxial stress components. The hydrostatic stress effectively changes the direct bandgap values without removing any band degeneracies. However, the uniaxial stress components lift the valence band degeneracy, resulting in the modified band structure shown in Fig. 4.14(b). The increased conduction and valence band overlap allows layer thicknesses ~25 Å to provide both the desired LWIR bandgap values and absorption coefficients $>10^3$ cm^{-1}. As with the Type III superlattice, the optical matrix element for the InAs/GaInSb system is less than that for the HgCdTe alloy with the same bandgap, but this is compensated by the large conduction band density of states associated with the electron effective mass values of the superlattice, which are typically $\sim 0.04\, m_o$. These large values of effective mass will also result in significantly lower values of tunnel current, but somewhat larger values of thermally generated dark current, for diodes fabricated on this materials system relative to the alloy HgCdTe, assuming that similar values of doping concentration are attainable, namely $\sim 1-5 \times 10^{14}$ cm^{-3}.

4.7.1 Minority carrier lifetime

As pointed out by several authors [52, 53], the bandgap engineering approach can also be used to good effect in reducing the contribution of band-to-band Auger mechanisms to the recombination and generation of minority carriers in this materials system. Minority carrier recombination via the Auger1 and Auger7 processes is shown in Fig. 4.15 for the HgCdTe alloy. The modification to the band structure of InAs and GaInSb caused by the strain in the layers is qualitatively shown in Fig. 4.16(a) for $Ga_{0.6}In_{0.4}Sb$ grown on the (111) direction, after Mailhiot and Smith [54]. The IR properties of the superlattice are determined by the layer quantization in the conduction band of InAs and the light and heavy

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

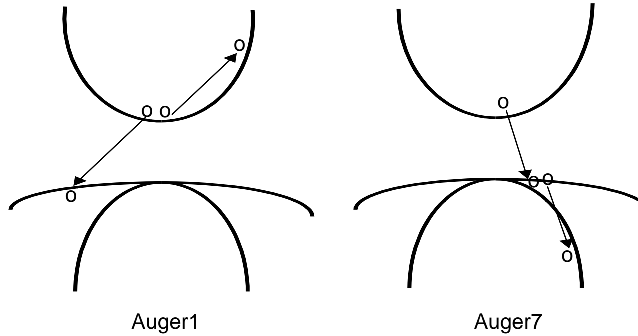


Figure 4.15 Auger1 and Auger7 recombination transitions in HgCdTe alloy.

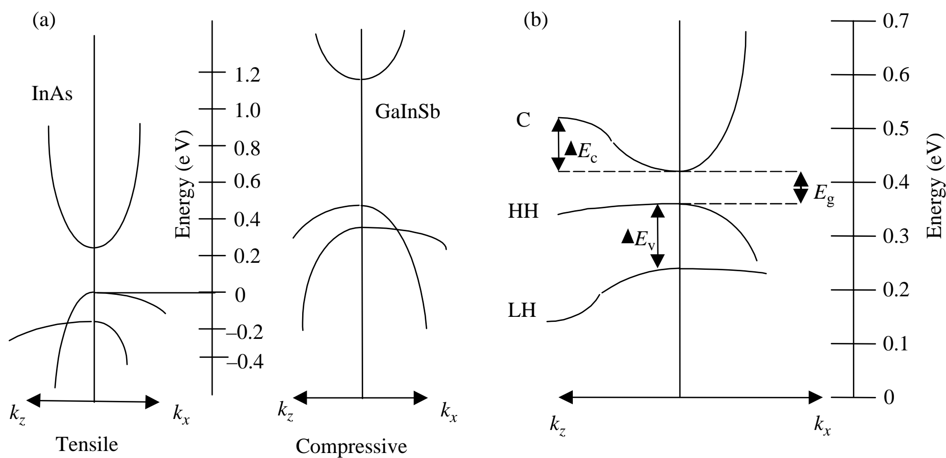


Figure 4.16 The qualitative effect of (a) biaxial strain on the band structure of the Type II superlattice bulk materials and (b) layer quantization on the relevant IR bands for the Type II superlattice, for the in-plane (x) and superlattice (z) directions.

hole bands of GaInSb. The band structure for an arbitrary superlattice is shown in Fig. 4.16(b).

For the Type II superlattice, the contribution of Auger1 to minority carrier generation and recombination, involving two electrons and a hole, can be reduced by making the width of the lowest electron sub-band in the z-direction, ΔE_c, smaller than the bandgap, E_g, of the superlattice. This limits the available final state of the electron to values $<E_g$ above E_{no}, and hence eliminates Auger1 transitions for momentum in the z-direction. The width of the superlattice conduction band in InAs varies inversely as the thickness of the GaInSb layers. In the plane of the layers, the heavy hole band assumes a light hole character due to the biaxial compression in the GaInSb layers and possesses considerable curvature, as shown in Fig. 4.16(b). Thus, the participating hole that simultaneously satisfies the momentum and energy conservation requirements of the Auger transition in the plane of the layers resides in a region of phase space that is not populated, resulting in a significant reduction in the Auger1 rate. This suppression of Auger1 recombination is achieved by engineering

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

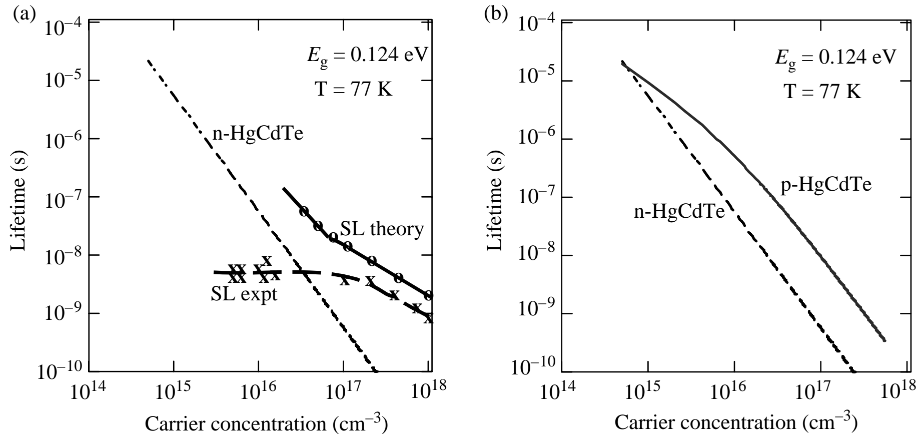


Figure 4.17 Measured lifetime at 77 K for (a) a LWIR Type II superlattice [55] and (b) n- and p-type 10 μm HgCdTe as a function of carrier concentration. Reused with permission from E. R. Youngdale, J. R. Meyer, C. A. Hoffman, F. J. Bartoli, C. H. Grein, P. M. Young, H. Ehrenreich, R. H. Miles, and D. H. Chow, *Applied Physics Letters*, 64, 3160 (1994). Copyright 1994, American Institute of Physics.

the band structure of the Type II superlattice to greatly reduce the volume of phase space available for the Auger transitions. Similar techniques can be employed to reduce the magnitude of Auger7. In this case, the splitting of the participating valence bands [in Fig. 4.16(b) these are the heavy and light hole bands of the GaInSb layer] due to both mismatch strain and layer quantization, ΔE_v, is engineered to be $> E_g$, again forcing at least one of the participating carriers (in this case a hole) into an unoccupied region of phase space. The splitting of the valence bands is increased by the use of narrow GaInSb layers and as such is not necessarily compatible with the Auger1 suppression requirement. It is not clear whether the $\Delta E_v > E_g$ condition can be met for the MWIR using acceptable values of layer mismatch, but it can certainly be achieved for LWIR.

The limited lifetime data available [55] on Type II superlattices indicate that the bandgap engineering model correctly accounts for Auger1 recombination in LWIR material at high electron concentrations, as shown in Fig. 4.17. However, at electron concentrations $<2 \times 10^{17}$ cm^{-3}, the lifetime appears limited by a Shockley–Read mechanism to a value of 5 ns, far below the reported values for LWIR n-HgCdTe, which is Auger1 limited to doping concentrations of $<10^{15}$ cm^{-3}, with typical $\tau \approx 2-10\,\mu$s at 77 K. The measured lifetimes for both n- and p-type 10-μm HgCdTe at 77 K, as a function of doping concentration, are shown in Fig. 4.17(b). State-of-the-art extrinsically doped p-HgCdTe is presently limited by a Shockley–Read mechanism for doping concentrations $<10^{16}$ cm^{-3}, but the Auger7 lifetime is considerably greater than Auger1. Recent estimates [56] of room-temperature Auger1 coefficients in various type II MWIR quantum wells, determined by photoconductive response and laser threshold measurements, indicate values an order of magnitude less than for MWIR Type I bulk and quantum well III-V semiconductors. Moreover, these results do not exhibit the predicted sensitivity of the Auger recombination rate to valence band engineering discussed above.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The Auger1 coefficients (defined as $\gamma_{Ai} = [2\tau_{Ai} \cdot n_i^2]^{-1}$) for bulk MWIR HgCdTe are given by Eqs. (4.4) and (5.4) and are virtually the same as for the corresponding III-V bandgap. Interestingly enough, as pointed out earlier, the same equations predict Auger coefficients for the ideal bulk lead salt alloys that are many orders of magnitude lower than both HgCdTe and Type II superlattices.

4.8 Direct Bandgap Materials: Conclusions

In the author's opinion, the following conclusions can be drawn regarding the current crop of available direct bandgap IR materials systems.

4.8.1 HgCdTe

HgCdTe currently has no equal as an IR detector material.

- Composition can be tuned to cover the complete IR spectrum from 1 to 20 μm.
- Phase fields of all alloy compositions are understood and defects controllable.
- Crystalline quality of epi-material is good, with dislocations of $\sim 10^4$–10^5 cm^{-2}.
- Band structure is well understood.
- Lifetime is limited primarily by band-to-band Auger recombination.
- Dark currents are diffusion-current dominated.
- Controlled n- and p-doping concentrations from 10^{14} to 10^{18} cm^{-3} are readily introduced.
- Surfaces are well understood.

HgCdTe diodes offer diffusion-limited performance over the complete IR spectrum, with lifetimes typically limited by Auger recombination. The only perceived issue may lie in the capability of this material to detect very long wavelength IR photons, past $\sim$20 μm, due to the potential limitation of tunnel currents associated with the narrow bandgap and small effective mass. This issue is discussed further in Chapter 8.

4.8.2 InSb

InSb is a good IR material for a very limited set of circumstances:

- No tunability;
- Well-understood band structure;
- Excellent crystalline quality, with dislocations $\sim$100 cm^{-2};
- Lifetime limited by S-R centers to values $\sim 3 \times 10^{-7}$ s. (S-R center is unknown);
- Depletion current dominated by these S-R centers, for $T < 125$ K;

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

- Bandgap decreases with temperature, making high-temperature operation difficult;
- Surfaces not easily passivated, as is the case with all III-V alloys;
- Low n- and p-doping readily achieved.

InSb diodes offer excellent MWIR performance at 77 K, but they are depletion-current dominated and hence are very limited in their capability for BLIP operation at higher temperatures, and/or lower background flux environments, compared to HgCdTe.

4.8.3 PbSnTe

PbSnTe is potentially an excellent IR material.

- It has excellent bulk crystalline quality.
- Phase fields are understood, but defect densities are not easily controlled.
- Doping concentrations are typically $>10^{17}$ cm^{-3}, although $\sim 10^{15}$ cm^{-3} n- and p-type have been reported after extensive phase field anneals.
- Band structure is well understood.
- Conduction and valence band effective masses are essentially equal and relatively large.
- Auger recombination is suppressed naturally by the very band structure of the material that is responsible for the equal effective masses.
- High dielectric constants provide wide depletion regions and potentially lower electric fields.
- Limiting lifetime for n- and p-material is $\sim 2 \times 10^{-8}$ s and is relatively temperature independent. S-R center is unknown.
- Limited knowledge [57] of surfaces.

PbSnTe was adjudged to be a non-player in the field of IR detection in the early 1970s, primarily because of its high dielectric constant and relatively poor lifetime. The material itself is also mechanically soft, presenting potential issues for FPA hybridization. However, it does offer the built-in capability of strong suppression of Auger recombination for both n- and p-type material. For this reason it has been utilized for IR lasers and emitters to wavelengths as long as 30 μm, employing large doping concentrations, such that the low value of τ_{SR} is unimportant. It is tantalizing to think that a bulk alloy material already exists that is capable of Auger suppression, with large values of effective mass and dielectric constant and hence immunity to tunneling. These attributes essentially constitute the rationale that has been put forward for the expenditure of considerable resources on Type II superlattices. It may well be easier from a physics and a resource point of view to address the three issues that are perceived to be delinquent in PbSnTe as an IR detector material, namely (1) understanding and improving the S-R lifetime limitation, (2) investigating the alloy phase fields, so as to reproducibly control defect concentrations down to the 10^{14} cm^{-3} level, and (3) accommodating the mechanical limitations of the material

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

in diode and FPA processing. Success in such an endeavor would offer the hope of room-temperature photon detection over a large fraction of the IR spectrum.

4.8.4 Type III superlattices

In the author's opinion, these superlattices, while offering a wealth of interesting physics for investigation, are the least practical of the direct bandgap materials, due to the huge issue of interdiffusion of the composite layers at temperatures that are typically required for diode and FPA processing.

4.8.5 Type II superlattices

From a physics point of view, the Type II superlattice is an extremely attractive proposition. It offers the potential for

- Bandgap engineering,
- Larger effective masses than HgCdTe for the equivalent bandgap, and
- The possibility of Auger suppression in both the conduction and valence bands.

These attributes again would lead to a capability for IR detection at very long wavelengths, potentially at high temperatures, and possibly even at room temperature. Early work on this materials system has been encouraging, but many obstacles remain. The lifetime values are S-R limited typically at <10 ns and are either bulk or surface related. Band-to-band limited lifetime values $>10\,\mu s$ are mandatory at doping concentrations that are comparable with HgCdTe, namely $<10^{15}\,cm^{-3}$. Current doping concentrations are an order of magnitude or more higher than this. Surface passivation of this materials system will undoubtedly also be an issue, as with all III-V systems.

4.8.6 Final thoughts

The above conclusions can be easily summarized. The materials system of choice is HgCdTe, with the possible exception of VLWIR detection at $>20\,\mu m$, although this is still an open question. InSb fills a niche for MWIR detection at ~ 77 K. This will gradually disappear as mandated by higher system operating temperatures. Type III superlattices form an interesting academic exercise but will play no role in future IR systems. A better case can be made for Type II superlattices, but significant materials challenges remain before they will ever displace HgCdTe. The joker in the pack may well be PbSnTe or some other variation of the lead salt alloys. The virtual absence of Auger recombination, combined with relatively heavy effective masses and high dielectric constants, suggests that this materials system may yet provide the only viable alternative to HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 5

HgCdTe: Material of Choice for Tactical Systems

HgCdTe is the material of choice for second-generation tactical IR systems. It offers the highest performance at the highest operating temperature of any materials system currently available, and it is easily adapted to advanced third-generation system concepts. Numerous treatises have been written on the growth and materials properties of HgCdTe. We will not attempt to repeat them here. In this chapter, we will concentrate on the properties of HgCdTe that are relevant to the design and fabrication of high-quality HgCdTe FPAs.

5.1 HgCdTe Material Properties

5.1.1 Material growth

HgCdTe today is grown primarily by epitaxial techniques, either from the liquid or vapor phases, on various substrate materials. An example of the phase field of HgCdTe is shown in Fig. 5.1 for a composition $x = 0.3$, together with the typical growth conditions for Te-rich and Hg-rich liquid phase epitaxy (LPE) and molecular beam epitaxy (MBE). Metal–organic chemical vapor deposition (MOCVD) is more difficult to quantify on a phase field as the growth technique typically embodies the growth, and subsequent interdiffusion, of alternating layers of CdTe and HgTe at $\sim$360°C.

HgCdTe is dominated by metal vacancies on both sides of the phase field, and for large n_{i} values, the vacancy concentrations at the phase field extremities are given by [58]

$$C_{\mathrm{xste}} = 2.81 \times 10^{22}(1 - x)\exp[-0.65q/kT] \ \text{(Te-rich), and}$$
$$C_{\mathrm{xshg}} = 2.50 \times 10^{23}(1 - x)\exp[-1.0q/kT] \quad \text{(Hg-rich).} \qquad (5.1)$$

n-Type material is obtained typically by doping with indium, but it obviously requires a subsequent anneal under Hg at a temperature in the 220°C range, to lower the vacancy concentration to a sufficiently low value.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

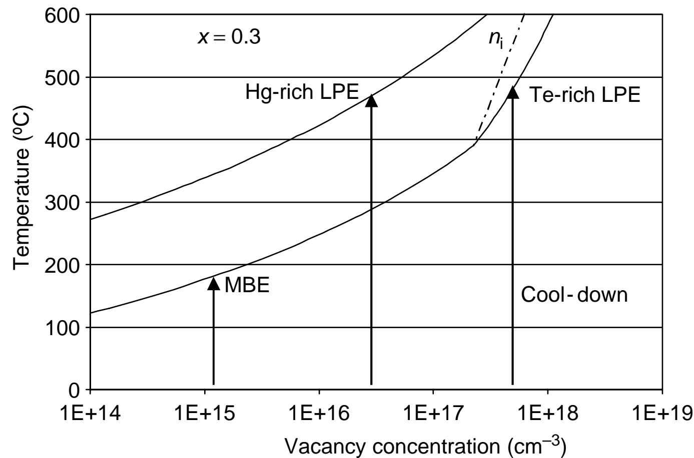


Figure 5.1 Phase field of $x = 0.3$ HgCdTe, illustrating LPE growth at 475°C under Hg-rich and Te-rich melts, and MBE growth at 170°C.

5.1.1.1 LPE

LPE growth can be achieved from melts on either the Hg-rich or the Te-rich side of the phase field and typically involves temperatures of 450−500°C. Lattice-matched CdZnTe substrates are used, typically with a (111)B or (111)A orientation. Growth of uniform epilayers from an Hg-rich melt is complicated by significant cadmium depletion of the melt. This can be alleviated to a certain extent by the use of large melts and high temperatures, although Hg vapor pressures can become high. Typical x gradients in Hg-rich layers are $\sim 5 \times 10^{-4}/\mu\text{m}$. Cooling down from the growth temperature on the Hg-rich side of the phase field is not an issue. As indicated in Fig. 5.1, the material can be programmed to remain within the phase field down to very low temperatures, and Te precipitation will not be a factor. The incorporation of impurities in Hg-rich epilayers is enhanced relative to Te-rich growth by larger segregation coefficients for many of the common impurities. Growth from a large Te-rich melt results in an order of magnitude shallower x-gradients than the Hg-rich case, typically $\sim 5 \times 10^{-5}/\mu\text{m}$. The segregation coefficients of many impurities in Te-rich melts are $\ll 1$. This aids in LPE layer purity but conversely can make impurity incorporation by doped melt growth difficult. The growth occurs on the Te-rich side of the phase field and cooling down from the growth temperature can result in Te precipitation in the layer if excursions outside the phase field are allowed during withdrawal. This in turn necessitates a subsequent high-temperature anneal on the Hg-rich side of the phase field to allow for the annihilation of these precipitates by in-diffusing Hg, prior to any final low-temperature anneal to set the required degree of stoichiometry in the layer.

Large melt Te-rich LPE is preferred for the growth of production material because of the purity and compositional uniformity of the grown layers. The segregation coefficient argument for Te-melt growth is illustrated by arsenic. The incorporation of 10^{15} cm^{-3} arsenic doping in MWIR HgCdTe can be achieved from

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

both Hg-rich and Te-rich melts. The required arsenic level in a 5 kg Hg-melt is ~10 μg, whereas the arsenic required in a similar Te-rich melt is ~1 atomic%.

It is of interest to consider the phase field of the substrate material CdZnTe. The melting point is ~1090°C and the phase field, unlike the case of HgCdTe, encompasses both sides of stoichiometry. It is thus possible to achieve metal- or Te-rich behavior. Substrates are grown by horizontal or vertical Bridgman and are typically Te-rich for use in standard LPE growth from Te-rich melts. The advantage of Te-rich substrates is that mobile impurities in the substrate material tend to migrate to Te precipitates, because of their low segregation coefficients, resulting in purer epitaxial layers. An example of the relevant impurities is shown in Table 5.1, for both DRS and NIMTEC CdZnTe, as given by glow discharge mass spectroscopy (GDMS), together with the impurities present in

Table 5.1 Glow discharge mass spectroscopy data of DRS and NIMTEC CdZnTe substrates, together with an arsenic-doped Te-rich LPE melt sample.

	CdZnTe NIM TEC	CdZnTe DRS	Te-melt R19-As		CdZnTe NIMTEC	CdZnTe DRS	Te-melt R19-As
Li	5	6	15	Br			
Be	<1	<1	<1	Rb	<0.4	<0.6	<0.9
B	10	<2	<3	Sr	<0.1	<0.1	<0.4
C	660	430	1000	Y	<0.1	<0.1	<0.3
N	45	60	130	Zr	<0.2	<0.3	<0.3
O	240	130	700	Nb	<0.2	<0.3	<0.3
F	<3	<3	<3	Mo	<0.8	<1	<0.7
Na	2	10	130	Pd			
Mg	18	21	<2	Ag	<20	<20	<3
Al	13	41	<1	Cd	Matrix	Matrix	0.30%
Si	2	4	<3	In	<30	<25	<50
P	4	4	<2	Sn	<20	<20	35
S	60	<4	<2	Sb	<20	<20	<20
Cl	5	23	4	Te	Matrix	Matrix	Matrix
K	<2	<2	12	I	<270	<270	<650
Ca	<8	<8	47	Cs	<2	<2	<2
Sc	<0.5	<0.5	<0.8	Ba	<2	<0.4	<8
Ti	<0.2	<0.3	3	La	<0.6	<0.6	<2
V	<0.2	<0.2	<0.2	Ce	<1	<1	<5
Cr	<0.8	<0.8	<1	Hf	<3	<4	<0.3
Mn	<10	<10	<0.8	Ta			
Fe	25	28	42	W	<0.9	<1	<0.5
Co	<0.3	<0.3	<0.4	Pt	<3	<5	<2
Ni	<2	<1	4	Au	<5	<5	<10
Cu	25	19	4	Ha	<2	<4	Matrix
Zn	1.80%	2.00%	63,000	TI	<0.4	<0.6	100
Ga	<1	<2	<1	Pb	<0.4	<0.8	<1
Ge	<10	<15	<15	Bi	<0.3	<0.3	<0.5
As	<10	<4	0.60%	Th	<0.6	<0.8	<0.1
Se				U	<1	<1	<1

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

an arsenic-doped Te-rich LPE melt sample. Cu, for example, is present at a level of $\sim 3-4 \times 10^{14}\,\text{cm}^{-3}$, which could result in significant contamination of a HgCdTe layer, particularly for thin layers grown on thick substrates. The disadvantage of Te-rich substrates is that the Te precipitates can act as nucleation sites for defective epitaxial growth. This effect is unimportant for thick LPE layer growth, which employs significant substrate melt-back, but is relevant for thin MBE layers. A degree of CdZnTe annealing in ingot form is employed to minimize the size of Te precipitates for MBE growth. A further issue regarding excess Te in the substrate occurs for melt growth of layers that are to be doped with fast-diffusing impurities, such as Au or Cu. Annealing such layers under Hg at low temperatures reduces the metal vacancy to the point where gettering of the impurity to the Te-rich substrate is possible. Such material annealing requires the use of substrates that are Cd annealed in slice form, so as to eliminate metal vacancies in the substrate. The price that you pay for this, however, is that any fast-diffusing impurities in the substrate are now readily available for doping the grown epitaxial layers, together with the dopant of choice. An example of the purity obtained from a Te-rich melt is given by the SIMS data of an LPE layer as shown in Fig. 5.2. Many of the impurities are at the SIMS detection limit.

5.1.1.2 MBE

Molecular beam epitaxy of HgCdTe takes place under extremely non-equilibrium conditions, using Hg, CdTe, and Te sources. The alloy composition is determined by the beam flux ratios. Typical growth occurs at ~175°C on (211)-oriented, lattice-matched CdZnTe substrates. Substrate temperature, surface preparation, and excess Te content are extremely important to the growth of MBE layers with minimal

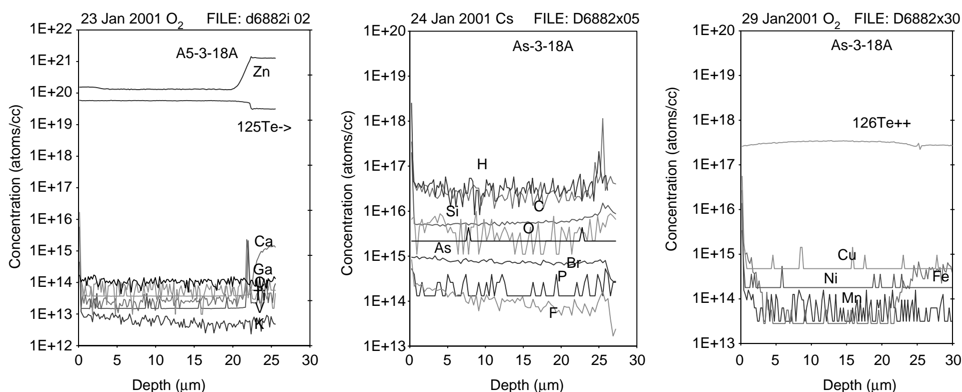


Figure 5.2 SIMS data on a Te-rich LPE layer for K, Ca, Ti, V, Cr, Zn, Ga, H, C, O, F, Si, P, As, Br, Mn, Cu, Fe, and Ni. Interestingly, the layer is $7 \times 10^{14}\,\text{cm}^{-3}$ arsenic doped, which is below the detection limit for the ions used in the SIMS analysis.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

defects such as voids and/or hillocks and dislocations. Optimal layer growth yields void densities of a few $100/cm^2$ and dislocation densities that reflect the value of the substrate, which are low 10^4 to low $10^5/cm^2$. n-Type doping is typically achieved with indium, and p-type with arsenic, although other impurities have been utilized. Arsenic must be incorporated onto the Te sublattice for acceptor activation and is thought to require a high-temperature anneal under metal-rich conditions. The electrical quality of layers, particularly p-type, can be compromised by interfacial properties, such as inversion layers at either the front or substrate interfaces. Two layer models are then needed to accurately interpret the electrical quality of the grown layer, together with any annealing sequence that is being utilized.

Growth on alternative substrates has been achieved, the most notable of which is Si/CdTe. Good electrical properties are reported; however, the dislocation content is high for HgCdTe layers grown on these substrates, typically a few $10^6/cm^2$. Such layers can be utilized for MWIR FPAs operating at 77 K, but are of limited use for LWIR, which cannot tolerate such high dislocation densities.

5.1.1.3 MOVPE

Metal–organic vapor phase epitaxy occurs via the thermal pyrolysis of metal alkyl molecules on a suitable substrate surface, at temperatures between 350 and 370°C. Alloy compositions are typically obtained by the growth of alternating layers of the appropriate thicknesses of CdTe and HgTe, and the subsequent interdiffusion of these layers. This allows for optimum growth conditions for each material and facilitates the incorporation of certain impurities, such as arsenic, which is introduced into the metal-rich CdTe layer growth phase, and is hence activated by accommodation on the Te sublattice. The best quality p-type material, as demonstrated by minority carrier lifetime measurements, is obtained utilizing tris-dimethylaminoarsene (TDMAAs), for arsenic doping, together with dimethylcadmium, di-isopropyltelluride, and elemental Hg. Iodine and indium have been used for n-type growth. As grown interdiffused layers are typically p-type due to metal vacancies, these are eliminated by an isothermal anneal in Hg at low temperatures, such as 225°C.

Successful growth has been reported on CdZnTe and GaAs/CdTe substrates, with orientations slightly off (100). Higher dislocation densities occur on the GaAs/CdTe substrates, but lower background doping concentrations are typically achieved.

5.1.2 HgCdTe annealing

The phase field of Fig. 5.1 must be considered in any subsequent annealing of material after growth and any subsequent device processing such as surface passivation. As was first pointed out by Destefanis [59], the creation and annihilation of metal vacancies in HgCdTe occurs at a finite rate. This rate can be extremely fast

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

in certain circumstances, and cannot be ignored, particularly when annealing on the Hg-rich side of the phase field. The appropriate diffusion coefficient to use is associated with Hg interstitials moving through vacancy-doped material as given by Chandra et al. [60].

$$D = f[2.25 \times 10^{19}/C_{\text{xste}}]\exp[-1.114q/kT], \tag{5.2}$$

where C_{xste} is the metal vacancy concentration, and f is 1 for $x = 0.22$, 0.12 for $x = 0.3$, and 0.03 for $x = 0.5$. Because of the inverse dependence of the diffusion coefficient on C_{xste}, the vacancy profile can be approximated by a step function, whose position in time, or space, is given by the diffusion relation $d^2 = Dt$, for any composition between $0.22 < x < 0.5$, utilizing Eqs. (5.1) and (5.2).

An example of vacancy kinetics is illustrated in Fig. 5.3(a) for LWIR material, which is annealed to equilibrium at 220°C in saturated Hg, and then the Hg pressure is allowed to drop to zero, simulating a faulty withdrawal from an annealing furnace. The vacancy profile is shown for various times at 220°C on the Te-rich side of the phase field. It is apparent that, even at these relatively low anneal temperatures, significant volumes of material can be converted to the Te-rich condition. For example, even 10 s at zero Hg pressure is sufficient to generate >4 μm of $5 \times 10^{15}\,\text{cm}^{-3}$ vacancy doped material. Thus, great care must be exercised with annealing and cool-down conditions on the Hg-rich side of the phase field. For thick LPE layers, this effect is not a major issue, as surfaces can be polished following the stoichiometry anneal. However, thin layers grown by MBE are more at risk. For this reason, it can be beneficial to grow large bandgap cap layers on the active layers, as illustrated in Fig. 5.3(b). The diffusion coefficient of $x = 0.6$ HgCdTe is sufficiently slow that it takes 14 s to deplete just a 1-μm cap layer, and thereafter the cap will still remain a diffusion bottleneck for Hg.

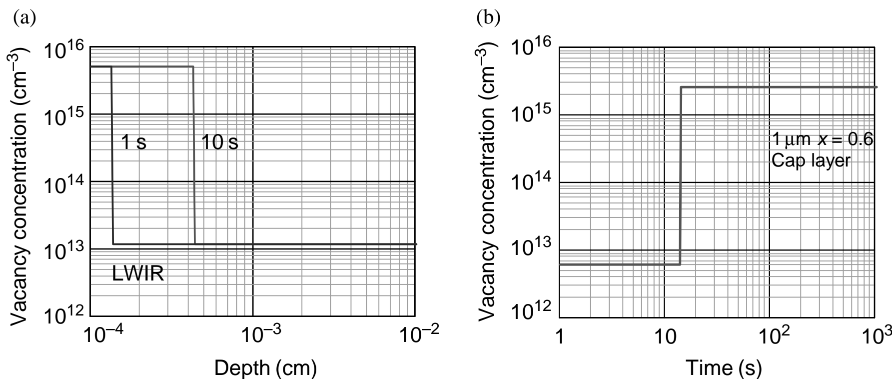


Figure 5.3 (a) Vacancy front after 1 s, 10 s, of 220°C anneal on the Te-rich side, after equilibrium anneal under Hg and (b) time taken to convert 1 μm of $x = 0.6$ HgCdTe cap layer from Hg- to Te-rich condition at 220°C.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

5.1.3 HgCdTe properties

5.1.3.1 Band structure

For $Hg_{1-x}Cd_xTe$ compositions of importance in the infrared, the band structure is direct gap at $k = 0$. The conduction band is spherical but non-parabolic; the band edge electron effective mass depends directly on the bandgap and is given by [61] $(m_c^*/m_o) \approx 7 \times 10^{-2} E_g$, where E_g is in eV. The valence band is degenerate heavy and light hole, with a split off spin–orbit band. The light hole mass is predicted to be $\sim m_c^*$, and the heavy hole is typically assumed $\sim 0.55 m_o$, and may be mildly dependent on composition across the complete range of x values.

The bandgap of $Hg_{1-x}Cd_xTe$ is a function of temperature and composition and is given by [62]

$$E_g = -0.302 + 1.93x + (5.35 \times 10^{-4})T(1 - 2x) - 0.81x^2 + 0.832x^3. \quad (5.3)$$

The intrinsic carrier concentration as a function of composition, x, and temperature, T, is given by [5]

$$n_i = [5.585 - 3.82x + (1.753 \times 10^{-3})T - (1.3641 \times 10^{-3})Tx] \times 10^{14} E_g^{0.75} T^{1.5} \exp(-E_g q/2kT), \quad (5.4)$$

where E_g is in eV and given by Eq. (5.3).

5.1.3.2 Optical properties

The low- and high-frequency dielectric constants for $x = 0.2$ HgCdTe are $\varepsilon = 19.5$, and $\varepsilon_{hf} = 14$, measured at 4.2 K. These values can be compared to CdTe with $\varepsilon \approx 10$, and $\varepsilon_{hf} \approx 7$. Intermediate x values are typically obtained by linear interpolation.

The absorption coefficient at the band edge, as a function of photon energy, E, is given by [7]

$$\alpha = 4 \times 10^4 [E(E - E_g)]^{1/2} \text{ cm}^{-1}, \quad (5.5)$$

where energy is in eV.

5.1.3.3 Transport properties

Care must be exercised in the interpretation of Hall data of HgCdTe due to the presence of mixed, and/or multilayer, conduction. Various techniques are available to enable this interpretation and typically involve the use of variable magnetic field measurements. The most quantitative technique is the quantitative mobility spectrum analysis (QMSA) [63], which is a fully automated technique for determining free-carrier densities and mobilities from magnetic field dependent Hall and resistivity data. However, for unambiguous interpretation, this technique requires μB

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

products to be $>10^8\,\text{cm}^2\text{G/Vs}$, which is difficult for the heavy hole in HgCdTe, even at the highest magnetic field values. An alternative, less complicated analysis is the two-layer model of Wong [64], which can utilize much smaller magnetic fields and is ideal for p-type HgCdTe with possible n-type interfaces.

The standard expression for the Hall coefficient in a single-layer semiconductor operating in the mixed conduction regime, measured in a magnetic field B, is given by [65]

$$R = [D_\text{n} + D_\text{p}]/B[(A_\text{n} + A_\text{p})^2 + (D_\text{n} + D_\text{p})^2],$$

where

$$\begin{aligned} A_\text{n} &= \sigma_\text{n}/(1 + \omega_\text{n}^2\tau^2), \\ D_\text{n} &= \sigma_\text{n}^2 R_\text{n}/[1 + \omega_\text{n}^2\tau^2], \end{aligned} \tag{5.6}$$

σ_n is the conductivity due to electrons $= nq\mu_\text{n}$, μ_n is the mobility $= q\tau/m^*$, τ is the lifetime for momentum scattering, and m^* is the effective mass of the electron. $R_\text{n} = 1/nq$, and $\omega_\text{n} = qB/m^*$. Similar expressions apply to hole carriers.

The Hall coefficient and conductivity in a two-layer system, composed of a high-mobility n-type layer and a low mobility p-layer, are given by

$$R = [X + Y]/[B^2(X + Y)^2 + Z^2]$$

and

$$\sigma = [t_1(\sigma_{\text{n}1} + \sigma_{\text{p}1})/(t_1 + t_2)] + [t_2(\sigma_{\text{n}2} + \sigma_{\text{p}2})/(t_1 + t_2)],$$

where

$$\begin{aligned} X &= t_1[D_{\text{n}1} + D_{\text{p}1}]/(t_1 + t_2), \\ Y &= t_2[D_{\text{n}2} + D_{\text{p}2}]/(t_1 + t_2), \\ Z &= [t_1(A_{\text{n}1} + A_{\text{p}1})/(t_1 + t_2)] + [t_2(A_{\text{n}2} + A_{\text{p}2})/(t_1 + t_2)], \end{aligned} \tag{5.7}$$

and t_1 and t_2 represent the layer thicknesses. For HgCdTe, in which $\mu_\text{n} \gg \mu_\text{p}$, the measured carrier concentration and mobility depend primarily on the n-type interface charge per unit area, the bulk layer doping concentration, and the mobility assumed for the interface electrons. Bulk layer mobilities as a function of composition are well known. Hall measurements at two values of magnetic field suffice to define the bulk layer doping concentration, interface charge density, and interface electron mobility.

An idea of the degree of difficulty in Hall data interpretation of p-HgCdTe is shown in Fig. 5.4, where a relatively low mobility ($1.5 \times 10^4\,\text{cm}^2/\text{Vs}$) n-skin layer on p-HgCdTe is modeled at 77 K. Figure 5.4(a) models the measured mobility, and Fig. 5.4(b) the measured carrier concentration, as a function of surface charge density in the n-layer. It is apparent that for surface charge densities that are typical of a native oxide ($\sim 10^{12}\,\text{cm}^{-2}$), even a $10^{16}\,\text{cm}^{-3}$ p-type sample will appear to be n-type with a carrier concentration of $\sim 10^{15}\,\text{cm}^{-3}$ and a mobility in excess of $10^4\,\text{cm}^2/\text{Vs}$. The use of Wong's model at two magnetic fields is sufficient to

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

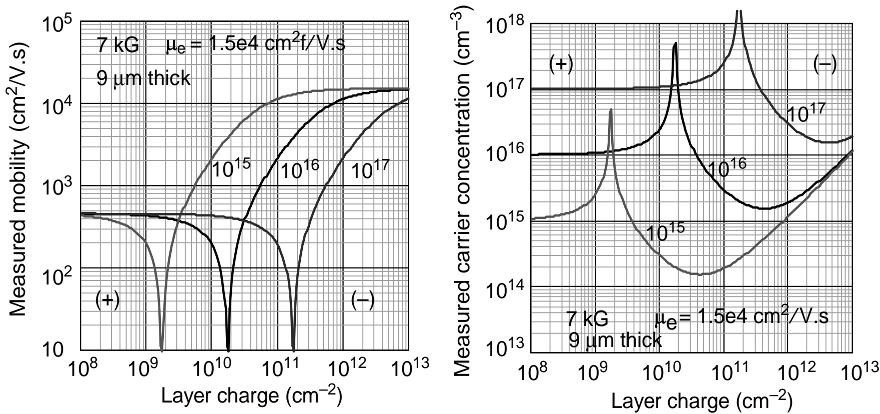


Figure 5.4 (a) Measured mobility and (b) measured carrier concentration vs. n-type surface charge for p-HgCdTe with doping levels of 10^{15}, 10^{16}, and 10^{17} cm^{-3}. The electron mobility is 1.5×10^4 cm^2/Vs, the magnetic field 7 kG, and thickness is 9 μm.

unambiguously identify the two carriers involved, with a minimum of assumptions. The predominant form of n-skin measured on p-HgCdTe is found to occur at the growth interface with the CdZnTe substrate, and it is characterized by low mobilities typical of large-x HgCdTe with charge densities $\sim 10^{11}$ cm^{-2}, which can render the characterization of low doped p-HgCdTe challenging.

