Carrier Transport and Related Effects in Detectors of the Cryogenic Dark Matter Search

by

Kyle Michael Sundqvist

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

and the Designated Emphasis

in

Nanoscale Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernard Sadoulet, Chair

Professor Peter Y. Yu

Professor Eugene E. Haller

Spring 2012

Carrier Transport and Related Effects in Detectors of the Cryogenic Dark Matter Search

Kyle Michael Sundqvist

Abstract

Carrier Transport and Related Effects in Detectors of the Cryogenic Dark Matter Search

by

Kyle Michael Sundqvist

Doctor of Philosophy in Physics

and the Designated Emphasis in Nanoscale Science and Engineering

University of California, Berkeley

Professor Bernard Sadoulet, Chair

The Cryogenic Dark Matter Search (CDMS) is searching for weakly-interacting massive particles (WIMPS), which could explain the dark matter problem in cosmology and particle physics.

By simultaneously measuring signals from deposited charge and the energy in non-equilibrium phonons created by particle interactions in intrinsic germanium crystals at a temperature of 40 mK, a signature response for each event is produced. This response, combined with phonon pulse-shape information, allows CDMS to actively discriminate candidate WIMP interactions with nuclei from electromagnetic radioactive background which interacts with electrons.

The challenges associated with these techniques are unique. Carrier scattering is dominated by the spontaneous emission of Luke-Neganov phonons due to zero-point fluctuations of the lattice ions. Drift fields are maintained at only a few V/cm, else these emitted phonons would dominate the phonons of the original interaction. The dominant systematic issues with CDMS detectors are due to the effects of space charge accumulation. It has been an open question how space charge accrues, and by which of several potential recombination and ionization processes.

In this work, we have simulated the transport of electrons and holes in germanium under CDMS conditions. We have implemented both a traditional Monte Carlo technique based on carrier energy, followed later by a novel Monte Carlo algorithm with scattering rates defined and sampled by vector momentum. This vector-based method provides for a full anisotropic simulation of carrier transport including free-flight acceleration with an anisotropic mass, and anisotropic scattering rates.

With knowledge of steady state carrier dynamics as a function of applied field, the results of our Monte Carlo simulations allow us to make a wide variety of predictions for energy dependent processes for both electrons and holes. Such processes include

carrier capture by charged impurities, neutral impurities, static dipoles, and capture forming "anion" (D^-/A^+) states. We also generate predictions for impact ionization of shallow impurities and of impact "neutralization" of D^-/A^+ states. We use measurements of carrier capture performed on CDMS detectors to validate a plausible model for electron and hole capture due to neutral shallow impurities and their charged D^-/A^+ states. This model, along with carrier drift and diffusion parameters from Monte Carlo simulation, can be used as the foundation for simulations of space charge evolution in CDMS detectors, simultaneously solving continuity equations with Poisson's equation.

To my parents, Jim & Carol Sundqvist

Contents

List of Figures

List of Tables

Acknowledgments

Many people gave me their support during my time at Berkeley, and I would like to express my gratitude.

I would like to thank my advisor, Bernard Sadoulet. Through Bernard's kind advice and thoughtful supervision, I take away a grounded and pragmatic view for approaching physics. Thank you for the long leash in letting me follow what I thought was important. It was an invaluable education.

A special thanks to my girlfriend, Maya. No one else would continue to give me such strong support, even after putting up with my constantly evolving mess of books and my pacing up and down the hall.

I would also like to thank the crew at Berkeley. Thanks to Arran Phipps for being my professional sanity check. Thanks to Bruno Serfass for his steadfast willingness to give help. Thank you to Dennis Seitz, who has a knack for straightforward advice. Thanks to Nader Mirabolfathi for his insight, and to fellow grad students Jeff Filippini and Miguel Daal for their humor and shared experience in figuring out how things work. It was also a pleasure to mentor undergraduates like Albert Lam and Nate Earnest. I enjoyed getting to know the newer grad students, Todd and Danielle. Thank you to Caryl Esteves, Mary MacCready, Elizabeth Arscott, and Rachel WInheld for making the group enjoyable and for cutting through red tape on my behalf. I also thank Avi Rosenzweig for his perspective on the university and how it works.

I enjoyed my time on the CDMS experiment. Thanks to Dan Bauer, Jeter, and Sten at FermiLab. It was a pleasure to work alongside Cathy Bailey, Mike Dragowski, and Dan Akerib at CWRU. Pondering electrical components with Bruce Hines and Martin Huber while in Colorado was a lot of fun. Thanks also to Steve Leman and Kevin McCarthy at MIT. Thanks to Matt, Walter, Paul, Betty, and Blas at Stanford, to Jim and Dave at Soudan, and to everybody else in the collaboration.

Lastly, I would like to thank my parents. Thank you for making me who I am, and for helping me not loose sight of it.

Chapter 1

Introduction to Dark Matter, the Cryogenic Dark Matter Search, and Non-Ideal Effects Relating to Charge

1.1 Introduction to dark matter and the Cryogenic Dark Matter Search

The existence of dark matter represents a central issue in modern cosmology. Several lines of evidence indicate dark matter comprises the majority of mass in the Universe, yet its nature remains unknown. Dark matter by definition does not emit or absorb electromagnetic radiation. As yet, its discovery and subsequent study have been permitted only through the observation of indirect gravitational effects.

Theories of particle physics, however, readily suggest the existence of yet undiscovered particles. Many such particles have properties known to be characteristic of dark matter. Provided they occur in adequate abundance, such particles could account for the mysterious dark matter in the Universe. Particular classes of dark matter candidates have been proposed, often generated within the framework of Supersymmetry (SUSY), to have interaction with the weak nuclear force. Particles of this category are known as Weakly Interacting Massive Particles (WIMPs). With an adequate coupling to ordinary matter via weak interaction, a direct and measurable response to dark matter may be possible. As the identification of dark matter is one of the highest priority goals in both cosmology and particle physics, a direct detection measurement may be the best option for substantial experimental progress.

This chapter will briefly review the arguments for the existence of dark matter, as well as general requirements for a direct detection experiment. We will introduce our experiment, the Cryogenic Dark Matter Search (CDMS). We will explain the design

and theory of operation of our detectors. We will then make the intent of this thesis work clear, as we then introduce some detector phenomenology demonstrating the need to understand charge transport processes in our detectors.

1.2 The weight of the Universe

In this section we discuss the mass budget of the Universe, pointing out the significance of the dark matter contribution.

1.2.1 The Friedmann equation and Ω

Cosmology is the study of the origin and development of the Universe. A good place to start is the Friedmann equation, which describes the expansion of the Universe. Under the assumption that the Universe is isotropic and homogenous, the Friedmann equation depicts the evolution of a *scale factor* within the context of general relativity.

$$H^2 + k\frac{c^2}{a^2} = \left(\frac{\dot{a}}{a}\right)^2 + k\frac{c^2}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} \tag{1.1}$$

Here, H is the Hubble constant, a is the scale factor, G the gravitational constant, k is the curvature term, ρ is the density of the Universe, and Λ is Einstein's cosmological constant. By inspection, we see that the Hubble constant represents a normalized expansion rate of the Universe.

We define a parameter Ω representing the right hand side of equation 1.1.

$$\Omega = \frac{8\pi G}{3H^2}\rho + \frac{\Lambda}{3H^2} \tag{1.2}$$

The Friedmann equation is then simply the following.

$$1 + k\frac{c^2}{H^2a^2} = \Omega \tag{1.3}$$

We find that this form of the Friedmann equation provides useful insights for interpreting the evolution of the Universe. This representation shows us that if Ω departs from 1, we must consider either an open ($k < 0$) or closed ($k > 0$) Universe. In these cases, the Universe would respectively either expand forever with acceleration, or eventually collapse in on itself.

It happens to be the remarkable case that observational evidence has found the present Universe to be currently of an $\Omega_0 = \Omega(t_0)$ extremely close to one ($k = 0$). This represents a "flat" Universe. Only $\Omega = 1$ is an equilibrium point (though an unstable one). If today's value Ω_0 were not exactly one today, then when the Universe was a second old, Ω at that time could differ from one by only about 10^{-16}. Avoiding

a fine-tuning problem to account for Ω at early times in the Universe is possible only if identically $\Omega = 1$. Fortunately, theories of *inflation* immediately following the Big Bang mandate subsequent values of Ω to be extremely close to one.

1.2.2 Contributions to Ω and the current cosmological picture

It brings insight to itemize Ω as independent contributions to the Universe's density. These terms evolve differently as the Universe develops. For instance, we know that contributions by photons contributed to the density budget of the early Universe, but this density diluted faster than non-relativistic matter due to relativistic effects on the radiation's momentum.

We introduce the normalized density contributions important to today's Universe. There are contributions by *matter*, and by the *vacuum energy* due to the cosmological constant.

$$\Omega_m = \frac{8\pi G\rho_0}{3H_0^2} \tag{1.4}$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2} \tag{1.5}$$

Since we say $\Omega_0 = 1$, we know that $\Omega_m + \Omega_\Lambda = 1$.

Developed from the convergence of many independent investigations, the current cosmological picture is called ΛCDM [21, 22] which is an acronym for "Λ-Cold Dark Matter." According to this model the Universe expands, but with a spatially flat geometry ($k = 0$). The mass component Ω_m is comprised of both normal baryonic matter and cold dark matter.

$$\Omega_m = \Omega_b + \Omega_{cdm} \tag{1.6}$$

Normal baryonic matter comprises only $\Omega_b = 0.0456 \pm 0.0015$ [23], leaving dark matter and vaccum energy ("dark energy") to dominate the remaining density as $\Omega_{cdm} + \Omega_\Lambda = (1 - \Omega_b)$. Dark energy has been determined to constitute a $\Omega_\Lambda = 0.726 \pm 0.015$ while dark matter is found to be $\Omega_{cdm} = 0.228 \pm 0.013$. While the nature of dark matter remains unknown, it has been studied by way of its gravitational effects on baryonic matter. While cosmological model is the prevailing paradigm, there are alternative hypotheses (MOND [24] and extensions to it, such as TeVeS [25]) which try to explain astronomical observations by modifying gravity at large distance scales. For our consideration, we focus on particle dark matter and assume Newtonian gravity holds across all scales.

1.3 Evidence for dark matter

We now discuss evidence for dark matter, motivating the ΛCDM model. Many lines of observational evidence are consistent with the ΛCDM model with Ω values described in the last section. Here, we give a brief outline of evidence specific to dark matter.

1.3.1 Dark matter as "missing mass"

Dark matter is inferred from the gravitational dynamics of luminous matter, indicating the presence of "missing" mass. Evidence for dark matter exists across many magnitudes of length scale.

Spiral galaxies and rotation curves

The most popular and traditional evidence of dark matter comes from measurements of angular velocities of spiral galaxies as a function of radius. Profiles of velocity as a function of galactic radius, or "rotation curves," provide the radial distribution of the galactic mass by equating gravitational and centripetal acceleration. Beyond the bulk mass of the galaxy which is optically visible, velocities should fall off in a characteristically $r^{-1/2}$ manner. Observations have consistently shown [1] that velocities of visible objects instead remain constant at large radius, as in figure 1.1. A constant velocity infers that mass as a function of radius continues to increase linearly, beyond the visible measurable galactic components. Measurements indicated that at least ten times the mass of the visible stars was present in spherical halos of unseen matter.

Elliptical galaxies and virial velocities

Elliptical galaxies also evidence dark matter. In elliptical galaxies, velocities have become virialized. This provides a relation between the gravitational potential energy and the kinetic energy. The kinetic energy provides the velocity *dispersion*. As $\langle E_k \rangle = -1/2 \langle V \rangle$, we can relate the velocity fluctuations as $\langle v^2 \rangle = GM/\langle r \rangle$. As with spiral galaxies, velocities in elliptical galaxies are also found to be too fast for the mass of visible contributions only [26]. Dark matter is again required to account for the long-term stability of elliptical galaxies.

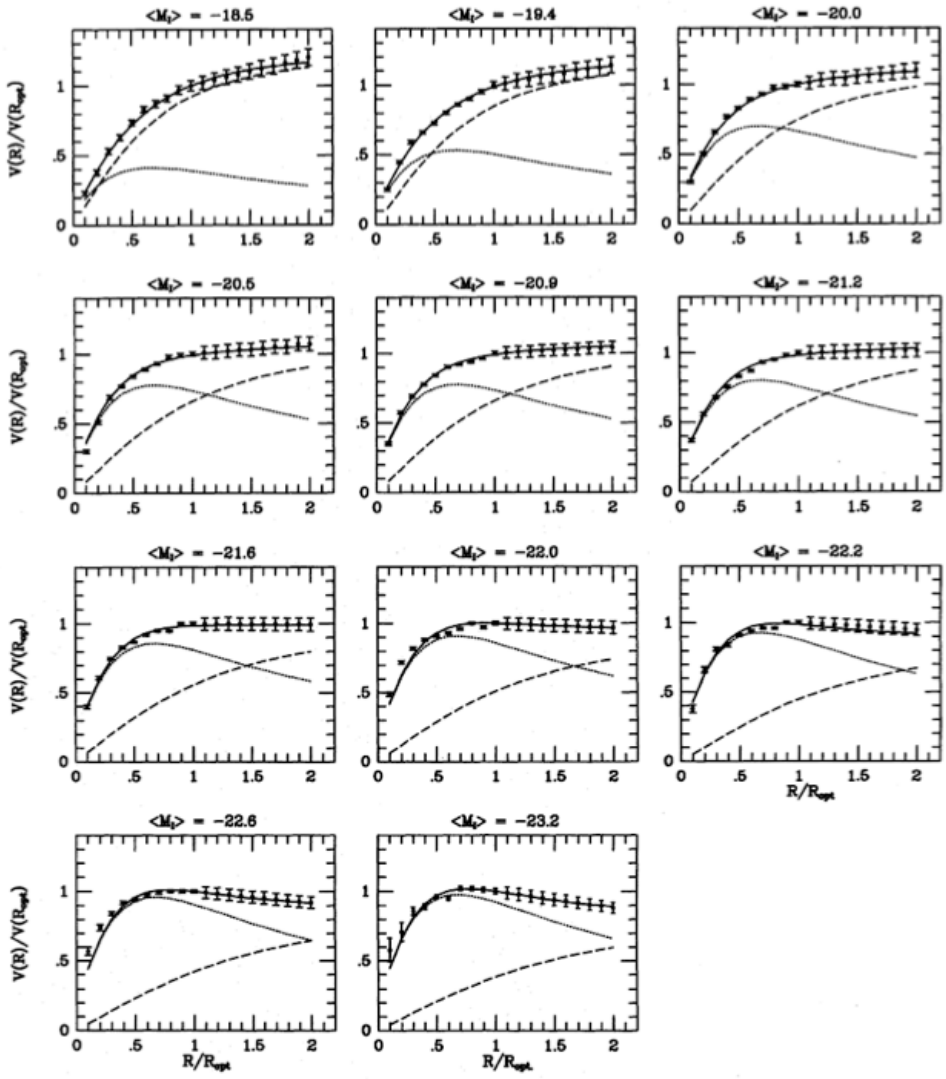

Figure 1.1: Best parameters for a two-species fit of mass distributions to galactic rotation curves, for a range of several magnitudes $\langle M_I \rangle$, from [1]. Dotted lines show luminous material, and dashed lines show implied dark halo.

Galaxy clusters

At the scale of galaxy clusters, Zwicky already in 1933 [27] pointed out that velocities in the Coma cluster also are too fast to be stable without the existence of additional mass. Futhermore, X-ray emission within galactic clusters shows [28] that equivalent temperatures of baryons are much higher, again using the virial theorem, than temperatures expected by the gravitational potentials of only the observed luminous matter.

A particularly striking line of evidence for dark matter comes from relatively recent observations of the "Bullet Cluster." The Bullet Cluster is actually two clusters that have collided. Using optical observations, the total mass distribution was reconstructed by accounting for gravitational lensing effects. With independent observations of X-ray emissions, the baryonic matter contribution was mapped. Comparing the total matter and baryonic matter profiles shows a stark difference between the baryonic and cold dark matter species [2]. This observation is consistent with the model that dark matter is collisionless, such that the dark matter halos of the original two clusters passed through each other unchanged. Meanwhile, hot baryonic gas distributions between the two structures interacted heavily during collision, impeding their motion.

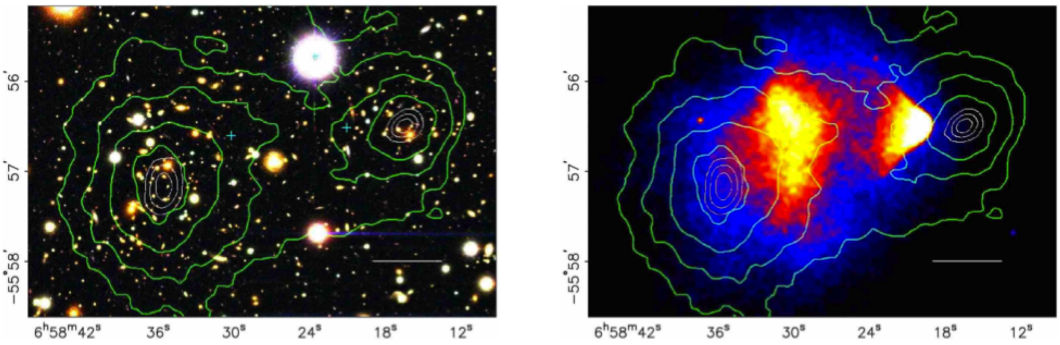

Figure 1.2: Evidence for dark matter in the Bullet Cluster. The baryonic mass distribution is distinctly different from the total mass distribution, inferring the presence of dark matter. **a.** Optical image from the Hubble Space Telescope. This shows mass contours reconstructed by gravitational lensing. **b.** Image from Chandra X-Ray Observatory, showing X-ray emission from baryonic matter. From Clowe, et. al. [2]

1.3.2 Evidence specific to dark matter as non-baryonic and cold

Other lines of evidence show that dark matter is both non-baryonic and non-relativistic.

Big Bang Nucleosynthesis

Minutes following the Big Bang, the Universe cooled to a temperature below the binding energy of typical nuclei, and light elements began to form. This continued for approximately 17 minutes, until the Universe cooled below the energy required for nuclear fusion [29].

With knowledge of nuclear cross sections, the conditions of the Universe at this time can be reconstructed. Predictions for Big Bang nucleosynthesis (BBN) account for the abundances of light elements seen today, and provide an accurate measurement of the baryon density of the Universe. Measurements of primordial deuterium happen to pin down the baryon density with great precision to only a few percent [30] of the critical density required for $\Omega = 1$. This requires the majority of dark matter to be non-baryonic.

Cosmic Microwave Background

The Cosmic Microwave Background (CMB) represents the oldest light that reaches Earth. In the hot environment of the early Universe, baryons coupled well to photons to such an extent that the Universe was opaque. This coupling provided a feedback mechanism between gravity and radiation pressure, leading to "acoustic oscillations" around the small potential wells formed by tiny density perturbations. These oscillations continued until the Universe expanded and cooled such that photons were no longer well coupled to baryons. However, information of these oscillations remain embedded in the anisotropies of cosmic microwave background today. The spectral power information of these anisotropies provide a wealth of precise information regarding the concentrations of baryonic matter, cold dark matter, and dark energy. So the CMB also supplies an excellent validation of the ΛCDM cosmological model.

Large Scale Structure Formation

Dark matter may also be studied through observational surveys of large scale structure in the Universe. Accurate studies of structure, particularly through comparing numerical simulations of structure growth with observation, provide a means of parameter estimation for components of cosmological models [31]. These comparisons show that relativistic (hot) dark matter produces more structure than is observed. However, non-relativistic (cold) dark matter creates approximately the structure seen today, when using a value consistent with the aforementioned ΛCDM

model [22]. Note that relativistic particles such as neutrinos are also outruled as dark matter candidates.

MACHO Searches

Massive Astrophysical Compact Halo Objects (MACHOs) are a general term for astronomical bodies, suggested to explain the presence of dark matter. Such bodies may include varieties of "failed stars," such as brown dwarfs or unassociated planets. They would be composed of normal baryonic matter, and would be hard to detect as they would emit little or no radiation.

Baryonic matter is ruled out by other lines of evidence as described above. However, it is worthwhile to point out that MACHOs searches using surveys for transient gravitational lensing events have also turned up far too few MACHO objects to account for dark matter [32]. This is another line of evidence to suggest a particle form of dark matter.

1.4 Cold dark matter candidates, and requirements for direct detection

Through the varieties of observational evidence we have pointed out, constraints have been placed on possible dark matter candidates. Dark matter is not comprised of baryons, nor neutrinos.

1.4.1 Requirements of a particle dark matter candidate

The situation dictates that dark matter must be a form of new particle, arising from particle physics models beyond the Standard Model. In fact, this connection to particle physics is an exciting prospect. A plethora of candidates suggested by particle theory are presently being searched for in a number of accelerator, indirect, and direct dark matter searches.

A viable dark matter candidate must have the following properties [33, 34].

Non-baryonic Can it account for dark matter and preserve the BBN baryon density?

Numerous Does it match the appropriate relic density?

Cold Is it non-relativistic during structure formation?

Non-interacting Is it neutral, and does not interact electromagnetically?

Stable Is it stable over time periods compatible with the age of the Universe?

One viable candidate are *axions*. Axions are proposed to solve the "strong CP problem" in QCD, which would explain the lack of observation of a neutron electric dipole moment [34]. Searches for axions are ongoing [35, 30].

Another viable candidate are Weakly Interacting Massive Particles (WIMPs). The WIMP is a hypothetical particle with mass and coupling strengths characteristic of weak interactions. The calculation of a dark matter relic particle density based on freeze-out following the Big Bang implies that such particles should have an annihilation cross section which happens to be on the order of the weak scale. This is appealing because this is a natural characteristic of WIMPs predicted under supersymmetric (SUSY) theories, which are popular extensions to the Standard Model. WIMPs are the candidate dark matter particle considered for this thesis.

1.4.2 Direct detection of dark matter

Three general categories of experiment may serve to evidence a particle dark matter candidate. *Accelerator* experiments seek to discover new particles and potentially dark matter candidates, in situ. Even after a new particle discovery, an identification as a dark matter candidate may not be straight forward. In this case, confirmation of such a particle as ambient dark matter would still be necessary. *Indirect searches* attempt to evidence dark matter by finding secondary products such as neutrinos, photons, or antimatter. Here, we are concerned with *direct detection* experiments, which seek to evidence interactions made by ambient dark matter particles.

Direct detection experiments share some common themes. Here, we focus on searches for WIMPs. The goal of direct detection experiments is to detect some number of WIMPs of an ambient flux which flow through the Earth. Although the WIMP interaction cross section may be small, the ambient flux is abundant enough to make measurement plausible. The putative WIMP flux through the Earth is of order [34]

$$F = 10^5 (100\ GeV/m_\chi)\ cm^{-2} s^{-1} \tag{1.7}$$

where m_χ is the WIMP mass. Such WIMPs would scatter elastically from nuclei in some absorbing material, and provide a measurable signal from nuclear recoils. The goal is to measure the event rate R, as a function of the nuclear recoil energy, E_R.

The differential WIMP event rate, commonly expressed in units $\mathrm{kg}^{-1}\ \mathrm{day}^{-1}\ \mathrm{keV}^{-1}$, is given as [36]

$$\frac{dR}{dE_R} = \frac{2}{\sqrt{\pi}} \frac{\sigma_0 \rho_0 v_0}{m_N m_\chi} \frac{1}{r E_0} \exp\left(\frac{-E_R}{r E_0}\right) \left(F(E_R)\right)^2 \tag{1.8}$$

where m_N is the mass of the target nuclei, ρ_0 is the WIMP density, v_0 is the WIMP velocity, E_0 the WIMP energy, σ_0 the cross section, $F(E_R)$ the nuclear form factor, and r the ratio $4 m_\chi m_N / \left(m_\chi + m_N\right)^2$.

Generally, there are scattering probabilities relevant for both spin-dependent and spin-independent cross sections. Generally, the spin-independent cross sections dominate. The spin-independent cross section may be given as [37]

$$\sigma_0^{SI} = \frac{4m_\chi^2 m_N^4}{\pi \left(m_\chi + m_N\right)^2} \left(\frac{f_n}{m_n}\right)^2 \tag{1.9}$$

where f_n is the WIMP-nucleon coupling.

Consider a generic, toy model as an illustration. We could adopt some common assumptions for galactic halo velocities (namely $v_0 \approx 220\ km/s$) and a cross section ($\sigma_0 \approx 10^{-42}\ cm^2$). One may then naively expect an integrated rate to be $\approx 1\ \mathrm{kg}^{-1}\mathrm{day}^{-1}$. So the overall design goal is to resolve nuclear recoils at this rate and lower.

Low cross sections amount to low event rates. The primary issue with direct detection experiments is the presence of background rates. Direct detection experiments require the ability to reject nuclear-scale background events from natural radioactive sources. Such sources may be γ-rays, X-rays, alpha particles, beta particles, cosmic rays, or neutrons induced by cosmic rays. Precautions to create a low-background environment should of course be taken, including such things as underground operation, passive and active shielding, and choice of radio-pure materials. However, it stands that the level of background events will determine the sensitivity of the experiment.

It becomes apparent that dark matter detectors require the ability to distinguish common electromagnetic background induced by natural sources of radioactivity, apart from putative dark matter interactions which are nuclear recoils. To that end, it is possible to discriminate between these two classes of interaction by simultaneous, independent measurements of quantities that reflect the difference in energy deposition density. That is to say for the same total deposited energy, two measured parameters can constrain both the energy deposited and difference between a nuclear recoil and an electron recoil.

When an elastic WIMP-nucleus collision generates a recoiling ion in a detector, this ion will dissipate its energy in a Markov chain of subsequent interactions. Secondary excitations are created, such as scintillation, phonons, or ionization, depending on the material. Some of these signals may be different if the initial interaction were electromagnetic in origin. So measurement of multiple kinds of signals can provide information about the type of initial interaction. Argon detectors produce ionization and scintillation, for example, as do liquid Xe detectors. Among experiments using cryogenic detectors, such as our Cryogenic Dark Matter Search, ionization and phonons supply constraining signals.

1.5 The Cryogenic Dark Matter Search and ZIP detectors

The Cryogenic Dark Matter Search utilizes simultaneous measurements of ionization and non-equilibrium phonons to distinguish nuclear recoils from background on an event-by-event basis. For the phase of the experiment known as "CDMS-II," ZIP (Z-sensitive Ionization and Phonon) detectors of high-purity germanium (and some silicon) were cooled to a temperature $\approx 40\ mK$. Typical germanium impurity densities were $\approx 10^{10}$ cm^{-3} with less than 5000 dislocations/cm^2. Substrate orientations were $\langle 100 \rangle$. Each detector is a disk $\approx 1\ cm$ thick and 7.6 cm in diameter. A ZIP detector is photolithographed with four phonon channels on one face. The opposite face (i.e., the "charge side") is divided concentrically into an inner circular region and an annular "guard" electrode. The phonon side also serves as a low impedance ground reference for the ionization electrodes, which are biased to supply electric fields of $\approx 3\ V/cm$ to serve as a drift field for the charge carriers created by a particle interaction. See figure 1.3.

The detectors for CDMS-II were arranged in five towers. They were shielded by passive lead and polyethylene, as well as an active plastic scintillator veto. Underground operation for the CDMS-II experiment took place at the Soudan Underground Laboratory (2090 meters water equivalent) in Soudan, Minnesota.

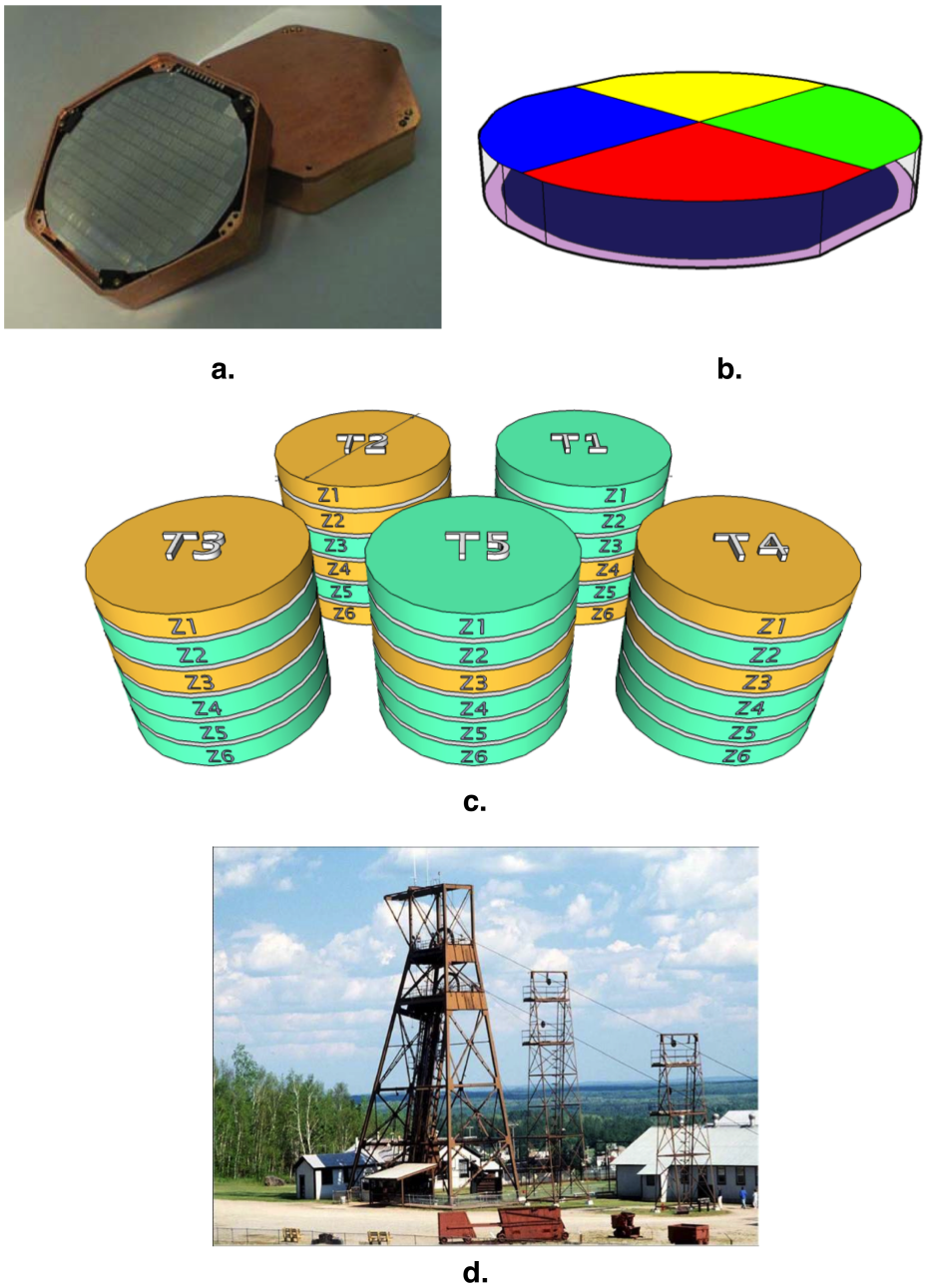

Figure 1.3: **a.** A CDMS-II ZIP detector, 1-cm thick, showing the "phonon side" **b.** A diagrammatic view of a ZIP detector. There are 4 phonon quadrants on the "phonon side." On the reverse, "charge side," there is an inner signal electrode with a concentric outer guard electrode **c.** The CDMS-II Tower assemblies (quantity 5). The yellow detectors represent silicon detectors, while the green are germanium. **d.** The minehead at the Soudan Underground Facility in Soudan, Minnesota

1.5.1 Ionization measurement

When an incident particle interacts within a CDMS detector, electrons and holes are generated in equal number from the germanium bulk. For a brief time, these hot carriers form a net neutral plasma which may screen an external field. While they may still diffuse locally, if left unperturbed they will eventually recombine to each other or to local imperfections of the crystal.

Under milliKelvin operation, there is no ionization by thermal fluctuations. As such, our substrates are automatically depleted. Producing a "depletion region" in our detectors is not a consideration as it would be for radiation detectors at warmer temperatures. So in order to measure ionized carriers, only a small external field is applied to drift these oppositely charged carriers away from each other to adequately avoid recombination. As we will soon discuss, the field should also not be too large as to limit the number of phonons produced by the carriers. So the necessary field is only of the order $\sim V/cm$ ($-3V$ bias for the 1-cm standard detectors). The field dependence of ionization collection has been studied in the past by both our Berkeley group and in the larger CDMS collaboration [38, 3].

As electrons and holes drift inside the detector, they induce image charge on the voltage-sourced charge electrodes. By Ramo's theorem, the image charge produced on an electrode of an ideal parallel plate detector is proportionate to the distance subtended between the electrodes. Provided an electron and hole pair both completely traverse their course to their respective electrodes, the integrated charge on one of the electrodes will be one electron charge. A transimpedance amplifier connected to the detector electrode nulls any change in charge through the use of feedback. When charge from a particle event is suddenly deposited on the detector electrode, the amplifier negates this charge by injecting the opposite amount of charge back onto the electrode, through its feedback network. On fast time scales, the amplifier produces an output voltage signal across its feedback capacitor. This signal is eventually dissipated through a parallel feedback resistor, with an $R_f C_f$ time chosen to maximize integration time and lower noise, though accommodating an expected event rate.

In figure 1.4, we see a diagram of our transimpedance amplifier. The first-stage transducer is a JFET. To reduce parasitic gate capacitance, the JFET is implemented within the dilution refrigerator, as close as possible to the detector electrode. The JFET is maintained at an operational temperature of $T \approx 150\ K$, such that its silicon substrate remains conductive above freeze-out. The JFET is read out in a cascode configuration by external warm electronics, which also supply feedback from the amplifier outputs.

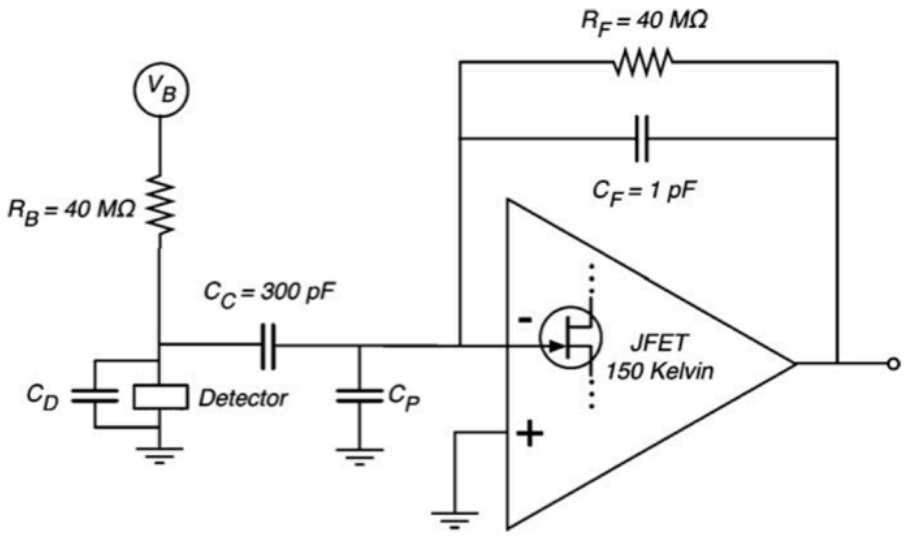

Figure 1.4: The CDMS charge amplifier, for the ionization measurement. Each ionization collection electrode has a similar amplifier. The "open-loop" portion of the amplifier has a JFET at $T \approx 150\ K$ as its first-stage input, which is followed by warm electronics.

1.5.2 Phonon measurement

Particle interactions in CDMS detectors also excite lattice vibrations, which are quantized as phonons. These phonons are measured out of equilibrium, meaning that they are measured long before equalizing across the detector bulk. Therefore, this measurement is not simply monitoring the bulk temperature change of the detector. The extra information held in non-equilibrium phonons provides a wealth of information regarding the lateral and depth locations of the event, the timing of the event, and contains sensitivity to the phonon spectral information. This additional event information allows CDMS to introduce a number of additional quality cuts, enabling further background rejection.

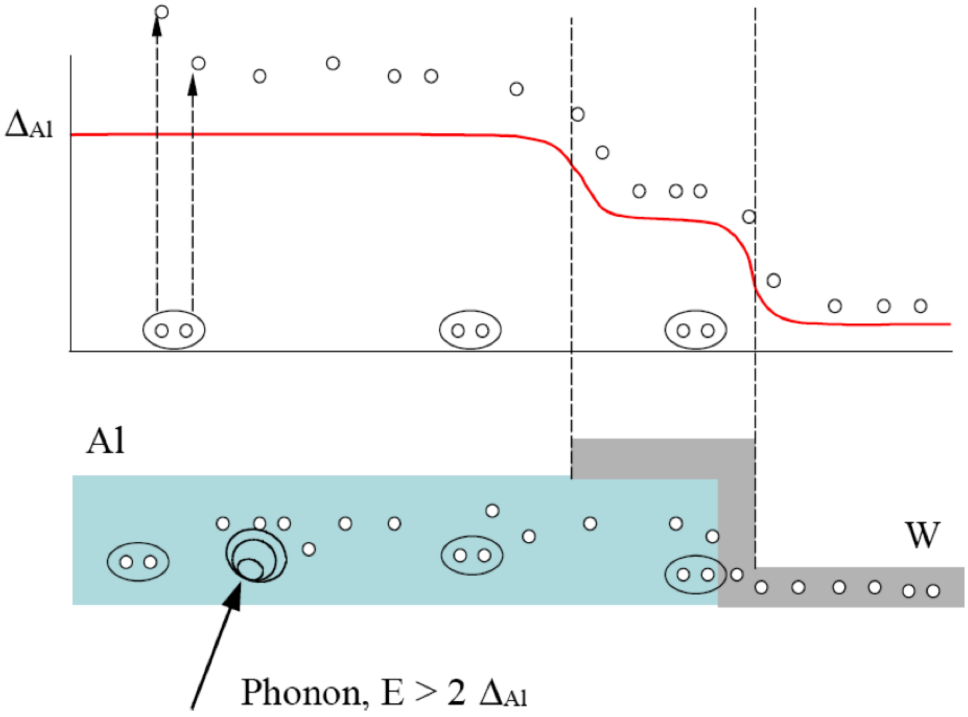

Figure 1.5: A representation of CDMS quasiparticle trap-assisted transition-edge sensors. At top, we see a plot of the superconducting gap as a function of position. At the bottom, we see a cartoon depiction the corresponding TES collection structure cross section. At the left, aluminum fins collect incident phonons. Phonons of energy $E > 2\Delta_{Al}$ will break Cooper pairs, generating quasiparticles. These quasiparticles will diffuse within the superconductor, but de-excite to the lower gap state of the "trap" interface region, and from there diffuse into the TES at the right.

The phonon signal is comprised of the following contributions.

primary phonons Phonons (primarily optical phonons, by energy) produced during the initial particle interaction

recombination phonons Phonons produced when carriers recombine, primarily to the contacts

Luke-Neganov phonons Phonons, assumed to be primarily acoustic phonons, emitted by carriers under drift

The presence of recombination phonons were evidenced by Shutt, et al [38]. The contribution of Luke-Neganov phonons [39, 40] will be discussed later. Here, while the evolution and collection of phonons is general, we focus on primary phonons.

Phonons simultaneously scatter and decay within the crystal. An elastic scattering mechanism is *Rayleigh* scattering, which is isotropic and has a phonon frequency dependence $\propto \omega^4$. The primary decay mechanism is *anharmonic decay* which allows a phonon to spontaneously decay into two lower-energy phonons. Anharmonic decay has a frequency dependence of ω^5. So we note that both processes are highly energy dependent.

A consequence of these highly energy-dependent rates is that high-energy phonons undergo many scatters before they are able to traverse significant distances. These high-energy phonons remain localized until their mean-free paths are approximately the length of the detector geometry. At this point, phonons are considered ballistic and will travel unimpeded until they reach a surface. Phonons incident on a bare surface have a high likelihood to reflect back into the crystal. However, phonons incident on superconducting aluminum collection fins of the CDMS phonon sensors will be absorbed and break Cooper pairs, provided their energy is adequate ($\hbar\omega > 2\Delta_{Al}$).

Tungsten transition-edge sensors (TESs) are used as the CDMS phonon transducers. These are superconducting sensors biased within their superconducting transition, and held in stable operation by virtue of being locally voltage biased (i.e., held in electro-thermal feedback). In order to maintain high sensitivity, the heat capacity of these TESs must kept small. Therefore, TESs cannot cover large areas of the detector surface. CDMS ZIP detectors utilize a two-stage phonon collection technique known as quasiparticle trapping, first proposed by Booth [41]. Tungsten TESs are lithographed onto large-area aluminum fins. Aluminum has a larger superconducting gap than tungsten. Phonons first enter the aluminum fins from the detector bulk. These phonons break Cooper pairs in the aluminum fins, generating quasiparticles. These diffusing quasiparticles then eventually end up in tungsten, but de-excite to the lower gap energy. This trapping of quasiparticles into the tungsten will preferentially force dissipation to occur there, driving the tungsten sensors into further into their normal state and providing signal. So in this scheme, phonons are collected over a large area, yet dissipation primarily occurs in the TES where such heating contributes to signal.

To cover the large detector area, each ZIP detector quadrant is actually comprised of 1036 individual TESs lithographed in parallel. This creates a low impedance, and

a low differential change in impedance. To read out this array, an input coil to a SQUID amplifier is included in series. This amplifier is discussed in appendix I.

1.5.3 Signal formation of an event

CDMS ZIP detectors measure ionization and phonons, but the processes initiated by incident radiation are the same as in traditional semiconductor detectors. Incident raditaition quanta, if they are of sufficent energy much greater than the semiconducting bandgap ($E_g = 0.76\ eV$ for milliKelvin germanium) may impart energy to individual atomic electrons in the substrate to form electron-hole pairs. These primary electrons and holes go on to produce further ionized carriers and phonons in a dissipative cascade of collisions. In fact, it is a good approximation that the deposited energy of an incident quanta produces a total number of electron-hole pairs related by an average ionization energy, as suggested by Klein [42] and others.

$$N_Q = {}^{W}/_{\varepsilon} \tag{1.10}$$

Here, W is the energy deposited by the ionizing particle or photon, N_Q is the number of pairs generated (the "yield"), and ε is an average ionization energy. The average ionization energy, ε, accounts for energy and momentum threshold effects in creating an electron-hole pair, plus dissipation which goes into producing phonons. In germanium, ε is quite close to 3 $eV/pair$, for events originating from *electromagnetic interations*. As an example, a 60 keV photon would generate $60,000/3 = 20,000$ electron-hole pairs.

Klein [42] suggests that the average ionization energy follows an empirical form

$$\varepsilon = \left(\frac{14}{5}\right) E_g + r\left(\hbar\omega_0\right) \tag{1.11}$$

Here, the E_g is the bandgap energy, $\hbar\omega_0$ is the optical phonon energy, and r is an adjustable parameter representing the number of phonons per generated pair. For the aforementioned bandgap, a phonon energy $\hbar\omega_0 = 37\ meV$, and an ε of 3 $eV/pair$, this would indicate ~ 24 optical phonons are generated for every electron-hole pair. For an incident photon or charged particle, this would mean 71% of the deposited energy is imparted to electron-hole pairs with the rest dissipated as optical phonons. A more sophisticated treatment of the cascade mechanisms between pair production versus phonon dissipation can be performed under Lindhard theory [43].

Recall that a WIMP-nucleus interaction produces a "nuclear recoil." In this case, a recoiling nucleus is considerably heavier than the primary electrons of the electromagnetic case. A recoiling nucleus transfers much more of its energy into the production of phonons. As a result, the required energy to produce a pair is notably higher for *nuclear recoils* than for electromagnetic events ("electron recoils"). Terms such as "quenching" are sometimes used to describe the fact that nuclear recoils have an average ionization energy larger than electron recoils.

1.5.4 Discrimination by ionization yield

The partition of deposited energy between electron-hole pairs and phonons gives CDMS the ability of event discrimination. With simultaneous measurements of ionization and phonons, every event can classified as either an electron recoil (background) or a nuclear recoil (possible signal) by a measurement of its yield.

A complication is that electron and holes while under drift will produce Luke-Neganov phonons. The energy emitted to these phonons will be the work done by the number of carriers, $N_Q = W/\varepsilon$, across the potential drop ΔV of the crystal.

$$E_{Luke} = W\left(\frac{e\Delta V}{\varepsilon}\right) \tag{1.12}$$

The measured phonon energy E_P is then the contribution of the primary phonons generated during the original particle recoil event E_r, plus the Luke phonon contribution.

$$E_P = E_r + E_{Luke} \tag{1.13}$$

In other words, we can deduce the original recoil phonon energy by the measured phonon energy minus a Luke correction based on measured ionization.

$$E_r = E_P - W\left(\frac{e\Delta V}{\varepsilon}\right) \tag{1.14}$$

CDMS defines its primary discrimation parameter as "ionization yield," taken to be the energy measured in ionization, divided by the energy measured in recoil phonons.

$$Y \equiv \frac{E_Q}{E_r} = \frac{E_Q}{E_P - E_Q\left(\frac{e\Delta V}{\varepsilon}\right)} \tag{1.15}$$

Here, E_Q is the amplitude of the ionization signal calibrated to a known source, such that $Y = 1$ for electron recoils. Electron recoils therefore have an ionization yield ~ 1, whereas nuclear recoils have ionization yield ~ 0.3 due to their proportionately larger production of optical phonons. Figure 1.6 shows ionization yield as a function of phonon recoil energy for CDMS calibration data. This shows the discrimination power attained by this parameter, rejecting electron recoils by $> 10^4$.

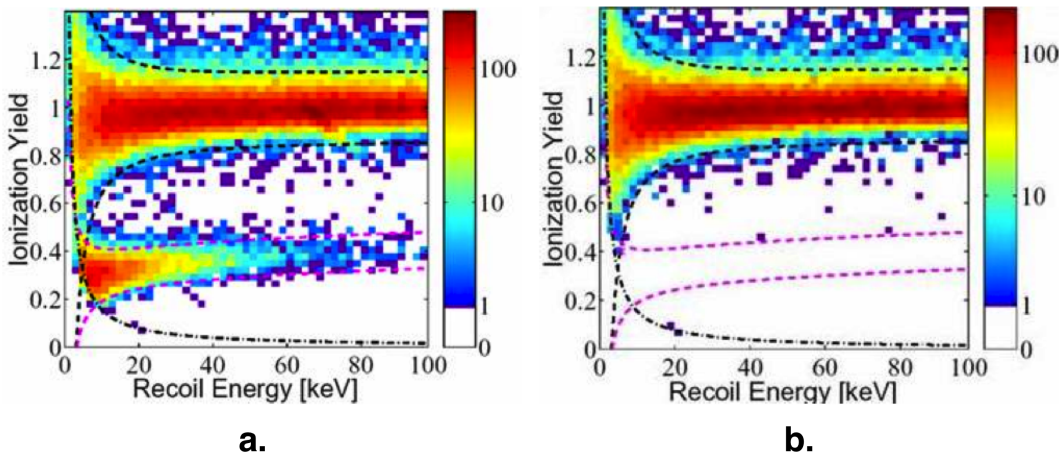

Figure 1.6: Density plots mapping ionization yield of CDMS data as a function of phonon recoil energy. The plotted curves represent the $\pm 2\sigma$ electron recoil ($Y \sim 1$) and nuclear recoil ($Y \sim 0.3$) bands determined by calibration data. **a.** Density plot of electron and nuclear recoil events from a ^{252}Cf neutron calibration source. **b.** Density plot of electron recoil events from a ^{133}Ba gamma calibration source. The nuclear recoil band is clear of any events, once the neutron source is removed. This illustrates the ionization yield electron recoil rejection power of $> 10^6$ for bulk.

1.6 Non-ideal effects, and the need to understand charge transport

We have introduced the CDMS experiment and its goal to evidence dark matter interactions in high-purity germanium detectors at milliKelvin temperatures. We now point out detector phenomenology that has added complications to CDMS operation, and motivates our desire to understand charge transport processes.

The CDMS approach requires discrimination between electron and nuclear recoils. Measuring a different amount of ionization other than what was actually created during an event will degrade our yield parameter of equation 1.15. An electron recoil with poor ionization collection begins to look like a nuclear recoil, and this degrades our background rejection. Furthermore, if internal electric fields accrue due to space charge, this represents a change in the work performed by carriers, disrupting our calibration of Luke phonons. Internal fields also lead to non-uniform operation in our detectors. This may lead to problems associated with avalanche break-down, or regions of poor ionization collection.

In short, we would like to understand processes that affect either: 1) the number of charges collected from a particle event, or 2) somehow affect the internal electric field of the crystal.

1.6.1 The neutralization process and space charge

When a detector is first cooled to milliKelvin base temperature, it is not operational. There is clearly space charge present that prevents ionization collection from a biased detector. No ionization signal is seen until a "neutralization" process is performed. Here, the detectors are grounded so that no external bias is applied. Detectors are equipped with infrared LEDs, which optically generate carriers in the crystal using shorts bursts of light over an extended time of perhaps 10 hours. Alternatively, nuclear sources have been used in the past which also generate carriers for this process. Aside from our empirical knowledge, the neutralization process itself is not well understood.

After some period of operation, the detector performance degrades and a detector must undergo a supplementary neutralization sequence. How and where exactly this space charge accrues and how exactly it affects our signals has not been well understood. Performance may be different from detector to detector, with polarity of bias, and even with the environment where it operates (at a surface facility versus underground). A notorious problem is that the charge state follows the history of detector operation, and hysteretic effects can make this difficult to study.

In figure 1.7, we see how the ionization signal from a calibration source degrades over time and needs to be refreshed with a period of grounding to neutralize the crystal.

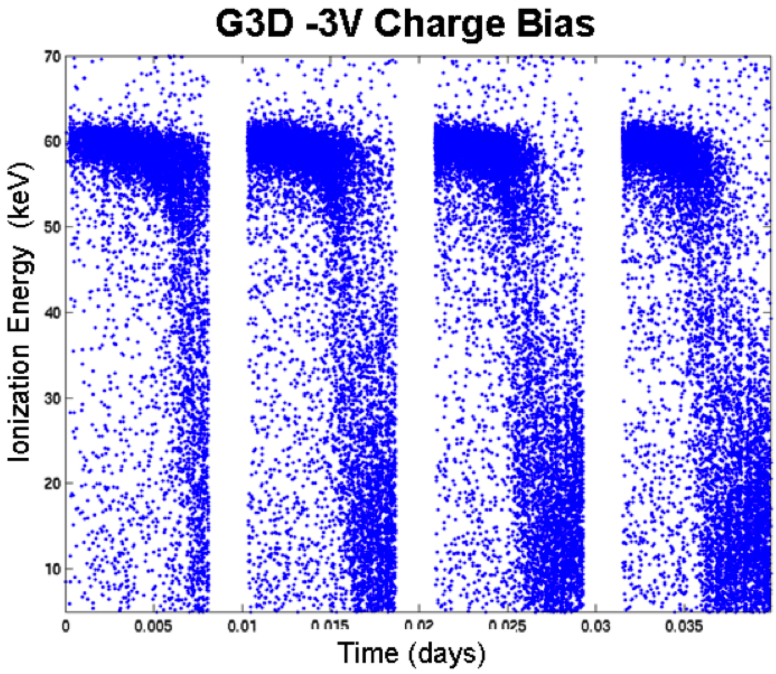

Figure 1.7: This shows the calibrated ionization collection of 60keV events from ^{241}Am, for a CDMS detector at a test facility. The collection response decays with time. Blank periods represent grounded periods of neutralization, after which full collection is rejuvenated. From C. Bailey [3].

Other specific examples are represented in figure 1.8. Here, collimated ^{241}Am check sources are incident on the CDMS detector faces. Using phonon parameters, the events from specific sources can be distinguished and are plotted in different colors. As a function of time and applied bias, a variety of behavior can ensue. In figure 1.8a., a negative potential is applied on the ionization electrodes. Events that represent holes drifting across the bulk of the detector (in green and blue traces) will bifurcate their response after a time. Some events are very poorly collected, some are unperturbed, and some gain an excess of charge. This infers that both capture and ionization rates are at work, in different regions of the detector, likely induced by fields due to the accumulation of space charge with time. In figure 1.8b., we see a "bounce" phenomenon in the blobs representing electrons (black, magenta, cyan traces). Here, electron collection is suppressed after a time, but mostly recovers later on. This is sort of "2nd-order" response is an indicator that multiple processes are at work in our detectors. These processes may be interactions between bulk and surface, and/or between generation and recombination rates. In general, we find we require more predictive power regarding possible processes for electron and holes under these conditions.

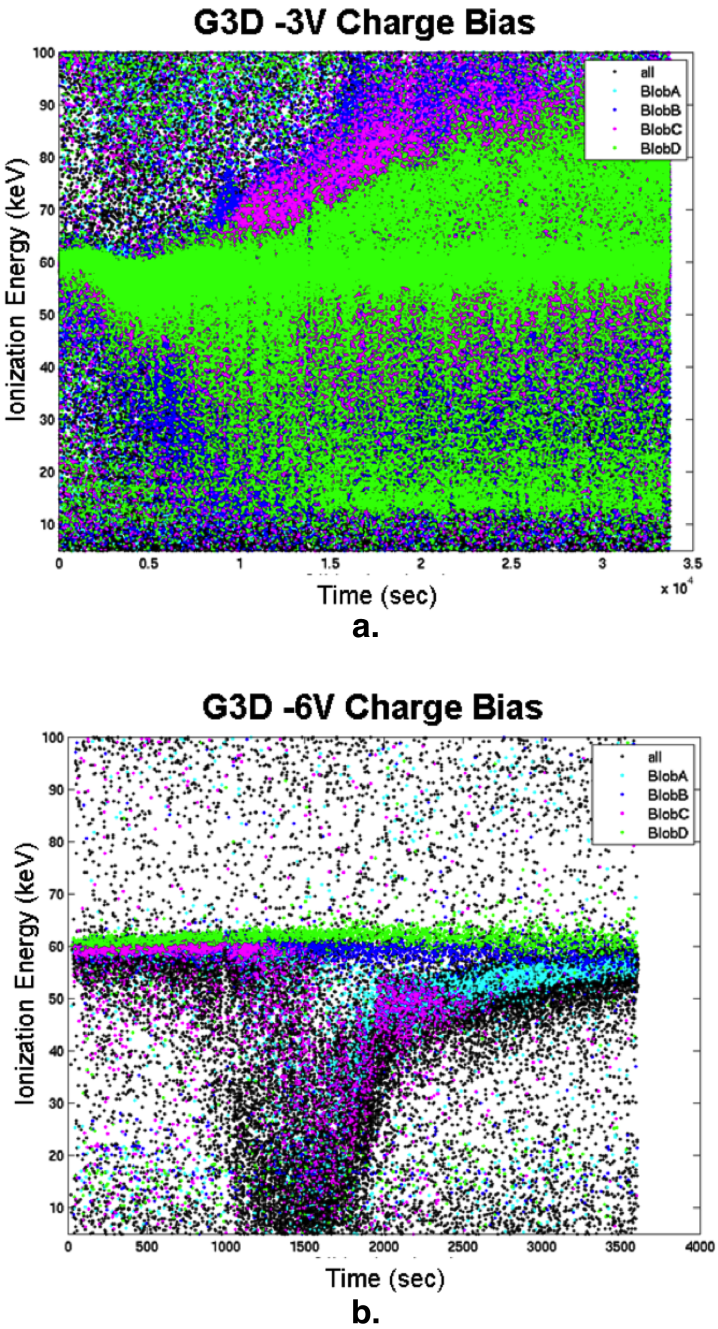

Figure 1.8: Charge collection of 60keV events from ^{241}Am, itemized by different collimated sources. In **a.**, after some time, events where holes traverse the bulk of the detector (green, blue) show bifurcation into poorly collected, well-collected, and excess charge collected regions. In **b.**, electrons (cyan, magenta) show poor collection after some time, but later seem to recover to some extent. From C. Bailey [3].

1.6.2 The dead layer

The dead layer problem is created by effects of surface trapping, which amount to near-surface events which will have incomplete charge collection. This surface trapping is caused by recombination to bound states at the contact, but is problematic due to effects of diffusion. As carriers from a particle event undergo random velocity fluctuations, an appreciable number of them may "back-diffuse" in the relatively small field we are using. This means that carriers that should traverse the crystal and terminate on the electrode on the opposite side simply recombine to the "wrong electrode." They do not induce the image charge that they should, and ionization collection is degraded.

The timing information of CDMS TES sensors provides some help in this respect. Near-surface events allow phonons to quickly decay near metalized surfaces. This means that near-surface events may show phonon timing which is faster than bulk events. Therefore, a quality cut can be made to discriminate against a number of near-surface events.

Regarding the contacts themselves, Shutt [44] made substantial progress in introducing a blockade layer of α-Si between the germanium bulk and electrode contact. The idea is that carriers are less probable to recombine to the contact if there is a wide-gap material in series. Should the wrong-sign carrier "back-diffuse" towards the wrong contact, there is less probability it will recombine. It will reflect from the surface and more likely continue onwards across the crystal. A problem with this tactic is that overall recombination rates are suppressed. This means that carriers are less able to leave the detector and space charge can accrue immediately at the contact. A remedy was to introduce boron into the α-Si, introducing an extra conduction path such that carriers will clear on long times scales. It is doubtful this situation is optimized, however.

1.6.3 Position dependence

Another problem is that there are examples of detectors with a lateral (xy) position dependence in the amount of charge collected. This phenomenon is not understood. The amplitude of the measured ionization signal to a calibrated source may be different by $\sim 10\%$ across the face of the crystal. This could be a consequence of the contacts or of the bulk. If this is a property of the contacts, then it could be a result of something geometric with the structure of the contact, a material property of work functions of the contacts, and/or somehow a difference in surface recombination rates.

So far, we correct for these effects with an empirical position correction. See figure 1.9. It is not clear how these corrections impact the phonon performance, or how these corrections should evolve in time.

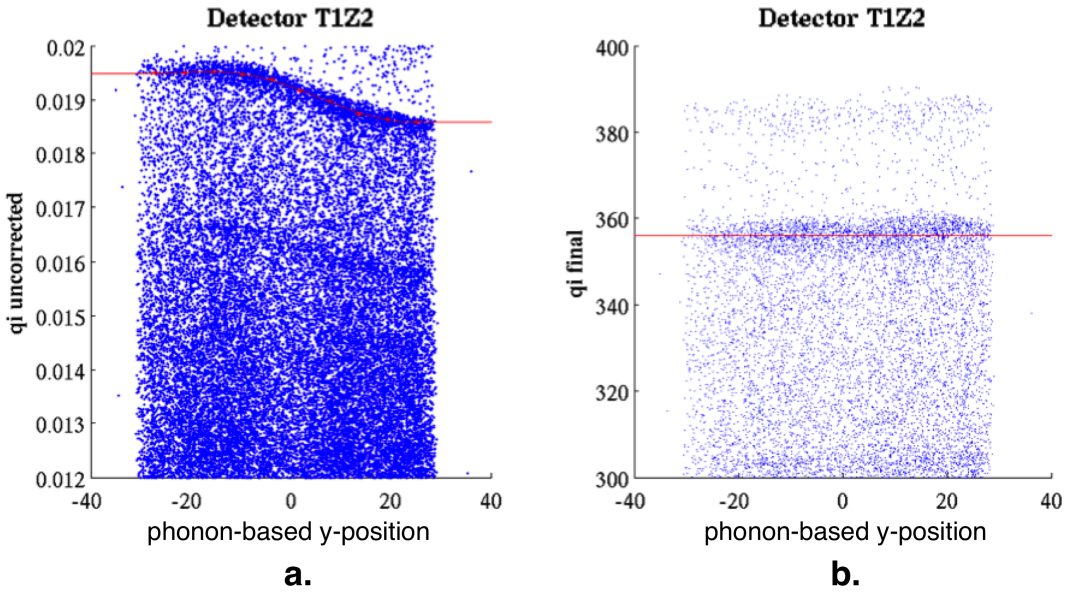

Figure 1.9: Lateral position dependence of charge collection. **a.** Charge collection, in arbitrary units, showing the detector response along the "y" lateral direction. The position was determined by triangulation of phonon parameters. **b.** The same data with a position correction used for subsequent analysis. This data is zoomed in, and calibrated in units of initial photon energy.

1.6.4 Concluding non-ideal effects

So we see that have a variety of detector phenomena in our CDMS detectors for which we must account. Most often, we have found ourselves with only speculation to guide us. Although germanium is a well understood semiconductor, milliKelvin operation with applied fields represents a difficult, non-equilibrium regime to make predictions based on textbook formulae. A more through understanding is required, which we now introduce.

Chapter 2

Band structure and Carrier Transport Processes in Milli-Kelvin Germanium

2.1 The need for simulating carrier transport

To further our understanding of CDMS detector phenomenology, we wish to understand the possible processes that electrons and holes may undergo while under transport in our extreme bias conditions.

To recap, CDMS measures both the ionized charge and the energy in "athermal" (non-equilibrium) phonons created by particle interactions in ultra-pure germanium crystals at a temperature of 40 mK. Ionization and phonon signals are measured simultaneously, allowing us to distinguish between electron and nucleon recoils. This measurement technique provides discrimination between the expected signal of WIMPs and electromagnetic background.