(a) Electron transport

The small electron effective mass results in relatively large values of electron mobility, as shown by the modeled data of mobility vs. temperature in Fig. 5.5(a)

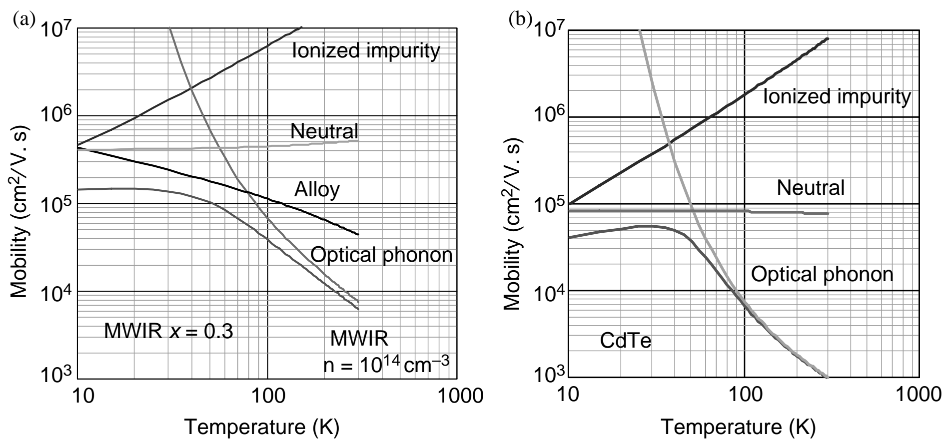


Figure 5.5 Electron mobility as a function of temperature, for (a) $x = 0.3$ HgCdTe and (b) CdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

for MWIR $x = 0.3$ HgCdTe, with $n_d = 10^{14}\,\mathrm{cm}^{-3}$. Measured values are in good agreement with the modeled data, which includes scattering by optical phonons [66], alloy disorder [67], ionized impurities [68], and neutral impurities [69]. The main components for pure n-type HgCdTe are optical phonons and alloy disorder.

Stratton's theory of optical phonon scattering for $T < \theta$, and energy $< k\theta$, gives

$$\mu_{\mathrm{opt}} = (k\theta/2m^*)^{1/2}[\exp(\gamma_o) - 1]/F_o, \tag{5.8}$$

where $k\theta$ represents the optical phonon energy, $\gamma_o = \theta/T_o$, T_o is the lattice temperature, and $F_o = [\varepsilon_{\mathrm{hf}}^{-1} - \varepsilon^{-1}]m^* qk\theta\pi/\varepsilon_o h^2$.

Alloy scattering at low energy, for a composition x, is given by

$$\mu_{\mathrm{alloy}} = qh/[4\pi^2 x(1-x)\Delta E_c^2 \rho(E) m^*], \tag{5.9}$$

where ΔE_c is the difference between the s atomic term values of the Hg and Cd atoms, and $\rho(E)$ is the density of states in the conduction band. The model in Fig. 5.5(a) utilized a value of $\Delta E_c = 1.5\,\mathrm{eV}$, in reasonable agreement with the conduction band difference between HgTe and CdTe.

Ionized impurity scattering is more relevant in highly doped material and is given by

$$\mu_I = [3.2 \times 10^{15}\varepsilon^2 T^{3/2}/N(m^*/m_o)^{1/2}]/[\ln(3.7 \times 10^{14} T^2 \varepsilon(m^*/m_o))/n - 1], \tag{5.10}$$

where N is the density of singly ionized impurities, and n the density of free carriers.

Neutral impurity scattering for low-energy electrons is given by

$$\mu_n = 1.44 \times 10^{22}(m^*/m_o)/\varepsilon N_n, \tag{5.11}$$

where N_n is the density of neutral centers.

As a point of interest, Fig. 5.5(b) shows the model for electron mobility in CdTe, for which alloy scattering is zero. The data are in excellent agreement with experiment [70], confirming both Stratton's model for optical phonon scattering in II-VI materials and the scattering due to alloy disorder of low-energy electrons in HgCdTe.

Donor impurities that are utilized for n-type doping in HgCdTe are typically Group III or Group VII elements substituting on the metal or Te sublattices, respectively. They tend to be very slow diffusers at typical processing temperatures. In high-quality Te-rich LPE, indium is typically used, substituting on the metal sublattice, for a background n-type dopant. Levels of $10^{14}\,\mathrm{cm}^{-3}$ are found to give classical Hall data and are consistent with SIMS analyses. The small electron effective masses across most of the composition range result in zero activation energies for these donors. Donor activation has been reported for CdTe, in agreement with a simple hydrogenic center at 0.014 eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

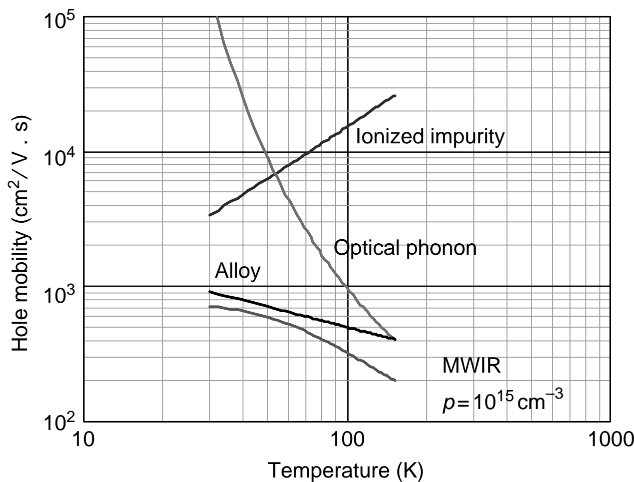


Figure 5.6 Hole mobility vs. temperature for $x = 0.3$ HgCdTe with $p = 10^{15}$ cm^{-3}.

(b) Hole transport

The transport of holes in HgCdTe is dominated by the heavy hole with an effective mass of $\sim 0.55 m_o$. Model data for $x = 0.3$ HgCdTe is shown in Fig. 5.6 for a doping concentration of 10^{15} cm^{-3}. These data are in good agreement with experimental data on both vacancy and arsenic-doped MWIR material. $\Delta E_v = 0.3$ eV is required to fit the alloy scattering component, in reasonable agreement with the valence band offset between HgTe and CdTe.

Activation energies are readily observed for acceptors in HgCdTe due to the large value of the heavy hole effective mass. The most common p-type behavior is associated with the vacancy on the metal sublattice. The concentration of vacancies is determined by the appropriate anneal temperature and Hg pressure, as indicated in the phase field diagram in Fig. 5.1. Theoretically, the vacancy should be doubly charged, and measurements of carrier concentration vs. temperature suggest this to be the case, as shown in Fig. 5.7(a) for LWIR HgCdTe. The model data are for a double acceptor with acceptor levels of 0.012 eV and 0.045 eV, and agree well with the indicated experimental data.

Group IB elements, namely gold, silver, and copper, have been introduced onto the metal sublattice with limited success. Group IB elements have a relatively high diffusivity but can be used as dopants for diode formation, provided appropriate steps are taken in the design and formation of the diode. Group IA elements such as Na, K, and Li are known acceptors exhibiting high diffusivities.

Group V elements, if incorporated on the Te sublattice, are acceptors. Depending on the method of introduction, uniform doping with Group V elements may require both diffusion and activation anneals. Arsenic is the most common Group V dopant in use today. It behaves as a monovalent acceptor, as shown in Fig. 5.7(b), giving an activation energy of ~ 0.011 eV. A similar activation energy is found for Au-doped MWIR HgCdTe. The simple hydrogenic acceptor for an effective

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

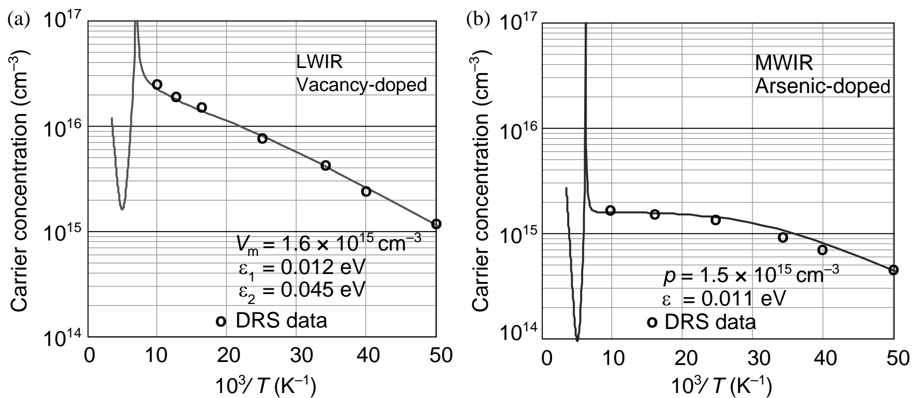


Figure 5.7 Carrier concentration vs. $10^3/T$ for (a) LWIR vacancy-doped and (b) MWIR arsenic-doped HgCdTe.

mass of $0.55m_o$ would exhibit an activation energy of $\sim$0.021 eV for MWIR HgCdTe.

Attempts have been made to investigate the light hole in p-type HgCdTe by measuring mobility and carrier concentration as a function of magnetic field at low temperatures. Figure 5.8 shows such measurements for the vacancy-doped LWIR sample illustrated in Fig. 5.7, at a temperature of 20 K over the magnetic field range 1–10 kG. The data can only be fitted to the model by assuming a small splitting of the valence bands $\sim$0.002 eV at $k = 0$, together with a finite increase of the effective mass to a value equal to approximately twice that of the electron. Such effects can be expected as a result of strain, which is possible for thin films that are bonded to sacrificial substrates. The required value of light hole mobility to provide the data fit shown in Fig. 5.7 is $\sim 2.5 \times 10^4$ cm^2/Vs. This value is probably consistent with neutral impurity scattering of the light hole in this sample, due to the freeze-out

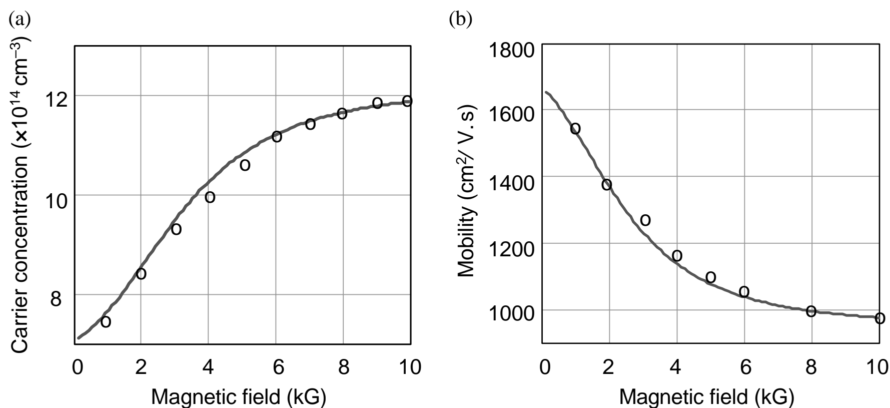


Figure 5.8 Magnetic field dependence of (a) carrier concentration and (b) mobility, for vacancy-doped LWIR HgCdTe at 20 K, compared to a model for light holes with mass $\sim 2\times$ electron mass, and a valence band splitting of $\sim$0.002 eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

of doubly ionized vacancies at these low temperatures. Other workers have found evidence of low-mobility light holes in similar circumstances. It remains to be seen whether the valence band distortion needed to explain this Hall data will impact other properties of HgCdTe, such as Auger7 recombination.

5.1.3.4 Minority carrier lifetime properties

There are three minority carrier lifetime mechanisms of potential importance in HgCdTe

- Auger,
- Radiative, and
- Shockley–Read.

The first two are band-to-band recombination mechanisms and are determined by the band structure of the material, and to a large extent are unavoidable, except in grossly non-equilibrium circumstances. The third is associated with defects, or impurities, located in the bandgap of the semiconductor material and is theoretically avoidable by the elimination of the defects. These mechanisms have been thoroughly discussed in Sec. 4.4. The current state of the art is shown in Fig. 5.9 for both n- and p-type MWIR and LWIR HgCdTe. Internal radiative generation–recombination is considered a non-issue in a well-designed photodiode.

The Auger7 lifetime in p-type material in Fig. 5.9 is arbitrarily chosen as equal to $12\times$ Auger1, as this value seems to agree with the experimental data at the present time. Auger1 has been observed at doping concentrations $>6\times10^{14}$ cm^{-3}; but at lower concentrations, lifetime values in n-type material appear to be limited by S-R centers of unspecified origin and vary somewhat. At doping concentrations $>10^{16}$ cm^{-3}, Auger1 will be suppressed in n-type material by band-filling effects,

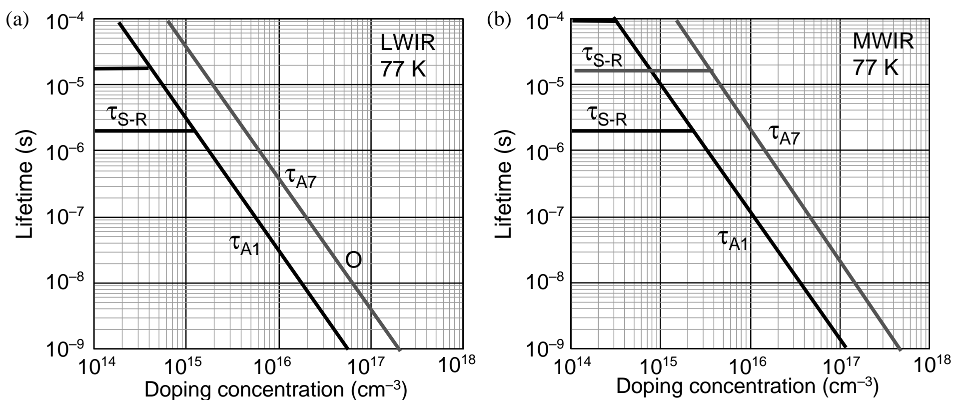


Figure 5.9 State-of-the-art lifetimes at 77 K for n- and p-type (a) MWIR and (b) LWIR HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

due to the small effective mass associated with the conduction band. No band filling will take place in the valence band due to the large effective mass of the heavy hole.

5.1.3.5 Physical properties of HgCdTe

Physical properties of importance in HgCdTe growth and processing are lattice constant, coefficient of thermal expansion (CTE), and dislocation density.

Lattice constant is important in the growth of epitaxial layers, in that dislocation-free growth can only be achieved on lattice-matched substrates. The lattice constant of HgTe is 6.46 Å, and CdTe 6.48 Å, with compositions in between obtained by linear interpolation. Growth of a dislocation-free epitaxial layer on a lattice-mismatched substrate can occur up to a thickness given by the Matthews criterion, namely

$$h_c = [(1 - \nu/4)b]/[2.31\pi f(1 + \nu)] \ln(ah_c/b), \tag{5.12}$$

where ν is Poisson's ratio (=0.33), b the Burgers vector of a misfit dislocation (=4.583 Å), a the dislocation core cutoff parameter assumed equal to a value between 1 and 4, and f the misfit strain. Once this critical thickness is exceeded, the accommodation of strain by misfit dislocations is energetically favored. The residual strain in the layer for any thickness is given by the strain associated with that thickness as given by the Matthews criterion.

Lattice matching is a necessary but not sufficient condition for misfit dislocation-free interfaces. CTE also plays a role. Perfect lattice matching at the growth temperature is not retained during cool-down due to differences in CTE of substrate and epilayer. The coefficient of thermal expansion of HgCdTe and CdTe, as a function of temperature, has been documented by Collins et al. and Smith [71, 72]. The data indicate that LPE growth on lattice-matched CdZnTe at 475°C should lead to the formation of misfit bands at the growth interface on cooling down to room temperature.

CTE differences also play a vital role in the successful hybridization of HgCdTe layers to silicon ROICs. The difference in CTE of HgCdTe and Si is well-documented and must be accommodated in the hybridization process. The concept is qualitatively illustrated in Fig. 5.10, where the strain introduced in HgCdTe due to thin-film epoxy-hybridization to Si is plotted as a function of temperature, assuming an epoxy cure temperature of T_c. The HgCdTe is effectively lattice-matched to the Si ROIC at the cure temperature, and strain is progressively introduced as the temperature is raised or lowered from that temperature. Equation (5.5) gives the maximum strain that can be accommodated by the Matthews' criterion for any thickness of HgCdTe. Thus, for a 10-μm thick layer epoxied to a Si ROIC, the critical strain is $\sim 5 \times 10^{-5}$. At temperatures where this strain is exceeded, misfit dislocations can be formed. This has been observed at DRS on the high-temperature side of the cure temperature, but not on the low side, indicating that perhaps the yield stress at temperatures approaching room temperature is high enough to inhibit misfit formation. This is not the case at higher temperatures. This raises an interesting issue regarding possible MBE growth of HgCdTe on Si/CdTe, or if possible, Si/CdZnTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

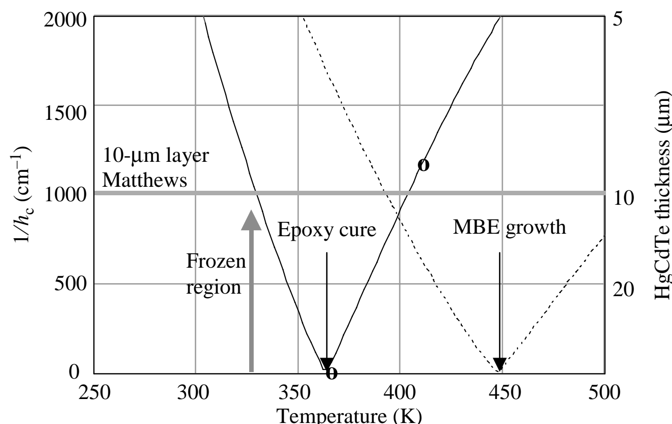


Figure 5.10 The relevance of CTE in HgCdTe/Si hybridization and HgCdTe MBE Growth on Si.

Growth at 175°C, even for a perfectly matched CdZnTe buffer layer on Si would be confronted with the problem of cool-down to room temperature. The CTE mismatch will generate the strain vs. temperature curves included in Fig. 5.10, and it indicates that for a 10-μm thick layer the Matthews criterion for tensional strain would be exceeded at ~125°C, at which temperature the yield stress would not be particularly high. There would thus be a tendency for misfits to form, even for this seemingly idealized growth situation. Providing no subsequent high-temperature processing was required, this issue could be improved by initiating growth on a lattice-mismatched CdZnTe buffer layer that induced compressional strain at the layer growth temperature, but which would be relieved on cooling down, before transitioning to a region of tensional strain at a temperature that is sufficiently low that the HgCdTe yield stress was high enough to inhibit misfit formation.

Current state of the art for HgCdTe MBE growth on Si/CdTe yields dislocation densities well in excess of $10^6/\text{cm}^2$. Such values cannot be tolerated by LWIR diodes operating at 77 K, which require dislocation densities $<2 \times 10^5/\text{cm}^2$ for low-defect FPAs. MWIR diodes, on the other hand, are more tolerant to dislocations at 77 K and provide good yield on HgCdTe grown on Si. Our experience at DRS with MWIR HgCdTe has been that FPAs that appear blemish-free at 77 K tend to highlight any mechanically induced defects at higher operating temperatures, suggesting that these higher temperatures may be more demanding of the dislocation density requirement.

5.2 HgCdTe Device Architectures

First-generation systems consisted entirely of HgCdTe n-type photoconductive elements in various scanning configurations. Signal processing and bias circuitry was located off the focal plane. Current-generation IR systems rely on large-area scanning and staring HgCdTe photodiode arrays hybridized to Si ROICs on the focal

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

plane itself. These device architectures fall primarily into two categories of diode type and two categories of hybridization.

Diode types are grown double-layer p^+/n hetero-junctions (DLHJs) and $n^+/n^-/p$ homo-junctions, where the diodes are formed either by ion implantation or ion etching of p-HgCdTe. In the latter case, the diode is formed by diffusing Hg interstitials, generated by the damage associated with the implant or etch process.

Hybridization of the DLHJ to the ROIC is achieved by the use of indium bumps. The ion-implanted homo-junction can be connected to the ROIC by indium bumps or by etching vias through the HgCdTe down to the ROIC. The vias can be an active part of the diode structure if so desired.

The four common competing HgCdTe device architectures are shown in Fig. 5.11. The DLHJ and bump-bonded implant architectures will be discussed only briefly here. The major discussion is reserved for the VIP and HDVIP devices. This is not necessarily indicative of the relevant importance of the various architectures, but it merely reflects the author's lack of intimate knowledge of bump-bond technologies.

5.2.1 DLHJ architecture

Figure 5.11(a) shows one particular format of the bump-bonded DLHJ, which utilizes mesa etched diode formation prior to surface passivation. The n-HgCdTe is typically doped with indium to the level of 5 to 20×10^{14} cm^{-3}, depending on the growth technique used, and the grown p^+ layer is of a wider bandgap material, heavily doped with arsenic to the level of 5×10^{17} to 10^{18} cm^{-3}. Another variation

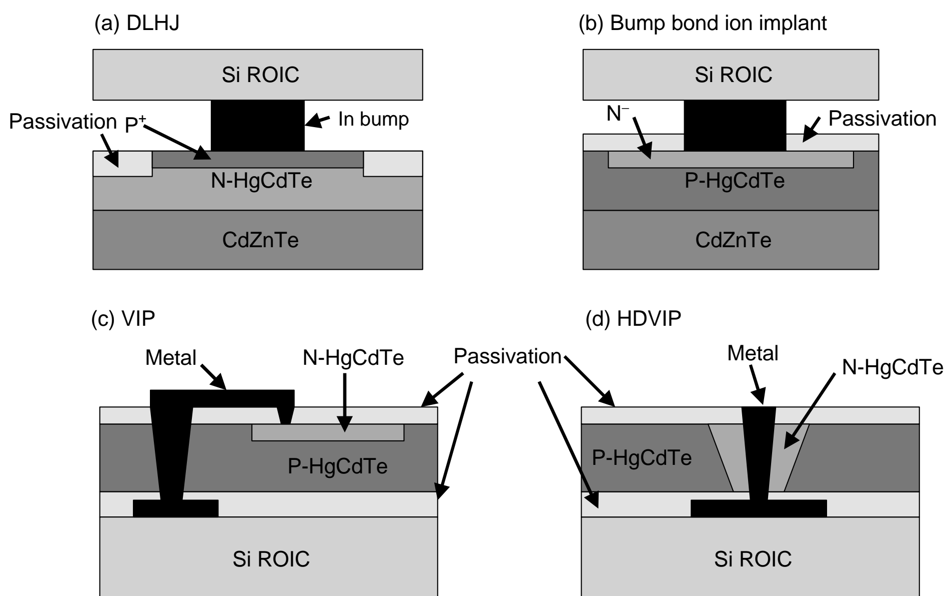


Figure 5.11 Four present-day HgCdTe device architectures.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

of the DLHJ avoids the mesa diode formation and utilizes implantation and diffusion of arsenic into a grown planar DLHJ to define the diode region. In either case, the p/n junction must be positioned in the narrow gap region of the diode to avoid setting up possible barriers to minority carrier flow, which result in reduced quantum efficiency. This can be achieved by interdiffusion of the arsenic from the wide gap into the narrow gap region of the structure.

The attraction of the DLHJ is that essentially no dark current is contributed by the p-region of the device due to its wide bandgap and hence lower value of n_i. The n-region is theoretically limited by the Auger1 diffusion current.

Due to the nature of the hybridization process, the FPA is illuminated from the backside, through the substrate, which is typically CdZnTe.

The major issue with bump-bond hybridization is the difference in thermal expansion coefficient between the detector growth substrate material and the Si ROIC. FPAs above certain critical sizes must be processed so as to minimize the stress associated with this CTE mismatch. This typically involves thinning of the Si ROIC and/or the detector substrate, or the use of intermediate substrates in the packaging process. For this reason, a great deal of effort has been expended on the development of growth of HgCdTe on Si. As discussed earlier, the resulting FPAs appear to be appropriate for MWIR operation at 77 K but are far from meeting the necessary criteria for successful LWIR operation, at any temperature. As materials technology improves, operating temperatures will increase towards room temperature, and the issue of CTE mismatch of the bump-bonded FPA with the Si ROIC will disappear.

5.2.2 Bump-bonded ion implant architecture

The diode in this architecture is planar and formed by the ion implantation of any element into CdTe pre-passivated, vacancy-doped HgCdTe containing a low background donor level of, say, indium. The damage generated by the implant process is heavily n^+ type and liberates Hg interstitials, which fill the metal vacancies to a finite depth, thus converting the region to n^- due to the residual donor background. This conversion process is portrayed in Fig. 5.12, for a typical planar B implantation $\sim 2 \times 10^{15}/\text{cm}^2$.

The homo-junction diode structure is indium bump-bonded to the underlying Si ROIC. Diffusion currents are dominated by thermal generation in the p-volume at levels typically higher than those in a DLHJ of the same cutoff wavelength. This architecture combines the undesirable attributes of high dark current from the vacancy-doped material, together with the CTE mismatch associated with HgCdTe bump-bonded to Si.

5.2.3 Vertically integrated photodiode (VIP and HDVIP) architectures

The vertically integrated photodiode shown in Fig. 5.13 consists of a thin layer of extrinsically doped p-HgCdTe, passivated on both surfaces with an interdiffused

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

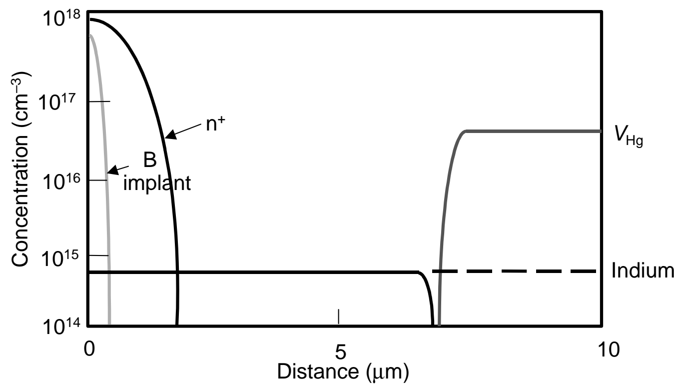


Figure 5.12 $n^+/n^-/p$ diode formation by boron implantation of vacancy-doped HgCdTe.

layer of CdTe, which is glued to an underlying Si ROIC with a uniformly thin bond-line of epoxy. Optimizing of the epoxy cure step, in the manner described in Sec. 5.1.3.5, ensures that the stress in the HgCdTe is uniform throughout the whole layer except for a small region around the periphery that is approximately the size of a layer thickness, thus enabling the fabrication of FPAs of essentially unlimited size.

Contact to the Si ROIC is achieved by etching vias through the HgCdTe down to landing pads on the silicon. In the HDVIP™ architecture, the n^+/n^- region of the diode is formed around the via, both by the etching process itself and a subsequent ion implant step. The other notable feature of the HDVIP™ architecture is the p^+/p non-injecting contact formed in each unit cell of the FPA. These contacts are joined electrically by a top surface metal grid as shown in Fig. 5.13, which results in negligible de-biasing across a large-area array, even at relatively high background

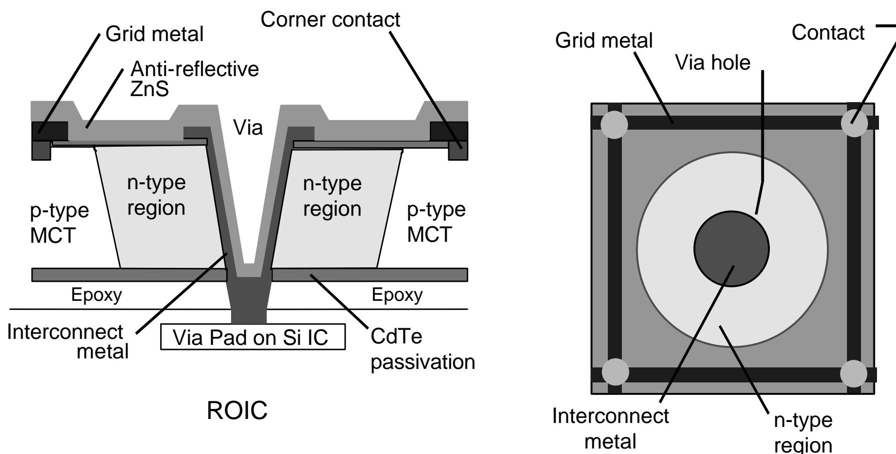


Figure 5.13 DRS HDVIP™ $n^+/n^-/p$ structure.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

flux currents. The fill factor of the HDVIP™ unit cell, associated with the metal grid and via hole, depends on pitch size but is typically $>80\%$, with resulting unit cell collection efficiencies in excess of 60%.

The materials and processing technology utilized in HDVIP detectors have a number of features that significantly impact diode performance, producibility, and reproducibility, including the following:

- Liquid phase epitaxial (LPE) material growth from a large-volume Te-rich melt. This results in inherently purer HgCdTe epitaxial layers, due to the smaller segregation coefficients of most impurities, as compared to a Hg-rich melt.
- Low background indium doping levels are readily achievable in Te-rich melt grown LPE layers. Indium concentrations of 1.5 to 5×10^{14} cm^{-3} are routinely used. These result in low built-in electric fields and capacitance values for HDVIP diodes.
- The diode formation process employed in HDVIP utilizes Hg interstitials to form the n^+/n^- region and is illustrated in Fig. 5.14. During this process, Hg vacancies are annihilated, and fast-diffusing impurities (such as Group IB elements Cu, Ag, and Au, and Group IA elements Na, K, and Li, which reside on the metal sublattice) are swept out of the n^- region of the diode by the flux of Hg interstitials. Due to the relatively high p-doping concentrations employed in HDVIP, the depletion region of the diode is formed

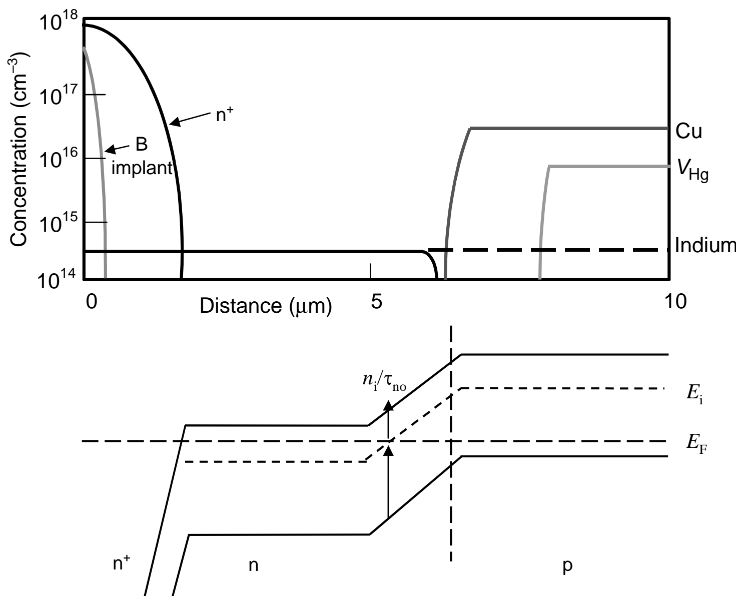


Figure 5.14 Typical VIP diode formation in p-HgCdTe doped with Cu and metal vacancies, and background doped with indium and the associated band diagram.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

almost entirely in the n-region, which by definition is devoid of fast-diffusing impurities.

- The HDVIP diode is immune to fast-diffusing impurities from the substrate, such as Cu, introduced during LPE layer growth, and subsequent high-temperature annealing, due to the purging action of the diode depletion region by the diode formation process.
- The HDVIP diode geometry is essentially cylindrical, with its depletion region formed perpendicular to the plane of the LPE layer, that is, parallel to the LPE layer growth direction. This geometry is particularly beneficial for minimizing effects due to threading dislocations from the substrate. Dislocations have been found to be a significant source of defects in FPAs, acting as macroscopic sources of tunnel current and $1/f$ noise, but are thought to be only a problem if they intersect the depletion region of the diode. In the HDVIP geometry, the depletion region presents an extremely small cross section to the threading dislocations, as indicated in Fig. 5.15, relative to planar diode geometries, such as the vertically integrated photodiode (VIP^{TM}) and the double-layer heterojunction (DLHJ). This results in an obvious advantage with regard to potential array defects. This advantage has been demonstrated by the fabrication of 480×5 scanning FPAs in both the VIP and HDVIP formats. Lower defect densities are observed with HDVIP.
- The surface passivation consists of interdiffused wide-bandgap CdTe layers formed prior to the diode process. The band bending associated with the interdiffused nature of these regions screens out fixed surface charge, and it also keeps minority carriers away from the surface region.
- The n^+/n^- interface acts as a very low recombination velocity surface due essentially to the conduction band filling associated with the n^+ region, suppressing Auger1 recombination in this volume of the diode.
- The photodiode is front-side illuminated, leading to near theoretical MTF performance of the HDVIP FPA.
- The hybridization technique to Si renders the device completely immune to rapid thermal cycling, as discussed in Sec. 5.1.3.5.

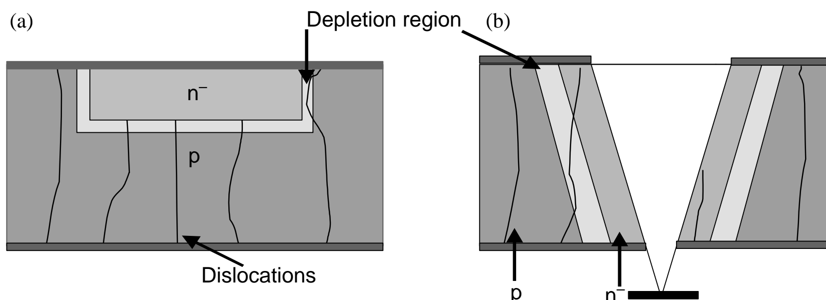


Figure 5.15 Impact of threading dislocations on (a) planar VIP and (b) HDVIP.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

It should be pointed out here that the HDVIP architecture possesses many of the same attributes as the so-called loophole technology developed in the UK and described in an excellent article by Baker and Maxey [73].

5.3 ROIC Requirements

It is not the purpose of this section to present a comprehensive description of the input requirements of ROICs. This has been accomplished in many other treatises. It is merely to point out that the input stage of the ROIC typically integrates photon-generated charge on an integrating capacitor via a MOSFET circuit. The type of circuit used depends to a large extent on the available detector impedance and the real estate available in the Si unit cell. For advanced FPAs, storage on the integration capacitor can become an issue for high flux levels and/or long integration times. These larger well capacities are typically achieved by utilizing a switched capacitor readout scheme (sometimes referred to as pixel accumulation (PA), or subframe averaging), as opposed to direct integration on the readout node. In this mode of operation, the output voltage associated with the charge integrated on the detector node C_{int}, after a finite integration time, which can be much less than a frame time, is switched onto an intermediate capacitor node C_{b} by closing switch 1, as shown in Fig. 5.16. The intermediate node is then floated by opening switch 1, and the detector is reset and then integrates again, and its output is again shared with C_{b}. This operation is repeated a number of times, R, in the allowed frame time. At the end of the frame time the voltage on C_{b} is read out by closing switch 2. The signal to noise ratio (SNR) on C_{b} follows from standard recursive filter theory.

The signal voltage is

$$V_{\text{s}} = Q_{\text{s}}(1 - \alpha^{\text{R}})/C_{\text{int}}. \tag{5.13}$$

The noise voltage is given by

$$V_{\text{n}} = Q_{\text{n}}[1 - \alpha^{2\text{R}}]^{1/2}[(1 - \alpha)/(1 + \alpha)]^{1/2}/C_{\text{int}}, \tag{5.14}$$

where

$$\alpha = C_{\text{b}}/(C_{\text{b}} + C_{\text{int}}). \tag{5.15}$$

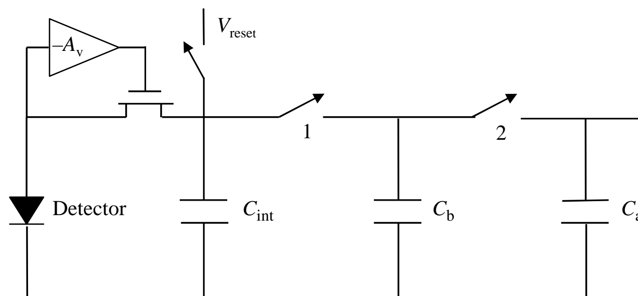


Figure 5.16 Switched capacitor filter readout of staring FPAs.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

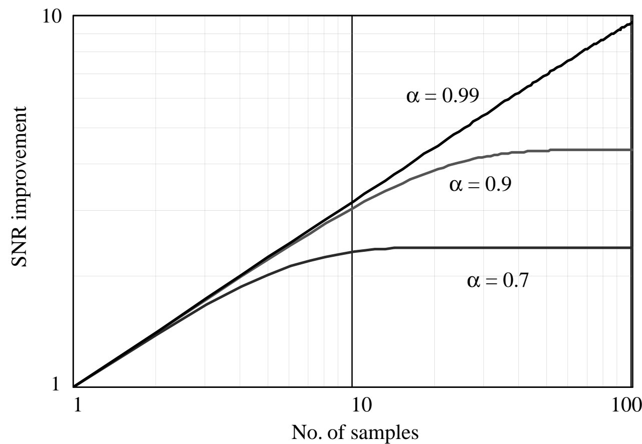


Figure 5.17 SNR enhancement from pixel accumulation.

The resulting enhancement in SNR vs. the number of samples, or accumulates, is shown in Fig. 5.17 as a function of capacitance ratio, α. The SNR is seen to vary as $R^{1/2}$, for α approaching unity. This enhancement effect is identical in magnitude to a time-delay-integrate (TDI) operation, with the exception that the observed signal voltage, at best, is equal to the signal voltage on C_{int} from one sample. The improvement in SNR is achieved by averaging down the observed noise on node C_{b}. Thus, care must be exercised in the input design to ensure not only that α approaches unity, but that the noise on C_{b} remains larger than any other noise source in the ROIC. Using these guidelines, it becomes possible to approach full frame integration with high photon collection efficiency in LWIR staring FPAs.

5.3.1 Detector performance: Modeling

The ultimate performance of an IR detector is determined by dark current and excess $1/f$ noise. Many treatises on photodiodes tend to use $R_{\text{o}}A$, the resistance–area product at zero bias, as a criterion for comparing diode performance. For diffusion-limited operation, in which $R_{\text{o}}A = kT/qJ_{\text{dif}}$, the criteria are equivalent; but, for other sources of current, the comparison becomes unnecessarily complicated, and the dark current criterion is much more meaningful. The only real requirement on diode impedance is that it be large enough that the detector noise dominates that from the ROIC.

5.3.2 Dark current in HgCdTe diodes

Possible sources of dark current in a reverse-biased HgCdTe diode are

- Thermally generated diffusion current from both sides of the junction;
- Thermal generation through bandgap states in the junction depletion region;

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

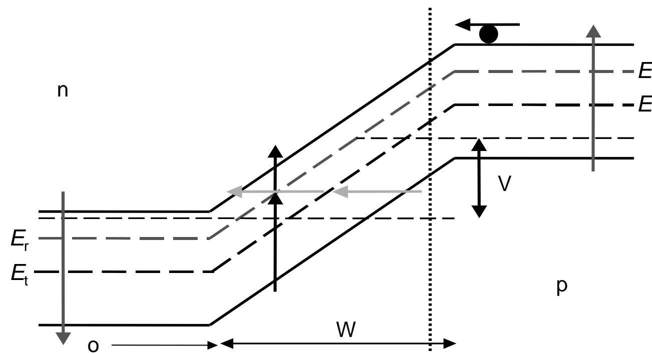


Figure 5.18 Dark current sources in a reverse-biased HgCdTe homojunction diode.

- Generation from the diode surfaces;
- Tunneling in the depletion region via bandgap states;
- Direct tunneling between the valence and conduction bands in the depletion region.

These dark current mechanisms, relevant to the band-structure of a HgCdTe homo-junction diode, are illustrated in Fig. 5.18. Because of the large difference in conduction band and heavy hole valence band effective masses, the intrinsic energy level is considerably above mid-gap in HgCdTe. This energy level is important because it represents the optimum stepping-stone for thermally induced transitions via bandgap states, E_r, in HgCdTe. On the other hand, tunnel transitions maximize between electron and light hole states, with approximately equal masses, and the optimum stepping-stone for tunneling is via states at approximately mid-gap, E_t.