In accomplishing this, charge collection potentials must remain at only a few volts, else emitted phonons from drifted carriers will dominate the phonons of the original interaction. At these drift fields, there are practically no thermal phonons and carrier transport is determined by phonon emission.

In absence of radiation there are no ambient carriers. When excited by an external mechanism such as a radioactive source, the excited carriers are not in equilibrium with the lattice while they are free. So, these carriers are extremely "hot." Furthermore, the concentration of impurities at $10^{10}\ cm^{-3}$ is quite low, such that it is a poor approximation to take a typical assumption that low-temperature scattering is dominated by impurity sites.

So CDMS detectors operate in an exceptional limit. Electrons and holes, over the range of typical bias values, are in a mobility *quantum limit*. This limit is also known as the *zero-point limit* [5]. It describes the case where scattering is predominantly

determined by quantum zero-point fluctuations of the lattice ions. This results in the only dissipation mechanism being the spontaneous emission of phonons. The lattice temperature, which induces stimulated phonon emission, has no significant impact. This zero-point limit is contrary to the *equipartition limit* [5], where phonons are emitted and absorbed in roughly equal number, stimulated by *thermal* excitations of the lattice.

As there are no ambient carriers, we must also reconsider other device assumptions. Under standard conditions, we might typically state Ohm's law as [20]

$$\mathbf{J} = \sigma \cdot \mathbf{F} \tag{2.1}$$

where conductivity is a tensor defined as

$$\sigma = ne^2 \left[\mathbf{m}^{*-1}\right] \langle \tau_m \rangle \tag{2.2}$$

with n, the ambient carrier density, e the electric charge, $\left[\mathbf{m}^{*-1}\right]$ the effective inverse mass tensor, and τ_m the momentum relaxation time. Without a substantial n in our freeze-out conditions, we have no "conductivity" of which to speak. Additionally, the momentum relaxation time itself has a substantial field dependence under our conditions, which would further complicate this linear form of Ohm's law. In fact, several scattering processes make up this rate, with each their own field dependence. To be formal, we can even say that rates with varying energy dependencies invalidates the use of the often-used Mathieson's Rule [45],

$$\frac{1}{\tau} = \Sigma_i \frac{1}{\tau_i}, \tag{2.3}$$

as it becomes possible in this case that rates are correlated. Therefore, these simple relations may not be useful ones for our considerations. We need to rethink precisely what quantities we are allowed to use, and base these on a microscopic understanding of charge transport processes.

It is worthwhile to note that a priori estimations of carrier distribution functions can be inaccurate. We know typical carriers energies under bias are large enough such that they emit acoustic phonons while under drift. This again is the Luke-Neganov effect [40, 39], which generates what we call "Luke phonons." For this to occur, the kinetic energy, E_k, of a carrier must be large enough to emit a phonon by momentum and energy conservation. This amounts to $E_k > \frac{1}{2} m_c v_s^2$, where m_c is carrier's conductivity mass, and v_s is the speed of sound. It also happens to be true that the average carrier energy is not only above the threshold for acoustic phonon emission, but also typically below the threshold for emission of *optical* phonons of energy $\hbar\omega_0$. In other words,

$$\frac{1}{2} m_c v_s^2 < \langle E_k \rangle < \hbar\omega_0 \tag{2.4}$$

In this scenario, the high-energy portion of the carrier energy distribution may substantially exceed the threshold for optical phonon emission. The emission of an optical phonon is an extremely inelastic process. If a carrier emits an optical phonon, its energy is "reset" to a quite low value. So, transport between these thresholds can be nonlinear. The resulting carrier distributions may be difficult to predict with simple analytical approximations based on perturbative, near-elastic scattering.

In this chapter, we will review some basics about semiconductor physics with an emphasis on germanium. We review the origin and implications of band structure, due to the lattice nature of semiconductors. We introduce carrier dynamical equations, as well as scattering mechanisms and rates. As will become useful in later chapters, we introduce the importance of energy and momentum conservation for implementing scattering rates.

2.2 A review of semiconductor physics

First, we briefly review the necessary semiconductor fundamentals relevant to the treatment of quasiparticles in germanium. We introduce the various kinds of potentials that carriers experience while within a crystal. We treat many-body effects and the periodicity of the lattice by introducing *band structure*. We then introduce the simplifying concept of effective mass and quasi-momentum, as well as density of states.

2.2.1 The independent-particle Schrödinger equation, and various potentials

Solids are comprised of a large number of interacting particles, of which the resulting dynamics would naively seem to pose an intimidating problem. Fortunately, a simplifying approximation is made where the motion of individual particles can be treated in a mean-field approach, approximating the effect of all the other particles as an average force. One may call this the independent-particle approximation, and it is used to justify the use of the Schrödinger equation for a single carrier.

$$i\hbar\frac{\partial}{\partial t}\psi_0(\mathbf{r},t) = -\frac{\hbar^2}{2m_0}\nabla^2\psi_0(\mathbf{r},t) + U(\mathbf{r},t)\psi_0(\mathbf{r},t) \tag{2.5}$$

Above, the subscript "0" is used to denote the properties of an electron wavefunction with no approximation applied to the potential, $U(\mathbf{r},t)$.

An insightful way to approach this problem is to appreciate that there are many contributions to the potential energy, $U(\mathbf{r},\mathbf{t})$. In this light, we follow an argument presented by Supriyo Datta [4]. We itemize the contributions to the potential energy, and then describe how we account for them independently.

$$U(\mathbf{r},t) = U_E(\mathbf{r},t) + U_L(\mathbf{r}) + U_S(\mathbf{r},t) \tag{2.6}$$

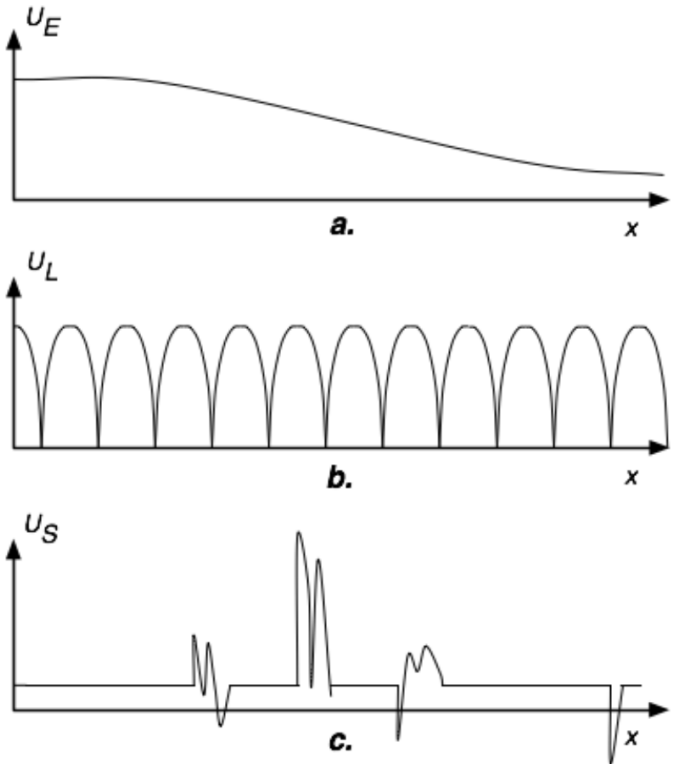

Figure 2.1: Here, we note three distinct categories of potential energy which appear in Schrödinger's equation. **a.** the externally applied potential. **b.** the periodic and stationary lattice potential. **c.** the random scattering potential. Reproduced from Datta [4].

The term $U_E(\mathbf{r}, t)$ is an external, macroscopic potential that may arise due to the presence of an externally applied potential or from space charge due to charged impurities. The *external* potential appears in the equations of motion which describes carrier acceleration. Next, we include the microscopic, periodic potential of lattice ions in the crystal by introducing the potential term, $U_L(\mathbf{r})$, to account for the lattice. Most concepts we use in depicting band structure and effective mass come from idealizing the effect of this term. Finally, we account for the microscopic fluctuations of the periodic crystal by $U_S(\mathbf{r}, t)$, which is a random and fluctuating term due to phonons, impurities, or crystal defects. We treat this term as the *scattering* term, and include it by way of perturbation theory and Fermi's golden rule.

In what follows, we treat band structure effects due to the lattice potential, and then include them into the equations of motion which include the external potential. We then describe scattering mechanisms where various scattering potentials play a role.

2.2.2 Dealing with the lattice potential, U_L: band structure, quasi-momentum, and effective mass

In the crystal lattice, the presence of all electrons is by and large accounted for in a mean-field, single-carrier approach by $U_L(\mathbf{r})$, the lattice potential. The lattice is responsible for the material's underlying electronic band structure arising from the many-body Schrödinger equation. A vast number of condensed matter basic references cover this material in depth. We point out only the most basic understanding required for the discussions that follow.

This periodic lattice potential provides solutions to Schrödinger's equation in the form of Bloch functions. The electron wavefunction for a given band ν and wavevector $\mathbf{k}$ is written as

$$\psi_0(\mathbf{r},t) = u_{\nu,k}(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})\exp(iE_\nu(\mathbf{k})/\hbar) \tag{2.7}$$

The function $u_{\nu,k}(\mathbf{r})$ shares the periodicity of the lattice potential. Incorporating these Bloch wavefunctions into to Schrödinger's equation, we find eigenvalue energy solutions. A thorough treatment, again common in a number of textbooks, shows that $E(\mathbf{k})$ is a function of $\mathbf{k}$. We call this relation of $E(\mathbf{k})$ to $\mathbf{k}$ the *band structure* or the *dispersion relation*. We recover the energy band gap, the location of conduction and valence band-edges, and in fact knowledge of the whole dispersion relation in k-space known as Brillouin Zones.

The dispersion relation

Considering that dispersion relations are generally parabolic near the band minimum (maximum) where free electrons (holes) are typically found, we can approximate

$$\mathrm{E}(\mathbf{k}) = \frac{\hbar^2}{2}\mathbf{k}^T\left[\mathbf{m}^{*-1}\right]\mathbf{k} \tag{2.8}$$

where $\left[\mathbf{m}^{*-1}\right]$ is the *inverse effective mass* tensor.

At times, a non-parabolic band structure is defined, adding a higher-order correction to the dispersion relation as carriers depart from the local band extremum. In this case, to first order, the dispersion relation will be defined as

$$\mathrm{E}(\mathbf{k})\,(1+\alpha\mathrm{E}(\mathbf{k})) = \frac{\hbar^2}{2}\mathbf{k}^T\left[\mathbf{m}^{*-1}\right]\mathbf{k} \tag{2.9}$$

The term α has been coined as the *nonparabolicity* factor. To note, this is a simple constant that attempts to account for the first non-parabolic term in an expansion. A full dispersion relation might require several terms to adequately account for a non-parabolic $\mathbf{E}(\mathbf{k})$, furthermore the directionality of these higher terms might need to be included as well. Analytical corrections above and beyond the α constant do not typically appear in the literature, however.

Quasi-momentum versus momentum: *group* velocity

The wavevector, $\mathbf{k}$, provides the *quasi-momentum* or *crystal momentum* of the carrier state. However, as pointed out in several standard textbooks such as [46], it is critical to understand that quasi-momentum is not identical to the momentum originating from the *group velocity*. The group velocity of a carrier is related to the dispersion relation [46] as

$$\mathbf{v_g} = \frac{1}{\hbar}\nabla_{\mathbf{k}}\mathrm{E}(\mathbf{k}) \tag{2.10}$$

Considering that this expression represents the gradient across surfaces of equal energy in k space, then carrier velocities must be perpendicular to this surface. Generally, the same is not true for any particular quasi-momentum wavevector, $\mathbf{k}$. Therefore, quasi-momentum is in general not co-linear with the actual group velocity. Instead, the following relation applies for the case of a parabolic band.

$$\mathbf{v_g} = \hbar\left[\mathbf{m}^{*-1}\right]\cdot\mathbf{k} \tag{2.11}$$

Above, off-diagonal elements in the inverse mass tensor may allow the quasi-momentum and group velocity to be not parallel. In summary, the effective mass accounts for the structure of the dispersion relation, $\mathrm{E}(\mathbf{k})$, which was deduced by accounting for the U_L term in Schrödinger's equation.

Next, we continue to develop an understanding of the effective mass by considering electron kinematics due to the externally applied potential, U_E.

2.2.3 Dealing with the external potential, U_E: carrier dynamics and specifics of the effective mass

Under the influence of an external force, $\mathbf{F}$, the wavevector of a Bloch-state electron can be shown by equation 2.5 to change as

$$\boxed{\mathbf{F} = -\nabla U_E = \hbar\dot{\mathbf{k}}} \tag{2.12}$$

As the wavevector changes, the corresponding acceleration changes the group velocity by way of equation 2.11.

In considering the valence band which is generally filled by electrons, an unfilled state acts also according to this acceleration equation. This leads to the concept of *holes*. They undergo similar dynamics to electrons but with motion in the opposite direction. They therefore act as particles in their own right, but with a positive charge.

The standard definition of effective mass

We now use classic arguments [46, 45, 47, 48] to define the effective mass, so that we may afterward introduce an important but lesser-known definition.

We start by energy conservation arguments, closely following those found in reference [49]. While under the influence of an external force, a Bloch electron's energy will increase infinitesimally as

$$dE = F\ v_g\ dt \tag{2.13}$$

One can therefore relate the change in energy to the acceleration equation as

$$|F| = \frac{1}{v_g}\dot{E} = \hbar\dot{k} \tag{2.14}$$

This, in turn, gives

$$\dot{v}_g = \frac{1}{\hbar}\frac{d}{dt}\left(\frac{dE}{dk}\right) = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt} = \frac{1}{\hbar^2}\frac{d^2E}{dk^2}F \tag{2.15}$$

We may relate the external force to the change of group velocity as

$$F = m^*\left(\dot{v}_g\right) \tag{2.16}$$

then we would find the effective mass as depicted in standard solid-state textbooks such as [46].

$$m^* = \left(\frac{1}{\hbar^2}\frac{d^2E}{dk^2}\right)^{-1} \tag{2.17}$$

In more general form, we can write the inverse tensor

$$\boxed{\left[m^{*-1}\right]_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}} \tag{2.18}$$

This definition is standard in the literature, and we label equation 2.18 as the *standard effective mass*, or *band effective mass*.

2.2.4 Ferry's effective mass

An important, straightforward, but widely unknown observation is made by David K. Ferry, in reference [50]. We briefly revisit his argument in order to define his expression for effective mass, which differs from the standard result.

Quite simply, we know that force is the time-derivative of momentum.

$$F = \dot{p} = m^*\left(\dot{v}_g\right) + (\dot{m}^*)v_g \tag{2.19}$$

Yet, in equation 2.16 above, we did not include the time-derivative of the mass itself. This is Ferry's critical point. It may often be the case that the effective mass has a dependence on energy or momentum, causing it to change over the course of a trajectory. If we see this through, the force

$$F = \left(m^* + \frac{dm^*}{dv_g}v_g\right)(\dot{v}_g) = \left(m^* + \frac{dm^*}{dv}v_g\right)\left(\frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt}\right) = \hbar\frac{dk}{dt} \tag{2.20}$$

can be combined with the simple relation that

$$\hbar k = m^* v_g \tag{2.21}$$

Solving this, Ferry determined that the effective mass for this equation must be related by

$$m^*_{Ferry} = \left(\frac{1}{\hbar^2}\frac{1}{k}\frac{dE}{dk}\right)^{-1} \tag{2.22}$$

In full tensor notation, this is generalized to

$$\boxed{\left[m^{*-1}_{Ferry}\right]_{ij} = \frac{1}{\hbar^2}\frac{1}{k_i}\frac{\partial E}{\partial k_j}} \tag{2.23}$$

We define equation 2.23 as the *Ferry effective mass*. We generally use the traditional defintion of effective mass in this work, but Ferry's treatment does suggest an interesting correction to consider.

2.2.5 The effective mass of electrons in an ellipsoidal valley

In general the effective mass dependence on a constant energy surface is anisotropic and can often be approximated as *ellipsoidal*, as is the case for the L-valleys in germanium [51].

Consider the energy-momentum relation to be parabolic. We have a longitudinal axis parallel to the principal axis, z, and two transverse directions along x and y. We have constants C_1 and C_2.

$$E(k_x, k_y, k_z) = C_1(k_x^2 + k_y^2) + C_2 k_z^2 \tag{2.24}$$

Along the z axis, we can define a longitudinal mass from either the standard definition of effective mass or the Ferry effective mass.

$$1/m_l = 1/m_{zz} = \frac{2C_2}{\hbar^2} \tag{2.25}$$

Likewise, we can define a transverse mass from either effective mass expression.

$$1/m_t = 1/m_{xx} = 1/m_{yy} = \frac{2C_1}{\hbar^2} \tag{2.26}$$

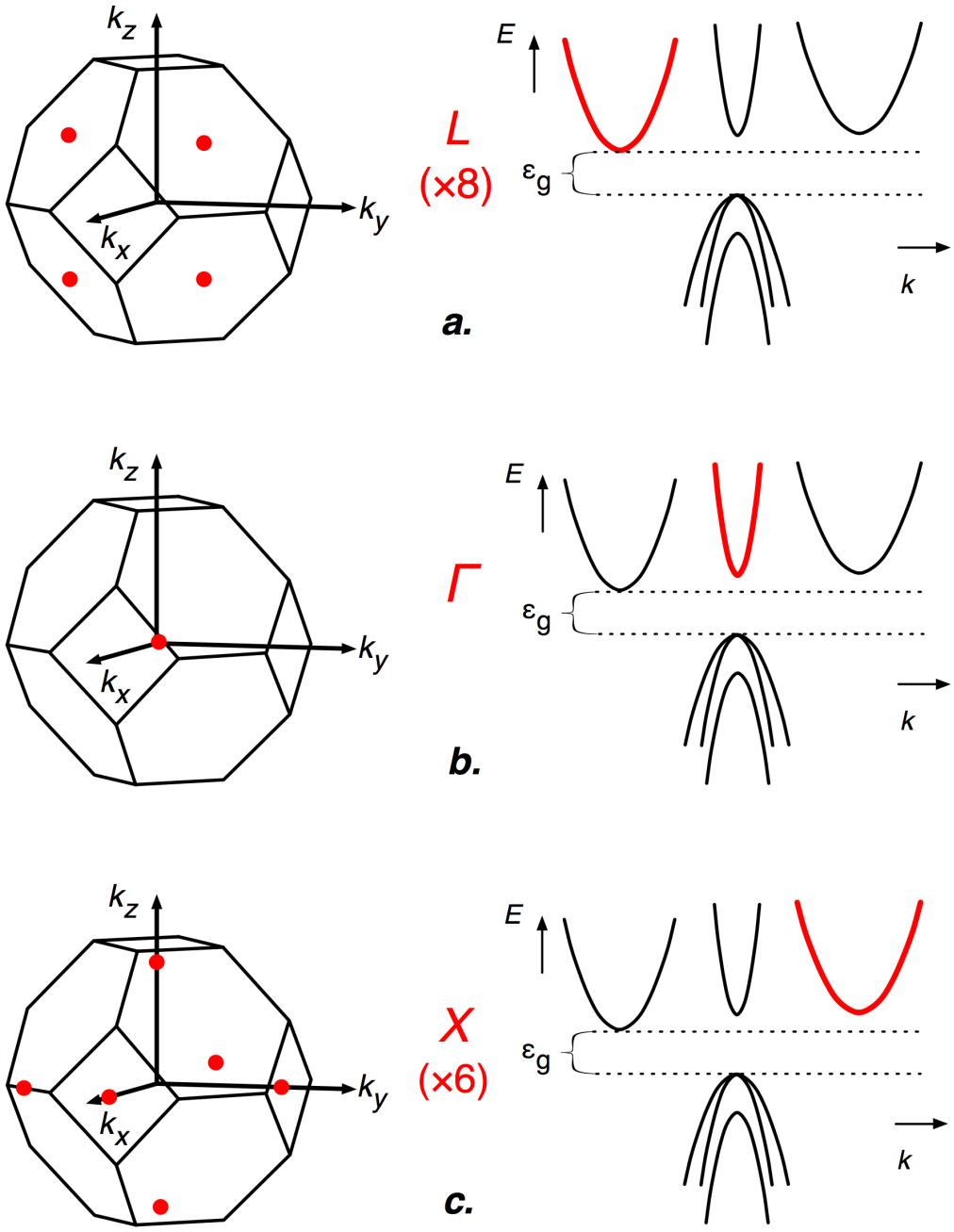

Figure 2.2: The multiple valley band structure of germanium. **a.** the L valleys, the lowest and dominant conduction band **b.** the Γ valley **c.** the X valleys

As a full tensor, the standard inverse mass tensor is therefore the following.

$$\left[\mathbf{m}^{*-1}\right] = \begin{pmatrix} 1/m_t & 0 & 0 \\ 0 & 1/m_t & 0 \\ 0 & 0 & 1/m_l \end{pmatrix} \tag{2.27}$$

In contrast, Ferry's effective mass expression provides non-trivial, off-diagonal components. These terms contain ratios of quasi-momentum components.

$$\left[\mathbf{m}^{*-1}_{Ferry}\right] = \begin{pmatrix} 1/m_t & \left(\frac{1}{m_t}\frac{k_y}{k_x}\right) & \left(\frac{1}{m_l}\frac{k_z}{k_x}\right) \\ \left(\frac{1}{m_t}\frac{k_x}{k_y}\right) & 1/m_t & \left(\frac{1}{m_l}\frac{k_z}{k_y}\right) \\ \left(\frac{1}{m_t}\frac{k_x}{k_z}\right) & \left(\frac{1}{m_t}\frac{k_y}{k_z}\right) & 1/m_l \end{pmatrix} \tag{2.28}$$

Rotation of a mass tensor

Now that we have introduced the concept of an inverse mass tensor *along an ellipsoid principal axis*, it is also worthwhile to point out the necessary rotational transformations. In germanium, the L-valleys are oriented along the directions $\langle 111 \rangle$. To rotate these valleys into the "crystal" or "lab frame," rotation matrices, R, are required. To note, an inverse mass tensor is a 2nd-rank tensor [51, 20], so it must rotate as such.

$$\left[\mathbf{m}^{*-1}\right]_{\text{lab}} = R\left[\mathbf{m}^{*-1}\right]_{\text{principle}} R^{-1} \tag{2.29}$$

These rotations align a mass tensor from the reference frame of the principle axis (associated with the longitudinal mass), to the lab frame. Note that distinct rotation matrices are required for each electron valley.

Isotropic approximations to the electron effective mass

In order to simplify our work, we often make use of isotropic assumptions. It is important to specify how exactly we can take averages of the inverse mass tensor of equation 2.27.

Conductivity mass The *conductivity effective mass* for electrons is the typical value of the mass tensor found during acceleration. Here, we use the harmonic mean of the effective mass elements. For kinematic considerations of an individual electron, we use this mass.

$$\frac{1}{m_c} = \frac{1}{3}\left(\frac{1}{m_l} + \frac{1}{m_t} + \frac{1}{m_t}\right) \tag{2.30}$$

Density-of-states mass Considering the occupation of electrons in a density of available states, we often encounter approximations which require a *density-of-states mass*. This mass is used for calculations which consider the number of states available to an electron [20].

$$m_{DOS} = \left(m_l m_t^2\right)^{1/3} \tag{2.31}$$

2.2.6 The effective mass for holes in different bands

We have three bands to consider with holes; heavy, light, and split-off, as in figure 2.3. The heavy and light bands share a degeneracy level at $k = 0$. The proximity of these levels to each other leads to interactions which make the dispersion relation complicated [52, 5]. In the absence of strain, the dispersion relations can be written as the following.

$$\begin{aligned} E_h(k) &= -\tfrac{\hbar^2}{2m_e}\left\{Ak^2 - \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2\right)}\right\} \\ E_l(k) &= -\tfrac{\hbar^2}{2m_e}\left\{Ak^2 + \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2\right)}\right\} \end{aligned} \tag{2.32}$$

Here, h stands for the heavy band, and l for the light band.

The split-off level is separated from the heavy and light bands [53]. It is often neglected. We treat later it in our full anisotropic simulation, but only with a simple scalar mass and parabolic dispersion relation.

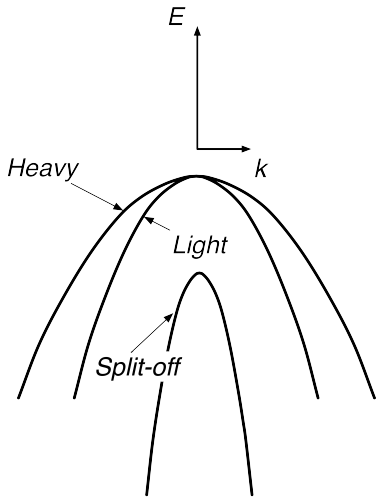

Figure 2.3: The (idealized) dispersion relation for holes. There are three bands: heavy, light, and split-off. As holes occupy the heavy band most of the time, it dominates the transport properties.

The effective mass for holes is complicated by the warped nature of these valence bands. We come across three definitions of mass for holes, in different contexts.

Isotropic effective mass

There are *isotropic* average masses defined for each hole band [53] (i.e., m_h, m_l, or m_s), which also serve as a density-of-states mass for the different hole bands [54].

Dispersion effective mass

Although the dispersion relations are complicated, we can also consider defining mass *without* derivatives, but simply to make a parabolic dispersion work out in a given direction. In other words, we can define a directional "dispersion mass" by the following.

$$\begin{aligned} \frac{\hbar^2 k^2}{2m_h(\theta_k,\varphi_k)} &= E_h(k) \\ \frac{\hbar^2 k^2}{2m_l(\theta_k,\varphi_k)} &= E_l(k) \end{aligned} \tag{2.33}$$

In some references, mass defined in this way has been used for all kinematic concerns [54], regardless of equation 2.18.

Standard, or kinematic effective mass

Consider the full dispersion-curve derivative for the inverse mass tensor from equation 2.18. This, we will call the *standard* mass definition. It is appropriate for how the wavevector $\mathbf{k}$ is evolving as a group velocity. Therefore, this definition of mass is appropriate for how holes accelerate, evolving kinematically in a differential sense. Since the dispersion relation is more complicated, the standard mass definition does not lend itself well to a matrix formalism as well as for electrons. In the inverse mass tensor using equation 2.18, each element depends on all $\mathbf{k}$-vector components. Although there is an analytical form for every element, it remains cumbersome to depict for this reason.

2.3 Scattering mechanisms: dealing with the random scattering potential, U_s

We now discuss carrier scattering mechanisms. As carriers propagate through the germanium medium, they encounter quantum mechanical processes which cause them to scatter from some initial momentum state $\mathbf{k}$ to a new momentum state $\mathbf{k}'$. Several different kinds of processes may cause these scattering events, namely the emission or absorption of a phonon, a collision with a defect or impurity, or collisions with other carriers.

Under the scattering conditions of the CDMS experiment, the temperature is low and the purity of the detector crystals is high. Well below carrier freeze-out temperatures, we do not consider carrier-carrier interactions. Any carrier-carrier interactions during the initial deposition of a nuclear-physics scale particle or photon are far from the field steady-state conditions we attempt to depict. Although we go into the details for carrier scattering on crystal impurities, we consider this rate as generally a small term for our high-purity detectors. We largely neglect it in our final treatments, but we do elaborate on this topic later on. We will describe phonon absorption and emission. By far, the scattering rates for electrons and holes is dominated by only phonon emission processes. As a reminder, this is indicative of the "zero-point" regime where phonons are emitted far more than they are absorbed.

In what follows, we describe the general categories of scattering processes. We then introduce Fermi's golden rule, and how it is used to depict scattering rates. This is similar to classic introductions as found in the popular references [7, 12, 45, 5, 48]. We will then analyze in depth the case of impurity scattering, followed by phonon scattering at substantial length.

2.3.1 General categories of scattering processes

There are a number of general interaction mechanisms that introduce scattering potentials, causing electrons and holes to stochastically scatter between momentum states. Here, we outline these physical processes particular to germanium. Broadly, scattering mechanisms are introduced by either

- some sort of defect or impurity site in the germanium crystal, or
- the probability to emit or absorb a phonon, or
- the probability for carriers to scatter off one another.

The general types of scattering processes are summarized by figure 2.4. We briefly summarize these categories, and the relevance of different kinds of scattering to the operating conditions in CDMS germanium detectors at $T = 40\ mK$.

Impurity scattering

Both electrons and holes may scatter while in the proximity to impurities or crystal defects. These potentials are localized and time-independent. Carrier scattering is assumed to be elastic in this case, as there is no momentum assumed to be imparted to the impurity or crystal as a whole. As a consequence, the primary outcome of impurity scattering is only to randomize the electron or hole distributions, as energy is not lost. These scattering processes are assumed to be due to either charged (*ionized*) impurities, or to *neutral* impurities. Of the ionized impurity processes, the

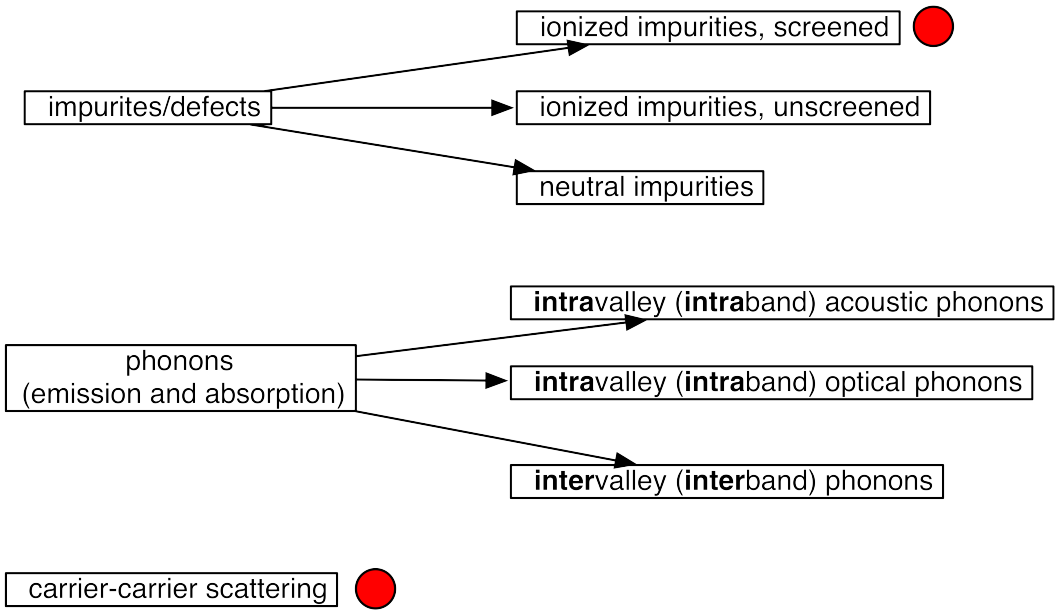

Figure 2.4: General categories of scattering processes for both electrons and holes. *Solid dots:* represent processes that need not be considered for sub-Kelvin operation. Also to note, phonon emission rates greatly outnumber absorption rates at sub-Kelvin temperatures.

scattering rate is treated with a different formalism depending on whether or not the ionized impurity potential includes screening effects.

It is worthwhile to consider the difference between screened and unscreened ionized impurity scattering. In room-temperature semiconductors, there are enough ambient free carriers throughout the bulk such that their presence can *screen* the Coulomb potential of ionized impurities. This is referred to as *Debye screening* [55, 7, 45]. In considering carrier scattering by screened impurities, rates are based on the *Brooks-Herring* [56] model of impurity scattering. At sub-Kelvin temperatures, the ambient carrier concentration is practically non-existant. Screening is not a concern. In this case, the *Conwell-Weisskopf* [57] treatment for unscreened ionized impurities is the appropriate formulation. The Conwell-Weisskopf approach basically reduces to a form of Rutherford scattering on the impurity's Coulomb potential. As an aside, formulations by Ridley [58] successfully bridge the screened and unscreened limits in a continuous formulation of ionized impurity scattering.

Neutral impurities also present a source of scattering potentials for incident electrons and holes. Although there is no Coulomb repulsion or attraction, there is a short-range interaction that can be modeled basically as a spherical square well potential. The earliest notable treatment of neutral impurity scattering is that of *Erginsoy* [59]. A slightly more sophisticated treatment is that of Sclar [60], which

includes a dependence based on the incident carrier's energy. For neutral impurity scattering, we use Sclar's formulation. However, a simple inspection under Erginsoy's formulation shows that neutral impurity scattering is a rare process for neutral impurity levels of $10^{10}\ cm^{-3}$. A carrier is expected to traverse nearly a meter before scattering on a neutral impurity at that concentration. So, although we can include neutral impurity scattering, it is not a significant contribution to the overall energy and momentum distributions of drifting electrons and holes.

Scattering by both ionized and neutral impurities has a strong energy dependence. Intuitively, as carriers have lower incident energy, they remain in proximity to the scattering potential for a longer period of time. This can be seen as why the scattering probability is higher when the carrier incident energy is lower. As it turns out, under the low-temperature operation of CDMS detectors, carriers can reach high energies with relatively low applied electric fields. For typical bias conditions with fields of $\sim V/cm$, carrier scattering by either ionized or neutral scattering is negligible for concentrations of $10^{10} - 10^{11}\ cm^{-3}$. This is fortunate because, a priori, we don't *know* what fractions of the impurity concentrations are ionized versus neutralized. We must calculate energy distributions first before we can make predictions for ionization and neutralization rates. Therefore, our overall approach is to neglect carrier scattering in our transport calculations. One could imagine that if impurity scattering effects (say for fields at the $\sim\ 10\ mV/cm$ scale) should be considered, an iterative approach between transport calculations and ionization calculations could be computed.

Phonon scattering

Phonons are quantized vibrations of the crystal lattice. These vibrations represent time-dependent displacements of the lattice ions. These local displacements serve to locally alter the relative band energies. As seen by free carriers, phonons present a time-dependent source of scattering potentials.

Phonon scattering is inelastic. Phonon absorption (emission) represents the absorption (emission) of phonon energy quanta by carriers, therefore increasing (lowering) carrier energy. Of the many different phonon processes, some processes are "intravalley" meaning they transfer electrons to new momentum states in the *same* valley. The term happens to be "intraband" for phonons processes that scatter holes within the same band. Phonon processes that scatter electrons (holes) to different valleys (bands) are referred to as "intervalley" ("interband") processes.

Phonons represent the vast majority of scattering processes needed to depict transport in high-purity crystalline germanium at sub-Kelvin temperatures. There are numerous individual phonon processes. There are *acoustic* phonons, and *optical* phonons. In a three-dimensional germanium crystal, there are separate *branches* of phonons (one longitudinal and two transverse) for both acoustic and optical phonons.

There are probabilities for carriers to both emit and absorb phonons, although there are essentially no ambient phonons around for carriers to absorb at sub-Kelvin

temperatures. In fact, it is this *absence* of ambient phonons (for the drift fields of interest) that determines the unique transport limit in which CDMS detectors operate. As a reminder, carrier scattering is dominated by spontaneous phonon emission (or collision with zero-point fluctuations of the lattice).

Calculating the transport conditions limited by phonon emission represents the majority of the work we must perform.

Carrier-carrier scattering

We do not treat carrier-carrier scattering in our treatment of electron and hole transport. Carrier-carrier scattering is the interaction of carriers with each other. This rate is dependent on the concentration of carriers generally occupying the same volume. Carrier-carrier scattering is more difficult to treat than scattering processes involving only one carrier, typically because the joint energy distribution between carriers needs to be established. Fortunately, as sub-Kelvin temperatures are well below the freeze-out temperature of carriers in germanium, we are safe in neglecting carrier-carrier scattering.

2.3.2 An introduction to Fermi's golden rule and scattering processes

In this section, we determine ideal scattering rates by perturbation theory, largely following Datta [4]. We point out the required understanding of the scattering potential and the conservation of energy and momentum in order to arrive at a total, integrated scattering rate.

The random scattering potential gives us Fermi's golden rule

The principle is that the random scattering potential, demonstrated in figure 2.1, is responsible for a first-order quantum mechanical perturbation coupling a carrier of state $\mathbf{k}$ to a state $\mathbf{k}'$. Therefore, we define the *matrix element* which couples an electron of wavevector $\mathbf{k}$ to a wavevector $\mathbf{k}'$ by this random scattering potential U_s.

$$\langle U_s \rangle_{\mathbf{k},\mathbf{k}'} \equiv \int \frac{d\mathbf{r}}{\Omega} \psi_{\mathbf{k}'} U_s \psi_{\mathbf{k}} \tag{2.34}$$

Fermi's golden rule is the first-order transition rate associated with this matrix element. This is formula is standard.

$$\Gamma(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} |\langle U_s \rangle_{\mathbf{k},\mathbf{k}'}| \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) \tag{2.35}$$

For Bloch waves, the matrix element can be broken into a *relative* scattering potential, $U_s(\mathbf{k} - \mathbf{k}')$, and an *overlap integral*, $I(\mathbf{k}, \mathbf{k}')$ [45].

$$\langle U_s \rangle_{\mathbf{k},\mathbf{k}'} = U_s(\mathbf{k}-\mathbf{k}')I(\mathbf{k},\mathbf{k}') \tag{2.36}$$

Separating the overlap factor from the potential term allows us to use plane waves instead of Bloch waves. In addition, for parabolic bands the overlap factor is typically near unity.

A total rate, integrated over final states

Equation 2.35 gives us the probability that a carrier in the original state $\mathbf{k}$ is scattered by the perturbation potential $\langle U_s \rangle_{\mathbf{k},\mathbf{k}'}$ to the particular final momentum state $\mathbf{k}'$. While this formulation is correct, it doesn't afford much of a physical understanding because there are typically a huge number of possible $\mathbf{k}'$ states in realistic situations. What is required for a physically useful rate is the *total* rate over all possible $\mathbf{k}'$ final states.

The total scattering rate is the sum over all possible $\mathbf{k}'$ final states.

$$\Gamma(\mathbf{k}) = \frac{2\pi}{\hbar}\sum_{\mathbf{k}'} |U_s(\mathbf{k}-\mathbf{k}')|^2 I(\mathbf{k},\mathbf{k}')^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}})) \tag{2.37}$$

There are in fact many possible final states to "choose from" in realistic situations concerning a macroscopic, bulk semiconductor. We can make use of the *density of states* Ω per unit volume in k-space to replace the discrete sum of equation 2.37 by a corresponding continuous integral. We make the following replacement.

$$\sum_{\mathbf{k}'} \rightarrow \frac{\Omega}{(2\pi)^3}\int d^3k' \tag{2.38}$$

By expressing the integral in terms of spherical coordinates in $\mathbf{k}'$-space, we find the following form.

$$\Gamma(\mathbf{k}) = \frac{\Omega}{(2\pi)^3}\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \Gamma(\mathbf{k},\mathbf{k}')\ k'^2\ dk'\ \sin\theta\ d\theta\ d\phi \tag{2.39}$$

The difficulties: we must investigate $U_s(\mathbf{k}-\mathbf{k}')$ and $\delta(E_{\mathbf{k}'} - E_{\mathbf{k}})$, to arrive at a total scattering rate

By replacing $\Gamma(\mathbf{k},\mathbf{k}')$ in equation 2.39, we can reconsider its constituent terms.

$$\boxed{\Gamma(\mathbf{k}) = \frac{\Omega}{(2\pi)^3}\frac{2\pi}{\hbar}\int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} |U_s(\mathbf{k}-\mathbf{k}')|^2 I(\mathbf{k},\mathbf{k}')^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) k'^2 \sin\theta dk' d\theta d\phi} \tag{2.40}$$

The above equation contains the delta function in energy, $\delta(E_{\mathbf{k}'} - E_{\mathbf{k}})$, which generally serves to simplify the $\mathbf{k}'$-space integral. Only the final states that conserve energy are allowed.

A basic understanding of a particular scattering process may be gained by interpreting it within the framework of equation 2.40. What changes from process to process is primarily the scattering potential matrix element, $U_s(\mathbf{k} - \mathbf{k}')$, as well as the particular form of the delta function $\delta(E_{\mathbf{k}'} - E_{\mathbf{k}})$. So we point out here that the details of solving these integrated rates often can be reduced to understanding the behavior of these two quantities. We will point examples of this in sections to come.

A caveat to Fermi's golden rule

It is worthwhile to a caveat to Fermi's golden rule. Our picture is semiclassical. We assume electrons or holes propagate freely in an electric field, but are interrupted by discrete, instantaneous scattering events. The delta function in Fermi's golden rule is a strict interpretation of energy conservation.

In some physical circumstances, it may be that scattering events happen with such great rapidity, that energy conservation is not strictly determined as per Uncertainty Principle limits. When the uncertainty in the final-state energy after a scattering event is not well depicted by Fermi's golden rule, we say that this transport scenario is affected by *collisional broadening*.

Under the conditions of CDMS detectors, carrier scattering is essentially limited by spontaneous phonon emission. In this case, we have no need to consider collisional broadening. Fermi's golden rule is always acceptable in the work that follows.

We now consider the application of Fermi's golden rule to specific scattering processes.

2.3.3 The impurity scattering rates

Here, we describe the integrated scattering rates appropriate to describe carrier scattering by both ionized and neutral impurities.

Ionized impurities

Ionized impurity scattering is considered as elastic. Under CDMS conditions which includes a temperature well below carrier freeze-out, we need only be concerned with *unscreened* Coulombic impurity potentials. Again, the appropriate treatment for this is the Conwell-Weisskopf [57] formulation.

The unscreened Coulomb interaction gives a scattering potential of

$$\langle U_s \rangle_{\mathbf{k},\mathbf{k}'} = \int \frac{d\mathbf{r}}{\Omega} \psi_{\mathbf{k}'} \left(\frac{Ze^2}{4\pi\kappa\varepsilon_0 r} \right) \psi_{\mathbf{k}} \tag{2.41}$$

Expressing in real-space spherical coordinates, and using Bloch plane waves for the incident and outgoing wavefunctions, we find the following.

$$\langle U_s \rangle_{\mathbf{k},\mathbf{k}'} = \frac{1}{\Omega} \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} (e^{-i\mathbf{k}'\cdot\mathbf{r}})(\frac{Ze^2}{4\pi\kappa\varepsilon_0})(e^{i\mathbf{k}\cdot\mathbf{r}}) r dr \sin\theta d\theta d\phi \tag{2.42}$$

Combining this with Fermi's golden rule, as well as multiplying by $N_I\Omega$ for the number of impurities in the plane-wave volume, one arrives at [45]

$$\Gamma(\mathbf{k},\mathbf{k}') = \frac{\pi N_I e^4 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}})}{8\hbar \left(\kappa\varepsilon_0\right)^2 \Omega k^4 \sin^4\left(\theta_{kk'}/2\right)} \tag{2.43}$$

where $\theta_{kk'}$ is the angle between $\mathbf{k}$ and $\mathbf{k}'$.

What Conwell and Weisskopf first performed [57] was the integrated rate based on this particular formulation of Fermi's golden rule. Their insight comes from the fact that the rate becomes infinite when the angular deflection $\theta_{kk'}$ is zero. To avoid this divergence, they introduced a threshold minimum angle into the integration limits. They declared the relevant scattering angle $\theta_{kk'}$ to be no smaller than the angle representing an *impact parameter* greater than the mean distance in between two impurities. In other words, the maximum impact parameter would be [45]

$$b_{max} = \frac{1}{2} N_I^{1/3} \tag{2.44}$$

where impact parameter and scattering angle are related kinematically by [45]

$$b = \frac{e^2}{8\pi\kappa\varepsilon_0 \left(\frac{\hbar^2 k^2}{2m^*}\right)} \cot(\theta_{kk'}/2) \tag{2.45}$$

Notice that an assumption of isotropy was used in expressing m^*, a scalar carrier effective mass.

As a result of introducing the *maximum* impact parameter and therefore a *minimum* deflection angle for integration, Fermi's golden rule can be performed. This results in [61]

$$\Gamma(\mathbf{k}) = \frac{\pi N_I}{\sqrt{2m^* E_{\mathbf{k}}^3}} \left(\frac{Ze^2}{4\pi\kappa\varepsilon_0}\right)^2 \log\left[1 + \left(\frac{4\pi\kappa\varepsilon_0}{Ze^2}\frac{E_{\mathbf{k}}}{N_I^{1/3}}\right)^2\right] \tag{2.46}$$

We conclude this discussion of ionized impurity scattering by pointing out that there is a better way to describe this scattering rate, freeing up the assumption of isotropy. We will revisit this later.

Neutral impurities

The scattering rates due to neutral impurity scattering are arguably quite negligible for our high-purity detectors. We include neutral impurity scattering for both completeness and to glean some insight. The idea is that, even without a Coulomb potential, there still exists a short-range hydrogenic scattering potential due to core effects. This potential is traditionally modeled as a spherical box potential, with a hydrogenic Rydberg energy serving as the square-well depth. For reasonable energies of incident free electrons or holes, this potential represents a large perturbation. A simple Born approximation is not adequate [5], and the scattering formalism actually reverts to solving the time-dependent Schrödinger equation for incident plane-wave scattering. Thus, Fermi's golden rule is not utilized in this case, but rather the phase-shift method [62] for quantum mechanical scattering.

The earliest notable formulation for neutral impurity scattering in semiconductors is that of Erginsoy [59]. He proposed a scattering rate [5] for the limit that incident carrier wavevector $ka_B < 0.5$ where a_B is the effective hydrogenic Bohr radius. The idea is that the phase-shift method in this low-energy limit results in scattering with a predominant spherical ($l = 0$) outgoing wave. Erginsoy's results for this situation gave a scattering cross-section as

$$\sigma_{Erginsoy} = 20a_B/k \tag{2.47}$$

Considering this cross-section needs a random carrier velocity and target impurity concentration, the corresponding isotropic scattering rate can be given as

$$\begin{aligned} \Gamma_{Erginsoy}(\mathbf{k}) &= N_I v \sigma_{Erginsoy} \\ &= N_I \left(\sqrt{2E_{\mathbf{k}}/m^*}\right)\left(\tfrac{20a_B}{k}\right) \end{aligned} \tag{2.48}$$

A more sophisticated use of the phase-shift method for quantum mechanical scattering was developed by Sclar [60]. Sclar found [5]

$$\sigma_{Sclar} = \frac{4\pi(\hbar^2/2m^*)}{\sqrt{(E_k + E_T)}} \tag{2.49}$$

where E_T is a particularly important *threshold* energy, unique to the square-well formulation. For shallow hydrogenic centers in germanium, the threshold energy is expected to be some small fraction of the Rydberg energy, perhaps $E_T \sim 0.5\ meV$ [5].

This gives a corresponding scattering rate as

$$\begin{aligned} \Gamma_{Sclar}(\mathbf{k}) &= N_I v \sigma_{Sclar} \\ &= N_I \left(\sqrt{2E_{\mathbf{k}}/m^*}\right)\left(\frac{4\pi(\hbar^2/2m^*)}{\sqrt{(E_k+E_T)}}\right) \end{aligned} \tag{2.50}$$

Sclar's formulation gives results that are generally the same order of magnitude as Erginsoy's result, but has a substantially different energy dependence. For our work, we choose to adopt Sclar's result as the slightly more sophisticated treatment.

Sclar's threshold energy, E_T, represents the energy of a singular, singlet *bound* state, which is precisely the same state we will revisit later in this thesis when we discuss the processes involving "anion," "H-," or "overcharged" states. It is interaction with this potential bound state which is the root cause of the neutral impurity scattering, which is now apparent in Sclar's cross-section.

This examination of neutral impurity scattering was useful on two counts. First, by equation 2.50, we appreciate the presence of a *bound* state, E_T, resulting merely from the quantum nature of a square-well potential. This will be important later on in this thesis, regarding capture to bound states. Second, by inspecting Sclar's cross-section of equation 2.49, we note that the threshold energy, E_T, promotes scattering only if it is *small*. This implies that the effect of shallow states are dominant because their threshold energies are small. It is interesting to note that, by this same formulation, *deep* neutral states could exist at extremely high concentrations while still contributing negligible impact on carrier transport properties.

2.3.4 The phonon scattering rates

We now introduce the scattering rates associated with phonon emission and absorption. Fermi's golden rule is applied to all cases. We will point out the possible phonons that may be emitted or absorbed by electrons and holes. We will then cover the origin of the electron-phonon scattering potential, U_s, for both acoustic and optical phonon absorption and emission. Scattering potentials are primarily determined by quantities called *deformation potentials*, which determine the amount of scattering potential created by lattice ion displacement. After describing these necessary ingredients, we finally assemble an integrated rate based on Fermi's golden rule. We will need to consider energy and momentum conservation to correctly manage the required phase-space integrals.

The phonon dispersion relation, and the varieties of possible phonons

There are a variety of phonons that exist due to the lattice structure of crystalline semiconductors. Here, we will briefly outline the phonon dispersion relations pertinent for germanium, and point out the common approximations used for them. We will point out the general classes of phonons that may be absorbed or emitted by electrons and holes.

We will not describe the origin of phonons as harmonic oscillations of the lattice ions. There are numerous, elementary solid-state references for that [63, 61, 4, 45, 5]. It will suffice to say that:

- there are two major classes of phonons; *acoustic* and *optical*,

- both acoustic and optical phonons each have three *branches* based on polarization; one longitudinal branch and two transverse.

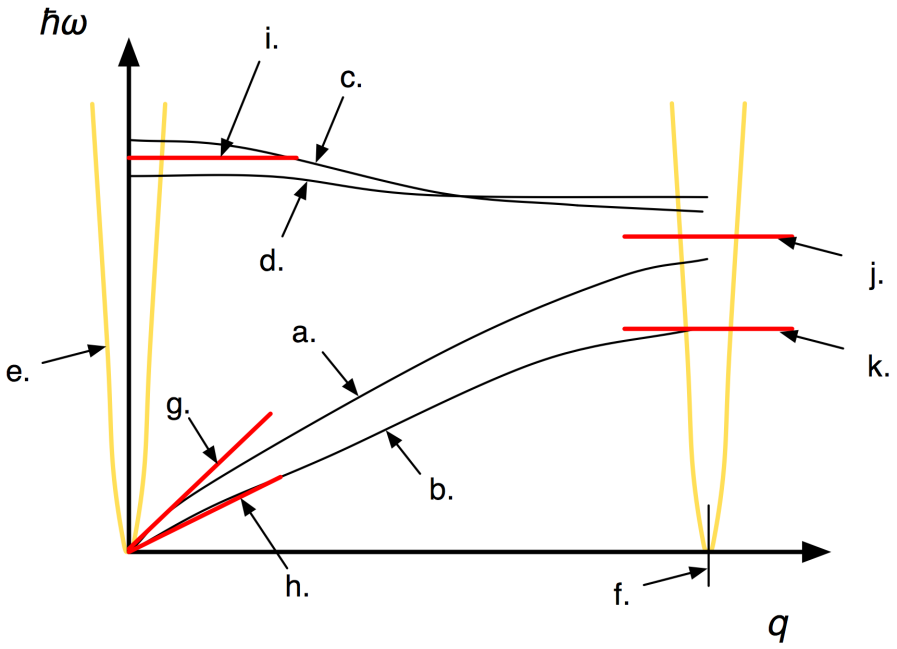

Figure 2.5: The phonon dispersion relation for germanium, idealized for a longitudinal branch and a single transverse branch for both acoustic and optical phonons. **a.** longitudinal acoustic branch (LA) **b.** transverse acoustic branch (TA) **c.** longitudinal optical branch (LO) **d.** transverse optical branch (TO) **3.** a superimposed dispersion relation for parabolic electrons (plotting the corresponding k vector for electrons – not to scale) **f.** the Brillouin Zone boundary **g.** slope for the longitudinal speed of sound **h.** slope for the transverse speed of sound **i.** the optical phonon energy, $\hbar\omega_0$ **j.** the intervalley phonon energy, $\hbar\omega_i$ **k.** the transverse acoustic intervalley phonon energy, $\hbar\omega_{it}$.

Figure 2.5 shows the dispersion relationship for phonons. For clarity, only one transverse branch is depicted for both acoustic and optical phonons. Note that we use variable q for the wavevector for phonons, analogous to k for electrons.

We see that acoustic phonons have zero energy ($\hbar\omega = 0$) at zero wavevector, q. Near the origin, acoustic phonons have a near-linear relationship between ω and phonon wavevector q. The slope is the *speed of sound*, v_s, for acoustic phonons.

$$v_s = \omega/q \tag{2.51}$$

The speed of sound is different for the three different branches of acoustic phonons. The longitudinal branch is fastest. The two transverse branches near the origin are

quite similar, so they are taken to have the same speed of sound. The speeds of sound even have a slight anisotropies, but we do not consider them in this treatment.

For optical phonons, although there are technically three branches, it is standard that all three branches are approximated as having the same energy near $q = 0$. Since the dispersion relation is rather constant near $q = 0$, we consider the optical phonon frequency to be a constant. We call this frequency ω_0.

We can justify the linear relationships for acoustic phonons, and the constant frequency for optical phonons. If we superimpose an electron (or hole) dispersion relation on top of the phonon dispersion relation at $q = k = 0$, we would find that the electrons assume extremely small wavevectors relative to phonons for the same energy. In other words, phonons as seen by electrons or holes are typically very "horizontal" or "carry a lot of momentum per unit energy" on the carrier dispersion curve. This makes sense as phonons represent the displacement of heavy lattice ions. Therefore, while conserving both energy and momentum, carriers typically only interact with small wavevector phonons (near $q \approx 0$). In figure 2.5, a carrier dispersion relation is superimposed at $q = k = 0$, and yet is still scaled larger along the wavevector axis in order to give it some visible width.

The picture we have established so far is further complicated when we consider that phonons can also transfer electrons (holes) in between valleys (bands). So far, we have only considered the *intravalley* (*intraband*) processes. It is also possible to conserve energy and momentum in transferring electrons between different valleys (intervalley scattering), and holes between other bands (interband scattering). This can be accomplished by any branch of phonon, acoustic or optical.

A particular asymmetry between electrons and holes has to do with the many-valley nature of the $L - valley$ conduction band in germanium, as depicted in figure 2.2. There are 8 half-valleys (or equivalently 4 full valleys) for the L points, all at the same conduction-band energy minimum. it is readily possible for electrons to transfer from one valley to another provided they can emit a phonon to adequately conserve momentum and energy. This particular case is interesting because the electron L-valleys (and even the X-valleys) are *indirect* minima, meaning they are offset from $k = 0$. So the emitted (or absorbed) phonons in the intervalley case have very large wavevectors. In fact, the k-space distance subtended between L-valleys represents the same q-momentum of an X-direction ($\langle 100 \rangle$) phonon across the Brillouin Zone. This process is quite similar to Umklapp processes, with the slight difference here being that these intervalley phonons represent large wavevector phonons oriented across the *same* unit cell, rather than an Umklapp phonon which is offset by the components of a single reciprocal lattice vector that would extend beyond the First Brillouin Zone.

Figure 2.5 also shows that the large-wavevector phonons responsible for intervalley processes are located at the Brillouin Zone boundary. The dispersion curves have changed across the Brillouin Zone. Furthermore, it is not clear if intervalley phonons are dominated by acoustic or optical processes at these large wavevectors. Generally, intervalley phonons are treated as optical phonons, with a modified frequency which

we call ω_i. The slowest transverse acoustic branch is substantially smaller than the other branches at the Brillouin Zone boundary, so it is specified separately in some references. We label the transverse acoustic branch for intervalley scattering as ω_{it}.

With this general picture in mind, we will now begin to specify the electron-phonon coupling.

The electron-phonon scattering potential

This section borrows heavily from the explanation of Lundstrom [45].

By analogy, phonons represent "a moving diffraction grating" as seen by electrons and holes. The lattice displacement generated by a phonon creates a stress or strain in the lattice, in turn disrupting the equilibrium bandgap energy by producing a scattering potential. As this lattice displacement is periodic, it may be represented (one-dimensionally) by the following form [45].

$$U_s(z,t) = C_q e^{\pm i(qz-\omega t)} \tag{2.52}$$

In contrast to impurity scattering, this scattering potential is *time-dependent*. It is also oscillatory over position.

A condition mentioned previously was that Fermi's golden rule had to be implemented in cases where scattering was *infrequent*, in order to avoid the complications of collisional broadening. So in the case of a time-dependent scattering potential, it suffices to take a limit as $t \to \infty$ to calculate the matrix element. In taking this limit, it is well known [45] that the factor of $e^{\pm i(\omega t)}$ leads to a modified delta function in Fermi's golden rule, appearing as $\delta(E_{\mathbf{k}'} - E_{\mathbf{k}} \mp \hbar\omega)$.

The oscillatory factor over position $e^{\pm i(qz)}$, leads to a matrix element which is suppressed unless $k' = k \pm q$. This is a statement of momentum conservation, and is implemented by multiplying the matrix element by delta function also in momentum, $\delta(k' - k \mp q)$.

So to summarize the consequences of a potential of the form of equation 2.52, it is important to note that

- the oscillatory time-dependence mandates that the carrier *must* either emit or absorb an energy quantum to conserve energy before and after a scatter
- the oscillatory position-dependence additionally forces momentum conservation.

To note, the one-dimensional nature of equation 2.52 also extrapolates directly to three dimensions.

We now will inspect the constituent ingredients of the prefactor term C_q in equation 2.52.

The phonon displacement and deformation potentials

Consider a phonon as an elastic wave with a quantum field,

$$u_{\nu,q}(\mathbf{r},t) = \widehat{e}_{\nu,q}\sqrt{\frac{\hbar}{2\rho\Omega\omega_q}}\left(a_{\nu,q}e^{i(q\cdot\mathbf{r}-\omega t)} + a^*_{\nu,q}e^{-i(q\cdot\mathbf{r}-\omega t)}\right) \tag{2.53}$$

where $\widehat{e}_{\nu,q}$ is the phonon polarization vector, and $a^*_{\nu,q}$ and $a_{\nu,q}$ are creation and annihilation operators, respectively. The mass density of germanium is ρ, and ω_q is the phonon frequency.

The phonon displacement of equation 2.53 drives the electron-phonon scattering potential. To show this, we need a relation between the displacement provided by a phonon and the change in potential it creates. The energy shift per displacement is given by *deformation potentials*. The perturbed energy due to a displacement in lattice constant can be expressed

$$U_s = \Xi\frac{\delta a}{a} \tag{2.54}$$

where Ξ is a deformation potential, in units eV.

The idea is that a strain $\nabla\cdot\mathbf{u}_{\nu,\mathbf{q}}$ represents the displacement on the lattice exerted by *acoustic* phonons, so that the acoustic phonon scattering potential goes as [45]

$$U_A(\mathbf{r},t) = \Xi_A\ \nabla\cdot\mathbf{u}_{\nu,\mathbf{q}}(\mathbf{r},t) \tag{2.55}$$

where $u_{\nu,q}(\mathbf{r},t)$ is from equation 2.53 and Ξ_A is the deformation potential explicitly for acoustic phonons.

Optical phonons represent oscillations where the motion of neighboring ions is opposite. Therefore, the *optical* phonon scattering potential goes as the displacement vector itself, not the gradient [45]

$$U_0(\mathbf{r},t) = \Xi_0\ u_{\nu,q}(\mathbf{r},t) \tag{2.56}$$

and the deformation potential for optical phonons is in units eV/cm.

In an anisotropic crystal, an acoustic or optical deformation potential is a *tensor*, and relates the energy shift in the lattice to a particular phonon of wavevector q and polarization vector $\widehat{e}_{\nu,q}$. In anisotropic valleys or bands, both shear and dilatational strains can lead to scattering potentials, coupling differently to longitudinal and transverse phonons. So, taking acoustic phonons as an example, a more thorough depiction of a deformation potential is the relation

$$U_s = \Xi_{ij}\frac{\partial u_i}{\partial x_j} \tag{2.57}$$

To relate back to the general form of equation 2.52, we can posit the following relation

$$|C_q|^2 = |\kappa_q|^2 |A_{\nu,q}|^2 \tag{2.58}$$

where the term κ_q represents the energy shift produced per unit lattice displacement, and $A_{\nu,q}$ is the displacement amplitude.

We make use of equation 2.53 to find the displacement amplitude. Note that we are taking scattering expectation values. The annihilation operator contributes to the *absorption* displacement amplitude, while the creation operator refers to the displacement amplitude appropriate for *emission*. Therefore, we should use the displacement amplitude

$$|A_{\nu,q}|^2 = \frac{\hbar}{2\rho\Omega\omega_q}(N_q) \tag{2.59}$$

for absorption processes, and

$$|A_{\nu,q}|^2 = \frac{\hbar}{2\rho\Omega\omega_q}(N_q + 1) \tag{2.60}$$

when considering emission processes. N_q represents the ambient phonon occupation number, following Bose-Einstein statistics.

Meanwhile, the coupling term κ_q is useful because it helps us use a consistent formalism while referring to either acoustic or optical phonons. For acoustic phonons, we should use the relation

$$|\kappa_q|^2 = q^2 \Xi_A^2 \tag{2.61}$$

For optical phonons,

$$|\kappa_q|^2 = \Xi_0^2 \tag{2.62}$$

Combining into a matrix element, the oscillatory dependence of equation 2.52 includes momentum conservation.

$$\boxed{(U_s)_{\mathbf{k},\mathbf{k}'} = |\kappa_q|^2 |A_{\nu,q}|^2 I(\mathbf{k},\mathbf{k}')^2 \delta(\mathbf{k}' - \mathbf{k} \mp q)} \tag{2.63}$$

We then find the rate

$$\boxed{\Gamma(\mathbf{k},\mathbf{k}') = \frac{2\pi}{\hbar}|\kappa_q|^2 |A_{\nu,q}|^2 I(\mathbf{k},\mathbf{k}')^2 \delta(\mathbf{k}' - \mathbf{k} \mp q)\delta(E_{\mathbf{k}'} - E_{\mathbf{k}} \mp \hbar\omega)} \tag{2.64}$$

This is the most basic formulation of the $\mathbf{k}$ to $\mathbf{k}'$ transition rate for phonon scattering. After a detour into some of the specifics of electron and hole deformation potentials, we will investigate the integration of equation 2.64 into a total scattering rate.

The vast majority of scattering rates depicted in our work use simple scalar deformation potentials. We are bound by what is known in the literature from transport

measurements and ab initio calculations. Although there are some constraints from non-transport experiments, deformation constants appropriate for every valley/band transition and for every phonon branch are not always known. Furthermore, many authors only use a scalar form.

Even at room temperature, transport properties are determined almost entirely by electron-phonon properties of the L-valleys, and by hole-phonon properties of the heavy and light bands. So for these most common phonon processes, the formalism has been improved.

In order to have a thorough naming convention, we have adopted a unique system of indices for specific deformation potentials. This is identified in figure 2.6. It is an attempt to give a consistent format to the numerous couplings which arise between carriers and phonons.

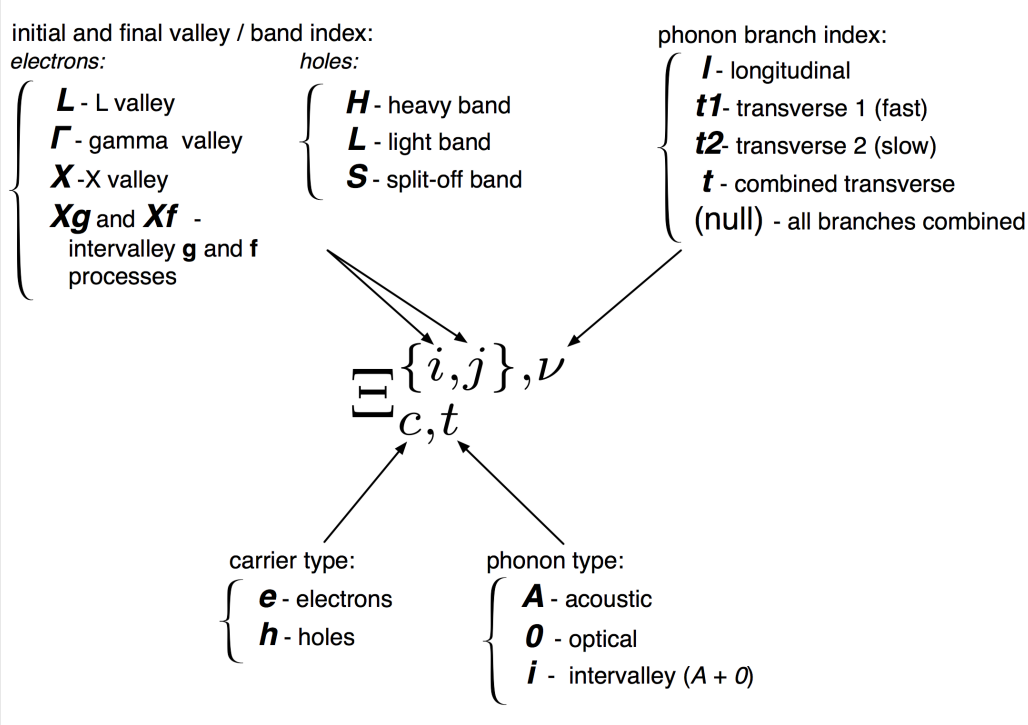

Figure 2.6: There are many labeling systems for deformation potentials throughout the literature. This figure is a pictorial guide to the indices created for this work to delineate specific constants. Intervalley phonons, although treated as optical phonons, are classified as a distinct phonon type because they may contain contributions from both high-wavevector acoustic and optical phonons. Furthermore, for intravalley and intraband processes or for the X-valley g and f processes, only one index is given for both $\{i, j\}$.

We will now visit the more sophisticated approximations used in these cases.

The deformation potentials for electrons

For electrons, scalar constants are typically used for acoustic and optical phonon branches as depicted in figure 2.5. For L-valleys, a common assumption is to consider all three phonon branches as a single, average, isotropic phonon branch, with effective speed of sound [7]

$$\langle v_s \rangle = \frac{2v_{st} + v_{sl}}{3} \tag{2.65}$$

The deformation potential for this branch is just assumed to be a scalar, call it Ξ_A.

Highly anisotropic semiconductors can produce electron-phonon scattering potentials due to both shear and dilational strains. Lacking any symmetry arguments, there could be as many as six independent matrix elements [50] for the required deformation potential tensor. For the case of L-valleys in germanium (and for X-valleys in silicon), fortunately there is a formulation [64] reducing the number of free constants to only two. These constants are the *dilational potential*, Ξ_d, and the *uniaxial shear potential*, Ξ_u. The coupling of these deformation potentials to the longitudinal acoustic and transverse acoustic phonons is a function of the direction of wavevector q, relative to the principle axis (the $\langle 111 \rangle$ orientations) of the ellipsoid valleys. For longitudinal acoustic phonons, the deformation potential as a function of q-vector polar angle is

$$\Xi_{e,A}^{L,l} = \Xi_d + \Xi_u \cos^2(\theta_{\mathrm{qp}}) . \tag{2.66}$$

If the two transverse acoustic branches are treated as one, the combined deformation potential (for a combined TA1 and TA2 rate) is given by

$$\Xi_{e,A}^{L,t} = \Xi_u \sin(\theta_{\mathrm{qp}}) \cos(\theta_{\mathrm{qp}}) . \tag{2.67}$$

For clarity, the polar angle θ_{qp} is labeled "qp" to denote the angle subtended between the phonon q-vector and the "principle axis" of the specific $\langle 111 \rangle$ (or "L-") valley.

We can see in figure 2.7 the angular dependency gives the longitudinal branch substantially stronger coupling than that of the transverse branches.

Although we do not isolate all three acoustic branches coupled to electrons, this approximation is the most sophisticated formulation readily available in the literature [5, 50]. In fact most of the literature, since the introduction [65] of an *average* deformation potential $\Xi_{e,A}^{L}$, representing electron coupling to all three acoustic branches in the L-valleys, uses this single constant rather than individual transverse and longitudinal rates, as in [7].

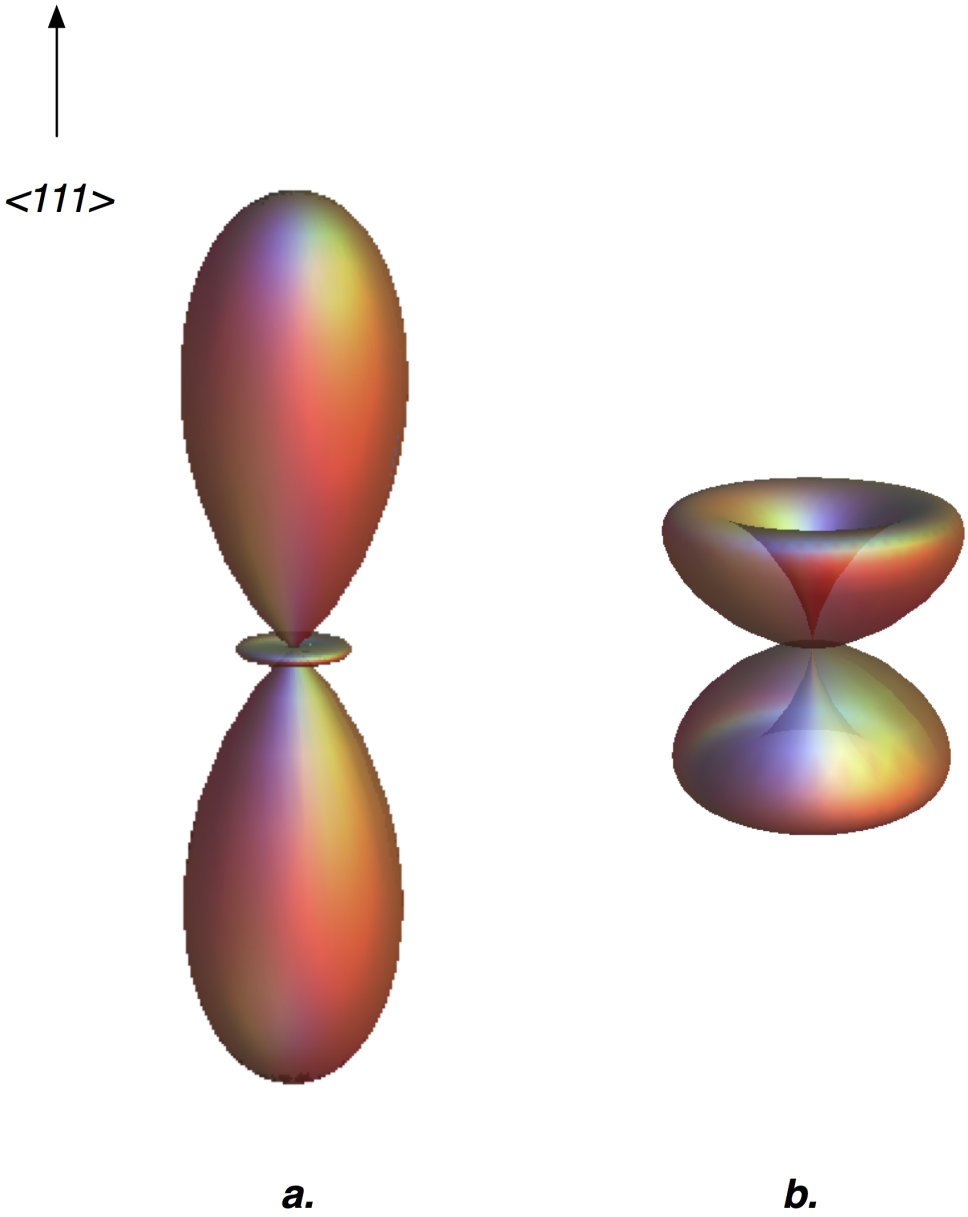

Figure 2.7: The directional dependence of the electron acoustic deformation potentials. Here, polar plots represent the *square* of the deformation potential matrix elements for electron-phonon acoustic scattering in the L-valleys. **a.** the squared longitudinal deformation potential, $(\Xi_{e,A}^{L,l})^2 = (\Xi_d + \Xi_u \cos^2(\theta_{\rm qp}))^2$ **b.** The combined $(TA_1 + TA_2)$ squared transverse deformation potential, $(\Xi_{e,A}^{L,t})^2 = (\Xi_u \sin(\theta_{\rm qp}) \cos(\theta_{\rm qp}))^2$

The deformation potential for holes

The near degeneracy between the heavy and light bands causes them to be substantially warped. This generates considerable difficulty in expressing analytical formulations for the electron-phonon interaction. The most common treatment, introduced by Lawaetz [66], is to group all three phonon branches together [7] with a single, isotropic constant as a deformation potential, $\Xi_{h,A}$, along with a single, averaged speed of sound as per equation 2.65. This simplified single-branch model is used in calculations for both heavy and light bands.

Of course, the interband processes that can occur for holes are only those near $k = 0$. Therefore, phonon wavevectors are typically small unless one were to consider Umklapp processes. So the phonon deformation constants appropriate for holes are those for acoustic and optical branches, all near $q = 0$.

A more rigorous consideration of the acoustic hole-phonon interaction must follow the interacting phonon's wavevector q, polarization vector, $\hat{e}_{\nu,q}$, and speed of sound for that branch. The deformation potential tensor must be depicted by still three independent constants in this case [5, 12]. Following Bir and Pikus [52, 67], the acoustic deformation constants for holes are usually called a, b, and d. The deformation constant a is associated with the acoustic dilational strain in heavy and light bands, while b and d are associated with acoustic shear strain terms [5]. Associated with these deformation potentials are the scattering rates as discussed above, with original implementations by Bir and Pikus [52] in addition to Tiersten [68]. Refinements by Madarasz and Szmulowicz [69, 70], with subsequent adaptions by Hinckley and Singh [71] and a thorough review by Fischetti and Laux [12], have made the representation of hole-phonon scattering rates accessible for use in modern Monte Carlo implementations.

Using the formalism presented by Hinckley and Singh [71], we can define a product of squared deformation potential and overlap factor, as

$$|q\Xi_{h,A}^{\{n,n'\},\nu}|^2 I(k',k)^2 = \left| \sum_{\alpha,\beta=1}^{3} q_\alpha \left(\sum_{j,j'=1}^{3} \langle \psi_{n'}(k')|\ j' \rangle\ D_{\alpha\beta}^{j'j}\ \langle j\ |\psi_n(k)\rangle \right) \hat{e}_{v,\beta}(q) \right|^2 \tag{2.68}$$

where $\hat{e}_{v,\beta}$ is the β-th component of the phonon polarization vector of the ν-th acoustic branch, with a wavevector of $\mathbf{q}$. The element q_α is the α-th component of $\mathbf{q}$. The matrix element $D_{\alpha\beta}^{j'j}$ is the deformation potential component of one of the following $D_{\alpha\beta}$ matrices[71] in x, y, and z coordinates. Symmetries [71] impose $D_{xy} = D_{yx}$, $D_{yz} = D_{zy}$, and $D_{zx} = D_{xz}$. These expressions in terms of a, b, and d are

$$
\begin{aligned}
D_{xx} &= \begin{bmatrix} a+2b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & a-b \end{bmatrix} \quad D_{yy} = \begin{bmatrix} a-b & 0 & 0 \\ 0 & a+2b & 0 \\ 0 & 0 & a-b \end{bmatrix} \\
D_{zz} &= \begin{bmatrix} a-b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & a+2b \end{bmatrix} \quad D_{zx} = \begin{bmatrix} 0 & 0 & \frac{\sqrt{3}}{2}d \\ 0 & 0 & 0 \\ \frac{\sqrt{3}}{2}d & 0 & 0 \end{bmatrix} \\
D_{xy} &= \begin{bmatrix} 0 & \frac{\sqrt{3}}{2}d & 0 \\ \frac{\sqrt{3}}{2}d & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad D_{yz} = \begin{bmatrix} 0 & 0 & \frac{\sqrt{3}}{2}d \\ 0 & 0 & 0 \\ \frac{\sqrt{3}}{2}d & 0 & 0 \end{bmatrix}
\end{aligned} \tag{2.69}
$$

Notice in equation 2.68 that we also require phonon polarization vectors, appropriate for each branch. From first principles, the three acoustic phonon branches may be depicted as a system of linear equations representing the lattice displacement equations of motion. The phonon *polarization vectors* for each branch are the system eigenvectors. The rigorous phonon polarization vectors do not have a manageable closed form [72]. As is common to the relevant literature [69, 70, 71], we choose to adopt the approximated form of phonon polarization vectors developed by Ehrenreich and Overhauser [73].

$$
\begin{aligned}
\hat{e}_L &= \frac{1}{|q|} \begin{pmatrix} q_x \\ q_y \\ q_z \end{pmatrix} \\
\hat{e}_{T1} &= \frac{1}{\sqrt{q_x^2+q_y^2}} \begin{pmatrix} q_y \\ q_x \\ 0 \end{pmatrix} \\
\hat{e}_{T2} &= \frac{1}{|q|\sqrt{q_x^2+q_y^2}} \begin{pmatrix} q_x q_z \\ q_y q_z \\ -\left(q_x^2+q_y^2\right) \end{pmatrix}
\end{aligned} \tag{2.70}
$$

With this information, we can implement equation 2.68 when we wish to use the full form of the acoustic deformation potential for holes. However, we see that the full matrix element formalism is quite complicated. To consider any possible scattering event in this case, we need full knowledge of the phonon wavevector and its branch, its polarization, as well as the final and initial hole **k**-vectors. This is somewhat intense for for standard transport treatments, although we do utilize it in our full, anisotropic Monte Carlo algorithm to be explained later in this work.

Considering now the *optical* phonon deformation potential tensor for holes, we again use the formalism of Hinckley and Singh [71]. In this case, we have the following coupling.

$$
|\Xi_{h,0}^{\{n,n'\},\nu}|^2 I(k',k)^2 = \left(\frac{3d_0^2}{2a_0^2}\right) \left| \sum_{j,j'=1}^{3} \langle \psi_{n'}(k')| \, j' \rangle \, A_\nu^{j'j} \, \langle j \, |\psi_n(k)\rangle \right|^2 \tag{2.71}
$$

Here, d_0 is the optical deformation potential for holes from the formalism of Bir and Pikus [52], a_0 is the germanium lattice constant, and matrices A_ν represent the permutation operator corresponding to the ν-th optical phonon branch. The operators A_ν are determined by lattice symmetry and are provided once again in reference [71].

$$A_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad A_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad A_z = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.72}$$

In equation 2.71, note the prefactor containing d_0, the valence-band optical deformation potential. This quantity is adapted to units of energy per unit displacement by a quadrature factor, $\frac{3}{2}\frac{d_0^2}{a_0^2}$. In spite of the apparent complexity of equation 2.71, when optical strains are averaged over all directions the sum of the coupling to all three optical modes is found to be isotropic [5, 53]. Therefore, we can use a single deformation potential appropriate to depict the *net* optical phonon coupling to holes, for a particular band.

$$|\Xi_{h,0}^{\{n,n'\}}|^2 \;=\; \sum_{\nu=1}^{3} \left(|\Xi_{h,0}^{\{n,n'\},\nu}|^2 I(k',k)^2 \right) \;=\; \tfrac{3}{2}\left(\tfrac{d_0}{a_0}\right)^2 \qquad \text{(isotropic)} \tag{2.73}$$

Summarizing this discussion of the deformation potentials specific to holes, we see that the formalism is sometimes complicated. For the optical phonons, the lattice symmetries simplify the problem if we consider all the optical phonon branches combined. However, transport properties at a small electric field are determined largely by the acoustic phonon coupling, and its depiction remains difficult. To determine the possible interaction between a hole and an acoustic phonon, the phonon's wavevector and polarization must be completely known. We must also simultaneously know the initial and prospective final wavevector components of the hole as well. This highlights the fact that complete energy and momentum conservation in our simulations is key if we wish to retain the sophisticated details of carrier-phonon coupling.

2.3.5 Towards a *total* rate, $\Gamma(\mathbf{k})$, for carrier-phonon scattering processes

We have been discussing the nature of the scattering potential between carriers and phonons. We established a scattering rate $\Gamma(\mathbf{k}, \mathbf{k}')$ by Fermi's golden rule in equation 2.64. We then discussed particular forms of the deformation potentials for electrons and holes. Now, we wish to *integrate* the state-to-state rate $\Gamma(\mathbf{k}, \mathbf{k}')$ across all possible final states to reach the physically meaningful *total* rate, $\Gamma(\mathbf{k})$. To accomplish this integration, what we will find is that we must understand energy and momentum conservation for various cases.

We could integrate the rate, $\Gamma(\mathbf{k}, \mathbf{k}')$, of equation 2.64 over $\mathbf{k}'$-space as in equation 2.40. However, if we take note of the simple functional form for deformation potential coupling, we notice that the phonon wavevector q readily appears. To proceed, it is straightforward if we perform an integration over the *phonon* ($\mathbf{q}$-space) density of states, rather than carrier states $\mathbf{k}'$. This change of representation is permissible because the correspondence between phonon and carrier states is one-to-one and unique [45]. So we wish to integrate the following.

$$\boxed{\Gamma(\mathbf{k}) = \frac{2\pi}{\hbar}\frac{\Omega}{(2\pi)^3}\int d^3q|\kappa_q|^2|A_{\nu,q}|^2 I(\mathbf{k},\mathbf{k}')^2\delta(\mathbf{k}'-\mathbf{k}\mp q)\delta(E_{\mathbf{k}'}-E_{\mathbf{k}}\mp\hbar\omega)} \quad (2.74)$$

To note, the delta function representing energy conservation in the above expression does not assume parabolicity.

What we observe in equation 2.74 is that performing the integration relies on satisfying the delta functions in momentum and in energy. In other words, the limits of integration are determined when the range of $\mathbf{q}$ conserving energy and momentum is determined. We must proceed by considering energy and momentum conservation for specific processes. Specifically, we need to consider energy and momentum conservation for acoustic phonon absorption and emission, as well as for optical phonon absorption and emission. These processes are represented in figure 2.8.

Unlike the standard treatments [5, 7, 45], we take care throughout these cases to consider energy and momentum conservation where the initial- and final-state carrier mass values (m_i and m_f) are not necessarily equal.

2.3.6 Conservation of energy and momentum for acoustic phonon emission

Our first condition is to consider a carrier that emits an acoustic phonon. We will investigate energy and momentum conservation.

To be completely general, we could first consider the non-parabolic case for energy conservation. Including nonparabolicity, to first order, the conservation of energy implies

$$\frac{\hbar^2k^2}{2m} = \gamma(E_k) \approx E_k(1+\alpha E_k) \quad (2.75)$$

Equating a truncated Taylor expansion for acoustic phonon emission, we have

$$\gamma(k') \approx \gamma(k) - \frac{d\gamma}{dE_k}\times(\hbar v_s q) \quad (2.76)$$

which amounts to the following.

$$\frac{\hbar^2k'^2}{2m_f} = \frac{\hbar^2k^2}{2m_i} - \hbar v_s q\left(1+2\alpha\frac{\hbar^2k^2}{2m_i}\right) \quad (2.77)$$

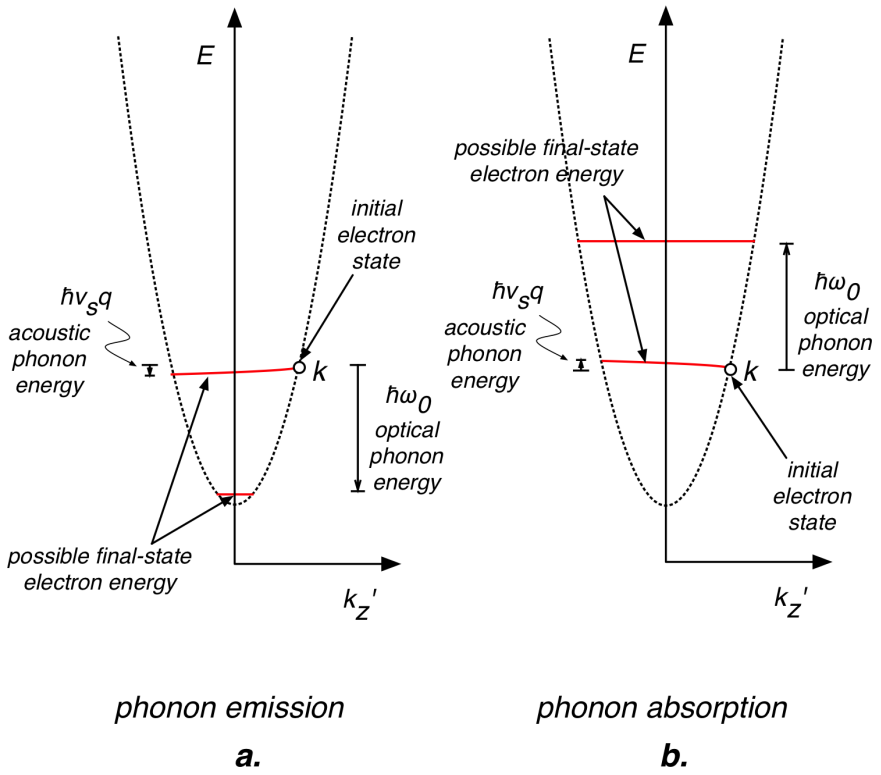

Figure 2.8: Dispersion relations depicting energy conservation for processes involving acoustic and optical phonons. **a.** Emission of an acoustic phonon with energy $\hbar v_s q$, or emission of an optical phonon with energy $\hbar\omega_0$ **b.** Absorption of an acoustic phonon with energy $\hbar v_s q$, or absorption of an optical phonon with energy $\hbar\omega_0$

We can solve this for k'.

$$k' = \sqrt{\frac{m_f}{m_i}}\left(k^2(1-2\alpha\hbar q v_s) - 2\frac{m_i q v_s}{\hbar}\right)^{1/2} \tag{2.78}$$

Next, to consider momentum conservation, it helps to refer to the wavevector composition of figure 2.9. As we are considering phonon emission, we find

$$k'^2 = k^2 + q^2 - 2\ k\ q\ cos\theta \tag{2.79}$$

where θ is the angle subtended between the initial k-vector and phonon q-vector.

Our immediate goal is to solve for $cos\theta$. Combining equations 2.79 and 2.78, we find

$$\boxed{\cos\theta = -\frac{k}{2q} + \frac{km_f}{2m_i q} + \frac{m_f v_s}{\hbar k} + \alpha\frac{m_f v_s}{m_i}\hbar k - \frac{q}{2k}} \tag{2.80}$$

Although our equation 2.80 is more cumbersome than one might typically find in an introductory reference, we still follow the standard line of reasoning for this problem

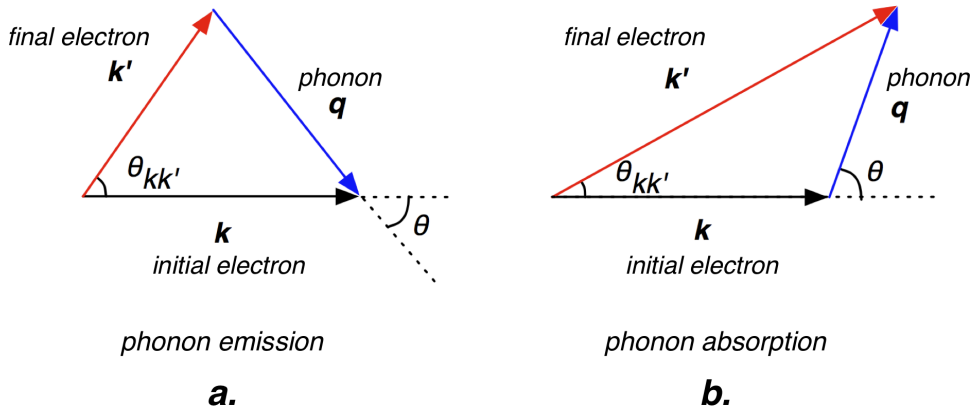

Figure 2.9: Wavevector diagrams depicting momentum conservation between a phonon of wavevector **q** and a carrier with initial wavevector **k** and final state **k′** [5]. **a.** *Emission* of a phonon **b.** *Absorption* of a phonon

as presented in [5, 45]. We solve for the limiting cases of q, where the right-hand side of equation 2.80 meets the physical extent of $cos\theta$ at limiting conditions

$$\begin{cases} \cos\theta &= 1 \\ &= -1 \end{cases} \tag{2.81}$$

The problem is constrained. We can solve for limiting cases of q if the variables are known. For clarity, we now assume parabolic bands by setting $\alpha \to 0$. We will now consider different scenarios relating the final and initial mass values.

For $m^* = m_f = m_i$:

Consider the carrier to be in an isotropic valley/band, so that the effective mass is the same in initial and final states. For the case $cos\theta = 1$, the momenta are oriented such that the emitted phonon is in the direction of the initial state and momentum transfer is optimum. We recover the standard result that, if $v = \frac{\hbar k}{m} \geq v_s$, then $q = q_{max} = 2k\left(1 - \frac{v_s}{v}\right)$. This is a *maximum* q value. If $v = \frac{\hbar k}{m} < v_s$, then q_{max} is negative and unphysical, and phonon emission cannot take place. Physical values of q never allow $cos\theta$ less than zero, so phonon emission is always forward. Figure 2.10 illustrates the allowed values of wavevector q.

For $m_f < m_i$:

We find parabolic roots. Phonon emission is still always in the forward direction, but the kinematics has also now introduced a *minimum* q value.

$$q_{\min} = k - \frac{m_f v_s}{\hbar} - \frac{\sqrt{\hbar^2 k^2 m_f m_i - 2\hbar k v_s m_f m_i^2 + v_s^2 m_f^2 m_i^2}}{\hbar m_i} \tag{2.82}$$

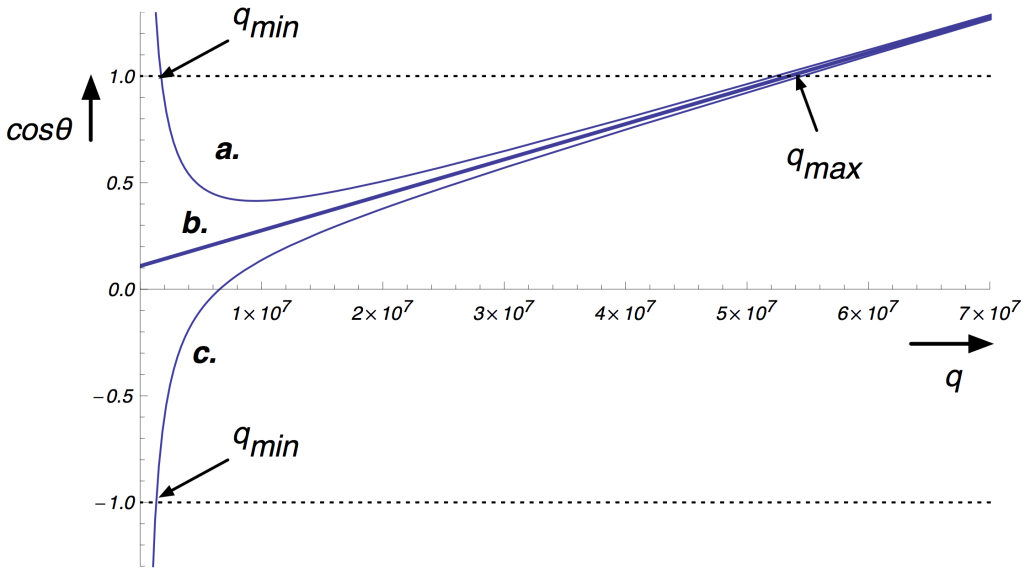

Figure 2.10: Plotting $cos\theta$ as a function of phonon wavevector q, using some typical values for germanium. **a.** For $m_f < m_i$, e.g., $m_f = 0.9m_i$. We see the appearance of a lower root, q_{min}, **b.** For $m_f = m_i$, **c.** For $m_f > m_i$, e.g., $m_f = 1.1m_i$. Here, we also find a lower allowed limit for q, but here q_{min} represents *backward* phonon emission.

Taking a series expansion in m_f about m_i, we find this to be a presumably small term of approximately

$$q_{\min} \approx \frac{\hbar k^2 \left(m_i - m_f\right)}{2m_f \left(\hbar k_i - m_f v_s\right)} \tag{2.83}$$

For $m_f > m_i$:

We again find parabolic roots, but this time the minimum q value corresponds to a phonon emitted *backwards* at $cos\theta = -1$. This is a unique situation. The backward emitted phonon results in an *increased* carrier wavevector amplitude.

We find the minimum allowed q value to be

$$q_{\min} = -k - \frac{m_f v_s}{\hbar} + \sqrt{\frac{m_f}{m_i}k^2 + 2\frac{k m_f v_s}{\hbar} + \frac{m_f^2 v_s^2}{\hbar^2}} \tag{2.84}$$

which, from another series expansion in m_f about m_i, is given by the approximate relation

$$q_{\min} \approx \frac{\hbar k^2 (m_f - m_i)}{2m_i(\hbar k - m_i v_s)} \tag{2.85}$$

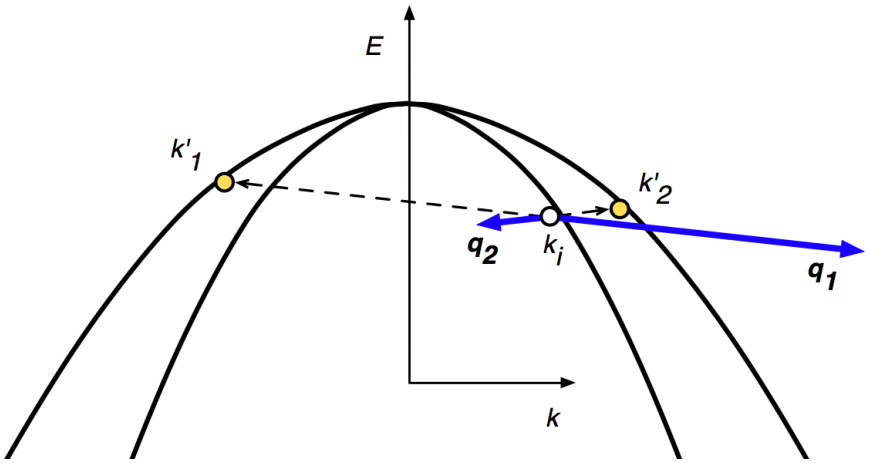

Figure 2.11: A dispersion relation for holes, assuming isotropy for the light (bottom) and heavy (top) bands. If the hole is initially in the light band, $m_f > m_i$. In this situation, the hole can *always* emit a phonon and relax to the heavy band. In addition to the typical forward phonon emission process represented by q_1, it is also possible for the hole to emit a phonon *backwards*, represented by q_2. The backward emission of a phonon results in a final hole state with larger k-vector.

Also note that, since the value of $cos\theta$ always ranges from -1 to 1, **there must always be an allowed range of possible final states when** $m_f > m_i$**.** So a light band carrier may always emit to a heavy band, provided all emission orientations are allowed. This is representative of the case of a hole occupying the light band. In figure 2.11, it is evident that a hole can emit a phonon both forward and backward.

Revisiting the delta functions:

What we have found in equation 2.80 is an expression that combines both energy and momentum conservation. We can use this to reconsider the energy and momentum conservation imposed by the delta functions in equation 2.74.

Using the identity

$$\delta(ax) = \delta(x)/|a| \tag{2.86}$$

and keeping the first non-parabolic term, we can write the product of the momentum and energy delta functions as a combined delta function with a prefactor.

$$\begin{aligned} &\delta_{k',k-\hbar v_s q}\delta\left(\gamma(k') - \gamma(k) + \tfrac{d\gamma}{dE_k}(\hbar v_s q)\right) = \\ &\tfrac{m_f}{\hbar^2 kq}\delta\left(-\cos\theta + \tfrac{k}{2q} - \tfrac{km_f}{2m_i q} + \tfrac{q}{2k} + \tfrac{m_f v_s}{\hbar k} + \alpha\tfrac{m_f}{m_i}\hbar k v_s\right) \end{aligned} \tag{2.87}$$

The prefactor $\frac{m_f}{\hbar^2 kq}$ appears in front of this delta function due to the identity of equation 2.86. This prefactor enters our expression for the integrated scattering

rate, and serves to preserve the correct factor for the density of states. While most treatments [5, 45] assume an isotropic mass ($m^* = m_f = m_i$), here we have shown that it is the **final mass state** m_f that appears in this expression. So all else equal, final states $\mathbf{k}'$ with heavier masses m_f contribute more to scattering than lighter final states. We will revisit this later.

Integration of the total rate:

Here, we assert parabolic assumptions. Rates incorporating nonparabolicity (though with isotropic assumptions) can be found in the literature [53].

We wish to integrate

$$\Gamma(\mathbf{k}) = \frac{2\pi}{\hbar}\frac{\Omega}{(2\pi)^3}\int d^3q \left(|\kappa_q|^2\right)\left(|A_{\nu,q}|^2\right) I(\mathbf{k},\mathbf{k}')^2 \delta(\mathbf{k}'-\mathbf{k}+\mathbf{q})\delta(E_{\mathbf{k}'}-E_{\mathbf{k}}+\hbar\omega) \quad (2.88)$$

with the substitutions that

$$\begin{array}{rcl} |\kappa_q|^2 & = & \left(q\Xi_A\right)^2 \\ |A_{\nu,q}|^2 & = & \frac{\hbar}{2\rho\Omega v_s q}\left(N_q+1\right) \end{array} \quad (2.89)$$

$$\delta(\mathbf{k}'-\mathbf{k}+\mathbf{q})\delta(E_{\mathbf{k}'}-E_{\mathbf{k}}+\hbar\omega) = \left(\frac{m_f}{\hbar^2 kq}\right)\delta\left(-\cos\theta+\frac{k}{2q}-\frac{km_f}{2m_i q}+\frac{q}{2k}+\frac{m_f v_s}{\hbar k}\right)$$

This last identity uses a joint expression for momentum and energy conservation, with an implicit assumption of integration over one of the delta functions.

We also consider the d^3q integral in spherical q-coordinates, with principle axis aligned along the k-vector.

Upon integration, the delta function acts as a unity multiplier for values of q that conserve energy and momentum. Effectively, this changes the limits of integration over amplitude q.

For now, if we consider $I(\mathbf{k},\mathbf{k}')^2 \to 1$, then we have symmetry over the polar and azimuthal angles. Our total rate simplifies to

$$\begin{array}{rcl} \Gamma(\mathbf{k}) & = & \frac{2\pi}{\hbar}\frac{\Omega}{(2\pi)^3}\left(\Xi_A^2\right)\left(\frac{\hbar}{2\rho\Omega v_s}\right)\left(\frac{m_f}{\hbar^2 k}\right)\int\limits_{q_{\min}}^{q_{\max}} q^2\left(N_q+1\right)dq \\ & = & \frac{m_f\Xi_A^2}{8\pi^2\rho v_s\hbar^2 k}\int\limits_{q_{\min}}^{q_{\max}} q^2\left(N_q+1\right)dq \end{array} \quad (2.90)$$

The Bose-Einstein distribution for phonon occupation can be approximated in high and low temperature limits. At room temperature, or whenever $k_BT \gg \hbar v_s q$, the lattice is considered to be in the *equipartition limit*. This is because $N_q+1 \approx N_q$, so emission processes approximately equal absorption processes. In the *zero-point limit*, $k_BT \ll \hbar v_s q$, phonon modes are barely occupied. Here, emission processes dominate,

but only by the spontaneous emission probability to emit phonons due to zero-point fluctuations of the lattice ions.

$$\begin{array}{rcccll} N_q & \approx & N_q + 1 & \approx & \frac{k_B T}{\hbar v_s q} & (\text{equipartition}: \ k_B T \gg \hbar v_s q) \\ N_q & \approx & 0 & & & (\text{zero}-\text{point}: \ k_B T \ll \hbar v_s q) \end{array} \tag{2.91}$$

The zero-point limit is generally appropriate for a CDMS base temperature of $T = 40\ mK$. Therefore, equation 2.90 simplifies. If we further assume $m^* = m_f = m_i$, we recover the simple zero-point limit for acoustic phonon emission [5].

$$\boxed{\begin{array}{rcll} \Gamma(\mathbf{k}) & = & \frac{m^* \Xi_A^2}{8\pi^2 \rho v_s \hbar^2 k} \frac{\left(q_{\max}^3 - q_{\min}^3\right)}{3} & (\text{zero}-\text{point}: \ k_B T \ll \hbar v_s q) \\ & = & \frac{m^* \Xi_A^2}{3\pi^2 \rho v_s \hbar^5 k} \left(\hbar k - m v_s\right)^3 & \end{array}} \tag{2.92}$$

2.3.7 Conservation of energy and momentum for phonon processes: acoustic phonon absorption

Although rare at a temperature of $T = 40\ mK$, the next condition we consider is for a carrier to absorb an ambient acoustic phonon. We will first investigate energy and momentum conservation, as per the previous section.

Conserving energy for acoustic phonon absorption, we find similar relations as for the case of acoustic emission. From

$$\frac{\hbar^2 {k_f}^2}{2m_f} = \frac{\hbar^2 {k_i}^2}{2m_i} + \hbar q v_s \left(1 + 2\alpha \frac{\hbar^2 {k_i}^2}{2m_i}\right) \tag{2.93}$$

we find

$$k' = \sqrt{\frac{m_f}{m_i}} \left(k^2 (1 + 2\alpha \hbar q v_s) + 2 \frac{m_i q v_s}{\hbar}\right)^{1/2} \tag{2.94}$$

Next, momentum conservation as depicted by figure 2.9 is

$$k'^2 = k^2 + q^2 + 2\ k\ q\ cos\theta \tag{2.95}$$

From equations 2.94 and 2.95, we find

$$\boxed{\cos\theta = \frac{k}{2q} - \frac{k m_f}{2 m_i q} + \frac{m_f v_s}{\hbar k} + \alpha \frac{m_f v_s}{m_i} \hbar k + \frac{q}{2k}} \tag{2.96}$$

Now, we plot equation 2.96 for the three scenarios of m_f versus m_i. As before, when considering the following subsections, we assume parabolicity so that $\alpha \to 0$.

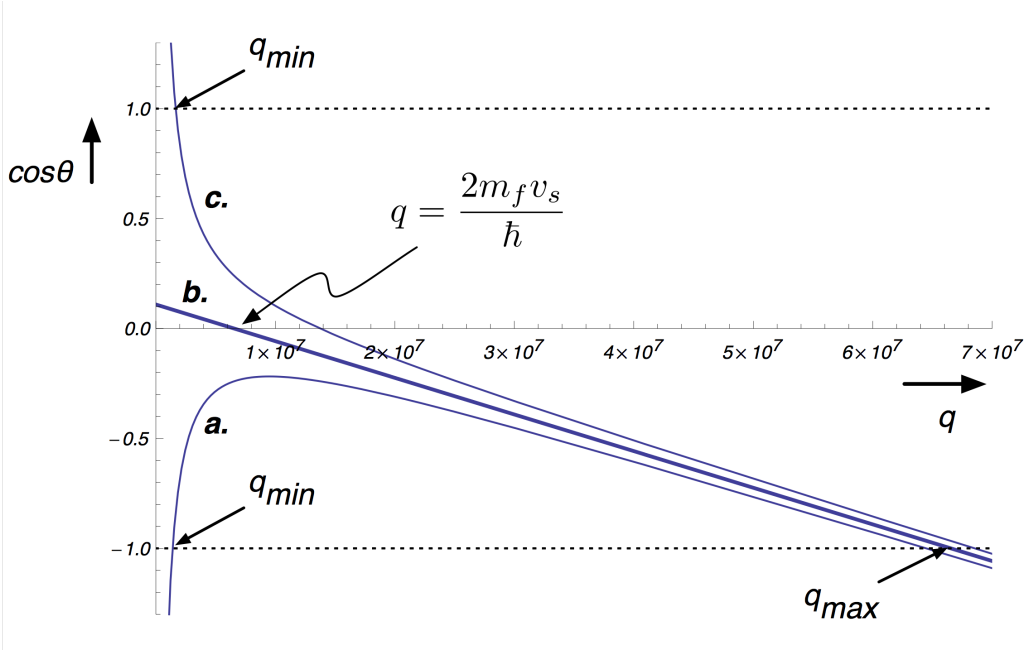

Figure 2.12: Plotting the transcendental function for $cos\theta$ for the case of acoustic phonon absorption. **a.** For $m_f < m_i$, which in particular happens to be $m_f = 0.9m_i$. We see the appearance of a lower root, q_{min}, that now takes on its value at $cos\theta = -1$, **b.** For $m_f = m_i$, **c.** For $m_f > m_i$, which in particular happens to be $m_f = 1.1m_i$. We also find a lower allowed limit for q in this case.

For $m^* = m_f = m_i$:

Assume the final and initial states to have the same value for effective mass. What is different in this case versus acoustic phonon emission is that momenta combine rather than subtract, so that there is a q maximum, $q = q_{max} = 2k\left(1 + \frac{v_s}{v}\right)$, now at $cos\theta = -1$. The minimum q is now typically at 0, although the situation does arise that, if the carrier is traveling slower than the speed of sound, the minimum q becomes non-zero. So we recover the result, as also found in [5],

$$q_{\min} = \begin{cases} 0 & v \geq v_s \\ 2k(\frac{v_s}{v} - 1) & v < v_s \end{cases} \tag{2.97}$$

For $m_f < m_i$:

If the final-state mass is less than the initial state mass for the case of acoustic phonon absorption, then we see the appearance of a minimum q arise at $cos\theta = -1$

as in figure 2.12. This lower q-value is

$$q_{\min} = k + \frac{m_f v_s}{\hbar} - \sqrt{\frac{m_f}{m_i}k^2 + 2\frac{k m_f v_s}{\hbar} + \frac{m_f^2 v_s^2}{\hbar^2}} \tag{2.98}$$

and in a series expansion of m_f about m_i, reduces to

$$q_{\min} \approx \frac{\hbar k^2 (m_i - m_f)}{2m_i(\hbar k + m_i v_s)} \tag{2.99}$$

It should be noted that when $\frac{\hbar k}{m_i} = v \ll v_s$, there occur a total of *four* roots for q; two at $cos\theta = 1$ and two at $cos\theta = -1$. Therefore, there are two regions where phonon absorption can occur, and these regions are bounded by these two sets of parabolic roots.

For $m_f > m_i$:

We also consider this $m_f > m_i$ case, where the minimum q value occurs at $cos\theta = +1$. This gives a lower q-value at

$$q_{\min} = -k + \frac{m_f v_s}{\hbar} + \sqrt{\frac{m_f}{m_i}k^2 - 2\frac{k m_f v_s}{\hbar} + \frac{m_f^2 v_s^2}{\hbar^2}} \tag{2.100}$$

which, approximated in our typical way, is

$$q_{\min} \approx \frac{\hbar k^2 (m_f - m_i)}{2m_i(\hbar k - m_i v_s)} \tag{2.101}$$

which we see holds as long as $\hbar k/m_i > v_s$. When $\hbar k/m_i < v_s$, the range of allowed q values becomes extremely narrow, centered at the value of $q = \frac{2m_f v_s}{\hbar}$.

The delta function and integrated rate for acoustic phonon absorption

Analogous to equation 2.87, we find the following for a combined delta function representing both energy and momentum conservation for acoustic phonon absorption.

$$\begin{gathered}\delta_{k',k+\hbar v_s q}\delta\left(E_{\mathbf{k}'} - E_{\mathbf{k}} - \hbar v_s q\right) = \\ \frac{m_f}{\hbar^2 k q}\delta\left(\cos\theta + \frac{k}{2q} - \frac{k m_f}{2m_i q} + \frac{q}{2k} - \frac{m_f v_s}{\hbar k}\right)\end{gathered} \tag{2.102}$$

Using the same assumptions as for equation 2.90, we can state the following total rate for an acoustic absorption process. Recall that the phonon occupation term appropriate for absorption is N_q rather than $N_q + 1$.

$$\begin{aligned}\Gamma(\mathbf{k}) &= \tfrac{2\pi}{\hbar}\tfrac{\Omega}{(2\pi)^3}\left(\Xi_A^2\right)\left(\tfrac{\hbar}{2\rho\Omega v_s}\right)\left(\tfrac{m_f}{\hbar^2 k}\right)\int_{q_{\min}}^{q_{\max}} q^2\left(N_q\right)dq \\ &= \tfrac{m_f \Xi_A^2}{8\pi^2\rho v_s \hbar^2 k}\int_{q_{\min}}^{q_{\max}} q^2\left(N_q\right)dq\end{aligned} \tag{2.103}$$

Of course we can solve this in the various limits we have depicted above. In the "zero-point limit" which is described in equation 2.91 and is the limit appropriate for our operating conditions, the phonon occupation is assumed to be *zero*. There are no ambient phonons to absorb, and the absorption rate is also *zero*.

In our numerical treatment to be described, we do include absorption terms. This allows us to simulate transport conditions at higher temperatures, for comparison to other experimental findings.

2.3.8 Conservation of energy and momentum for phonon processes: optical and intervalley phonons

We must also consider energy and momentum conservation for optical and intervalley processes. We treat these processes considering only parabolic bands. A prescription for treating these processes with a nonparabolicity factor does exist [5], but we opt not to use it. For most our range of interest, optical phonons do not dominate transport processes. Furthermore, even near the optical phonon threshold, the change in the electron dispersion relation due to nonparabolicity is quite small. We do not consider non-parabolic bands for holes, in any case.

To note, the L-valleys in germanium are located at the edge of the Brillouin Zone. Consequently, intervalley phonons are identical to Umklapp processes, in that a phonon wavevector is emitted which spans the Brillouin Zone (an inverse lattice vector). Here, we conserve energy and momentum for optical processes and intervalley processes equivalently, neglecting the inverse lattice wavevector of the Brillouin Zone. In practice, the inverse lattice wavevector is added to the q wavevectors for the intervalley processes solved for here.

Where $\pm$ terms represent *absorption* and *emission* processes, Energy conservation is given by

$$\frac{\hbar^2 k'^2}{2m_f} = \frac{\hbar^2 k^2}{2m_i} \pm \hbar\omega_0 \tag{2.104}$$

where $\hbar\omega_0$ represents the energy of the optical or intervalley phonon.

Momentum conservation, as always, is given by

$$k'^2 = k^2 + q^2 \pm 2kq\cos\theta \tag{2.105}$$

Even though we treat this case entirely in terms of parabolic bands, our derivation still differs from standard treatments such as [5, 45] because we have chosen to maintain the distinction between m_f and m_i in equation 2.104.

By combining equations 2.104 and 2.105, we find the relationship

$$\boxed{\cos\theta = \pm\frac{k}{2q}\frac{m_f - m_i}{m_i} \mp \frac{q}{2k} + \frac{m_f\omega_0}{\hbar kq}} \tag{2.106}$$

We go on to solve for the extrema of allowed q wavevectors as we have established in the previous cases for acoustic phonon processes. These roots are somewhat cumbersome. It is worthwhile to mention that solving equation 2.106 for $\cos\theta = 1$ and $\cos\theta = -1$ results once again in *parabolic* roots.

For the case of emission processes, most conditions promote forward emission of optical phonons near threshold. When $m^* = m_f = m_i$, we recover the standard case [5] that

$$\begin{aligned} q_{\min} &= k\left(1 - \sqrt{1 - \tfrac{\hbar\omega_0}{E_k}}\right) \qquad (\cos\theta = 1) \\ q_{\max} &= k\left(1 + \sqrt{1 - \tfrac{\hbar\omega_0}{E_k}}\right) \qquad (\cos\theta = 1) \end{aligned} \tag{2.107}$$

For the case of the absorption processes, minimum and maximum q-values are generally non-trivial at forward ($\cos\theta = 1$) and backward ($\cos\theta = -1$) emission angles. When $m^* = m_f = m_i$, we find the standard case [5] that

$$\begin{aligned} q_{\min} &= k\left(\sqrt{1 + \tfrac{\hbar\omega_0}{E_k}} - 1\right) \qquad (\cos\theta = 1) \\ q_{\max} &= k\left(\sqrt{1 + \tfrac{\hbar\omega_0}{E_k}} + 1\right) \qquad (\cos\theta = -1) \end{aligned} \tag{2.108}$$

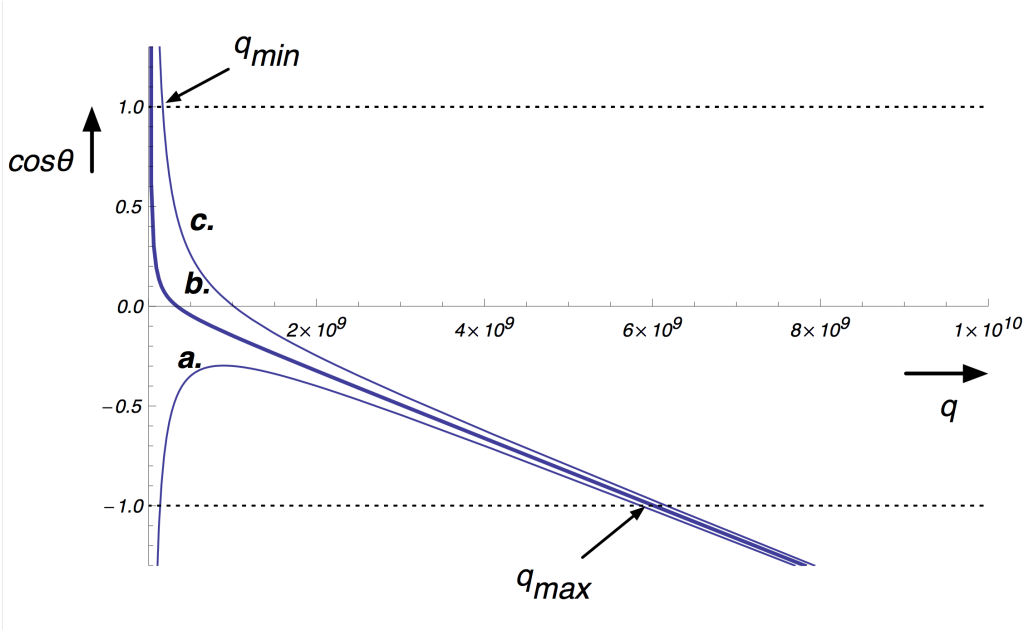

Figure 2.13: Plotting the transcendental function for $cos\theta$ for the case of optical and intervalley phonon absorption. To dramatize the possible range of $\cos\theta$, kinematic values were chosen that E_k exceeds $\hbar\omega_0$ by a factor of 90. **a.** For $m_f < m_i$, which in particular happens to be $m_f = 0.9m_i$, **b.** For $m_f = m_i$, **c.** For $m_f > m_i$, which in particular happens to be $m_f = 1.1m_i$.

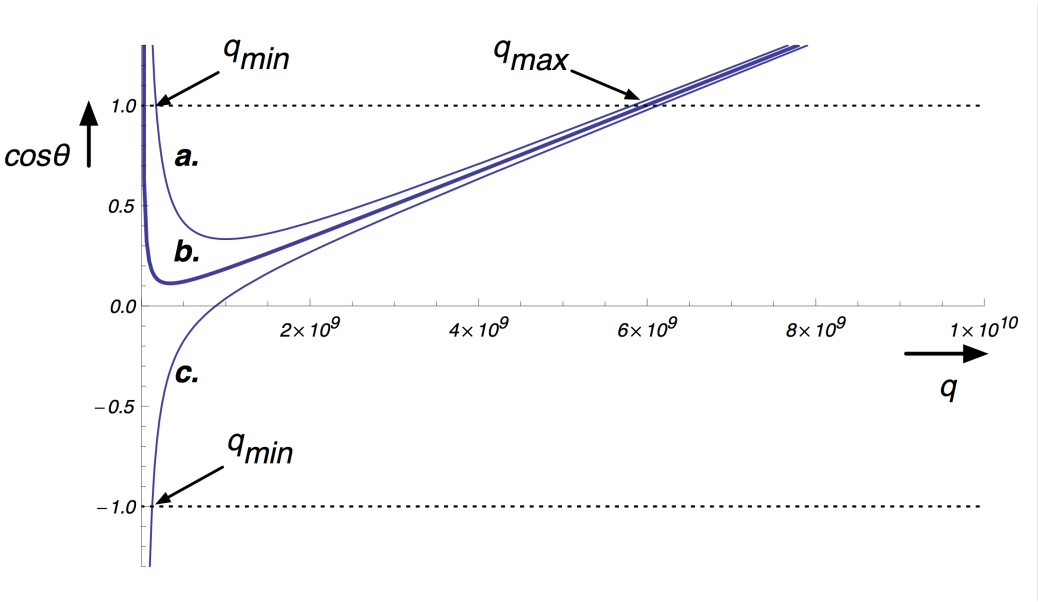

Figure 2.14: Plotting the transcendental function for $cos\theta$ for the case of optical and intervalley phonon emission. To dramatize the possible range of $\cos\theta$, kinematic values were chosen that E_k exceeds $\hbar\omega_0$ by a factor of 90. **a.** For $m_f < m_i$, which in particular happens to be $m_f = 0.9m_i$, **b.** For $m_f = m_i$, **c.** For $m_f > m_i$, which in particular happens to be $m_f = 1.1m_i$.

For both emission and absorption processes, the consideration that $m_f \neq m_i$ allows some freedom for the range of q-values and of phonon angles. Figure 2.14 demonstrates the transcendental representation of equation 2.106 for optical emission processes, and figure 2.13 represents the equivalent expression for optical absorption processes.

We now move on to establish the delta function and total rate for optical and intervalley phonon processes.

The delta function and total rate for optical and intervalley processes

From equation 2.106, we arrive at the following delta function which combines both energy and momentum conservation.

$$\boxed{\begin{array}{c}\delta_{k',k\pm\hbar\omega_0}\delta\left(E_{k'} - E_k \mp \hbar\omega_0\right) = \\ \frac{m_f}{\hbar^2 kq}\delta\left(\pm\cos\theta + \frac{k}{2q} - \frac{km_f}{2m_i q} + \frac{q}{2k} \mp \frac{m_f\omega_0}{\hbar kq}\right)\end{array}} \tag{2.109}$$

For a full solution of the optical phonon total rate,

$$\begin{aligned}\Gamma(\mathbf{k}) \quad &= \quad \tfrac{2\pi}{\hbar}\tfrac{\Omega}{(2\pi)^3}\int d^3q\left\{(|\Xi_0|^2)\left[\tfrac{\hbar}{2\rho\Omega\omega_0}\left(N_0+\tfrac{1}{2}\mp\tfrac{1}{2}\right)\right]\right. \\ &\left.\times\, I(\mathbf{k},\mathbf{k}')^2\left(\tfrac{m_f}{\hbar^2kq}\right)\delta\left(\pm\cos\theta+\tfrac{k}{2q}-\tfrac{km_f}{2m_iq}+\tfrac{q}{2k}\mp\tfrac{m_f\omega_0}{\hbar kq}\right)\right\}\end{aligned} \tag{2.110}$$

we require the identity for the included delta function. After switching to spherical coordinates in **q**-space, the delta function selects the range of q-values which conserve energy and momentum. Note that, since the optical phonon dispersion relation is independent of q, the phonon occupation term N_0 is also independent of q.

$$\begin{aligned}\Gamma(\mathbf{k}) \quad &= \quad \tfrac{2\pi}{\hbar}\tfrac{\Omega}{(2\pi)^3}2\pi\int_{q_{\min}}^{q_{\max}} qdq\left\{(|\Xi_0|^2)\left[\tfrac{\hbar}{2\rho\Omega\omega_0}\left(N_0+\tfrac{1}{2}\mp\tfrac{1}{2}\right)\right]\left(\tfrac{m_f}{\hbar^2k}\right)I(\mathbf{k},\mathbf{k}')^2\right\} \\ &= \quad \tfrac{2\pi}{\hbar}\tfrac{\Omega}{(2\pi)^3}2\pi\left(|\Xi_0|^2\right)\tfrac{\hbar}{2\rho\Omega\omega_0}\left(N_0+\tfrac{1}{2}\mp\tfrac{1}{2}\right)\left(\tfrac{m_f}{\hbar^2k}\right)\int_{q_{\min}}^{q_{\max}} qdq\left(I(\mathbf{k},\mathbf{k}')^2\right) \\ &= \quad \tfrac{m_f\Xi_0^2}{4\pi\rho\omega_0\hbar^2k}\left(N_0+\tfrac{1}{2}\mp\tfrac{1}{2}\right)\int_{q_{\min}}^{q_{\max}} qdq\left(I(\mathbf{k},\mathbf{k}')^2\right)\end{aligned} \tag{2.111}$$

So assuming $I(\mathbf{k},\mathbf{k}')^2\to 1$, assuming $m^*=m_f=m_i$, and using the appropriate limits of either equation 2.107 or 2.108, we recover the accepted[5, 45], isotropic optical rate for optical or intervalley phonon absorption or emission.

$$\begin{aligned}\Gamma(\mathbf{k}) \quad &= \quad \tfrac{\pi\Xi_0^2}{\rho\omega_0}\left(N_0+1\right)\left(\tfrac{2m^*(E_k-\hbar\omega_0)^{1/2}}{4\pi^2\hbar^3}\right) \qquad \text{(emission)} \\ &= \quad \tfrac{\pi\Xi_0^2}{\rho\omega_0}\left(N_0\right)\left(\tfrac{2m^*(E_k+\hbar\omega_0)^{1/2}}{4\pi^2\hbar^3}\right) \qquad \text{(absorption)}\end{aligned} \tag{2.112}$$

The last terms in parentheses can be identified with an energy *density of states* [5], as

$$g(E_k)=\left(\frac{2m^*\left(E_k\right)^{1/2}}{4\pi^2\hbar^3}\right) \tag{2.113}$$

So we can rewrite equations 2.112 as

$$\boxed{\begin{aligned}\Gamma(\mathbf{k}) \quad &= \quad \tfrac{\pi\Xi_0^2}{\rho\omega_0}\left(N_0+1\right)g(E_k-\hbar\omega_0) \qquad \text{(emission)} \\ &= \quad \tfrac{\pi\Xi_0^2}{\rho\omega_0}\left(N_0\right)g(E_k+\hbar\omega_0) \qquad \text{(absorption)}\end{aligned}} \tag{2.114}$$

So we have recovered the traditional rates for optical and intervalley phonon absorption and emission. However, the main exercise of this subsection was to present the fact that the limits of integration are sometimes complicated when $m_f\neq m_i$. To account for the full case correctly is most easily done in the numerical work, where an algorithm can select the proper physical root for q_{min} and q_{max}, depending on which term is physical for the kinematics at hand.

In concluding this subsection, note equation 2.113. It is important to point out that the m^* associated with this density of states is the *density of states mass*, so $m^* = m_d$. This $m^* = m_d$ identity holds true for the acoustic rates, as well. Recall that this mass first appeared due to the identity of the delta function. As we took care to keep track of the particular masses used in our derivation as in equation 2.87, we were able to identify this mass as m_f. So something else we have learned is that the information regarding any particular mass of a given final state is isotropically integrated away in the integrated rates. By identifying the many possible m_f states in some average fashion as m_d, we make the integrals tractable. This is an argument used later on as to why we care to preserve directional information from scatter to scatter.