5.3.2.1 Diffusion current

Thermally generated diffusion current from both the n- and p-regions is given by

$$J_{\text{dif}} = qn_i^2 t[(1/n\tau_n) + (1/p\tau_p)],$$

where

$$\tau_{n,p} = [[(1/\tau_A) + (1/\tau_{S-R})]_{n,p}]^{-1}, \tag{5.16}$$

t is the thickness of the diode, typically $\sim 1/\alpha$, and τ_A and $\tau_{S\text{-}R}$ are the Auger and S-R lifetimes, respectively, discussed in Sec. 5.1.3.4. The dark current as a function of applied voltage is given by

$$J_{\text{dif}} = qn_i^2 t[(1/n\tau_n) + (1/p\tau_p)]\{\exp(qV/kT) - 1]. \tag{5.17}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

5.3.2.2 Depletion current

Thermal generation in the depletion region through an S-R center at E_r is given by

$$J_{dep} = qn_i^2 W/[\tau_{no} p_r + \tau_{po} n_r], \quad (5.18)$$

where $\tau_{no} = 1/\gamma_n N_r$, $\tau_{po} = 1/\gamma_p N_r$, $n_r = N_c \exp(-E_r/kT)$, $p_r = N_v \exp[(E_r - E_g)/kT]$, and W is the width of the depletion region. N_c and N_v are the conduction and heavy hole valence band density of states, respectively, γ_n and γ_p are recombination coefficients for electrons and holes, and N_r is the density of S–R centers at E_r. For the optimum levels at E_i, the intrinsic energy level, $n_r = p_r = n_i$, and $J_{dep} = qn_i W/(\tau_{no} + \tau_{po})$, the familiar expression for diode depletion current. The bias dependence of J_{dep} is contained in the depletion width, W, which varies as $\sim V^{1/2}$ for an abrupt junction of unlimited extent.

5.3.2.3 Surface generation current

Surface generation currents are typically represented by

$$J_s = qn_i s/2, \quad (5.19)$$

where s is the surface recombination velocity, with no explicit bias dependence. Surface recombination velocity is a term that is used to cover a multitude of possible recombination mechanisms that can occur at the surface of a semiconductor. These mechanisms include

- Recombination via fast surface states located within the semiconductor bandgap at the surface;
- Recombination through S-R centers located within a depletion region located at the surface of the semiconductor, either thermally or by tunneling.

Analytical expressions [74] for the surface recombination velocity associated with these transitions indicate that the recombination velocity depends upon the surface potential of the semiconductor. It will be minimal for accumulated surfaces and considerably larger for intrinsic and depleted surfaces. It is obviously relatively easy to minimize s for either the n- or p-type sides of the junction but not for both at the same time. This is the reason for the popularity of interdiffused semi-insulating wider bandgap semiconductors, such as CdTe, at the surfaces of the detecting layer. It represents an attempt to keep minority carriers away from the surfaces on both sides of the p-n junction using a common passivation.

5.3.2.4 Tunneling via bandgap states

Tunneling via N_t bandgap states at E_t, for a parabolic barrier, in a uniform electric field, with doping N, gives a bias-voltage-dependent current

$$J_{tsr} = [10^{-13} N_t V/E_g] \exp[-1.5 \times 10^{10} \pi m_e^{*\,1/2} E_g^{3/2}/N^{1/2}/(E_g + V)^{1/2}], \quad (5.20)$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

where V is the applied voltage, and $m_e^* = 7 \times 10^{-2} E_g$; E_g is in eV, and N and N_t are in cm^{-3}.

5.3.2.5 Direct tunneling

Direct tunneling between the conduction and valence bands, for a parabolic barrier, in a uniform electric field, with doping N, gives a bias-dependent current

$$J_{tdir} = 1.2 \times 10^{-2} V [N(E_g + V)]^{1/2} \exp[-9.43 \times 10^{10} m_e^{*\,1/2} E_g^{3/2} / [N(E_g + V)]^{1/2}]. \quad (5.21)$$

These expressions for tunneling based on a uniform field should give an overestimate of tunnel current for standard abrupt junction diodes. Allowances for this effect can be made by the inclusion of a factor, f, in the exponents in Eqs. (5.19) and (5.20). Modeling suggests that $1 < f < 2$.

5.3.2.6 Total dark current

Dark current in any HgCdTe diode architecture can be modeled using Eqs. (5.16) through (5.21), provided the relevant parameters are known. As discussed earlier, the Auger generation lifetime τ_A is a reasonably well-understood concept. The largest uncertainty involves the density of S-R centers, their position in the bandgap, and their associated capture coefficients.

Two S-R centers are typically used in the modeling of p-HgCdTe. One is associated with the metal vacancy and has been described in Sec. 4.4.1.2. The S-R level is donor-like at ~0.03 eV below the conduction band, with a density that is proportional to the vacancy concentration. The electron lifetime associated with the donor S-R center is given by $\tau_{vac} = 5 \times 10^9 n_1 / p / N_{vac}$, where $n_1 = N_c \exp(-0.03q/kT)$, and N_c is the appropriate conduction band density of states. The other S-R center is located at the intrinsic energy level and is considered independent of the extrinsic doping concentration. It can be expressed, in general, as $\tau_{ext} = [\tau_{no}(n_i + p) + \tau_{po}(n_i + n)]/(n + p)$. p-HgCdTe typically contains a mixture of vacancies and extrinsic doping, and their lifetime contributions can be added in parallel, in the normal manner. n-HgCdTe is typically limited by the Auger1 mechanism for donor concentrations down to ~5×10^{14} cm^{-3}. Lower donor densities are limited by unspecified S-R centers and give measured lifetime values in the $2-50$ μ s range, depending to a certain extent on composition. Figure 5.19 shows a model of the above dark current components for the specific case of a homo-junction HgCdTe diode with a cutoff wavelength of 10.5 μm at 77 K. A suitable surface passivation is assumed, thus providing a negligible surface recombination velocity. The parameters in the model are typical of high-quality p-type material utilized in the fabrication of ion-implanted $n^+/n^-/p$ diodes. The indicated S-R lifetimes, and tunnel center densities in the depletion region, are reasonably conservative values. The f-factor in the tunneling exponents is assumed to be 1.7.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

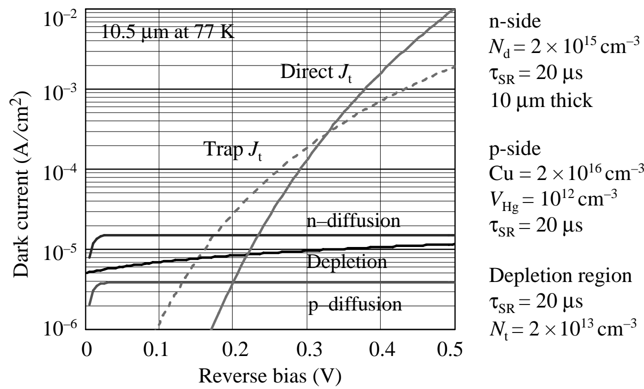


Figure 5.19 Model of dark current components for LWIR at 77 K with assumed parameters. The assumed *f*-factor for tunneling is 1.7.

The indicated diffusion current from the p-volume is strongly dependent on the metal vacancy concentration. Figure 5.20 shows the dependence of p-volume diffusion current as a function of temperature for various extrinsic doping and vacancy concentrations, again for a constant bandgap of 10.5 μm, with a $\tau_{SR} = 20\,\mu$s. Also included is a calculation of one-sided depletion region current for $\tau_{SR} = 20\,\mu$s, assuming a donor concentration of 2×10^{15} cm^{-3}. It is apparent that background-limited performance can easily be achieved in LWIR HgCdTe at temperatures <80 K, even for a 10^{16} cm^{-3} vacancy concentration. It should be pointed out that the above dark current models allow for performance predictions of any operating HgCdTe diode provided that the appropriate S-R lifetimes are defined throughout the unit cell. This capability will be demonstrated in later chapters of this book dealing with advanced detector concepts.

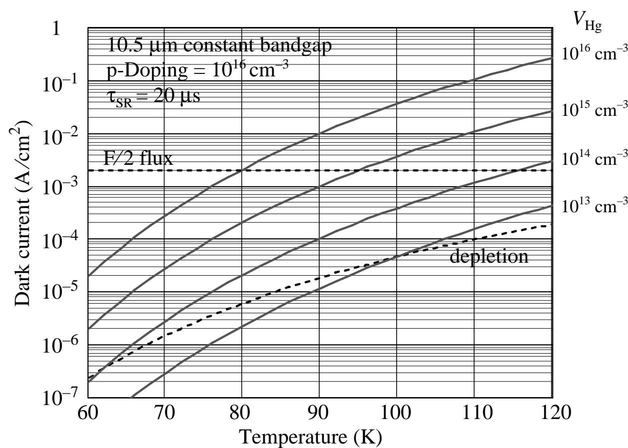


Figure 5.20 p-Region LWIR HgCdTe diffusion and depletion current vs. temperature for various extrinsic and metal vacancy concentrations.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

5.3.3 1/*f* noise

Excess $1/f$ noise is observed in HgCdTe photodiodes from a variety of sources, seemingly depending upon the nature of the dominant dark current component. One interesting aspect of $1/f$ noise in the narrow gap composition range is that, at all reasonable operating temperatures, the thermal transition rates to states in the bandgap are too fast to support the spectral distribution of time constants required to explain the $1/f$ spectral noise density typically observed. In other words, it is unlikely to be "bulk" in nature. Models of this $1/f$ noise thus typically appeal to surface modulation arguments of the dominant noise source. One such model, developed by Schiebel [75], is relevant to HgCdTe diodes that are diffusion current limited and is worth discussion here. The model assumes a surface recombination velocity, s, which is modulated by a fluctuation in surface potential, ψ_s, which is in turn caused by charge tunneling from surface traps, N_{IT}. The traps are located at various tunneling distances from the HgCdTe surface and provide the required $1/f$ spectrum. The noise spectral density of current is given by

$$S_I = \left(\frac{\partial I}{\partial s}\right)^2 \cdot \left(\frac{\partial s}{\partial \psi_s}\right)^2 \cdot \left(\frac{\partial \psi_s}{\partial N_{IT}}\right)^2 \cdot S_{N_{IT}}. \tag{5.22}$$

The individual components are easily modeled for the simple case of a depleted surface region containing S-R centers that control the bulk lifetime, τ_{SR}, and also provide a vehicle for surface recombination.

The p-side diffusion current for a diode of area A, and thickness t < minority carrier diffusion length, is $I = Aqn_i^2 t/p[(1/\tau_{SR}) + (s/t)]^{-1}$, giving

$$\frac{\partial I}{\partial s} = \frac{Aqn_i^2}{p}. \tag{5.23}$$

The dependence of surface recombination velocity on surface potential obtained from a simple consideration of thermal transitions through S-R centers in the depleted surface region is given by

$$\frac{\partial s}{\partial \psi_s} = \frac{1}{\tau_{SR} E_s}, \tag{5.24}$$

where $E_s = [2pq\psi_s/(\varepsilon\varepsilon_o)]^{1/2}$ and represents the electric field in the surface depletion region.

The surface potential is related to the charge in surface traps, N_{IT}, by the simple relationship

$$\frac{\partial \psi_s}{\partial N_{IT}} = \frac{q}{A\left(C_{ins} + C_{dep}\right)}, \tag{5.25}$$

where C_{ins} and C_{dep} are capacitances per unit area associated with the surface insulator and depletion regions.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The insulator trap noise spectral density, S_{NIT}, was first considered by McWhorter [76], and further by Anderson and Hoffman [77], and is given by $S_{NIT} = N_T A/f$, where N_T is the effective density of insulator traps per unit area. Thus,

$$I_n = \frac{qn_i^2}{p\tau_{SR}} \cdot \frac{q}{E_s(C_{ins} + C_{dep})} \cdot \left[\frac{N_T A}{f}\right]^{1/2} = \frac{J_{dif}\ [N_T A]^{1/2}}{tf^{1/2}} \frac{q}{E_s(C_{ins} + C_{dep})}, \tag{5.26}$$

where we have assumed that $t/s \gg \tau_{SR}$. The $1/f$ component of the noise current density thus depends directly on diffusion current density and the surface density of traps as $N_T^{1/2}$. For a simple semiconductor surface, the term $q/E_s/(C_{ins} + C_{dep}) = 1/p$, if $C_{dep} > C_{ins}$. However, the surfaces of HgCdTe photodiodes are normally passivated with thin layers of CdTe, which is also interdiffused to a greater or lesser degree, and this simplification is probably not justified. The band bending at the surfaces of such devices is somewhat complex, but some relevant generalizations can be made. The surface of the semiconductor is unlikely to be either accumulated or inverted, as this will lead to shunted or tunnel current dominated diode performance, which is not the case. Thus, by a process of elimination, we may conclude that the most likely surface condition is one of depletion, at least on one side of the junction. It can also be said that the interdiffusion of CdTe with HgCdTe will result in a voltage $\sim\Delta E_g$ being dropped across the resulting depletion region, which is $\sim$1000 Å thick for the annealing temperatures used, according to interdiffusion models.[4]

For the case of a surface depletion region, the $1/f$ modulation of diffusion current of Eq. (5.26) can be arrived of in a somewhat simpler manner than described above. Consider the cross-section geometry of the n-side of a diode. In the spirit of the above argument, we can utilize the depletion width W as the variable, instead of surface recombination velocity, s. Thus, Eq. (5.22) becomes

$$S_I = \left(\frac{\partial I}{\partial W}\right)^2 \cdot \left(\frac{\partial W}{\partial \psi_s}\right)^2 \cdot \left(\frac{\partial \psi_s}{\partial N_{IT}}\right)^2 \cdot S_{N_T}, \tag{5.27}$$

and I is $qn_i^2(t - W)/N\tau$, giving

$$\frac{\partial I}{\partial W} = -\frac{Aqn_i^2}{N\tau} = -\frac{I}{t}, \tag{5.28}$$

for $t \gg W$.

By definition,

$$\frac{\partial W}{\partial \psi_s} = \frac{1}{E_s}. \tag{5.29}$$

Combining Eqs. (5.27), (5.28), (5.29), and (5.25), together with $S_{NT} = [N_T A/f]$, and $J_{dif} = I/A$, gives for the noise spectral density of the current

$$I_n = \frac{J_{dif}\ [N_T A]^{1/2}}{tf^{1/2}} \frac{q}{E_s(C_{ins} + C_{dep})}, \tag{5.26}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

which is identical to Eq. (5.26).

The $1/f$ component of the noise current density thus depends directly on diffusion current density and the surface density of traps as $N_T^{1/2}$. As pointed out by Schiebel, this model indicates the importance of looking at noise as the derivative of a dark current component with respect to a fluctuating parameter, and the principle should have general applicability.

5.4 Detector Performance

The VIP and HDVIP technologies are currently employed at DRS for production manufacturing of a wide variety of scanning and staring HgCdTe FPAs and detector dewar cooler assemblies (DDACs). These arrays are listed in Table 5.2, together with some relevant performance parameters. The scanning FPAs have the VIP architecture, and the staring FPAs are HDVIP. The products of other manufacturers, utilizing other architectures, are comparable, in that all current-generation FPAs are essentially background noise limited. The scanning arrays have been in volume production for a number of years, and the operabilities are typically $>98\%$. They are typically built in the TDI architecture, which easily allows for pixel substitution. The 240×2 FPA is manufactured for the JAVELIN Commander's Launch Unit (CLU), and the 480×5 for SADA II. The units meet, or exceed, all system specifications, which are classified. Collection efficiencies are typically $>75\%$, and MTF is theoretical, due to the front-side illumination of the FPA, combined with the use of an opaque field plate to define the optical area.

The operabilities of the staring arrays are typically $>99\%$, with minimal or no clusters. When clusters are observed, they are mostly associated with the array mounting process and not due to grown in material defects. Excellent responsivity uniformity is obtained, indicating cutoff wavelength control across the FPA, and the collection efficiency (i.e., quantum efficiency $\times$ unit cell fill factor) is $>66\%$. The dark current for MWIR HgCdTe is essentially not measurable at 77 K.

Table 5.2 Current-generation FPAs manufactured at DRS.

Array format	240 × 2	240 × 4	480 × 5	256 × 256	256 × 256	480 × 640	480 × 640
Spectral band (μm)	8–10	8–10	8–10	3–5	8–10.5	3–5	8–10.5
Pixel size (μm)	33 × 22	28 × 25	38 × 28	40 × 40	40 × 40	25 × 25	25 × 25
Pitch (μm)	50	50	50	40	40	25	25
Coll. efficiency	>75%	>75%	>75%	>60%	>55%	>60%	>60%
MTF @ Nyquist	Theory	Theory	Theory	>90%	>90%	>90%	>90%
Operating temp.	77	77	77	77	77	77	77
ROIC input	CTIA	CTIA	CTIA	BDI	BDI/PA	BDI	BDI/PA
NEDT/pixel (mK)		30		8	13	10	13
Frame rate (Hz)				120	120	60	60
Dyn. range (dB)	72	72	72	74	74	74	74

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Similar behavior is observed for the large-area LWIR HgCdTe FPAs at 77 K. Dark currents at 77 K are diffusion limited and typically $\sim 100\,\mu A/\text{cm}^2$, for a cutoff wavelength of 10.3 μm, in agreement with the dark current models.

Excess $1/f$ noise in HgCdTe diodes has been found to statistically depend on dislocation density of the HgCdTe material. At 77 K, the requirement on dislocation density for LWIR material [78] is $<2 \times 10^5/\text{cm}^2$. MWIR, on the other hand, can tolerate higher densities of dislocations at 77 K, but evidence is mounting that this may no longer be true at higher operating temperatures. Measurements taken at DRS [79] of $1/f$ noise of CdTe passivated $n^+/n^-/p$ VIP and HDVIP diode structures indicate a dependence on device dark current density, as shown in Fig. 5.21, for HgCdTe material with nominally $<2 \times 10^5/\text{cm}^2$ dislocations. The $1/f$ noise figure is expressed as the measured noise current at 1 Hz normalized to the square root of

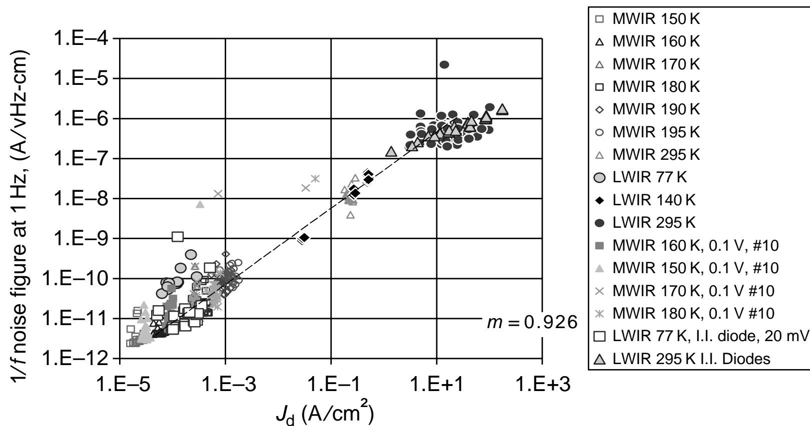


Figure 5.21 $1/f$ noise figure vs. dark current density for various HgCdTe devices.

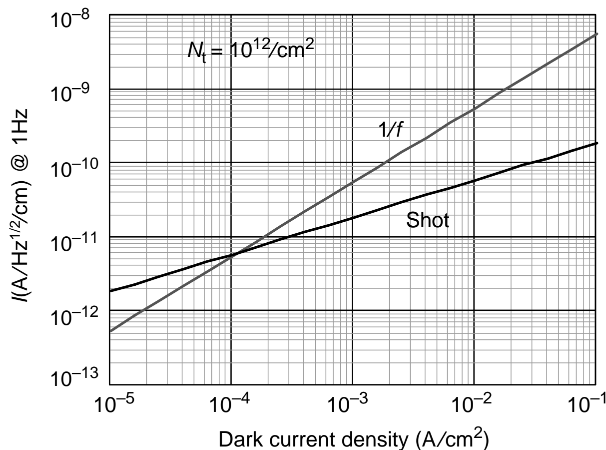


Figure 5.22 $1/f$ noise relative to shot noise for HgCdTe photodiodes with $N_T = 10^{12}/\text{cm}^2$.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

the area. The measured dependence of $1/f$ noise, for dark currents dominated by diffusion, depends upon the absolute value of dark current, regardless of HgCdTe x-value or operating temperature. This is in keeping with $1/f$ noise theory discussed earlier, based on a modulation of diffusion current by fluctuating surface potential associated with tunneling of charge from traps. The required surface trap density to fit the experimental data is $N_T \approx 10^{12}/\text{cm}^2$.

Equation (5.26) allows us to meaningfully predict $1/f$ noise performance for any diffusion-limited photodiode. The shot noise on the flux and dark currents at 1Hz is given by $[2q(J_{\text{flux}} + J_{\text{dif}})]^{1/2}$ A/Hz$^{1/2}$/cm, and this can be compared to the noise figure, $I_{1/f}$ given by Eq. (5.26). Such a comparison is shown in Fig. 5.22, for a surface trap density of $10^{12}/\text{cm}^2$. The corresponding noise corner is given by $f_c \approx [I_{1/f}^2/2q(J_{\text{flux}} + J_{\text{dif}})]$.

5.5 HgCdTe: Conclusions

HgCdTe has an extremely well understood technology database and is the material of choice for second-generation IRFPAs. A variety of architectures are available for FPA fabrication, depending to a certain extent on the desired cutoff wavelength and size of the FPA. The larger the array, and the longer the cutoff wavelength, the more difficult it is to utilize bump-bonded architectures. HDVIP is in fact the only architecture to successfully yield BLIP 10.5 μm cutoff LWIR 640 × 480 25 μm pitch FPAs operating at 77 K. The success of this technology is due to its capacity to accommodate the mismatch in coefficients of thermal expansion of HgCdTe and the silicon ROIC.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 6

Uncooled Detection

The ultimate IRFPA will operate at room temperature and exhibit background limited performance (BLIP), regardless of the operating cutoff wavelength. Currently the primary vehicle for room-temperature IR detection is the thermal detector, but it is essentially limited to the detection of LWIR radiation, at relatively slow frame rates, or integration times. These limitations were discussed briefly in Sec. 2.2 for the optimized perfect thermal detector. The successful room-temperature detection of the complete IR spectrum, with BLIP performance, will require the utilization of other detector concepts. The HOT detector of Elliott and Ashley [80] is just such a concept, the potential of which is limited entirely by the quality of the direct bandgap semiconductor utilized to detect the IR radiation and the contacts and passivation applied to that material. The HgCdTe HOT detector concept is discussed here, together with a consideration of relevant material properties within the active volume of the device for successful room-temperature operation, taking into account the all-important issue of Shockley–Read (S-R) centers in HgCdTe. The HOT concept is not limited to the HgCdTe materials system alone, but it can be implemented in any direct bandgap semiconductor material or its bandgap engineered equivalent.

6.1 Thermal Detection

The principle of the thermal detector is illustrated in Fig. 6.1. The detector element, of heat capacity C_{th}, is thermally coupled to a heat sink by a conductance G_{th}. The thermal detector element is micromachined onto the surface of a silicon readout integrated circuit (ROIC) and the resulting pedestal is connected to the ROIC by thin metal buss leads. The pedestal can be fashioned from any one of a number of temperature-sensitive materials that are process compatible with silicon ROIC processing. We will assume here that the temperature sensitivity of the element is sufficiently large that all noise sources other than temperature fluctuations of the detector element can be ignored. As discussed in Sec. 2.2, the power fluctuation of the element is then given by $\Delta W_{\mathrm{n}} = [4kT^2 G_{\mathrm{th}} \Delta f]^{1/2}$, where Δf is the noise bandwidth, which translates into a noise temperature fluctuation, $\Delta T_{\mathrm{n}} = [4kT^2 R_{\mathrm{th}} \Delta f]^{1/2}$. The detector signal is

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

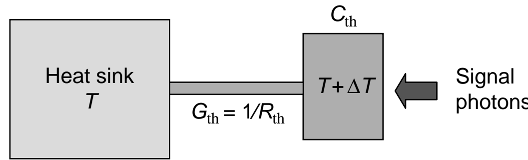


Figure 6.1 IR thermal detector.

given by $\Delta T_s = \eta R_{th}(dP/dT)_\lambda \Delta TA/(1+4F^2)$, where η is the quantum efficiency, $(dP/dT)_\lambda$ the differential change in radiated power per scene temperature change in the spectral region of interest, A the detector area, and F the $F/\#$ of the optical system. Thus, the detectivity of the detector, defined as [signal/flux power][$A^{1/2}/(\text{noise}/\Delta f^{1/2})$], is given by

$$D^* = [\eta^2 A/(4kT^2 G_{th})]^{1/2}. \tag{6.1}$$

If G_{th} is radiatively limited, then $G_{th} = 4PA/T$, where $P = \sigma_B T^4$. Thus, for $\eta = 1$, $D^* = 2\times 10^{10}\,\text{cm}\,\text{Hz}^{1/2}/\text{W}$ and is independent of area.

This value of D^* is comparable to the theoretical limit of a LWIR photon detector, but it provides only a limited insight into the suitability of the thermal detector for specific IR systems. As pointed out in Sec. 2.2, signal and spectral bandwidth requirements are also important, and to this end it is meaningful to consider the available noise equivalent temperature difference, NEΔT.

Correlated double sampling [4] of the bandwidth-limited temperature fluctuation, with an integration time τ_{int}, gives $\Delta T_n = \{2kT^2[1-\exp(-\tau_{int}/R_{th}C_{th})]/C_{th}\}^{1/2}$, where we have assumed $\Delta f = 1/4R_{th}C_{th}$. Equating this to a signal temperature change given by $\Delta T_s = (dP/dT)_\lambda \Delta TA[1-\exp(-\tau_{int}/R_{th}C_{th})](1+4F^2)R_{th}$ gives

$$\text{NE}\Delta T = \left\{2kT^2/C_{th}/\left[1-\exp(-\tau_{int}/R_{th}C_{th})\right]\right\}^{1/2}(1+4F^2)/(dP/dT)_\lambda/AR_{th}. \tag{6.2}$$

Image smearing considerations require $\tau_{int} \approx 2R_{th}C_{th}$, resulting in

$$\text{NE}\Delta T = [2kT^2/C_{th}]^{1/2}(1+4F^2)/(dP/dT)_\lambda/AR_{th}.$$

For $R_{th} < R_{rad}$

$$\text{NE}\Delta T \approx [2kT^2 C_{th}/\tau_{int}^2]^{1/2}(1+4F^2)/(dP/dT)_\lambda/A, \tag{6.3}$$

and for $R_{th} = R_{rad}$

$$\text{NE}\Delta T \approx [32k/C_{th}]^{1/2}(1+4F^2)/[(dP/dT)_\lambda/P]. \tag{6.4}$$

This predicted dependence of NEΔT on integration time, τ_{int}, for LWIR and MWIR detection is shown in Fig. 6.2 for a 25-μm pixel, operating with $F/1$ optics, and $\eta = 1$.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

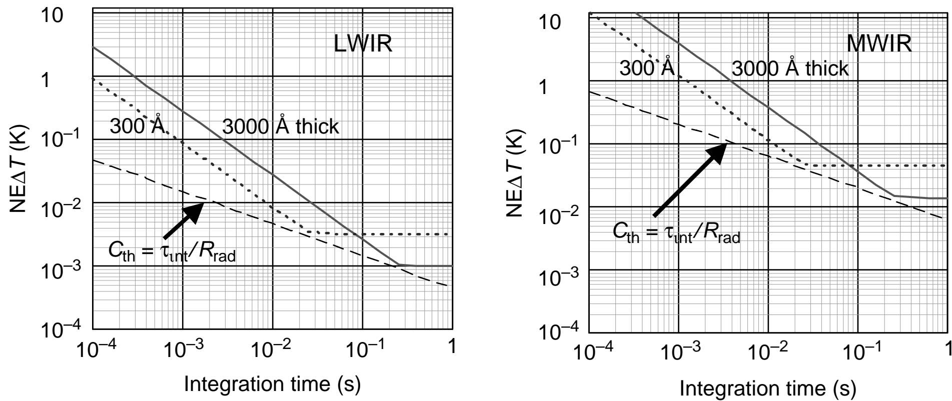


Figure 6.2 NEΔT vs. integration time for perfect LWIR and MWIR thermal detectors for 25 μm pixels, $F/1$ optics, and $\eta = 1$.

The curves labeled $C_{th} = \tau_{int}/R_{rad}$ represent the optimum mode of detector operation, assuming that (1) the detector thermal conductance is limited by radiation to its surroundings, and (2) the thermal capacitance can always be made small enough to satisfy the requirement of $\tau_{int} \approx R_{rad}C_{th}$. In practice, the available thermal capacitance is limited by the thickness of the bolometer pedestal and material specific heat. Realistic values for bolometer thickness are in the range 300–3000 Å, and the volumetric specific heat is 1–2 J/cm^3/K. Examples of fixed bolometer thickness values of 300 Å and 3000 Å are shown in Fig. 6.2, encompassing both ranges of R_{th} described by Eqs. (6.3) and (6.4). Bolometers today are typically 3000 Å thick, and their performance approaches the values shown in Fig. 6.2 for LWIR at relatively long integration times, such as 15–30 ms. However, it is apparent that high performance in a snapshot mode ($\tau_{int} \approx 2$ ms) will not be realized. At best, NEΔT values can approach 0.1 K for LWIR under these conditions. The NEΔT achieved by thermal detectors in the MWIR is mediocre at best. This is due to the low value of $(dP/dT)_\lambda$ available for signal in this spectral band, relative to the temperature noise, which is determined by all of the black-body flux absorbed by the bolometer. Theoretically, this situation could be improved by resonant tuning of the pedestal to its surroundings for the same IR spectral band as the desired signal; however, this would not only be difficult but, if successful, would also increase R_{rad} significantly and result in considerably longer thermal time constants. For slower optics the situation is exacerbated, varying as $(F/\#)^2$. Thinner detector elements improve the NEΔT as $t^{1/2}$.

6.2 Photon Detection

6.2.1 HOT detector theory

The high operating temperature (HOT) photon detector concept, first proposed by Elliott and Ashley, is depicted in Fig. 6.3 for an $n^+/\pi/p^+$ architecture, where π

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

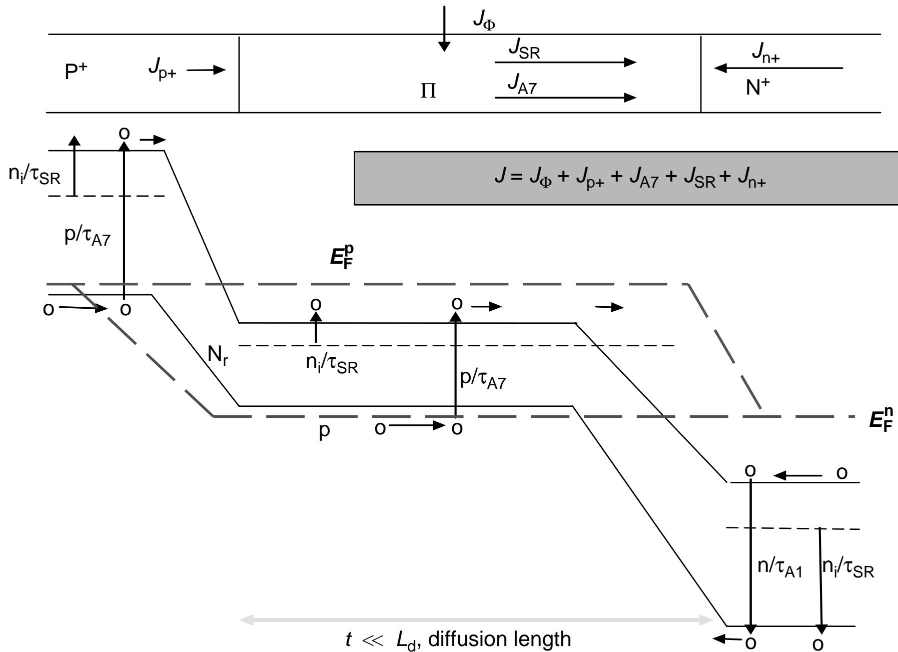


Figure 6.3 HOT detector concept with dark current sources.

designates an intrinsic region containing a p-type background dopant. The intrinsic IR absorbing volume is connected to both a minority carrier contact and a majority carrier contact. The geometry of the active volume is small relative to a minority carrier diffusion length. It is operated in strong non-equilibrium by reverse biasing the minority carrier contact to completely extract all of the intrinsically generated minority carriers. Charge neutrality in the active volume is violated, creating an electric field to sweep out the intrinsically generated majority carriers until the majority carrier concentration equals the background dopant concentration. The active volume thus consists of a depletion region, with a width determined by the doping concentration and applied bias, and a field-free region with a majority carrier concentration determined by background doping. The elimination of minority carriers throughout the device means that minority carrier recombination is a non-issue, and dark current is determined entirely by generation mechanisms within the active volume of the device, plus possible components associated with injection from the contact and surface regions. The relevant dark current mechanisms within the active volume of the HOT detector are (1) Auger generation associated with majority carriers in the field-free region, and (2) thermal generation through S-R centers throughout the whole non-equilibrium volume. As discussed earlier, the arguments of Humphreys suggest that radiative mechanisms are not important in this device.

Auger generation–recombination in direct-gap semiconductors is a well-understood phenomenon involving the interaction of three carriers. The Auger1

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

lifetime in n-type material is due to the interaction of two electrons and a heavy hole and is given by

$$\tau_{A1} = 2\tau_{Ai1} n_i^2 / n(n+p), \tag{6.5}$$

$$\tau_{Ai1} = 7.6 \times 10^{-18} \varepsilon^2 (1+\mu)^{1/2} (1+2\mu) / ((m_e/m_o)|F_1 F_2|^2 (kT/E_g)^{3/2}) \times [\exp((1+2\mu)E_g/(1+\mu)kT)], \tag{6.6}$$

and $\mu = (m_e/m_h)$. τ_{Ai1} is defined as the intrinsic Auger1 lifetime. The largest uncertainty in the Auger model lies in the calculation of the overlap integral $F_1 F_2$, and empirical values, provided by a comparison to experimental data [81], are typically used. For HgCdTe, the empirical value for $F_1 F_2$ is ~0.22. In p-type material, Auger recombination involves two holes and an electron and is referred to as Auger7. Calculations by Casselman [82] for HgCdTe give $\tau_{Ai7} \approx 6\tau_{Ai1}$, although experimental data would indicate that it may be somewhat higher, namely $\sim 10\tau_{Ai1}$. This longer lifetime suggests that p-HgCdTe is the material of choice for the HOT detector active volume with regard to ultimate dark current performance. The Auger generation rate/unit volume in the field-free region of the HOT detector, under non-equilibrium conditions, is proportional to the majority carrier concentration and is given by

$$G_{A7} = n_a / 2\tau_{Ai7}. \tag{6.7}$$

In p-HgCdTe under non-equilibrium conditions, the minority carrier generation rate/unit volume associated with an S-R center of density N_r, located E_r from the conduction band, is given by

$$\begin{aligned} G_{SR} &= n_i^2 / [\tau_{pr} n_1 + \tau_{nr}(p + p_1)], \\ n_{1r} &= N_c \exp(-E_r q/kT), \\ p_{1r} &= N_v \exp(-(E_g - E_r)q/kT), \end{aligned} \tag{6.8}$$

where $\tau_{nr} = 1/\gamma_n N_r$, $\tau_{pr} = 1/\gamma_p N_r$. γ_n and γ_v are the capture coefficients for electrons and holes into N_r, and N_c and N_v are the conduction-band and valence-band density of states, respectively. In the field-free region of the detector, $p = n_a$, whereas in the depletion region $p = 0$.

The dark current density in a simple planar geometry HOT detector of thickness t, with a depletion region width W, is thus given by

$$J = qn_a(t - W)/2\tau_{Ai7} + \sum^{r} qn_i^2 \{W/(\tau_{pr} n_{1r} + \tau_{nr} p_{1r}) + (t - W) / [\tau_{pr} n_{1r} + \tau_{nr}(n_a + p_{1r})]\}, \tag{6.9}$$

where the summation is carried out over all the S-R centers in the device. The materials technology development required for HOT detector optimization is clear:

- Reduce the majority carrier concentration in the active volume of the device, and
- Reduce the density of S-R centers in the active volume to a minimum.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Reducing the majority carrier concentration is beneficial in two respects. It not only results in a reduction in Auger generation rate, but it also results in a larger depletion region for the diode at a constant bias, and hence a smaller required field-free volume. For an applied bias of 1 V, a depletion region width of 5 μm is achieved for doping concentrations of $10^{14}\,\mathrm{cm}^{-3}$. Such a depletion region width will provide ample absorption of the IR radiation, with good quantum efficiency, and completely eliminates Auger generation due to the total absence of majority carriers. The device is then essentially a P-I-N diode, and either p-type or n-type material can be employed for the intrinsic region. Interestingly enough, doping concentrations of $1–2 \times 10^{14}\,\mathrm{cm}^{-3}$ are already available using indium background doping in LPE layers grown in Te-rich melts. A P-I-N diode is shown in Fig. 6.4, where the p^+ contact is shown as wide-bandgap HgCdTe, to avoid the possibility of minority carrier injection into the intrinsic active volume of the HOT device. The band-filling in the n^+ contact is assumed deep enough to inhibit Auger transitions in that region, and hence minority carrier generation and injection. A wide-bandgap n^+ region could also be utilized for this purpose. The dark current in such a device is given by

$$J_{\mathrm{d}} = \sum_{\mathrm{r}} q n_{\mathrm{i}}^2 [W/(\tau_{\mathrm{pr}} n_{1\mathrm{r}} + \tau_{\mathrm{nr}} p_{1\mathrm{r}}], \tag{6.10}$$

where the summation is over all the S-R centers, and the depletion width is equal to the device width. The most efficient S-R centers for thermal generation are normally located close to the intrinsic energy level in the bandgap, for $\tau_{\mathrm{pr}} \approx \tau_{\mathrm{nr}}$. In HgCdTe, the intrinsic level is located far above mid-gap due to the fact that the conduction band effective mass is much less than the heavy hole mass.

Thus, for the well-designed P-I-N device, operating under reverse bias, the dark current is determined solely by the density and distribution of S-R centers in the bandgap of the active volume of the diode, and the success or failure of the HOT detector depends entirely upon the management of these bandgap states.

It is of interest to ascertain the magnitude of N_{r} required to enable background-limited performance at room temperature for a particular cutoff wavelength.

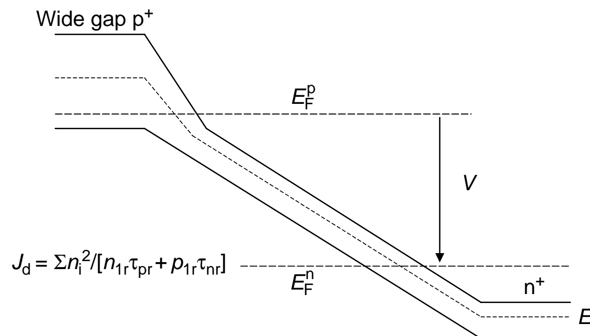


Figure 6.4 A reverse-biased P-I-N diode.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

For neutral S-R centers, this requires that

$$G = W n_i \gamma N_r < \Phi_B,$$

where we have assumed unit quantum efficiency. Further assuming that $\gamma \approx 10^{-9}\,\text{cm}^3/\text{s}$, which is typical of neutral S-R centers, $W = 5 \times 10^{-4}\,\text{cm}$, then, for LWIR HgCdTe at 295 K, $n_i \approx 5 \times 10^{16}\,\text{cm}^{-3}$, $\Phi_B = 4 \times 10^{17}/\text{cm}^2/\text{s}$, and the requirement on N_r is $<1.6 \times 10^{13}\,\text{cm}^{-3}$. This refers to the sum total of S-R centers in the material. The maximum allowed density of S-R centers for BLIP operation of a P-I-N diode at 295 K, as a function of cutoff wavelength, is shown in Fig. 6.5. Under these conditions, in the important MWIR/LWIR spectral regions, it is apparent that N_r needs to be $<10^{12}\,\text{cm}^{-3}$, which represents a significant materials challenge, but with enormous potential payoff. For a recombination coefficient of $10^{-10}\,\text{cm}^3/\text{s}$, the requirement on N_r is relaxed by an order of magnitude. The presence of S-R centers away from the intrinsic energy level will modify the predictions of Fig. 6.5, but not too significantly, due to the magnitude of kT relative to the bandgaps involved. This requirement on N_r is at the limit of sensitivity for most impurity detection techniques such as glow discharge mass spectroscopy (GDMS) and secondary ion mass spectroscopy (SIMS).