2.4 Concluding arguments regarding scattering mechanisms

In this chapter, we have described germanium band structure, the effective mass approximation, and Fermi's golden rule for scattering rates. We went into detail regarding the variety scattering mechanisms, and the particular approximations employed. When considering phonon scattering rates, an integration of Fermi's golden rule over all possible final states required a thorough examination of energy and momentum conservation.

We summarize this chapter by displaying two simple figures. Figure 2.15 represents several total scattering rates, evaluated in isotropic approximation, for electrons as a function of carrier energy at a temperature of $T = 40\,mK$. Figure 2.16 is a similar plot, for representative scattering processes appropriate for holes.

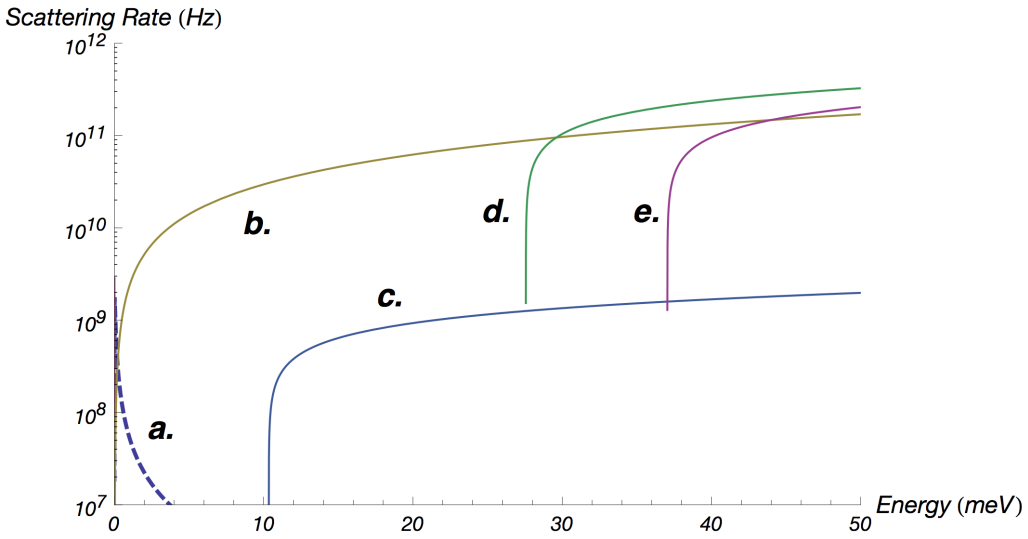

Figure 2.15: Total scattering rates used for ELECTRONS at $T = 40\ mK$, calculated under isotropic approximations. **a.** Conwell-Weisskopf ionized impurity scattering rate at $N_I = 10^{10}\ cm^{-3}$ **b.** acoustic phonon emission **c.** slow-transverse intervalley phonon emission **d.** intervalley phonon emission **e.** optical phonon emission

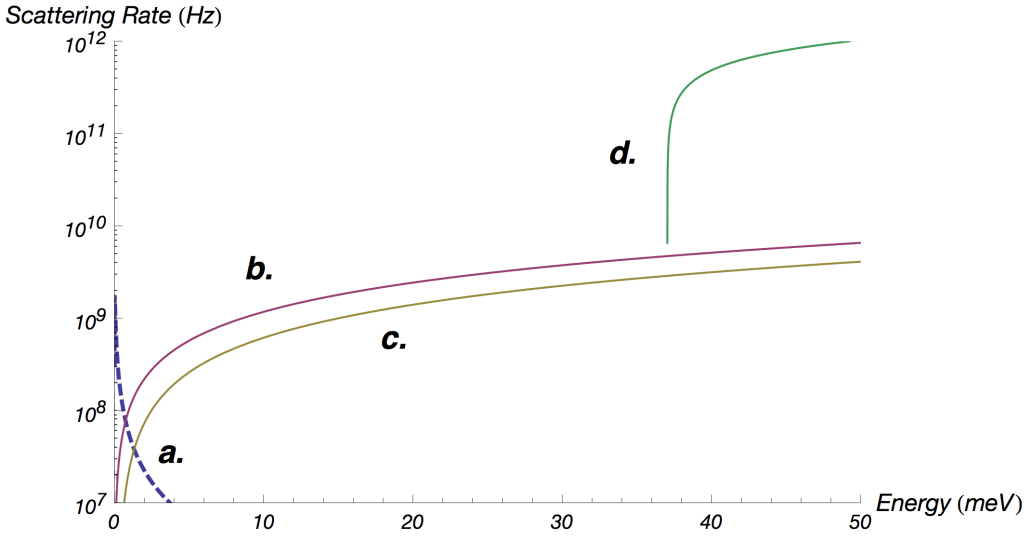

Figure 2.16: Some select total scattering rates used for HOLES at $T = 40\ mK$, calculated under isotropic approximations. **a.** Conwell-Weisskopf ionized impurity scattering rate at $N_I = 10^{10}\ cm^{-3}$ **b.** acoustic phonon emission for heavy-to-heavy band transitions **c.** acoustic phonon emission for heavy-to-light band transitions **d.** optical phonon emission for heavy-to-heavy band transitions

Now, we consider whether or not to include the scattering rate for ionized impurities. We do not know, *a priori*, the percentage of impurities which are ionized under our low-temperature conditions. Also recall that impurity scattering is considered to be a completely elastic process, serving only to randomize rather than to "cool" hot carriers. Also note that this randomizing effect is largely suppressed because the ionized-impurity scattering rate promotes predominantly small-angle deflections, owing to its strong angular dependence ($\sim \sin^{-4}(\theta_{kk'}/2)$). Furthermore, for CDMS operating conditions where fields are of a few V/cm, simulations yet to be described show that electrons reach an average of a few meV while holes reach an average of $\sim 10\ meV$. FIgures 2.15 and 2.16 show that the Conwell-Weisskopf scattering rate for $10^{10}\ cm^{-3}$ ionized impurities is inconsequential for such energies. So for this multitude of reasons, we choose to neglect scattering rates due to impurities for the work that follows. We will find that the effect of impurities is extremely important when considering *recombination* and *ionization* processes, but generally *not* for *transport* processes under modest applied fields.

In the next chapter, we describe how the knowledge of the scattering rates we have established in this chapter is incorporated into a Monte Carlo algorithm to predict transport properties.

Chapter 3

Monte Carlo Simulation: the Traditional Technique and Anisotropic Approaches

3.1 Introduction

Now that we have established carrier scattering processes in the previous chapter, it is possible to assemble these rates in a transport simulation. To bypass the analytical assumptions and difficulties in solving the Boltzmann Transport Equation [45], we implement the *Monte Carlo method*. Variations of an established Monte Carlo method for carrier transport has been developed and utilized for decades [7, 74, 75, 76]. It is a straightforward way to model the transport effects of non-equilibrium carriers in semiconductor devices. In this work, we refer to the relatively standard Monte Carlo implementation as the *traditional* Monte Carlo method. Later in this chapter, we will show how we have adapted and built upon the general Monte Carlo framework to more precisely preserve momentum information and crystal anisotropy within the scattering rates. This new Monte Carlo framework requires both a restructured computational algorithm, as well as a revisitation of Fermi's golden rule.

Our implementations of the Monte Carlo method allow us to examine how carriers are "hot," relative to the crystal lattice taking into account the anisotropy of the crystal to a much greater extent than standard treatments while rigorously imposing energy and momentum conservation at each scatter, contrary to the existing treatments. The typical carrier energies reached in steady state are far greater than the thermal energy ($\frac{3}{2}k_BT_L$) expected if they were in *thermal equilibrium* with the crystal lattice. The Monte Carlo method is a stochastic simulation technique to reproduce these physics at the microscopic scattering level, so that we may then examine overall distribution characteristics.

To outline this chapter, we will cover the following topics.

Traditional Monte Carlo: We introduce and summarize the techniques traditionally used in the Monte Carlo method applied to carrier transport in semiconductors.

Time steps: We explain how time steps are chosen for free-flight acceleration in the traditional Monte Carlo. We also introduce Rees' "self-scattering" term.

Selection of a scattering process: We explain how a scattering process is chosen to occur in simulation, using rates based on energy.

Selection of a final state: Once a physical scattering event has occurred, a separate routine must choose the specific final state after the scatter.

Vector-based Monte Carlo: We explain how we adopt a momentum(vector)-based representation of Fermi's golden rule, and implement a Monte Carlo based on sampling momenta. Scattering selection and final-state selection are now one and the same.

Electrons and holes: We explain how we implement our vector-based Monte Carlo, with some specifics regarding electrons and holes.

3.1.1 Regarding constants used

Regarding the constants used in this work, by default we use those of the Monte Carlo review by Jacoboni and Reggiani [7]. For other constants not mentioned there, we use Fischetti's work [76]. For anisotropic deformation potentials for both electrons and holes, we use values from Fischetti and Laux [12]. We also use the anisotropic constants for the valence band (A, B, C) from Ridley [5]. We account for $T \sim 0$ conditions by adjusting the germanium minimum band gap to 0.76 eV and the higher, direct gap to 0.89 eV where needed.

3.2 The traditional, isotropic implementation of the Monte Carlo algorithm

First, we introduce the traditional Monte Carlo algorithm. Under the simplest assumptions, namely those of an isotropic model, an accepted application of the Monte Carlo method to semiconductor transport has been established over many years, across a range of literature [7, 53, 74, 75, 76, 45, 48, 77].

In a semiconductor, the random scattering of carriers redistributes their energy and momenta over time. Meanwhile, external forces continuously accelerate the carriers. Transport properties at the macroscopic scale are dominated by the overall stochastic nature of carrier scattering processes occurring at the individual, microscopic level. Therefore, simulating transport processes can be accomplished by an

algorithm which models the random motion of carriers subject both to an applied homogeneous field as well as the known, stochastic scattering processes. The Monte Carlo algorithm accomplishes this by the use of both semiclassical kinematics in conjunction with quantum-mechanically determined, random scattering events.

Our purpose is to design a Monte Carlo routine to replicate the carrier transport dynamics of a $T_L = 40\ mK$ germanium lattice, in a homogeneous applied electric field. Our implementation of the traditional Monte Carlo algorithm is executed by simulating individual carrier trajectories over position and momentum as they travel in the crystal medium. The algorithm begins with a single electron (or hole) starting at given initial conditions for position and momentum. A resulting trajectory is generated by tracking a carrier governed by semiclassical equations of motion, including the acceleration produced by an externally-applied electric field. The carrier propagates only over brief periods of time, however. The length of these time steps, or "free flights," is randomly determined by the combined scattering probability from the total of many scattering processes described by Fermi's golden rule. So the periods of semiclassical free-flights are terminated by instantaneous scattering events. A separate subroutine within the Monte Carlo randomly determines which particular type of scattering will occur following a free flight. A further subroutine then utilizes random number generation to determine the *final momentum state*, $\mathbf{k}'$, of the carrier immediately deflected after a scatter. In this way, the carrier is subject to continuous periods of semiclassical acceleration balanced by randomizing scatters. The resulting random walk can be tracked for some sufficiently long period of time. Moreover, several identical carriers can be regenerated from the same initial conditions and tracked again to develop another sufficiently long dynamical history. The final states of these carriers are recorded, resulting in distributions of parameters such as energy and momenta. These distributions are then a good approximation to the expected distributions of actual carriers in a steady state with the given electric field. From these final ensembles, statistics can be performed to determine transport properties.

In figure 3.1, we provide a flow chart which outlines the Monte Carlo method as traditionally applied to semiconductor transport. This reproduces the random walk particle trajectories in the crystal which provide data about the final ensemble of carriers. We now describe the methodology needed to complete each of the steps depicted in this algorithm.

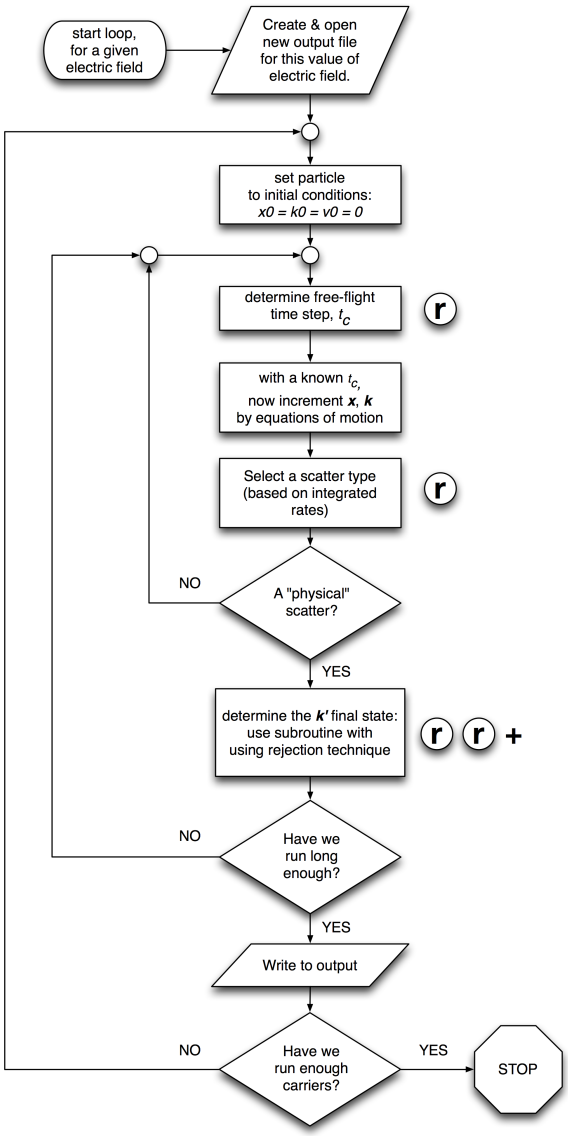

Figure 3.1: A flow chart representing the traditional Monte Carlo algorithm. After initializing a carrier's position and momentum, a continuous loop of small time steps chooses a time step based on a random number, increments momentum according to acceleration, decides whether to scatter or to free flight, and decides the outcome of any possible scatter. Every scatter event requires at least 4 random numbers, as marked. The system is ergodic, but many such particles may be run independently.

3.2.1 Establishing initial conditions for carrier propagation

We typically choose to simulate carriers originating from a "cold start." That is to say, we set initial conditions as zero for momenta and position coordinates. While there may be more efficient initial conditions, these conditions are at least consistent with all possible external conditions to be simulated.

We wish to find carrier ensembles that are well equilibrated to the applied electric field. It is important to ensure that we do not mistakenly include transient behavior as a relic of the initial conditions. Therefore, we run each carrier in the electric field long enough such that transient behavior has subsided. We choose a criterion such that on average each carrier runs a total time, T, that is several *energy relaxation lengths*. For this purpose, the energy relaxation length is calculated from the final carrier ensemble. We define it for this simulation as

$$\frac{1}{\tau_E} = \frac{1}{\tau_a}\left(\frac{\langle \hbar v_s q \rangle}{\langle E \rangle}\right) + \frac{1}{\tau_0}\left(\frac{\hbar\omega_0}{\langle E \rangle}\right) \tag{3.1}$$

where $\langle E \rangle$ is the average carrier energy from the final ensemble, $\langle \hbar v_s q \rangle$ is the average energy emitted to an acoustic phonon, τ_a is the acoustic phonon emission scattering time, $\hbar\omega_0$ is the optical phonon energy (or intervalley phonon energy when considering electrons) and τ_0 is the optical (intervalley for electrons) phonon emission scattering time. Under our zero-point scattering conditions, scattering processes are dominated by these two emission rates. Therefore it is appropriate that the combined rate for these two major processes, as in equation 3.1, is an indicative measure of the typical energy relaxation time.

3.2.2 Generation of carrier free-flight: Rees scattering

While tracking the dynamics of an electron or hole within the Monte Carlo simulation, much of the time the carrier is simply accelerating in the external field. Carriers accelerate in the field by a change in momentum

$$\hbar\dot{\mathbf{k}} = \mathbf{F} = \mp e\mathbf{E} \tag{3.2}$$

for an electron or hole, respectively. If some initial wavevector is $\mathbf{k_0}$, then after accelerating for a small time step t_c, the adjusted wavevector $\mathbf{k}$ is taken simply as

$$\mathbf{k} = \mathbf{k_0} \mp \frac{e\mathbf{E}}{\hbar} t_c \tag{3.3}$$

the resultant velocity and change in position are also incremented for this small time step t_c, as per the band structure dynamics explained in the previous chapter. The question is: what time increment do we take for t_c?

Determination of the incremental time step

We know that scattering events are stochastic in nature. If we knew the rate of their occurrence, we should be able to adjust time steps large enough to represent simply the time between randomly chosen scatters. In other words, if the combined rate of all scattering mechanisms were Γ_0, the the probability $P(t)$ to scatter in an incremental time dt is

$$P(t)dt = \left[\Gamma_0 \exp\left(-\int_0^t \Gamma_0 \; dt'\right)\right] dt \tag{3.4}$$

So we should be able to sample this distribution with a uniformly generated random number, r, thereby finding a random time step t_c. Using the following,

$$r = \int_0^{t_c} P(t)dt = \int_0^{t_c} \left[\Gamma_0 \exp\left(-\int_0^t \Gamma_0 dt'\right)\right] dt = 1 - e^{-\Gamma_0 t_c} \tag{3.5}$$

we must invert the expression to find t_c.

$$t_c = -\frac{\ln[1-r]}{\Gamma_0} \tag{3.6}$$

Since r is a uniform random number spanning $(0, 1)$, choosing a particular random number r is equivalent to choosing one from the distribution $1 - r$. Therefore we can slightly simplify the expression for t_c, so it becomes

$$t_c = -\frac{\ln[r]}{\Gamma_0}. \tag{3.7}$$

This determination of the random time step as in expression 3.7 is a commonly derived quantity. The above argument mirrors similar arguments posed in [45].

Inclusion of the *time-dependent* scattering rate, by Rees' self-scattering

We claimed that Γ_0 is the scattering rate appropriate for a carrier undergoing stochastic processes. However, we did not consider that this rate changes *dynamically* with the energy of the carrier. As the carrier energy increases or decreases, the number of possible final states with which to scatter also changes.

So as the carrier energy changes with time, so does the scattering rate. Rather than the constant Γ_0, we should call the combination of all scattering rates a *time-dependent* quantity, $\Gamma(t)$. We must revisit equation 3.4.

$$P(t)dt = \left[\Gamma(t) \exp\left(-\int_0^t \Gamma(t')dt'\right)\right] dt \qquad (\text{ using } \Gamma(t) \;) \tag{3.8}$$

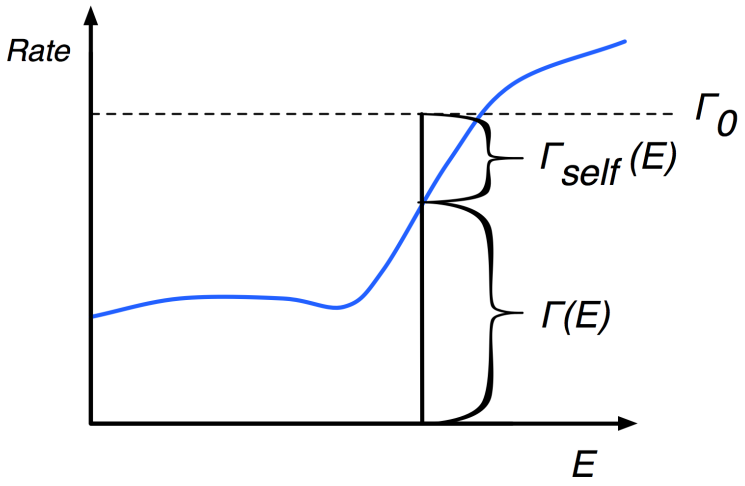

Figure 3.2: The total sampling rate Γ_0 remains constant, but the proportion between the net physical scattering rate $\Gamma(E)$ versus self-scatters $\Gamma_{\text{self}}(E)$ may change. For physically meaningful results, $\Gamma(E)$ must always remain less than Γ_0.

Unless we knew the precise dynamics and could recover a functional form for $\Gamma(t)$, we might otherwise be stuck at this point.

A way out, which is commonly utilized but rarely attributed to Rees [78] who first proposed it [75], is to intentionally redefine the sampling rate as a constant. While the time-dependent rate $\Gamma(t)$ is the cumulative rate of all physical scattering processes,

$$\Gamma(t) = \sum_{i=1}^{N} \Gamma_i(t) \tag{3.9}$$

we introduce an extra, fictitious rate to recover a composite, but *constant*, sampling rate Γ_0.

$$\Gamma_0 = \Gamma(t) + \Gamma_{\text{self}}(t) \tag{3.10}$$

Here, $\Gamma_{\text{self}}(t)$ is commonly referred to as *self-scattering*. Self-scattering represents no physical process. It is simply a placeholder. By equation 3.11, we simply choose an adequately large Γ_0 and define $\Gamma_{\text{self}}(t)$ by the difference with $\Gamma(t)$. Although the introduction of the self-scattering term seems to only add extra, seemingly useless samples to the Monte Carlo, it is a small price to pay to recover a constant Γ_0. This allows us to regain the simple relation of equation 3.7, allowing us to directly generate efficient time intervals, t_c.

Physical limits on Γ_0, and the Iterative-Γ method

To represent physically meaningful results, the cumulative rate $\Gamma(t)$ must always be less than the chosen Γ_0. As we can see in the representative figure 3.2, a sharp energy threshold in scattering rate makes this difficult. Such a threshold, which is due to the onset of optical (or intervalley) phonon emission, often determines where Γ_0 must be set. Even if carriers spend most of the simulation time in lower-energy regions where the scattering rate may be magnitudes smaller, even brief carrier excursions into a region where $\Gamma_0 < \Gamma(t)$ may produce unphysical results. On the other hand, if Γ_0 is chosen to be too large, then computational power is lost on unnecessary sampling.

One may ask, "is there a way to change Γ_0 during the simulation, adaptively?" A priori, this is worrisome because we know that the sampling rate should remain constant with time, otherwise equation 3.7 would be void. However, a method depicted in the textbook of Moglestue [75], suggests a methodology to adaptively alter Γ_0 in a permissible way. The consequence is that the probability for a scattering event to be correctly sampled depends on the integral, $\exp\left(-\int_0^t \Gamma_0 dt'\right)$. It is permissible to alter Γ_0 before entering the bounds of this integral. In other words, $t' = 0$ occurs immediately at the instant of the last physical scatter. It is possible to make fine-tuned adjustments for a more efficient Γ_0 by making adaptive changes immediately after a physical scattering event has occurred, but before subtending further time steps. *In simplest terms,* we are allowed to change Γ_0 at each physical scatter, but not during the acceleration phase. This is called the *Iterative-*Γ *Method*. We do not use this method within the framework of the traditional Monte Carlo, but we do take advantage of it in our later methods.

3.2.3 Selection of scattering process

We have determined *when* scattering events should happen. Yet we also must choose which *type* of process occurs. Here, we show how a scattering type is selected within the Monte Carlo, following a free-flight segment.

Since the sampling rate is the sum of the self-scattering term plus all N number of physical processes,

$$\Gamma_0 = \sum_{i=1}^{N} \Gamma_i(E) + \Gamma_{self}(E) \tag{3.11}$$

we can use a random number and subroutine to select among these weighted probabilities. For a uniform random number on $(0, 1)$, we select the j-th process by comparing successive sums such that

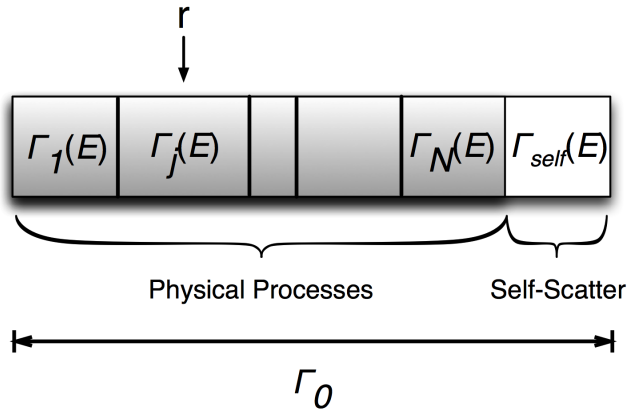

Figure 3.3: The selection of the scattering type based on the relative rates. A random number, r, selects the j-th process.

$$\frac{\sum_{i=1}^{j-1}\Gamma_i(t)}{\Gamma_0} \leq r < \frac{\sum_{i=1}^{j}\Gamma_i(t)}{\Gamma_0} \tag{3.12}$$

This is depicted in figure 3.3.

If

$$r \geq \sum_{i=1}^{N}\Gamma_i(E)/\Gamma_0 \tag{3.13}$$

then the scattering event at hand is a *self-scatter*, so no action is taken regarding the final state of the carrier.

If the chosen process is a physical process, further random number generation must then occur to select a proper final state.

3.2.4 Selection of final state

Having chosen a physical process for a scattering event, another subroutine must now stochastically determine the final momentum $\mathbf{k}'$-state. Generally, this is the most elaborate subroutine of the simulation. The new momentum state must be generated according to the differential rate of the specific process.

The random selection methods: To randomly select some dependent parameter a from the distribution determined by the total scattering rate $\Gamma(\mathbf{k})$, we need to express the probability density in terms of a differential rate.

$$P(a)\,da = \frac{1}{\Gamma(\mathbf{k})}\frac{\partial\Gamma(\mathbf{k})}{\partial a}\,da \tag{3.14}$$

Now, over the (uniform) space of possible random numbers r, a possible outcome might be value r_1. We need to map this particular random number r_1 to a representative choice of parameter a, which we call a_1.

$$r_1 = \int_0^{r_1} dr = \int_{a_{\min}}^{a_1} P(a)\,da \tag{3.15}$$

So, in best cases, we can find a tractable form for a_1 by inverting equation 3.15. In these cases, a value for a_1 can be chosen directly from a randomly generated value r_1 by some function. This method is therefore called the *direct method*.

$$a_1 = f(r_1) \qquad \text{(direct method)} \tag{3.16}$$

Unfortunately, it is often the case that a needed form of equation 3.15 cannot be inverted for a direction solution of a_1. This happens when the implementation of equation 3.15 is not integrable. In this case, we must use the probability density of equation 3.14 and produce joint random trials of parameter a and the differential probability $P(a)$ from equation 3.14. A random pair $\{a_1, P(a_1)\}$ is generated, and then rejected unless a_1 is found to belong to the distribution $P(a_1)$. If rejected, as in figure 3.4, new random pairs will be produced until an acceptable value a is found. This is the *rejection technique*, described in detail in the popular references [53] on the topic. Tailored adaptions based on approximate forms for $P(a)$ can improve the efficiency of the rejection technique, and these are called *combined techniques* [53].

When we consider the complexity of the band structure and scattering rates for most processes, it turns out that we usually turn to some form of the *combined* technique to select final states in our Monte Carlo implementations. Even then, as we will show, energy and momentum conservation in an anisotropic medium make it difficult to do this correctly within the traditional Monte Carlo framework.

For a completely isotropic and elastic scattering process, the selection of a final $\mathbf{k}'$ state would be straightforward. The $\mathbf{k}$-to-$\mathbf{k}'$ scattering rate, $\Gamma(\mathbf{k}, \mathbf{k}')$, does not depend on angle. Therefore, evaluation of 3.14 is straightforward and a direct selection method is possible for both polar and azimuthal angles. The final state can be determined by just two random numbers, one for each angle. This is the best possible scenario. Generally, the use of the rejection or combined technique requires several more random numbers per state selection for each scattering event.

We do not always choose *angles* to specify the final state. Sometimes a wavevector amplitude leads to a better formulation. Within the constraints of the traditional Monte Carlo, we tailor our approach to final-state selection based on the best symmetries of the effective mass and the scattering rate at hand. This leads us to treat equivalent processes for electrons and holes in rather separate ways.

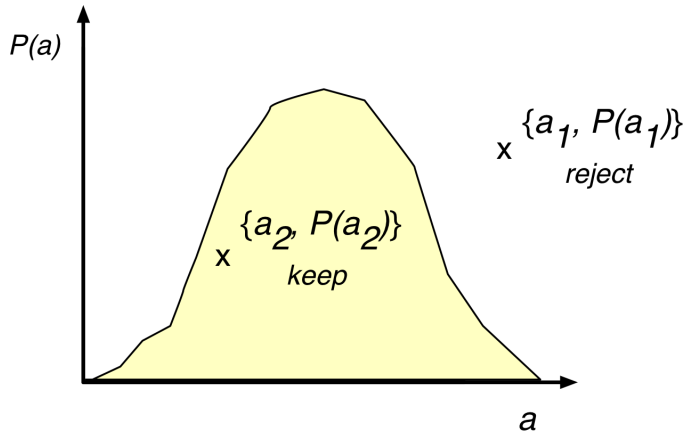

Figure 3.4: This figure represents the rejection technique. Random pairs of an appropriate range are generated. If these trial coordinates are found to belong to the desired distribution, they are utilized. If not, these points are rejected and new trial coordinates are generated.

Next, we describe specific aspects of how we implemented the Monte Carlo in this traditional framework which have been describing up to this point.

3.2.5 Implementation of the traditional Monte Carlo for electrons and holes

Here, we summarize the methods and approximations employed to simulate electron and hole transport within the traditional Monte Carlo framework.

The traditional Monte Carlo implementation for transport of electrons:

When using the traditional Monte Carlo framework, we included electron nonparabolicity in the calculation of scattering rates and in incrementing position and momentum. However, we used only scalar deformation potentials. Furthermore, we did not treat the mass anisotropy for electrons. As we use $\langle 100 \rangle$ geramanium crystals, the L-valley orientations are symmetric relative to the applied field. This symmetry decreases the influence of anisotropy compared to other orientations. So, in this case, we treated electrons with the spherical-band conductivity mass for acceleration and a density of states mass for rates.

The final-state selection for electrons was determined by the generation of the polar and azimuthal angles for the $\mathbf{k}'$ vector relative to $\mathbf{k}$.

The emission of intervalley and optical phonons were handled isotropically.

While it may be straightforward to include mass anisotropy for electron acceleration during free-flight, the anisotropy requires *absolute* coordinates relative to the

crystal orientation. This severely complicates the scattering probability needed for final-state selection. The Herring-Vogt transformation [64, 5], was designed as rough method to account for this fact. However, within the Herring-Vogt transformation, energy and momentum conservation are only approximate. Moreover, this approximation offers no simplification for the use of directional matrix couplings.

The traditional Monte Carlo implementation for transport of holes:

Holes spend only a few percent of the time occupying the light band before decaying to the lower-energy heavy band. Consequently, it is a common practice to neglect all but the heavy band. However, in this particular simulation, both heavy and light bands are considered in a parabolic way. In a manner similar to [8, 7, 53], this requires four independent scattering rates (intraband and interband) for each phonon scattering process. Warped bands have been accounted for. A direction-dependent effective mass is used to increment position and velocity. Similar to [8, 7, 53], wavefunction overlap integrals are included in calculating scattering rates, and in selection of a final k-state. This makes final state selection somewhat more complicated than the electron case. A random number search using a rejection technique is employed to select a suitable magnitude of phonon wavevector, q. Final values for hole wavevector and scattering angle are then constrained by energy and momentum conservation. Some common assumptions and approximations have been re-examined in order to accommodate our extreme low-temperature, low-bias case. Unlike assumptions in previous work [8, 7, 53], there is no approximation of near-elastic energy conservation for phonon magnitudes. Available scattering angles are dened explicitly and series-expansions are not used to satisfy energy conservation.

Within the traditional Monte Carlo framework, we did not implement the split-off band.

The warped nature of the heavy and light bands makes the treatment of holes complicated. This warping makes the heavy and light effective masses directional relative to the crystal frame, and creates a substantial overlap factor as a function of angle $\theta_{kk'}$

As with the electrons, the use of a directional deformation potential coupling relative to the crystal frame is again too complicated for use in the traditional Monte Carlo algorithm.

3.2.6 Problems with the traditional Monte Carlo

The approximations and idealizations we were forced to make under the traditional Monte Carlo, both for electrons and holes, stem from the reliance on integrated scattering rates and the difficulties encountered during final-state selection. Within scattering rates which are integrated over all *relative* orientations, it is extremely difficult to include scattering quantities which depend on the carrier's *absolute* ori-

entation within the crystal. It also difficult to include complicated integration limits which include properties about the final state, such as m_f, because we don't know prior to final-state selection what state m_f might represent. If we are serious about energy and momentum conservation, we are faced with the fact that the integrated scattering rate over all possible final states may change as the absolute orientation of the carrier changes, even with the same initial energy.

3.3 Monte Carlo simulation: a new vector-based approach

Isotropic assumptions are fundamental to the techniques of the traditional Monte Carlo, and the choice of scattering type is distinctly separate from the choice of final-state momentum. In this section, we explain how we have implemented a new algorithm that completely preserves anisotropy, despite any complicated formulation of a directional mass or matrix element. This algorithm also combines the selection of scattering type with the selection of final-state momentum.

The traditional Monte Carlo method, when determining properties of a scattering event, requires separate subroutines using separate random numbers for the determination of the type of scattering process versus the determination of the final-state momentum. Moreover, it is generally difficult to completely express scattering probabilities in terms of coordinates of the absolute reference frame of the crystal. The selection of the scattering process traditionally requires a comparison of *integrated* rates based on carrier energy. Since there may not be any analytical closed form preserving all of the anisotropic effects for the integrated rate, simplifying assumptions of isotropy are used. Even full-band Monte Carlo routines, which take care to numerically discretize the full electron band structure, will still rely on isotropic assumptions for the choice of scattering type. If we intended to take advantage of the directional information within the full matrix formulation of the deformation potentials, or if we wish to re-examine the scattering rate independent from a simple density-of-states mass, we are at a disadvantage.

We took care to rethink the traditional Monte Carlo algorithm and developed a modified Monte Carlo algorithm that preserves the precise $\mathbf{k}$-to-$\mathbf{k}'$ scattering rate, $\Gamma(\mathbf{k}, \mathbf{k}')$, at every step. This is accomplished by selecting first the direction of a possible final momentum vector, reframing Fermi's golden rule in terms of the chosen momentum variables, and then selecting between possible scattering rates. A self-scattering process is also included, as in the traditional Monte Carlo. When a physical scatter is chosen to occur, the momentum orientation, i.e., the final state, has also been determined as a consequence.

We can understand the selection process by realizing that emission or absorption processes are *three-body* collisions. In a thee-body interaction, there are nine vector components to constrain, as illustrated in figure 3.6a. The initial-state of the electron

is already known. This leaves us to determine possible values for all the final-state electron and phonon momentum vector components (six in all). Energy conservation as well as momentum conservation across three dimensions, gives us the ability to constrain four of the six needed components. The basis of this method is therefore to use a random number selection process to sample *bivariate pairs* of the remaining two momentum components. With all the momenta for a specific interaction, we can compute the rate. Our algorithm constantly samples specific final states to make the decision whether a physical scatter takes place. If a physical scatter *does* take place, the final state has already been determined, and a separate subroutine is not needed. So average quantities such as the total rate for the given initial state and the "density-of-states mass," m_{dos}, are not required in this method. We will need of course to sample the two momentum components according to their probability. We use a combined rejection method as explained in section 3.3.1.

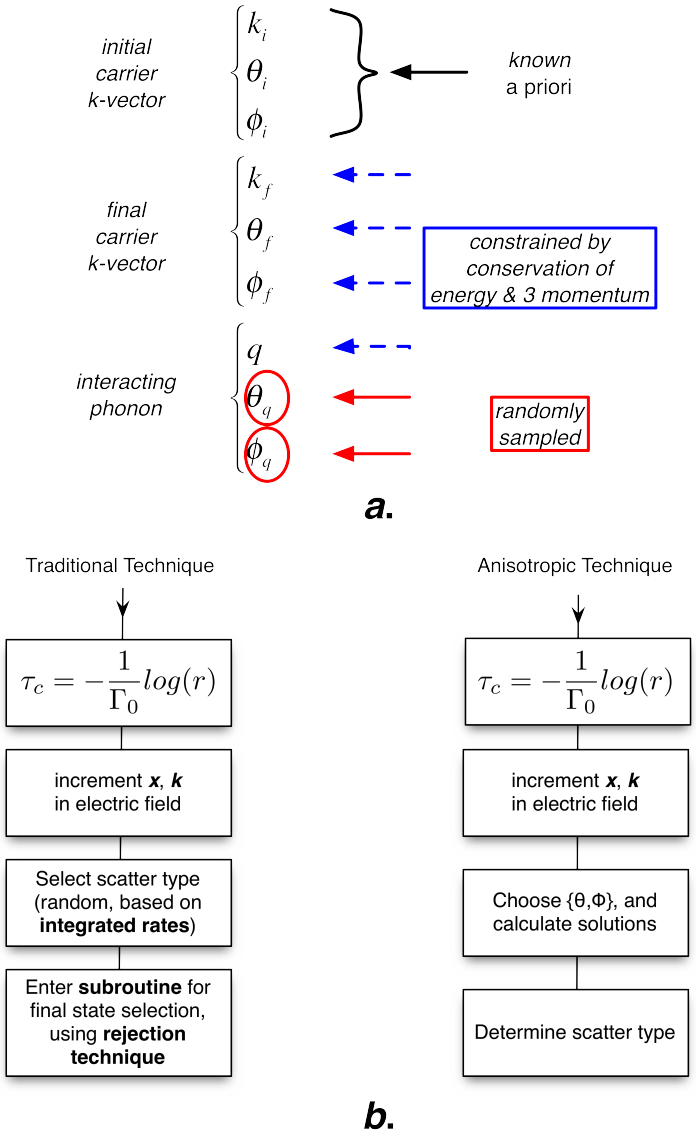

Figure 3.5: Simulating electrons with our new method. **a.** We have nine vector components in a three-body interaction. Three are known from the initial k-vector. Four are constrained by energy and momentum conservation. For electrons, we choose to sample the remaining two variables as phonon angles. **b.** A flow-chart comparison with the traditional Monte Carlo. On the left, the important steps of a carrier propagation sequence are revisited. In comparison, at the right are the steps of our "Anisotropic Technique." In the Anisotropic Technique, vector coordinate pairs $\{\theta, \phi\}$ are chosen by random number generation. Once the orientation is chosen, the resulting possible final state is calculated for all types of scattering. Rates for this particular orientation are tested to determine the scatter type. Therefore, when a physical scattering event is chosen, we also already have full knowledge of the final state.

In figure 3.5, we outline the general steps of the carrier propagation sequence in comparison to the traditional algorithm. First, a random time step based on net scattering probability is taken. This rate is based on the *total* scattering rate over all directions, which happens to be an analytical expression and easy to compute with our method. We then increment momentum and position, as in the traditional technique. We then sample a random direction, determined by a pair of random numbers chosen as we will soon describe. We then determine to see if a scattering event occurs. If a physical scatter does take place, the sampled direction has already determined the energy/momentum final state. The final state is simply implemented, and the transport propagation repeated.

3.3.1 Anisotropic electrons

Here, we discuss how our anisotropic Monte Carlo technique is applied to the treatment of electrons in germanium. We start with an explanation of our vector-based incarnation of Fermi's golden rule. This requires solving for possible solutions to phonon amplitude, q, using energy and momentum conservation. After describing the rate as a function of prospective angles θ_q and ϕ_q, we go on to describe our random sampling technique during free flight.

The vector formulation of Fermi's golden rule

In order to simplify the formalism we reformulate Fermi's golden rule for this method. In equation 2.74, we expressed the phase-space integral required for a total scattering rate, $\Gamma(\mathbf{k})$. We repeat it here.

$$\Gamma(\mathbf{k}) = \frac{2\pi}{\hbar}\frac{\Omega}{(2\pi)^3}\int d^3\mathbf{q}|\kappa_q|^2|A_{\nu,q}|^2 I(\mathbf{k},\mathbf{k}')^2\delta(\mathbf{k}'-\mathbf{k}\mp\mathbf{q})\delta(E_{\mathbf{k}'}-E_{\mathbf{k}}\mp\hbar\omega) \quad (3.17)$$

Furthermore, equations 2.66 and 2.66 informed us that acoustic deformation couplings for electrons were also found to be a function of *phonon direction* relative to the principle axis. We choose the local reference frame to be the electron valley, oriented with the principle axis along the z-axis. In this frame, the polar and azimuthal angles of equations 2.66 and 2.66 apply directly.

Next, we choose to perform the integral for Fermi's golden rule *only* over the range of allowed modulus for the momentum of emitted/absorbed phonon, q. In other words, we integrate the rate only over the *radial* component in the $d^3\mathbf{q}$ phase space. This gives the partial rate,

$$\frac{\partial^2\Gamma(k,\theta_q,\phi_q)}{\partial\theta_q\partial\phi_q} = \frac{2\pi}{\hbar}\frac{\Omega}{(2\pi)^3}\int q^2 dq\,(\sin\theta_q)\,|\kappa_q(\theta_q,\phi_q)|^2\,|A_{v,q}|^2\,I(k,k')^2\,\delta(f(q)) \quad (3.18)$$

where $\delta(f(q))$ is our shorthand for the combined delta functions over momentum and energy, expressed as a function of the phonon amplitude q. To evaluate the integral over q, we need to express the delta function in a useable way. We use the identity,

$$\delta(f(q)) = \sum_i \frac{\delta(q - q_i)}{|f'(q_i)|} \tag{3.19}$$

where the q_i are the roots for q which conserve both energy and momentum. The $f'(q_i)$ is the derivative of $f(q)$ with respect to q, evaluated at q_i.

Momentum and energy conservation, to find solutions of q

Considering parabolic bands, energy and momentum conservation give parabolic solutions to the amplitude of the phonon wavevector, q. $f(q)$ is the form

$$f(q) = a\ q^2 + b\ q + c \tag{3.20}$$

By definition, coefficients are determined by joint energy and momentum conservation to constrain

$$f(q) = 0 \tag{3.21}$$

The roots of q are

$$q_i = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{3.22}$$

For acoustic phonon processes, there is only one physical root when considering absorption and emission processes separately. Both roots are physical for optical phonons.

For acoustic phonon processes, we find solutions for q following

$$\frac{\hbar}{2}\ \mathbf{k}' \cdot \mathbf{m}^{*-1} \cdot \mathbf{k}' = \frac{\hbar}{2}\ \mathbf{k} \cdot \mathbf{m}^{*-1} \cdot \mathbf{k} \pm \hbar v_s |\mathbf{q}| \tag{3.23}$$

where $\pm$ is for absorption or emission respectively. Optical processes follow

$$\frac{\hbar}{2}\ \mathbf{k}' \cdot \mathbf{m}^{*-1} \cdot \mathbf{k}' = \frac{\hbar}{2}\ \mathbf{k} \cdot \mathbf{m}^{*-1} \cdot \mathbf{k} \pm \hbar\omega_0 \tag{3.24}$$

These equations provide the roots for equation 3.20 by supplying the constants a, b, and c. These constants depend on the components of $\mathbf{k}$ as well as angles θ_q and ϕ_q. For optical phonons, it is necessary to first find $\mathbf{q}$ by $\mathbf{k} \pm \mathbf{k}'$. While these constants are cumbersome and occupy several lines to write down, they do represent analytical, closed-form solutions.

Assembling the vector-based rate

With solutions of q as a function of initial electron $\mathbf{k}$-vector and phonon angles θ_q and ϕ_q, we can determine the overall partial rate of equation 3.18. Here, we show how the partial rate is assembled, piece by piece.

We illustrate using the example of acoustic phonon processes. Considering only one physical root q_i, the delta function depicting energy and momentum conservation can be expressed as the following.

$$\delta\left(f(q)\right) = \frac{\delta\left(q - q_i\right)}{\left|2aq_i + b\right|} \tag{3.25}$$

The partial rate of equation 3.18, for acoustic phonon processes, then becomes the following.

$$\begin{aligned}&\tfrac{\partial^2\Gamma(\mathbf{k},\theta_q,\phi_q)}{\partial\theta_q\partial\phi_q} = \\ &\tfrac{2\pi}{\hbar}\tfrac{\Omega}{(2\pi)^3}\int q^2 dq\left(\sin\theta_q\right)\underbrace{\left(\left|q\Xi_{e,a}^{\{i,j\},\nu}\left(\theta_q\right)\right|^2\right)}_{|\kappa_q(\theta_q,\phi_q)|^2}\underbrace{\left(\frac{\hbar\left(N_q+\frac{1}{2}\mp\frac{1}{2}\right)}{2\rho\Omega v_s q}\right)}_{|A_{v,q}|^2}\underbrace{(1)}_{I(\mathbf{k},\mathbf{k}')^2}\underbrace{\left(\frac{\delta\left(q-q_i\right)}{\left|2aq_i+b\right|}\right)}_{\delta(f(q))}\end{aligned} \tag{3.26}$$

After applying the delta function, we find

$$\frac{\partial^2\Gamma(k,\theta_q,\phi_q)}{\partial\theta_q\partial\phi_q} = \frac{\sin\,\theta_q}{8\pi^2\rho v_s}\left|\Xi_{e,a}^{\{i,j\},\nu}\left(\theta_q\right)\right|^2\left(N_{q_i}+\frac{1}{2}\mp\frac{1}{2}\right)\frac{q_i^3}{\left|2aq_i+b\right|} \tag{3.27}$$

which is actually a simple expression.

Solid-angle sampling manifold and global time step

In the previous section, we described how we have determined the exact rates for phonon processes considering an electron in state $\mathbf{k}$ that may emit or absorb a phonon in a specific direction. Now, we describe how we may stochastically sample physical processes from these rates in a numerical way, while incrementing time steps correctly.

The partial rates depicted by equation 3.18 represent interaction rates *per solid angle*. To be useful in our Monte Carlo, we need a method by which to sample this solid angle, and stochastically determine which processes occur. Furthermore, we need to be able to increment a global time step according to an integrated rate.

In order to implement a two-dimensional combined rejection method, we devise *a sampling manifold*, (an ellipse in the scattering angle space) which is a simple function with similar units of rate per solid angle. The principle is simply a multi-dimensional analogy to the scalar sample rate, Γ_0 (equation 3.11), from the traditional Monte Carlo. In the traditional Monte Carlo, it is important to pick a constant Γ_0 with a

rate which is greater in any direction than the sum of the physical rates than the net physical scattering rate, for all times. We depict this sampling manifold in figure 3.6, and define it $\Gamma_0(\theta_q, \phi_q)$ such that

$$\frac{\partial^2 \Gamma_0(\theta_q, \phi_q)}{\partial \theta_q \partial \phi_q} = \underbrace{\sum_{i=1}^{N} \frac{\partial^2 \Gamma_i(\mathbf{k}, \theta_q, \phi_q)}{\partial \theta_q \partial \phi_q}}_{\text{net physical scattering}} + \underbrace{\frac{\partial^2 \Gamma_{Self}(\mathbf{k}, \theta_q, \phi_q)}{\partial \theta_q \partial \phi_q}}_{\text{self-scattering}} \tag{3.28}$$

Simultaneously, we choose a global increment time in analogy with the Γ_0 of the traditional Monte Carlo.

$$\Gamma_0 = \int_0^{2\pi} \int_0^{\pi} \frac{\partial^2 \Gamma_0(\theta_q, \phi_q)}{\partial \theta_q \partial \phi_q} \sin\theta_q d\theta_q d\phi_q \tag{3.29}$$

We can still use

$$\tau_c = -\frac{1}{\Gamma_0} log(r) \tag{3.30}$$

Since electrons spend most of their time in the L-valleys, an empirically efficient sampling manifold represents an ellipse of a major/minor-axis ratio of ≈ 5. This is close, though somewhat larger, than the mass ratio $\sqrt{m_l/m_t}$.

Scatter determination for electrons

Just as in the traditional Monte Carlo, this algorithm increments momentum and position for electrons during free flights of time duration t_c. These free flights terminate in probability trials to determine if a scattering event takes place. Now, we describe the process that occurs at the termination of a free-flight, to determine the outcome of a scatter trial.

An ordinate pair of possible phonon angles $\{\theta_q,\ \phi_q\ \}$ is generated, chosen accordingly by the functional form of $\frac{\partial^2 \Gamma_0(\theta_q, \phi_q)}{\partial \theta_q \partial \phi_q}$ to ensure physically uniform sampling.

For scatter determination, the electron momentum vector is rotated into the reference frame where the valley's principal axis is oriented along the z-axis. Note figure 3.7. The electron momentum vector is then transformed into *polar* coordinates, still relative to the valley orientation. Next, our calculation of energy/momentum conservation determines the value of possible phonon amplitudes, q, for all processes. If a process is not allowed by energy or momentum for a particular instance and $\{\theta_q,\ \phi_q\ \}$, it will return a non-physical root of q and is disregarded.

A uniform random number $(0, 1)$ is generated. It is used to determine the outcome of this scattering trial, based on the sampled value of $\frac{\partial^2 \Gamma_0(\theta_q, \phi_q)}{\partial \theta_q \partial \phi_q}$ and the partial scattering rates. This is a multi-dimensional analogy to the selection method with self-scattering used in the traditional Monte Carlo, as expressed in equations 3.12 and 3.13.

If a scatter takes place, the determined root of q and angles are recorded as they are now considered the physical outcome. The final $\mathbf{k}'$-vector is chosen by momentum conservation, and the electron is put into its new state. The electron's new $\mathbf{k}'$-state is transformed into its final-state valley (depending whether the interaction was an intra- or inter-valley process), and put into cartesian, lab-frame coordinates. Another free-flight may now ensue, and the process repeats. Properties of the scattering events, phonons, and the electron's position, momentum, velocity are accordingly recorded to output.

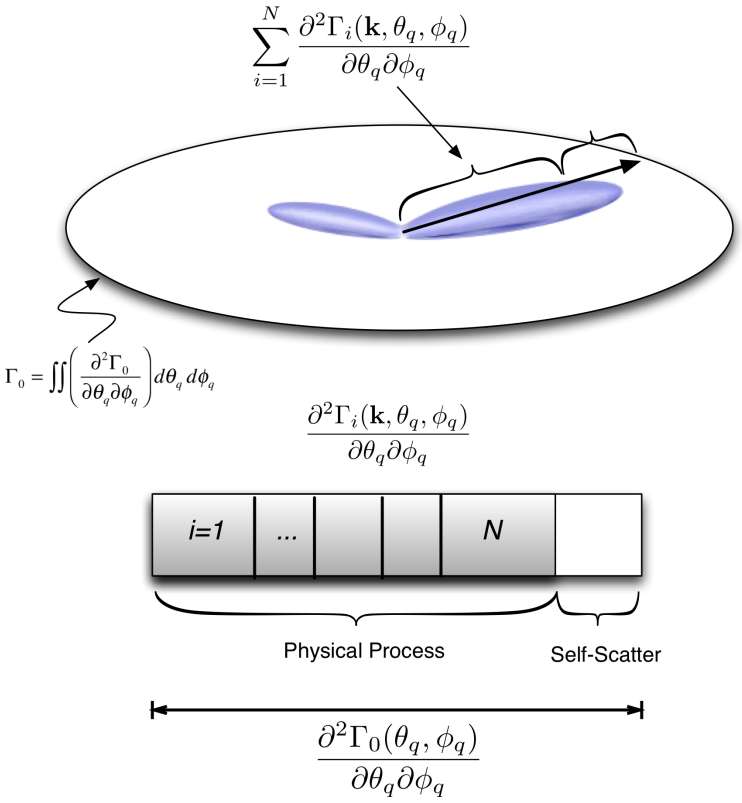

Figure 3.6: In the anisotropic case, the *vector*-based probability creates a manifold over angular coordinates θ_q and ϕ_q. While sampling from a "sampling manifold" (the encompassing ellipse), both the scattering process and final state are determined in the same step.

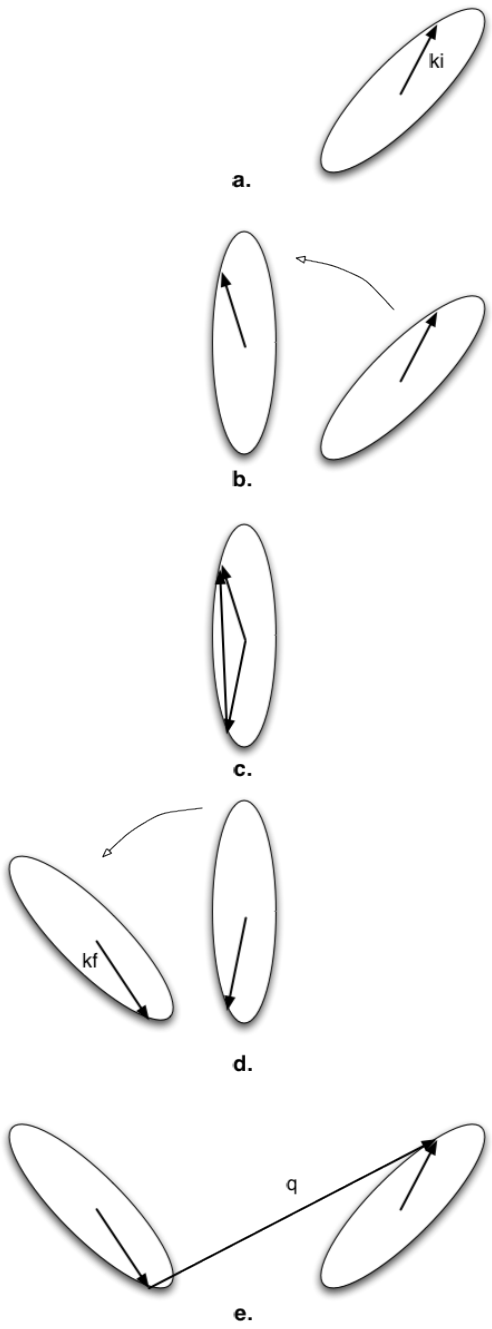

Figure 3.7: Transformations carried out when sampling electrons for scatter events. **a.** initial valley and k-vector, **b.** rotation to valley frame, **c.** conserve energy and momentum for a potential scattering event, **d.** if a scatter is carried out, rotate into the final-state valley (it will be a different valley, if an *inter*-valley process), **e.** if an intervalley process, resolve for phonon wavevector between $\mathbf{k}$ and $\mathbf{k}'$ states

3.3.2 Anisotropic holes

We have just explained the anisotropic simulation for electrons. There is less to explain for holes, as it is largely analogous. The primary difference is that holes are more complicated in their overlap factor and effective mass. So for simulating holes, it is easier to sample angles *of the final-hole state*, $\mathbf{k}'$. In brief, the same mechanics holds true for holes as for electrons, except that energy and momentum are solved for roots of the final state vector, $\mathbf{k}'$.

The mass relation for holes is complicated, due to the dispersion relation. Therefore roots of k' are chosen by the following, along with momentum conservation.

$$\frac{\hbar^2}{2}\frac{k'^2}{m\left(\theta_f,\phi_f\right)} = \frac{\hbar^2}{2}\frac{k^2}{m\left(\theta_i,\phi_i\right)} \pm \hbar\omega \tag{3.31}$$

Above, $\hbar\omega$ is general for either acoustic or optical phonons, and "$\pm$" as always stands for absorption or emission, respectively. The mass used is the *dispersion* mass, which directly relates the dispersion relation appropriate for energy conservation. Since $m\left(\theta_f,\phi_f\right)$ is a function of the final-state angles, it makes sense to generate these angles in our sampling of the phase space. Otherwise, to express these angles as representations of the other parameters is lengthy and taxing.

As explained in the previous chapter, we use matrix-based formalism for holes. This formalism requires all cartesian vector components of $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{q}$. Therefore, after solving for a root of k', a vector addition/subtraction with $\mathbf{k}$ takes place to also recover $\mathbf{q}$. This must be performed for every prospective value of k', for every $\{\theta_f,$ $\phi_f\}$ sampled, to evaluate the partial rates.

As a reminder, when we accelerate holes during free flight we use the *kinematic* effective mass. This was defined by

$$m_{ij}^{*-1} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}. \tag{3.32}$$

The warped nature of the holes is complicated in the $\{\theta_f,\ \phi_f\}$ parameter space. So for the sampling manifold $\frac{\partial^2\Gamma_0(\theta_q,\phi_q)}{\partial\theta_q\partial\phi_q}$ for holes, a perfectly spherical manifold was used.

3.4 Some end remarks regarding Monte Carlo techniques

In this chapter, we have reviewed the Monte Carlo techniques developed for this work. Originally, we used an algorithm based on traditional techniques. Where energy and momentum conservation is only approximate and only true on average. Furthermore, although final-state selection routines in the traditional method attempt

to "put back" anisotropy into simulation, there is still a reliance on integrated rates based on Fermi's golden rule.

Ultimately, we developed a vector-based version of Fermi's golden rule. A simulation routine based on a phase-space sampling of such rates takes place over samples of simultaneous angles. This enabled us to incorporate full knowledge of the deformation potentials and matrix formalism than has been developed for germanium over the past decades, without need for approximation.

Approximations which remain in this implementation include parabolic bands. The nonparabolicity factor used to adjust the bands in traditional methods is itself isotropic. One could imagine a nonparabolic correction to the vector-based method, and now even contemplate the directional dependence of such a term. Other corrections could include the anisotropy of the phonon "slowness surface" (i.e., the anisotropy of the speed of sound), but this would make determination of the required roots quite difficult. In any case, the anisotropy in the germanium speed of sound is much less than that of the carrier dispersion relation, at least in considering electrons. We also included only one effectively "lumped" transverse phonon branch for electrons. The size of this effect is small though, as the "fast" and "slow" TA branches are quite similar in magnitude for germanium near $q \approx 0$.

In the next chapter, we report on the results of these efforts.

Chapter 4

The Presentation and Interpretation of Monte Carlo Simulation Results

In previous chapters, we explored the band structure and scattering mechanisms for electrons and holes in germanium. We also reviewed the traditional Monte Carlo technique, and introduced our own novel, vector-based algorithm.

In this chapter, we explore the results of simulation based on this foundation. We will see how electrons and holes behave as a function of electric field. We compare results to experiment, and show how these results form the basis for yet more theoretical predictions. In a nutshell, validating carrier kinematics through a comparison between simulated and measured drift velocities allows us to predict the probabilities for other physical mechanisms in our detectors. These results continue to be a valuable tool in understanding the behavior of our detectors. In particular, our understanding of the mechanisms behind the evolution of space charge and the rate and spectral emission of Luke-Neganov ("Luke") phonons benefit from these developments.

We begin this chapter by reviewing some results from our original, isotropic Monte Carlo simulations based on the traditional algorithm. Next, we introduce the effects of anisotropy as explained in the previous chapter. The resulting distributions of wavevector, velocity, and position show that anisotropy has a pronounced effect on carrier transport. We then report on macroscopic transport properties based on our simulation results. These macroscopic quantities, based on ensemble averages, further our understanding of the underlying phenomena as a function of applied field.

We treat electrons and holes on a more or less even footing. In this way, we juxtapose the dynamical behavior of the two species for deeper insight. Among other properties, we report on results for energy distributions as a function of applied external field. Knowledge of the carrier energies enable us to make further predictions, necessary in the remainder of this thesis. We will also discuss the phonon emission distributions that take place.

We conclude this chapter by reporting on the particular case of electrons drifted in the $\langle 111 \rangle$ direction, discussing the new predictions this simulation has produced for drift physics at $T = 40\ mK$.

4.1 Results from the isotropic Monte Carlo

Our original implementation of a traditional, isotropic Monte Carlo simulation served well in advancing our understanding of general transport properties under our non-equilibrium conditions. As shown in our reference [6], we reproduced the theory and experimental results for both holes and electrons at $T = 8\ K$ [8, 13, 7, 74], and made a viable prediction for both species for our $T = 40\ mK$ base temperature.

For electrons, the results from this simulation predicted a further increase in drift velocity between $T = 8K$ and $T = 40mK$ by a factor of over 3 for electrons, at low fields. As we utilize $\langle 100 \rangle$ germanium, the influence of the Brillouin zone L-valleys is symmetrical relative to the electric field. The elliptical nature of the L-valleys was assumed to be averaged out, and a spherical approximation was used. So for this first model, we simply used a single valley approximation. However, non-parabolicity was included in calculation of the scattering rates and the acceleration kinematics. Emission and absorption rates were included for the acoustic, intervalley, and optical phonon processes.

Previous work by other authors [8, 74, 53] on hole simulation did not match well to data in the low-temperature and low-field range. We attempted an isotropic Monte Carlo simulation, taking great care in conserving energy and momentum in our scattering rates and final-state selection. A split-off band was not yet included for this model. Scattering rates were based on scalar density-of-states masses. When incrementing holes in the field, we included a directional mass based on inverting the dispersion relation for the heavy and light bands (i.e., the *dispersion mass*). We later found this definition of mass was also utilized in reference [54]. These improvements allowed us to reproduce the data of [8, 7] at their lowest published temperature of 8K, even where previous simulation did previous simulation did not perform well in this low-temperature regime. Extending our simulation to T=40 mK allowed us to predict that drift velocities were expected to change below 3 V/cm. This would be later confirmed to excellent precision by our subsequent experimental drift velocity measurements [10] at T 40mK as in figure 4.1.

Although holes in the heavy band are considerably heavier than L-valley electrons, they have considerably lower scattering rates. The deformation potentials for holes smaller than for electrons, and the warped nature of the valence bands contributes to a reduced wavefunction overlap integral. This signifies a reduced phase-space overlap with the density of possible hole final states. Therefore, holes "brake" to the lattice less than electrons, and are allowed to accelerate and carry much more energy for a given field than electrons. This was apparent in this simulation. For

fields above 3 V/cm, hole velocities should be determined predominantly by optical phonon emission, for both T=8 K or 40mK. That is, optical phonons should begin to dominate the dissipation of power for holes above about 3V/cm. Substantial emission of optical phonons at such low fields was not previously anticipated. The results of later anisotropic simulation, to be described below, also reveal this same behavior.