Knowledge of S-R centers in HgCdTe is limited but improving. As discussed in Sec. 4.4.1.2, metal vacancies and extrinsic p-type dopants have been characterized by Hall and minority carrier lifetime data over a limited range of $Hg_{1-x}Cd_xTe$ x-values, primarily in the range of $0.2 < x < 0.3$ (LWIR to MWIR).

Metal vacancies are characterized by a donor-like S-R center located ~30mV from the conduction band for compositions in the $0.2 < x < 0.3$ range. Lifetime measurements indicate an electron lifetime for the vacancy given by

$$\tau_{vac} = 5 \times 10^9 [(n + n_1)/p]/N_{vac}. \tag{6.11}$$

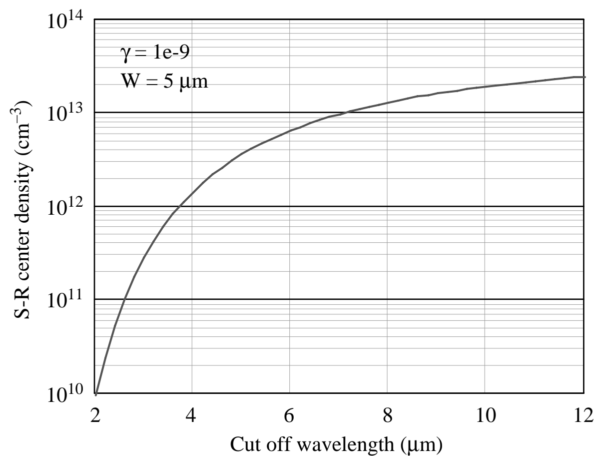


Figure 6.5 Maximum density of allowed S-R centers in a P-I-N diode for BLIP at 295 K.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

For extrinsically doped HgCdTe, with no vacancies, S-R centers are characterized by neutral-like behavior with a relatively low density at the intrinsic energy level in the bandgap. The measured electron lifetime in this case is given by

$$\tau_{ext} = \tau_{no}[n_i + p]/p. \tag{6.12}$$

The dependence of lifetime on inverse temperature and doping concentration described in Eqs. (6.11) and (6.12) is shown in Fig. 6.6 for LWIR HgCdTe.

Other S-R centers that have a dramatic effect on lifetime at low temperatures have been characterized in HgCdTe. The characteristic of such a center as a function of inverse temperature is shown in Fig. 6.7, for 10^{15} cm^{-3} arsenic-doped MWIR HgCdTe, and compared to the intrinsic level alone. The lifetime is independent of

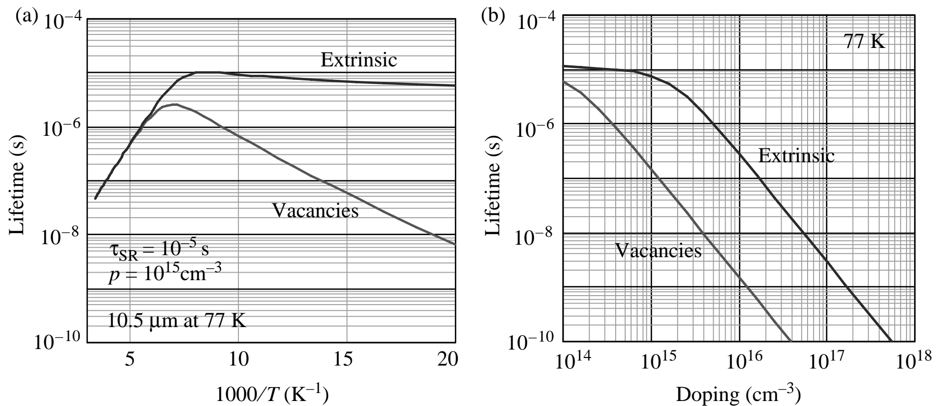


Figure 6.6 Electron lifetime as a function of (a) $10^3/T$ and (b) doping concentration for S-R centers associated with metal vacancies and extrinsic dopants in LWIR HgCdTe.

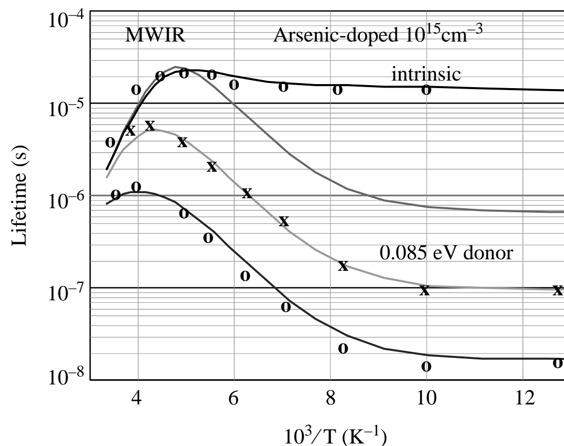


Figure 6.7 Lifetime vs. inverse temperature for a lifetime-inhibiting S-R center in MWIR arsenic-doped HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

the doping concentration, and the center is again donor-like, located 0.085 eV from the conduction band.

It is apparent from the signatures of these various S-R centers as a function of temperature that the lifetime of HgCdTe can be greatly compromised at low temperatures. High performance at low temperatures mandates the use of pure, extrinsically doped HgCdTe. However, it is interesting to note that at high temperatures the lifetime characteristics tend to merge, and the only relevant factor is the total density of S-R centers. This is to be expected from our simple theory invoked above for the non-equilibrium P-I-N diode.

The generation rate in the non-equilibrium P-I-N diode is given by

$$G = \sum^{r} n_{\rm i}^2 [W/(\tau_{\rm pr} n_{\rm 1r} + \tau_{\rm nr} p_{\rm 1r}] = \sum^{r} W n_{\rm i}^2 \gamma_{\rm n} \gamma_{\rm p} N_{\rm r}/(\gamma_{\rm n} n_1 + \gamma_{\rm p} p_1). \tag{6.13}$$

For neutral centers, the capture coefficients for electrons and holes are approximately equal, and $\gamma_{\rm n} = \gamma_{\rm p} = \gamma$. Equation (6.13) will then maximize for centers at the intrinsic level, giving $G = W n_{\rm i} \gamma N_{\rm r}$. For a donor center, the denominator of Eq. (6.13) will be dominated by $\gamma_{\rm n} n_1$, and $G = W p_1 \gamma_{\rm p} N_{\rm r}$, where $\gamma_{\rm p}$ will be $\sim\gamma$. In this case, p_1 may be larger than $n_{\rm i}$, depending on the position of the donor level relative to the intrinsic level. Similar arguments would apply for acceptor-like S-R centers.

6.2.2 HOT detector data

Arsenic-doped MWIR HgCdTe has been grown at DRS by MBE, Te-rich, and Hg-rich LPE. The longest lifetimes at 77 K have been achieved with Te-rich LPE and MBE, and some of the data are shown in Fig. 6.8. The values agree with a combination of S-R, internal radiative, and Auger7 lifetimes, according to the

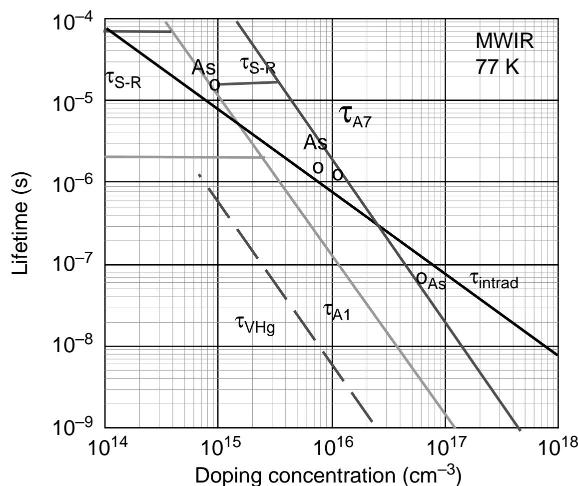


Figure 6.8 Lifetime at 77 K vs. arsenic doping for Te-LPE and MBE MWIR HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

models presented earlier. Auger1 and S-R lifetime models appropriate for n-type HgCdTe are also included for comparison purposes. These values agree with those of other investigators [83, 84].

HDVIP diode mini-arrays have been fabricated on both arsenic- and Au-doped material utilizing the structure shown earlier in Fig. 5.13. The n^+ regions of the diodes were generated by ion etching and implantation of $10^{15}/cm^2$ of boron at 150 keV. The periphery of the diode was protected by a reverse-biased guard diode to exclude dark current generated outside the volume of interest, as illustrated in Fig. 6.9.

The dark current performance of diodes fabricated on both Au- and arsenic-doped material is shown in Fig. 6.10 as a function of temperature, and the

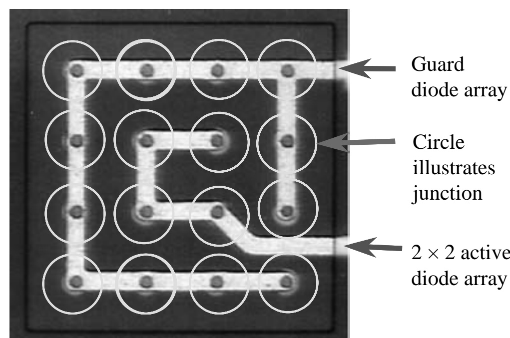


Figure 6.9 Geometry of a unit cell with a diode mini-array and guard diode ring.

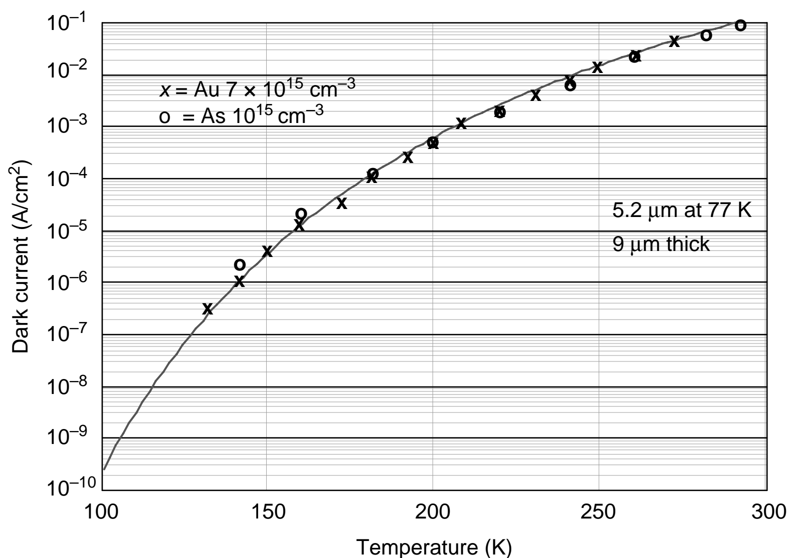


Figure 6.10 Measured diode dark-current density vs. temperature for two HOT material layers, one with an arsenic at 10^{15} cm^{-3}, and the other with Au at 7×10^{15} cm^{-3}.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

performance is compared to a model for the expected diffusion-current components calculated from the measured lifetime and carrier concentration values. The material composition is $x \approx 0.29$, with an estimated cutoff at 240 K of 4.6 μm.

Diffusion current is seen to dominate over the whole temperature range, and it is thought to be due to a small number of residual vacancies left in the unit cell after diode processing. One surprising aspect of the data is the lack of a depletion-current component at low temperatures, where it should prevail and exhibit its customary n_i slope. The absence of this component allows for an estimate of the minimal S-R lifetime that must prevail within at least the n-region of the HDVIP diodes. The total absence of depletion current at 130 K is consistent with a value for $\tau_{sr} > 500\,\mu s$, which is considerably larger than measured values for MWIR n-type material at these temperatures. It is not clear whether this discrepancy is due to the vagaries of the measurement technique or a result of the diode processing. MWIR FPAs have been fabricated with these dark current values. Similar dark current data have been reported by other investigators [85].

6.2.3 HOT detector contacts

The majority and minority carrier contacts to the HOT diode must not introduce minority carriers into the active volume of the device. The obvious way to achieve this is by utilizing wide-bandgap interdiffused compositions of HgCdTe in these regions, resulting in low n_i values and hence low generation rates for minority carriers. For p-type material, the minority carrier contact is n^+. The conduction band effective mass is relatively small in HgCdTe; hence, significant band-filling occurs for n^+ doping concentrations of $10^{18}\ cm^{-3}$, which are easily achieved by ion implantation and ion etching. If the band-filling approaches a bandgap in value, then a significant improvement in minority carrier lifetime will occur due to the inhibition of Auger recombination, and wide-bandgap material may not be necessary. For the p^+ contact, the effective mass is large, and no such band-filling occurs. The only choices for this contact are wide-bandgap HgCdTe or the use of very low volume contacts within the unit cell of the device. Remote contacts can also be used if the substrate current is not too large, enabling the avoidance of substrate debiasing issues.

6.2.4 HOT detector options

Theoretically, any direct-gap semiconductor, or its bandgap-engineered equivalent, can be utilized for HOT detection. Because of their inherent band structure, with a built-in immunity to Auger recombination, the Pb salt alloy systems come to mind. Another alternative to HgCdTe that is currently receiving a lot of press is the Type II superlattice, typified by the InAs/GaSb system, as shown in Fig. 6.11(a). The conduction band edge of InAs lies below the valence band edge of GaSb; however, layer quantization is employed to produce a positive bandgap between the InAs and GaSb. Bandgaps of 0.1 eV are achieved with layer thicknesses in the 50–100 Å

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

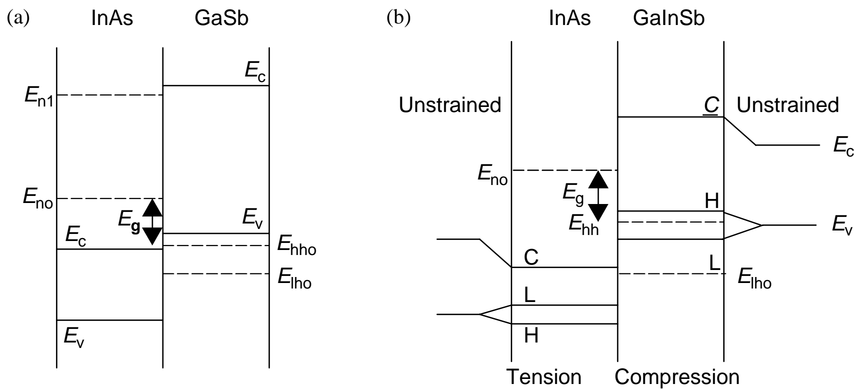


Figure 6.11 Type II superlattices (a) InAs/GaSb and (b) InAs/GaInSb, with strain effects.

range. However, the electrons and holes are localized in different layers, and the optical matrix element for photon absorption depends directly on the degree of electron and hole wavefunction overlap in the layers. For layer thicknesses appropriate for LWIR detection in the InAs/GaSb system, this overlap is small, resulting in low absorption coefficients. This problem is eliminated by the proposal of Smith and Mailhiot [86] to replace GaSb with the alloy GaInSb, thus utilizing the effects of strain associated with the increasing lattice constant of $Ga_{1-x}In_xSb$ with x, relative to the lattice constant of InAs. The strain places the InAs in biaxial tension and the GaInSb in biaxial compression. These biaxial stresses can be represented by equivalent hydrostatic and uniaxial stress components. The hydrostatic stress effectively changes the direct bandgap values without removing any band degeneracies. However, the uniaxial stress components lift the valence band degeneracy, resulting in the modified band structure shown in Fig. 6.11(b). The increased conduction and valence band overlap allows layer thicknesses ~25 Å to provide both the desired LWIR bandgap values together with absorption coefficients $> 10^3$ cm^{-1}. The optical matrix element for the InAs/GaInSb system is less than that for the HgCdTe alloy with the same bandgap, but this is compensated by the large conduction band density of states associated with the electron effective mass values of the superlattice, which are typically $\sim 0.04\, m_o$. These large values of effective mass will result in

- Significantly lower values of tunnel current in diodes fabricated on this materials system relative to alloy HgCdTe, provided that similar values of doping concentration are attainable, namely $\sim 1\text{–}2 \times 10^{14}$ cm^{-3};
- Larger values of n_i, and hence larger values of thermally generated dark currents relative to HgCdTe with similar lifetimes.

It has been pointed out that the bandgap engineering approach can also be used to good effect in reducing the contribution of band-to-band Auger mechanisms to the recombination and generation of minority carriers in this materials system. This is achieved by tailoring the layer quantization and strain effects in the superlattice to significantly reduce the phase space available to the electrons and holes participating

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

in the relevant Auger transition. The phase space is constrained by the requirement of energy and momentum conservation in the Auger transition.

This of course would be of limited use in the proposed P-I-N version of the HOT diode, where Auger generation is absent due to a total absence of free carriers in the active volume of the device. The most stringent requirement on the bandgap-engineered device, as for HgCdTe, will be the allowable density of S-R centers in the material.

6.3 Uncooled Photon vs. Thermal Detection Limits

In view of the current emphasis on uncooled detection, it is of interest to compare the fundamental limits of uncooled photon and thermal detectors. The most suitable parameter for comparison purposes is NEΔT, as defined earlier in Chapter 2. Considering an uncooled photon detector with unity gain, quantum efficiency, and optics transmission, with an integrating node, then, for zero dark current, the noise is given by $N^{1/2} = [\Phi_B \tau_{int} A]^{1/2}$. Equating this to the signal of $\Delta T (d\Phi_B/dT)\tau_{int} A/(1 + 4F^2)$, we obtain

$$\mathrm{NE}\Delta T = [N^{1/2}C]^{-1}(1 + 4F^2), \tag{6.14}$$

where C is the scene contrast $(d\Phi_B/dT)/\Phi_B$. This is to be compared with the thermal detector for the radiatively limited thermal conduction case given by Eq. (6.4), namely

$$\mathrm{NE}\Delta T = [32k/C_{th}]^{1/2}(1 + 4F^2)/[(dP/dT)_\lambda/P]. \tag{6.4}$$

The heat capacity of the bolometer material can be expressed in terms of lattice vibration modes, giving

$$\mathrm{NE}\Delta T = 4(1 + 4F^2)/N_\omega^{1/2}/[(dP/dT)_\lambda/P], \tag{6.15}$$

where N_ω is the number of lattice modes in the bolometer element. The similarity between Eqs. (6.14) and (6.15) is striking. We can qualitatively equate the number of carriers integrated in the photon detector with the number of lattice modes of the thermal detector, and the scene contrast, C, expressed in terms of photons, with $[(dP/dT)_\lambda/P]$, the normalized differential change in radiated power per scene temperature change in the spectral region of interest.

The predictions of Eqs. (6.14) and (6.15) are shown in Fig. 6.12 with the following assumptions. The system parameters are $F/1$, with a 25-μm pixel size, operating at 290 K. The responsivity of the thermal detector is large enough that it is limited purely by temperature fluctuations of the bolometer pedestal, and its heat capacity can always be reduced to satisfy the condition $\tau_{int} = C_{th}R_{rad}$, where R_{rad} is determined by the complete IR spectral bandwidth; that is, it is perfectly black. Dark current in the photon detector is negligible relative to background flux, and the integrating node is large enough to store all of the charge.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

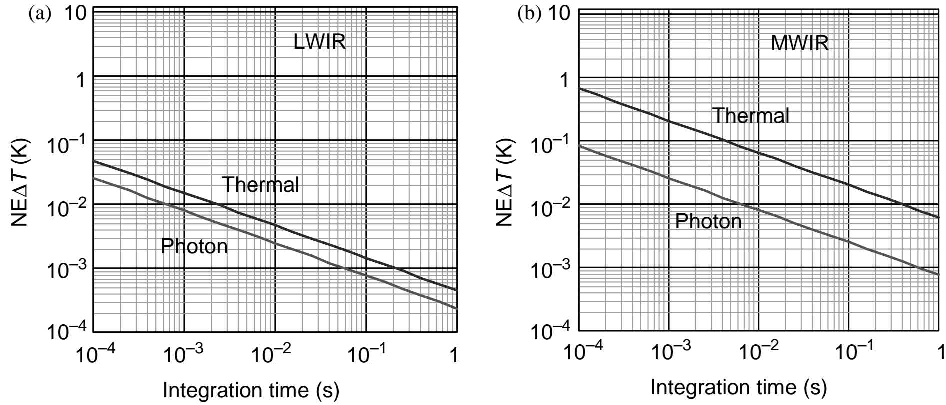


Figure 6.12 Theoretical NEΔT of uncooled photon and thermal detection at LWIR, and MWIR vs. integration time.

The fundamental limits of performance are seen to be essentially the same in the LWIR region of the spectrum, and the difference in Fig. 6.12 is probably due to the use of correlated double sampling in deducing Eq. (6.4). The photon detector is superior in the MWIR, due to the limitation on spectral bandwidth inherent in the device material. Real-world limitations on achieving the fundamental performance levels of Fig. 6.12 are severe for both thermal and photon detection.

For the thermal detector, the challenge is to reduce the thermal capacitance to values consistent with $\tau_{int} = C_{th} R_{rad}$. A practical limit for today's MEMS technology is ~500 Å, which for $R_{rad} \sim 4 \times 10^8$ K/W (for a 25-μm pixel) gives a minimum usable integration time of 25 ms.

For the photon detector, the challenge is not only to reduce the dark current such that $J_{dark} < J_{\Phi B}$ but also to store the charge accumulated during an integration time. The magnitude of both the dark current and charge storage issues is illustrated in Fig. 6.13. The S-R and Auger7 components of thermally generated dark current are

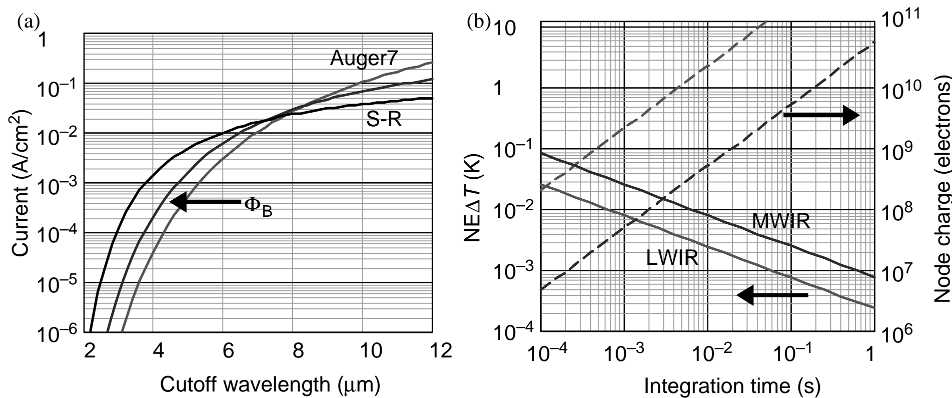


Figure 6.13 (a) Current density vs. cutoff wavelength for Auger7, state-of-the-art S-R, and background flux and (b) the required charge storage on the detector node vs. integration time for MWIR and LWIR.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

modeled for a 5×10^{14} cm^{-3} extrinsic doping concentration. The required storage capacity for a 25-μm LWIR pixel, in a 30-ms frame time, is seen to be $\sim 6 \times 10^{10}$ electrons, and for MWIR it is 1.4×10^9 electrons. These are both well beyond the current state of the art in effective well capacity for FPA ROIC technology, which is $\sim 2 \times 10^8$ carriers.

6.4 Uncooled Detection: Conclusions

Uncooled detection is currently the realm of the thermal detector, as typified by the MEMS bolometer FPA. If $1/f$ noise is not an issue, then the performance of LWIR thermal FPAs is limited ultimately by the thickness of the pedestal employed in the unit cell. For operation at relatively slow frame rates ($\sim$60 Hz), the thermal detector offers excellent performance and will probably never be surpassed by the photon detector. However, for LWIR operation at fast frame rates, and MWIR operation at any frame rate, the uncooled photon detector is preferred. The realization of the uncooled photon detector is dependent on the solution of two major issues: (1) improved material quality in the form of a significant reduction in S-R defects, such that the dark current is less than that due to the background flux, and (2) an improved ROIC capability to handle the storage of background flux generated charge.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 7

HgCdTe Electron Avalanche Photodiodes (EAPDs)

When electrons and holes traverse a semiconductor depletion region under large reverse bias they can acquire enough energy to impact ionize a secondary carrier. This effect, known as avalanche multiplication, has been widely investigated for the more commonplace semiconductors, such as Si, Ge, and GaAs. The modeling [87–89] of avalanche multiplication has been primarily phenomenological, and it is based on considerations of the electron distribution function obtained from equating the energy gained by the carrier from the applied electric field to the energy lost by interaction with optical phonons. The distribution function obtained in this way then has a fraction of carriers with sufficient energy to impact ionize electron–hole pairs. The probability of impact ionization is assumed to be a function of applied electric field only and in no way depends on the position or history of the carrier. The band structure of the specific semiconductor is essentially ignored.

There are a number of issues concerning the validity of applying standard APD theory to HgCdTe $n^+/n^-/p$ diodes at 77 K and perhaps even at higher temperatures.

1. Optical phonon scattering is well-defined, but not that strong, even at elevated temperatures;
2. Carrier densities are not that large, calling into question the use of randomized distribution functions;
3. Impact ionization is highly unlikely to be history independent; and
4. The values of physical parameters needed to form-fit standard APD models to theory tend not to be self-consistent.

This last point is quite important in that the three parameters employed for model fitting, namely, ionization threshold energy, optical phonon mean free path, and optical phonon energy, are quite well-known for HgCdTe.

It is tempting to take the opposite approach for HgCdTe, model the probability of impact ionization based on the specific band structure of the alloy, and then compare this ionization rate with the various well-documented scattering rates in HgCdTe. This approach has been utilized with significant success by Fischetti et al. [90] for Ge, Si, and GaAs, utilizing Monte Carlo calculations. We

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

will attempt to accomplish the same end here in HgCdTe, utilizing a mildly phenomenological model involving one adjustable parameter to replace the Monte Carlo calculations, and enable a correlation between electron energy and applied voltage.

For discussion purposes, it will first be beneficial to define the terms used to characterize APD performance. This will be done in Sec. 7.1, utilizing the generic concepts developed by McIntyre [91] to describe avalanche gain and excess noise factors in a diode exhibiting both electron and hole impact ionization.

7.1 McIntyre's Avalanche Photodiode Model

McIntyre's model assumes a reverse-biased P-N junction of width W and arbitrary electric-field profile. Electrons and holes traverse the junction depletion region with ionization probabilities of α and β per unit length, respectively. The ionization probabilities are assumed to be functions of the electric field only. The ramifications of this assumption are significant. The ionization probability of a carrier is a function of its energy and momentum. A carrier moving in a constant electric field experiences a whole range of momentum values defined by a momentum probability distribution for that value of electric field, F. Only carriers at the top end of the distribution are capable of an ionizing collision. The assumption that α or β is a function of F only, and not of the previous carrier history, implies that the momentum of an individual carrier will sample a large number of values throughout the complete momentum probability distribution within a distance that is small compared to the mean distance between ionizing events, given by E_I/qF, where E_I is the threshold energy for ionization. Thus, the implication is that the applied voltage drop across this distance needs to be $V \gg E_I/q$. In other words, the path of the carrier is far from ballistic, and the carrier suffers multiple collisions prior to achieving ionizing status.

A consideration of the electron current in the APD for electrons injected at $x = 0$ gives a gain

$$M_n = \frac{\exp[\int_0^W (\alpha - \beta)\mathrm{d}x']}{1 - \int_0^W \alpha \cdot \exp\left[\int_x^W (\alpha - \beta)\mathrm{d}x'\right]\mathrm{d}x}. \tag{7.1}$$

For α and β that are dependent only on F, and a uniform electric field such as in an ideal P-I-N junction, Eq. (7.1) gives

$$M_n = \frac{(1 - \beta/\alpha)\exp\left[\alpha W(1 - \beta/\alpha)\right]}{1 - (\beta/\alpha)\exp\left[\alpha W(1 - (\beta/\alpha)\right]}. \tag{7.2}$$

The ratio of ionization coefficients for electrons and holes, $\beta/\alpha = k$, varies for different semiconductors, and the dependence of avalanche gain, M_n, on k is shown in Fig. 7.1 for a particular set of parameters. In these phenomenological models,

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

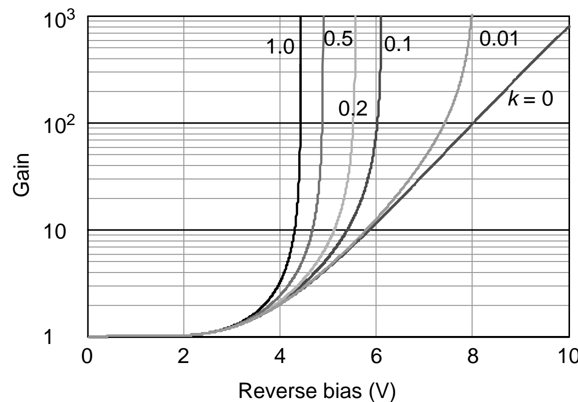


Figure 7.1 APD gain vs. voltage for various values of ionization coefficient ratio, k, assuming an electron ionization coefficient, $a = 3 \times 10^5 \exp(-1.5 \times 10^5/F)\,\mathrm{cm}^{-1}$; $W = 1\,\mu\mathrm{m}$.

the electron ionization coefficient, α, is assumed to depend only on electric field, with a form $\alpha = a_1 \exp(-a_2/F)$. For the particular case shown in Fig. 7.1, $\alpha = 3 \times 10^5 \exp(-1.5 \times 10^5/F)\,\mathrm{cm}^{-1}$, where $F = V/W$, V is the applied voltage, and W the junction width, assumed to be $1\,\mu\mathrm{m}$. It is apparent from Fig. 7.1 that for the case of $k = 0$, that is, when the ionization coefficient for holes is zero, then the dependence of avalanche gain on applied voltage is a pure exponential, and there is no avalanche breakdown exhibited in the gain–voltage relationship. However, for k-values that are nonzero, the gain–voltage characteristic exhibits a runaway, or avalanche breakdown, condition. This is due to the avalanche effect associated with the holes, and the denominator of Eq. (7.2) tends to zero, driving M_n to infinity.

McIntyre's format can also be utilized to model the dependence of excess noise in an APD on the ionization coefficient ratio, assuming a perfectly random generation of carriers. The result is normally quoted as the ratio of actual modeled noise in the APD to $2qI_\mathrm{inj}M^2$, where I_inj is the injected current, and M is the current gain. The excess noise factor is strongly dependent on where the current is generated. For pure electron injection into the high field region at $x = 0$, McIntyre's model gives for the excess noise factor for electrons

$$F_\mathrm{n} = M_\mathrm{n}\left[1 - (1 - k)\left(\frac{M_\mathrm{n} - 1}{M_\mathrm{n}}\right)^2\right]. \tag{7.3}$$

The dependence of F_n vs. gain for various k-values is shown in Fig. 7.2. It is apparent that the lowest value of excess noise factor that can be achieved for this perfectly random process is 2, for $k = 0$. For semiconductors with nonzero values of k, then significant amounts of excess noise will be observed as the gain is increased. For $k = 1$, the excess noise factor is for all intents and purposes given by M_n. For $k > 1$, the excess noise factor is even larger than M_n, but in this case the APD architecture should be modified to optimize hole injection, which is obviously favored.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

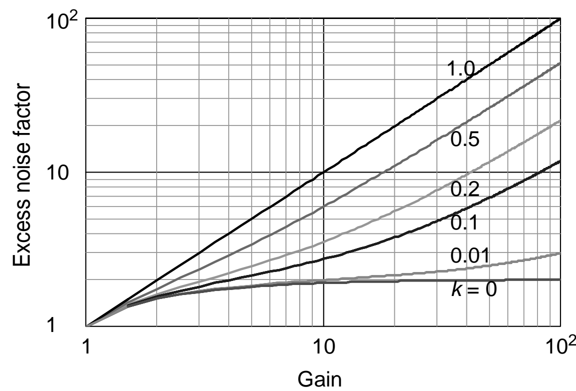


Figure 7.2 Excess noise factor vs. gain for various values of $k(=\beta v/\alpha)$.

The phenomenological model described above has been used for a number of years but has obvious shortcomings, not the least of which is the assumption that each ionizing event is completely random, dependent only on the value of applied electric field and independent of its history. McIntyre [92] has attempted to address this issue with his "dead-length" model, which assumes that a cold carrier must travel a finite distance before it acquires enough energy to impact ionize.

With this relatively simple backdrop we can now take a look at the very recent avalanche photodiode data that has been achieved in HgCdTe.

7.2 Physics of HgCdTe EAPDs

Exponential gain values well in excess of 1000 have been observed [93, 94] in a wide variety of $Hg_{1-x}Cd_xTe$ $n^+/n^-/p$ diodes fabricated at DRS, with excess noise factors approaching unity. This gain exhibits a total absence of avalanche breakdown. All of these characteristics are indicative of not only single-carrier multiplication but also a deterministic impact ionization process, that is, a history-dependent gain mechanism. Qualitative arguments can be made based on the band structure of $Hg_{1-x}Cd_xTe$ to eliminate the holes from avalanche considerations. The heavy hole mass is $\sim 0.55\, m_o$ across the entire HgCdTe composition range, resulting in low hole mobility values and a significant degree of optical phonon scattering. Thus, the heavy hole acquires energy from applied fields very inefficiently, and it readily loses what it does gain to optical phonons. The effective mass of the electron, on the other hand, is very small and is $\sim 7 \times 10^{-2} E_g$, where E_g is the bandgap in eV. Its mobility is high, particularly at low temperatures, and scattering by optical phonons is weak, which results in significant energy gain at even modest applied fields. The conduction band of HgCdTe is also devoid of any low-lying secondary minima, which allows for large electron energy excursions deep into the band and hence the high probability of impact ionization, with the generation of electron–hole pairs. The asymmetry of electron and hole masses also results in an ionization

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

threshold energy for electrons $\sim E_g$, and for holes $\sim 2E_g$, which in turn further exacerbates the difference between electron and heavy hole multiplication. Simple band theory would suggest that light holes might be expected to play a role in hole avalanche multiplication in HgCdTe, despite the large density of states difference compared to the heavy hole. However, band structure models of HgCdTe indicate that at the higher energies involved in avalanche multiplication, the curvature of the light hole valence band is distorted, possibly due to an interaction with the split-off spin–orbit band, and is consistent with a large effective mass, somewhat similar to the heavy hole. It is also interesting to note that evidence for light hole transport in p-type material even at low hole energies is not strong. The measured mobility values observed in magnetic field dependent transport measurements at low temperatures, discussed in Sec. 5.1.3.3.b, indicate that some lifting of the light and heavy hole valence band degeneracy is required, together with a larger average light hole effective mass (approximately, twice the conduction band effective mass), to explain the data. Thus, the occurrence of electron-only avalanche in HgCdTe, in the appropriate diode geometry, is perhaps not surprising. These qualitative arguments concerning electron-only avalanche can be put on a firmer footing by a quantitative examination of the relevant scattering rates in HgCdTe for high-energy electrons and holes, together with a consideration of impact ionization rate theory for HgCdTe.

7.2.1 High-energy scattering rates

An indication of the relative scattering rates for high-energy electrons and holes in HgCdTe is given by the measured mobility of low-energy carriers in n-type and p-type material. This was discussed earlier in Sec. 5.1.3.3 and is shown again in Figs. 7.3(a) and (b), for a MWIR composition. Various contributions to the scattering rates were considered, including optical phonons, ionized impurities, neutral impurities, and alloy disorder. Ionized impurities are seen to be relatively unimportant at these doping concentrations, as is neutral impurity scattering. The two main scattering mechanisms are due to interactions with optical phonons, which are inelastic, and alloy disorder, which is elastic.

Optical phonon scattering of high-energy electrons of energy E is given by Eq. (5.8) to be

$$\tau_{op} = [2m^* E]^{1/2}/[qF_o(2N_o + 1)], \tag{7.4}$$

where $F_o = [\varepsilon_{hf}^{-1} - \varepsilon^{-1}]m^* qk\theta\pi/\varepsilon_o h^2$, and ε_{hf} and ε are the high- and low-frequency dielectric constants of the material, $k\theta$ is the optical phonon energy, and $m^* = m_o^*[1 + (E/E_g)]$ is the energy-dependent effective mass. $N_o = [\exp(\theta/T_o) - 1]^{-1}$.

Alloy scattering at low energy, assuming weak scattering that is also isotropic and taking place on a constant energy surface, is given by

$$\tau_{alloy} = h/[4\pi^2 x(1 - x)\Delta E_c^2 \rho(E)], \tag{7.5}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

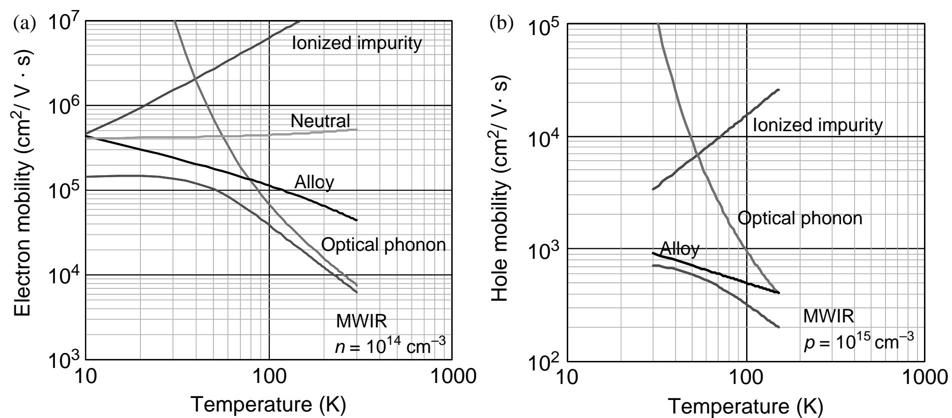


Figure 7.3 (a) Electron mobility and (b) hole mobility vs. temperature for MWIR HgCdTe with low doping concentrations.

where ΔE_c is the difference between the s atomic term values of the Hg and Cd atoms. For low-energy and spherical, parabolic energy surfaces, $\rho(E)$ is $4\pi[2m^*/h^2]^{3/2}E^{1/2}$. The mobility of electrons in Fig. 7.3(a) is consistent with a value for τ_{alloy} as given by Eq. (7.5) with $\Delta E_c = 1.5\,\text{eV}$. Higher electron energies will be complicated by the fact that scattering will no longer be weak, and will also be anisotropic, by virtue of the fact that the energy surfaces are no longer spherical for this elastic interaction.

The holes are also dominated by the same two scattering mechanisms, with a ΔE_v required for alloy scattering of $\sim$0.3 eV, in reasonable agreement with the proposed valence band offset [95] for HgTe and CdTe. The mobility of holes is much less than that of electrons, due entirely to their effective mass of $0.55m_o$. This difference in mobility and effective mass enables the electron to gain energy from an applied electric field at a far greater rate than the hole. This is illustrated for MWIR HgCdTe in Fig. 7.4, where we have modeled the simple ballistic gain of

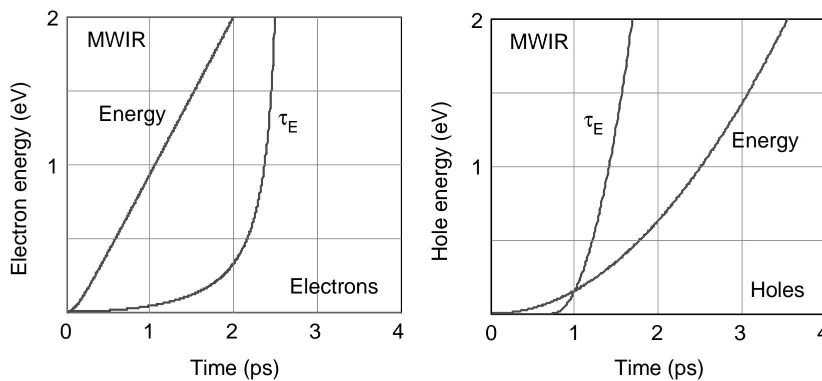


Figure 7.4 Energy gained from applied field of 10^4 V/cm by electrons and holes in MWIR HgCdTe vs. time, relative to the optical phonon lifetime as a function of energy at 77 K.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

energy by Kane-like conduction band electrons and parabolic valence band holes in a uniform electric field of 10^4 V/cm. This rate of energy gain for electrons and holes is compared to the time taken to interact with an optical phonon at 77 K as a function of energy.