It is a curious fact that subsequent drift velocity measurements at T=40 mK do not appear to evidence the onset of optical phonon emission from holes as predicted, although this phenomena gave excellent consistency with the T=8 K data. This remains an issue to be resolved, no doubt aided by ongoing drift measurements at $T \sim 40\ mK$ now underway.

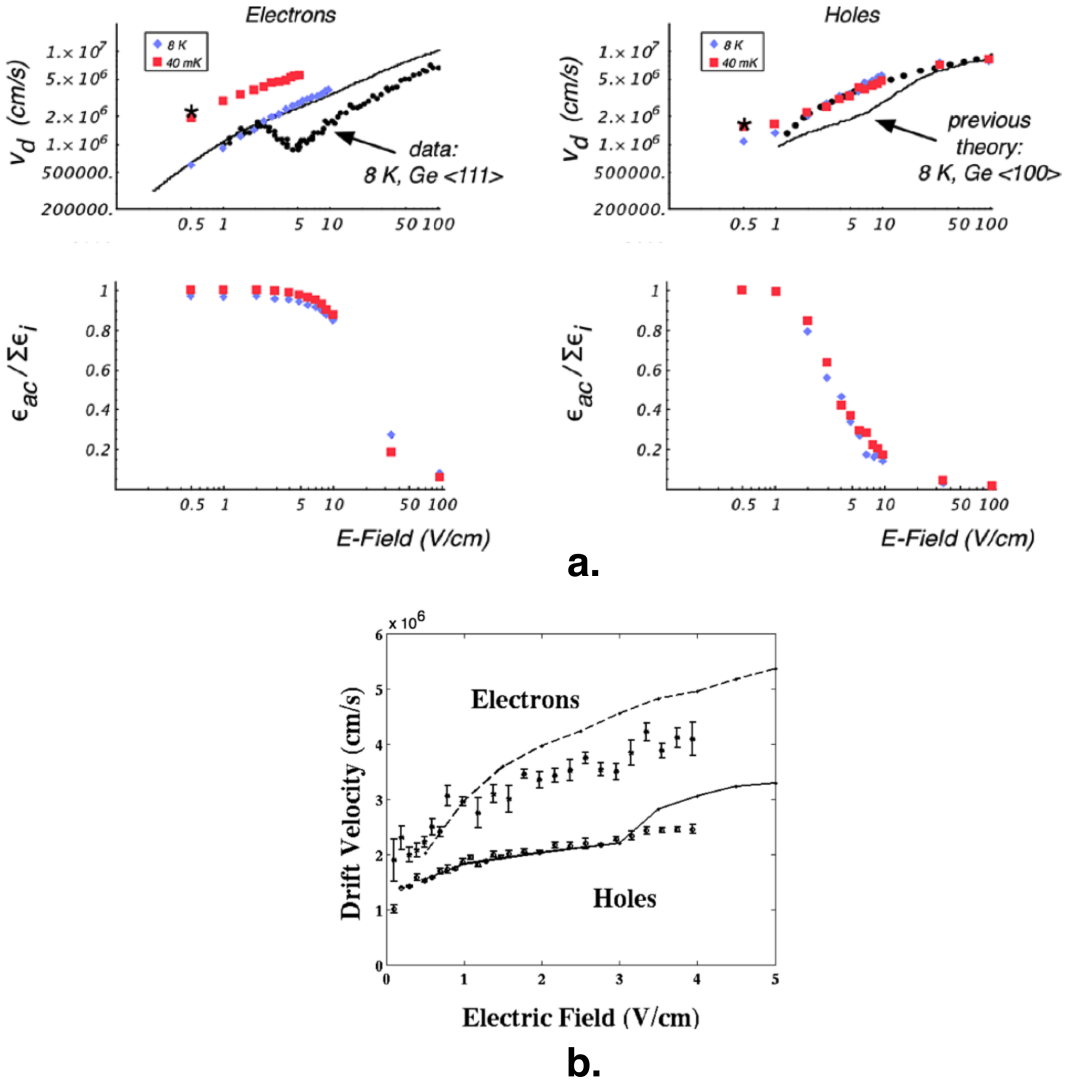

Figure 4.1: Results from our traditional Monte Carlo, from [6], using energy-based scattering rates and a separate subroutine for final-state selection. **a.** Predicted drift velocities for electrons and holes, for T=8 K and for T=40 mK. *fine trace:* T=8 K theory from [7], and *round dots*: data from [8, 7]. Note that we already improved momentum conservation in simulation for low-energy holes (needed for low-temperature, low-bias conditions). Markers, "*," denote $T = 20\,mK$ data points from [9]. Also shown is typical energy emitted by phonons as a ratio of acoustic phonons to the total. This shows that power dissipated by holes to optical phonons is equal to power dissipated to acoustic phonons above approximately 3 V/cm. **b.** Same T=40 mK predictions, overlain with subsequent drift velocity data taken at $T = 31\ mK$, from [10]. Note that there is little indication of the optical phonon "knee" for holes near 3 V/cm.

4.2 Simulation with full anisotropy ("vector-based Monte Carlo")

This initial simulation was followed by a full anisotropic, vector-based simulation which allowed us to couple carriers to all phonon branches with no approximation in the scattering rates or final-state selection. We simulated transport knowing that expressions for band structure effective mass and Fermi's golden rule were fully preserved for every scatter and every final-state selection. We review below the resulting statistical distributions pointing out the physics involved.

4.2.1 Anisotropic electrons: k-vectors, velocities, and displacement

In this first section, we report on the statistical distributions for electrons using our vector-based anisotropic simulation.

The majority of literature regarding semiconductors is based on silicon. In silicon, the conduction band minima are the X-valleys. The electron anisotropy factor, or the ratio of longitudinal to transverse effective masses $K = m_l/m_t$, of these silicon X-valleys is $K = 5.2$ [20]. This is far less than the anisotropy of the L-Valley minima in germanium, $K = 20$ [20]. As a consequence, isotropic assumptions are less applicable for treating electron transport in germanium. Furthermore, anisotropy effects are heightened at low temperature and low bias. In this case, electrons are more prone to remain isolated in their particular L-valleys without becoming randomized. Their overall movements will stay correlated with their specific L-valley for longer periods of time. So, there is particular need to treat electron anisotropy correctly for germanium in the low-temperature limit. Our Monte Carlo development is therefore particularly useful in this situation. Not only do electrons accelerate with an appropriate mass to an applied field vector, but any directionality in the scattering response is also preserved without a reliance on integrated rates.

As a demonstration, consider figure 4.2. In figure 4.2a., we see a representation of the iso-energy surface, in **k**-space, defined by the dispersion relation of the electron L-valley in the $\langle 111 \rangle$ direction. This iso-surface represents a particular kinetic energy above the L-valley minimum. As we have seen by equation 2.11, the wavevectors **k** must be perpendicular to this iso-surface when energy is the same in all directions. We see by plot 4.2b. that the final, steady-state distribution of simulated wavevectors follows a profile similar to the iso-surface. This shows that kinetic energy is rather equally partitioned in all directions. We also know, also by equation 2.11, that a physical velocity is the projection of a momentum wavevector upon the inverse mass tensor. In figure 4.2d., we see a final distribution of velocities in this valley. These velocities, indicative of the ongoing velocity fluctuations in the steady state, follow the form of a k-vector projected onto the inverse mass tensor (also depicted in sub-plot

a.). There is a non-trivial *average* velocity for these distributions, however. Electrons following the randomized velocity *fluctuations* of figure 4.2d. are also moving with the average velocity, which is harder to notice. In plot 4.2c., we see a distribution of electrons in this valley after drifting for $\approx 2 \times 10^{-7}$ seconds. Their *dispersion* is somewhat saucer-shaped somewhat like the velocity distribution of plot 4.2d. However, we also notice the tensor-nature of the inverse mass, in that the *mean* displacement is not parallel to the field. This *oblique propagation* is characteristic of the low field transport at very low temperature. Our work confirms the result of previous analytical work. To some approximation this effect can be treated as a "mobility tensor" phenomenon with the analytical infrastructure presented in classic textbooks such as [20, 51]. It has been been simulated with traditional Monte Carlo techniques with inverse mass tensors in the zero-point limit as in Aubry-Fortuna, et al. [79], and Cabrera, et al. [80]. This behavior is compatible with our experimental results.

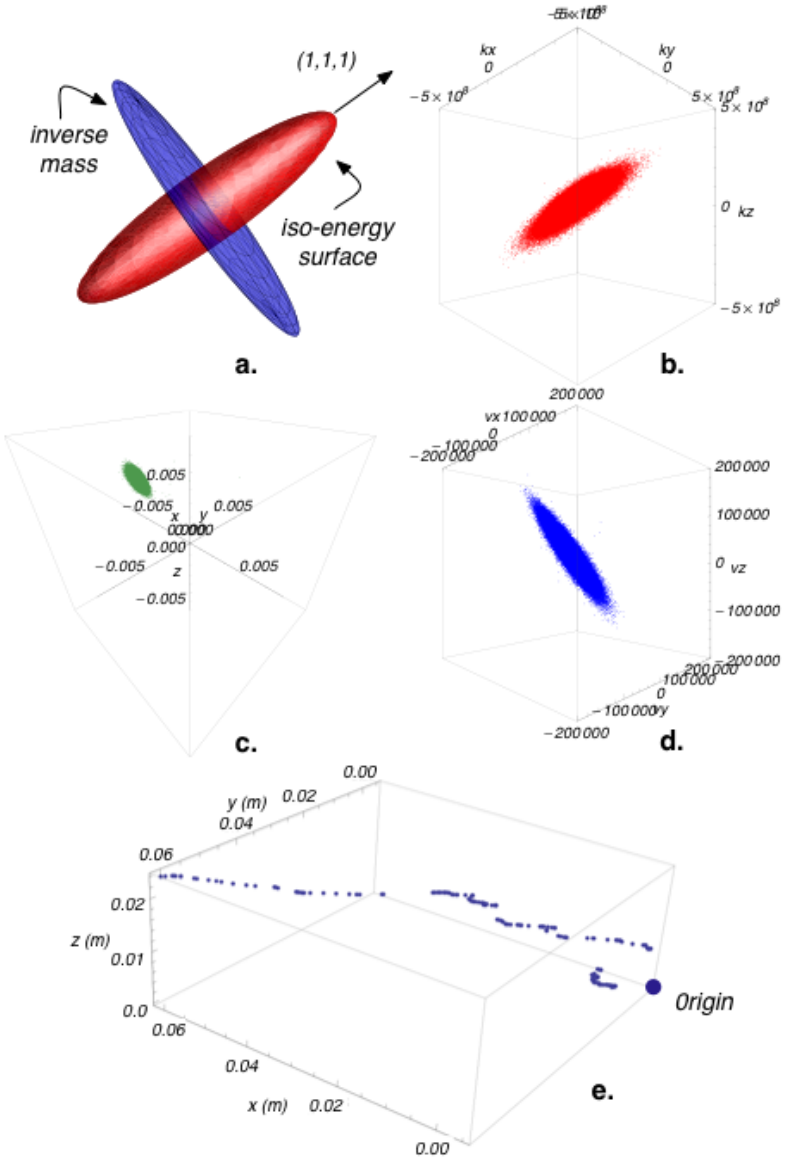

Figure 4.2: An example of electron distributions in the $\langle 111 \rangle$ L-valley for $F = -1\ V/cm$ along z. All units are MKS, in this instance. **a.** the iso-energy surface, and the momentum projection of the inverse mass tensor, both re-scaled. **b.** the ensemble of steady-state wavevectors. **c.** the ensemble of final real-space positions after $\approx 2 \times 10^{-7}$ seconds. **d.** the ensemble of steady-state velocities. **e.** a single electron starting from rest in this valley, and propagating over time.

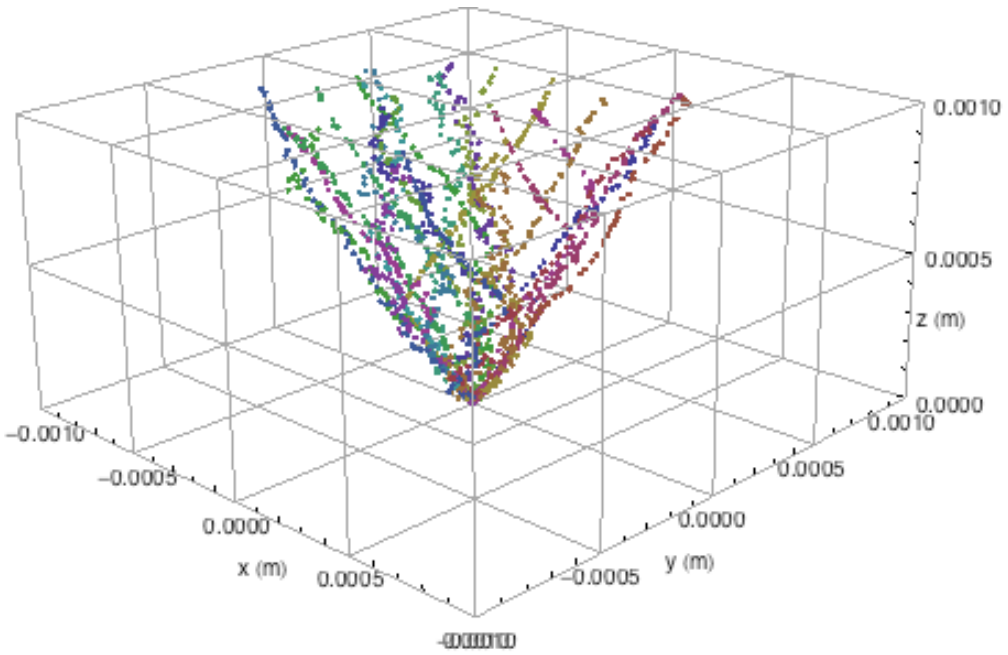

Figure 4.3: Electrons in germanium ⟨100⟩ starting from rest in a strong, z-oriented electric field. Intervalley transitions in **k**-space result in different inverse mass tensors with different favored directions relative to the field.

4.2.2 Anisotropic holes: k-vectors, velocities, and displacement

Considering now holes, the fully anisotropic simulation that we performed along the method described in the previous chapter displays strong anisotropic effects, although the propagation is on average no longer oblique.

From the dispersion relations, we find the warped valence bands are also far from isotropic. In our original, "isotropic" Monte Carlo, we used a directional *dispersion mass*. Contrary to other references which also use this definition for hole mass [54], in our full anisotropic simulation we use the rigorous definition of kinematic mass as $m_{ij}^{*-1} = \hbar^{-2}\ \partial^2 E/\partial k_i \partial k_j$. The resulting expressions are complicated by the fact that the warped valence bands for the heavy and light holes depend on all three wavevector components in a complex way. The resulting inverse mass tensor for the heavy whole band is given in figure 4.4 part a.

We present distributions for holes analogous to the previous section for electrons. The electric field (now $F = +1V/cm$) is parallel to the (001) direction. As holes energetically favor the occupation of the heavy(top) band, the heavy-hole inverse mass tensor dominates the kinematic response for holes at low fields. We present statistical ensembles from our Monte Carlo simulation. In figure 4.4b., wavevectors are distributed in a somewhat boxy relation due to the energy iso-surface of the heavy band. The steady-state velocities of figure 4.4d. map well to the inverse mass tensor, again dominated by occupation of the heavy band. The resulting distribution of holes in real space (figure 4.4c.) tends to be very spread out along the vertical axis where the heavy mass tends to be lightest.

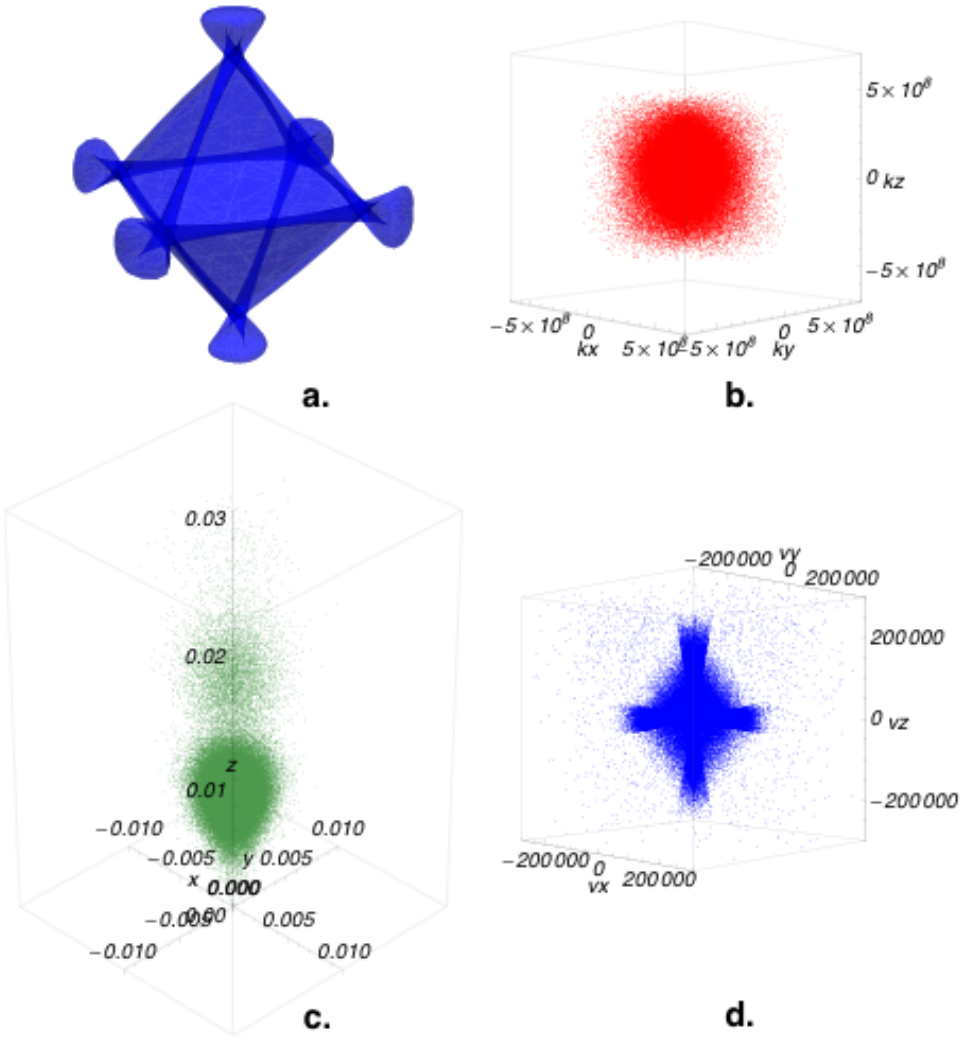

Figure 4.4: An example of holes distributions for $F = 1\ V/cm$ along z. All units are MKS, in this instance. **a.** the momentum projection of the inverse mass tensor for the heavy band, defined by k **b.** the ensemble of steady-state wavevectors. **c.** the ensemble of final real-space positions after $\approx 5 \times 10^{-7}$ seconds. Note the propensity for z-axis diffusion. **d.** the ensemble of steady-state velocities, which are dominated by occupation of the heavy hole band.

4.3 Introduction to macroscopic properties

In the previous sections, we introduced some of the general characteristics of steady-state carrier ensembles.

Here, we introduce important average properties of electrons and holes as a function of applied field. We generally consider the case of an electric field applied in the z-direction, although later we investigate electrons in the $\langle 111 \rangle$ orientation.

The idea is to define macroscopic parameters that will be useful to our understanding and future simulation work. The idea is that, if we can assume that carriers are in a steady-state with the electric field, then particular macroscopic parameters will be useful for higher-level work.

We take advantage of the fact that this system is assumed to be *ergodic*, as per Boltzmann's hypothesis for gaseous particles. In this case, the general properties of a single particle in the steady state is characteristic of an ensemble of similar particles. This means that we could take *correlation functions* at different points in time with the same carrier, *or equivalently* we could take statistical moments of an ensemble of many such carriers. As we have taken the approach of simulating several particles until they have reached a steady state with the field, we take moments of that ensemble and associate them with probabilistic behavior of any such particle. Our motivation here is to depict our simulation output as usable, macroscopic expectation values of processes occurring in the germanium bulk. The purpose is that this will help us make predictions for processes occurring within our high-purity, milliKelvin CDMS detectors.

In forming appropriate definitions for useful quantities, a constant theme is that the expectation value of multiplied quantities is in general not the same as the multiplication of expected quantities. In other words,

$$\langle AB \rangle \neq \langle A \rangle \langle B \rangle \tag{4.1}$$

As we proceed in this chapter, we point out our assumptions in this regard.

As most properties are applicable to both electrons and holes, we present results for both carriers side by side.

4.4 Velocities

We need to be careful about defining quantities related to velocity. We make three distinctions for velocity.

Drift velocity Drift velocity is equivalent to the *average* velocity.

Thermal velocity Velocity *fluctuations* are generally much higher in speed than the small average velocity represented by drift. The average "deviation" of velocities gives the thermal velocity.

Total velocity Typically, the thermal velocity is quite nearly *all* of the velocity since it represents dwarfs the average velocity. However, to be general, we are particular enough in this work to make a distinction in definitions. We define the "total velocity" the mean modulus of velocity.

For the next three subsections, we take some effort to define these three velocity quantities more specifically.

4.4.1 Drift velocity

Carriers accelerate in an applied electric field, but dissipate power by inelastic scattering processes. In the steady state, charge carriers have reached some terminal, average velocity which we call the *drift velocity*. The drift velocity is extracted from simulation simply by finding the average velocity vector.

$$\mathbf{v_d} =< \mathbf{v} > \tag{4.2}$$

A note on drift velocity for electrons: There are the four, full L-valleys representing the conduction band minimum. Electrons may be initially excited into any one of these L-valleys. While the L-valleys represent the same energy minimum in the Brillouin zone, they may manifest different dynamics depending on orientation of the applied field. To accurately describe electron behavior in the general case, we would need to specify electrons occupying specific valleys.

The case of a $\langle 100 \rangle$ (specifically (001)) electric field orientation affects all four L-valleys symmetrically. Therefore, we can present results for the z-axis drift velocity and know that it is the same for all L-valleys.

4.4.2 Thermal (or "random") velocity

As carriers propagate, their motions are largely stochastic. They are undergoing random motion in just about any arbitrary direction. The average velocity is typically only a small offset compared to the velocity fluctuations ongoing at any instant. So the *thermal velocity* generally describes the randomized velocity of carriers from point to point, between instantaneous scattering events. Whereas the *drift* velocity represents the mean velocity, the *thermal* velocity represents the velocity distribution's *standard deviation*. Assuming that the carrier distributions are depicted well by *displaced Maxwellian*, a mean and a standard deviation in velocity (or momentum) would completely quantify the form. In general this is not the case and higher moments would be necessary. The thermal velocity is equivalent to the standard deviation of velocity.

$$v_{th} = \sigma_v \tag{4.3}$$

The term "thermal" could reasonably imply thermal equilibrium, which need not be the case. "Random" may be a better term. In the case of hot carriers, this thermal/random velocity is the best metric for something resembling temperature, provided the distribution profile resembles a Maxwellian distribution.

In 3 dimensions, this velocity distribution width can be described by a covariance matrix.

$$\Sigma^2_{v,ij} = \mathrm{cov}(\delta v_i, \delta v_j) =< \delta v_i,\ \delta v_j > \tag{4.4}$$

where

$$\delta v_i = v_i - < v_i > \tag{4.5}$$

To recover a velocity quantity with statistical relevance in 3D, we can take a *matrix decomposition* of the covariance tensor. Several methods of matrix decomposition are possible.

$$\mathbf{v}_{th} = \mathbf{\Sigma}_v = \mathbf{U} \tag{4.6}$$

where $\mathbf{v}_{th}$ and $\mathbf{U}$ are still matrices.

$$\mathbf{\Sigma}^2_v = \mathbf{U}^T\mathbf{U} \tag{4.7}$$

Now, for the sake of defining a single expected thermal velocity, taking into consideration all directions, we can define a quantity based on the matrix *trace*. The matrix trace is invariant under rotation.

$$\langle v_{th} \rangle = \sqrt{\mathrm{Trace}(\mathbf{\Sigma}^2_v)} \tag{4.8}$$

4.4.3 Total velocity

We were quite specific in defining the thermal velocity in terms of the velocity *variance*, which also defines temperature in a displaced Maxwellian distribution. Often, we may want to connect this to the total velocity.

If a carrier velocity (or momentum) distribution is well-depicted by a *displaced Maxwellian* profile, then a randomly sampled total velocity should be related as

$$\mathbf{v} = \mathbf{v}_d + \mathbf{v}_{th}\ \mathbf{r} \tag{4.9}$$

where $\mathbf{r}$ is random number array, sampled from a gaussian $\sigma = 1$ distribution.

Again, the thermal/random velocity fluctuations are typically much higher in amplitude compared to the drift velocity contribution. In most cases, the average total velocity should quite nearly equal to the thermal/random velocity. So comparing average total velocity to thermal velocity is a useful check.

$$v_{tot} =< \sqrt{\mathbf{v} \cdot \mathbf{v}} > \tag{4.10}$$

As we are sampling an ensemble of particles, we use this simple and straightforward definition to define total velocity.

We note that we took care in defining a statistical "thermal velocity" separate from a "total velocity." The thermal velocity is taken from the velocity variance, and has its roots based in the *width* of the velocity distribution. Since the velocity deviation so often is much greater than the norm of the average velocity, the "total velocity" and "thermal velocity" are almost always quite nearly the same and are often associated with each other.

4.5 Mean-free path

Since we know the total velocity as a function of applied field, we can combine this with typical scattering times (to be described more, shortly). This simple scalar definition is a useful characteristic scattering length.

$$l_{mfp} = \langle v \rangle \langle \tau \rangle \tag{4.11}$$

Here, $\langle \tau \rangle$ is the mean scatter time between physical scatters.

This scattering length as a function of field is shown in figure 4.6b. We itemize specific lengths for both electrons and holes, using the sum rate of all acoustic phonon processes (leading to l_{ac}) and also the sum rate of all optical and intervalley rates (leading to l_{op}). So we can see from figure 4.6 it is a good approximation that acoustic and optical scattering processes combine, as per Mathieson's rule.

$$1/l_{mfp} = 1/l_{ac} + 1/l_{op} \tag{4.12}$$

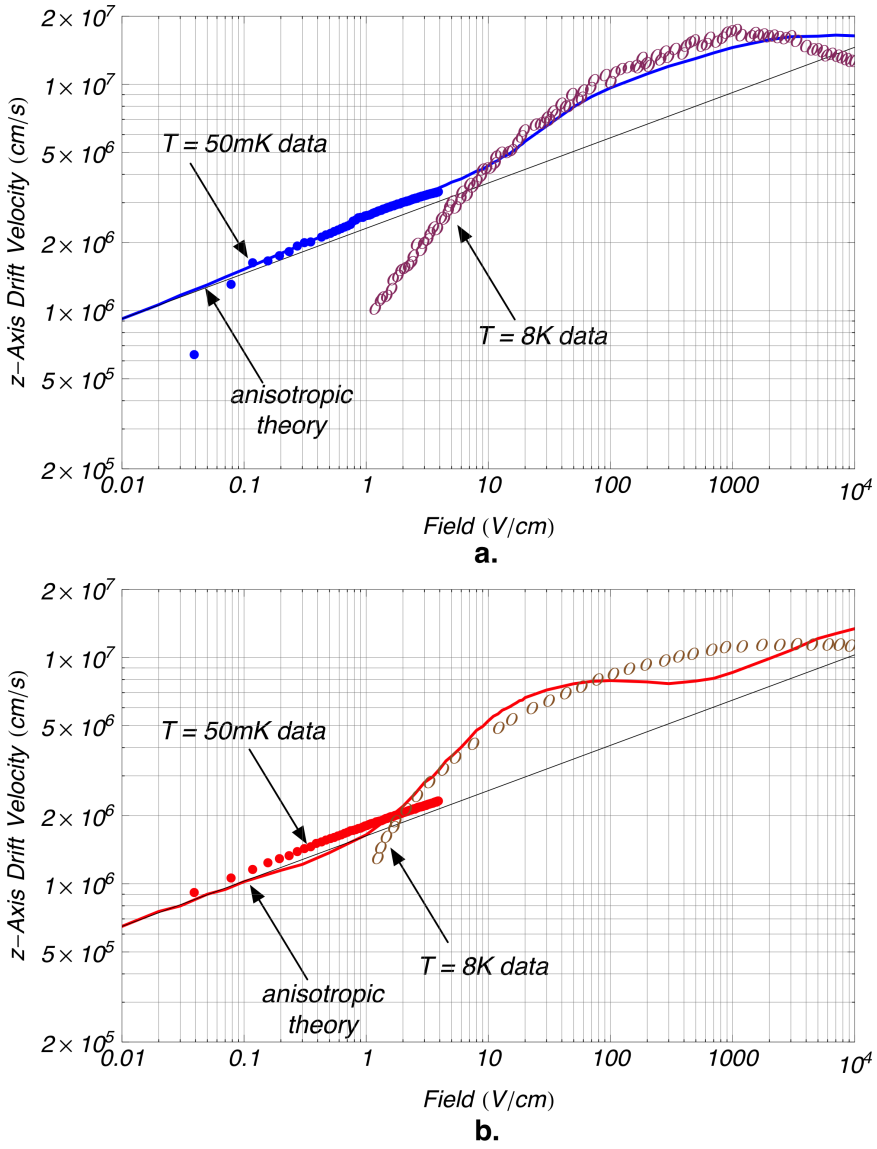

Figure 4.5: Drift velocities from the new anisotropic (vector-based) Monte Carlo technique, along with the latest data at T=50mK from [11]. **a.** For electrons, the increase in drift velocity from T=8K to a simulated T=40mK is just as predicted. The lowest data point near 40 mV/cm is most likely a systematic of inherent built-in fields, which are identifiable when grounded. **b.** For the full anisotropic simulation of holes, the values used are from literature and predominantly from [12]. Drift velocities are $\approx 9\%$ lower than data for $F \leq 1\ V/cm$, but there is subtle coupling to other bands not present in the previous simulation at these low fields. The threshold for optical phonon dissipation to dominate over acoustic emission occurs at the same field ($F = 2.48\ V/cm$) consistent with our traditional Monte Carlo simulation. The corresponding effect on simulated hole velocities also remains, and is consistent with the T=8K data. Thin black lines match simulation at $F = 0.01\ V/cm$ and follow a $F^{1/5}$ power law as per appendix B.

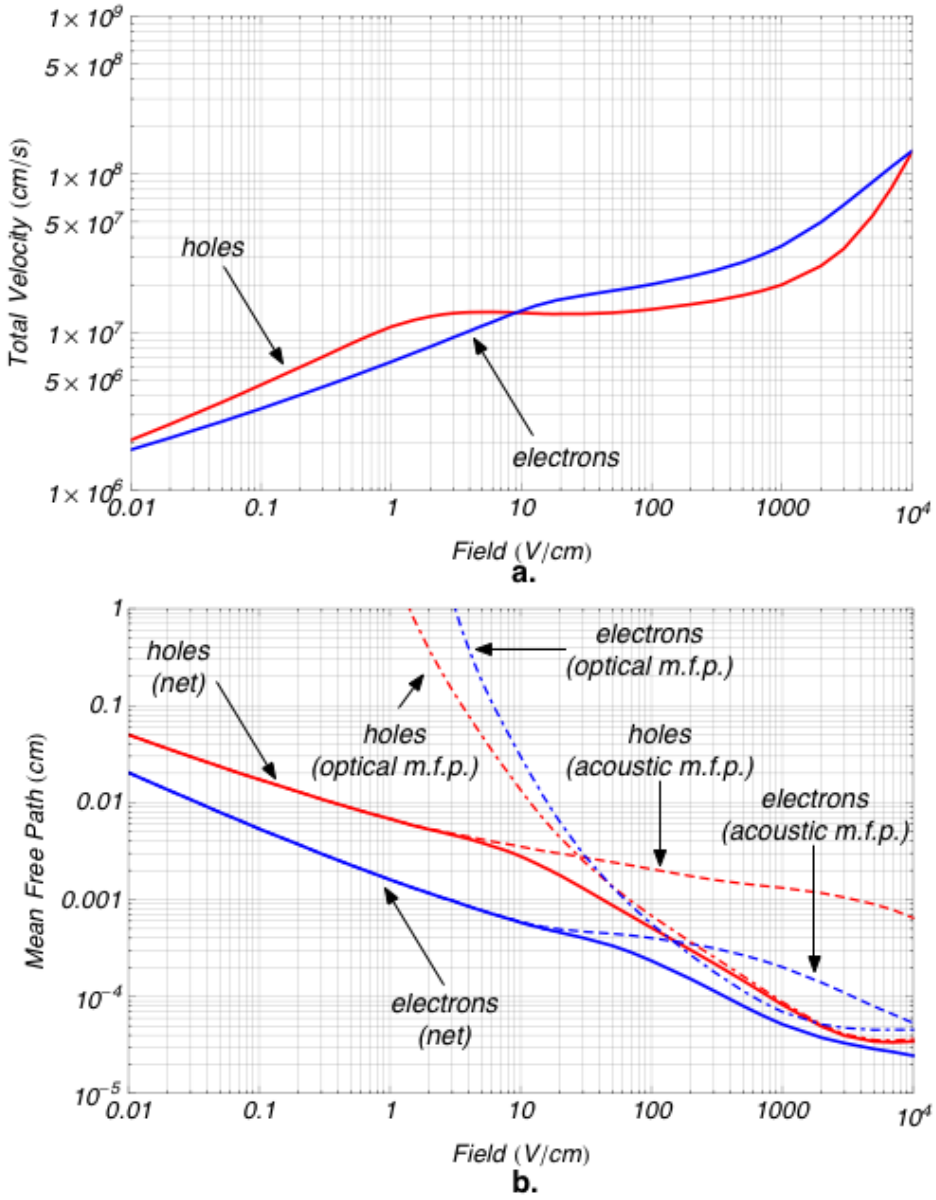

Figure 4.6: The total velocity and mean-free path. **a.** the total velocity for electrons and holes, which we distinguish as a separate quantity from the *drift* or *thermal* velocities. **b.** the mean free path for electrons and holes, itemized into both total acoustic and total optical mean free paths.

4.6 Carrier temperature

Another quantity tied to the definition of velocities is the effective carrier temperature. Once again, because our velocity distributions need not necessarily be Maxwellian, a temperature may not be defined in the strictest sense. Nonetheless, we can take the statistical variance of the velocities of our simulated particle ensembles and empirically associate it with a temperature quantity. Since we have been preserving the anisotropy, we may want to specify a tensor notation for this temperature quantity as per [45]. We can specify a temperature for every degree of freedom of our system, and in general relate the velocity distribution "width" to components which essentially depict the temperature in different directions.

When we introduce a mass tensor into our work, we need to be careful about our definitions. A simple path is to recognize that energy can also be defined the dot product of momentum and velocity. In this case, we can be sure that we can pick a correct form of the "thermal tensor" with a correctly preserved and invariant trace.

$$\frac{1}{2}k_B T_{ij} = \frac{\hbar}{2}\mathrm{cov}\left(\delta k_i, \delta v_j\right) \tag{4.13}$$

We are free to rotate this energy tensor into any reference frame we wish. As the trace of this matrix is invariant, it makes a sound definition to define a scalar temperature as the contribution of all three degrees of freedom from our 3×3 temperature matrix. In general, random fluctuations evenly distribute energy quite rapidly among the three degrees of freedom. Taking the trace and defining a scalar diagnostic temperature is not a terrible assumption for a diagnostic parameter. We therefore define the (scalar) carrier temperature as the following.

$$\frac{3}{2}k_B T_C = \mathrm{Trace}\left(\frac{1}{2}k_B T_{ij}\right) \tag{4.14}$$

We plot the scalar carrier temperature for both electrons and holes in figure 4.7.

To note, some [74] would refer to our definition of temperature as the "noise temperature." This is because we used the velocity *fluctuations* as our definition. If we had subscribed to this definition, then the "carrier temperature" is a separate temperature quantity proportionate to the mean *total* velocity squared. We prefer to retain our definition of temperature by fluctuations because it relates to the *Wannier relation*, which we will define shortly. To emphasize, these two definitions of temperature are in practice typically about the same, anyway.

4.7 Diffusion

We have introduced the concept of average, random, and total velocities. The random, or "thermal," velocity can be naively associated with an empirically defined temperature, although the underlying velocity distribution need not be truly

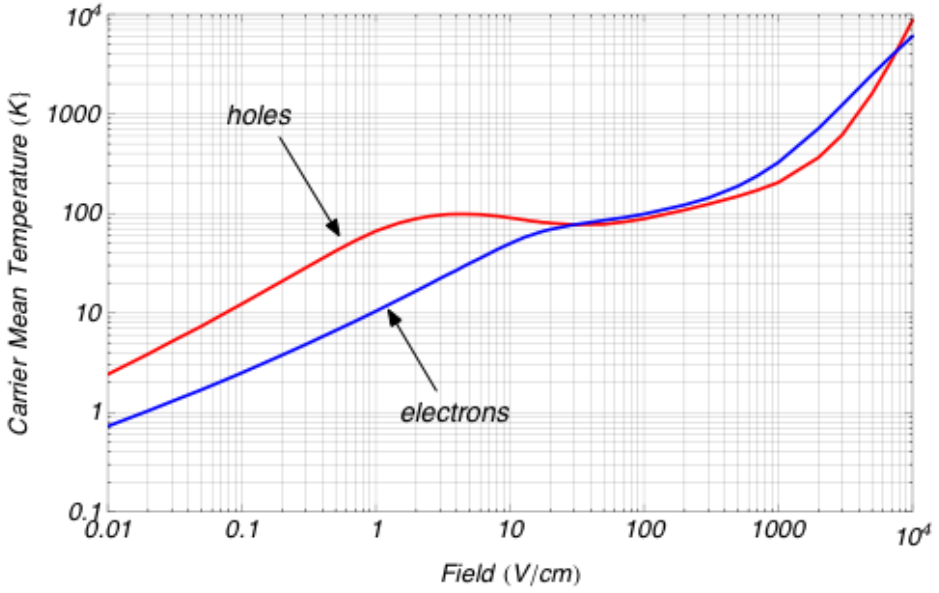

Figure 4.7: The effective (scalar) carrier temperature for electrons and holes, defined by the trace of the statistical temperature tensor.

Maxwellian. The random velocity was associated with elements of an energy tensor that allowed us to define corresponding components of a temperature tensor. These tensors represent energy held in *stochastic* fluctuations of the distribution under drift. It is informative to continue to explore this concept of energy held in directional degrees of freedom. To this end, in this section we will associate the property *diffusion* with this concept.

Diffusion is the stochastic motion of carriers undergoing random motions. From numerous statistical mechanics references, such as Reif [81], we can define a diffusion constant according to the mean square displacement of radomizing particles undergiong velocity fluctatations. Defining the mean square displacement as

$$R(t) = \left\langle |x(t) - x(t_0)|^2 \right\rangle \tag{4.15}$$

we can define a diffusion constant by

$$R(t) = 2Dt \tag{4.16}$$

In our simulation, we once again used the ergocity of our system to sample a large number of carriers in the steady state. After a particle was equilibriated to the field but still $\sim 10\%$ before the end of a simulation, we would record a time and position as a new origin. At the end of the trial, we would find the mean square displacement and time subtended from this new origin. Using ensemble averages, we establish a

3×3 diffusion matrix as the following.

$$\Sigma^2_{x,ij} = 2D_{ij} \langle \Delta t \rangle \tag{4.17}$$

or

$$\boxed{D_{ij} = \frac{\Sigma^2_{x,ij}}{2 \langle \Delta t \rangle}} \tag{4.18}$$

Here, $\Sigma^2_{x,ij}$ is the real-space displacement covariance taken from the sample origin, and $\langle \Delta t \rangle$ is the average time subtended since the sample origin. This is our primary definition of diffusion.

Now, we make an exercise out of equation 4.1. We can also consider diffusion elements along specific directional degrees of freedom as a random walk problem and recover the standard three-dimensional result [81],

$$R(t) = \left(\frac{2}{3} \left\langle v^2 \right\rangle \tau \right) t \tag{4.19}$$

which leads to the isotropic three-dimensional result,

$$D = \frac{1}{3} \left\langle v^2 \right\rangle \tau. \tag{4.20}$$

To make this appropriate for tensorial notation with individual directional degrees of freedom, we define the following making use once again of our velocity fluctuation covariance and average scatter time.

$$\text{velocity based: } D_{ij} = \Sigma^2_{v,ij} \langle \tau \rangle \qquad \text{(incorrect)} \tag{4.21}$$

We notice the comparison of diffusion definitions of 4.18 and 4.21 are plotted in figure 4.8. These two definitions of diffusion do NOT depict the same result. The difference is that equation 4.21 would serve as a better representation of position displacements if it were instead the expectation quantity,

$$\text{velocity-rate based: } D_{ij} = \left\langle \Sigma^2_{v,ij} \tau \right\rangle = \operatorname{cov} \left(\delta v_i \, , \tau_j \delta v_j \right). \tag{4.22}$$

So we consider that the velocity-based equation 4.21 is *not* a good definition of diffusion. The position-based formulation of equation 4.18 is correct. The problem is that our velocity-based definition suffers from a lack of correlations needed in the true definition of $D_{i,j} = \operatorname{cov}(\delta v_i, \delta x_j)$, (or equivalently, equation 4.22).

Something we can glean from these subtleties in diffusion and diffusion-like quantities is an understanding of the value and directionality of the diffusion-based relaxation time. The $\langle \tau \rangle$ we used in 4.21 was simply the mean net scattering rate from our statistical enembles. We may be curious how long velocity fluctuations characteristically last before again being randomized. By comparing the variance in real-space

displacement and in velocity, we can define a τ tensor. We could then inspect this as a function of direction. We equate the position-based and velocity-based diffusion elements, and solve for a $\langle\tau\rangle$-like relaxation tensor.

$$[\tau] = \left[\Sigma_v^2\right]^{-1} \cdot [D] \tag{4.23}$$

The diffusion tensor above, $[D]$, is the position-defined tensor of equation 4.18, and $[\sigma_v^2]^{-1}$ is the inverse matrix of the velocity covariance.

Let us add a few further words about diffusion as defined by simulation data. In figure 4.8, we have plotted the "axial" diffusion elements along the main diagonal. The fact that electrons typically occupy one of the independent and rather different L-valleys complicates their diffusive motion as well as their drift velocities. In an isotropic case, the diffusion component along the electric field (along the z-axis) we could easily identify as the "longitudinal" diffusion coefficient. Here, "longitudinal" might be a more appropriate term when aligned along the resulting average drift velocity vector rather than the field orientation. So we depict the "D_{zz}" component. We also plot the $D_{xx} = D_{yy}$ components, which are identical as they are symmetric to the field and inverse mass tensors, but which we likewise refrain from calling the "transverse" components. All the diffusion constants are smaller at higher fields, as the carrier gases are becoming "hotter," so the relaxation time is becoming shortened. Velocity fluctuations are not as long lived at higher fields. So our position-defined diffusion has resonant peaks at threshold fields where the onset of large (optical or intervalley) phonon emission starts to occur.

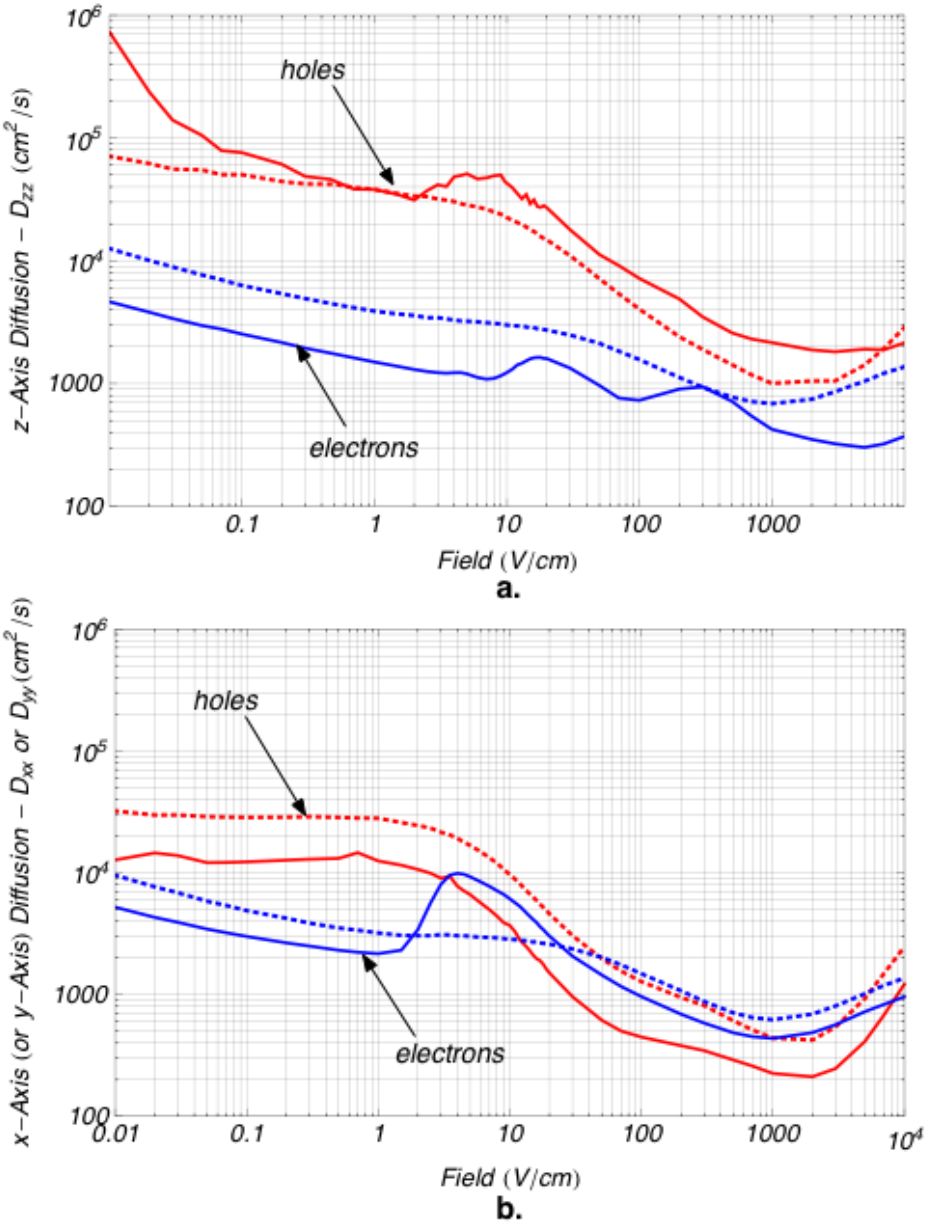

Figure 4.8: Diffusion for electrons and holes **a.** The z-axis diffusion element, D_{zz}, for both the position (*solid*) and velocity (*dashed*) definitions used for the diffusion tensor. **b.** The x-axis (y-axis) diffusion elements, D_{xx} (D_{yy}), with both position (*solid*) and velocity (*dashed*) definitions. Note that we consider the *position-based* definition of diffusion to be correct; our velocity-based definition is for illustrative purposes.

4.8 Relaxation times

When we explored diffusion, we noticed that we could define an expectation value for a "time" quantity related to distance and velocity fluctuations. It turns out there are several definitions for time, related to averaged ensemble quantities. Two important definitions are *momentum relaxation time* and *energy relaxation time*. We may find excellent textbook discussions of these quantities based on analytical underpinnings, such as in Lundstrom [45] and Ridley [5].

Empirically, formulations for relaxation times have also been proposed by Shur [82]. These expressions have had success in characterizing a number of semiconducting materials [83]. We can extract these relaxation times from our transport simulation.

We can define an empirical momentum relaxation time as the following.

$$\langle \tau_m \rangle = \frac{\langle m^* \rangle \langle v_d \rangle}{eF} \tag{4.24}$$

Here, $\langle m^* \rangle$ is the steady-state average mass, $\langle v_d \rangle$ is the steady-state drift velocity, e the electric charge, and F the field.

An empirical energy relaxation time is defined as the following.

$$\langle \tau_E \rangle = \frac{\langle E \rangle - E_0}{eF \langle v_d \rangle} \tag{4.25}$$

In this case, $\langle E \rangle$ is the mean carrier energy, and E_0 is the average energy if no field were applied. Typically, one would just assume that $E_0 = 3/2\ k_B T_L$, the lattice equilibrium temperature. In our case we completely neglect E_0, as it is insignificant for the range considered.

These relaxation times are related to the power added by the electric field to the carriers under drift. The average power added to drifting carriers is $P = e\,|\mathbf{F}|\,|\mathbf{v_d}| \cos\theta$, where θ is the angle between field and drift velocity. Considering that we have obliquely propagating electrons, we project the inverse mass tensor along z and take only the z-component of the drift velocity for these computations. We compute relaxation times this way for both electrons and holes.

We show results for relaxation times in figure 4.9. Note that the energy relaxation time is longer than the momentum relaxation time. This is because it typically takes several inelastic scatters for the energy to decay, whereas momentum relaxation only requires the randomization of the momentum's initial direction. As a ratio of the momentum relaxation time, the energy relaxation time is comparatively shorter when scattering is highly inelastic. This occurs, for instance, around the onset for optical phonon emission.

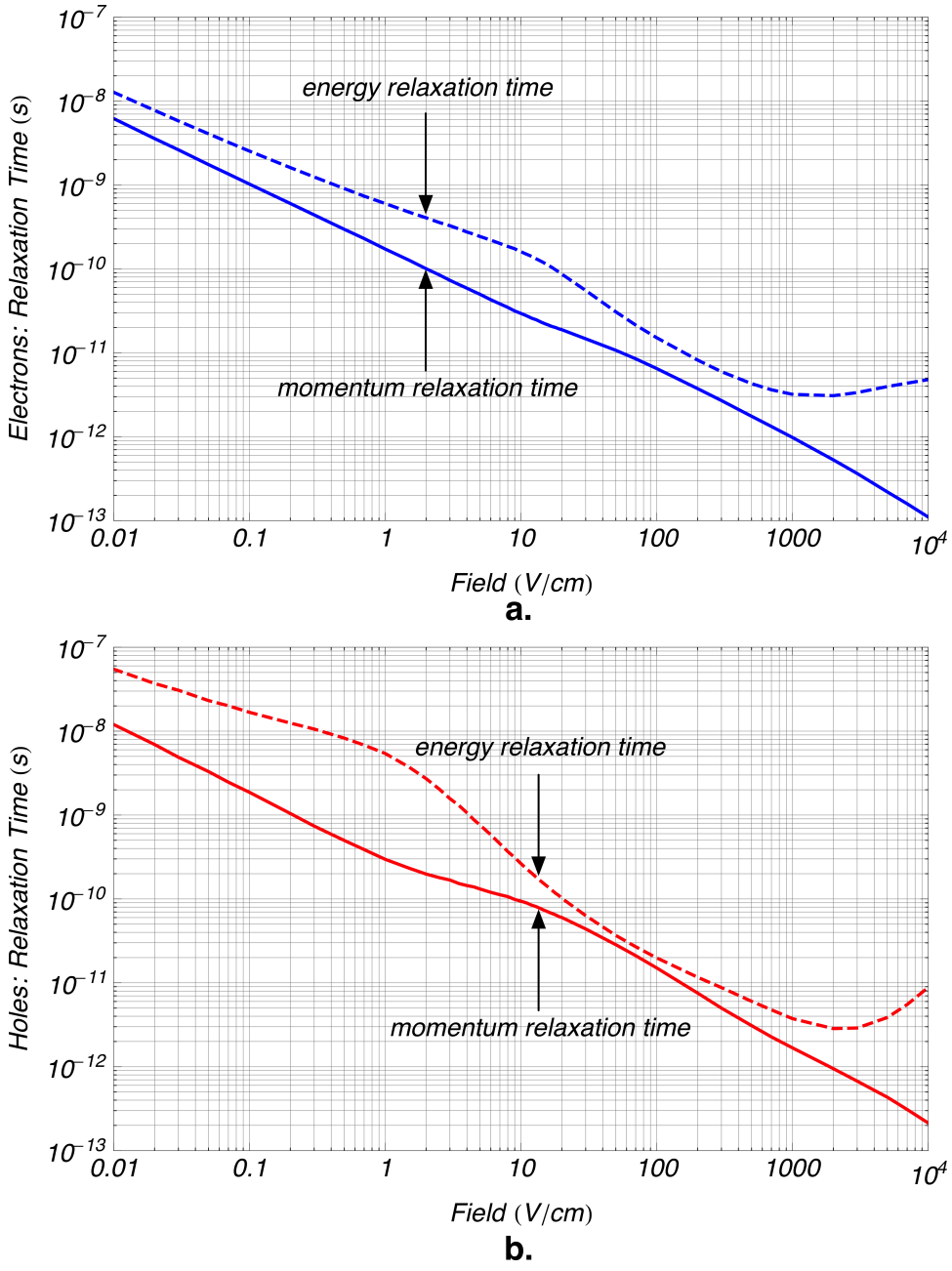

Figure 4.9: The relaxation times for electrons and holes. **a.** Energy and momentum relaxation times for electrons **b.** Energy and momentum relaxation times for electrons.

4.9 Carrier energy, and the Wannier relation

The most important steady-state quantity in describing carriers under our "hot" conditions is energy. We have defined a thermal energy tensor already. We know that there is an average drift velocity vector that may carry kinetic energy, as well. We define a *drift energy* tensor, as in Lundstrom [45]. We define this energy as one-half of the Kronecker product between the average ("drift") momentum and drift velocity.

$$W_{ij}^{drift} = \frac{\hbar}{2}\ k_{di}\ v_{dj} = \frac{\hbar^2}{2}\ k_{di}\ m_{ij}^{*-1}\ k_{dj} \tag{4.26}$$

Our purpose here to compare this drift energy to the thermal (or "diffusive") energy. In 1951, Wannier [84] used Boltzmann transport formalism to show that the expected energy of drifted, charged ions in gases follows a particular relation, later coined the "Wannier relation" as by Robson [85]. In appendix C, we verify that the Wannier relation returns the expected energy in the case where the functional form of the energy distribution is a displaced Maxwellian.

The Wannier relation states that the average energy is the combination of both the drift and diffusive/thermal terms.

$$< E >= \underbrace{\frac{3}{2}k_B T_C}_{\varepsilon_{diffusive}} + \underbrace{\frac{1}{2}m^* v_d^2}_{\varepsilon_{drift}} \tag{4.27}$$

For our case with full anisotropy, this is

$$< E >= \text{Trace}\left(\frac{1}{2}k_B T_{ij}\right) + \text{Trace}\left(\frac{\hbar}{2}\ k_{di}\ v_{dj}\right) \tag{4.28}$$

In figure 4.10, we plot the average energy for electrons and holes as a function of field. We also plot the contributions from the diffusive and drift energies. Even if our distributions may depart from the form of a displaced Maxwellian because they cross the inelastic threshold for optical (or intervalley) phonon emission, the Wannier relation remains a good approximation. Perhaps more importantly, we verify that the energy held in random velocity fluctuations greatly exceeds the energy held in the average drift velocity. This justifies arguments that the random velocity is quite close to the total velocity, or that the quantity defined as the carrier temperature is a good estimator for the mean energy. Yet another way to put this, is that the "noise temperature" is about equal to the "carrier temperature," as per Reggiani's terminology [74].

In figure 4.11, we also show the energy probability distribution functions for our simulated ensembles of electrons for a log sweep in $\langle 001 \rangle$-aligned electric fields. Note the scale spans many orders of magnitude, and that for a wide range of field amplitude the distribution centroid often occupies the $10 - 100$ meV scale where large phonons are emitted. In figure 4.12 we plot the energy probability distribution for electrons

and for holes at an applied field of 1 V/cm. We also superimpose the respective displaced Maxwellian distribution functions for the empirically determined drift and diffusion energy.

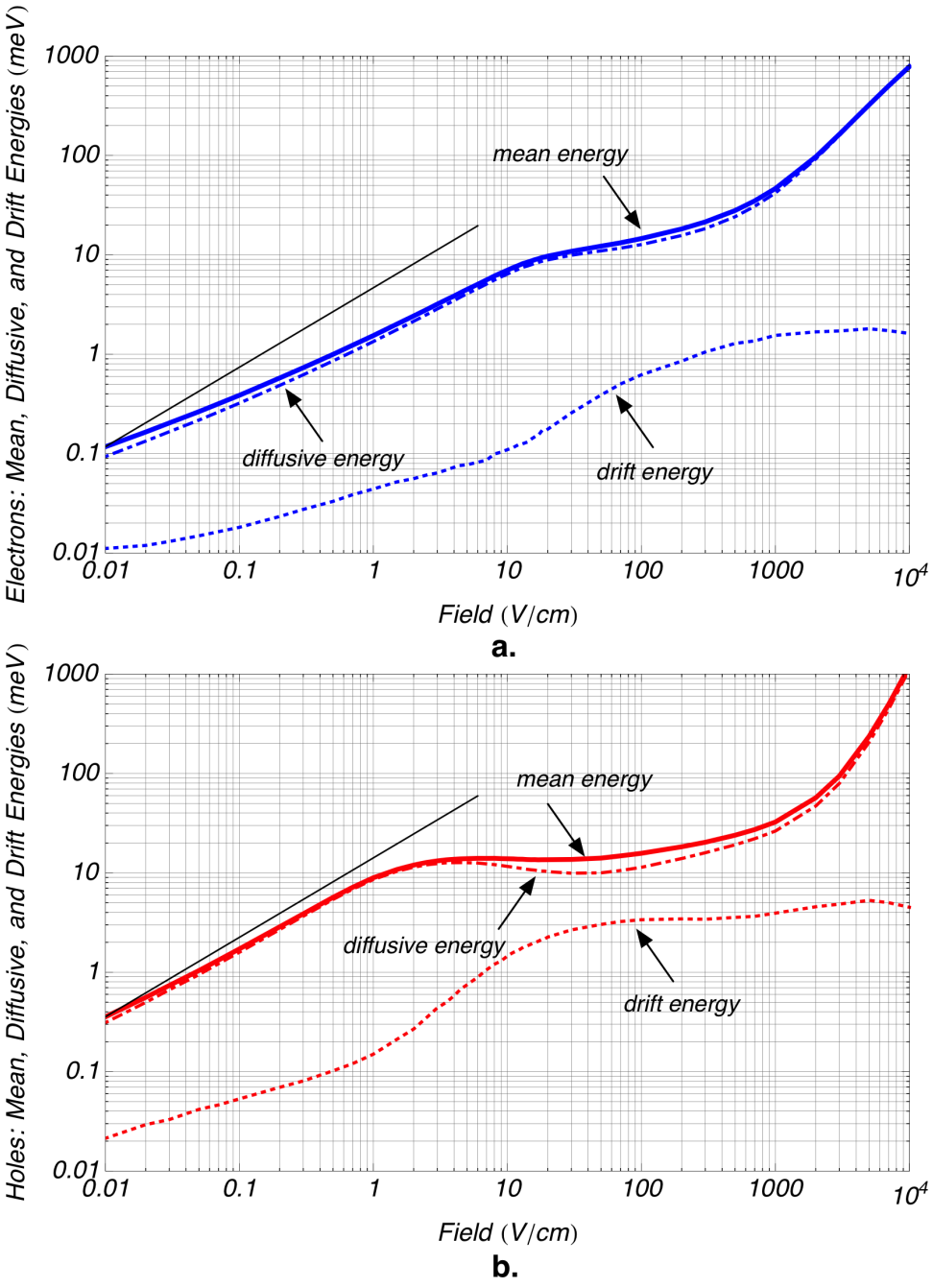

Figure 4.10: Carrier energies for electrons and holes, validating the empirical use of the Wannier relation. Clearly, the thermal ("diffusive") energy dominates the mean energy. The thin black traces are trend lines that intersect the mean energy at $F = 0.01\ V/cm$ and follow a power law of $F^{4/5}$ as per appendix B. The anisotropy of the electrons likely accounts for their larger discrepancy from this trend.