The electron is seen to penetrate deep into the conduction band prior to interacting with an optical phonon, whereas the hole scarcely acquires an optical phonon energy (~0.017 eV). This simple model predicts hot electrons that are easily capable of impact ionization (energy required $>E_g$) and cold holes that are not (energy required $>2E_g$). It must be pointed out that alloy scattering, although elastic, can also theoretically be responsible for electron energy dispersion. However, the experimental evidence discussed later suggests that this scattering mechanism does not play a strong role in HgCdTe APDs.

7.2.2 Electron impact ionization rate in HgCdTe

The general band structure of $x = 0.3$ HgCdTe [95a, 95b] is shown in Fig. 7.5(a), with an expanded view of the E–k region around the Γ point in k-space in Fig. 7.5(b), assuming spherical symmetry. In the region around $k = 0$, the band is parabolic with mass m_o^*, whereas at higher energies in the conduction band, the effective mass becomes energy dependent and is given by the Kane expression $m^* = m_o^*[1 + E/E_g]$, resulting in a linear E–k relationship. The quantitative expression for electron energy in the conduction band is given by

$$E = [-E_g + E_g[1 + (h^2k^2/2\pi^2 E_g m_o^*)]^{1/2}]/2. \tag{7.6}$$

Thus, at high energies in the band, $E \approx E_g[h^2/(8\pi^2 E_g m_o^*)]^{1/2}k = E_g[k/k_g]$, where $k_g = [8m_o^* E_g \pi^2/h^2]^{1/2}$.

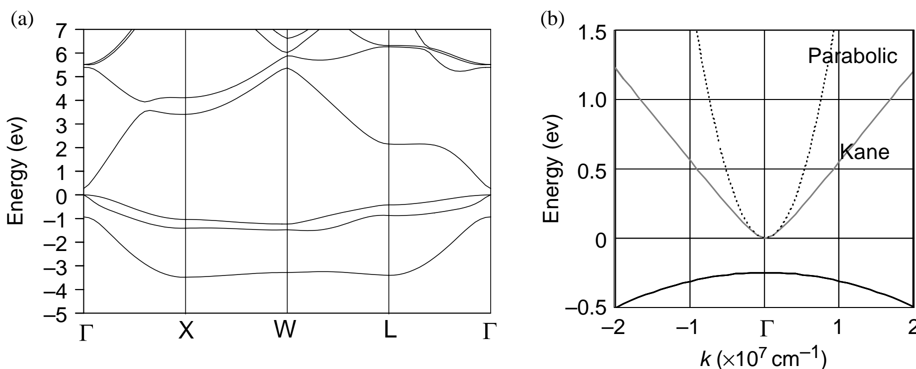


Figure 7.5 (a) Band structure of $x = 0.3$ HgCdTe and (b) an expanded view of the Γ point in k-space assuming spherical symmetry.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The probability of impact ionization by a conduction band electron at k_1, to produce a heavy hole at k_2 and an electron at $k_{2'}$, is given by

$$P = (4\pi^2/h)^* \sum_{\mathrm{Ec}}^{*} \sum_{\mathrm{Ev}} [|H|^2 \rho(E_c k_{1'}) \rho(E_c k_{2'}) \rho(E_v k_2)], \tag{7.7}$$

where $k_{1'}$ is the final state of the ionizing electron, $|H|^2$ represents the matrix element associated with the electron–electron interaction, and $\rho(E)$ is the density of states of the various product carriers. The summations in Eq. (7.7) are typically limited to transitions that conserve both energy and momentum, except where umklappe processes are involved in the electron–electron interactions, such as in indirect gap semiconductors like Si and Ge, or semiconductors with a low ($\ll E_g$) secondary conduction band minimum, such as GaAs. The matrix element is that of the screened Coulomb interaction between electrons and is similar to that used in Auger recombination theory [96], given by

$$|H|^2 = [Vq^4/\varepsilon^2\varepsilon_o^2][F_1F_2]^2/[\lambda^2 + g^2]^2, \tag{7.8}$$

where V is the volume of a unit cell, ε the dielectric constant, F_1F_2 are the overlap integrals of Auger theory, λ is the screening parameter of the Coulomb interaction, and $g = \Delta k = k_1 - k_{1'} = k_2 - k_{2'}$.

The conservation of energy and momentum requirements, coupled with the format of the interaction matrix element, greatly limit the volume of k-space available for impact ionization, as can be seen from a consideration of the simplified HgCdTe energy bands given by Eq. (7.6), shown in Fig. 7.6. The heavy hole valence band is approximated by a parabolic band with essentially infinite mass, and the conduction band for high-energy electrons by the linear E–k format of Kane, from $-E_g/2$ up to some energy value at which it transitions into a secondary heavy mass conduction band. The band structure for HgCdTe suggests that this transition occurs at least $\sim$2 eV into the band, depending on orientation. ΔE represents the energy lost by the hot electron when it impact ionizes. The presence of the Δk^4 term in the denominator of $|H|^2$ heavily weights transitions for any ΔE greater

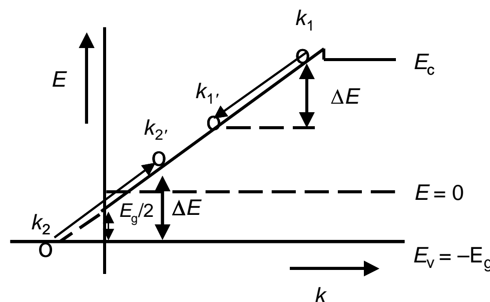


Figure 7.6 Simplified energy band structure around the Γ point in HgCdTe.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

than the threshold energy for ionization, E_g, to Δk values along the direction of the applied field, as shown in Fig. 7.6. This means that the density of states for the hole at k_2 is essentially a delta function in the summation over E_v space in Eq. (7.7). This density of states is approximately given by the linear k-value required at $E = -E_g$ by Eq. (7.6), incorporated into the correct parabolic expression for the heavy hole band. Thus, $\rho(E_v k_2) = [E_g m_o^* m_v^2/2]^{1/2} 8\pi/h^3$, where m_v is the heavy hole mass. The conduction band density of states, for spherical symmetry, is given by $\rho(k) = (k^2/\pi^2)/[dE/dk] = [(2E_n + 1)^2 - 1]^{1/2}[2E_n + 1](k_g)^3/E_g\pi^2$, with $E_n = E/E_g$, where the value of E for $k = k_{1'}$ is $E - \Delta E$, and for $k = k_{2'}$ it is $\Delta E - E_g$. The summation over the conduction band in Eq. (7.7) for any $\Delta E(>E_g)$ is replaced by an integral, and the impact ionization rate is given by

$$P(E_n) = \frac{3.2 \times 10^7 V q^4 m_o^* [F_1 F_2]^2 \rho(E_v k_2)}{h^3 \varepsilon^2 \varepsilon_o^2} \int_1^{E_n} \frac{[[2(E_n - x) + 1]^2 - 1]^{1/2}[2E_n - x) + 1][2x - 1][(2x - 1)^2 - 1]^{1/2}}{[(\lambda/k_g)^2 + x^2]^2} dx, \tag{7.9}$$

where $x = \Delta E/E_g$, and we have utilized $g = \Delta k \approx k_g \Delta E/E_g$ at high energies. The function under the integral is shown in Fig. 7.7 for various primary electron energies, as a function of the energy lost ($>E_g$) during impact ionization. $F_1 F_2$ is assumed to be constant, and the screening parameter negligible. The function is independent of bandgap. For large E_n, the impact ionization probability varies as $\sim E_n^2$.

The optimum electron transition is seen to be $\sim 1.3E_g$, regardless of initial electron energy, and this is due to the strong initial increase in the density of available states for $x > 1$, modified by the denominator of the screened electron Coulomb matrix element term, which varies as $1/x^4$. The optimum transition in this regime

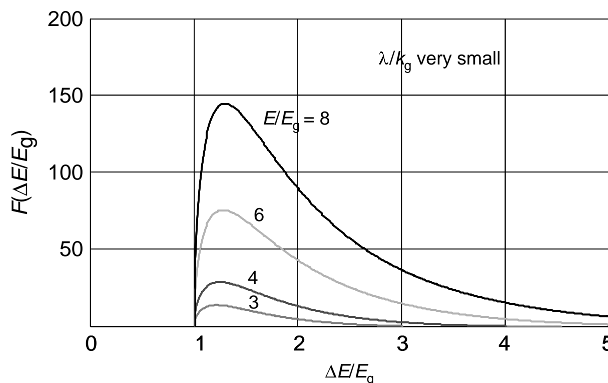


Figure 7.7 The function in the integrand of Eq. (7.9) for impact ionization.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

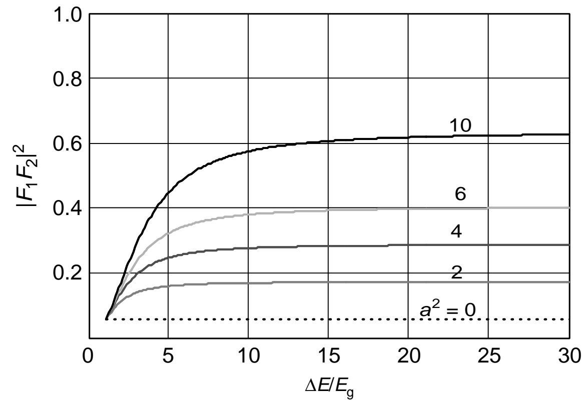


Figure 7.8 Overlap integral dependence on $\Delta k(=k_g\Delta E/E_g)$ for various values of a^2.

thus generates an essentially cold electron, with energy $\sim 0.3E_g$ above the conduction band edge at $k = 0$, together with a warm electron that remains deep in the conduction band. The assumption of a constant overlap integral $|F_1F_2|^2$ equal to that utilized in low-energy Auger theory is appropriate to first order, but a better approximation would take into account that $|F_1F_2|^2$ varies as Δk^2 for small Δk values. Previous modeling [97, 98] suggests that, for a small fixed value of hole momentum, k_2, we may take the overlap integral to vary as $|F_1F_2|_o^2(a^2+1)\Delta k^2/[a^2k_g^2+\Delta k^2]$, where $|F_1F_2|_o$ represents the value of the overlap integral used in Auger theory. The dependence of the overlap integral on $\Delta k/k_g$ is shown in Fig. 7.8, for various values of a^2. Comparing the resulting $|F_1F_2|^2$ with simple overlap integrals calculated from a one-dimensional Kronig–Penney model by Beattie and Landsberg [95], a value of $a^2 = 10$ appears appropriate. Inclusion of this k-dependence of the overlap integral modifies the integrand of Eq. (7.9) and its dependence on $\Delta E/E_g$. This is

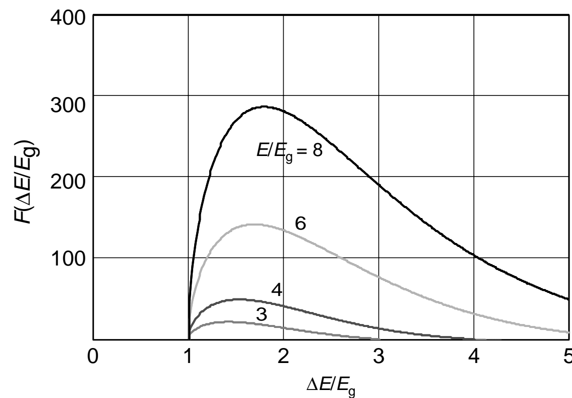


Figure 7.9 The modified integrand of Eq. (7.9) as a function of energy loss at impact ionization, for various initial electron energies.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

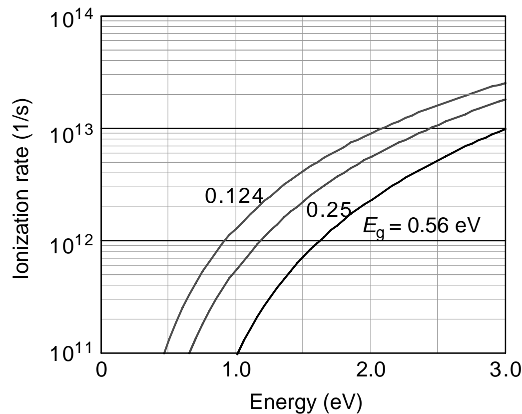


Figure 7.10 Conduction band electron impact ionization rate for SWIR, MWIR, and LWIR HgCdTe as a function of electron energy.

shown in Fig. 7.9 for various values of initial electron energy. It is apparent that the modified overlap integral has two effects. First, the values of the integrand are increased, and second, the most probable electron transition for impact ionization is shifted to $\sim 2E_g$, and the spectrum of transitions is broadened considerably.

The impact ionization rates given by the modified Eq. (7.9) for SWIR, MWIR, and LWIR HgCdTe are shown in Fig. 7.10 as a function of primary electron energy. Meaningful ionization rates $\sim 10^{12}$/s (providing ionization coefficients $\sim 10^4$/cm for a saturation velocity of 10^8 cm/s) are seen to be achieved at relatively low electron energies of ~ 1 eV.

It should be pointed out here that HgCdTe APD behavior has been the subject of a large body of research [99, 100] for a number of years, for possible optical communications applications. However, that research has been confined to architectures that, strangely enough, rely on hole multiplication in this short-wavelength region of the IR spectrum. The reason for this is again associated with the somewhat peculiar band structure of HgCdTe. The spin–orbit split-off band is $E_{so} \approx 0.9$ eV from the heavy hole valence band edge at $k = 0$, corresponding to a spectral response ~ 1.35 μm, which lies in the currently favored bands for optical communication. The accepted mode of operation is with a HgCdTe composition where $E_g \approx 0.9$ eV, such that the bandgap is resonant with E_{so}, as shown in Fig. 7.11. The probability of carrier ionization is given by Eqs. (7.7) and (7.8), but now $g \approx 0$, as the transitions are essentially vertical in k-value. Even though the collective density of states of the product carriers (two heavy holes and one electron) is low, as all carriers are at band edges, the transitions are described as "resonant," and theoretically one might have cause to expect a large hole ionization coefficient relative to the electron, and hence low noise gain. The problem of course is that one needs to create a hot hole in the split-off band to initiate the interaction, which will require in turn a hot heavy hole energy E_g and sufficiently strong optical phonon interactions to allow the hot heavy hole to transition to the split-off band. This probably accounts for the magnitude of the voltages required to initiate this APD architecture, being ~ 75 to 100 V.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

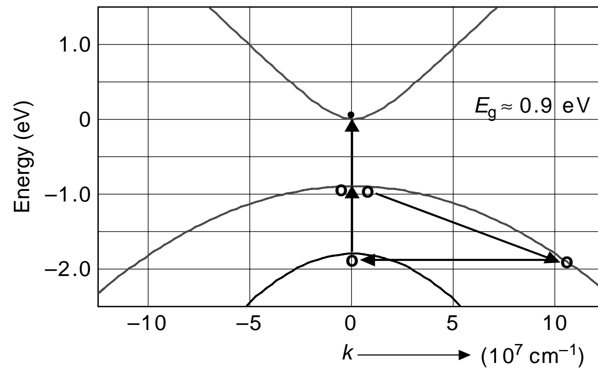


Figure 7.11 The "resonant" transition from the split-off valence band to the heavy hole valence band, originated by a hot heavy hole.

It is interesting to note that, provided the conditions exist to generate hot holes with energy ~ 0.9 eV, then transitions from the split-off band to the heavy hole valence band can occur for all compositions and bandgaps of HgCdTe, with the corresponding generation of a hot electron by the appropriate vertical $\Delta k = 0$ transition from the heavy hole band to the conduction band, with a larger density of states than the so-called resonant composition.

The above discussions indicate that electron impact ionization will be observed in HgCdTe for the correct geometry long before the advent of hole ionization, thus offering capability of low noise, large, exponential gain, without avalanche breakdown. However, the reverse is not true. Hole ionization may be geometrically favored with the correct architecture, but the advent of hole impact ionization must be accompanied very shortly thereafter by electron impact ionization, and hence excess noise, and eventually avalanche breakdown.

Before leaving the subject of the dependence of impact ionization coefficients on semiconductor band structure, it is worth re-visiting the case of the indirect bandgap semiconductor and the direct bandgap semiconductor with a very low secondary conduction band minimum. The ionization coefficient is again determined by Eqs. (7.7) and (7.8); however, in this case the electron distribution in the conduction band is not only dominated by optical phonon interactions, but those interactions are in turn dominated by umklappe processes associated with transitions between subsidiary minima involving large momentum changes of approximately a reciprocal lattice vector $2\pi/a_o$. Fischetti et al. [90] have analysed this situation and find that, in these circumstances, the "random-k approximation" [101] can be used. The randomization of momentum via umklappe processes is so strong that the many momentum-conserving delta function-like transitions can be replaced by a uniform integration of the collective density of states of the product carriers over the entire k-space of the Brillouin zone, assuming an essentially constant matrix element for the Coulomb interaction. In these circumstances, the ionization coefficient is dominated by the collective density of states of the product carriers, and it is reflected in a much stronger dependence of ionization coefficient on energy than for the HgCdTe case.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

7.3 Empirical Model for Electron Avalanche Gain in HgCdTe

Empirically, it has been determined by Beck [102] that the EAPD gain, *M*, in HgCdTe at 77 K exhibits a dependence on applied voltage given approximately by

$$M = 2^{\frac{(V - V_{th})}{V_{th}/2}} + 1, \tag{7.10}$$

with $V_{th} \approx 6.8E_g$ for all compositions from $0.2 < x < 0.5$. This empirical dependence of gain on applied voltage is shown in Fig. 7.12, together with experimental data points taken at DRS and from the published literature [103]. The experimental data have an uncertainty of $\sim$2, due to uncertainties in the optically active p-volume of the device.

Equation (7.10) suggests the following phenomenological "dead voltage" model for EAPD gain in HgCdTe. A cold electron is introduced into the depletion region of the APD biased with an applied voltage *V*. The electron will gain energy to some threshold energy E_{th}, directly related to the threshold voltage, V_{th}, at which point it will impact ionize, generating two product warm electrons and a cold hole. For the time being, we will assume no electron scattering so that $E_{th} = V_{th}$. For electron energy dispersion $E_{th} = \alpha V_{th}$, and this case is treated later. The threshold energy should be directly related to the impact ionization rate given in Fig. 7.10 and the characteristics of the diode depletion region. The energy spectrum of the product electrons is determined by energy and momentum considerations and is given by the function $F(\Delta E/E_g)$ of Fig. 7.9. The product warm electrons will then be capable of subsequent impact ionization after falling through the necessary "dead voltage," when they again attain the required threshold energy, V_{th}. For reasons discussed earlier, the cold hole can be ignored at the values of applied voltage under consideration here. The gain process is illustrated in Fig. 7.13 for an arbitrary situation in which the energy dispersion is assumed to be zero, and the electron gains an energy

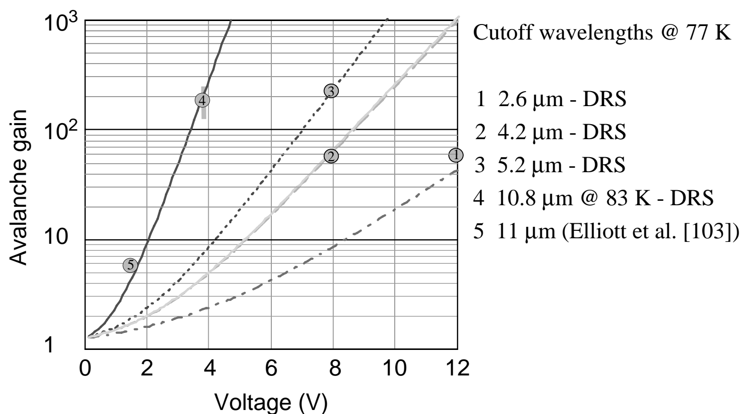


Figure 7.12 Empirical plot of gain vs. applied voltage for various cutoff wavelengths at 77 K with experimental data points taken at $\sim$77 K.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

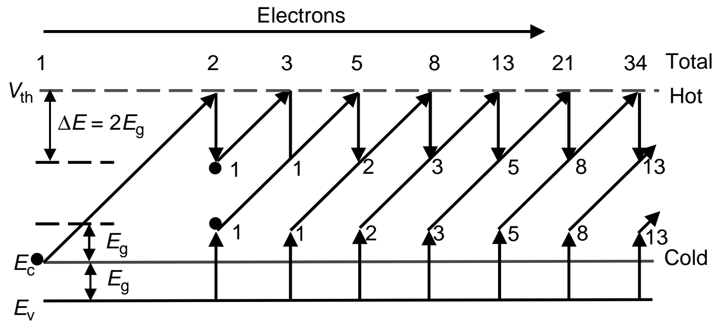


Figure 7.13 Avalanche gain due to two unequally warm electrons.

equal to the potential drop associated with the applied voltage. For this particular example, V_{th} is assumed to be $5E_g$, and the product electrons are assumed to have energies of E_g and V_{th}-$2E_g$, relative to E_c, as required by energy and momentum conservation. The gain is obviously determined by the number of dead voltage increments for each warm electron contained in the applied voltage, V, and can easily be calculated for any combination of V_{th} and ΔE, by the simple use of an Excel spreadsheet. It is interesting to note that the modeled avalanche gain along different energy channels (that is, different values of ΔE for the same threshold energy) is relatively independent of ΔE, except for ΔE values approaching E_g, for which the conduction band density of states, and hence probability of occurrence, is very low. This result is perhaps not surprising, as for two unequally warm electrons the increased probability of the warmer electron impact ionizing (i.e., to reach V_{th}) is offset by the decreased probability of the less warm electron achieving V_{th}. Two equally warm electrons thus represent a somewhat optimum case. This effect is illustrated in Fig. 7.14 for MWIR HgCdTe ($E_g = 0.25$ eV) with an assumed threshold energy of $\sim 7E_g$, and various ΔE values. This relative independence of

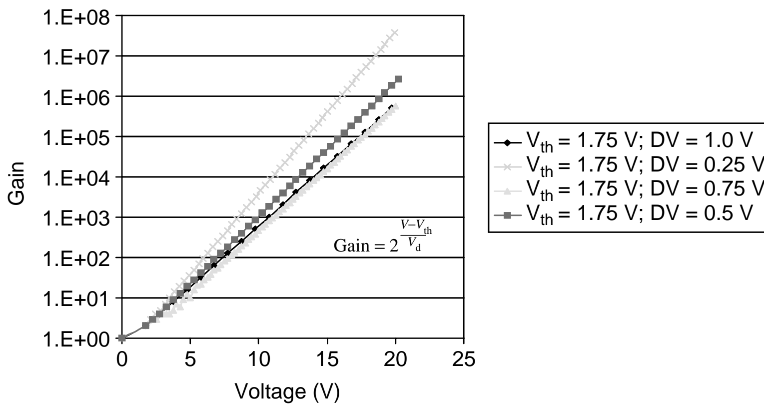


Figure 7.14 Avalanche gain modeling for various warm electron combinations with a threshold energy of $7E_g$, where $E_g = 0.25$ eV.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

avalanche gain across the electron distribution given by $F(\Delta E/E_g)$ suggests a convenient phenomenological model for APD gain that can be expressed in simple analytical terms.

The gain experienced by two equally warm electrons, as given by Fig. 7.14, will be characteristic of the electron distribution as a whole, to within a factor ~3. Thus, for a threshold energy, V_{th}, the associated dead voltage, V_d, for these equally warm electrons, by geometrical arguments, will be $[V_{th} + E_g]/2$, and the APD gain will be given approximately by $2^{(V-V_{th})/V_d}$.

Making allowances for the slightly higher gain associated with electron energy values between $E_g < \Delta E < 2E_g$, we can say that the average electron APD gain of a HgCdTe diode is given approximately by

$$\text{Gain} = 2^{\frac{V-V_{th}}{V_d}}, \tag{7.11}$$

where $V_d = V_{th}/2$; that is, the factor E_g has been dropped from the dead voltage to compensate for the slightly larger gain expected for electrons with low ΔE. The format of Eq. (7.11) is seen to agree exactly with the empirical dependence of the experimental data noted by Beck et al. at 77 K, and defined in Eq. (7.10), provided $V_{th} \approx 6.8E_g$.

The threshold energy should be related directly to the ionization probability modeled earlier. The instantaneous ionization coefficient at any energy E is given by $P(E_n)/\upsilon_{sat}$, where $P(E_n)$ is given by Eq. (7.9), and υ_{sat} is the electron saturation velocity. The electron saturation velocity in a Kane-like conduction band is given by

$$\upsilon_{sat} = \sqrt{\frac{2E}{m_o^*[1 + E/E_g]}}, \tag{7.12}$$

where m_o^* is the effective mass at $k = 0$ in the conduction band, $m_o^* \approx 7 \times 10^{-2} E_g m_o$ over a large composition range in HgCdTe, where m_o is the free electron mass. Thus, at high electron energies, where $E > E_g$, Eq. (7.12) gives $\upsilon_{sat} = [2E_g/m_o^*]^{1/2} \approx 2 \times 10^8$ cm/s, regardless of cutoff wavelength. It should be remembered here that, although the electron velocity is saturated, the electron still gains energy from the applied electric field due to the fact that its mass is increasing.

The probability of impact ionization over a distance y is given by

$$Q(E_n, y) = \int_0^y \frac{P(E_n, y)}{\upsilon_{sat}} \mathrm{d}y. \tag{7.13}$$

Assuming a fixed depletion width W, which is commonplace in many APD geometries, a uniform electric field $F = V/W$, and an electron energy equal to that gained from the applied voltage V, then at a distance y from the p-side, the electron energy is given by

$$E = \Phi = (V + E_g)y/W. \tag{7.14}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

After some manipulation, the integral in Eq. (7.13) can be transformed into one of energy, namely

$$Q(E_n) = \int_1^{E_n} \frac{P(E_n) \cdot W}{v_{sat} \cdot (E_n + 1)} dE_n. \tag{7.15}$$

The threshold energy of the EAPD can then be defined as the electron energy at which Eq. (7.15) is unity. The depletion width is then just sufficient to provide a probability of unity for impact ionization under the applied bias voltage, with the production of two warm electrons and a cold hole. It is apparent from Eq. (7.13) that V_{th} is determined by the electron energy gained from the applied voltage, but it is not explicitly dependent on the magnitude of the applied voltage. Thus, the initial impact ionization will occur after the electron has acquired the necessary threshold energy, regardless of the magnitude of the applied voltage. Any subsequent impact ionization events will also depend only on the electron energy.

The implication of Eqs. (7.14) and (7.15) can be appreciated by a quantitative comparison with 77 K data for a group of MWIR APDs, with a cutoff wavelength of 4.4 μm at 77 K, which is shown in Fig. 7.15(a). The threshold voltage is seen to be $\sim$2 V (voltage at which gain is $\sim$2), and subsequently the gain increases by a factor of 2 for every increment of 1V. A model of diode potential versus distance for various applied voltages, for a doping concentration 10^{14} cm^{-3}, is illustrated in Fig. 7.15(b). The chosen geometry is cylindrical HDVIP as used by Beck et al. The central via with its associated n$^+$ region has a diameter of 7 μm, with a finite n$^-$ region of 3.5 μm width bounded by a p-region of 4×10^{16} cm^{-3} doping. A uniform electric field is seen to be a good approximation. At a bias voltage of 2 V, the electron can achieve the required threshold energy to impact ionize. At a bias of 4 V, the initial electron achieves threshold energy in a much shorter initial distance and generates two equally warm electrons. These two electrons then fall through the following dead voltage of $\sim$1 V, achieving enough energy from the

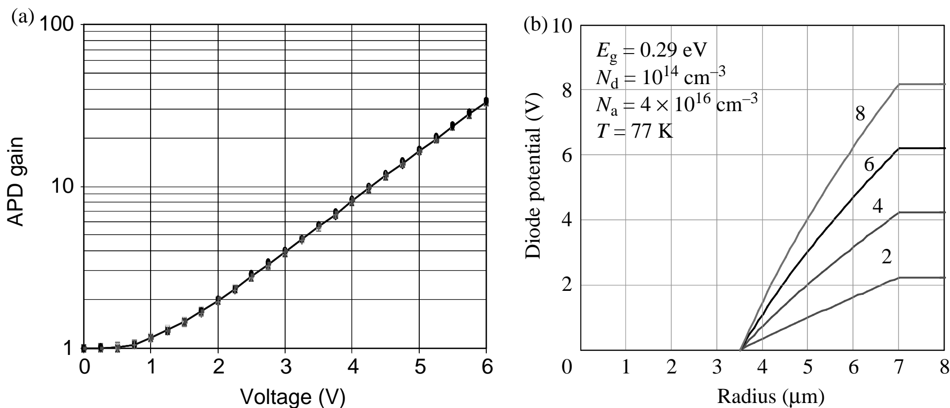


Figure 7.15 (a) Measured gain vs. voltage for MWIR HgCdTe APDs at 77 K, and (b) cylindrical diode potential vs. radius for various applied voltages, with $n = 10^{14}$cm^{-3}.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

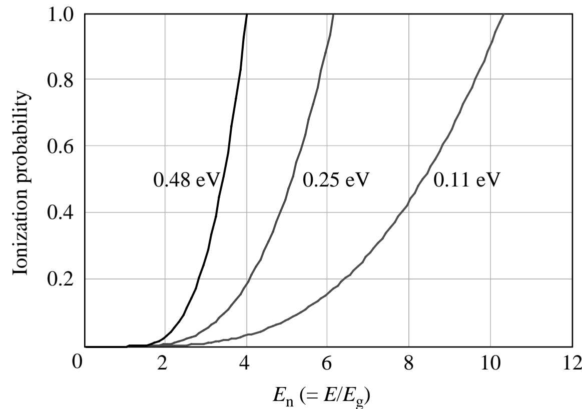


Figure 7.16 Modeled threshold energies for $E_g = 0.11$ eV, 0.25 eV, and 0.48 eV using Eq. (7.15) with $W = 3.5$ μm.

diode potential to impact ionize again, resulting in four equally warm electrons. These electrons again fall through a further dead voltage ~1 V before exiting the depletion region, acquiring just enough energy to impact ionize for a third time, resulting in an overall gain of 8 at this applied voltage, and so on for higher applied voltages.

In this picture, the "dead voltage" defined earlier is determined to be the energy lost by the hot electron to produce two approximately equally warm electrons, namely $\sim V_{th}/2$, and the threshold energy V_{th} should be given by Eq. (7.15). Threshold energies predicted by Eq. (7.15), for a finite depletion width of 3.5 μm, are shown in Fig. 7.16 for the three different bandgaps that approximate the experimental values of Fig. 7.12.

It is immediately apparent from Fig. 7.16 that

- The calculated threshold energy for $E_g = 0.11$ eV is $\sim 10E_g$, far in excess of Beck's empirical value for this bandgap of $V_{th} \approx 6.8E_g$, and
- The threshold energy is not linearly dependent on bandgap, as suggested by Beck's data.

Two issues are relevant to this comparison of model with experiment, namely the approximations that were used to estimate the electron impact ionization rate and the non-consideration of electron scattering mechanisms, which will lead to an electron energy that is given by $E = \alpha(E)V$, where $\alpha(E) < 1$.

The absolute value of the impact ionization probability, which is closely related to the electronic overlap integrals of Auger theory, can be normalized to experimental values of the threshold voltage provided the measurement in question is sufficiently well defined. In this regard, it is imperative to ascertain the relationship between electron energy and applied bias voltage, and to this end it is relevant to consider the predictions of ballistic electron theory. For an electron traveling in a uniform electric field F, given by V/W, with a saturation velocity υ_{sat}, the energy

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

gained before a momentum-destroying collision occurs at a time $\tau(E)$ is given by

$$\begin{aligned} E &= \tau(E) v_{\mathrm{sat}} V / W \\ &= [\lambda_{\mathrm{e}}(E)/W] V, \end{aligned} \tag{7.16}$$

where $\tau(E)$ is by definition the electron momentum scattering lifetime, and λ_e is the associated electron mean free path. We thus have a relationship between electron energy (in eV) and applied voltage given by

$$E = [\lambda_{\mathrm{e}}(E)/W] V = \alpha(E) V. \tag{7.17}$$

The assumption of a uniform electric field is not correct for a standard abrupt junction, but it is approximately true for most APD geometries in which the extent of the depletion region is limited by heterogeneous doping profiles. This expression for $\alpha(E)$ makes perfect physical sense in that it says that a voltage V applied across the depletion region W allows the electron to acquire only an energy E within the allotted mean free path $\lambda_e(E)$ prior to a momentum-destroying collision.

The earlier discussion of electron mobility in HgCdTe indicated that the dominant scattering mechanisms are due to optical phonons and alloy disorder. It is of interest to consider their contributions to $\lambda_e(E)$. Stratton's theory of polar phonon hot electron effects indicates that the optical phonon relaxation time for hot electrons is given by Eq. (7.4) to be $\tau(E)_{\mathrm{op}} = [2m_o^*(1 + (E/E_g))E]^{1/2}/qF_o^*[1 + (E/E_g)](2N_o + 1)$, where $F_o^* = [\varepsilon_{\mathrm{hf}}^{-1} - \varepsilon^{-1}]m_o^* qk\theta\pi/\varepsilon_o h^2$, $N_o = [\exp(\theta/T_o) - 1]^{-1}$, and $k\theta$ is the optical phonon energy. Thus, at high electron energies $\tau(E)_{\mathrm{op}}$ is independent of energy as illustrated in Fig. 7.4(a) and given by $\tau_{\mathrm{op}} = [2m_o^* E_g]^{1/2}/[qF_o^*(2N_o + 1)]$. Utilizing Eq. (7.12) for v_{sat}, we have a dispersion coefficient $\alpha(E)$ for optical phonon scattering given by

$$\begin{aligned} \alpha(E) &= \frac{2E_g}{F_o^*(2N_o + 1)W}, \\ &= \frac{1.44 \times 10^{-3}}{W\left[\dfrac{1}{\varepsilon_\infty} - \dfrac{1}{\varepsilon}\right]\theta\,(2N_o + 1)}. \end{aligned} \tag{7.18}$$

It is apparent that for optical phonon scattering, $\alpha(E)$ is independent of both electron energy and HgCdTe bandgap and depends only on the dielectric parameters of the material. The lack of energy dependence of α is consistent with Beck's empirical observation that the dead voltage is $\sim V_{\mathrm{th}}/2$, namely one-half of the threshold voltage, for two equally warm electrons. The modeled variation of $\alpha(E)$ with HgCdTe cutoff wavelength is shown in Fig. 7.17 for various diode depletion widths. For $\alpha(E) > 1$, the electron energy is limited to a value equal to the associated potential drop. As a matter of interest, the value of λ_e for MWIR HgCdTe, with $\lambda_c = 5\,\mu$m, due to optical phonon scattering at 77 K is $\sim 2.33\,\mu$m. For depletion widths $W < \lambda_e$ then optical phonon scattering is irrelevant.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

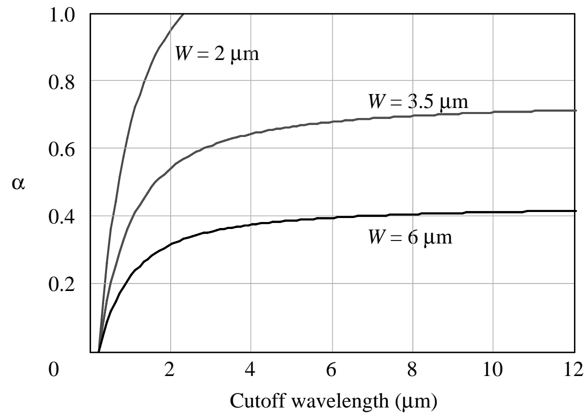


Figure 7.17 The energy–voltage dispersion coefficient, α vs. cutoff wavelength at 77 K for various fixed depletion widths.

Alloy disorder scattering for hot electrons is not well modeled across the HgCdTe alloy range at this time. The normal low-energy assumption of weak scattering in a spherically symmetric conduction band environment is certainly not valid at high electron energies. It is also not clear how to treat the alloy disorder parameter, ΔE_c of Eq. (7.5), deep into the band, where band structure calculations [104, 105] indicate that the conduction band energy is relatively composition independent. ΔE_c may well be at least an order of magnitude lower than required to explain low-energy electron transport. The density-of-states factor in the alloy scattering lifetime will lead to a strong dependence of lifetime on energy [106], which in turn is not consistent with the simple model of $E = \alpha\Phi$.

If we assume that the dispersion coefficient $\alpha(E)$ is limited by optical phonon scattering as given by Eq. (7.8), and assume a diode spread $W \approx 6\,\mu$m, which is reasonable for standard HDVIP diodes formed in material doped with 10^{16}cm^{-3} metal vacancies and 4×10^{16} cm^{-3} Cu, then $\alpha \approx 0.38$ for the $\lambda_c = 4.4\,\mu$m data shown in Fig. 7.15(a). The measured threshold voltage of 2 V is thus equivalent to an electron energy of 0.76 eV. Normalization of Eq. (7.15) then requires a factor $\sim$10 to give a threshold energy of 0.76 eV for this cutoff wavelength. This normalization factor is expected to be relatively independent of cutoff wavelength, diode geometry, and temperature, and so it can be used in general HgCdTe APD modeling.

Utilizing this empirical value for the impact ionization coefficient, we are now in a position to compare modeled threshold voltages for different cutoff wavelengths at 77 K with Beck's experimental data, assuming a fixed finite diode spread of $\sim$6 μm, and utilizing the optical phonon limited energy dispersion coefficient defined above. The predicted threshold voltages normalized to the APD bandgap are shown in Fig. 7.18. The threshold voltage values for the 0.28 eV and 0.48 eV bandgaps are seen to be in agreement with Beck's empirical data, but the LWIR bandgap of 0.11 eV is far above the $6.8E_g$ value given by Eq. (7.10). The reason for this is to be found in the diode geometries utilized in the comparison plot of Fig. 7.12. The diode data from Elliott et al. were obtained from diodes with a quoted diode

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

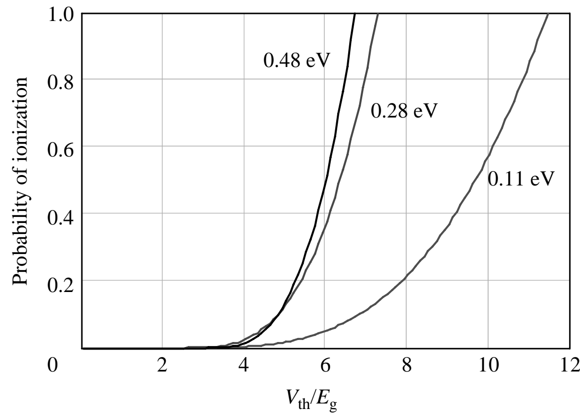


Figure 7.18 Modeled threshold voltages normalized to E_g for various values of E_g, for a depletion width of 6 μm.

spread of ∼3 μm. The DRS 10.8 μm data at 83 K were obtained on diodes with an essentially unlimited n-volume available for diode spread and hence a quadratic dependence of potential on distance. Both of these geometries result in larger values of the energy dispersion coefficient, $\alpha(E)$, and hence result in electron energy values much closer to the actual voltages used to generate threshold; that is, α tends to unity. As an example, for a confined 3 μm diode spread the predicted $\alpha(E)$ for phonon scattering is 0.83, resulting in a predicted threshold voltage $\sim 6.8E_g$, in agreement with Beck's analysis of Elliott's data.