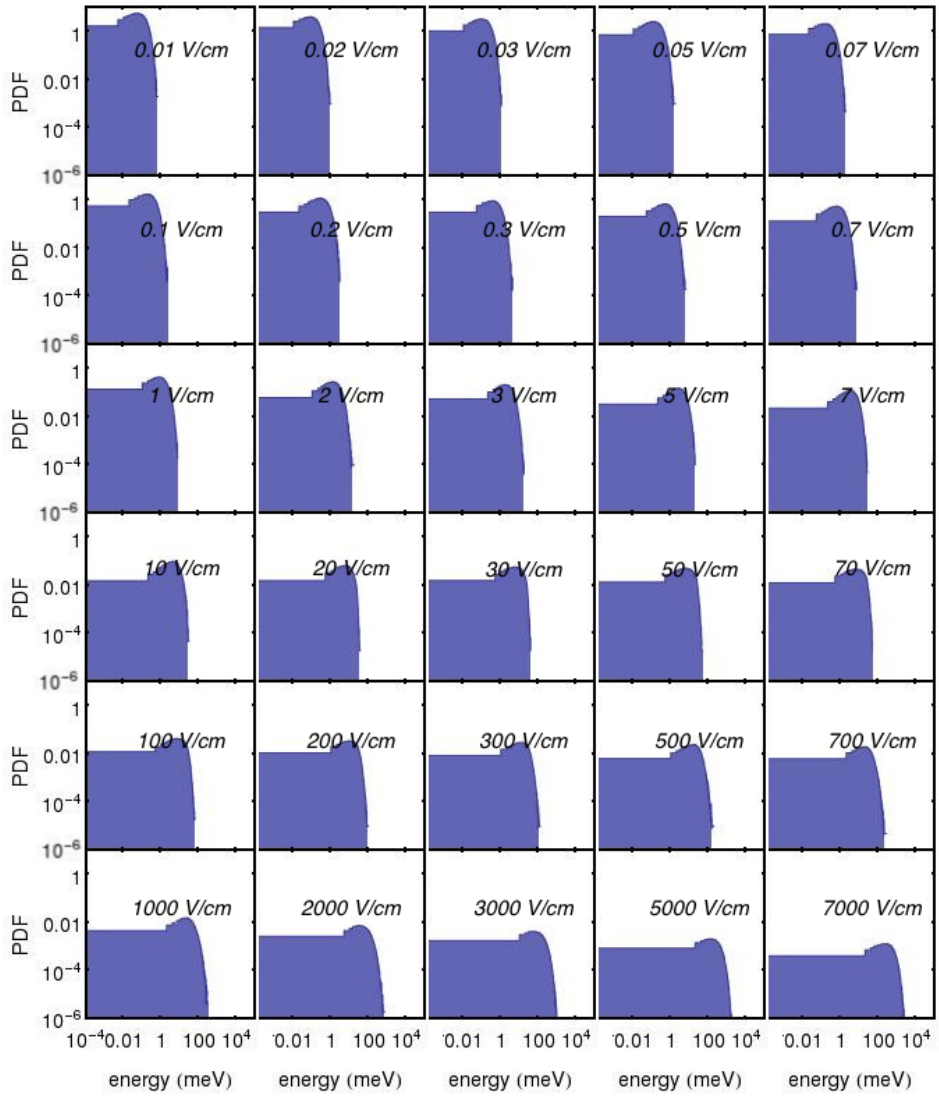

Figure 4.11: Log-log representations of the energy probability distribution function for electrons, across a log-sweep of applied fields.

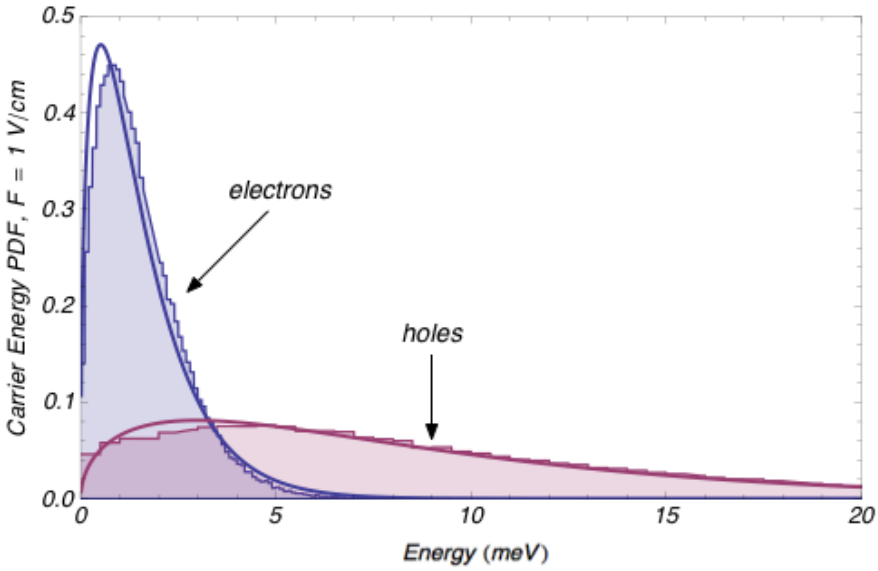

Figure 4.12: For a given field of 1 V/cm, we juxtapose the energy probability distribution for electrons with that for holes. This is to emphasize that holes carry substantially more carrier energy in the steady state. The traces overlain are displaced Maxwellian distributions using isotropic effective masses, in addition to statistically deduced drift velocities and effective carrier temperatures.

4.10 Luke-Neganov phonons

So far, we have reported on properties of electrons and holes under drift, with an emphasis on the energy distribution functions as a function of electric field. Here we examine the reciprocating distributions of emitted phonons, which are responsible for of the field-dependent structure of the quantities we have covered.

Knowledge of emitted phonons is important for understanding the phonon signal response in our CDMS dark matter detectors. Phonons created during the initial nuclear impact and electron/hole production represent the primary calorimetry signal. The energy of these phonons is dominated by optical phonon production. The collection and measurement of these phonons by our CDMS transition-edge sensors creates a unique measurement of the particle event. As charged particles or photons interact electromagnetically, their interactions in germanium at nuclear energy scales are strongly coupled to the electronic structure. Per unit deposited energy, the interactions initiated by charged particles (and photons) interact electromagnetically and ionize more electrons and holes than do the the nuclear recoils initiated by neutral particle interactions. Neutral particles include potential dark matter candidates. Therefore, the measured ratio between charge and phonon signals provides an excellent handle for background event discrimination. CDMS detectors further have added discrimination capability inherent in measuring arrival time physics of the phonon signals.

We attempt to understand and take advantage of phonon phenomena inherent in our cold, intrinsic germanium detectors. Aside from the primary phonons of the initial particle interaction, it is also important to understand the phonon spectra and production rates elsewhere in the crystal. Relaxation of carriers to the Fermi level at the interface has been evidenced in our detectors to produce relaxation phonons [38]. The dominant contribution to the phonon signal, however, are the phonons produced under drift. These phonons are the aforementioned Luke-Neganov ("Luke") phonons. At our low bias fields, Luke phonons have always been assumed to be acoustic phonons. These simulations suggest this may not always be the case. In particular, the Luke phonon spectra from holes are substantially higher in energy than from electrons, owing to the low hole-lattice interaction and high carrier energy of holes. Optical Luke phonons would mimic the optical phonons of the primary signal, and likely affect our timing measurement of the non-equilibrium phonon arrival physics at the contacts. The phonon effects suggested in these simulations will continue to be studied.

We have simulated acoustic phonon production to LA and a sum TA mode for anisotropic electrons. Electrons also experience intervalley coupling to high-q phonons at the Brillouin zone boundary. The slow TA branch is low enough to be treated separately at the Brillouin zone boundary. Otherwise, intervalley phonons are treated as a single degenerate branch, lumping the contributions of acoustic and optical branches. Electrons can also couple to intravalley optical phonons with $q \approx 0$. Holes have cou-

pling to three acoustic and three phonon braches, including both intra- and inter-band transitions between all three bands.

In the following plots, we present results for steady-state phonon: emission rates, mean power dissipation, average acoustic wavevectors, and some LA and TA PDFs.

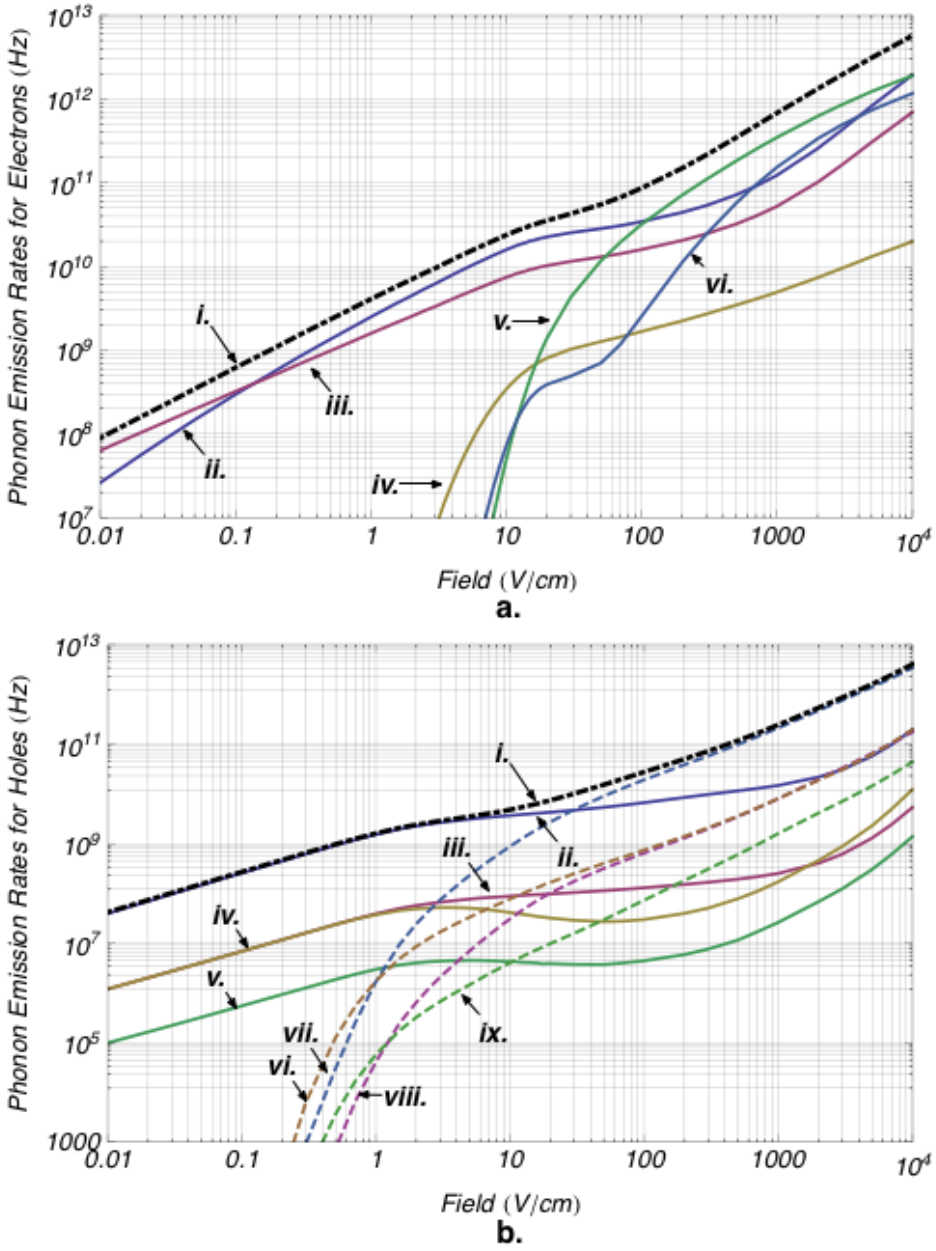

Figure 4.13: Rates of phonon emission **a.** For electrons, **i.**total rate, **ii.** LA, **iii.** TA, **iv.** Slow TA intervalley, **v.** intervalley, **vi.** optical phonons. For plot **b.**, holes, we have **i.**total rate, **ii.** all H-to-H acoustic phonons, **iii.** all H-to-L acoustic phonons, **iv.** L-to-H acoustic, **v.** L-to-L acoustic, **vi.** L-to-H optical, **vii.** H-to-H optical, **viii.** H-to-L optical, **ix.** L-to-L optical.

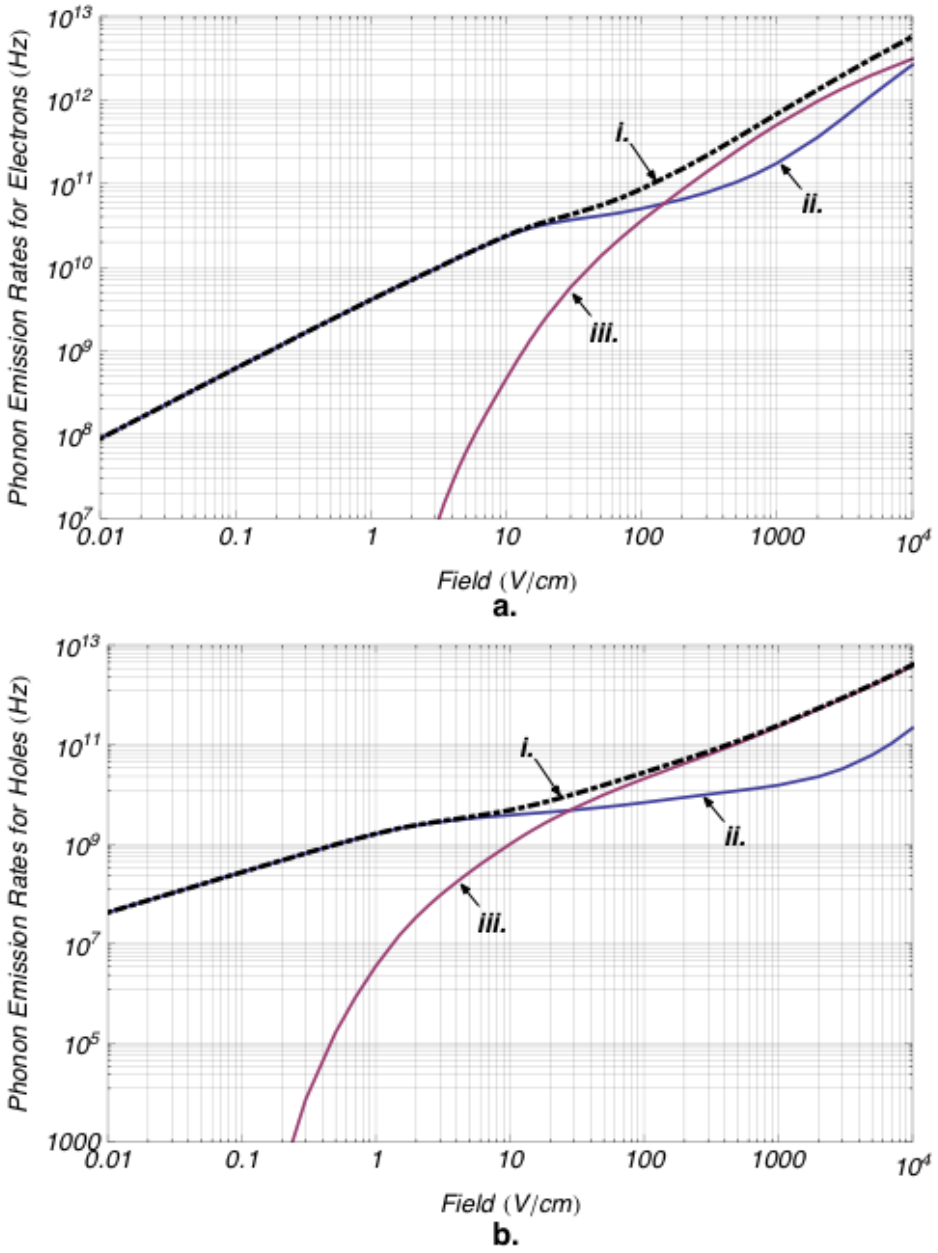

Figure 4.14: Phonon emission rates for electrons and holes, grouped into acoustic and into optical (and intervalley) phonons. **a.** For electrons, **i.** total emission rate, **ii.** acoustic emission rate, **iii.** optical and intervalley emission rate. Plot **b.**, for holes; **i.** total emission rate, **ii.** acoustic emission rate, **iii.** optical emission rate. For electrons, the cross-over field where the net acoustic rate equals the net optical rate is $F = 147\ V/cm$. The rates cross-over field for holes is at $F = 27.9\ V/cm$.

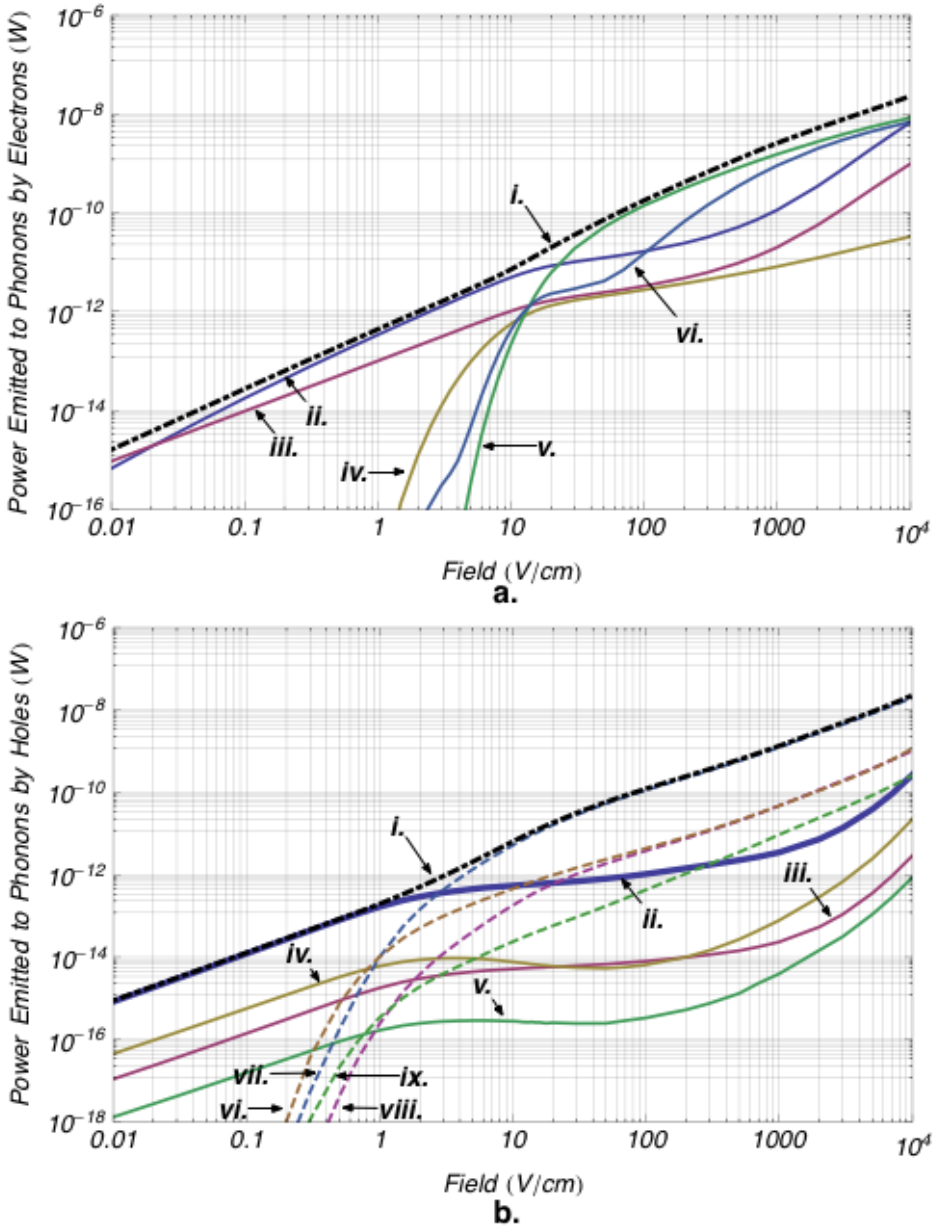

Figure 4.15: Emitted phonon power **a.** For electrons, **i.**total, **ii.** LA, **iii.** TA, **iv.** Slow TA intervalley, **v.** intervalley, **vi.** optical phonons. For plot **b.**, holes, we have **i.**total, **ii.** all H-to-H acoustic phonons, **iii.** all H-to-L acoustic phonons, **iv.** L-to-H acoustic, **v.** L-to-L acoustic, **vi.** L-to-H optical, **vii.** H-to-H optical, **viii.** H-to-L optical, **ix.** L-to-L optical.

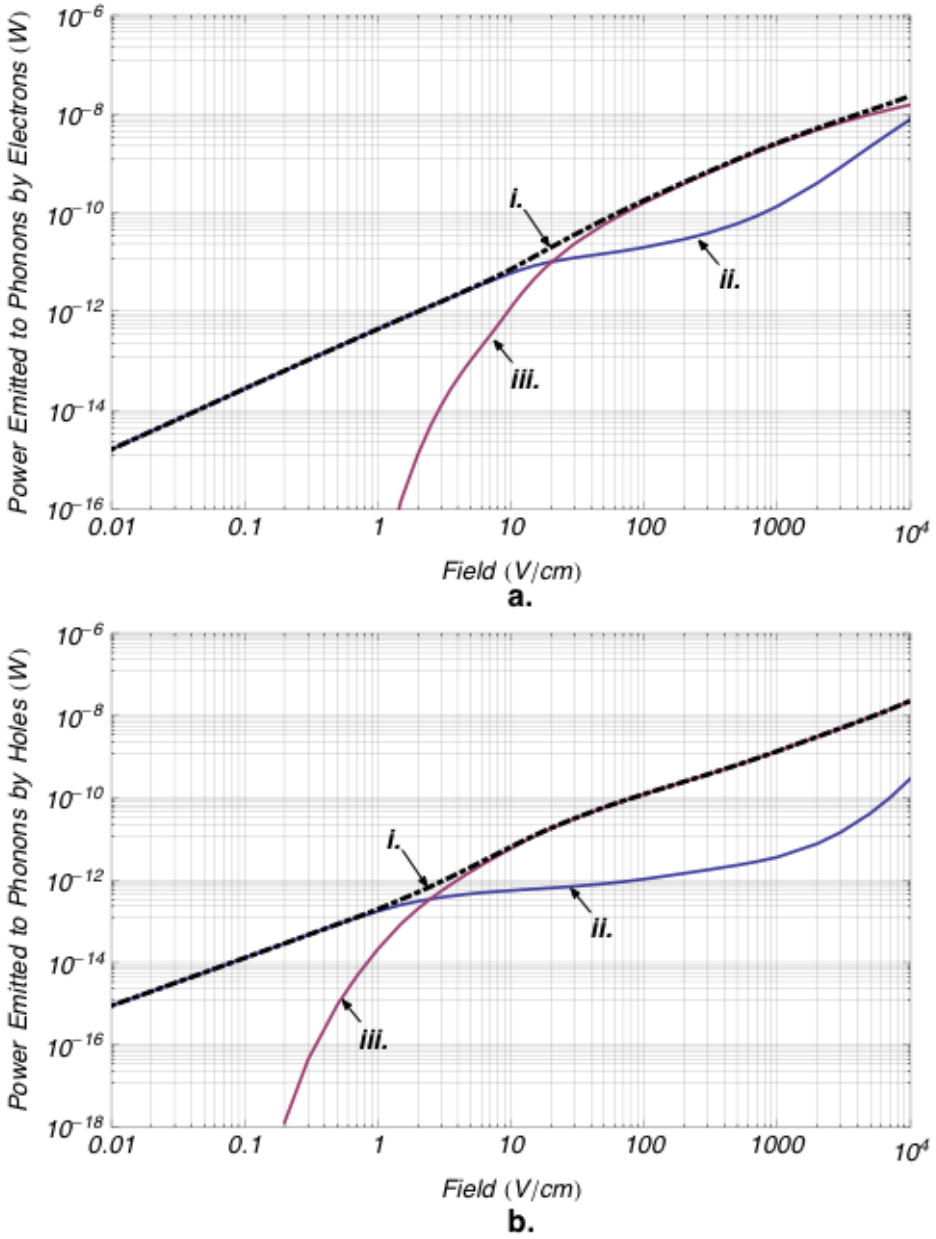

Figure 4.16: The mean power emitted to phonons for electrons and holes, grouped into acoustic and into optical (and intervalley) phonons. **a.** For electrons, **i.** total emitted power, **ii.** acoustic emitted power, **iii.** optical and intervalley emitted power. Plot **b.**, for holes; **i.** total emitted power, **ii.** acoustic emitted power, **iii.** optical emitted power. For electrons, the cross-over field where the net acoustic power equals the net optical power is $F = 20.6\ V/cm$. It is $F = 2.48\ V/cm$ for holes.

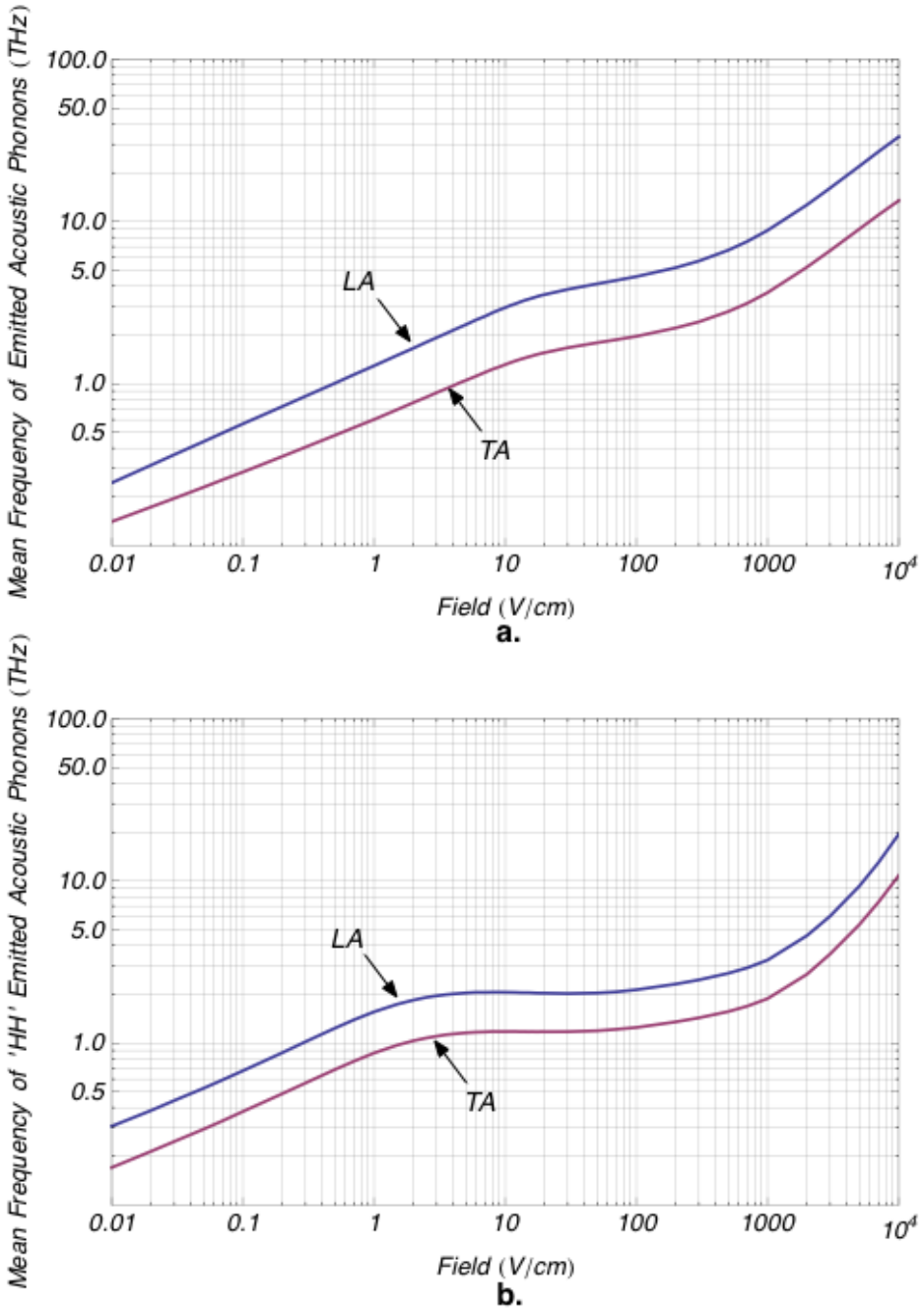

Figure 4.17: The mean frequency emitted to acoustic phonons. **a.** For electrons, LA and grouped TA modes. **b.** For holes. LA and TA modes.

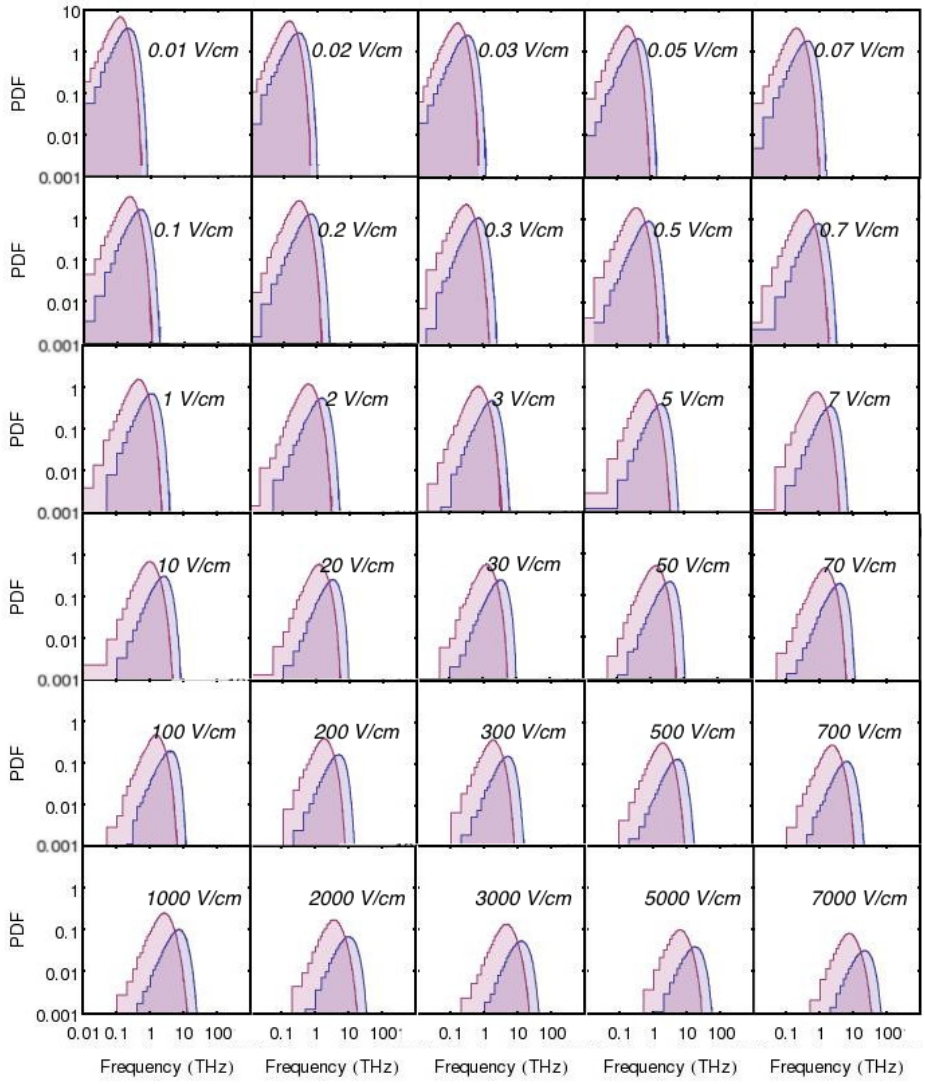

Figure 4.18: Log-log representations of the frequency probability distribution function for acoustic phonons emitted by electrons, across a log-sweep of applied fields. Blue are LA phonons, Magenta are TA phonons.

4.11 Electrons with field applied in the $\langle 111 \rangle$ direction

The general phenomenon behind the anisotropy of electrons in germanium has been known for some time. Without the randomizing influence of high lattice temperatures or high fields, anisotropy effects become most apparent under low-temperature and low-bias conditions such as ours.

The most striking evidence for germanium anisotropy may be the phenomenon of a negative differential mobility relation in for samples in the $\langle 111 \rangle$ orientation. This effect only appears at low temperatures, perhaps at $T = 20\ K$ and below [13]. Bias under the $\langle 111 \rangle$ orientation is such that it introduce a large assymetry between L-valleys in the angle subtended between the applied field direction and the principle axes of the valleys. In the case of $\langle 111 \rangle$, there will be one L-valley that is aligned with its principle axis parallel to the electric field, where the tensor mass is completely longitudinal and quite heavy. The other three L-valley ellipsoids intersect the field vector closer to their semi-minor axes, which is the transverse, or "light," orientation.

We can link drift velocity to the inverse mass tensor as the following.

$$\begin{aligned} \mathbf{v_d}(F) &= [\mu(F)] \cdot \mathbf{F} \\ &= e[\tau_m(F) m_{eff}^{*-1}(F)] \cdot \mathbf{F} \end{aligned} \tag{4.29}$$

Consider that the average drift velocity is weighted by all populations of electrons occupying any of the possible conduction minima. We limit the scope in these arguments to only the four full L-valleys. In equation 4.29, the thought is that we need to average the inverse mass tensors by weighted occupation. The main assumption here is that the momentum relaxation time is a simple isotropic scalar, and also not particularly different between valleys. If these assumptions were true, then

$$v_d \propto \left[m_{eff}^{*-1}(F) \right] \tag{4.30}$$

and this effective mass tensor would be weighted by the relative occupation, 0 to 1, between all four L-valleys. In other words,

$$\left[m_{eff}^{*-1}(F) \right] = \sum_{i=1}^{4} N_i(F) \left[m^{*-1} \right]_i \tag{4.31}$$

where the occupation distributed between only L-valleys,

$$\sum_{i=1}^{4} N_i = 1. \tag{4.32}$$

It is a worthwhile aside to examine the inverse mass tensors. The valleys are numbered counterclockwise, looking down in the $-z$ direction, starting with valley $(+1, +1, +1)$.

$$
\begin{aligned}
\left[m_1^{*-1}\right] &= \tfrac{1}{3}\begin{pmatrix} \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) \\ \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) \\ \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) \end{pmatrix} \\
\left[m_2^{*-1}\right] &= \tfrac{1}{3}\begin{pmatrix} \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) \\ \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) \\ \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) \end{pmatrix} \\
\left[m_3^{*-1}\right] &= \tfrac{1}{3}\begin{pmatrix} \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) \\ \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) \\ \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) \end{pmatrix} \\
\left[m_4^{*-1}\right] &= \tfrac{1}{3}\begin{pmatrix} \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{1}{m_l}-\frac{1}{m_t}\right) \\ \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) \\ \left(\frac{1}{m_l}-\frac{1}{m_t}\right) & \left(\frac{1}{m_t}-\frac{1}{m_l}\right) & \left(\frac{2}{m_t}+\frac{1}{m_l}\right) \end{pmatrix}
\end{aligned}
\tag{4.33}
$$

Under the assumption of an equal population between the L-valleys, equation 4.31 results in an isotropic effective mass – the conductivity mass! – due to symmetry.

$$
\left[m_{eff}^{*-1}\right] = {}^{1}\!/_{m_c}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{iff} \quad \left(N_1 = N_2 = N_3 = N_4 = {}^{1}\!/_{4}\right) \tag{4.34}
$$

The averaged effective mass is isotropic if we have an equal distribution between valleys. If the other assumptions are correct, then drift velocity would be the same on average, regardless of the field orientation. Anisotropy effects enter when we disrupt this uniform distribution.

When the field is aligned along the $\langle 111 \rangle$ direction, as function of increasing field amplitude, electrons in the three "light" valleys can reach higher energies. They will reach the threshold for intervalley phonon emission where the heavy valley will be energetically disfavored. When electrons randomize between valleys, the "heavy" valley acts like a cold trap. Eventually, the light electrons randomly transferring between valleys will eventually become trapped in the heavy valley. The heavy valley's population will grow to near unity. This effect is called *repopulation*. It is well evidenced in literature on germanium [86, 87, 88, 89, 13].

What we find in simulation is surprising. Recall in the ⟨001⟩ case, pre-existing literature values supplied an excellent match to recent $T = 50\ mK$ velocity measurements. Also, according to equation 4.30, the mass tensor should determine the drift velocity. If occupation is uniform, the net averaged mass should be symmetric. Therefore, these assumptions would say that at lowest fields, drift velocities in any orientation should match those measured and simulated for the ⟨001⟩ case. However, simulation for ⟨111⟩ at the lowest fields actually shows a drift velocity substantially lower (factor ≈ 4) than in the ⟨001⟩ case. Yet, the population distributions are uniform. It seems the "light" mass valleys do not appear to be occupying particularly light wavevectors on average. By taking the ensemble mass for each valley as

$$\langle m \rangle = \left\langle \left(\frac{1}{k^2} \right) \mathbf{k}^T \cdot \left[m^{*-1} \right] \cdot \mathbf{k} \right\rangle^{-1} \tag{4.35}$$

we find that the "light" valley electrons are demonstrating quite nearly the same expected mass as do electrons in the symmetric valleys in the ⟨001⟩ case. The "light" valley drift velocity comes out even less than that of the ⟨001⟩ case. Meanwhile, the heavy valley acts to pull down even this lowered average. Something unforeseen is occurring in momentum relaxation rate. Within the simulation code, each valley is rotated into the frame of the principle axis while random sampling the possibility for a vector-based scattering event. The crystal orientation only matters when implementing propagation in the electric field. The rotations in the code have been checked, so it remains that this could be a physical phenomenon. It does stand to reason that wavevectors representing "heavy" mass states are preferred as final states during a scatter due to the higher density of states in those directions. A carrier already oriented along a light direction has all the more heavy final states with which to scatter.

Furthermore, the drift velocities simulated at high fields after the negative differential velocity region are also low compared to $T = 8\ K$ data. However, there is an experimental dependence on charged impurities with this region of the velocity curve as described in the original reference [13]. We do not account for impurity scattering in this simulation, so it may explain low velocity in this region.

Concluding our results for ⟨111⟩ electrons, we predict a lower velocity than has yet been observed in the ⟨111⟩ orientation for fields both below the onset of repopulation, and at the maximum repopulated field point ("the dip"). Perhaps this difference is due to the absence of impurity scattering, and/or a scattering-based anisotropy resultant from our anisotropic (vector-based) Monte Carlo implementation.

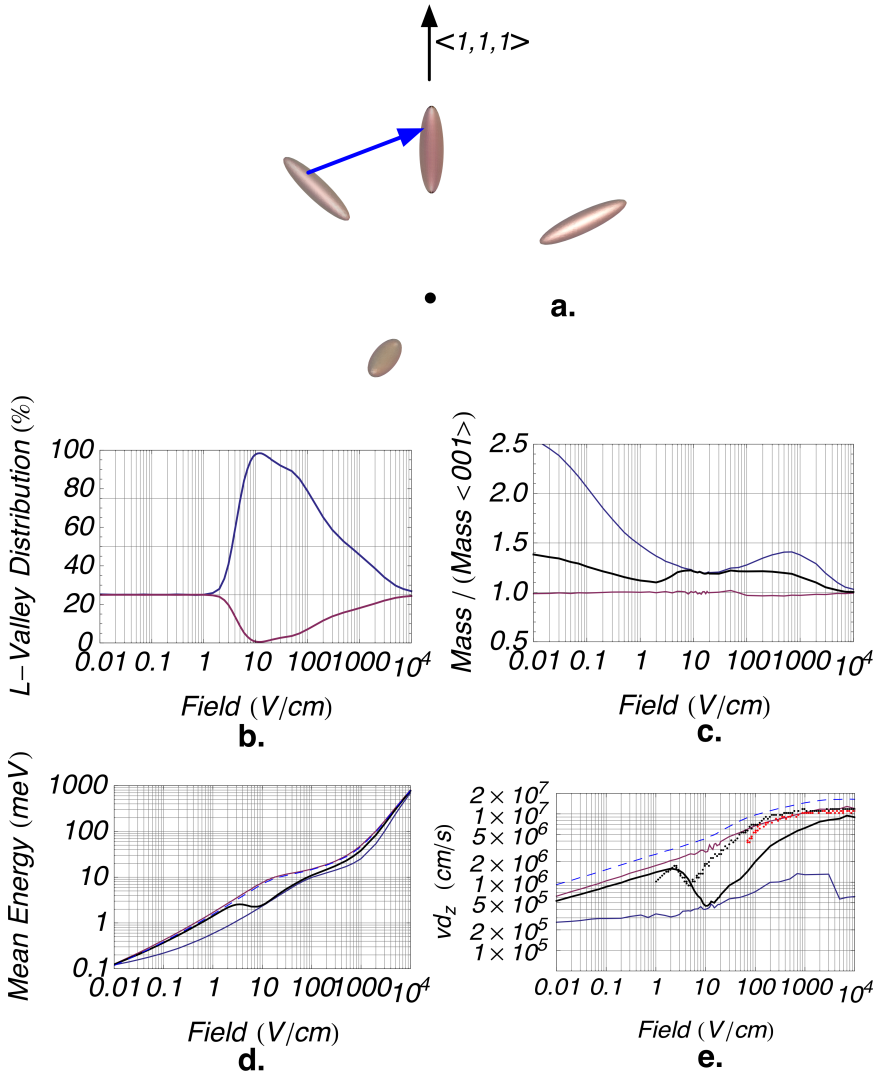

Figure 4.19: Electrons propagating with field applied in the ⟨111⟩ direction. **a.** The L-valley ellipsoids, with the heavy-mass ellipsoid oriented along ⟨111⟩. A vector represents a random intervalley transfer into the heavy ellipsoid. **b.** The percentage distribution of the electrons occupying the L-valleys as a function of electric field. The heavy valley (*blue*) has maximum occupation near 10 V/cm. **c.** The mean value of effective mass, normalized to the baseline case with ⟨001⟩ field applied. *Black* is net average, *blue* is the heavy ellipsoid. The "light" valleys (*magenta*) do not appear to favor a particularly light mass.**d.** Mean energy, net average energy transfers between light-favored to heavy-favored energy during population inversion. *Blue, dashed* is the case with field aligned ⟨001⟩. **e.** Drift velocities, net and separate valleys. *Blue-dashed:* ⟨001⟩ case. *Black dots:* $T = 8K$ data, $N = 3 \times 10^{13}/cc$. *Red dots:* $T = 8K$ data, $N = 2 \times 10^{12}/cc$, $from$[13].

4.12 Concluding remarks regarding our simulation work

In this chapter, we have presented and interpreted the results of our Monte Carlo simulation studies. The implementation of many details regarding scattering and the correct conservation of energy and momentum have resulted in a working simulation code, and simulated particle ensembles which have produced measurable predictions. In concert with our lab's efforts to produce milliKelvin drift velocity measurements, we have been able to validate the underlying kinematics of electrons and holes in our detectors. This automatically affords us some predictions for phonon emission properties.

In the rest of this thesis, we will apply the understanding presented in this chapter to make further predictions about detector phenomenology. Specifically, we will focus on the energy distributions introduced in this chapter to predict energy-dependent processes possible to electrons and holes. In the next chapter, we introduce possibilities for generation and recombination mechanisms.

Chapter 5

Capture Processes in CDMS Detectors

We now turn our focus to understanding carrier capture processes in our detectors. Capture (and possibly generation) processes represent the dominant cause of systematic effects in our detectors and in the CDMS experiment. This is because such processes represent a rate of change of the space charge density of the bulk, resulting in a change in internal field. Here, we make quantitative predictions for some capture processes we believe dominate under CDMS conditions. This includes their energy and field dependence. Whereas our previous understanding of these effects and their role in space-charge evolution has been completely empirical, we can now utilize the results presented here within higher-level computational simulations for a more substantive understanding of space charge.

In previous chapters, we have determined and validated charge carrier dynamics and distribution functions. An understanding of drift and diffusion dynamics is in itself important and useful. Yet with knowledge of energy distributions of electrons and holes in an electric field, we can make predictions for experiment-limiting processes which we know are extremely energy dependent. By taking expectation values with our numerical distributions, are able to integrate away the energy dependence in the capture and ionization quantities at hand. This lets us make predictions for capture rates expressed solely as a function of electric field, which are extremely useful in any macro-level simulation of space charge evolution and detector performance.

As an overview, in this chapter we cover the following topics.

Introduce CDMS capture data: We note that both electrons and hole data show two, competing capture processes. These capture processes follow distinctly different power laws.

Bound-state neutral capture, forming D^-/A^+ states: We see that the higher-field capture rate is consistent with the over-charging of neutral hydrogenic

states into D^-/A^+ states, which are stable at the CDMS operational temperature of $T_L = 40\ mK$.

Review the concept of capture by phonon cascade: The remaining capture process in our data is likely due to the emission of low-energy acoustic phonons which de-excite a carrier into an impurity site. We review the established knowledge of this cascade capture process from Lax based on Thomson's model of capture in gases. We then point out the extensive work done by Abakumov.

Review possible attractive potentials: We review the attractive potentials which may allow a cascade capture mechanism for carriers on impurities/defects.

Review possible impurities/defects: We review the possible impurities/defects that might be considered, for possible capture of carriers.

Consider Coulomb capture, near the low-T limit: We revisit Abakumov's work for *equilibrium* carrier capture on **charged** sites for the low temperature limit ($kT_L \ll mv_s^2$). We then find the appropriate value for the *equilibrium* capture cross section at the CDMS value of $T_L = 40\ mK$.

Consider Coulomb capture out of equilibrium: With the *equilibrium* cross section in hand, we can then utilize Abakumov's hot-carrier *correction*. This now depicts Coulomb capture in the *high* carrier temperature, non-equilibrium limit ($kT_C \gg mv_s^2$).

Conclude CDMS capture data: Finally, with our non-equilibrium Coulomb capture cross section in hand, we can compare to data. We explain both capture processes, for both carriers. We suggest an interpretation for space charge evolution using a 2-carrier, 2-impurity balance of rates.

5.1 Comparing to data from CDMS detectors: inverse capture length

For many years, our group has been able to profile charge collection efficiency for electrons and holes as a function of applied bias [38, 3]. On one face of a detector, a collimated radioactive source is incident. After grounding for a long period of time to ensure a well-neutralized state, a voltage bias is applied and pulse data is acquired. After adequately high statistics are acquired, the detector is once again grounded. A neutralization cycle ensues before repeating the procedure again at another bias value. Plotting the resulting amplitudes as a function of applied voltage maps the collection efficiency. These plots have earned the nickname "seagull" plots for their appearance on a linear plot, as in figure 5.1. It is clear that electrons and holes recombine over a drift length in germanium which becomes asymptotically larger

with applied field. This provides a measurement of the field dependence of the net electron or hole *capture length* (or possibly *avalanche length*, to be general).

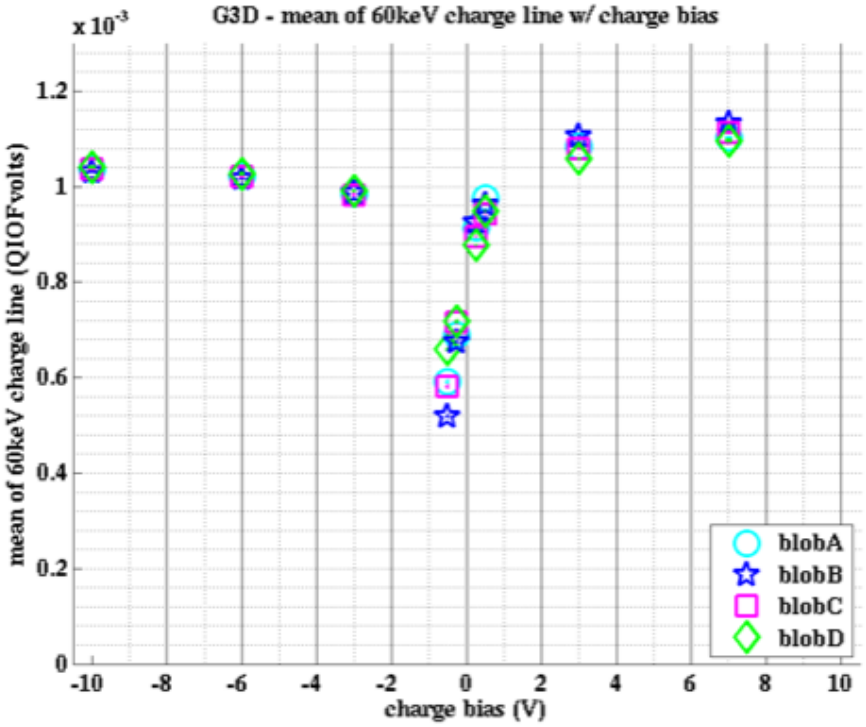

Figure 5.1: A "seagull" plot, showing the collected ionization of 60 keV ^{241}Am photons, incident from different collimated sources corresponding to "blobs" A-D. These sources were incident on the "charge side" of a 1-inch CDMS detector "G3D." Such measurements of ionization collection as a function of applied voltage affords a mapping between net carrier capture length and applied electric field. From C. Bailey [3].

Measurements in Tom Shutt's thesis [38], supply a good data set of measured capture lengths. He used the ionization response of 60 keV photons from a ^{241}Am source incident on a 0.95 cm p-type germanium $\langle 100 \rangle$ crystal. The temperature was $T = 20\ mK$. The substrate had a p-type shallow impurity concentration, measured to be in the range $0.6 \times 10^{11} cm^{-3} < |N_A - N_D| < 2.2 \times 10^{10} cm^{-3}$. This data set we will call "Shutt 1993."

More recently, Arran Phipps in our Berkeley group was able to utilize fiber optic pulses incident on a detector at $T = 50\ mK$ [11]. Using a custom, fast amplifier with a synchronized trigger, he was able to find capture lengths by measuring well-averaged time-domain dynamics of a laser-pulsed carrier response. A data set which we will call "Phipps 2011" was taken on a 1-inch $\langle 100 \rangle$ n-type germanium detector with a shallow impurity concentration of $|N_A - N_D| = 1.3 \times 10^{10} cm^{-3}$.

To map the macroscopic behavior of capture (and generation) processes, it is necessary to supply a few definitions.

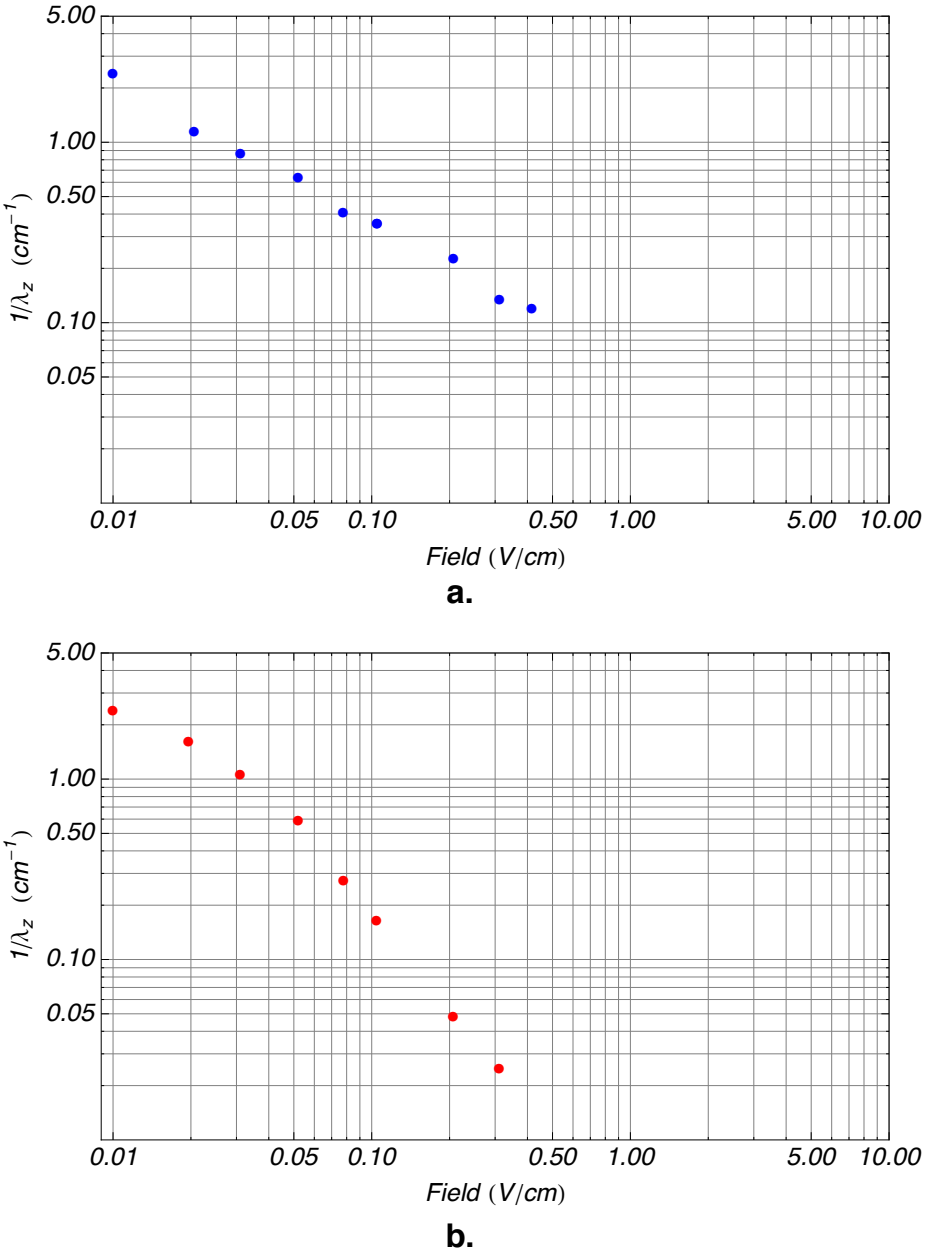

Figure 5.2: Carrier capture data in CDMS detectors, from the "Shutt 1993" data set. *a.* The inverse capture length for *electrons* with field. *b.* for *holes*.

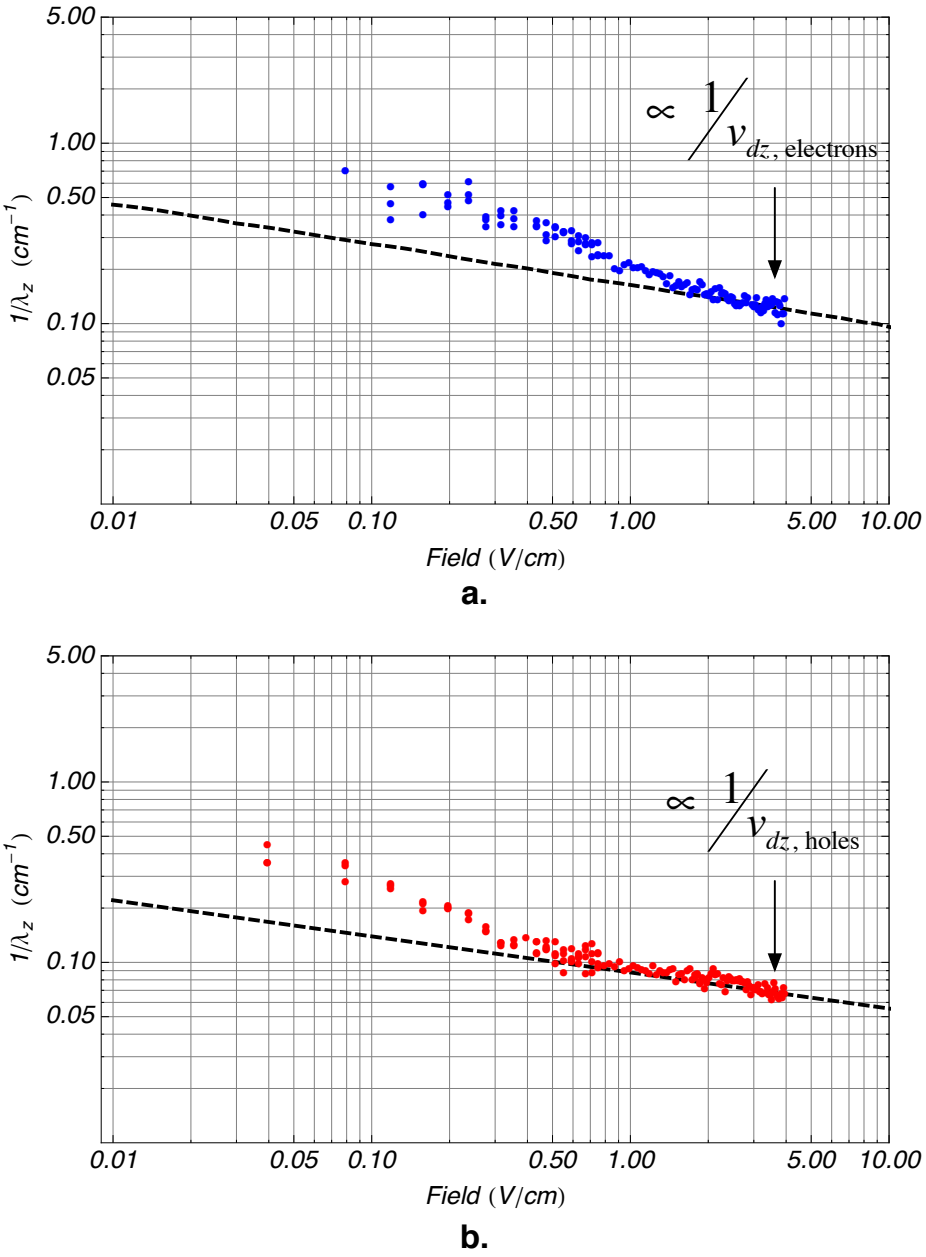

Figure 5.3: Carrier capture data in CDMS detectors, from the "Phipps 2011" data set. Note the 20 points at highest field in both the electron and hole data sets were fitted for a power law. These high-field power-laws are consistent with a field dependence proportional to the inverse drift velocity. *a.* The inverse capture length for *electrons* with field. *b.* for *holes*. Note also the clearly pronounced transition in power-law response for the holes in this data.

5.1.1 Capture rates, cross sections, and inverse lengths

Capture rates, cross sections, and inverse lengths are related quantities. These quantities are based on ensemble means, and their utility is to represent predictable and measurable quantities in experiment. The mean deficit (or possible increase) of an initial charge carrier population reveals information about recombination (and ionization) processes occurring as a function of field.

Considering an energy distribution and random velocity distributions, $g(E)$ and $v_{tot}(E)$, we may define a total *rate per density* for a capture process. The function $g(E)$ includes the energy density of states.

$$C = \frac{1}{N\langle\tau\rangle} = \langle v_{tot}(E)\sigma(E)\rangle = \frac{\int v_{tot}(E)\sigma(E)g(E)dE}{\int g(E)dE} \tag{5.1}$$

This is the *total* cross section of this quantity over the typical carrier distribution functions. The "target" density is depicted by concentration, N. These rates per density, defined by several variants of "C," will be useful in solving drift-diffusion equations when recombination terms are included.

To preserve the standard definition of a total capture cross section,

$$C = \langle v_{tot}\rangle\langle\sigma\rangle, \tag{5.2}$$

we define the total cross section as

$$\langle\sigma\rangle = \frac{\langle v_{tot}\sigma(E)\rangle}{\langle v_{tot}\rangle} \tag{5.3}$$

To compare to experimental CDMS data in particular, the macro-scale quantities useful here are the *inverse capture lengths*, which may be defined as the following.

$$\frac{1}{\lambda_z} = \frac{C}{\langle v_z\rangle}N = \frac{1}{\langle v_z\rangle\tau_c} \tag{5.4}$$

Above, $\langle v_z\rangle = v_{dz}$ is the mean velocity in the direction of drift assumed to be along z. The pertinent capture time is denoted by τ_c.

5.1.2 Multiple contributions to the capture length

The inverse capture length is given by the sum of all capture (and generation) processes. This can also be expressed by capture rates or cross sections.

$$\begin{aligned}\tfrac{1}{\lambda_z} &= \sum_i \tfrac{1}{\lambda_{z,i}} = \sum_j \tfrac{1}{\lambda_{z,j}} - \sum_k \tfrac{1}{\lambda_{z,k}} \\ &= \sum_j \tfrac{N_j C_j}{v_{dz}} - \sum_k \tfrac{N_k E_k}{v_{dz}} \\ &= \sum_j N_j\sigma_j\tfrac{\langle v_{tot}\rangle}{v_{dz}} - \sum_k N_k\sigma_k\tfrac{\langle v_{tot}\rangle}{v_{dz}}\end{aligned} \tag{5.5}$$

Here, we provided for "negative" capture lengths in the summation over "k-th" terms. These negative capture lengths correspond to possible carrier generation processes which will presented in the next chapter.

5.1.3 CDMS conditions: a 2-impurity, 2-carrier capture model

Our experimental results appear to be dominated by capture. Of course other processes may also occur, but we limit ourselves to what we believe are dominant terms. To limit the scope of this discussion, we consider only a 2-species capture model. We suggest the following form, representing capture due to charged/Coulomb capture (index "q") and to capture forming D^-/A^+ states (index "0").

$$\frac{1}{\lambda_z} \approx \left(N_q \sigma_q \frac{\langle v_{tot} \rangle}{v_{dz}} + N_0 \sigma_0 \frac{\langle v_{tot} \rangle}{v_{dz}} \right) \tag{5.6}$$

Here, N_q is the density of charged capture centers, N_0 the density of neutral centers, σ terms are cross sections, and v_{dz} represents the z-axis drift velocity.

5.2 Neutral capture, forming D^-/A^+ states

As we see most pronounced in the "Phipps 2011" data set of figure 5.3, the inverse capture length varies with a small but well-defined field dependence at higher drift fields. Specifically, we can perform a numerical fit of the 20 points of the inverse capture length at highest measured fields. For both electrons and holes of the "Phipps 2011" dataset, this fit to a power law $\propto F^a$ gives an exponent range $-0.3 < a < -0.2$ for both electrons and holes at highest fields. Now consider that this inverse capture length is dominated by a single capture process, as $\frac{1}{\lambda_z} = \frac{N_0 C_0}{v_{dz}}$. If this capture rate, $N_0 C_0$, were *energy-independent*, there should no dependence on carrier temperature in the presence of an electric field. Most likely, this capture rate should then be *field independent* as well. Therefore, the field dependence of the measured inverse capture length would be due to only the drift velocity. The drift velocity ideally goes as $F^{1/5}$ (as in Appendix B), assuming that acoustic phonon emission dominates transport and the effective mass is constant. For this interpretation, we consider electrons to follow the drift velocities predicted by our anisotropic transport simulation of the last chapter. Since we believe, from measured data, that holes are *not* emitting optical phonons in this measured range of fields, we interpret holes to follow closely the $F^{1/5}$ dependence depicted in figure 4.5.