It is of interest at this point to compare the empirically derived ionization rate with the electron scattering rates associated with optical phonons and alloy disorder. The scattering rates, $1/\tau_{op}$, as given by Eq. (7.4) for high-energy Kane-like electrons, and $1/\tau_{alloy}$, as given by Eq. (7.5) for parabolic electrons of energy E, are shown in Fig. 7.19, together with the impact ionization rate given by Eqs. (7.7), (7.8), and (7.4), modified by the empirical factor of 10 obtained from the comparison with Beck's APD data, for both MWIR and LWIR HgCdTe at 77 K. It is apparent that for high electron energies, impact ionization is much more probable than scattering by optical phonons for both cutoff wavelengths. Alloy scattering on the other hand appears more relevant, particularly for shorter cutoff wavelengths. However, as discussed earlier, Eq. (7.5) is not appropriate for deep excursions into the conduction band and can only be considered a crude guide for these high electron energies. The experimental data on threshold and dead voltages as a function of cutoff wavelength discussed above suggest that electron scattering at 77 K is not a concern, over and above the arguments presented for the effects due to optical phonons. It would thus appear that alloy scattering, possibly for the reasons alluded to earlier, is a non-issue at 77 K and probably also at higher temperatures.

Equation (7.18) can be used to optimize the gain of any HgCdTe APD. It is obvious that higher gain should be obtained for a fixed applied bias voltage by operating with a smaller depletion width, W, that is until W becomes $<\lambda_e$, the phonon-limited mean free path. At this point, electron energy will no longer be

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

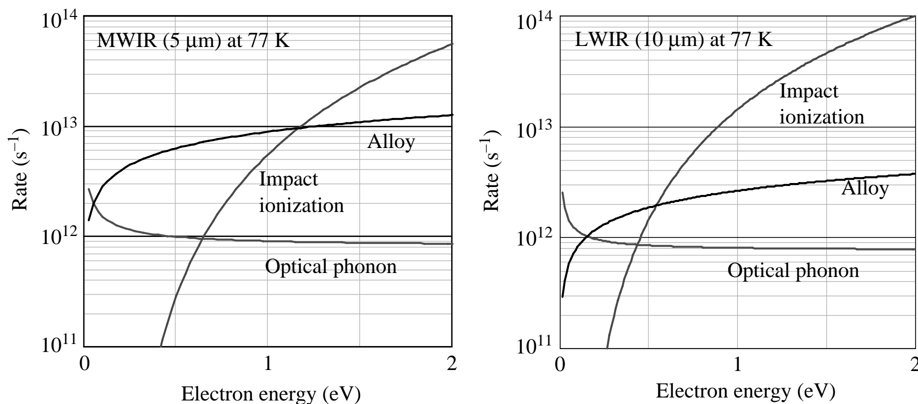


Figure 7.19 Comparison of electron impact ionization and scattering rates as a function of electron energy, for MWIR ($E_g = 0.25$ eV), and LWIR ($E_g = 0.124$ eV) HgCdTe at 77 K.

dispersed by optical phonons and the gain for a fixed bias voltage should saturate. For smaller values of *W*, the gain at a fixed bias voltage should in fact decrease due to the smaller gain volume.

The speed of the EAPD will be determined by the transit of the slower carrier, in this case the heavy hole. Unlike the case of the electron, the hole will achieve a saturation velocity given by its rate of energy loss to optical phonons $\sim(\varepsilon_{opt}/m_h^*)^{1/2} \approx 8 \times 10^6$ cm/s, for a hole mass $\sim 0.55 m_o$.

The picture that emerges from all of this is that HgCdTe APDs can be reasonably well explained by ballistic arguments for electron transport, with the assumption of a correlation between electron energy and applied voltage, given by $E = \alpha(E)\Phi$, where Φ is directly related to the applied bias voltage. Theory for optical phonon scattering indicates that α(E) varies directly with the dielectric properties of the material composition, and inversely with the depletion width, and is essentially independent of electron energy.

7.4 Room-Temperature HgCdTe APD Performance

It is of interest to utilize the above model to predict the behavior of SWIR HgCdTe at room temperature. The scattering rates given by Eqs. (7.4) and (7.5), together with the impact ionization rate given by Eq. (7.9), are shown in Fig. 7.20(a) for $E_g = 0.56$ eV ($\lambda_c = 2.2$ μm) at 295 K. It is again apparent that alloy scattering should dominate optical phonon scattering, but all evidence points to the contrary, and it will be ignored in the remaining analysis. For $W = 6$ μm, the threshold energy as defined earlier is shown in Fig. 7.20(b) and is ~1.2 eV. The energy dispersion coefficient due to optical phonons, $\alpha(E) = \lambda_e/W$, for a 2.2 μm cutoff, is shown in Fig. 7.21(a) for both 77 K and 295 K. It is apparent that λ_e is considerably shorter at room temperature due to the increased density of optical phonons, and for a 6 μm diode spread $\alpha \approx 0.125$. Thus, the threshold electron energy of 1.2 eV for impact

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

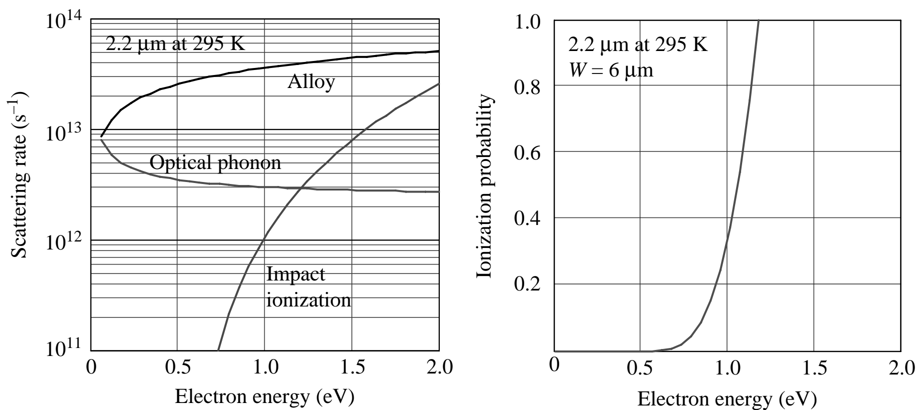


Figure 7.20 (a) Scattering and impact ionization rates for SWIR HgCdTe at 295 K and (b) impact ionization probability as a function of electron energy for $W = 6$ μm.

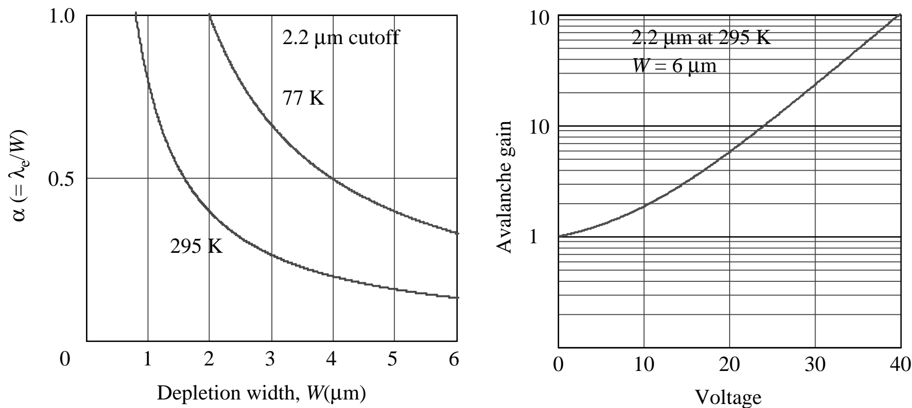


Figure 7.21 (a) Energy dispersion coefficient vs. depletion layer width at 77 K and 295 K, and (b) avalanche gain vs. bias voltage for $W = 6$ μm at 295 K, for 2.2 μm cutoff wavelength HgCdTe diodes.

ionization requires an applied bias voltage of $\sim 1.2/0.125 \approx 10$V. The modeled avalanche gain vs. applied bias voltage utilizing this threshold voltage is shown in Fig. 7.21(b). For comparison purposes, the published data of Beck et al. [107] for 2.2 μm cutoff APDs at room temperature are shown in Fig. 7.22. The agreement with model is seen to be excellent.

It is obvious from the above model that significant avalanche gain can be achieved at lower values of applied bias voltage by utilizing narrower depletion regions at any temperature. Figure 7.21 suggests that 2.2 μm HgCdTe requires depletion widths ~ 1 μm to achieve significant gain at 295 K at relatively small applied bias. The use of such small depletion region widths in turn requires somewhat larger electron threshold energies to impact ionize, due essentially to the smaller active volume. A calculation for $\lambda_c = 2.2$ μm, with $W = 1$ μm at 295 K is

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

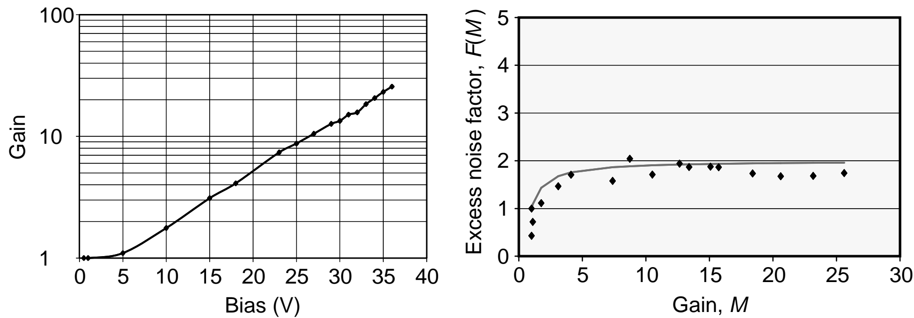


Figure 7.22 Measured avalanche gain vs. bias, and excess noise factor vs. gain, for 2.2 μm cutoff HgCdTe at room temperature.

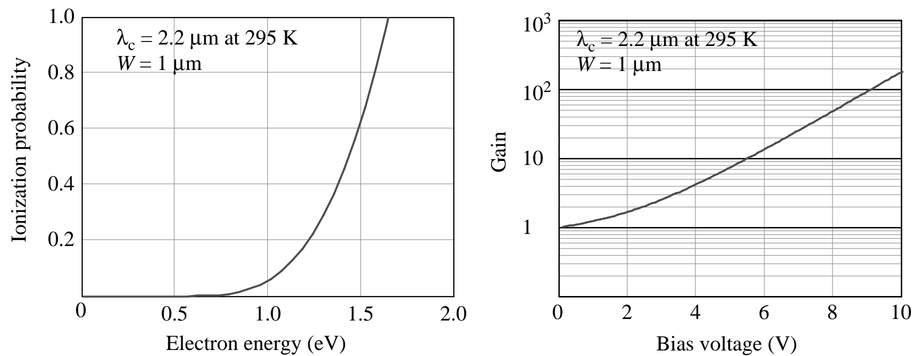


Figure 7.23 (a) Impact ionization probability as a function of electron energy for $\lambda_c = 2.2\,\mu m$ at 295 K with a depletion width of 1 μm and (b) predicted gain vs. applied voltage for the same geometry.

shown in Fig. 7.23. The threshold energy for impact ionization under these circumstances is >1.6 eV, which is deep in the conduction band, but electron dispersion should be essentially eliminated, and α will approach unity. These energy values represent relatively uncharted waters for the HgCdTe conduction band, and at some point one might expect to run into the secondary minima predicted by the band structure calculations described earlier. Significant scattering mechanisms may then come into play to more efficiently randomize electron motion.

7.5 Monte Carlo Modeling

Nothing has been said to this point concerning the expected noise for this avalanche model. However, the use of a ballistic model implies that avalanche gain is history-dependent, and hence that the excess noise factor in an electron-controlled HgCdTe APD will be less than the value 2 that is typically associated with completely random gain occurrence in a semiconductor in which only

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

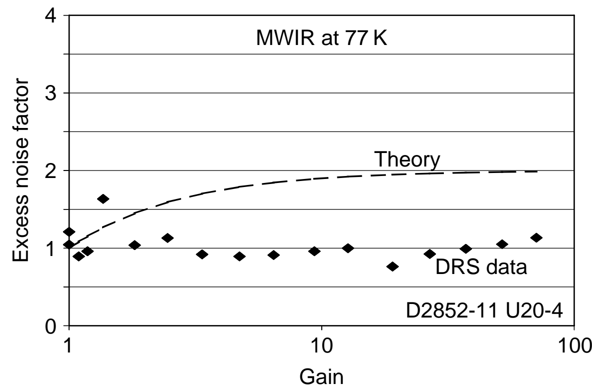


Figure 7.24 Excess noise factor vs. gain for MWIR APD at 77 K.

one carrier is capable of impact ionization. This is in qualitative agreement with the data shown in Fig. 7.24, in which the theoretical prediction of McIntyre [91] for random gain is compared to DRS noise data taken on cooled MWIR HgCdTe APDs at 77 K. A quantitative assessment of expected excess noise requires the use of more sophisticated analytical modeling such as Monte Carlo techniques.

Monte Carlo modeling [108,109] of the HgCdTe EAPD has been carried out at the University of Texas at Austin, utilizing standard expressions for scattering. Impact ionization was represented by a Keldysh-like expression containing an energy-dependent term $[(E/E_{\mathrm{th}}) - 1]^r$, where $r \sim 3$. The band structure of HgCdTe was essentially simulated by a two-band model with a valence band maximum at $k = 0$, and conduction band minimum at $k = 0$, and satellite valleys far out in the Brillouin zone at a height of 2 eV, with appropriate effective masses.

An example of modeled gain versus voltage for HgCdTe with $E_g = 0.29$ eV, at 77 K, is shown in Fig. 7.25(a). The model is in good agreement with data taken at

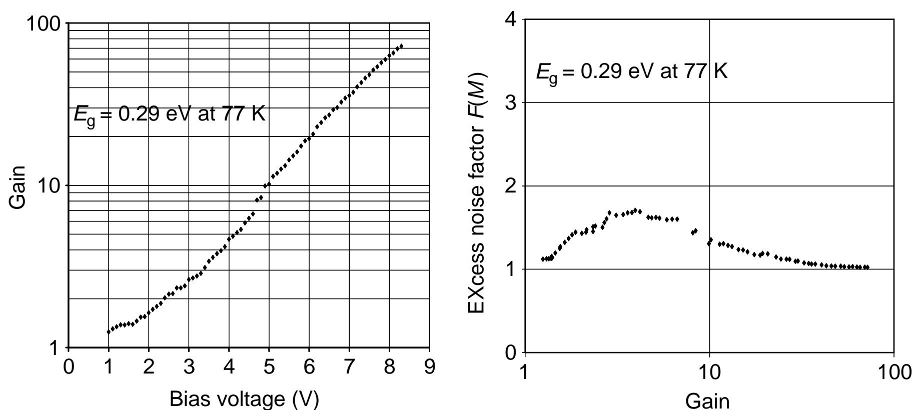


Figure 7.25 Monte Carlo model of gain, and noise factor, for 4.3 μm HgCdTe at 77 K.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

DRS on 4.3 μm cutoff HgCdTe at 77 K. The excess noise factor predicted for these same diodes is shown in Fig. 7.25(b) as a function of avalanche gain and is seen to approach unity at large values of gain. With regard to excess noise it is interesting to note that the measured noise factor as a function of gain for SWIR HgCdTe at room temperature as shown in Fig. 7.22 approaches McIntyre's theoretical model. It appears that the excess noise factor is now larger than for the low-temperature MWIR case, implying that optical phonon scattering is indeed playing a more significant role at room temperature for wide-bandgap HgCdTe with diode geometries that are considerably in excess of a phonon limited mean free path. Avalanche gain is still solely due to the electron, at the bias voltages used, but it now appears more randomized due to the increased nature of the electron–optical phonon interaction.

7.6 Conclusions

HgCdTe electron avalanche photodiodes (EAPDs) are in their infancy but show great promise for applications across the whole IR spectrum, in both military and commercial arenas. The band structure of the HgCdTe alloy system is unique in that the conduction band allows excursions of the electron deep into the band, but the valence band does not offer the same privilege to the hole, resulting in large values of exponential gain (>1000) with essentially zero excess noise. No III-V semiconductor alloy offers this capability, with the sole exception of possibly the binary alloy, InSb, which does have the desired band structure but with a somewhat lower heavy hole mass, which may be an issue. The current favorite in the commercial arena is 1.7 μm cutoff InGaAs, but the band structure of this alloy exhibits relatively low secondary conduction band minima, and a spin–orbit splitting that is narrower than the bandgap, leading to both electron and hole avalanche multiplication. This allows only modest values of avalanche gain, with considerable excess noise.

It remains to be seen what the limitations are for device bandwidth, but this should really be an architectural issue. Electron mobility values are reasonable at all temperatures, implying that high-speed EAPDs should be possible. Dark currents in HgCdTe are well understood, and HgCdTe material properties can be optimized to accommodate any cutoff wavelength at any operating temperature.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Chapter 8

Future HgCdTe Developments

New applications of IR devices are continually evolving. Properties that will be required of future FPAs are:

- Multispectral, multifunctional system operation;
- Larger FPAs with higher density pitch;
- Low background operation at the highest possible operating temperature;
- Higher operating temperatures for tactical systems, preferably room temperature;
- Longer cutoff wavelengths;
- High bias operation for avalanche photodiodes (APDs) of all wavelengths.

All of these requirements basically push HgCdTe into extreme modes of operation for currently available devices, in which all the warts of the material and processing are exposed. Dark currents must be reduced, dramatically in some cases, requiring a quantum jump in material performance, together with the introduction of innovative detector architectures. The discussion of the HOT detector requirements in Chapter 6 indicates the required path, namely a reduction in the defect and impurity concentrations in starting materials, together with lower extrinsic doping concentrations. This chapter expands the HOT detector discussion to address detector architectures that, if properly implemented, could meet all of the above IRFPA requirements. We will begin with a discussion of the dark current components on both sides of a standard HgCdTe homo-junction photodiode, with a view to optimizing the diode architecture for specific applications. Examples will be given of projected performance for the above applications utilizing HgCdTe. It will become apparent that one particular architecture, above all others, can meet all of the above IRFPA requirements.

8.1 Dark Current Model

The relevant dark current components, discussed in Sec. 5.6, for a one-sided abrupt $n^+/n^-/p$ homo-junction with $n \ll p$, are

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

- Diffusion currents from the n and p volumes to a depth $\sim L$, a minority carrier diffusion length (these have Shockley–Read (S-R) and Auger components);
- Generation through S-R centers in the depletion region located in the n-side of the diode;
- Surface generation;
- Tunneling across the depletion region of the diode, both direct band-to-band, and via mid-gap traps.

It is assumed in subsequent discussion that surface passivation techniques allow for the neglect of surface generation currents. This, by and large, is a good assumption for HgCdTe when CdTe passivation is used.

The various dark current components are listed below for the n- and p-sides of the junction. Non-equilibrium operation is assumed, such that majority carrier concentrations are given by the relevant doping, and minority carrier concentrations are essentially zero. We will also assume that the relevant S-R centers are neutral in nature and that $\tau_{po} \approx \tau_{no} \approx \tau_{sr}$, where τ_{po} and τ_{no} represent times for the center to capture holes and electrons. Further, we can assume that for a diode formed by ion implantation, or ion etching, of p-HgCdTe, the values for τ_{sr} are the same on either side of the junction.

8.1.1 N-side

Diffusion current associated with S-R centers located at the intrinsic energy level is given by

$$J_{\mathrm{nsr}} = qn_{\mathrm{i}}^2 t/[n_{\mathrm{d}} + 2n_{\mathrm{i}}]\tau_{\mathrm{sr}}. \tag{8.1}$$

Diffusion current associated with Auger1 generation is given by

$$J_{\mathrm{A1}} = qn_{\mathrm{d}} t/2\tau_{\mathrm{Ai1}}, \tag{8.2}$$

where this component may also be associated with the n^+ contact region to the diode.

Depletion current associated with thermal generation through S-R centers at the intrinsic energy level is given by

$$J_{\mathrm{dep}} = qn_{\mathrm{i}} W/[2\tau_{\mathrm{sr}}], \tag{8.3}$$

where W is the width of the depletion region. For a planar diode, W will vary as $[2\varepsilon\varepsilon_{\mathrm{o}}(V + E_{\mathrm{g}})/qn_{\mathrm{d}}]^{1/2}$, until the applied bias voltage is large enough that W consumes the entire n-region. This situation invariably applies to APD operation.

Tunneling via N_{t} bandgap states located at E_{t}, for a parabolic barrier, in a uniform electric field, with doping n_{d}, gives a bias voltage dependent current density of

$$J_{\mathrm{tsr}} = [10^{-13} N_{\mathrm{t}} V/E_{\mathrm{g}}] \exp[-1.5 \times 10^{10} \pi m_{\mathrm{e}}^{*1/2} E_{\mathrm{g}}^{3/2}/n_{\mathrm{d}}^{1/2}/(E_{\mathrm{g}} + V)^{1/2}], \tag{8.4}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

where V is the applied voltage, and $m_e^* = 7 \times 10^{-2} E_g$, E_g is in eV; n_d and N_t are in cm^{-3}.

Direct tunneling between the conduction and valence bands, for a parabolic barrier, in a uniform electric field, with doping n_d, gives a bias dependent current of

$$J_{tdir} = 1.2 \times 10^{-2} V [n_d(E_g + V)]^{1/2} \exp(-9.43 \times 10^{10} m_e^{*\,1/2} E_g^{3/2} / [n_d(E_g + V)]^{1/2}. \quad (8.5)$$

These expressions for tunneling based on a uniform field should give an overestimate of tunnel current for standard abrupt junction diodes. Allowances for this effect can be made by the inclusion of a factor, f, in the exponents in Eqs. (5.19) and (5.20). Modeling suggests that $1 < f < 2$.

8.1.2 P-side

Diffusion current associated with S-R centers located at the intrinsic energy level is given by

$$J_{psr} = q n_i^2 t / [n_a + 2n_i] \tau_{sr}. \quad (8.6)$$

Diffusion current from the donor-like S-R centers associated with the V_M metal vacancy, located 0.03 eV from the conduction band, and defined by a lifetime, $\tau_V = 5 \times 10^9 / V_M$, is given by

$$\begin{aligned} J_{pvac} &= q n_i^2 t / n_1 \tau_V, \\ &= 2 \times 10^{-10} q p_1 V_M t. \end{aligned} \quad (8.7)$$

Diffusion current associated with Auger7 generation is expressed as

$$J_{A7} = q n_a t / 2\tau_{Ai7}, \quad (8.8)$$

where again there may be a contribution from any p^+ contact that may be utilized in the diode architecture.

It is a simple matter to calculate these dark currents for any defined diode geometry, whether planar (as for VIP and DLHJ) or cylindrical (HDVIP) in nature, utilizing the lifetime values for n- and p-HgCdTe discussed earlier in Chapter 5.

Consider the case of a planar MWIR diode operating at a temperature of 140 K with a cutoff wavelength of 5 μm. We can calculate all of the above components as a function of doping concentration, for specific values of S-R lifetime, and density of traps, N_T, for tunneling. We will neglect direct tunneling relative to the trap assisted variety, which is normally justified for reasonable densities of mid-gap states ($>10^{12}/cm^3$), and also assume a factor of $\sim$1.7 for the tunneling parameter to allow for a non-uniform electric field in the junction. The dark current components for the n- and p-sides of the junction are shown in Fig. 8.1, assuming equal thicknesses for the n- and p-volumes of 3 μm, and a common S-R lifetime of 100 μs.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

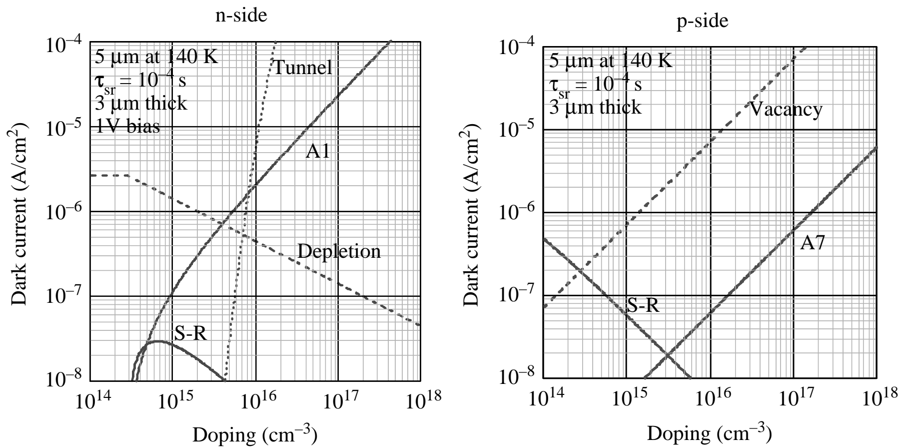


Figure 8.1 Dark current components for MWIR 5 μm HgCdTe photodiode at 140 K.

A number of issues are immediately apparent.

- Vacancies are to be avoided at all costs if optimum diode dark current performance is required.
- In the absence of vacancies, the p-volume dark current optimizes at a doping concentration ~mid-10^{15} cm^{-3} for the assumed S-R lifetime of 100 μs. Longer lifetimes will optimize at lower doping concentrations as S-R currents vary as $1/p$, and Auger7 as p.
- The last thing that you need in your IR absorption volume is a depletion region. It is a source of significant dark current relative to diffusion current, even for the S-R limited case. For the example chosen here, depletion current is two orders of magnitude higher than any optimized diffusion current component. This is because the S-R limited diffusion current/unit volume of Eq. (8.1) varies as $n_i^2/[n_d + 2n_i]\tau_{sr}$, whereas the S-R limited depletion current of Eq. (8.3) varies as $n_i/2\tau_{sr}$. For operating temperatures where $n_d \gg 2n_i$, then diffusion current $\ll$ depletion current. This will be true regardless of the value of τ_{sr}. For $n_d < 2n_i$, that is, at high temperatures, diffusion current is of exactly the same form as depletion current.
- Low doping concentrations are mandatory on the n-side to minimize the Auger1 diffusion component. The n^+ region Auger1 is suppressed by band filling.
- Auger1 can also be eliminated by utilizing a 1 V bias on the diode. The depletion region can then be made to extend throughout the n-volume, and the n-side is dominated by depletion current. This is shown in Fig. 8.1 at n-side doping levels $<10^{15}$ cm^{-3}.
- Tunneling considerations indicate that doping concentrations in the depletion region must be $<5 \times 10^{15}$ cm^{-3} for the applied 1 V bias.

The issues listed above are significant at all cutoff wavelengths, and for any homojunction architecture fabricated in any semiconductor material, the depletion current

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

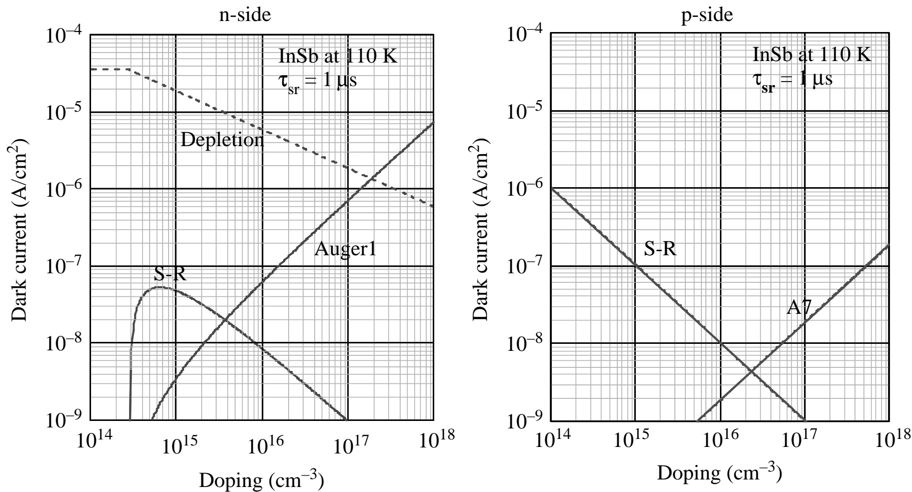


Figure 8.2 Dark current components for a p^+/n^- InSb homo-junction at 110 K.

limitation is undesirable and yet impossible to avoid. The diode will never achieve its maximum potential performance, which is a limitation associated with pure diffusion current from the IR absorbing volume of the device.

An example of this dark current modeling for another semiconductor is shown in Fig. 8.2, where the various thermally generated dark current components are shown for the commonplace InSb p^+/n^- homo-junction architecture, at an operating temperature of 110 K. This temperature is acknowledged as an upper limit for InSb operation in a background-limited environment, and the reason is clear. The S-R limited lifetime reported for InSb is not particularly good and limited to values $<1\ \mu s$. This results in a very large depletion current component at 110 K $>10\ \mu A/cm^2$ for the n-side doping concentrations typically used in InSb FPAs and prohibits BLIP performance for high f-number systems at higher temperatures.

8.2 The Separate Absorption and Detection Diode Structure

There is one architecture that addresses all of the issues raised above, and that is the separate absorption and detection (SAD) structure shown in Fig. 8.3. The SAD diode consists of a field-free p-type IR absorbing region coupled to a wider gap low-doped n (or p) detecting region by a graded bandgap transition. The dark currents in the p absorbing volume are pure diffusion-current limited and can be optimized by an appropriate doping level that is consistent with the S-R lifetime of the p-volume. The minority carriers in this volume must be extracted by applying a suitable bias voltage to the n^+ contact of the diode. This voltage is required to overcome the barrier that exists in the transition region of the SAD structure. The composition of the n (or p) region can be selected so that the dark current is dominated by the p absorption volume diffusion currents. If the doping of the wide gap region is low

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

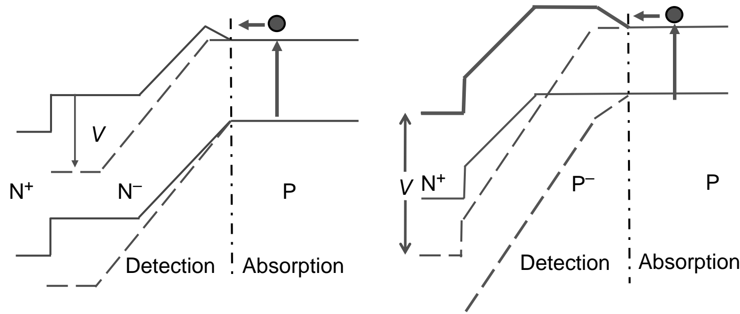


Figure 8.3 $N^+/N^-/P$ and $N^+/P^-/P$ separate absorption and detection (SAD) structures, with applied bias voltage *V*.

enough relative to the doping in the absorbing volume, then essentially none of the applied voltage will penetrate the p-absorber, and the depletion region will be confined to the wide gap volume.

The gradient of the bandgap, F_t, in the transition region must satisfy two conditions:

1. F_t must be large enough that the depletion current from the graded gap region is less than the diffusion current from the p-absorber. This condition is given by $(2kT/qF_t)n_i/2\tau_o < 2n_i^2 t/p\tau_o$, which reduces to the condition $F_t > (kT/q)(p/2n_i t)$, where t is the thickness of the absorber.
2. F_t must not be so large as to cause tunnel breakdown in the wide bandgap diode depletion region. This can be avoided by a suitable choice of transition region thickness and hence a value for E_g in the wide gap region that avoids tunneling.

We can easily calculate the requirements of the SAD structure that ensure p-side limited performance, and this is illustrated in Fig. 8.4 with an n-volume bandgap = 0.335 eV (cutoff wavelength of 3.7 μm). The p-volume is modeled with no metal vacancies, and the magnitude of dark current from the n-side is less than that from the IR absorbing p-volume.

It should be noted that the $N^+/P^-/P$ version of the SAD structure is virtually identical to the separate absorption and multiplication (SAM) structure used in avalanche photodiode (APD) technology. However, in the SAD diode the emphasis is on the low dark current aspect of the device.

The SAD structure is an interesting concept in that the p-volume acts virtually as a photoconductive substrate whose minority carriers are extracted by a reverse-biased heterojunction. Another aspect of the optimized SAD structure is that it represents the very antithesis of one of the main HgCdTe photodiode approaches of the last twenty years, namely the double-layer heterojunction, described in Sec. 5.4. The double-layer heterojunction technology, as practiced in the USA, is wide bandgap p^+/n, where n represents the IR absorbing volume. Great pains are normally taken to interdiffuse the heterojunction so that no barrier is present between the p^+ and n-layers, and the depletion region resides entirely within the

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

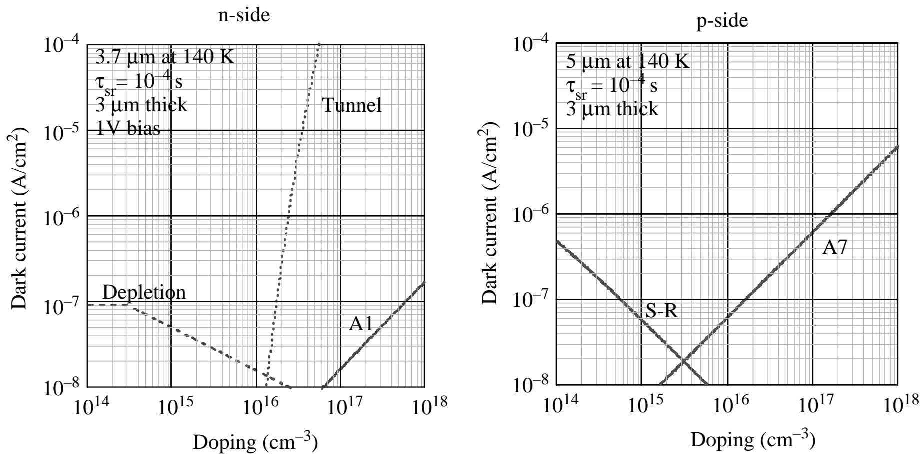


Figure 8.4 Dark current components for a MWIR 5 μm SAD structure operating at 140 K. The n-side has a cutoff wavelength of 3.7 μm.

IR absorbing volume. All of these attributes are exactly what one does not need for an optimized IR heterojunction photodiode. It would appear that perhaps we have been going in the wrong direction these twenty years with regard to the race for the ultimate photon detector.

The SAD architecture can be applied to many of the desired advanced FPA requirements described earlier in this chapter. Let us now do this and generate performance projections, together with associated materials requirements, for the various advanced FPAs.

8.3 Multicolor and Multispectral FPAs

Multispectral FPA detection has a number of very desirable features, among which are possible target discrimination and identification. A multispectral capability can be achieved in two ways. Optical systems employing either dispersive or interferometric elements can be used in conjunction with broadband FPAs, or alternatively the FPAs can be tuned to detect a finite number of specific spectral bands. The first option is optically complex, requiring considerable system computing horsepower for real-time analysis, but it employs a very simple FPA approach. The second option is optically simple, but this is achieved at the expense of greatly increased complexity in the FPA. The optical solution is best used for complex spectral analysis, whereas the tuned FPA solution is best suited for a small number of spectral bands for which essentially simultaneous temporal and spatial detection are required.

Two architectures are currently being employed to address the tuned IRFPA option, namely, the multilayer DLHJ, and the multilayer HDVIP. The DLHJ architecture, as in the monocolor case, is bump-bonded to the Si ROIC, and true simultaneous detection requires two bumps per pixel. A simpler version operates sequentially, utilizing back-to-back diode operation, and only requires one bump

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

per pixel, a quality that can be important in high-density IRFPAs. The multilayer HDVIP architecture simply consists of two separate monocolor HDVIP arrays that are physically glued on top of one another. These two simultaneous architectures are shown in Fig. 8.5. For sequential DLHJ operation, the bump to MCT2 can be eliminated, but a backside contact will be required for the array. It is apparent that the two-color DLHJ is more complex from a materials growth point of view and probably requires a vapor-phase growth technique for implementation. The two-color HDVIP is merely an evolutionary technology relative to monocolor HDVIP. It requires no advances in materials technology and can utilize either liquid- or vapor-phase growth techniques. It is, however, a somewhat more complex structure from a fabrication point of view, requiring an insulated via from the top layer through the second layer to the ROIC. The spectral response of early two-color FPAs is shown in Fig. 8.6, for both MW/MW and MW/LW. Composite operabilities >99% have been achieved for MW/MW and MW/LW FPAs with <5% spectral crosstalk.

Multispectral operation on a high-density pitch is obviously challenging because of the number of contacts required per unit cell. Even the one bump per pixel for pseudo-simultaneous operation of the DLHJ can be problematic because of the associated forward bias on the non-detecting shorter wavelength diode, which can then act as an IR emitter, raising possible cross-talk issues. The SAD device

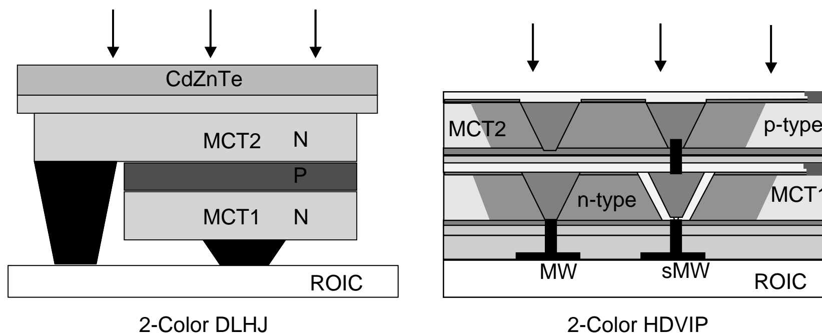


Figure 8.5 Two-color HgCdTe architectures.

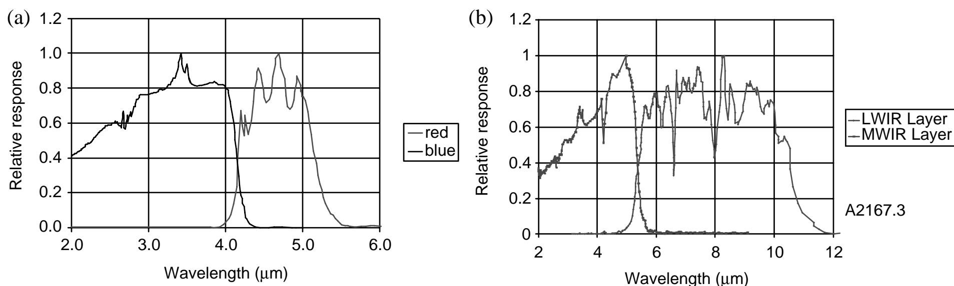


Figure 8.6 Spectral response of (a) MW/MW and (b) MW/LW 2-Color HDVIP HgCdTe FPAs.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

offers the possibility of 2-color operation with no forward bias within the unit cell. The p-volume can constitute the longer cutoff wavelength absorbing region and the n-volume the shorter of the two cutoffs. The n-side will be biased between two values of reverse bias, namely one lower value that simply extracts photo-generated carriers from the wide gap n-region, and a larger value of reverse bias that extracts photo-generated carriers from both the n- and p-regions. The longer cutoff spectral information is obtained by a subtraction of the two signals.

8.4 High-Density FPAs

High-density monocolor FPAs will be required with pitches that possibly border on or exceed the diffraction limit of the system optics. The significant advantage in reducing the pitch and the size of the FPA lies in the reduction in cost of the overall IR system. Two major limitations on pixel size are (1) the physical size of indium bumps and the pressure required to bond a large number of pixels, and (2) containing the spread of diodes formed by ion implantation or ion etching. The SAD structure will not lend itself per se to the achievement of higher densities, but the use of an absorbing p-volume will allow for the utilization of arsenic doping. Activated arsenic resides on the Te sublattice and as such is not moved around by Hg interstitials, thus theoretically enabling the formation of very small diodes limited essentially to the size of the n^+ region associated with the etched via and/or the implant region. This concept has been verified by fabricating arsenic-doped MWIR HgCdTe test diodes with minimal diode spread [110].