Our observed capture data in CDMS detectors seems to be consistent with an ideally (or nearly) energy-independent rate at the highest measured fields. The only nearly energy-independent capture process reported readily in the literature [15] is one which represents a direct, quantum-mechanical transition from continuum to a bound, ground state.

In fact, a possible capture state is to form D^- and A^+ states from neutral D_0 or A_0 states. These "over-charged" states are analogous to the negative hydrogen "anion" (H^-) bound state, which can carry two bound electrons. The negative hydrogen ion was first predicted by Bethe [90], and the semiconductor impurity analogs have been evidenced in germanium [91, 92, 15]. In this case, the polarizability of the neutral impurity allows for a bound, "anion" state to exist. As the bound state energy is extremely small, we can expect a particularly different capture (and ionization) behavior from these states versus the "standard" hydrogenic charge states (D^+/A^-). Neutral capture occurs near the band edge in this case, rather than across the gap to form D^+/A^- states. The "same-side" neutral capture state is so shallow, capture is dominated by the direct/"resonant" capture of section 5.2, and not by the "cascade capture" mechanism which we will later describe.

The extreme low-temperature CDMS conditions of the crystal which would allow such weakly-bound D^- and A^+ states to remain stable. This does not out rule the possibility for other kinds of states to be the cause of this capture mechanism in our data, such as some variety of deep state. D^- and A^+ states are natural candidates to consider, however, since they are associated with known shallow impurity concentrations and would be stable for our milliKelvin operation.

We now follow the treatment for neutral capture into these "anion" states.

5.2.1 The D^-/A^+ / "anion" wavefunction

We use the wavefunction of the "zero-radius-potential model" of Demkov [93], as reported by Abakumov [15].

$$\psi = \sqrt{\frac{\kappa}{2\pi}}\frac{B}{r}e^{-\kappa r} \tag{5.7}$$

where B is a normalization, and

$$\kappa = \sqrt{\frac{2m_c\epsilon_{anion}}{\hbar^2}} \tag{5.8}$$

Note that the effective mass of the bound carrier enters, so this is an aspect how D^- and A^+ states differ.

The normalization factor B used is equal to 1.1, as from [15], which accounts for the normalization with core correction in this approximation.

5.2.2 Capture rate to an anion

Abakumov treats neutral capture to anion states with a quantum-mechanical coupling integral [15]. His equation 4.34 of [15], depicts the direct anion-state coupling as a capture rate per unit density.

$$C = \frac{\pi^3 B^2}{16 l_0} \sqrt{\frac{2\epsilon_{anion}}{m}} \left(\frac{\hbar}{m v_s}\right)^3 \Psi(E_k) \tag{5.9}$$

The prefactor B again is the normalization for the anion wavefunction of equation 5.7. The kinetic energy of the incident carrier is E_k.

The scalar l_0 is adopted from Abakumov's formalism [15], and represents a scattering length constant derived from Fermi's golden rule for acoustic phonon emission. It groups scattering parameters such as the deformation potential into one term. We express it as the following.

$$l_0 = \frac{\pi \hbar^4 \rho}{2 m_d^3 \Xi^2 I^2} \tag{5.10}$$

Above, ρ is the crystal mass density, Ξ is the acoustic deformation potential, I^2 is the wavefunction overlap factor, and m_d is the density of states effective mass.

The anion bound state energy is ϵ_{anion}, and the coupling factor $\Psi(E_k)$ ranges between zero and unity. The coupling factor $\Psi(E_k)$, not to be confused with the wavefunction ψ, is given by the following.

$$\Psi(E_k) = \frac{4}{\pi^2} \left(\arctan \frac{1}{x} - \frac{x}{x^2+1}\right)^2 \tag{5.11}$$

Here, the variable x is the ratio of bound-state anion wavevector to the putative acoustic phonon emitted in the capture process.

$$x = \frac{\kappa}{q} = \frac{\sqrt{2 m v_s^2 \epsilon_{anion}}}{\epsilon_{anion} + E_k} \tag{5.12}$$

Equation 5.11 is strictly appropriate for $\sqrt{E_k 2 m v_s^2} \ll (\epsilon_{anion} + E_k)$, else the coupling factor tends to unity if $E_k \gg \epsilon_{anion}$ for higher kinetic energies. A high kinetic energy is often the case in our system of hot carriers, particularly at higher fields where neutral trapping processes can actually be noticed over Coulomb capture. At zero kinetic energy, this factor is still simply $\approx 1/2$. Therefore, for simplicity, we evaluate equation 5.9 under the coupling factor of equation 5.11 set to unity. We treat neutral anion capture to both donors and acceptors with a very shallow ($\epsilon_{anion} = 0.7\ meV$ [92]) binding energy.

5.2.3 Concluding neutral capture to form "anion" (or at least "resonant") states

Concluding this section, we can say that because anion states may be thermally stable at $T = 40\ mK$, their production by carrier capture presents a large capture cross section depicted by equation 5.9.

As a caveat, the model proposed here does not out rule the possibility that other, similar states may exist which would act in the same way but with slightly different binding energies. For instance, the term "deep level" traditionally refers to a non-perturbative state located deep in the band gap. However, many varieties of localized defects in germanium may produce strong interactions with the Bloch electrons of the host crystal, yet their "binding energy" is still on the order of meV [47]. Generally, when a deep center energy level overlaps with continuum band states, *resonant states* are formed [47].

For the purposes of this work, we proceed to consider neutral capture using only with the benchmark model presented above. We continue to refer to neutral capture as forming the traditional, hydrogenic "anion" states.

5.2.4 Capture rates and the steady-state, charged impurity ratio

So far, we have established a plausible, energy-independent capture process to explain our high-field capture data for electrons and holes. We must also understand the *other* capture process in figure 5.3, which dominates at *low fields*.

At this point, it is worthwhile to introduce another concept; the steady-state, charged impurity ratio. Given net impurity concentrations, this ratio determines how many of these impurities should be ionized, assuming a balance of capture rates and net bulk neutrality. This ratio is determined completely by theoretical considerations of balancing rates in a 2-carrier, 2-impurity capture model. For details, refer to appendix D, regarding arguments of defining a non-equilibrium chemical potential through a balance of capture rates.

We can use the identity from appendix D that we can use the following approximate relation.

$$\boxed{N_D^- = N_A^+ \approx \sqrt{\left(\frac{C_{\mathrm{nA0}}\ C_{\mathrm{pD0}}}{C_{\mathrm{nA+}}\ C_{\mathrm{pD-}}}\right) N_A N_D}} \tag{5.13}$$

Here, the C terms are the appropriate recombination rates per density for the four capture processes in this 2-species, 2-center model. Although this defines the steady state ratio of charged impurities, these rates may be field-dependent.

Therefore, this ratio is an important consideration when considering possibilities for steady-state built-in fields and capture mechanisms.

5.3 Cascade capture: from Thomson, to Lax, to Abakumov

Since germanium is an indirect semiconductor, capture mediated by *photon* emission (e.g., "radiative capture") is a rare and negligible effect [94]. The other known, non-radiative capture mechanisms fall under the category of *cascade capture* processes. This is where carrier capture is assisted by the emission of a *phonon* while in proximity to a capture site. This is a "cascade" process because de-exciting a free carrier from the continuum all the way to the bound, ground state of an impurity typically requires the emission of many phonons. There are many varieties of cascade capture, involving different forms of electromagnetic interaction. The kinematics of phonon-assisted capture by various interaction potentials introduces different forms of carrier energy (field) dependence. In this section, we describe the principle of cascade capture processes, and outline the typical attractive potentials which may be involved.

5.3.1 A consideration of phonon emission kinematics

Generally, cascade capture is an emission process requiring multiple phonons (depending on the temperature scale). To see why this is necessary, we briefly explore the kinematics of the situation.

We primarily consider *acoustic* phonons. These phonons are much more readily emitted by free carriers than are optical phonons. While compared to *photons* traveling at the speed of light, *phonons* propagate at comparatively slow speeds of sound. So while these quanta propagate with large wavevectors (e.g., large momenta), acoustic phonons represent *small* quanta in energy.

Consider a free carrier to have kinetic energy $E_k \gg 1/2 m_c v_s^2$. If this carrier has a positive total energy E and emits an acoustic phonon while in proximity to a capture site of potential U, we know the maximum phonon wavevector that can be emitted represents a complete backscatter of the electron.

$$q_{max} \approx 2k = 2\frac{\sqrt{2m_c\,(E-U)}}{\hbar} \tag{5.14}$$

The *change* in *carrier energy* due to the emission of this phonon is small; typically on order of a few factors times $1/2 m_c v_s^2$. So it is actually a relatively rare process for a carrier to emit an adequately large phonon to become bound to an impurity site. Figure 5.4 shows this situation.

For a more thorough investigation of these capture kinematics, please see appendix section A.4.

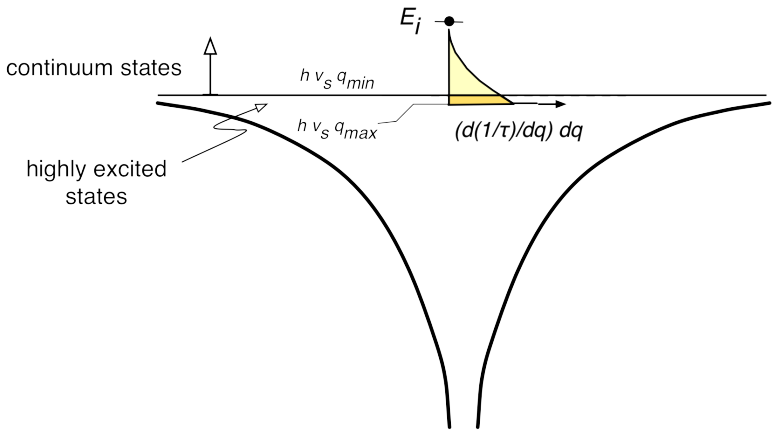

Figure 5.4: A free carrier occupying a continuum state, with initial energy E, is depicted along with its differential acoustic emission rate. This illustrates the probability to de-excite the carrier into a highly-excited bound state upon emission of an acoustic phonon. Even a maximum amplitude phonon, which would completely backscatter the carrier, still dissipates little energy. This is one reason why cascade capture processes must be considered carefully.

5.3.2 Lax's modification of the Thomson capture cross section

Germanium was actually the material studied in low-temperature ($T < 10\ K$) experiments [95] that demonstrated how *large* capture cross sections may actually be in semiconductors. These large cross sections, of order $10^{-12}\ cm^2$, were a mystery since the cross sections predicted by carrier emission of a single photon or phonon directly to the ground state are many magnitudes smaller. Melvin Lax [96] first proposed the theoretical underpinning to explain these large cross sections as the result of a *multi*-phonon, or a *phonon cascade*, mechanism. The principle is that an incident charge carrier may recombine to some impurity or defect site by the emission of a single, relatively small energy phonon. However, for this carrier to become ultimately bound, the carrier subsequently emits a successive chain of phonons transitioning from a nearly-ionized state to the ground state. Lax used primarily semi-classical concepts to construct his model, but his predicted cross sections were the first to match those from experimental measurements.

Lax points out [96] that his work "bears a close resemblance" to J.J. Thomson's theory of recombination in gases [97]. Thomson's theory depicts a thermal equilibrium case where an ionic capture cross section is depicted by a critical radius, r_c, and is determined by an energy equality.

$$\frac{Ze^2}{(4\pi\epsilon\epsilon_0)r_c} = \frac{3}{2}kT. \tag{5.15}$$

Thomson's insight was that ions with kinetic energy greater than the thermal equilibrium average of $3/2\ kT$ typically lose energy upon subsequent scattering, whereas ions with energies much lower than this will gain energy on average. So the volume determined by this critical radius represents a sphere of influence where ionic collisions are effective in producing recombination. The average distance across this spherical volume is $4r_c/3$. Given a mean free scattering path, λ, the probability of collisions within the sphere is $4/3r_c/\lambda$. An incident cross section for this volume can be described in 2-D as πr^2. Combining these terms, Lax points out that Thomson's proposed cross section is

$$\sigma = \frac{4\pi}{3}\frac{r_c^3}{\lambda} \tag{5.16}$$

What Lax accomplished was to produce a derivation of carrier capture in semiconductors in a framework similar to Thomson's cross section. In doing so, Lax points out that one of the of the limits in Thomson's derivation is to assume the limit that $r_c/\lambda \ll 1$. That is to say that scattering events are infrequent during the carrier transit of the trapping volume, so that capture must be dominated by single scattering events. To examine this, we can take an example. Typical mean free paths in the milliKelvin limit are at the 10 μm scale, as per chapter 4. On the other hand, empirical capture cross sections at fields under $\sim 1\ V/cm$ are at the $10^{-11}\ cm^2$ level. If we compare the scattering length to the cross section radius, we see many magnitudes of difference. It is unlikely that more than one scatter will contribute to a capture event in these conditions.

Lax points out that other limiting cases exist. A *diffusion-limited* regime could exist, where $r_c/\lambda \gg 1$. This is a case Lax attributes to Langevin. Here, several stochastic scatters occur along the trajectory in and out of the capture volume. Trapping would be limited by drift into the central $1/r$ potential, balanced by diffusion out of the trapping volume. The cross sections proposed by this mechanism are far larger than measured, and it can be concluded that the Langevin condition does not apply to typical semiconductor conditions.

Lax also notes that Wannier had proposals of carrier relaxation to recombination centers. In Wannier's model, capture would not be limited by the concentration of trap sites, but rather on the incident distribution of free carriers able to impart their energy to traps by acoustic phonon emission. By energy and momentum considerations, these carriers would have typical velocities at about $2v_s$, or twice the speed of sound. Although observed trapping probabilities *do* have a dependence on trap concentration contrary to this theory, the insight that the kinematics of the incident carriers is an important one. We will visit similar kinematic arguments in the derivations to come, later in this chapter.

5.3.3 Lax's "Sticking Probability"

Lax's theory of phonon cascade capture determines that highly excited impurity states provide an important trapping mechanism. To understand this requires a consideration of the transitions a free carrier must undergo between occupying a continuum state and occupying the impurity ground state. A carrier "cascades" down the ladder of impurity states, emitting phonons all the while. It is possible, though, that a carrier may subsequently *escape* from a nearly-ionized state by absorbing a phonon. So we find that there are a number of possible combinations to either emit or absorb phonons as a carrier climbs up and down the ladder of transitions, ultimately to be bound or re-ionized.

Definitions of energy

Precise definitions of rates and energies play an important role in defining our problem. What does energy relaxation rate mean while the carrier is within the potential well? This relaxation time is based on the phonon emission rate due to *kinetic* energy. To keep our definitions clear, recall from classical mechanics the definitions of total, kinetic, and potential energy.

$$E = E_k + U(r) \tag{5.17}$$

Note that $U(r)$ is defined as a *negative* quantity. Although a simple fact, this can often be a source of confusion. The bracketed energies illustrated in figure 5.5 may be helpful.

So as the carrier moves into the trap region, a non-zero value for $U(r)$ causes the values of total energy, E, and kinetic energy, E_k, to diverge. Consider an initial carrier distribution based on total energy, where the potential energy is zero. As a carrier enters an area of non-zero $U(r)$, the kinetic energy increases and this should be accounted for within our capture integral.

Accounting for the random velocity as a function of position

If our calculation were performed in bulk, we would already have velocity distributions numerically calculated by Monte Carlo. For the purpose of cascade capture to a local potential well, however, we wish to account for the relaxation kinematics as a function of position. In this completely semiclassical treatment, we always assume the bulk effective mass is valid. Given an incident energy has total energy $E_k = E$ while deep in the bulk, we assume that its random velocity as a function of position will become

$$\mathrm{v}(E,r) = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2\left[E - U(r)\right]}{m}} \tag{5.18}$$

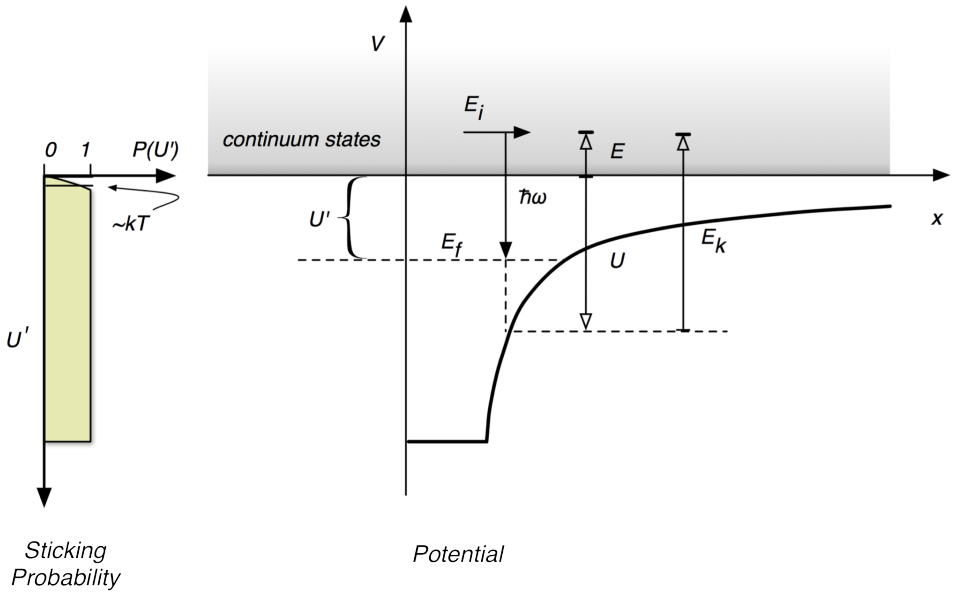

Figure 5.5: A depiction of a carrier capture process by way of phonon emission into an attractive potential is presented. The meanings of the various energies are illustrated. In particular, the binding energy U' and sticking probability $P(U')$ are depicted, as well as total energy (E), kinetic energy (E_k), and potential energy (U) of the initial state.

Introducing the sticking probability

To evaluate the probability for a carrier in an excited state to remain permanently captured may seem to be a daunting exercise in Markov processes. However, Lax was able to show that the balance of subsequent transition rates could be modeled as a *state function*. Lax defines a "sticking probability," ranging from zero to unity, which can be thought of as the probability for a carrier to ultimately remain bound to the impurity. The virtue of the sticking probability is that it is a memoryless state function of energy, bypassing laborious path integrals of emission and absorption probabilities. The sticking probability is a function, $P(U')$, where U' is the "binding energy." As in figure 5.5, the binding energy is simply the potential energy measured into the potential well, as measured from the continuum band edge.

Ambient phonons typically carry an energy $\sim kT$. Since this is the typical energy an absorption event may impart to a carrier, it sets the scale of binding energy for a carrier to re-ionize. Therefore, the sticking probability will vary from zero to unity as a function of the "binding" energy as measured from the conduction band in the pure crystal (near the band edge). The transition width is of order $\sim kT$. Figure 5.5 depicts the sticking probability for a thermal equilibrium case.

Several comprehensive reviews of capture cross sections [94, 98, 15] have appeared

in the literature. They propose functional forms appropriate for sticking probabilities in various limits. Some treatments even forgo the use of sticking probabilities altogether [99], but must retain complicated integrands or numerical treatments to retain the dynamics of emission and absorption probabilities.

The energy scale for the sticking probability is determined by the lattice thermal energy, kT. This leads to a common, naive assumption of a step function to approximate $P(U')$. That is, the sticking probability is assumed to be zero while $U' < kT$, and unity for $U' \geq kT$ [15].

If we had a continuous form for the sticking probability $P(U')$ to consider, we would further need to consider that an energy-dependent cross section also depends on binding energy U' (the energy below the conduction band in the free crystal). We would require a weighted integral over binding energy [96].

$$\sigma(E) = \int \sigma(E, U') P\left(U'\right) dU' \tag{5.19}$$

Where $\sigma(E, U')$ is dependent on both the incident carrier energy and binding energy.

5.3.4 Cascade capture, revisited by Abakumov and others

Analytical formulae for characteristic processes have been developed, commonly with ties back to Lax's innovative work on cascade capture [96]. In this regard, notable achievements are attributable to Abakumov, Reggiani, Ridley, Landsberg, and their collaborators [15, 100, 5, 94]. Such formulae are typically expressed as a function of lattice temperature. Since equilibrium is assumed, incident carriers are treated under Maxwell-Boltzmann statistics using the lattice temperature. While there exist some corrections for the presence of an applied field [100, 15, 94], carrier equilibrium with the lattice is often –but not always– still assumed.

Abakumov formulation, for $kT \gg mv_s^2$

The seminal work in capture cross sections performed by Abakumov and collaborators [14, 15] is particularly useful for our purposes. He revisits the formulations of Lax, but notes that at large temperatures ($kT \gg mv_s^2$), cascade capture need not be dominated by single phonons of energy $\approx kT$. Abakumov uses similar underpinnings as Lax, but reformulates a cross section for high temperatures without a sticking probability. He adopts a method based on what he calls "the method of Pitaevskii [101]," which amounts to a kinematic "Fokker-Planck," or what we call a "Boltzmann Transport," based formulation.

Here, we give a brief overview of Abakumov's method for this limit $kT \gg mv_s^2$. This is method is developed in references [14, 15]. In summary, Abakumov solves

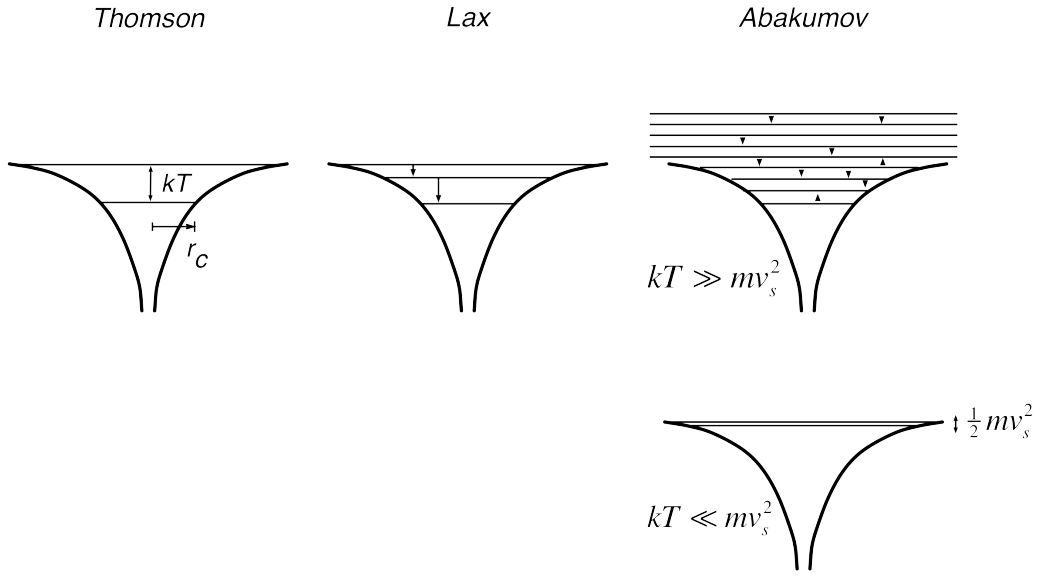

Figure 5.6: This compares successive approaches to carrier capture. Thomson, in considering capture in gases, uses an argument of thermal/kinetic energy balancing capture by an attractive potential, as in equation 5.15. Lax considers consecutive collisions as a cascade capture process, requiring a "sticking probability" to depict the probability of re-emitting a carrier before final capture (as also detailed in figure 5.5). Abakumov treats the capture problem in separate limits. In the limit $kT \gg mv_s^2$, Abakumov uses a "Boltzmann Transport" approach, solving for cascade capture by way of a "collision integral" and a self-consistent carrier distribution function. For Abakumov's limit $kT \ll mv_s^2$, the appropriate integral represents capture limited by only one phonon emission event while in proximity to a capture site. Abakumov's capture rates are integrations performed over possible initial and final-state energies, as well as real-space volume. This figure is reproduced and extended from reference [14].

the Boltzmann Transport equation for the particular case of a free carrier within the potential of a local impurity. This is depicted in figure 5.6.

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f - e\frac{\partial U}{\partial \mathbf{r}} \cdot \nabla_{\mathbf{p}} f = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{5.20}$$

Above, U is the local impurity potential, f is the carrier distribution function, $\left(\frac{\partial f}{\partial t}\right)_{\text{coll}}$ is the "collision integral," and the partial derivatives of the left-hand side represent the time-derivative of the distribution function, as well as the "flow" of the distribution function through real and momentum phase spaces [14, 15, 45, 102]. To conserve energy, he multiplies applies a delta function to both sides of the equation, $\delta\left(E - E_k - U(r)\right)$, where E is the total carrier energy and E_k its kinetic energy.

He then integrates over all real-space ($d^3\mathbf{r}$) and momentum-space ($d^3\mathbf{p}$) to find the relation

$$R\left(E\right)\frac{\partial f}{\partial t}=-\frac{\partial C}{\partial t} \tag{5.21}$$

where $R(E)$ is the "density of states in total energy space,"

$$R\left(E\right)=\int\rho\left(E_k\right)\delta\left(E-E_k-U(\mathbf{r})\right)dE_k d^3\mathbf{r} \tag{5.22}$$

and $\rho(E_k)$ is the density of states in (kinetic) energy space. C is "the flux of particles reaching a single isolated center under steady-state conditions," amounting to a capture rate per density. This capture rate per density is further found to be

$$C=B\left(E\right)\left[f\left(E\right)+kT\frac{df\left(E\right)}{dE}\right] \tag{5.23}$$

where $B(E)$ is a quantity known as "the coefficient of dynamic friction," and is an integral over energy and real-space of the energy loss rate of the carrier.

$$B\left(E\right)=\int\frac{E_k}{\tau_{E_k}}\rho\left(E_k\right)\delta\left(E-E_k-U(\mathbf{r})\right)dE_k d^3\mathbf{r} \tag{5.24}$$

The τ_{E_k} is an energy relaxation time based on scattering rate.

The boundary condition for the distribution function $f(E)$ is given as

$$f\left(-E_1\right)=0,\quad -E_1\gg kT \tag{5.25}$$

where E_1 is the ground state of the potential. Abakumov further assumes that the distribution function is in equilibrium for energies free of the potential, $E>0$.

$$f\left(E\right)=Ae^{-E/kT},\quad A=\frac{1}{2}\left(\frac{2\pi\hbar^2}{mkT}\right),\quad E>0 \tag{5.26}$$

This is enough information to express a solution for the non-equilibrium distribution function within the potential.

$$f\left(E\right)=\frac{C}{kT}\exp\left(-\frac{E}{kT}\right)\int_{-E_1}^{0}\frac{\exp\left(E'/kT\right)}{B\left(E'\right)}dE' \tag{5.27}$$

We also know the capture rate per density is related to random velocity and the cross section.

$$C=\sigma\left\langle v_{tot}\right\rangle \tag{5.28}$$

Abakumov also assumes a Boltzmann Transport derivation of the random velocity to be the following.

$$\left\langle v_{tot}\right\rangle=\sqrt{\frac{8kT}{\pi m}} \tag{5.29}$$

The velocity of equation 5.29 is about 8% less than the equipartition assumption of $\sqrt{3kT/m}$.

Finally, the capture cross section in this $kT \gg mv_s^2$ limit can be equated as

$$\boxed{\sigma_{\text{Hi-T}} = \frac{\pi^2\hbar^3}{2kTm}\left[\int_{-E_1}^{0} \frac{\exp\left(E'/kT\right)}{B\left(E'\right)}dE'\right]^{-1}} \tag{5.30}$$

This expression is general for any type of impurity potential, which appears in the expression for $B(E')$ in equation 5.24.

Abakumov formulation, for $kT \ll mv_s^2$

At extremely low temperatures, Abakumov reformulates capture based on the rate of a *single emitted phonon* to de-excite the passing carrier into a bound state [14, 15]. In this low-temperature limit, both the need for a "sticking probability" as by Lax and the self-consistency needed for the Boltzmann Transport formulation of the previous section are bypassed. There are no ambient phonons for absorption, and so the rate per density integral becomes an evaluation of positive initial energy carriers de-exciting into a final energy state which is negative. This integral again requires an integral over initial and final carrier energies, as well as over the real-space volume of the impurity.

Abakumov uses the following terminology. We can identify his equation 4.1 of [15] with the numerator of equation 5.28.

$$\boxed{\sigma_{\text{Lo-T}} = \frac{C}{\langle v_{tot}\rangle} = \frac{1}{\langle v_{tot}\rangle}\int d^3r \int dE_k \int dE_k' \;\; \mathcal{F}(E_k, r)\; \rho(E_k)\; w\left(E_k, E_k'\right)} \tag{5.31}$$

Here, we have maintained our own conventions for energy quantities as per chapter A. However, $\mathcal{F}(E_k, r)$ is an energy and position-dependent distribution function for incident carriers. Also, $\rho(E_k)$ is the density of states, *referenced from the bottom of the capture potential*, $U(r)$. Since the density of states weighs the low-energy side of the total carrier distribution, we can think of this treatment with similarities to the non-ballisitic limit depicted in figure A.4. Shortly, we will examine the particular coupling rate used here, $w(E_k, E_k')$.

The idea behind equation 5.31 is what Abakumov thinks of as a *carrier flux per center*. It represents the flux of incident carriers of initial kinetic energy E_k, representing positive total energy, which couple within the real-space integral to final kinetic energies E_k' which are bound (representing negative total energy). The rate $w(E_k, E_k')$ determines the coupling of carriers of initial kinetic energy E_k to possible final states of E_k'.

We will now introduce the relevant quantities and integral limits for this formulation.

Carrier distribution, density of states, and coupling rate for Abakumov's $kT \ll mv_s^2$ limit Here, we point out the relevant forms for the quantities in the integrand of equation 5.31. We can group the density of states and distribution function into the "displaced Maxwellian" joint distribution we discuss elsewhere, as in Appendix C. Ignoring the "drift energy" term, we recover the standard Maxwell-Boltzmann distribution follows.

$$\mathcal{F}\left(E_k, r\right)\rho(E_k) = \frac{2}{\sqrt{\pi}\left(kT\right)^{3/2}}\sqrt{E_k}\exp\left[-\frac{\left(E_k + U(r)\right)}{kT}\right] \tag{5.32}$$

Take note that we have introduced the exponential term as referred to the *total* energy, $E = E_k + U(r)$.

The coupling rate, $w(E_k, E_k')$, while seemingly straightforward can actually be a source of frustration. This is *not* the integrated rate, but a partial rate based on an integrated energy relaxation time.

Define a power loss by phonon emission, as

$$P = \frac{dE}{dt} = \int \left(\hbar v_s q\right)\left(\frac{\partial\Gamma(q)}{\partial q}\right)dq \tag{5.33}$$

where $\frac{\partial\Gamma(q)}{\partial q}$ is the acoustic emission probability before integrating with q, as can be found in our previous chapter or as in Lundstrom equation 2.81 [45].

Abakumov's equation D.12 [15] is

$$-\frac{dE}{dt} = \int w\left(E_k, E_k'\right)(E_k - E_k')dE_k \tag{5.34}$$

By identifying the integrands of 5.34 and 5.33, we find the necessary connection allowing us to coincide Abakumov's rate with our scattering rate formalism of equation 2.90.

$$w\left(E_k, E_k'\right) = \frac{\left(E_k - E_k'\right)^2}{8\sqrt{2}l_0\sqrt{E_k}m^{5/2}v_s^4} \tag{5.35}$$

Here, l_0 is defined in equation 5.10. We now have the constituent parts to repeat Abakumov's calculation of equation 5.31.

5.3.5 Concluding cascade capture

In this section, we have introduced the concept of phonon cascade as a capture mechanism to relax free carriers. We reviewed the insights made by Thomson, Lax, and others. Of most importance, we have investigated the generally applicable expressions for Abakumov's cross sections in two important temperature limits.

5.4 Possible attractive potentials related to cascade capture

In the last section, we developed a formalism for capture cross sections which requires knowledge of the attractive interaction between impurity and carrier. In this section, we introduce the possible impurity potentials which may contribute to the cascade capture process.

Mapped as a one-body attractive potential, the general form goes as

$$U(r) = \begin{cases} -Ar^{-n} & for\ r\ \geq\ r_{min} \\ -A_0 & for\ r\ <\ r_{min} \end{cases} \tag{5.36}$$

Here, r_{min} is a minimum radius, for which the potential has "bottomed out" to the assumed ground state energy, $-A_0$. This is a common, semiclassical approximation for attractive potentials, which accounts for core effects [102].

As an example, a shallow, Coulomb center would have $A = Ze^2/(4\pi\varepsilon\varepsilon_0)$, $n = 1$, $r_{min} = a_B = \frac{\hbar^2(4\pi\varepsilon\varepsilon_0)}{Ze^2m} \approx 7$ nm (the effective Bohr radius), and $A_0 = \frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)a_B}$.

We consider that these local trap potentials are present in an applied electric field. At low temperatures, the effect of barrier lowering due to an applied electric field is an important concern. To approximate the impact of the electric field, we "lower" the barrier by a constant factor. We find the correction as follows.

We know that the externally applied electric field contributes to the potential energy in the following way,

$$U_{tot}(r) = U(r) - eFz \tag{5.37}$$

where z is the z-coordinate, symmetric about the trap potential. The electric field is F along the z-axis, and the carrier charge e. We find where the derivative is lowest along r, hence along z, and find where the derivative is zero. This gives the maximum barrier height where the potential has been lowered, as in figure 5.7.

$$\partial_r U_{tot}(r) = 0 = Anr^{-n-1} - \mathrm{eF} \tag{5.38}$$

We find the "barrier" radius, r_b, for which this occurs,

$$r_b = \left(\frac{\mathrm{eF}}{An}\right)^{\frac{-1}{1+n}} \tag{5.39}$$

and subsequently find the value of the local potential at this distance.

$$U(r_b) = -A(n+1)\left[\left(\frac{\mathrm{eF}}{An}\right)^{\frac{-1}{1+n}}\right]^{-n} \tag{5.40}$$

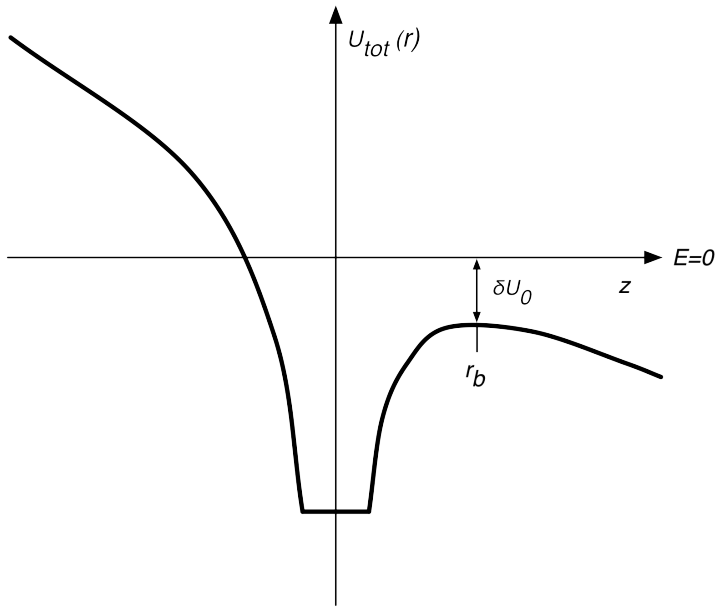

Figure 5.7: This figure demonstrates a (Coulomb) potential subject to an applied field. Here, the total potential $U_{tot}(r) = U(r) - eFz$ is plotted as a function of radial coordinate, taken here along the field's z-axis. This figure shows the truncation of the potential at the ground state, which is assumed to be flat with or without an applied field. Note how the applied field lowers the local barrier. This is described by the barrier lowering parameter δU_0, occurring at barrier radius, r_b.

So what this leads to is a correction to the local potential. We account for the effects of the external electric field by a correction factor to the local potential. Note, that we define this *correction* to the potential as a *positive* term.

$$\delta U_0 = -U(r_b) = A(n+1)\left[\left(\frac{\mathrm{eF}}{An}\right)^{\frac{-1}{1+n}}\right]^{-n} \tag{5.41}$$

We see that we can now express the classically bound states within the potential as contained up to an energy of $\approx |U(r)| - \delta U_0$. As an external field is increasingly applied, the radius at which these potentials "tip over" in the external field depends upon the spatial dependence of the potential.

5.4.1 Charged capture: Coulomb potential

Charged impurities exhibit a standard Coulomb potential. This is well understood as a standard $1/r$ potential,

$$U(r) = -\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)r} \tag{5.42}$$

We follow our prescription for truncating the potential at the ground state, which will correspond to some minimum r_{min}, depending on the type of impurity.

Our definition of the barrier lowering becomes the following.

$$\delta U_0 = 2\sqrt{\frac{Ze^3F}{4\pi\varepsilon\varepsilon_0}} \tag{5.43}$$

5.4.2 Neutral capture: the polarization potential

Now we describe the potential appropriate for neutral capture into "standard" (D^+/A^-) hydrogenic charge states. This process occurs "across the band gap," as opposed to the neutral capture process which forms anion (D^-/A^+) states.

The polarization potential may be less familiar to the reader, so we take more care to explain its origin.

Neutral centers may polarize

To be glib, neutral centers are pith balls. Consider a neutral, hydrogenic impurity in $T = 0$ germanium. For example, consider an acceptor. At the central core, an electron is localized by a strong chemical valence bond. A positive charged hole orbits this central core at a Bohr radius scaled for the hole's effective mass and for the germanium dielectric constant.

Now consider an applied electric field. The effect on the energy states due the electric field is known as the *Stark effect*, and the perturbation Hamiltonian is

$$H' = -e\mathbf{F}\cdot\mathbf{r} \tag{5.44}$$

The occupied atomic eigenstates have definite parity, and the first order perturbation term vanishes due to inversion symmetry. The next, second-order term is the *quadratic Stark effect*. It represents the induced dipole moment produced by an external field. Being quadratic, an electric field force eF produces an energy shift of eF^2. While this resultant energy shift can be determined quantum mechanically, we choose follow Lax [96] by adopting the simple, classical treatment which also happens to match the accepted result.

The induced polarization potential

Consider our neutral impurity acceptor as a simple, classical atom. This treatment follows [103]. The "orbiting" hole is represented by a uniform charge density, which integrated totals one electron charge (+e). This charge density occupies a spherical

volume of Bohr radius a_H. At some far distance, r, a passing electron introduces an electric field (F) into the atomic region. In this external electric field, the nucleus and hole are respectively repelled and attracted to the far off electron. Consequently, a displacement distance, d, is induced between the nucleus and the centroid of the spherical hole density.

Knowing that the total charge for one hole is, of course, $+e$, a simple treatment of Gauss' Law within a spherical and uniform charge density goes as

$$F(4\pi d^2) = \frac{q_{enc}}{\varepsilon\varepsilon_0} = \frac{(+e)}{\varepsilon\varepsilon_0}\frac{d^3}{a_B^3} \tag{5.45}$$

We find an equilibrium between the externally applied field to this self-induced displacement field.

$$F_{ext} = F = \frac{(+e)}{4\pi\varepsilon\varepsilon_0}\frac{d}{a_B^3} \tag{5.46}$$

As seen by the electron, an axially aligned dipole is induced.

$$\wp = ed = (4\pi\varepsilon\varepsilon_0)\, a_B^3 F \tag{5.47}$$

The relationship between the dipole moment and the electric field is linear, so it lends itself to depiction by an *atomic polarizability*, α.

$$\wp = \alpha F \tag{5.48}$$

$$\alpha = (4\pi\varepsilon\varepsilon_0)\, a_B^3 \tag{5.49}$$

The electric field of a dipole is

$$F_{dip} = \frac{1}{4\pi\varepsilon\varepsilon_0}\left(\frac{3(\boldsymbol{\wp}\cdot\hat{\boldsymbol{r}})\hat{\boldsymbol{r}} - \wp}{r^3}\right) \tag{5.50}$$

The dynamical, induced dipole described here is always aligned to be $\hat{\boldsymbol{r}} = \hat{\boldsymbol{\wp}}$. After substituting a point-charge electric field for the induced dipole expression, we find the attractive potential for this interaction to be

$$U(r) = -\frac{e^2\alpha}{2\,(4\pi\varepsilon\varepsilon_0)^2\, r^4} \tag{5.51}$$

So for our model potentials of equation 5.36, we can say for the case of neutral centers that

$$\begin{array}{rcl} A & = & \frac{e^2\alpha}{2(4\pi\varepsilon\varepsilon_0)^2} \\ n & = & 4 \end{array} \tag{5.52}$$

In our simple expressions, the difference in effective masses between electrons and holes, orbiting in their neutral atomic states, are assumed to be the only difference in atomic polarizability between donors and acceptors. Polarizabilities are scaled by the Bohr radius dependence, a_B^3, appropriate for orbiting electrons or for holes. By this argument, the polarizability between acceptors (capturing electrons) and donors (capturing holes) has the following relation.

$$\frac{\alpha_A}{\alpha_D} = \frac{m_e^3}{m_h^3} \approx 0.187 \tag{5.53}$$

Barrier radius and barrier lowering for neutral traps

Following our prescription for a barrier radius of equation 5.39, we find the polarization potential gives

$$r_b = 2^{1/5} \left(\frac{F \left(4\pi\varepsilon\varepsilon_0\right)^2}{e\alpha} \right)^{-1/5} \tag{5.54}$$

Note this field dependence of $F^{-1/5}$ is considerably weaker than for a dipole. So this barrier radius, which is responsible for defining the cross section, changes very weakly as a function of electric field compared to the $F^{-1/2}$ dependence found in the Coulomb trapping critical radius. We should therefore expect the field dependence for neutral capture to be much weaker than for charged capture, because the polarization potential is harder than the Coulomb potential to "tip over" by application of external field.

For the corresponding barrier lowering, equation 5.40 gives the polarization potential barrier lowering as

$$\delta U_0 = \frac{5}{2^{9/5}} \frac{\left(e^2\alpha\right)^{1/5} \left(eF\right)^{4/5}}{\left(4\pi\varepsilon\varepsilon_0\right)^{2/5}} \tag{5.55}$$

The polarization potential depicts the *self*-polarization of a neutral site due to the induction of a dipole moment. We may also consider the case where a dipole moment remains static.

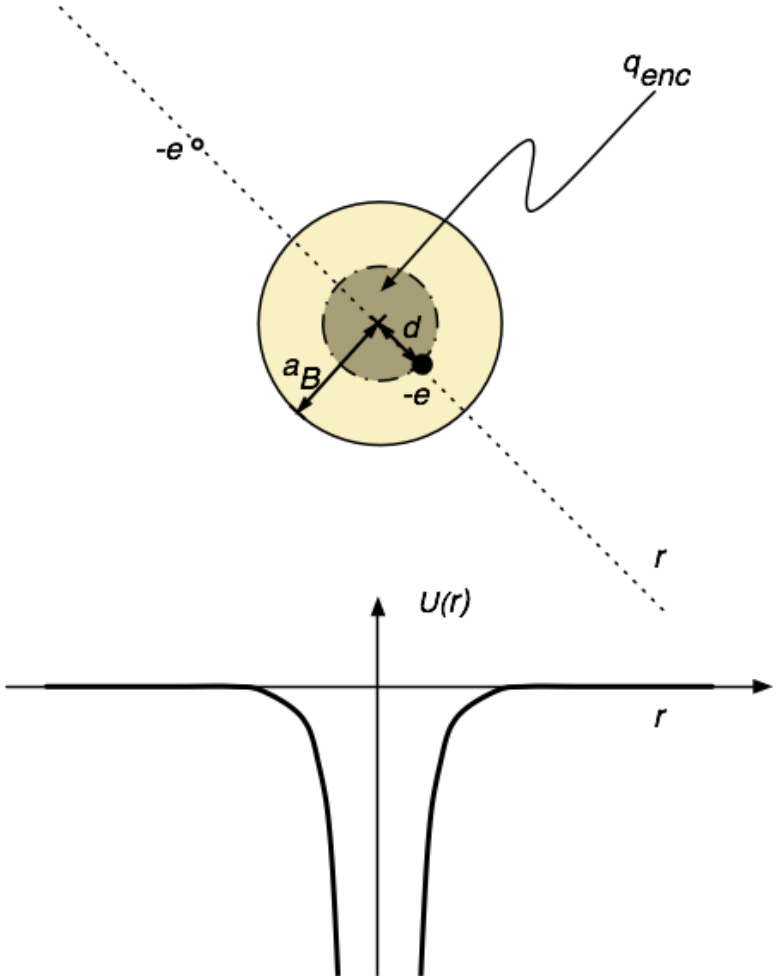

Figure 5.8: A passing electron induces an axially aligned dipole moment at a neutral acceptor. The polarizability of the atom is due to the displacement between the acceptor's nucleus and its (assumed) uniform hole density. The resulting polarization potential experienced by the passing electron can be mapped as a one-body potential, going as $1/r^4$.

5.4.3 The static dipole potential

The case may arise where charged sites occur in some proximity to each other. We operate our CDMS detectors well below freeze-out. If carrier capture is correlated during freeze-out, it may lead to the production of charged sites acting as static dipoles [15, 104], or even complexes of higher moments.

We can write the dipole potential as

$$U(r) = -\frac{e^2 d \cos\theta_d}{4\pi\varepsilon\varepsilon_0 r^2} \tag{5.56}$$

where d is the dipole moment, and θ_d is the angle relative to the dipole vector. So we have an angular dependence in this potential.

As the dipole angle is constant, we follow our prescription for the barrier radius and barrier lowering.

$$r_b = 2^{1/3}\left(\frac{ed}{F}\frac{\cos\theta_d}{4\pi\varepsilon\varepsilon_0}\right)^{1/3} \tag{5.57}$$

$$\delta U_0 = \frac{3}{2^{2/3}}\left(\frac{e^2 d\cos\theta_d}{4\pi\varepsilon\varepsilon_0}\right)^{1/3}(eF)^{2/3} \tag{5.58}$$

The dipole moment's angular dependence adds a complication with which to work. Within our volumetric integral method, we established a framework using cylindrical coordinates to calculate a spherically symmetric potential under the influence of a z-aligned electric field. While the complication of finding the allowed boundary by energy/momentum conservation already made this a numerical calculation, we didn't provide for a non-spherical local potential. We make a naive attempt at the dipole cross section by simply taking an expectation value over dipole angle.

When the $cos\theta_d$ term is negative, the potential is positive and repelling. We do not consider that a carrier may recombine to a repulsive potential, so we limit our scope to the range $\left(0 \leq \theta_d \leq \frac{\pi}{2}\right)$.

5.5 An introduction of the impurities/defects considered

In this section, we list the variety of impurities or defects that we will consider. The general character of the situation is characterized by the impurity potential profiles of figure 5.9.

We will first introduce the class of impurity, then itemize and label the possible processes under consideration.

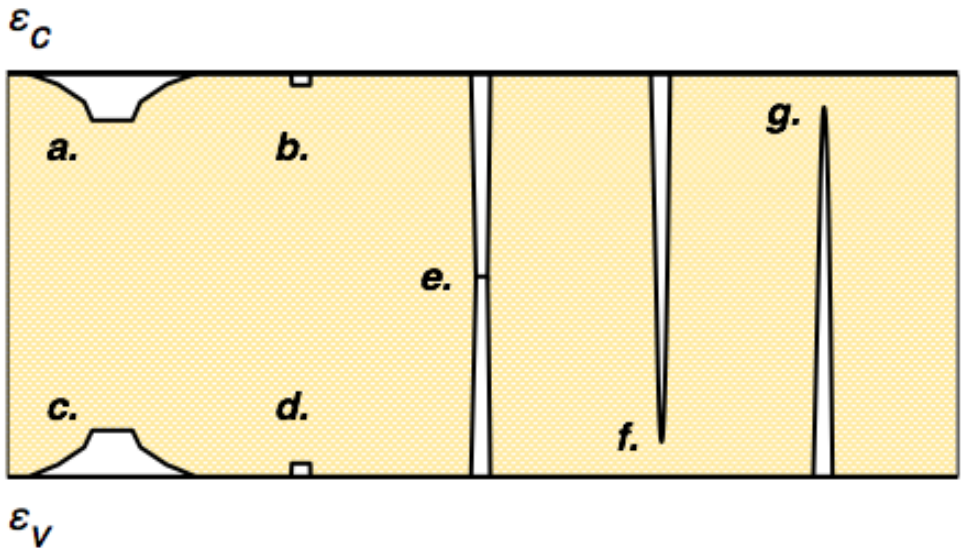

Figure 5.9: This depicts the energetic placement of impurity potentials relative to the band gap, although not to scale. The following benchmark cases to be considered are: a. charged shallow donor, b. neutral anion donor state, c. charged shallow acceptor, d. neutral anion acceptor state, e. neutral deep center, f. neutral shallow acceptor, g. neutral shallow donor. Not shown are the cases of charged deep centers, anion "over-charged" (D^-/A^+) states, or static dipoles.

5.5.1 Shallow / hydrogenic states

Standard shallow, hydrogenic impurities have usually been our first candidates considered regarding issues with impurities. We have come to adopt the term "trap" for these states, although at milli-Kelvin working temperatures these states serve as recombination centers and not temporary traps.

Donors behave well to a simple Bohr model. Due to the warped nature of the valence band, acceptors are more complicated [61]. Density functional theory was used to explain the ground state energy for shallow acceptors, and this matches well to experimental data [61]. We use an experimental value for gallium as our effective Rydberg level for shallow acceptors, as per [61]. As the acceptor's radius determined by this effective Rydberg does not correspond so well to a simple Bohr radius, we furthermore adopt an experimental radius determined for gallium in germanium [105].

5.5.2 Deep centers

Deep levels may occur deep in the band gap. Deep levels are not hydrogenic in nature. They may be caused by severe defects at the level of the lattice spacing, and represent strong departures from a near-continuum, semiclassical understanding. For

our best estimate, we specify a benchmark case.

We will consider that a deep level is associated a disturbance the size of a lattice spacing, rather than a hydrogenic Bohr radius. This means that our benchmark model will have considerably smaller polarizability ($\propto a_0^3$) and therefore a smaller neutral capture cross sections. We place the deep level bound state directly mid-gap (although this may vary). We do not calculate impact ionization cross sections for deep centers. We assume a polarizability, based on the lattice spacing radius. However, deep traps are known to be narrow in their spatial extent. They may also be approximated by a delta function [47].

5.5.3 Anions / "H^-" / "over-charged" states

As already mentioned at the beginning of this chapter, it is possible for an already neutral shallow donor to bind an additional electron forming a D^- state. A hole can bind to a neutral acceptor, it can produce a A^+ state.

We have now reviewed the major impurity types, for completeness.

5.6 Extending Abakumov's equilibrium, low-T limit ($kT \ll mv_s^2$) for the Coulomb cross section

The charged/Coulomb cross section tends to be large. Furthermore, the $1/r$ Coulomb potential also has a large field dependence in the capture rate versus the r^{-4} dependence of the polarization potential. We also expect shallow impurity concentrations to be low (certainly $< 10^{12}\ cm^{-3}$). If we consider static dipoles between opposite-sign impurities, we can assume their dipole moment distance goes as $(N)^{-1/3}$. This is a long dipole moment, so dipoles for these low concentrations act essentially as well-isolated single-point charges.

The the large cross section and field dependence of the Coulomb cross section is consistent with the observed power law behavior at lowest fields in figures 5.2 and 5.3.

Therefore, we treat the low-field capture in CDMS as dominated by charged / Coulomb capture. We proceed by finding an appropriate *equilibrium* expression for charged capture based on Abakumov's formalism. In section that follows, we will use this result within Abakumov's hot-carrier correction, extending this treatment to the *non-equilibrium* Coulomb capture of carriers under drift.

5.6.1 Abakumov's *thermal equilibrium* capture rate for $T_L \approx 40\ mK$

We reconstruct Abakumov's cross section appropriate for low temperatures. We do this in order to make a better evaluation of the Coulomb cross section appropriate

for $T = 40\ mK$, which is not particularly in the limit $kT \ll mv_s^2$.

Using the Coulomb potential with no applied field, we evaluate the capture integral of equation 5.31. We recover the following.

$$\begin{aligned} \sigma &= \frac{8}{45}\frac{\pi}{l_0}\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)kT_C}\right)\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)mv_s^2}\right)^2 + \\ &\frac{4\pi}{15l_0}\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)mv_s^2}\right)^3 + \frac{8\pi}{15l_0}\left(\frac{kT_C}{mv_s^2}\right)\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)mv_s^2}\right)^3 \end{aligned} \tag{5.59}$$

Note that we set $T = T_C$ to specifically indicate a *carrier* temperature. So we have three terms. These are proportionate to:

1. $1/T_C$,
2. constant
3. T_C

To note, Abakumov finds a Coulomb capture expression appropriate for $kT \ll mv_s^2$.

$$\sigma = \frac{8}{45}\frac{\pi}{l_0}\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)\,kT}\right)\left(\frac{Ze^2}{(4\pi\varepsilon\varepsilon_0)\,mv_s^2}\right)^2 \qquad \text{(Low-T limit)} \tag{5.60}$$

It is important to point out that T in this case is *only* the carrier temperature. The *lattice temperature* is effectively zero here, since a sticking probability was not used. Treating lattice temperature as zero here is permissible as most of the capture turns out to occur near the center of the trap. Near the center of the trap, the sticking probability has less impact relative to the additional capture states. Should we reintroduce lattice temperature, it helps to define it as $T = T_L$ to avoid confusion with an effective carrier temperature.

In the limit of T_C approaching zero, eventually only the $1/T_C$ term will dominate. In this case, our three-term expression of equation 5.59, recovers Abakumov's expression of equation 5.60. However, at a carrier temperature $kT_C \approx mv_s^2$, all three terms in equation 5.59 are approximately equal to each other, as Abakumov must have realized. With our benchmark $T_L = 40\ mK$, we see that the lattice temperature is not all that far below mv_s^2/k. Therefore, we keep our formula of equation 5.59 for use in the non-equilibrium expressions we will now show.

5.6.2 Exploring Abakumov's Coulomb capture cross section: non-equilibrium for an applied field

For the case of cascade capture under spontaneous emission, we recovered Abakumov's approximation for temperatures below mv_s^2. The quantity $kT = mv_s^2$, using the

density-of-states mass in germanium, gives a temperature equivalent to $T = 221\ mK$ for electrons, and $T = 353\ mK$ for holes. At an applied field of just 10 mV/cm, Monte Carlo simulation shows electrons are already hot relative to the lattice at $T_C = 908\ mK$, and holes $T_C = 2.76\ K$. So, we are not in Abakumov's low-temperature limit.

Fortunately, Abakumov suggests a form for a carrier capture rate under applied field [14]. For capture by acoustic phonon emission, Abakumov finds the following for a rate per density.

$$C_1 = C_{eq}\left(\frac{T_L}{T_C}\right)^{3/2}\left(1 + 0.98\frac{\delta U_0}{kT_L}\right)^{-1} \tag{5.61}$$

Here, T_L is the lattice temperature, T_C the effective carrier temperature, δU_0 is the Coulomb barrier lowering equivalent to equation A.4, and C_{eq} is the carrier capture rate per trap density when in thermal equilibrium with the lattice. We will use

$$C_{eq} = \left(\sqrt{\frac{8kT_L}{\pi m}}\right)^{-1}\sigma_{eq} \tag{5.62}$$

where we adopt cross section σ_{eq} to be the equilibrium cross section of equation 5.59.

We interpret Abakumov's summary of equation 5.61. An evaluation of of the Maxwell-Boltzmann distribution near zero results in a temperature dependence that goes as $\propto (kT_C)^{-3/2}$. Since capture is dominated by low-energy carriers within a few factors of mv_s^2, we therefore should expect the capture rate to also fall off by the same factor of $(kT_C)^{-3/2}$. The factor containing δU_0 accounts for the bound states near the continuum which can now leak off newly "captured" carriers due to the presence of the lowered barrier.

We define Abakumov's field-dependent cross section in the typical way; by dividing rate per density by total velocity. Considering equations 5.29, 5.59, 5.61, and 5.62, we find

$$\boxed{\sigma_1 = \left(\sqrt{\frac{8kT_C}{\pi m}}\right)^{-1} C_1 \quad (\text{"Abakumov field-dependent"})} \tag{5.63}$$

We have finally arrived at the needed, *non-equilibrium* cross section, which we can use for a lattice temperature $kT_L < mv_s^2$ with carrier temperatures $kT_C > mv_s^2$.

5.7 Returning to CDMS capture data: interpreting the results

We see, particularly in the Phipps 2011 data set, that we have distinctly different modes of capture. Coulomb capture at these lower fields is determined primarily by the carrier temperature dependence of $T_C^{-3/2}$. From our appendix B, we can determine in a simple analytical model that carrier temperature (which is related to mean energy) should have a roughly $\propto F^{4/5}$ dependence on field while transport is dominated by spontaneous emission of acoustic phonons. Note the total and drift velocity dependence of the capture length in equation 5.6. So by including also a $F^{1/5}$ dependence for drift velocity, also from appendix B, Coulomb capture should have a field dependence of

$$\begin{array}{l} 1/\lambda \propto F^{-7/5} \\ \sigma \propto F^{-8/5} \quad \text{(for Coulomb capture)} \\ C \propto F^{-6/5} \end{array} \tag{5.64}$$

The neutral capture rate to form anion states is energy independent. Therefore, the rate is also energy and field independent. The *cross section* then only depends on a denominator proportionate to the total velocity $\propto T_C^{1/2}$. Therefore, we have the following field dependence.

$$\begin{array}{l} 1/\lambda \propto F^{-1/5} \\ \sigma \propto F^{-2/5} \quad \text{(neutral capture into } D^-/A^+ \text{ states)} \\ C \propto F^{0} \end{array} \tag{5.65}$$

So we can see from the field dependence why Coulomb capture dominates at low fields. This happens to be *how* detectors neutralize when they are grounded. When the Coulomb capture cross section is much larger than the neutral capture cross section, it is proportionately easier for carriers "to find" the charged centers and neutralize them. Therefore, there will be less charged centers in the grounded steady state.

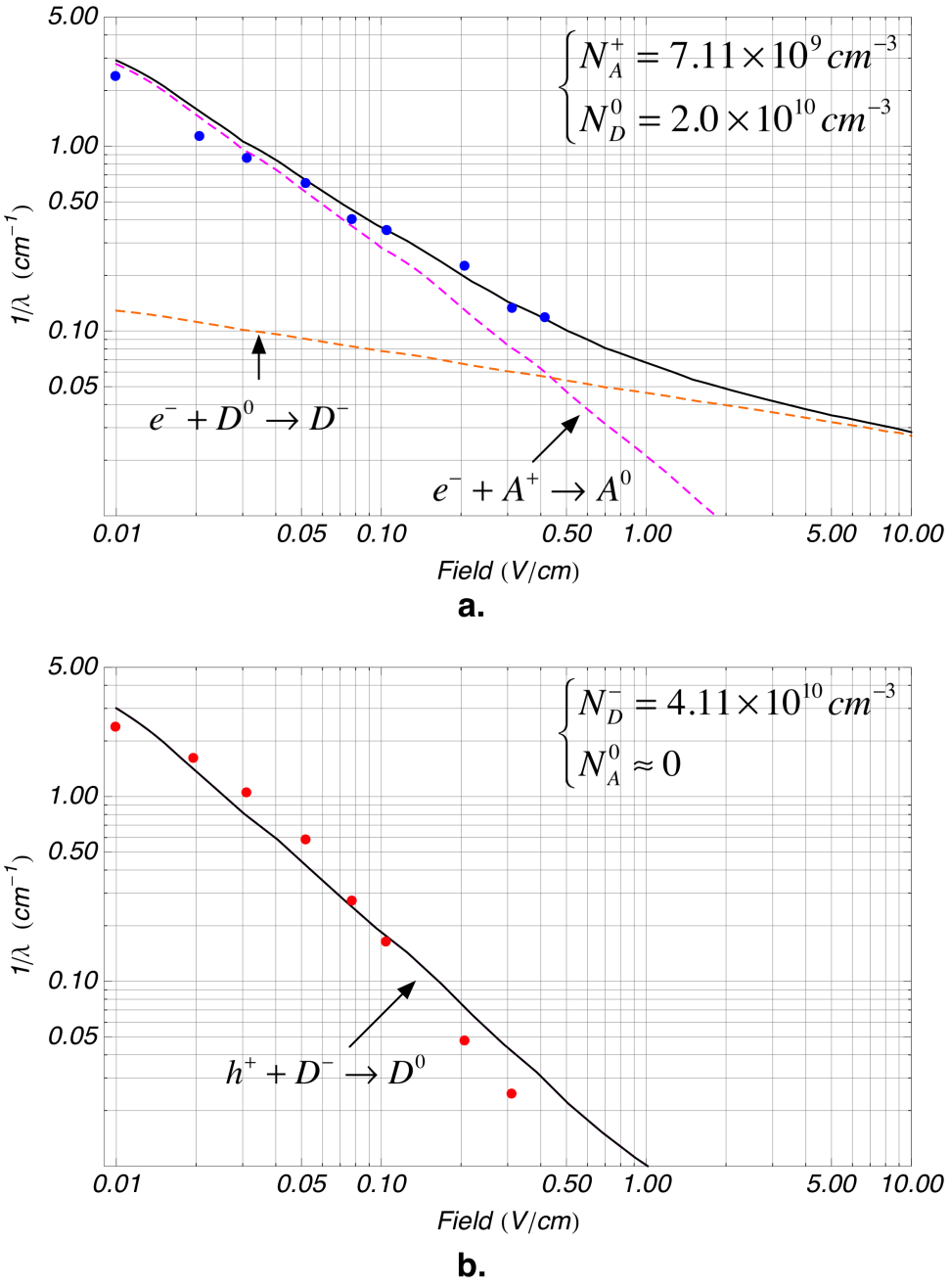

Figure 5.10: Comparison with carrier recombination data in CDMS detectors, from the "Shutt 1993" data set. Theory is for 2-species recombination between Coulomb trapping and neutral capture into anion (D^-/A^+) states. *a.* The inverse capture length for *electrons* with field. *b.* for *holes* – no neutral capture was resolvable in this case.