8.5 Low Background Operation

Surveillance in space requires the use of IR detectors that are close to BLIP at very low background flux levels. The more challenging of these scenarios occur at the longer cutoff wavelengths $>12\ \mu m$. It is of interest to look at operating the HgCdTe SAD structure in two situations:

1. Operation at 40 K with a cutoff wavelength $>14\ \mu m$ and
2. Operation at a cutoff wavelength of 25 μm.

8.5.1 LWIR 14 μm at 40 K

The optimum FPA performance will be obtained when the p-volume diffusion currents dominate the unit cell. Let us model this limit by assuming a reasonable S-R lifetime throughout the unit cell of $10-100\ \mu s$. The p-side dark current components due to S-R and Auger7 for an absorbing region at 14 μm are shown in Fig. 8.7 at an operating temperature of 40 K, for a thickness of 6 μm, together with the dark current components of a 10 μm cutoff n-type detection region. The bias is assumed to be $\sim 0.2\ V$.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

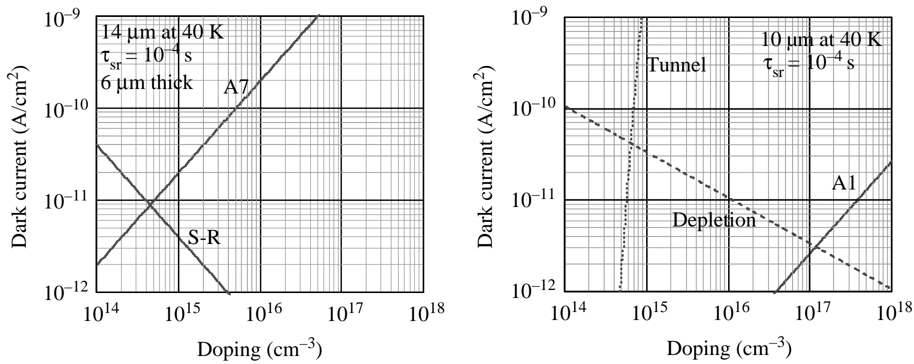


Figure 8.7 Dark current components in a SAD structure with absorption at 14 μm, and detection at 10 μm, for an operating temperature of 40 K.

It is apparent that doping concentrations on the n-side are limited to $<5 \times 10^{14}\ \mathrm{cm}^{-3}$ by tunneling considerations. The limiting dark current component is $\sim 5 \times 10^{-11}\ \mathrm{A/cm}^2$, which will yield an $R_o A \approx 2 \times 10^9\ \Omega\,\mathrm{cm}^2$.

8.5.2 Low background operation at a cutoff of 25 μm

Some space surveillance is carried out at cutoff wavelengths as long as 25 μm. The current state-of-the-art detection utilizes extrinsically doped Si operating at temperatures <10 K. It is of interest to model the potential performance of a HgCdTe SAD diode operating at these wavelengths. Let us consider the absorbing p-volume to have a somewhat conservative S-R lifetime of 10 μs and a thickness of 10 μm. Initially, we choose an operating temperature of 25 K, and the modeled data is shown in Fig. 8.8.

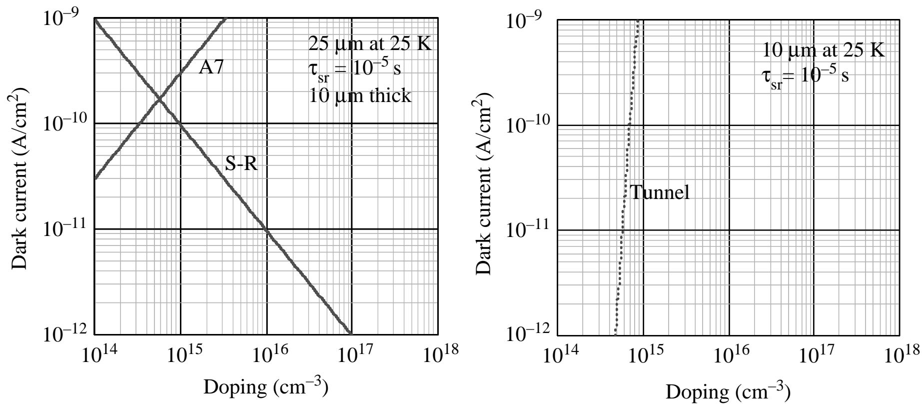


Figure 8.8 Dark current components for a SAD diode at 25 K with a 25 μm cutoff p-absorber and a 14 μm cutoff n-detector.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

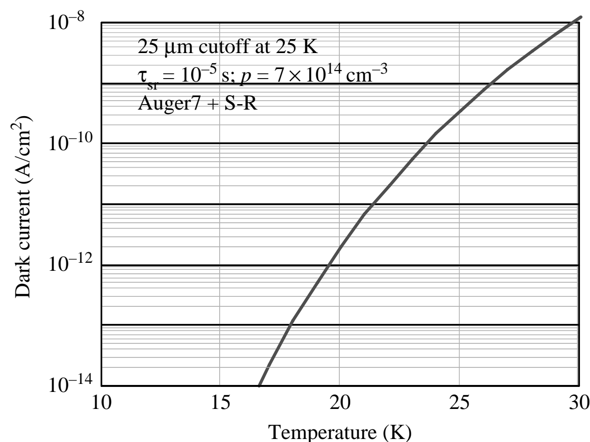


Figure 8.9 Auger7 + S-R limited dark current vs. temperature for 25 μm cutoff HgCdTe.

The absorber is seen to optimize at a doping level of $\sim 10^{15}$ cm^{-3}, providing a dark current $\sim 3 \times 10^{-10}$ A/cm^2. The major limitation on the n-side is imposed by tunneling considerations, and for this reason the cutoff wavelength of the n-side diode region is chosen to be 10 μm, allowing a manageable n-side doping level of $\sim 1-7 \times 10^{14}$ cm^{-3}. The dependence of p-absorber dark current on temperature for the S-R lifetime $\tau_{sr} = 10$ μs and a p-doping level of 7×10^{14} cm^{-3} is shown in Fig. 8.9.

8.6 Higher Operating Temperatures

The requirements for high-temperature operation of photon detectors were discussed in Chapter 6. It is meaningful to re-examine that analysis with a view to the possible implementation of the SAD structure. Consider the 5 μm MWIR planar homo-junction discussed in Sec. 8.1, but at an operating temperature of 253 K, which is in the two-stage thermo-electric cooler range. The dark current components for the n- and p-sides of the homo-junction are shown in Fig. 8.10, for a bias voltage of 1 V.

Examination of Fig. 8.10 shows that at this operating temperature there is no obvious benefit to utilizing a SAD structure. For $n_i > n_d$, n_a you do not pay a penalty for the presence of a depletion region in the absorbing volume, as the dark current generation rate per unit volume through S-R centers is identical in both the depletion and diffusion regions. The only realistic benefit to a SAD structure in this case occurs if there is any excess Auger1 generation from the n^+ contact region of the diode, which will be greatly reduced if the n^+ region is wide gap. This is not thought to be the case for relatively narrow bandgaps, as the n^+ concentrations are $\sim 10^{18}$ cm^{-3}, and significant band-filling occurs due to the low density of states in the conduction band. This band-filling effectively quenches Auger1 transitions at energies close to the band edge. This will not necessarily be the case for wider bandgaps.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

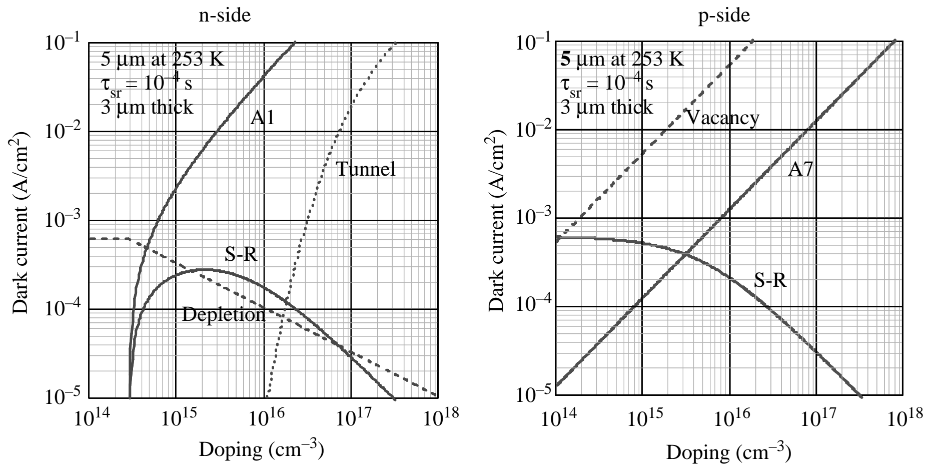


Figure 8.10 Dark current components for 5 µm cutoff MWIR at 253 K, for $\tau_{sr} = 10^{-4}$ s.

As stated in Chapter 6, optimum high-temperature performance will be achieved with an n^+/p^- homo-junction. For sufficiently low doping, the application of a reverse bias can drive the depletion region to fill the complete detector volume, in which case Auger7 generation is completely suppressed due to the elimination of majority carriers. This case is essentially a P-I-N diode and represents the lowest dark current situation for HOT operation. The performance will then be limited by the total dark current generation through all of the S-R centers of the active volume. The dark current components for a 5 µm cutoff MWIR n^+/p^- homo-junction at 295 K are shown in Fig. 8.11 for a 1 V bias. It is apparent that at the applied bias of 1 V the depletion region on the p-side can punch through the whole volume. The remaining dark current components under these operating conditions are Auger1

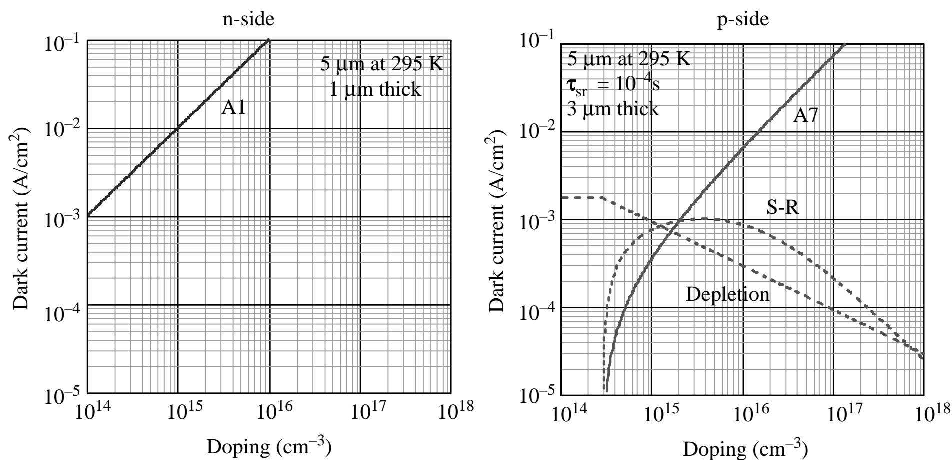


Figure 8.11 Dark current components for a 5 µm cutoff MWIR n^+/p^- homo-junction at 295 K. The assumed $\tau_{sr} = 10^{-4}$ s, and the applied bias $V = 1.0$ V.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

from the n^+ volume and S-R from the p-side depletion region. Assuming complete suppression of Auger1 generation in the n^+ region, then the limiting dark current will be given by $qn_i t/2\tau_{sr}$, where $t \approx 1/\alpha$, and α is the absorption coefficient. For $\tau_{sr} = 10^{-4}$ s and $t = 3\,\mu$m, we have $J_{dep} = 1.8 \times 10^{-3}$ A/cm^2; similarly, for $\tau_{sr} = 10^{-3}$ s (values as large as are observed), $J_{dep} = 1.8 \times 10^{-4}$ A/cm^2. The required doping concentrations will be $<5 \times 10^{14}$ cm^{-3}, with arsenic as the prime candidate. For comparison purposes, a 2π FOV flux current at 295 K is 2×10^{-3} A/cm^2.

8.6.1 High-gain APDs

The requirements for high gain with low excess noise in HgCdTe homo-junction APDs for all cutoff wavelengths were discussed at length in Chapter 7. The ultimate performance, particularly at SWIR cutoff wavelengths, will be provided by the SAD structure, or in this APD case, using its appropriate name, the separate absorption and multiplication (SAM) diode. This structure has been in use in the $In_{0.53}Ga_{0.47}As$ (with a cutoff wavelength of 1.7 μm) world for a number of years because it lends itself to lattice matching throughout the heterostructure by the use of InP, both as the growth substrate and as the epilayer in which avalanche gain takes place. The other, little talked about benefit to this SAM structure is the avoidance of a depletion region in the absorbing region, resulting in a huge reduction in dark current at the operating temperature in question.

It is of interest to consider the use of a SAM structure in SWIR HgCdTe at room temperature. The dark current components for a 1.7 μm absorber on a 1.1 μm avalanche region are shown in Fig. 8.12. Again, a minimal number of vacancies can be tolerated $<10^{14}$ cm^{-3}. The dark current in the absorber will optimize at $\sim 5 \times 10^{-11}$ A/cm^2, for p-type doping concentrations $\sim 10^{16}$ cm^{-3}. The depletion

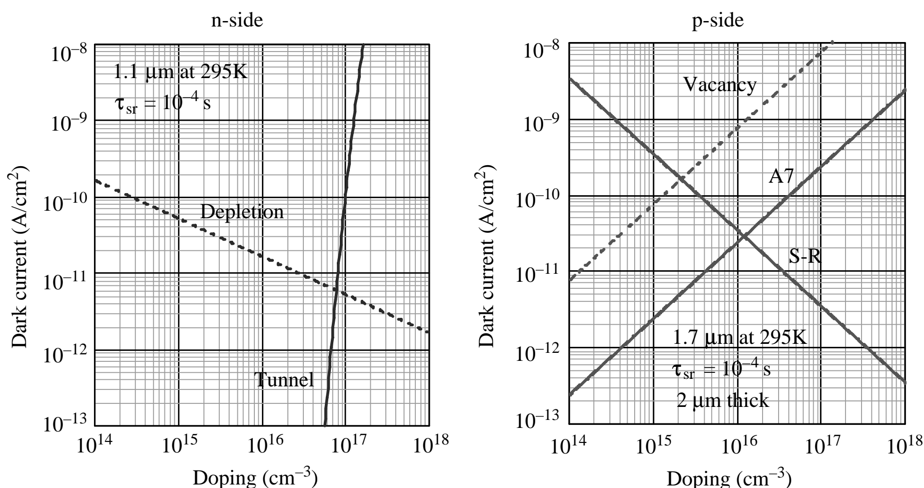


Figure 8.12 Dark current components at 295 K for a HgCdTe SAM structure with the absorber at 1.7 μm, and the avalanche region at 1.1 μm.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

current on the n-side is less than this for the 1.1 μm cutoff wavelength chosen. These values compare very favorably with published data on InGaAs, with the added benefit of greatly reduced excess noise for the HgCdTe structure.

8.7 Conclusion

In this chapter, we have addressed a number of advanced-generation focal plane issues utilizing the HgCdTe separate absorption and detection (SAD) structure. The versatility of this concept is impressive, offering significant opportunities of improvement over the current state of the art. The implementation of this concept will be a challenge, in view of the doping concentrations and S-R lifetimes required, but the potential reward is enormous.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Epilogue

This book represents the collective thoughts of an Oxford University physicist who had the privilege of experiencing 35 years of research and development in infrared physics within the confines of the central research laboratories of one of the major corporate entities in this field, namely Texas Instruments. I have been fortunate enough to work with giants of the corporate R&D world, in an environment that was both challenging and invigorating, and at times even academic. The benefit of working in a corporate central research laboratory was to be found in the degree of objectivity that could be applied to any particular problem being addressed. The customer was not necessarily the in-house systems group who might benefit from the effort, or the government agency that might have been funding the program, but in reality the customer was truth. Truth that manifested itself in the physics of the problem at hand and the considerable brainpower that was available to address it. Politics, both in-house and external, played little or no role in the outcome. Unfortunately, this environment no longer exists in the USA today. The days of the corporate central research laboratory are gone, and with it much of the objectivity and originality of R&D. The majority of the current IR&D efforts in corporate America fall under the heading of D, at best, and are carried out from a jaundiced point of view, reflecting the prejudices of either in-house systems groups or external government agencies with built-in agendas. This modus operandi does not serve America well and never will. I have written this book with a view to extolling the merits of the IR scene. I am sure that my innate prejudices have played a role in the opinions rendered herein, but these opinions are honest and represent the truth as I see it and have not been influenced by anything other than intellectual reasoning.

Michael A. Kinch
Chief Scientist
DRS Infrared Technologies
Dallas, Texas

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix A: Mathcad Program for HgCdTe Diode Dark Current Modeling

This program models the dark current in a HgCdTe diode for various combinations of vacancies and extrinsic dopants.

We first calculate the x-value dependence on wavelength and temperature, and the corresponding bandgap E_g, after Hansen, Schmit, and Casselman.

Legend

λc = cutoff wavelength at a defined temperature T

T1 = operating temperature

d = thickness of diffusion volume

τ = lifetime of S-R centers

nV = vacancy concentration

na = extrinsic acceptor concentration

Er = S-R center relative to conduction band edge

$q := 1.6 \cdot 10^{-19} \qquad Eg(\lambda c) := \dfrac{1.24}{\lambda c} \qquad h := 6.6256 \cdot 10^{-34}$

$mo := 9 \cdot 10^{-31} \quad k := 1.4 \cdot 10^{-23}$

$\mathbf{A}(T) := 2.00376 - 1.2861 \cdot 10^{-3} \cdot T$

$B(T, \lambda c) := 0.32144 + 2.25665 \cdot 10^{-4} \cdot T - 1.2019Eg(\lambda c)$

$$\mathbf{C}(T, \lambda c) := \sqrt{\frac{B(T, \lambda c)^2}{4} + \frac{A(T)^3}{27}}$$

$$ARG1(T, \lambda c) := -\frac{B(T, \lambda c)}{2} + C(T, \lambda c)$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$\mathrm{ARG2(T, \lambda c) := -\frac{B(T, \lambda c)}{2} - C(T, \lambda c)}$$

$$\mathrm{z(T, \lambda c) := if(ARG1(T, \lambda c) < 0, -1, 1)}$$

$$\mathrm{ARG3(T, \lambda c) := z(T, \lambda c) \cdot (|ARG1(T, \lambda c)|)^{\frac{1}{3}}}$$

$$\mathbf{z}\mathrm{(T, \lambda c) := if(ARG2(T, \lambda c) < 0, -1, 1)}$$

$$\mathrm{ARG4(T, \lambda c) := z(T, \lambda c) \cdot (|ARG2(T, \lambda c)|)^{\frac{1}{3}}}$$

$$\mathrm{x(T, \lambda c) := ARG3(T, \lambda c) + ARG4(T, \lambda c) + 0.3245:}$$

$$\mathrm{Eg1(\lambda c, T, T1) := [0.832 \times (T, \lambda c)^3 - 0.81 \times (T, \lambda c)^2 + 5.34 \cdot 10^{-4}}$$

$$\mathrm{\cdot T1 \cdot (1 - 2 \cdot x(T, \lambda c)) + 1.93 \times (T, \lambda c)] - 0.302}$$

The corresponding intrinsic carrier concentration is given by

$$\mathrm{hh(\lambda c, T, T1) := Eg1(\lambda c, T, T1)^{0.75} \cdot T1^{1.5} \cdot exp\left(\frac{-Eg1(\lambda c, T, T1) \cdot q}{2 \cdot k \cdot T1}\right)}$$

$$\mathrm{ni(\lambda c, T, T1) := (5.585 - 3.82 \cdot \times(T, \lambda c) + 1.753 \cdot 10^{-3} \cdot T1 - 1.364 \cdot 10^{-3}}$$

$$\mathrm{\cdot T1 \cdot \times(T, \lambda c)) \cdot 10^{14} \cdot hh(\lambda c, T, T1)}$$

We now calculate various carrier properties of HgCdTe.

For p-type

The valence band density of states assuming a hole mass of $0.55m_o$ is

$$\mathrm{Nv(T1) := 2 \cdot 10^{-6} \cdot \left(\frac{2 \cdot \pi \cdot 0.55 \cdot mo \cdot k \cdot Ti}{h^2}\right)^{1.5}}$$

For graphing purposes let

$$\mathrm{na(y) := 10^y}$$

$$\mathrm{nV(x) := 10^x}$$

If the acceptors are always ionized, which is not too far from the truth for tactical temperatures

$$\mathrm{p(\lambda c, T, T1, y, x) := \frac{(na(y) + nV(x))}{2} \cdot \left[1 + \left[1 + \frac{4 \cdot ni(\lambda c, T, T1)^2}{(na(y) + nV(x))^2}\right]^{0.5}\right]}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$n(\lambda c, T, T1, y, x) := \frac{ni(\lambda c, T, T1)^2}{p(\lambda c, T, T1, y, x)}$$

In most of modeling below, we assume that we are dealing with reversed biased diode regions in which p $\sim$ na + NV, that is, we ignore n_i majority carriers even at high T.

For n-type

The conduction band effective mass is approximately given by

$$me(\lambda c, T, T1) := 7 \cdot 10^{-2} \cdot Eg1(\lambda c, T, T1)$$

Thus, the conduction band density of states is given by

$$Nc(\lambda c, T, T1) := 2 \cdot 10^{-6} \cdot \left(\frac{2 \cdot \pi \cdot me(\lambda c, T, T1) \cdot mo \cdot k \cdot T1}{h^2}\right)^{1.5}$$

Minority Carrier Lifetime Modeling

The arguments of Humphreys suggest that the radiative lifetime can be ignored for most circumstances.

$$c := 2.99792510^{10}$$

The total number of photons emitted between 1 and 100 μm from a blackbody of temperature $T1$:

$$Qtot(T1) := \int_1^{100} \frac{2 \cdot \pi \cdot c}{\lambda^4 \cdot 10^{-16}} \cdot \frac{1 \cdot 10^{-4}}{\exp\left(\frac{h \cdot c}{\lambda \cdot 10^{-4} \cdot k \cdot T1}\right) - 1} d\lambda$$

Photons emitted between 1 and λ_c

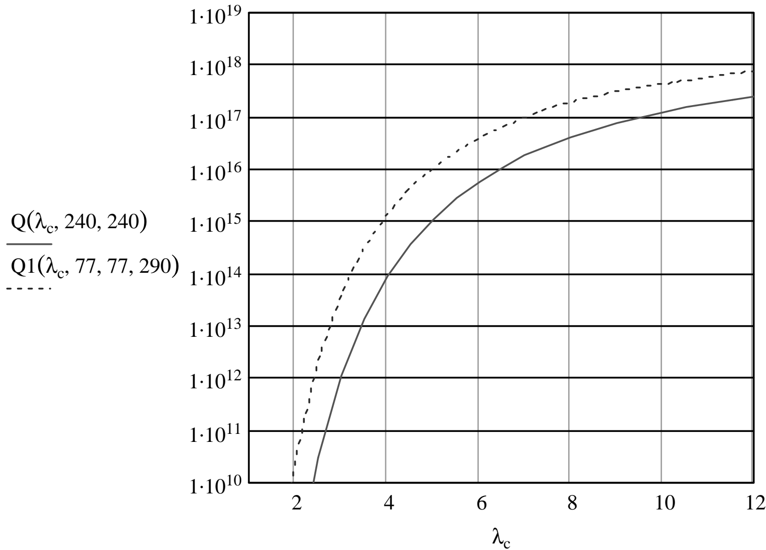
$$(\lambda c, T, T1) := \int_1^{\frac{1.24}{Eg1(\lambda c, T, T1)}} \frac{2 \cdot \pi \cdot c}{\lambda^4 \cdot 10^{-16}} \cdot \frac{1 \cdot 10^{-4}}{\exp\left(\frac{h \cdot c}{\lambda \cdot 10^{-4} \cdot k \cdot T1}\right) - 1} d\lambda$$

Total photons received by a detector at temperature T_1 from a blackbody at temperature T_2:

$$Q1(\lambda c, T, T1, T2) := \int_1^{\frac{1.24}{Eg1(\lambda c, T, T1)}} \frac{2 \cdot \pi \cdot c}{\lambda^4 \cdot 10^{-16}} \cdot \frac{1 \cdot 10^{-4}}{\exp\left(\frac{h \cdot c}{\lambda \cdot 10^{-4} \cdot k \cdot T2}\right) - 1} d\lambda$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$\lambda c := 1, 1.5..12$



The radiative lifetime, according to Humphreys, is given by

$$\tau r(\lambda c, T, T1, y, x, d, \eta) := \frac{ni(\lambda c, T, T1)^2 \cdot d}{p(\lambda c, T, T1, y, x) \cdot \eta \cdot Q(\lambda c, T, T1)}$$

where d is thickness of the diode and η is quantum efficiency.

Electron effective mass is given by

$$me(\lambda c, T, T1) := 7 \cdot 10^{-2} \cdot Eg1(\lambda c, T, T1)$$

The radiative lifetime according to van Roosbreck and Shockley is approximately given by

$$B(\lambda c, T, T1) := 5.8 \cdot 10^{-13} \cdot 14.0^{0.5} \cdot \left(\frac{1}{0.55}\right)^{1.5}$$

$$\cdot \left(1 + \frac{1}{me(\lambda c, T, T1)} + \frac{1}{0.55}\right)\left(\frac{300}{T1}\right)^{1.5} \cdot Eg1(\lambda c, T, T1)^2$$

$$\tau Ri(\lambda c, T, T1) := \frac{1}{2 \cdot B(\lambda c, T, T1) \cdot ni(\lambda c, T, T1)}$$

$$\tau rmic(\lambda c, T, T1, y, x) := [B(\lambda c, T, T1) \cdot (p(\lambda c, T, T1, y, x) + n(\lambda c, T, T1, y, x))]^{-1}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Auger Lifetime

The Auger1 lifetime in n-type material is given by numerical factor is $3.8e-18\varepsilon 2Eg3/2/[7e-2Eg] = 3.8e-18.4e2.1e2/7 = 2.2e-14$ if we use $\varepsilon = 20$

$FF := 0.47$

$$\tau Ai1(\lambda c, T, T1) := \frac{2.12 \cdot 10^{-14} \cdot \sqrt{Eg1(\lambda c, T, T1)} \cdot \exp\left(q \cdot \frac{Eg1(\lambda c, T, T1)}{K \cdot T1}\right)}{FF^4 \cdot \left(\frac{k \cdot T1}{q}\right)^{1.5}}$$

For p-type material, Beattie's model, and Casselman's, suggest that $\tau Ai7 \approx 6\tau Ai1$. Experiment suggests a larger factor than 6. Recent data [6/9/03] suggests the factor is as high as 12 for LWIR.

Consider

$$\tau Ai7c(\lambda c, T, T1) := 12 \cdot \tau Ai1(\lambda c, T, T1)$$

$$\tau A7c(\lambda c, T, T1, y, x) := \frac{2 \cdot \tau Ai7c(\lambda c, T, T1) \cdot ni(\lambda c, T, T1)^2}{p(\lambda c, T, T1, y, x)^2 + ni(\lambda c, T, T1)^2}$$

$$\tau A1(\lambda c, T, T1, y, x) := \frac{2 \cdot \tau Ai1(\lambda c, T, T1) \cdot ni(\lambda c, T, T1)^2}{n(\lambda c, T, T1, y, x)^2 + ni(\lambda c, T, T1)^2}$$

$$\tau Ac(\lambda c, T, T1, y, x) := \left(\frac{1}{\tau A7c(\lambda c, T, T1, y, x)} + \frac{1}{\tau A1(\lambda c, T, T1, y, x)}\right)^{-1}$$

Let us consider the various forms of possible S-R recombination. We have for S-R recombination, and few excess carriers,

$$\tau n = [\tau po(n + n1) + \tau no[p + p1 + Nrp1/(p + p1)]/[n + p + Nrpp1/(p + p1)2]$$

$$\tau p = [\tau no(p + p1) + \tau po[n + n1 + Nrp/(p + p1)]/[n + p + Nrpp1/(p + p1)2]$$

For an acceptor $\tau no \gg \tau po$, thus   For donors $\tau po \gg \tau no$

$\tau n = \tau no(p + p1 + Nrp1/(p + p1))/p$ $\tau n = \tau po(n + n1)/p$

$\tau p = \tau no(p + p1)/p$ $\tau p = \tau po(n + n1 + Nrp/(p + p1))/p$

Thus, we have four possibilities

1. Acceptor-like center; $p1 > p$, then

   $\tau n = \tau no(p1 + Nr)/p$ and $\tau p = \tau no(p1/p)$

2. Acceptor-like center; $p > p1$, then

   $\tau n = \tau no(p + Nrp1/p)/p$ and $\tau p = \tau no$.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

3. Donor-like center, n1 > n, then

$$\tau n = \tau po(n1/p) \quad \text{and} \quad \tau p = \tau po(n1 + Nr)/p$$

4. Donor-like center, and n > n1, then

$$\tau n = \tau po(n/p) \quad \text{and} \quad \tau p = \tau po(n + Nrp/p1)/p$$

and if pc lifetime is measured then

$$\tau pc = (\mu n \tau n + \mu p \tau p)/(\mu n + \mu p)$$

Acceptor-like center, $\tau no \gg \tau po$, or neutral center with $\tau no \sim \tau po$.

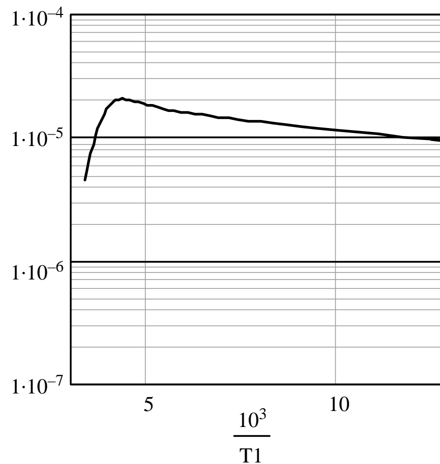
Assume a lifetime that is independent of doping from a low density of S-R centers, with τ as a variable parameter. Assume that we are in non-equilibrium so that majority carrier concentrations are given by background doping concentrations in n- and p-type regions.

$$\tau ni(\lambda c, T, T1, y, x, \tau) := \tau \cdot \frac{(2 \cdot ni(\lambda c, T, T1) + na(y) + nV(x))}{na(y) + nV(x)}$$

$$\tau in(\lambda c, T, T1, y, x, \tau) := \left(\frac{1}{\tau ni(\lambda c, T, T1, y, x, \tau)} + \frac{1}{\tau Ac(\lambda c, T, T1, y, x)}\right)^{-1}$$

$T1 := 30, 35..300$

$\tau in(5, 77, T1, 15.4, 10, 20 \cdot 10^{-6})$

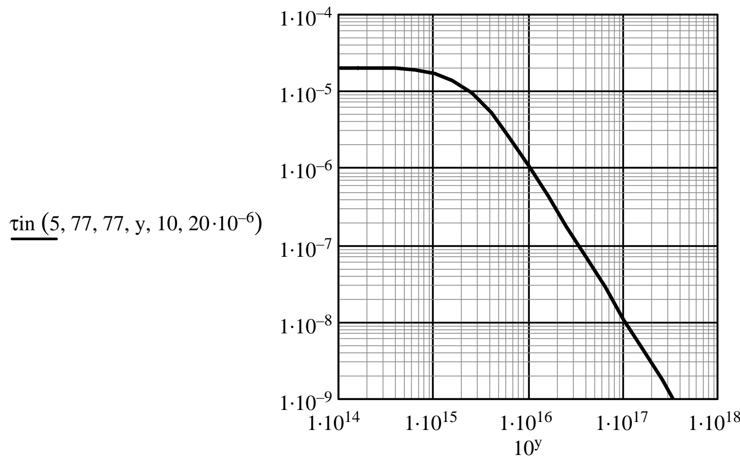


$y := 14, 14.2..18$

Donor-like center, $\tau po \gg \tau no$. This should represent the vacancy-doped case.

$$\mathbf{me}(\lambda c, T, T1) := 7 \cdot 10^{-2} \cdot Eg1(\lambda c, T, T1)$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use



$$\mathrm{Nc}(\lambda c, T, T1) := 2 \cdot 10^{-6} \cdot \left(\frac{2 \cdot \pi \cdot \mathrm{me}(\lambda c, T, T1) \cdot \mathrm{mo} \cdot k \cdot T1}{h^2}\right)^{1.5}$$

$$\mathrm{Efn}(\lambda c, T, T1, nd) := k \cdot \frac{T1}{q}. \ln\left(\frac{\mathrm{Nc}(\lambda c, T, T1)}{nd}\right)$$

If we are considering a donor-like center then we must ask ourselves the question as to whether the trap energy level is tied to the valence or conduction band. The activation energy for lifetime in vacancy-dominated $x = 0.2$ and $x = 0.3$ material appears to be similar and $Er = 0.03$ eV. This would suggest that it is tied to the conduction band.

$$n1(\lambda c, T, T1, Er) := \mathrm{Nc}(\lambda c, T, T1) \cdot \exp\left(-\frac{Er \cdot q}{k \cdot T1}\right)$$

$$\mathrm{Nv}(T1) := 2 \cdot 10^{-6} \left(\frac{2 \cdot \pi \cdot 0.6 \cdot \mathrm{mo} \cdot k \cdot T1}{h^2}\right)^{1.5}$$

$$p1(\lambda c, T, T1, Er) := \mathrm{Nv}(T1) \cdot \exp\left[\frac{-(\mathrm{Eg1}(\lambda c, T, T1) - Er) \cdot q}{k \cdot T1}\right]$$

In the following equations, we assume that reverse bias is applied and non-equilibrium applies within a diffusion length of the diode. Thus, as we approach intrinsic temperatures, the excess minority carriers are swept out and charge neutrality is preserved so p remains at na. The same argument applies to the earlier case.

$$\tau n2(\lambda c, T, T1, Er, y, x) := 5 \cdot 10^{-6} \cdot \frac{n1(\lambda c, T, T1, Er)}{na(y) + nV(x)} \cdot \frac{1 \cdot 10^{15}}{nV(x)}$$

$$N(y, x) := na(y) + nV(x)$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$\tau p2(\lambda c, T, T1, Er, y, x) := \frac{5.0 \cdot 10^9}{nV(x)} \cdot \frac{n1(\lambda c, T, T1, Er) + \frac{nV(x)}{16} \cdot \frac{N(y, x)}{N(y, x) + p1(\lambda c, T, T1, Er)}}{N(y, x)}$$

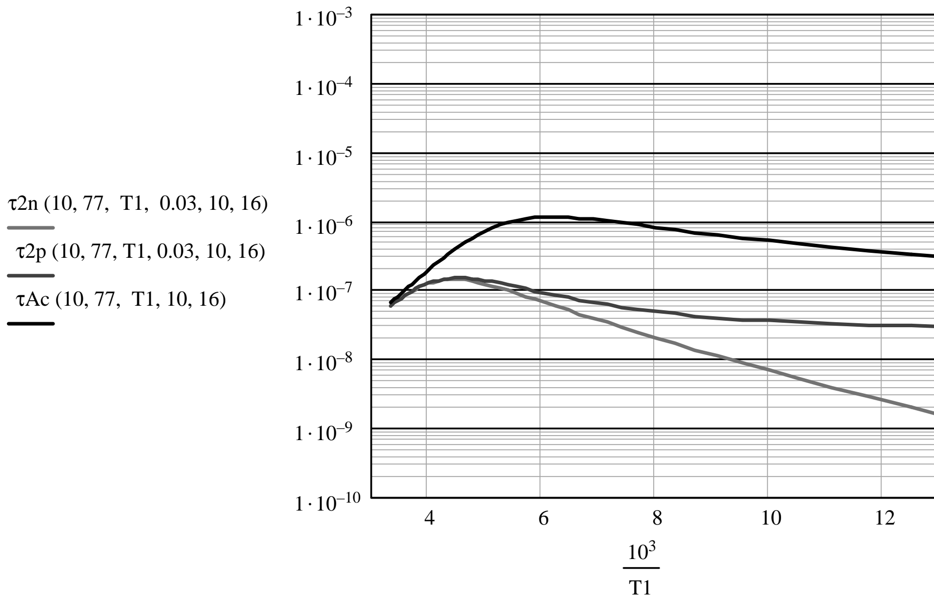
In the expression for τp2, we have assumed that the density of S-R centers is nV/16, which fits the experimental data for hole lifetimes in p-type material.

$$\tau 2p(\lambda c, T, T1, Er, y, x) := \left(\frac{1}{\tau Ac(\lambda c, T, T1, y, x)} + \frac{1}{\tau p2(\lambda c, T, T1, Er, y, x)}\right)^{-1}$$

$$\tau 2n(\lambda c, T, T1, Er, y, x) := \left(\frac{1}{\tau Ac(\lambda c, T, T1, y, x)} + \frac{1}{\tau n2(\lambda c, T, T1, Er, y, x)}\right)^{-1}$$

$$\tau n2(\lambda c, T, T1, Er, y, x) := \frac{n1(\lambda c, T, T1, Er)}{na(y) + nV(x)} \cdot \frac{5 \cdot 10^9}{nV(x)}$$

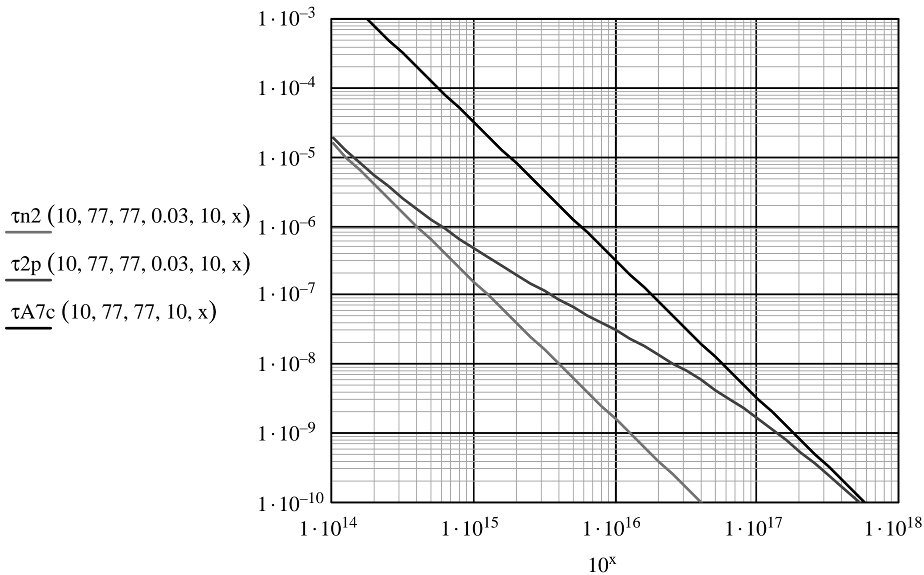
$$\tau ni(\lambda c, T, T1, y, x, \tau) := \tau \cdot \frac{(2 \cdot ni(\lambda c, T, T1) + na(y) + nV(x))}{na(y) + nV(x)}$$



x := 14, 14.1 . . . 18

We will ignore radiative from here on as it is irrelevant for the case of a well-designed unit cell.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use



Consider the possibilities for non-equilibrium diode operation. In the case of P-I-N we have for the intrinsic level S-R, assuming $\tau no = \tau po = \tau$

$$\mathrm{Jpin}(\lambda c, T, T1, \tau, w) := \frac{q \cdot w \cdot ni(\lambda c, T, T1)}{2 \cdot \tau}$$

If there is a region containing majority carriers, then the diffusion current components are given by $J = qn_i^2 d/N_{maj}^* \tau$.

Using the above lifetimes we have, on p-side, considering the individual components separately,

$$N(y, x) := na(y) + nV(x)$$

$$\mathrm{Jpsr}(\lambda c, T, T1, \tau, d, y, x) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2}{\tau \cdot (N(y, x) + 2 \cdot ni(\lambda c, T, T1))}$$

$$\mathrm{Jpvac}(\lambda c, T, T1, d, Er, x) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2 \cdot nV(x)}{n1(\lambda c, T, T1, Er) \cdot 5 \cdot 10^9}$$

$$\mathrm{JpA}(\lambda c, T, T1, d, y, x) := \frac{q \cdot d \cdot N(y, x)}{2 \cdot \tau Ai7c(\lambda c, T, T1)}$$

On the n-side, ignoring vacancies, the dark current components are

$$\mathrm{Jn}(\lambda c, T, T1, \tau, d, nd) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2}{\tau \cdot (nd + 2 \cdot ni(\lambda c, T, T1))} + \frac{151q \cdot d \cdot nd}{2 \cdot \tau Ai1(\lambda c, T, T1)}$$

$$nd(z) := 10^z$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

and the individual components again are

$$\mathrm{Jnsrl}(\lambda c, T, T1, d, \tau, z) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2}{\tau \cdot (10^z + 2 \cdot ni(\lambda c, T, T1))}$$

$$\mathrm{JnA1}(\lambda c, T, T1, d, z) := \frac{q \cdot d \cdot 10^z}{2 \cdot \tau Ai1(\lambda c, T, T1)}$$

Depletion current per unit volume from the n-side for a planar one-sided diode is given by the S-R expression above with nd = 0 and d = 1,

$$\mathrm{Wn}(\lambda c, T, T1, V, z) := \sqrt{\frac{3.6 \cdot 10^{-12} \cdot (Eg1(\lambda c, T, T1) + V)}{q \cdot 10^z}}$$

where we have assumed $\varepsilon \approx 18$. This is fine for LWIR and MWIR but needs to be modified for larger x-values. V is reverse bias, and we have assumed a built-in diode voltage $\sim$Eg, the bandgap.

$$\mathrm{Jdepn}(\lambda c, T, T1, \tau, V, z) := \frac{q \cdot ni(\lambda c, T, T1) \cdot Wn(\lambda c, T, T1, V, z)}{2 \cdot \tau}$$

If one-sided on the p-side then

$$\mathrm{Wp}(\lambda c, T, T1, V, y, x) := \sqrt{\frac{3.6 \cdot 10^{-12} \cdot (Eg1(\lambda c, T, T1) + V)}{q \cdot N(y, x)}}$$

$$\mathrm{Jdepp}(\lambda c, T, T1, \tau, V, y, x) := \frac{q \cdot Wp(\lambda c, T, T1, V, y, x) \cdot ni(\lambda c, T, T1)}{\tau}$$

Tunnel breakdown in HgCdTe

$$m(Eg) := 7 \cdot 10^{-2} \cdot Eg \quad nd(Z) := 10^z$$

$$\mathrm{Jtvdm}(V, Eg, N, f) := 4.4 \cdot 10^{-2} \cdot \left(\frac{m(Eg)}{Eg}\right)^{0.5} \cdot N^{0.5} \cdot (V) \cdot (Eg + V)^{0.5} \cdot \exp\left[\frac{-f \cdot 9.433\ 10^{10} \cdot m(Eg)^{0.5} \cdot Eg^{1.5}}{N^{0.5} \cdot (Eg + V)^{0.5}}\right]$$

$$\mathrm{Jtvsrm}(V, Eg, z, Nr, f) := \frac{10^{-13} \cdot Nr \cdot (V)}{Eg} \cdot \exp\left[\frac{-f \cdot 3 \cdot 10^{10} \cdot m(Eg)^{0.5} \cdot Eg^{1.5} \cdot \pi}{2 \cdot nd(z)^{0.5} \cdot (Eg + V)^{0.5}}\right]$$

where f is the fudge factor, which equals unity for a diode with a uniform electric field. For other than a uniform field, the factor will lie between 1 and 2, to account

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

for the fact that the electric field is a decreasing function as you move away from the junction n–p interface. For the cylindrical diode that represents the HDVIP case, then a uniform field is a reasonable approximation.

The following situations are the separated n- and p-sides of a junction, for S-R, Auger, depletion, and tunnel currents. For the purposes of graphical display of the individual dark current components on the same plot vs. doping concentration, we will use 10^z as a doping concentration for donors, vacancies, and extrinsic acceptors in the above expressions for the various dark current components.

Thus, we have

$$N(z) := 10^z$$

$$Jpsr(\lambda c, T, T1, \tau, d, z) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2}{\tau \cdot [N(z) + 2 \cdot ni(\lambda c, T, T1)]}$$

where for this component z refers to the sum of vacancies and extrinsic acceptors

$$Jpvac(\lambda c, T, T1, d, Er, z) := \frac{q \cdot d \cdot ni(\lambda c, T, T1)^2 \cdot nV(z)}{n1(\lambda c, T, T1, Er) \cdot 5 \cdot 10^9}$$

where for this component z refers to the vacancy concentration.

$$JpA(\lambda c, T, T1, d, z) := \frac{q \cdot d \cdot N(z)}{2 \cdot \tau Ai7c(\lambda c, T, T1)}$$

where for this component z refers to the sum of vacancies and extrinsic acceptors.

We need to consider that the depletion region will be entirely on the n-side of the junction and will consume an increasing portion of the n-volume as the n-side doping decreases. To accommodate this, we introduce the condition that the thickness available for diffusion current is d-W. We also have to limit the depletion region width to a maximum value of d, which we do by using the “if” condition.

$$W1(\lambda c, T, T1, V, z, d) := if(Wn(\lambda c, T, T1, V, z) < d, Wn(\lambda c, T, T1, V, z), d)$$

$$Jnsr(\lambda c, T, T1, V, d, \tau, z) := \frac{q \cdot (d - W1(\lambda c, T, T1, V, z, d)) \cdot ni(\lambda c, T, T1)^2}{\tau \cdot (10^z + 2 \cdot ni(\lambda c, T, T1))}$$

$$JnA(\lambda c, T, T1, V, d, z) := \frac{q \cdot (d - W1(\lambda c, T, T1, V, z, d)) \cdot 10^z}{2 \cdot \tau Ai1(\lambda c, T, T1)}$$

and the depletion component

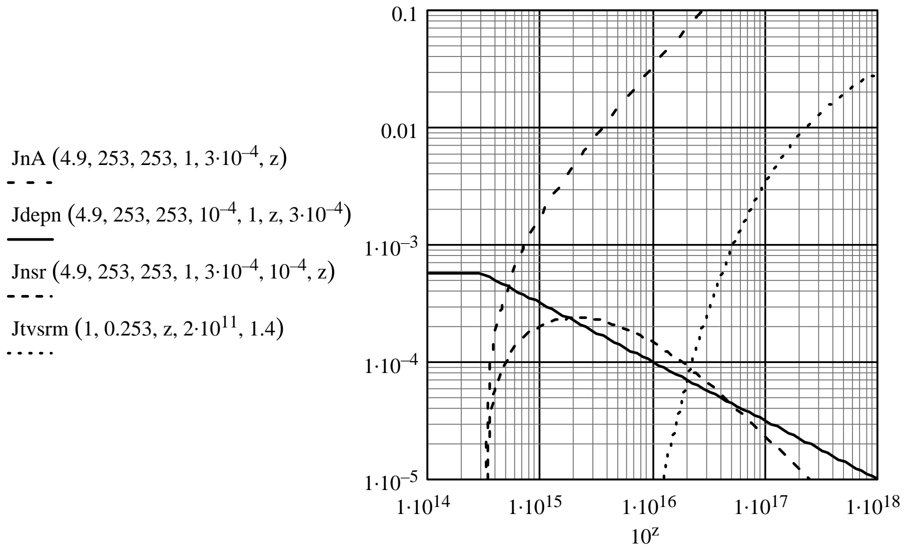
$$Jdepn(\lambda c, T, T1, \tau, V, z, d) := \frac{q \cdot ni(\lambda c, T, T1) \cdot W1(\lambda c, T, T1, V, z, d)}{2\tau}$$

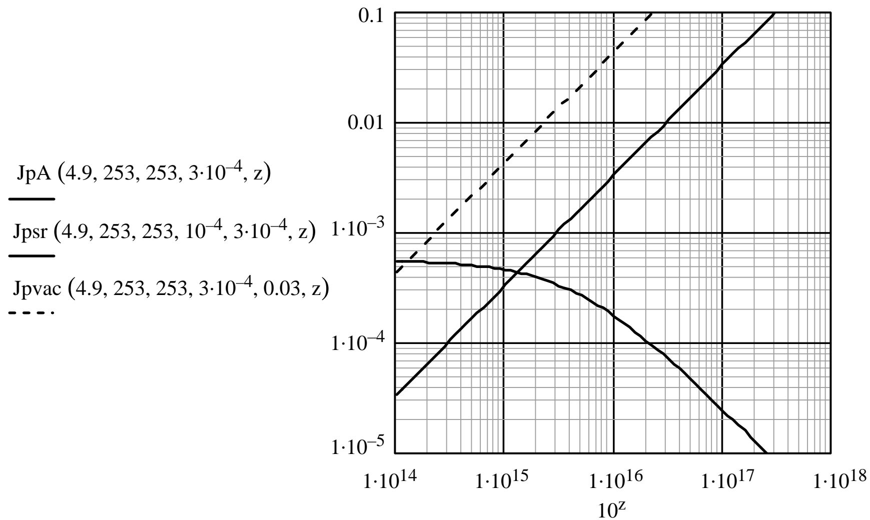
The expressions for the tunnel component remain the same.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Consider the case of a planar junction with n-side = p-side = 3μm thick. Use f = 1.4 for this planar diode.

z := 14, 14.05..18

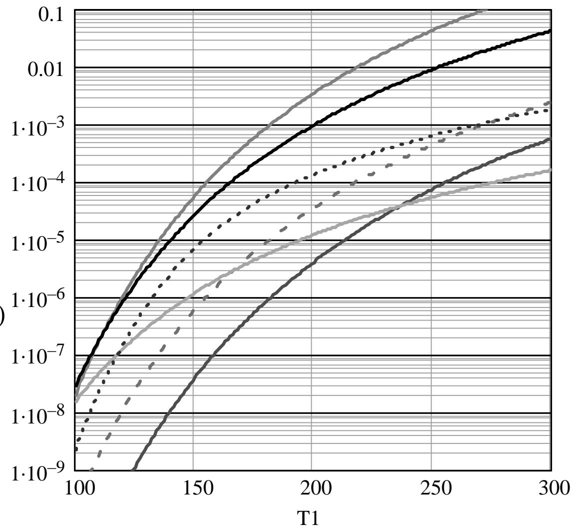




The variation of these dark current components with temperature is given by

T1 := 100, 101..300

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use



Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

References

1. W.D. Lawson, S. Nielsen, E.H. Putley and A.S. Young, *J. Phys. Chem. Solids*, vol. 9, pp. 325, 1959.
2. N.K. Abrikosov, K.A. Duldina and F.A. Dasilyan, *Zh. Neorgan. Khim.*, vol. 3, pp. 1632, 1958.
3. D. Long, *Infrared Phys.*, vol. 7, pp. 169, 1967.
4. R.W. Brodersen and S.P. Emmons, *Proc. Int. Conf. Applications of CCDs*, San Diego, CA, Naval Electronics Laboratory Center, 1975.
5. G.L. Hansen and J.L. Schmit, *J. Appl. Phys.,* vol. 54, pp. 1639, 1983.
6. M.A. Kinch, M.J. Brau and A. Simmons, *J. Appl. Phys.*, vol. 44, pp. 1649, 1973.
7. F. Stern, "Solid State Physics," F. Seitz and D. Turnbull (eds.), vol. 15, pp. 366, Academic Press, New York, 1963.
8. M.A. Kinch, S.R. Borrello and A. Simmons, *Infrared Phys.*, vol. 17, pp. 127, 1977.
9. R.A. Smith, "Semiconductors," pp. 90, Cambridge University Press, 1968.
10. M.M. Blouke, C.B. Burgett and R.L. Williams, *Infrared Phys.*, vol. 13, pp. 61, 1973.
11. N. Sclar, *Infrared Phys.*, vol. 16, pp. 435, 1976.
12. Y. Darviot, A. Sorrentino and B. Joly, *Infrared Phys.*, vol. 7, pp. 1, 1967.
13. B.K. Ridley, *J. Phys.*, vol. C15, pp. 5899, 1982.
14. R. Stratton, *Proc. Roy. Soc.*, vol. A246, pp. 406, 1957.
15. B.F. Levine, C.G. Bethea, G. Hasnain, J. Walker and R.J. Malik, *Appl. Phys. Lett.*, vol. 53, pp. 296, 1988.
16. J. Phillips, *J. Appl. Phys.*, vol. 91, pp. 4590, 2002.
17. D.E. Mercer and C.R. Helms, *J. Appl. Phys.*, vol. 65, pp. 5035, 1989.
18. J.J. Quinn, *Phys. Rev.*, vol. 126, pp. 1453, 1962.
19. J.M. Mooney, *J. Appl. Phys.*, vol. 64, pp. 4664, 1988.
20. W.H. Parker and W.D. Williams, *Phys. Rev. Lett.*, vol. 29, pp. 924, 1972.
21. A. Rothwarf and B.N. Taylor, *Phys. Rev. Lett.*, vol. 19, pp. 27, 1967.
22. W. van Roosbroeck and W. Shockley, *Phys. Rev.*, vol. 94, pp. 1558, 1954.
23. R.G. Humphreys, *Infrared Phys.*, vol. 26, pp. 337, 1986.
24. T.N. Casselman and S. Krishnamurthy, *J. Electron. Mater.*, vol. 29, pp. 828, 2000.
25. A.R. Beattie, private communication.
26. M.Y. Pines and O.M. Stafsudd, *Infrared Phys.*, vol. 20, pp. 73, 1980.
27. S.R. Jost, V.F. Meikleham and T.H. Myers, *MRS Symp. Proc.*, vol. 90, pp. 429, 1987.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

28. M.A. Blessinger, R.C. Fischer, C.J. Martin, C.A. Niblack, H.A. Timlin and G. Finger, *Proc. SPIE*, vol. 1308, pp. 194, 1990.
29. J.T. Wimmers, R.M. Davis, C.A. Niblack and D.S. Smith, *Proc. SPIE*, vol. 930, pp. 125, 1988.
30. D. Long and J.L. Schmit, in *Semiconductors and Semimetals*, R.K. Willardson and A.C. Beer (eds.), vol. 5, pp. 175, Academic Press, New York, 1970.
31. M.C. Chen, L. Colombo, J.A. Dodge and J.H. Tregilgas, *J. Electron. Mater*., vol. 24, pp. 539, 1995.
32. P. Mitra, T.R. Schimert, F.C. Case, S.L. Barnes, M.B. Reine, R. Starr, M.H. Weiler and M. Kestigian, *J. Electron. Mater*., vol. 24, pp. 1077, 1995.
33. M.J. Bevan, M.C. Chen and H.D. Shih, *Appl. Phys. Lett*., vol. 67, pp. 3450, 1995.
34. R.T. Bate, D.L. Carter and J.S. Wrobel, *Phys. Rev. Lett*., vol. 25, pp. 159, 1970.
35. F.A. Junga, K.F. Cuff, J.S. Blakemore and E.R. Washwell, *Physics Letters*, vol. 13, pp. 103, 1964.
36. I.Melngailis and T.C. Harman, in *Semiconductors and Semimetals*, R.K. Willardson and A.C. Beer (eds.), vol. 5, pp. 111, Academic Press, New York, 1970.
37. M.R. Johnson, R.A. Chapman and J.S. Wrobel, *Infrared Phys*., vol. 15, pp. 317, 1975.
38. I. Melngailis and T.C. Harman, *Appl. Phys. Lett*., vol. 13, pp. 180, 1968.
39. J.N. Schulman and T.C. McGill, *Appl. Phys. Lett*., vol. 34, pp. 663, 1979.
40. G. Bastard, *Phys. Rev*., vol. B24, pp. 5693, 1981.
41. G. Bastard, *Phys. Rev*., vol. B25, pp. 7584, 1982.
42. D.L. Smith, T.C. McGill and J.N. Schulman, *Appl. Phys. Lett*., vol. 43, pp. 180, 1983.
43. P.M. Hui, H. Ehrenreich and N.F. Johnson, *J. Vac. Sci. Technol*., vol. A7, pp. 424, 1989.
44. D.L. Smith, *Solid State Comm*., vol. 57, pp. 919, 1986.
45. D.L. Smith and C. Mailhiot, *Phys. Rev. Lett*., vol. 58, pp. 1264, 1987.
46. M.A. Mattson, T.H. Myers, M. Richards-Babb and J.R. Meyer, *J. Electron. Mater*., vol. 26, pp. 578, 1997.
47. J.W. Matthews and A.E. Blakeslee, *J. Cryst. Growth*, vol. 27, pp. 118, 1974; vol. 29, pp. 273, 1975; vol. 32, pp. 265, 1976.

47a. M.A. Kinch and M.W. Goodwin, *J. Appl. Phys.*, vol. 58, pp. 4455, 1985.

47b. E.O. Kane, in Semiconductors and Semimetals, R.K. Willardson and A.C. Beer (eds), vol. 1, pp. 91, Academic Press, New York, 1966.

48. K.A. Harris, R.W. Yanka, L.M. Mohnkern, A.R. Reisinger, T.H. Myers, Z. Yang, Z. Yu, S. Hwang and J.F. Schetzina, *J. Vac. Sci. Technol*., vol. B10, pp. 1574, 1992.
49. K.A. Harris, T.H. Myers, R.W. Yanka, L.M. Mohnkern and N. Otsuka, *J. Vac. Sci. Technol*., vol. B9, pp. 1752, 1991.
50. R.J. Koestner, M.W. Goodwin and H.F. Schaake, *J. Vac. Sci. Technol*., vol. B9, pp. 1731, 1991.
51. D.L. Smith and C. Mailhiot, *J. Appl. Phys*., vol. 62, pp. 2545, 1987.
52. C.H. Grein, P.M. Young, M.E. Flatte and H. Ehrenreich, *J. Appl. Phys*., vol. 78, pp. 7143, 1995.
53. C.H. Grein, P.M. Young and H. Ehrenreich, *Appl. Phys. Lett*., vol. 61, pp. 2905, 1992.
54. C. Mailhiot and D.L. Smith, *J. Vac. Sci. Technol*., vol. B5, pp. 1268, 1987.
55. E.R. Youngdale, J.R. Meyer, C.A. Hoffman, F.J. Bartoli, C.H. Grein, P.M. Young, H. Ehrenreich, R.H. Miles and D.H. Chow, *Appl. Phys. Lett*., vol. 64, pp. 3160, 1994.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

56. J.R. Meyer, C.L. Felix, W.W. Bewley, I. Vurgaftman, E.H. Aifer, L.J. Olafson, J.R. Lindle, C.A. Hoffman, M.-J. Yang, B.R. Bennett, B.V. Shanabrook, H. Lee, C.-H. Lin, S.S. Pei and R.H. Miles, *Appl. Phys. Lett*., vol. 73, pp. 2857, 1998.
57. H. Zogg, W. Vogt and H. Melchior, *Infrared Phys*., vol. 25, pp. 333, 1985.
58. D. Chandra, H. Schaake, J. Tregilgas, F. Aqariden, M. Kinch and A. Syllaios, *J. Electron. Mater*., vol. 29, pp. 729, 2000.
59. G.L. Destefanis, *J. Crystal Growth*, vol. 86, pp. 700, 1988.
60. D. Chandra, H.F. Schaake and M.A. Kinch, *Proc. 2002 U.S. Workshop on Phys. and Chem. Of II-VI Materials*.
61. M.A. Kinch and D.D. Buss, *J. Phys. Chem. Solids Suppl*., vol. 32, pp. 461, 1971.
62. G.L. Hansen, J.L. Schmit and T.N. Casselman, *J. Appl. Phys*., vol. 53, pp. 7099, 1982.
63. J. Antoszewski, D.J. Seymour, L. Faraone, J.R. Meyer and C.A. Hoffman, *J. Electron. Mater*., vol. 24, pp. 1255, 1995.
64. T. Wong, M.Sc. Thesis, MIT, 1974.
65. R.A. Smith, *Semiconductors*, Cambridge University Press, pp. 106, 1968.
66. R. Stratton, *Proc. Roy. Soc*., vol. A246, pp. 406, 1958.
67. S. Krishnamurthy, A. Sher and A.-B. Chen, *Appl. Phys. Lett*., vol. 47, pp. 160, 1985.
68. H. Brooks, *Adv. In El. And El. Phys*., vol. 7, pp. 85, 1955.
69. C. Erginsoy, *Phys. Rev*., vol. 79, pp. 1013, 1950.
70. B. Segall, M.R. Lorentz and R.E. Halstead, *Phys. Rev*., vol. 129, pp. 2471, 1963.
71. J.G. Collins, G.K. White, J.A. Birch and T.F. Smith, *J. Phys*., vol. C13, pp. 1649, 1980.
72. T.F. Smith and G.K. White, *J. Phys*., vol. C8, pp. 2031, 1975.
73. I.M. Baker and C.D. Maxey, *J. Electron. Mater*., vol. 30, pp. 682, 2001.
74. M.A. Kinch, *Properties of Narrow Gap Cadmium-based Compounds*, P. Capper (ed.) INSPEC, IEE, London, UK, pp. 291, 1994.
75. R.A. Schiebel, *IEEE Trans. Electron Devices*, vol. ED-41, pp. 768, 1994.
76. A.L. McWhorter, *Semiconductor Surface Physics*. Univ. Pennsylvania Press, Philadelphia, pp. 207, 1957.
77. W.W. Anderson and H.J. Hoffman, *J. Vac. Sci. Technol*., vol. A3, pp. 183, 1985.
78. S.M. Johnson, D.R. Rhiger, J.P. Rosbeck, J.M. Peterson, S.M. Taylor and M.E. Boyd, *J. Vac. Sci. Technol*., vol. B10, pp. 1499, 1992.
79. M.A. Kinch, C.-F. Wan and J.D. Beck, *J. Electron. Mater*., vol. 34, pp. 928, 2005.
80. C.T. Elliott and T. Ashley, *Electronics Letters*, vol. 21, pp. 451, 1985.
81. M.A. Kinch, M.J. Brau and A. Simmons, *J. Appl. Phys*., vol. 44, pp. 1649, 1973.
82. T.N. Casselman and S. Krishnamurthy, *J. Electron. Mater*., vol. 29, pp. 828, 2000.
83. C.D. Maxey, et al., *Proc. SPIE*, vol. 3122, pp. 453, 1997.
84. M. Bevan, *Appl. Phys. Lett*., vol. 67, pp. 3650, 1995.
85. R.S. Hall, N.T. Gordon, D.C. Herbert, A.M. White, C.D. Maxey, C.L. Jones, M. Ahmed, N.E. Metcalfe, A. Scholes and R.A. Catchpole, 2000 US II-VI Workshop Abstracts, pp. 179.
86. C. Mailhiot and D.L. Smith, *J. Vac. Sci. Technol*., vol. B, pp. 1268, 1987.
87. P.A. Wolff, *Phys. Rev*., vol. 95, pp. 1415, 1954.
88. W. Shockley, *Solid State Electron.*, vol. 2, pp. 35, 1961.
89. G.A. Baraff, *Phys. Rev*., vol. 128, pp. 2507, 1962.
90. M.V. Fischetti, N. Sano, S.E. Laux and K. Natori, www.ieee.org/products/online/journal/tcad, 1997.
91. R.J. McIntyre, *IEEE Trans on Electron Devices ED-13*, pp. 164, 1966.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

92. R.J. McIntyre, *IEEE Trans. Electron Devices*, vol. 46, pp. 1623, 1999.
93. J.D. Beck, C.-F. Wan, M.A. Kinch and J.E. Robinson, *SPIE Proc.*, vol. 4454, pp. 188, 2001.
94. M.A. Kinch, C.-F. Wan and J.D. Beck, unpublished data.
95. P.M. Hui, H. Ehrenreich and N.F. Johnson, *J. Vac. Sci. Technology*, vol. A7, pp. 424, 1989.
95a. C. Grein, private communication, 2002.
95b. R.D. Graft, *J. Vac. Sci. Technol.*, vol. 21, pp. 146, 1982.
96. A.R. Beattie and P.T. Landsberg, *Proc. Roy. Soc.*, vol. A249, pp. 16, 1958.
97. E. Antoncik and J. Tauc, in *Semiconductors and Semimetals* Willardson, R.K. and Beer, A.C. (eds), Academic Press, New York, vol. 2, pp. 245, 1970.
98. A.R. Beattie and P.T. Landsberg, *Proc. Roy. Soc.,* vol. 258A, pp. 486, 1960.
99. B. Orsal, R. Alabedra, M. Valenza, G. Pichard and J. Meslage, *J. Crystal Growth*, vol. 72, pp. 496, 1985.
100. M. Jack, et al., *Proc. SPIE*, vol. 4454, pp. 198, 2001.
101. E.O. Kane, *Phys. Rev.*, vol. 159, pp. 624, 1967.
102. J.D. Beck, Private communication, 2003.
103. C.T. Elliott, N.T. Gordon, R.S. Hall and C. Grimes, *J. Vac. Sci. Technol.*, vol. A8, pp. 1251, 1990.
104. A.-B. Chen and A. Sher, *J. Vac. Sci. Technol.*, vol. 21, pp. 138, 1982.
105. R.D. Graft, *J. Vac. Sci. Technol.*, vol. 21, pp. 146, 1982.
106. M.A. Berding, S. Krishnamurthy, A. Sher and A.-B. Chen, *J. Cryst. Growth*, vol. 86, pp. 33, 1988.
107. J.D. Beck, C.-F. Wan, M.A. Kinch, J.E. Robinson, P. Mitra, R.E. Scritchfield, F. Ma and J.C. Campbell, *Proc. SPIE*, vol. 5564, pp. 44, 2004.
108. F. Ma, X. Li, J.C. Campbell, J.D. Beck, C.-F. Wan and M.A. Kinch, *Appl. Phys. Lett.*, vol. 83, pp. 785, 2003.
109. F. Ma, *Monte Carlo Simulation of Gain, Noise and Speed of Low Noise and High Speed Photodiodes*, Ph. D. Thesis, University of Texas, Austin, August 2003.
110. M.A. Kinch, D. Chandra, H.F. Schaake, H.-D. Shih and F. Aqariden, *J. Electron. Mater.*, vol. 33, pp. 590, 2004.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Index

absorption coefficient, 49–51
absorption quantum efficiency, 7
α-silicon thermal detector technologies, 11
annealing, 65–66
antenna structures, 22
APD. *See* avalanche photodiodes.
Auger lifetime, 155–163
Auger minority carrier lifetime properties, 73–74
Auger recombination, 33–34
Auger transitions, 56
Avalanche photodiodes (APDs), 109
  high-gain, 147–148
  theory, 109

Background limited performance (BLIP), 6, 93, 143
band offsets and strain, 47–48
band structure, 67
bandgap engineering, 19
bandgap materials, direct, 57
bandgap states, tunneling via, 84–85
binary compounds, 35–37
  indium antimonide, 35–37
BLIP. *See* background limited performance.
bump-bonded ion implant architecture, 77

coefficient of thermal expansion (CTE), 74–75
CTE. *See* coefficient of thermal expansion.
cutoff wavelengths, 144–145

dark current
  HgCdTe diodes and, 82–86
  Mathcad program modeling, 151–163
    Auger lifetime, 155–163
    minority carrier lifetime modeling, 153–154
    n-type, 153
    p-type, 152–153
  model
    diffusion current, 136
    infrared devices and future development of, 135–139
      N-side, 136–137
      P-side, 137–139
  PbSnTe direct bandgap alloy and, 43–44
  total, 85–86
dead voltage model, 121–125
depletion current, 84
detector contacts, HOT theory and, 103
detector data, HOT theory and, 101–103
detector performance, 82
diffusion current, 83, 136, 137
diode dark current models, 34–35
direct bandgap materials, 57–59
  HgCdTe, 57
  InSb, 57–58
  PbSnTe, 58–59
  superlattices
    type II, 59
    type III, 59
direct tunneling, 85
dislocation density, 74–75
DLHJ diodes. *See* double layer heterojunction diodes.
doping, *p*-type mercury cadmium telluride and, 40–41

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

double layer heterojunction diodes (DLHJ), 38
architecture, 76–77
double-layer p^+/n heterojunctions. *See* double layer heterojunction diodes.
DRS manufacturing, 89–91
HDVIP technologies, 89
VIP technologies, 89

EAPDs. *See* electron avalanche photodiodes.
effective mass, 51–52
electron avalanche photodiodes, 109
electron impact ionization rate, 115–120
Kronig-Penney model, 118
electron transport, HgCdTe and, 69–70
extrinsic semiconductor, 16–18
extrinsically doped *p*-type mercury cadmium telluride, 40

Focal plane array (FPA), 2
high density, 143
multicolor, 141–143
multispectral, 141–143
FPA. *See* focal plane array.

Glow discharge mass spectroscopy, 63

HDVIP™ architecture, 77–81
liquid phase epitaxial material growth, 79
technologies, 89
HgCdTe annealing, 65–66
HgCdTe avalanche photodiodes, room-temperature, 129–131
HgCdTe device architecture, 75–81
bump-bonded ion inplant, 77
DLHJ, 76–77
vertically integrated photodiode, 77–81
HgCdTe diode dark current, Mathcad program modeling, 151–163
Auger lifetime, 155–163
minority carrier lifetime modeling, 153–154
n-type, 153
p-type, 152–153
HgCdTe diodes, dark current and, 82–86
sources of,
depletion current, 84
diffusion current, 83
surface generation current, 84
total dark current, 85–86
tunneling, 84–85
HgCdTe electron avalanche photodiodes (EAPDs), 109–137
APD theory, 109
empirical model for, 121–129
dead voltage model, 121–125
McIntyre's model, 110–112
Monte Carlo modeling, 131–133
physics of, 112–120
electron impact ionization rate, 115–120
high energy scattering rates, 113–115
HgCdTe materials, 61–91
detector performance, 89–91
glow discharge mass spectroscopy, 63
growth of, 61–65
Liquid phase epitaxy (LPE), 61, 62–64
Metal–organic chemical vapor deposition (MOCVD), 61, 65
Metal–organic vapor phase epitaxy (MOVPE), 65
Molecular beam epitaxy (MBE), 61, 64–65
physical properties, 74–75
coefficient of thermal expansion, 74–75
dislocation density, 74–75
lattice constant, 74
properties of, 61–75
band structure, 67
minority carrier lifetime properties, 73–74
Auger, 73–74
radiative, 73–74
Shockley–Read, 73–74
optical properties, 67
transport, 67–73
electron transport, 69–70
hole transport, 71–73
ROIC requirements, 81–89

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

HgTe/CdTe superlattices, 48
interdiffusion, 48
high density FPA, 143
high energy scattering rates, 113–115
high gain APDs, 147–148
High operating temperature (HOT) photon detection, 95–101
detector contacts, 103
detector data, 101–103
options, 103–105
high temperature operations, infrared devices and, 145–148
high-gain APDs, 147–148
High temperature superconductor (HTC), 26–27
hole transport, HgCdTe and, 71–73
HOT. *See* high operating temperature.
HTC. *See* high temperature superconductor.

impact ionization rate, 115–120
Indium antimonide (InSb), 35–37
IR detection material, 57–58
Infrared (IR) detectors, 1–3
absorption quantum efficiency, 7
DRS manufacturing, 89–91
early types, 1
first generation models, 1
focal plane array (FPA), 2
future development of, 135–148
dark current model, 135–139
high-density FPA, 143
higher operating temperatures, 145–148
low background operation, 143
multicolor FPA, 141–143
multispectral FPA, 141–143
separate absorption and detection structure, 139–141
required properties, 135
future models, 2
performance criteria, 5–11, 89–91
photon type, 5–9, 13–29
scanning arrays, 2
second generation models, 1–2
sensitivity of, 7–9
noise equivalent irradiance, 7
signal flux, 7
thermal detectors, 9–11
time, delay and integrate (TDI) techniques, 2
InSb. *See* indium antimonide.
interdiffusion, HgTe/CeTe superlattices and, 48
intrinsic direct bandgap semiconductor, 13–16, 27, 31–59
binary compounds, 35–37
diode dark current models, 34–35
materials, HgCdTe, 57
InSb, 57–58
PbSnTe, 58–59
type II superlattices, 59
minority carrier lifetime, 32–34
PbSnTe, 42–44
ternary alloys, 37–42
type II superlattices, 53–57
type III superlattices, 45–53, 59
IR. *See* infrared.
ionization rate, 115–120

Kronig–Penney model, 118

lattice constant, 74
Liquid phase epitaxy (LPE), 61, 62–64, 79
low background operation, 143
LWIR 14 μm at 40 K, 143–144
25 μm cutoff wavelength, 144–145
LPE. *See* liquid phase epitaxy.
LWIR
14 μm at 40 K, 143–144
minority carrier lifetimes and, 56
spectral bands, 37

majority carrier, 5–7
detection, photoconductive, 5
Mathcad programming modeling for dark current, 151–163
MBE. *See* molecular beam epitaxy.
McIntyre's avalanche photodiode model, 110–112
mercury cadmium telluride, 37–41
double layer heterojunction diodes, 38
LWIR spectral bands, 37

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

mercury cadmium telluride (*Continued*)
MWIR spectral bands, 37
n-type, 37–39
p-type, 39–41
Metal organic chemical vapor deposition (MOCVD), 61
Metal organic vapor phase epitaxy (MOVPE), 65
microlenses, 22
minority carrier devices, 5–7
detection, responsivity and noise, 5
lifetime, 32–34
Auger
recombination, 33–34
transitions, 56
LWIR, 56
MWIR, 56
PbSnTe direct bandgap alloy and, 43–44
properties
Auger, 73–74
HgCdTe and, 73–74
radiative, 73–74
Shockley–Read, 73–74
radiative recombination, 32–33
Shockley–Read recombination, 34
type II superlattices and, 54–57
type III superlattices and, 52–53
misfit dislocations, 48–49
MOCVD. *See* metal organic chemical vapor deposition.
MOVPE. *See* metal organic vapor phase epitaxy.
modeling, Mathcad program for dark current, 151–163
Molecular beam epitaxy (MBE), 61, 64–65
Monte Carlo modeling, 131–133
multicolor FPA, 141–143
multispectral FPA, 141–143
MWIR
minority carrier lifetimes and, 56
spectral bands, 37

NEI. *See* noise equivalent irradiance.
NEΔT. *See* noise equivalent temperature.
Noise equivalent irradiance (NEI), 7
Noise equivalent temperature (NEΔT), 9
noise, minority carrier detection and, 5
N-side, dark current model and, 136–137
N type
dark current Mathcad program modeling, 153
mercury cadmium telluride and, 37–39
$1/f$ noise, 87–89

operating temperatures, infrared devices and, 145–148
optical properties, 67

P type dark current Mathcad program modeling, 152–153
PbSnTe direct bandgap alloy, 42–44
dark currents, 44
as IR detection material, 58–59
minority carrier lifetime, 43–44
photon conductive detection, 5
photon detectors, 5–9, 95–105
background limited performance, 6
extrinsic semiconductor, 16–18
high temperature superconductor, 26–27
HOT theory, 95–101
detector
contacts, 103
data, 101–103
intrinsic direct bandgap semiconductor, 13–16
majority carrier, 5–7
minority carrier, 5–7
quantum well types, 18–23
silicon Schottky barrier type, 23–25
thermal generation rate, 6–7
photon uncooled detection vs. thermal detection limits, 105–107
P-side, dark current model and, 137–139
p-type mercury cadmium telluride, 39–41
extrinsically doped *p*-type, 40
general doping, 40–41
vacancy-doped *p*-type, 39–40

Quantum well IR photodetectors (QWIPs), 18–23
bandgap engineering, 19

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

superlattice, 18
thermal generation rate, 22
QWIPs. *See* quantum well IR photodetectors.

radiative lifetime, 153–154
radiative minority carrier lifetime properties, 73–74
radiative recombination, 32–33
resonant structures, 22
responsivity, minority carrier detection and, 5
ROIC requirements, 81–89
detector performance, 82
HgCdTe diodes, 82–86
$1/f$ noise, 87–89
room-temperature HgCdTe APD performance, 129–131

SAD. *See* separate absorption and detection.
SB. *See* Schottky barriers.
scanning arrays, 2
Shockley–Read minority carrier lifetime properties, 73–74
Schottky barriers (SB), 23–25
semiconductors
extrinsic, 16–18
intrinsic direct bandgap, 13–16, 31–59
PbSnTe bandgap alloy, 42–44
separate absorption and detection (SAD) structure, 139–141
Shockley–Read recombination, 34
signal flux, 7
Silicon Schottky barrier (SB) detectors, 23–25
superlattices
bandstructure, 45–47
HgTe/CdTe, 48
type II, 53–57, 59
type III, 45–53, 59
types, 18–19
surface generation current, 84

TDI. *See* time, delay and integrate.
ternary alloys, 37–42
ternary alloys, mercury cadmium telluride, 37–41
thermal detection, 93–95
limits vs. photon uncooled detection, 105–107
thermal detectors, 9–11
noise equivalent temperature, 9
technologies
α-silicon, 11
thin-film ferroelectrics, 11
vanadium oxide, 11
thermal generation rate, 6–7, 22
antenna structures, 22
microlenses, 22
resonant structures, 22
thin film ferroelectrics, 11
Time, delay and integrate (TDI) techniques, 2
total dark current, 85–86
transport properties, HgCdTe and, 67–73
tunneling via bandgap states, 84–85
tunneling, direct, 85
type II superlattices, 53–57, 59
minority carrier lifetime, 54–57
Auger transitions, 56
type III superlattices, 45–53, 59
absorption coefficient, 49–51
band offsets and strain, 47–48
bandstructure, 45–47
effective mass, 51–52
HgTe/CdTe, 48
minority carrier lifetime, 52–53
misfit dislocations, 48–49

uncooled detection, 93–107
BLIP performance, 93
photon, 95–105
thermal detection limits vs., 105–107
thermal, 93–95

vacancy-doped p-type mercury cadmium telluride, 39–40
vanadium oxide thermal detector technologies, 11
vertically integrated photodiode (VIP), 77–81, 89
VIP. *See* vertically integrated photodiode.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Dr. Kinch, an internationally known specialist in IR device and materials physics, is the Chief Scientist at DRS Infrared Technologies located in Dallas, Texas. He was educated at Wadham College, Oxford University, where he took a first-class degree in physics in 1960. He subsequently studied at the Clarendon Laboratory under Dr. B. V. Rollin and obtained a D. Phil. degree in 1965. In 1966 he joined the Central Research Laboratories of Texas Instruments, where he was engaged in research into the physics of far-infrared materials. He transitioned to infrared materials in 1969 and thereafter became fully embroiled in the many facets of the physics of HgCdTe. He was elected a TI Research Fellow in 1985 and became the leader of the Sensor Physics Branch in the Sensors and Infrared Laboratory (part of TI's Corporate R&D Laboratory) until the demise of TI's Central Research Laboratories in the mid-1990s. At that time, the infrared R&D effort was transitioned into TI's Defense Electronics Group, which was later acquired by Raytheon in 1997. Upon Raytheon's further acquisition of Hughes Aircraft Corporation, the Dept. of Justice mandated the divestment of the original TI IR sensors group to DRS Technologies in 1998. Dr. Kinch subsequently became Chief Scientist of DRS Infrared Technologies in 2000.

For the past 40 years, he has been engaged in (and published seminal papers on) all facets of HgCdTe material, processing, and detector research and development for first-, second-, and third-generation IR systems. He has authored, or co-authored, in excess of 200 publications, mostly in the open literature, in internationally recognized journals, and has been awarded several patents in the field of IR sensors. In 1987, together with D. D. Buss and R. A. Chapman, he received the IEEE Jack A. Morton Award for outstanding work in semiconductor devices, involving their research into HgCdTe monolithic IR charge-coupled and charge-injection devices. Dr. Kinch was elected a Fellow of the American Physical Society in 1988 and a Fellow of the Military Sensors Symposium in 2002. He is currently engaged in an effort to increase the operating temperature of HgCdTe detectors of all cutoff wavelengths into the uncooled range, as well as programs to fully develop the capability of HgCdTe electron avalanche photodiodes for all regions of the IR spectrum.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 28 Mar 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use