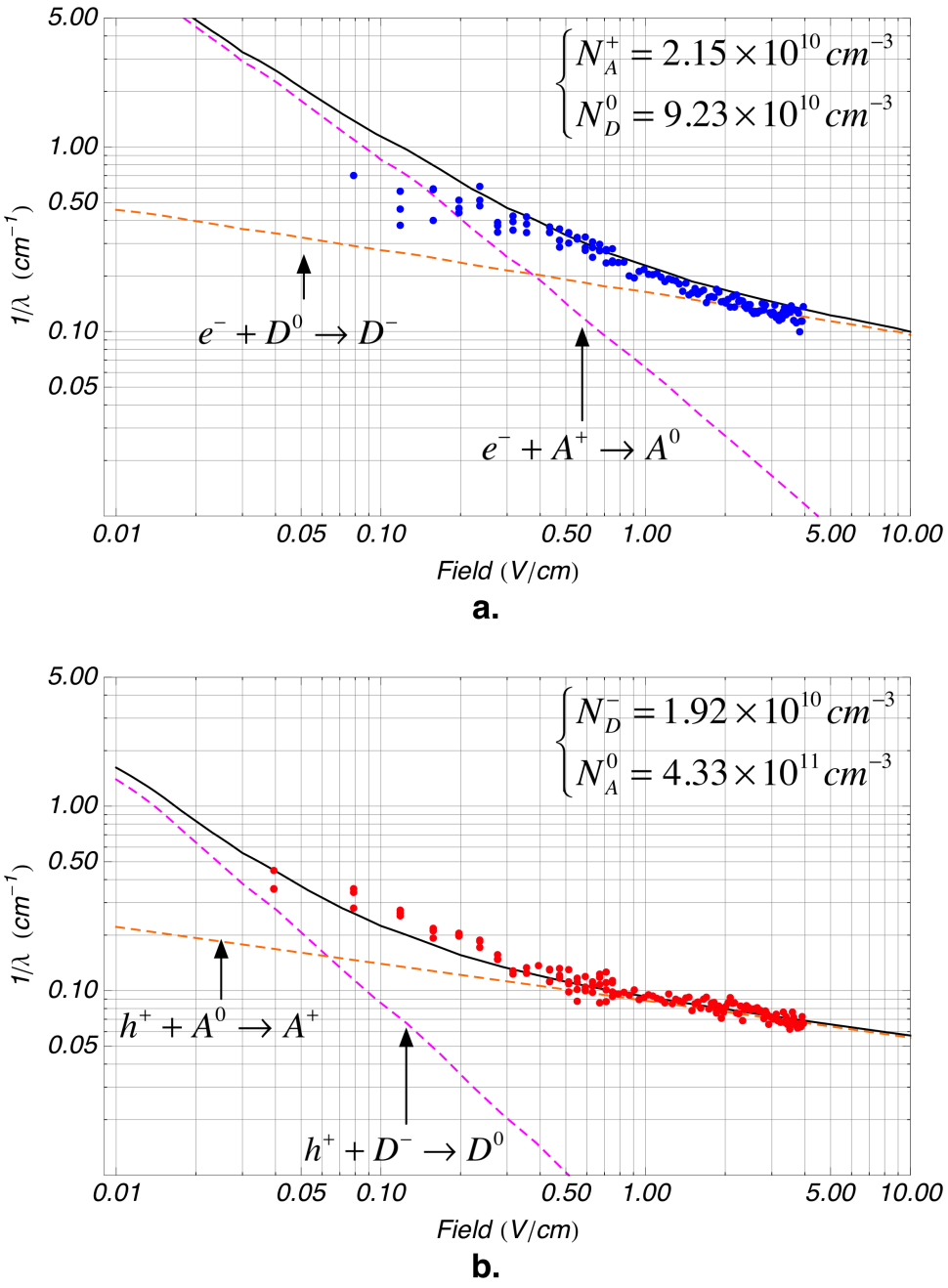

Figure 5.11: Comparison with carrier recombination data in CDMS detectors, from the "Phipps 2011" data set. Theory is for 2-species recombination between Coulomb trapping and neutral capture into anion (D^-/A^+) states. *a.* The inverse capture length for *electrons* with field. *b.* for *holes*. Note how there is a clear transition in the data for holes in this case.

From inspecting these data sets, we can learn a great deal. Note how there are "cross-over" points, where Coulomb-dominated capture at low fields gives way to neutral capture at higher fields. For the Shutt 1993 data, electrons have a cross-over point at 0.56 V/cm. Electrons in the Phipps data have a cross-over point at 0.38 V/cm. Holes in the Phipps data cross over at 0.064 V/cm. The fact that holes are much hotter than electrons for a given field is consistent with the substantially lower cross-over field for holes. This means that a field bias in between the cross-over point for electrons and for holes will be dominated by capture to acceptors for both species. Our typical CDMS bias of $F = 1\ V/cm$ lies above cross-over fields for both electrons and holes. This means that the detector *must* accrue more charged centers with time as recombination is dominated by neutral capture for both carrier species.

What is also reassuring is that the density of charged centers in the "Phipps 2011" data set is just about the same for both species. This coincides with the fact that charged centers must be net neutral to a high degree, as an imbalance of only $\approx 10^8\ cm^{-3}$ electron charges is enough to form serious internal fields that would affect $\sim V/cm$ biasing.

Consider the charged impurity ratio of equation 5.13. In the Phipps 2011 data set, we can find an average of N_D^- and N_A^+ as $2.04 \times 10^{10}\ cm^{-3}$. Comparing this to the quantity $\sqrt{N_A N_D} = \sqrt{(N_A^0 + N_A^+)(N_D^0 + N_D^-)}$, we find that

$$\frac{N_D^-}{\sqrt{N_A N_D}} \approx \frac{N_A^+}{\sqrt{N_A N_D}} = 0.09 \tag{5.66}$$

This expression represents the approximate steady-state ratios of charged impurities. Provided that a uniform "neutralization" process is adequately used when grounded, the actual neutralization mechanism should not affect this steady-state value, at least to first order. Interesting future experiments regarding neutralization may be had in applying a small nulling field when neutralizing crystals.

As per simple field dependence models, the prefactor $\sqrt{\left(\frac{C_{nA0}C_{pD0}}{C_{nA+}C_{pD-}}\right)}$ of equation 5.13 scales as $F^{6/5}$ – nearly linearly. To find the fraction of charged centers found in 5.66, our simulated rates would require a field of 104 mV/cm. While this may be a slightly high value for an assumed built-in field, we should note it is true that we readily find built-in fields in our grounded detectors of several 10's of mV/cm. The built-in field when grounded is an indicator of the number of charged, balanced centers in the bulk in the steady state. Yet for equation 5.66 to be valid, net neutrality must approximately hold true. This assumes the built-in steady state field which has developed in a detector is not too big, such that it is due to only a slight imbalance in charged densities. For greater precision, one can revisit the more complicated cases of rate balance as presented in appendix D.

So we have found a simple, 2-term model that describes the recombination we have measured in our detectors. When detectors are biased, it is now understandable why they must ionize, by a simple imbalance of capture rates. A complication under

bias is that free carriers drift apart from each other, and equation 5.13 is too ideal. What is required is a self-consistent solution to Poisson's equation and the continuity equations for electrons, holes, and all species of dominant recombination centers. Examples of the continuity equations (or "drift-diffusion" equations) can be found in the appendices.

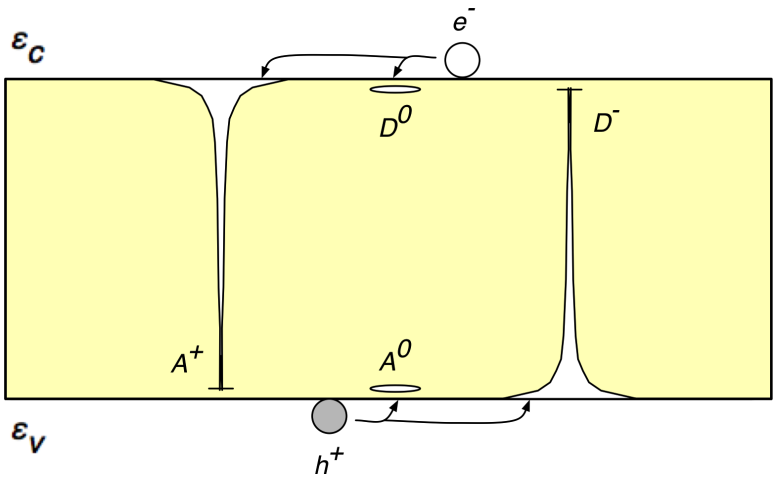

Figure 5.12: With our extracted Monte Carlo output for mean temperatures and velocities, two formulae from [15] give decent fits to data as in figures 5.10 and 5.11. This is consistent with a 2-carrier, 2-impurity model involving D^-/A^+ states.

We also note that impact ionization processes, introduced in the next chapter, should still have consideration. The probability to impact "neutralize" an anion state is an unusual circumstance and should be recognized. While we do not see evidence of these rates implicit in our capture length measurements, they may play a role in other phenomenology.

5.8 Concluding remarks

In this chapter, we have introduced CDMS detector data regarding the capture of carriers as a function of applied field. Viewed as processes following power laws, it appears that this capture data contains (at least) two capture processes. Based on the foundations of transport mechanisms described in previous chapters, we are now in a position to make some quantitative predictions about these capture processes.

The process which dominates electron and hole capture at higher measured fields appears to have an energy independent capture rate. We claim this as only the drift velocity appears to be contributing significantly to the field dependence of the inverse capture length. A strong candidate for this capture mechanism would be the presence of D^- and A^+ states. Capture into these states represents a large capture rate, considering that the resulting charged states are stable at our operating temperature. Variations of similar states may also be candidates (such as resonant states), but the mechanism of capture directly to the ground state makes a likely capture process.

All other known, nonradiative capture processes are grouped as cascade capture processes. In considering the capture mechanism which appears in our data at low fields, the pronounced field dependence and large cross section suggests the Coulomb capture as a candidate. So in this chapter, we visited the foundations of cascade capture theory, focusing primarily on contributions by Lax and Abakumov. We determined an equilibrium Coulomb capture cross section based on Abakumov's capture integral for single-quantum capture. We then applied his correction for hot-carrier effects in an applied field. These predicted capture rates are then consistent with both our capture data and the impurity concentrations of the crystals.

We propose an interpretation for our measured capture phenomena, using a simplified 2-carrier, 2-impurity model. From this model, it makes sense how detectors neutralize. When applied fields are low enough to substantially favor Coulomb capture, the population of charged sites will be suppressed. Balancing competing capture rates, as explained in appendix D, gives us an idea what charged densities should be in the steady state.

One outstanding question is a consideration of impact ionization. The impact ionization of neutral states – and moreover the somewhat unique situation of "impact neutralization" of anion states – is a big question. For typical D^- and A^+ states, the energy (field) threshold to impact "neutralize" should be quite low. Yet, we did not need to include this mechanism to interpret capture data. This suggests future capture measurements should be performed at higher drift fields.

In the next chapter, we will examine generation (e.g., ionization) processes. Generation processes include the impaction ionization of localized states (or even "impact neutralization" for anions). We will compare those predicted cross sections with the capture cross sections substantiated in this chapter.

Chapter 6

Generation Processes: Impact Ionization and Autoionization of Localized States

This chapter is concerned with possible generation (or "ionization") mechanisms in our CDMS detectors. Generation processes are those which generate carriers within the crystal. This may be due to either processes of existing populations of free carriers and localized sites (impact ionization), or processes involving only localized sites (auto-ionization / field emission).

Considering that the measured data presented in the last chapter seems to favor *capture* processes, it is an outstanding question to what extent we (should) measure generation processes. If neutral capture is dominated by anion states as we suggest, then the small binding energy of the "over-charged" bound carriers should be subject to future collisions and ionization of this state. Furthermore, we also find that the standard hydrogenic impurity (Rydberg) level should also be subject to a substantial level of impact ionization at only several V/cm.

In this chapter, we do not consider photo generation processes, contact/surface injection processes, or other radiation-based generation processes. We also do not consider impact ionization of carriers across the band gap.

6.1 Impact ionization of a localized state

The impact ionization of localized states is the process where a free carrier may collide with a neutral impurity state, imparting enough energy such that it may ionize a bound electron or hole. This is a generation process that could lead to avalanche. Note that this is different from the impact ionization of carriers *across the band gap* usually considered for fully depleted devices e.g., in detectors operating at $T = 77K$. We depict in this chapter its energy dependence.

Considering bound carriers in localized ground states, it should be noted that wavevector k is not a valid quantum number in bound-state Hamiltonians. While there is angular momentum, translational momentum does not enter. Therefore, when solving for the kinematics of this process, we do not consider momentum conservation [94]. As an aside, if we were instead to consider impact ionization *of carriers across the band gap*, those carriers lie in the continuum states and we *would* need to impose momentum conservation [94].

The ionization kinematics of this process are discussed at length in Landsberg [94]. However, a piece-wise simplification to these kinematics was proposed by Palmier [106], as reported in reference [107]. Palmier's result produces a fine approximation to Landsberg's work [94].

The expression of Palmier [106] is as follows.

$$\sigma(E) = \begin{cases} 0 & \text{for} \quad E < \varepsilon_i \\ \sigma_0 \left(\frac{E - \varepsilon_i}{\varepsilon_i} \right) & \text{for} \quad \varepsilon_i \leq E < 2\varepsilon_i \\ 2\sigma_0 \left(\frac{\varepsilon_i}{E} \right) & \text{for} \quad E \geq 2\varepsilon_i \end{cases} \tag{6.1}$$

Here, the threshold energy ε_i is the ground state energy of the localized state. Furthermore, σ_0 is the area of the localized ground state. The Bohr area is used for standard hydrogenic impurities. For anions, an effective area was determined by a wavefunction expectation integral. We use the wavefunction of Demkov, as in equation 5.7. We therefore adopt the area to be the following, noting it is simply proportional to $1/\kappa^2$.

$$\sigma_0 = \frac{\pi}{2} \left(\frac{\hbar^2}{2 m_c \varepsilon_{\text{anion}}} \right) \tag{6.2}$$

We now give a physical interpretation for this formula. We understand that an incident carrier must have an energy of at least the threshold energy to be able to ionize the carrier bound to the localized state. At threshold, the incident carrier can only barely ionize the bound carrier, and this probability is essentially still zero. As a function of increasing incident energy, more and more final states for this process are available, so impact ionization can occur more readily. At maximum ionization probability, which is at twice the threshold energy, the cross section for this interaction is the full area subtended by the ground state of the bound carrier (the Bohr area, for instance). As incident energy continues to increase much above twice the threshold energy, the incoming particle is becoming faster and spends less time in proximity with the bound state to ionize it with the same probability. So for the highest energy range, the cross section falls off by $1/E$. This behavior is the same as in the Bethe-Bloch formula for particle stopping powers in nuclear physics [108].

We should be careful to consider the effect of barrier lowering by the external field. We make adjustments to the impact ionization potential by adjusting the required ionization energy, as the following.

$$\varepsilon_i(F) = \varepsilon_i(0) - \delta U(F) \tag{6.3}$$

The δU is the barrier lowering, appropriate for the attractive potential that is localizing the bound carrier. For standard hydrogenic impurities, the Coulomb potential binds the localized carrier. For D^-/A^+ states which we will soon discuss, we consider the polarization potential keeps the extra charge bound (we exclude many-body effects).

The impact ionization cross sections of "standard" hydrogenic impurities is considered, as is the somewhat unusual condition of impact ionizing a stable, bound anion (D^-/A^+) state. The effective radius of the anion state is assumed to be the inverse wavevector associated with the binding energy.

6.1.1 Comparing results between capture and impact ionization

Now, we present the capture cross sections introduced in the last chapter, along with the impact ionization cross sections predicted by the theory in this section. Note that the hole drift velocities used in plot 6.2, unlike in chapter 5, are now taken directly from transport Monte Carlo output. Recall that the Monte Carlo simulation for transport had included the theoretical optical phonon effects which were not observed in the drift velocity data of chapter 4.

Nonetheless, one can still consider these predictions for holes at lowest fields (below the theoretical optical emission threshold for holes at $\approx 2.5\ V/cm$). This still does not explain why we should not observe more impact ionization in our hole data if the theory of this section is correct. It is possible that the 2-carrier, 2-impurity model of chapter 5 may simply be too simple to describe all the physics at hand, but these impact ionization processes require further study in any case.

These predictions suggest future measurements should be performed at higher electric fields. It is possible that deeper states are at work, which do not allow for impact ionization at lower carrier energies (fields). Another parsimonious hypothesis could be that simply that the Bohr area and/or the anion area mentioned above may not represent the best area estimates for impact ionization processes.

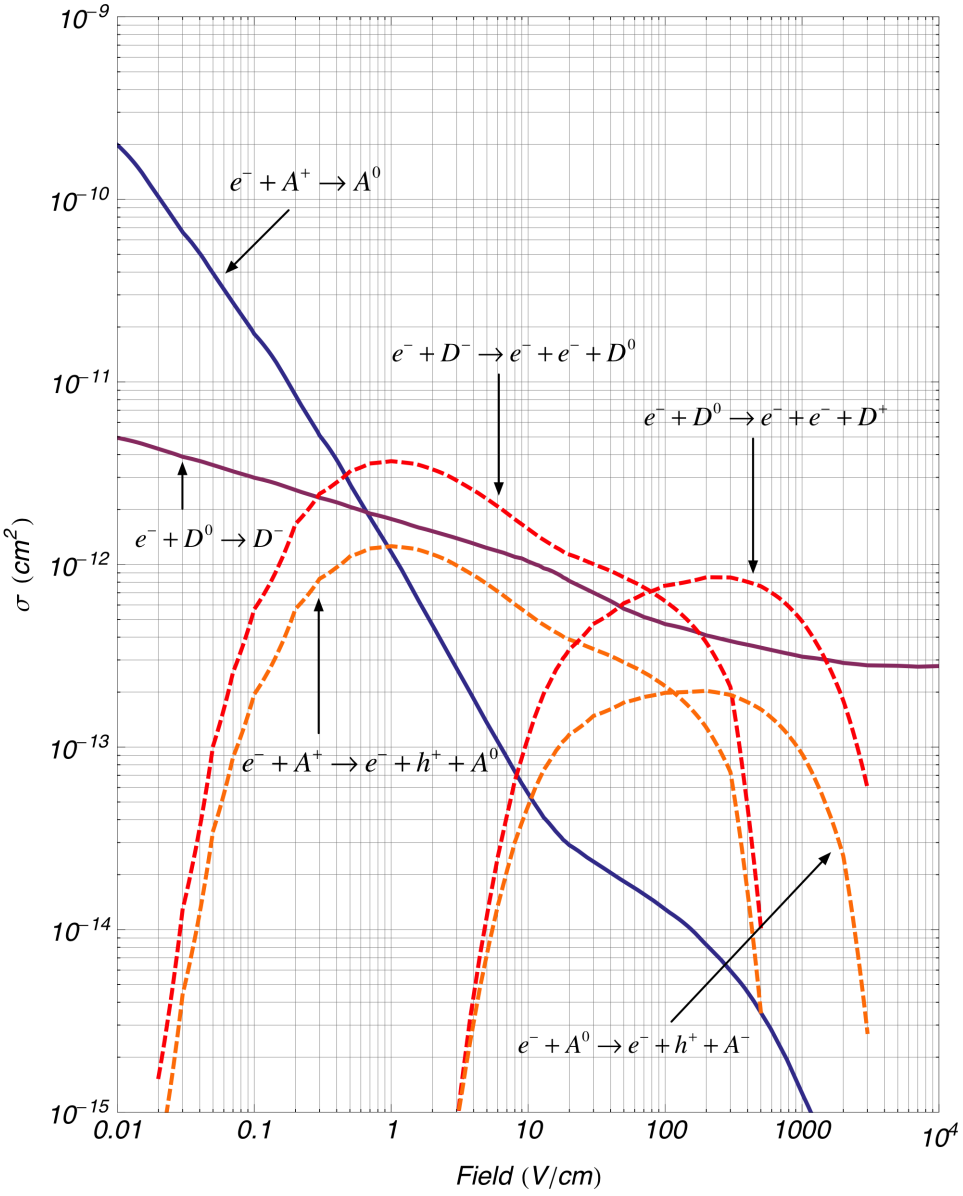

Figure 6.1: ELECTRONS: Predicted impact ionization cross sections for standard hydrogenic states and anions, along with the two capture processes depicting the capture data presented in the last chapter.

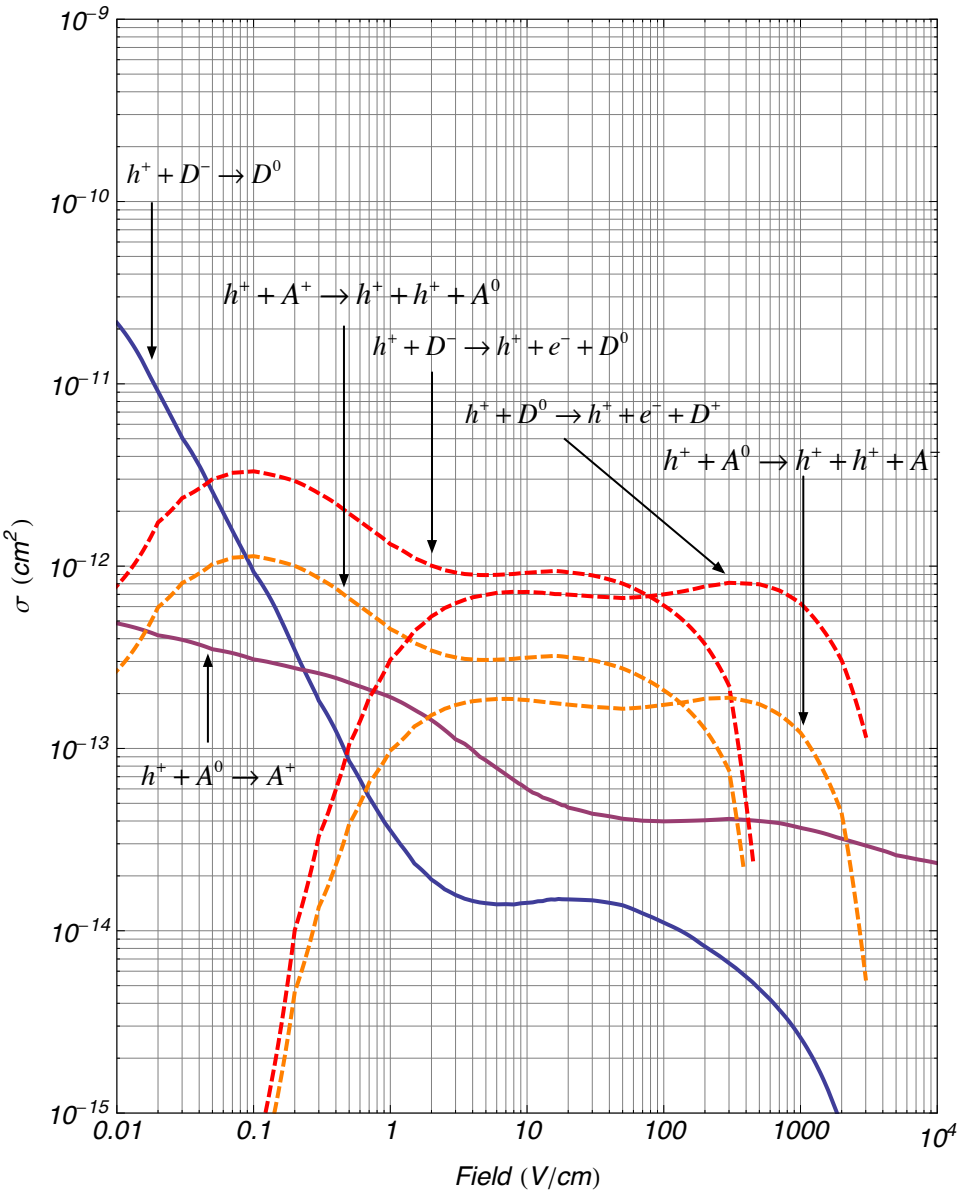

Figure 6.2: HOLES: Predicted impact ionization cross sections for standard hydrogenic states and anions, along with the two capture processes depicting the capture data presented in the last chapter. Note that hole drift velocities used in this plot, unlike in chapter 5, are now taken directly from Monte Carlo output and include theoretical optical phonon effects not observed in the drift velocity data of chapter 4.

6.2 Autoionization

We calculate escape rates for *autoionization* processes. This is where impurities may ionize due to only the presence of an external electric field. We calculate these rates for both hydrogenic donors and acceptors, as well as for assumed values of "anion" or "overcharged" donors (D^-) and acceptors (A^+) (assuming they exist). We construct the needed tunneling integral using the WKB approximation in 3D. We then show numerical results. Due to their small binding energy, anion states predictably ionize at modest fields of perhaps $15 - 30\ V/cm$.

The term "autoionization" is common in terminology regarding the physics of ionized gasses. In the terminology of solid-state physics, the term "field emission" is also used. The principle behind autoionization is that a neutral impurity can become ionized in an external electric field [61, 94]. The orbiting carrier is freed from the impurity due to the reduction of the local potential energy barrier. The charged impurity core is left behind, which contributes to space charge.

6.3 Overview of autoionization processes

While a charge carrier could escape an impurity *classically*, by overcoming the top of the potential barrier, this process must be assisted by thermal fluctuations. Consequently, this "thermionic" escape mechanism [109] is exceedingly small at low temperatures. Instead, the predominant autoionization mechanism is quantum mechanical tunneling [94]. The possibility for tunneling arises due to the tilted energy bands representing the external electric field. The escaping electron or hole is able to tunnel through the potential energy barrier formed between the localizing, attractive potential and the region of the external continuum states. This is depicted in figure 6.3.

We will describe autoionization as a tunneling process, using the WKB approximation. This is not unlike Gamow's model of alpha decay [110], in which an alpha particle tunnels out of the potential well of an atomic nucleus. For this case of semiconductor impurities, the applied electric field has a considerable effect on the height and width of the potential barrier. We wish to make this field dependence quantitative.

To this end, it is important that tunneling calculations are performed in all three spatial dimensions. In textbook examples, it is a common occurrence to find tunneling calculations performed in only one dimension. The exponential sensitivity of the tunneling process, however, creates a tremendous difference between the scenario of a three-dimensional spherical potential with an emission hot spot," versus a simple planar approximation representing a carrier escaping from one side of a box.

To describe our calculation more specifically, we use a one-dimensional WKB transmission coefficient to depict the radial transmission of a bound carrier. Next,

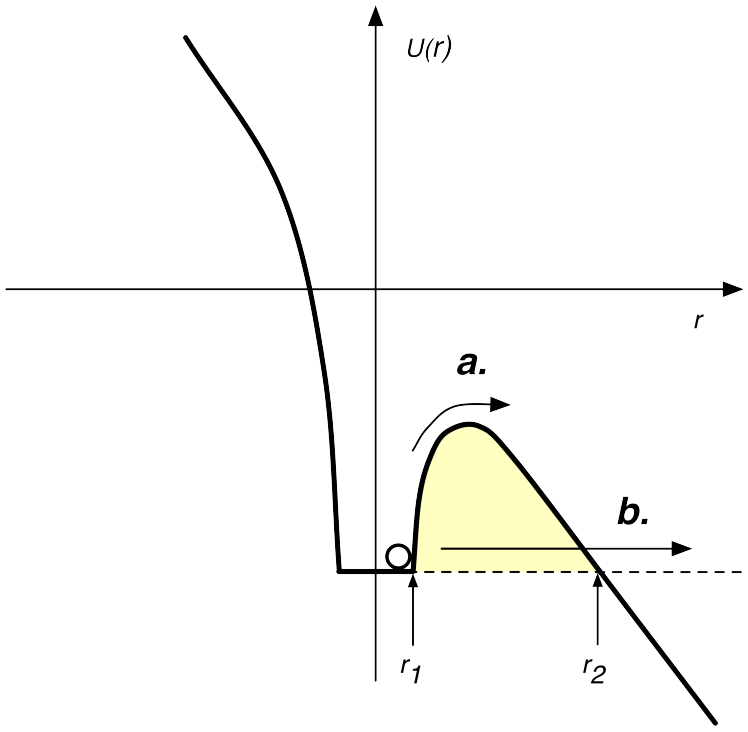

Figure 6.3: A neutral impurity in an external field, able to autoionize the bound carrier in its ground state. The radial coordinates of the classical turning points are at r_1 and r_2, representing respectively the "inner" and "outer" turning points. Process **a.** represents classical, "thermionic" emission, which is negligible. Process **b.** represents escape by tunneling through the potential barrier.

we use the virial theorem. The virial theorem allows us to use to compute an appropriate velocity, thereby a flux density, of the bound particle incident on the internal boundary of the potential barrier. Multiplying the internal flux density by the WKB transmission coefficient gives the escape probability for a particular radial direction. This radial escape probability is then integrated as a function of polar and azimuthal angles, over the internal area of the impurity. This results in an overall escape rate, or equivalently an expected lifetime, for the impurity to remain neutral in the electric field. This lifetime may be much longer than the age of the universe at very low fields. At fields where the barrier reduction nears the scale of the ground state energy, the transmission coefficient tends to unity. In this case, the impurity will quickly ionize at a timescale of roughly the impurity diameter divided by the bound carrier's virial velocity. The goal of this section is to make predictions for autoionization escape rates as a function of applied electric field.

6.4 The WKB approximation

The WKB (Wentzel-Kramers-Brillouin) approximation – or sometimes the WKBJ (Wentzel-Kramers-Brillouin-Jeffreys) approximation – is the workhorse calculation most often employed for semiclassical tunneling rates. Basically, the Schrodinger wavefunction, along with its matching conditions between barrier regions, is incorporated into an integral expression for the phase difference across a barrier. This is then incorporated into an approximation for the wavefunction's transmission coefficient. As the WKB approximation is derived in numerous textbooks such as [77, 111, 112], we summarize only the needed conclusions.

The wavefunction phase difference, g, is given by

$$g = \int_{r1}^{r2} \sqrt{\frac{2m_c}{\hbar^2}\left[U(r) - E\right]}dr \tag{6.4}$$

The WKB transmission coefficient is given by

$$T = \frac{e^{-2g}}{\left(1 + \frac{1}{4}e^{-2g}\right)^2} \tag{6.5}$$

It should be noted that these WKB conclusions are expressed in one dimension. As pointed out in the example of p. 279 of reference [113], the WKB transmission properties can be carried over directly into three dimensions. We neglect any angular contribution to the tunneling probability.

To apply the transmission coefficient, we need to know the escaping carrier's energy and the expression for the total potential energy. We also need to know the lower and upper limits of the phase integral, which represent the coordinates of the classical turning points of the potential. We will apply these quantities for specific cases, shortly.

Note that we also assume the bound carrier occupies *only one possible state*. We are considering low temperature operation, and we assume the level spacing to be large near the ground state of any type of impurity. Therefore, in this work we assume full occupation of the ground state.

6.5 The bound carrier flux density

We also need to know the incident flux density of the bound carrier while *inside* the neutral impurity. In some semiclassical sense, this represents the rate of the carrier "knocking on the door" in an attempt to escape the impurity. We define this flux, $\mathcal{F}$, from the definition for number current density.

$$\mathcal{F} = n_0 v_r \tag{6.6}$$

To define the bound carrier flux density, we see that we need to define both an *incident velocity*, v_r, as well as a *probability density*, n_0.

For the *incident velocity*, we make an approximation using the virial theorem. The virial theorem states that the expected kinetic energy of a particle in a central-force, power-law potential is related to its potential energy by a simple relation.

$$\langle E_k \rangle = \frac{m_c}{2} v_v^2 = -\frac{n}{2} \langle U(r) \rangle \tag{6.7}$$

where v_v is the virial velocity. The parameter n is the power-law exponent of the potential, which takes the typical form

$$U(r) = -\frac{A}{r^n} \tag{6.8}$$

We assume the velocity distribution is isotropic. We require only the outward, radial component of this velocity, which we are calling v_r. We therefore divide the virial velocity, v_v, by six, to account for three spatial degrees of freedom with both forwards and backwards motion.

$$v_r = \frac{1}{6} v_v \tag{6.9}$$

So we have established a useable incident velocity for the bound carrier inside the impurity potential well.

For the *probability density* of the bound carrier, we try two different approximations.

6.5.1 Probability density, approximation 1

We can define the probability density as a quantum mechanical quantity by introducing the bound carrier's wavefunction, ψ.

$$n_0(r_1) = \psi^*(r_1)\psi(r_1) \tag{6.10}$$

Here, some of our limitations start to become apparent. We intend to only use ideal and unperturbed wavefunctions, evaluated at the inner classical turning point.

6.5.2 Probability density, approximation 2

As a double-check, we introduce a second approximation for the probability density. Adopting a more semiclassical perspective, we fix the bound carrier density to be the inverse of the volume defined by the inner classical turning point.

$$n_0(r_1) = \left(\tfrac{4}{3}\pi r_1^3\right)^{-1} \tag{6.11}$$

For simplicity, we evaluate this volume at zero applied field. We can understand this approximation as an assertion that the bound carrier is somewhere strictly inside the classically allowed volume, defined by the ground state energy.

6.6 The autoionization rate, assembled

To summarize the framework we have established, here is the full integral expression for the autoionization escape rate.

$$\begin{aligned} {}^{1}\!/_{\tau} &= \int \mathcal{F}\ T\ dA \\ &= 2\pi \int_{0}^{\pi} \mathcal{F}\left(r_1,\theta\right)\ T\left(r_1,\theta\right)\ r_1^2\ sin\theta\ d\theta \end{aligned} \tag{6.12}$$

For generality, we have added the functional dependence of the radial and polar coordinates to both the flux density and transmission coefficient. We assume azimuthal symmetry until further notice. The "inner" classical turning point, r_1, is itself a function of angle, but this dependence was omitted for clarity.

Now we move on to set up this integral for specific cases, before evaluating it numerically.

6.6.1 Autoionization of shallow hydrogenic impurities

Here we examine the characteristics of shallow hydrogenic impurities for the purpose of calculating the autoionization escape rate.

The bound carriers for neutral hydrogenic states experience the binding form of the Coulomb potential. Therefore, the appropriate attractive potential at zero field is of course

$$U(r) = \frac{e^2}{4\pi\varepsilon\varepsilon_0 r} \tag{6.13}$$

Considering that we have ground state energy ϵ representing either ϵ_D for donors or ϵ_A for acceptors, it's straightforward to evaluate the WKB phase difference integral. Recall that $E = -\epsilon$.

$$g = \int_{r1}^{r2} \sqrt{\frac{2m_c}{\hbar^2}\left[-eFr\cos\theta - \frac{e^2}{4\pi\varepsilon\varepsilon_0 r} + \epsilon\right]}\, dr \tag{6.14}$$

We need to evaluate the radial coordinates of the inner and outer classical turning points. They are the following.

$$r_1 = \frac{\epsilon 4\pi\varepsilon\varepsilon_0 - \sqrt{4\pi\varepsilon\varepsilon_0}\sqrt{\epsilon^2 4\pi\varepsilon\varepsilon_0 - 4e^3 F\cos\theta}}{2eF\cos\theta(4\pi\varepsilon\varepsilon_0)} \tag{6.15}$$

and

$$r_2 = \frac{\epsilon 4\pi\varepsilon\varepsilon_0 + \sqrt{4\pi\varepsilon\varepsilon_0}\sqrt{\epsilon^2 4\pi\varepsilon\varepsilon_0 - 4e^3 F\cos\theta}}{2eF\cos\theta(4\pi\varepsilon\varepsilon_0)} \tag{6.16}$$

With this, the WKB transmission coefficient can be directly evaluated.

To consider the hydrogenic flux density, let us first consider the appropriate velocity. Application of the virial theorem gives us the outward radial velocity component.

$$v_r = \frac{1}{6}\sqrt{\epsilon/m_c} \tag{6.17}$$

For the bound state probability density in these hydrogenic states, we can use our quantum mechanical "approximation 1" by way of the following ground state wavefunction.

$$\psi = \frac{1}{\sqrt{\pi}a_B^{3/2}}e^{-r/a_B} \tag{6.18}$$

From this, we recover the following probability density.

$$n_0 = \frac{1}{\pi a_B^3}e^{-2r/a_B} \tag{6.19}$$

For our alternative semiclassical "approximation 2," we can find the following probability density from the Bohr radius. We apply this to hydrogenic acceptors as well, although they are not perfect representations of a Bohr atom.

$$n_0 = \left(\frac{4}{3}\pi a_B^3\right)^{-1} \tag{6.20}$$

Now that we have assembled the constituent parts, we can construct the full autoionization rate. We present results after constructing the same rate for the anion states.

6.6.2 Autoionization of anion states

For the case of anion states, a donor may overcharged with an extra electron to become a D^- minus state. Conversely, and acceptor may become and A^+ state with the addition of an extra hole. We assume that D^- minus and A^+ states are equally probable, with relatively equal energies. This may not be necessarily so, however. In any case, we proceed in a way similar to the previous section.

Bound carriers in the anionic potential experienced the polarization of the otherwise neutral atomic state. This again is the quadratic Stark effect, which as a reminder goes as

$$U(r) = -\frac{e^2\alpha}{2\left(4\pi\varepsilon\varepsilon_0\right)^2 r^4} \tag{6.21}$$

Therefore the WKB phase integral goes as

$$g = \int_{r1}^{r2} \sqrt{\frac{2m_c}{\hbar^2}\left[-eFr\cos\theta - \frac{e^2\alpha}{2\left(4\pi\varepsilon\varepsilon_0\right)^2 r^4} + \epsilon\right]}dr \tag{6.22}$$

Due to the $n = 4$ dependence of this potential, there are five roots when solving the barrier height expression for the classical turning points. It is best to solve numerically for the turning point coordinates, as needed.

We can evaluate the virial velocity assuming that the average potential energy is the ground state energy. Taking into account the $n = 4$ dependence of the potential, we recover an outward radial velocity of

$$v_r = \frac{1}{3}\sqrt{\epsilon/m_c} \tag{6.23}$$

For the bound carrier probability density, we again use the anion wavefunction of [93], as in section 5.2.1

In the framework of our "approximation 1," this gives the following probably density.

$$n_0 = \frac{\kappa B^2}{2\pi r^2}e^{-2\kappa r} \tag{6.24}$$

On the other hand, our semiclassical "approximation 2" gives a radius

$$r_a = \frac{1}{\sqrt{4\pi\varepsilon\varepsilon_0}}\left(\frac{e^2\alpha}{2\epsilon}\right)^{1/4} \tag{6.25}$$

resulting in an expression for probability density as

$$n_0 = \left(\frac{4}{3}\pi r_a^3\right)^{-1} \tag{6.26}$$

]

This is enough information that we can evaluate the full escape rates, as before.

6.7 Autoionization results

We present numerical results for the integrated autoionization rates we have constructed. These include the rates for hydrogenic donors(D^0) and acceptors(A^0). These also include the rates for anion donors(D^-) and acceptors(A^+). These four rates were calculated both under our "approximation 1" and "approximation 2" methods. This makes for eight distinct rates in total. We plot these rates as a function of applied electric field in figures 6.4 and 6.5.

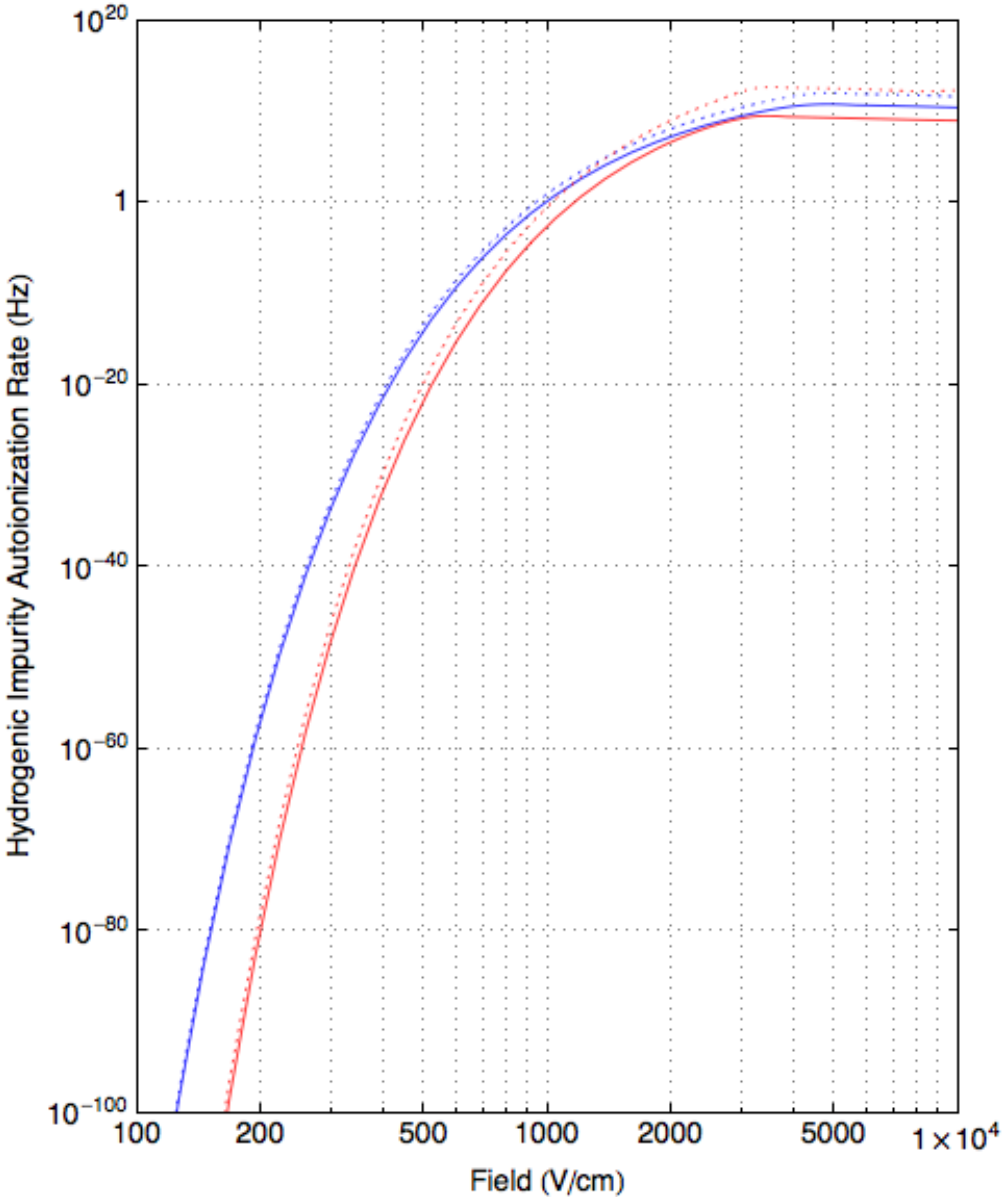

Figure 6.4: The autoionization rate is plotted for hydrogenic donors (D^0) in *blue*. Hydrogenic acceptors (A^0) are plotted in *red*. "Approximation 1" models, where bound carrier density was related to the wavefunction, are in *solid* lines. "Approximation 2" models, where bound carrier density was determined by the semiclassical volume, are in *dashed* lines.

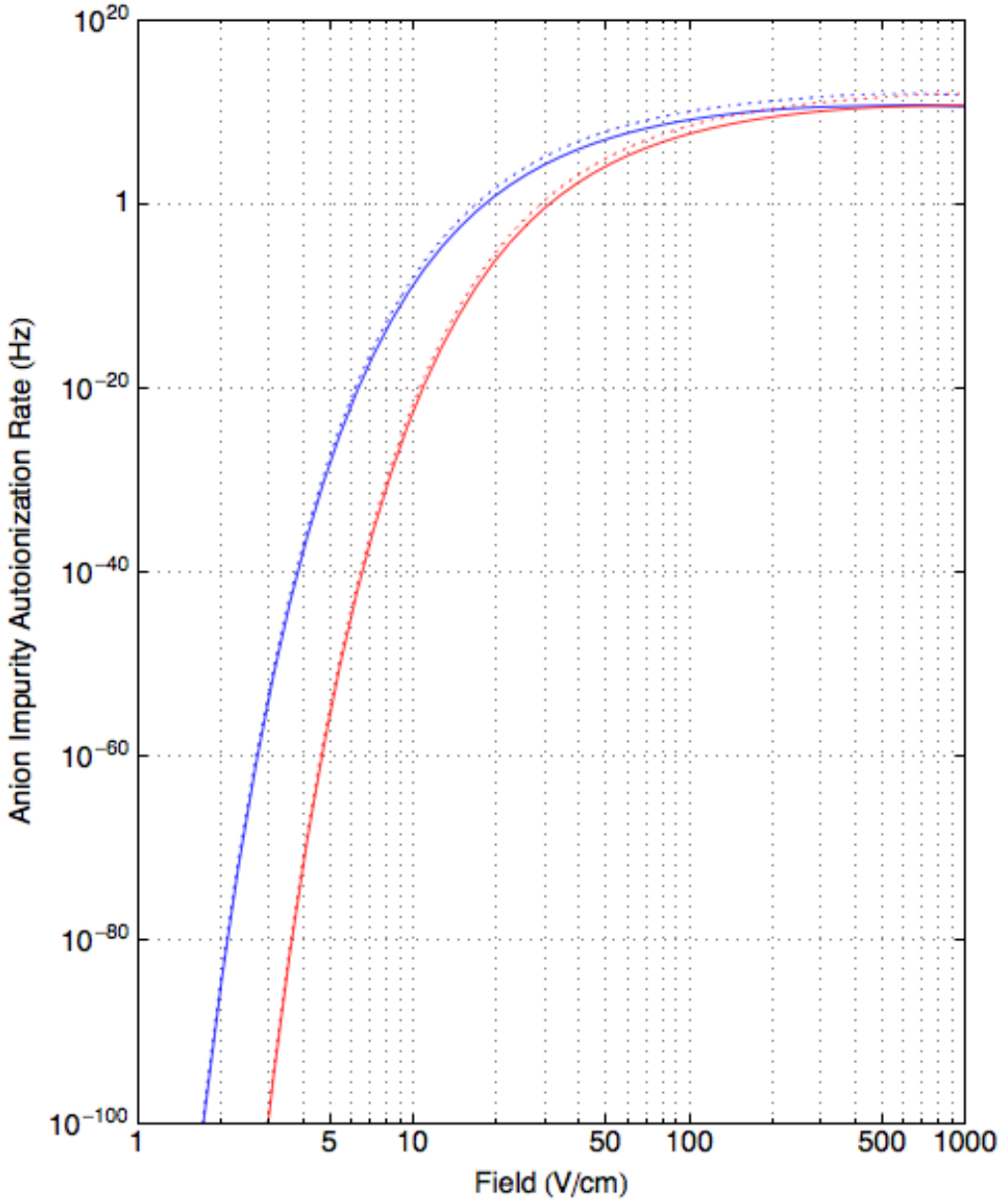

Figure 6.5: The autoionization rate is plotted for anion donors (D^-) in *blue*. Hydrogenic acceptors (A^+) are plotted in *red*. "Approximation 1" models, where bound carrier density was related to the wavefunction, are in *solid* lines. "Approximation 2" models, where bound carrier density was determined by the semiclassical volume, are in *dashed* lines.

Model	F (V/cm)@10^{-6} Hz	F (V/cm)@10^{-3} Hz	F (V/cm)@1Hz
D^0, approx. 1	695.415	815.329	985.461
D^0, approx. 2	669.836	779.731	933.35
A^0, approx. 1	845.053	975.886	1,156.11
A^0, approx. 2	773.036	879.963	1,021.95
D^-, approx. 1	11.584	14.115	18.035
D^-, approx. 2	11.026	13.298	16.725
A^+, approx. 1	19.832	24.164	30.877
A^+, approx. 2	18.884	22.776	28.638

Table 6.1: This is a selection of electric field strengths for all models, solved for at the indicated escape rates. For instance, hydrogenic donors (D^0) under "approximation 1" would autoionize with a lifetime of 10^6 seconds ($\sim$ 11.6 days) at 695.415 V/cm, 10^3 seconds ($\sim$ 17 minutes) at 815.329 V/cm, and 1 second at 985.461 V/cm.

Across a modest range of electric field, the autoionization rate spans a huge number of magnitudes. Even on log-log plots, it is difficult to determine some useful values. Therefore, we summarize some practical values of applied electric field in table 6.1.

6.8 Conclusions for autoionization

We have calculated plausible autoionization rates for hydrogenic and anion donors and acceptors. The exponential dependence of the tunneling transmission probability makes this problem extremely sensitive to the exact form of the local impurity potential. Non-ideal effects may have substantial impact on the actual rate. One source of nonideal effects may be the presence of dipole/multipole contributions due to neighboring, oppositely charged impurities. It is worthwhile to point out, though, that nonideal core effects in the potential are unlikely to contribute to transmission properties because the ground state energies of the impurities we have considered are so small (and therefore at high radius).

Fortunately, there is less sensitivity to the bound-state carrier density inside the impurity. The exact form of the bound carrier density within the impurity does not impact the escape rate nearly as much as the WKB transmission probability. Therefore, the use of approximations for the bound carrier density are rather forgiving in the calculation of the overall rate.

We see there is a great difference between the autoionization escape rates of hydrogenic impurities versus anion states. This is due to the extreme difference in their ground state energies. For this reason, the presence of anion states in our detector crystals should be an important consideration. If anion states exist in large abundance while the bulk is in a net neutral space-charge condition, then there is the possibility

of field ionization in a slightly peculiar way. That is, the "extra charge" bound to the anion state can be ripped away in rather modest electric fields. While the particular anion state would then itself become neutral after "ionizing," this would disrupt the global charge-compensated conditions of the crystal.

To some approximation, the exponential dependence of these escape rates allow us to consider autoionization as a sharp-threshold effect. We can approximately define a critical field where impurities will autoionize at some particular timescale, say over a few minutes. Then, we can approximate that impurities will have immediately autoionized in fields above this critical field. If impurities exist in adequate abundance, it may be that they maintain internal fields to be just at the critical field through ionizing. Where fields would have otherwise been very high, space charge can suddenly appear by way of autoionization. This may be in analogy to a phase transition. Rather than of some *critical temperature*, a *critical field strength* could greatly influence how impurities should be charged or uncharged.

Chapter 7

Conclusions: Evolving Space Charge

In seeking to find dark matter in the form of WIMPs, the Cryogenic Dark Matter Search operates high-purity germanium detectors at a temperature of 40 mK. Charge and phonons are measured from every particle event, allowing the discrimination of nuclear recoils (signal) from electron recoils (background). The voltage applied represents fields of only $\sim V/cm$, just enough to fully collect charge but small enough to keep secondary Luke-Neganov phonons from dominating the phonon signal. At this temperature, there are no ambient electrons or holes in the lattice, and the Fermi-level is therefore an ill-defined concept. Any free carriers are out of equilibrium by definition. Space charge can accrue, and degrade detector performance. A wide variety of detector phenomena have been measured and attributed to space charge with little more than an empirical understanding.

In this dissertation, we have established a framework necessary to begin to understand how space charge develops in our detectors. To do this, we revisited electron and hole transport from first principles. We implemented Monte Carlo techniques, both traditional and of our own design. We developed the ability to implement full anisotropy in both mass and scattering probabilities. We extracted useful quantities from our Monte Carlo simulations, including macroscopic averages as well as energy distributions we could use for future numerical integrals. We introduced some theory for cross sections and presented a number of possible recombination and generation mechanisms. While many models may still be possible, we present one simple model which was consistent with recombination data measured in CDMS detectors.

The work left to go is to put these concepts together to form an understanding of evolving space charge. This can be performed in a number of ways. A general approach is to consider the carriers and recombination centers as variables in *continuity equations*. Continuity equations (sometimes specifically "drift-diffusion" equations) depict volume densities of a species. They include dynamical terms like drift and diffusion, as well as generation and recombination rates.

In ending this dissertation, we demonstrate a self-consistent model which can be used to model some dynamical charge behavior of both free and bound charge in a macroscopic detector.

7.1 A self-consistent model

We briefly present how to put together a model depicting how space charge develops. The purpose is to solve continuity equations and Poisson's equation together, in a self-consistent manner. We use this method as it is general and free of assumptions regarding thermal equilibrium. We need to express a continuity equation for every species present in the model. For a two carrier, two impurity model, we need four continuity equations. Specifically regarding the model proposed in the last chapter, we would have the following.

$$\frac{d}{dt}\begin{pmatrix} n \\ p \\ N_D^- \\ N_A^+ \end{pmatrix} = \begin{pmatrix} v_{nd}\frac{\partial n}{\partial z} & +D_n\frac{\partial^2 n}{\partial z^2} & -R_n & +G_n \\ -v_{pd}\frac{\partial p}{\partial z} & +D_p\frac{\partial^2 p}{\partial z^2} & -R_p & +G_p \\ & & -R_A & +G_A \\ & & -R_D & +G_D \end{pmatrix} \tag{7.1}$$

where v_d is drift velocity, D is diffusion, an R term represents recombination of a species, G terms represent generation.

The expressions for G and R terms may depend on the concentrations of the other species. In this way, we introduce a coupling and form of feedback between the equations. For example, terms within R_n may go as $C \times n \times N$, where C is a rate per density as in the end of chapter A, n is the electron concentration, and N is the density of one of the other species (N_D^-, for example).

Within the expression for G for electrons and holes, we can also specify an illumination term for pair production by external radiation. At the surface, cosmic rays may induce an illumination of $g = 10^4\ s^{-1}cm^{-3}$.

We also know absolute densities of donors and acceptors. If we ignore the standard D^+/A^- charge state, we could simplify the concentrations as

$$\begin{aligned} N_D &= N_D^- + N_D^0 \\ N_A &= N_A^+ + N_A^0 \end{aligned} \tag{7.2}$$

We also need to solve Poisson's equation, which determines the electrostatics of the situation including the charge densities of the concentrations. With V for potential, we have the following.

$$\nabla^2 V = -\frac{\rho}{\varepsilon\varepsilon_0} = -\frac{1}{\varepsilon\varepsilon_0}\left(p + N_A^+ - n - N_D^-\right) \tag{7.3}$$

Also note that every differential equation requires a boundary condition. For continuity equations, most generally we can specify a *surface recombination velocity*,

as described in appendix F. For Poisson's equation at an electrode, we specify a sourced potential.

Using drift, diffusion, and the recombination terms described earlier in this work, we can self-consistently solve this system using a commercial finite-element solver (*Comsol*) for the underlying concentrations and potential as a function of position in a one-dimensional (1-inch) detector. Some results are in figure 7.1. Note that surface recombination velocities were set with a transmission of around 0.1 using equation F.12. Results are displaced in increments of $40k$ seconds, and total impurity densities were $N_D = 10^{11}\ cm^{-3}$ and $N_A = 10^{10}\ cm^{-3}$.

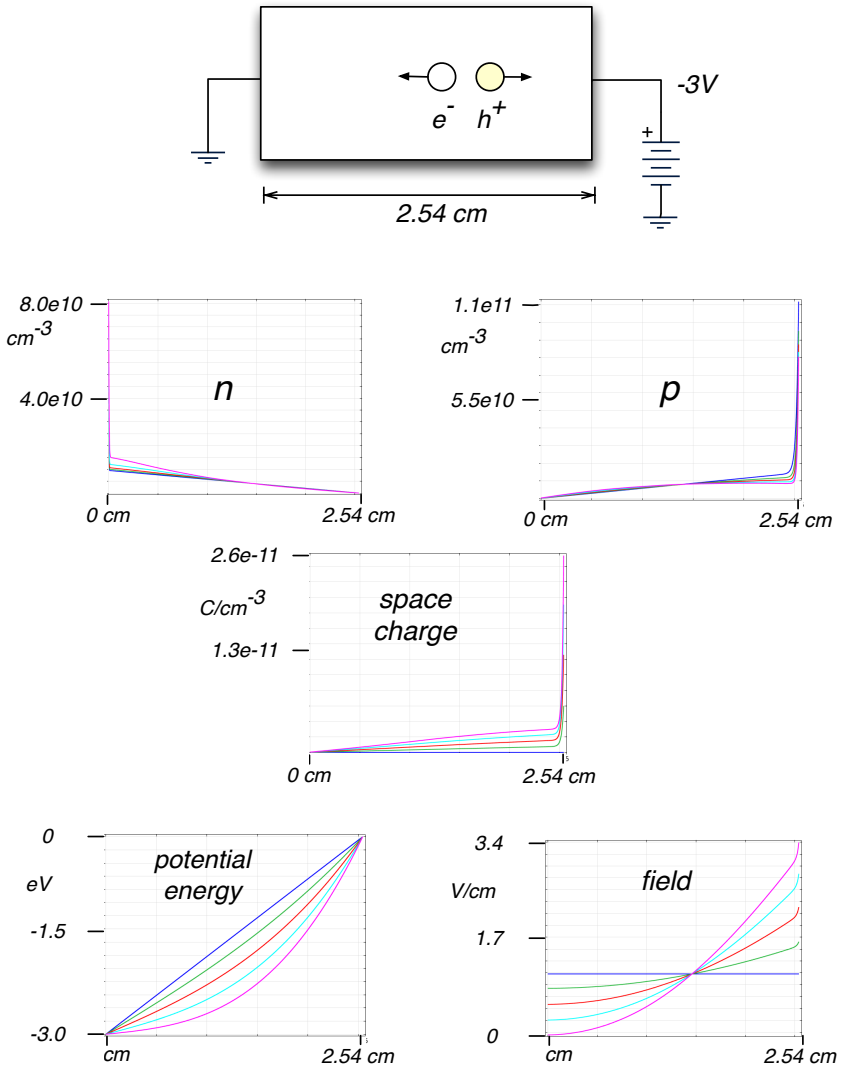

Figure 7.1: A depiction of a space charge model, self-consistently solving Poisson's equation and continuity equations for electrons, holes, and charged impurities. Note how space charge accrues and affects the internal potential and field.

7.2 Conclusion

We can see how the accumulation of space charge affects the internal potential and carrier concentrations. For more intuition, see the number of appendices regarding these equations. These profiles are a simple glance at how space charge should form in our detectors.

In concluding this dissertation, we show that we have formed a framework to solve the macroscopic problem of space charge evolution in CDMS detectors, incorporating the underpinnings we developed investigating transport processes and capture processes. One can imagine the number of ways this can be compared to experiment regarding detector phenomenology. Under this framework, a number of future models can be explored and validated.

Bibliography

[1] M. Persic. The universal rotation curve of spiral galaxies - i. the dark matter connection. *Mon. Not. R. Astron. Soc.*, 281:27–47, 1996.

[2] D. Clowe et al. A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006.

[3] C.N. Bailey. *The Cryogenic Dark Matter Search: First 5-Tower Data and Improved Understanding of Ionization Collection*. PhD thesis, Case Western Reserve University, 2009.

[4] S. Datta. *Quantum phenomena*. Modular series on solid state devices, v. 8. Addison-Wesley, Reading, Mass., 1989.

[5] B.K. Ridley. *Quantum Processes in Semiconductors*. Oxford University Press, 3rd edition, 1993.

[6] K.M. Sundqvist and B. Sadoulet. Detectors of the Cryogenic Dark Matter Search: Charge Transport and Phonon Emission in Ge $\langle 111 \rangle$ Crystals at 40mk. *J. Low Temp. Phys.*, 151:443–447, 2008.

[7] C. Jacoboni and L. Reggiani. The monte carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Rev. Mod. Phys.*, 55(3):645–705, Jul 1983.

[8] L. Reggiani, C. Canali, F. Nava, and G. Ottaviani. Hole drift velocity in germanium. *Phys. Rev. B*, 16:2781, 1977.

[9] B. Censier. *Etude et optimisation de la voie ionisation dans l'experience Edelweiss*. PhD thesis, IPN Orsay, 2006.

[10] K.M. Sundqvist, A.T.J. Phipps, C.N. Bailey, P.L. Brink, B. Cabrera, M. Daal, A.C.Y. Lam, N. Mirabolfathi, L. Novak, M. Pyle, B. Sadoulet, B. Serfass, D. Seitz, A. Tomada, and J. J. Yen. A measurement of electron and hole drift velocities in a germanium $< 100 >$ CDMS detector, at a temperature of 31 millikelvin. *AIP Conference Proceedings*, 1185:128–131, 2009.

[11] A.T.J. Phipps, K.M. Sundqvist, A.C.Y. Lam, and B. Sadoulet. Optically induced measurement of electron and hole drift velocities in a germanium $< 100 >$ CDMS detector at 50 mk. *Journal of Low Temperature Physics*, (in publication).

[12] M. V. Fischetti and S. E. Laux. Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys. *Journal of Applied Physics*, 80(4):2234–2252, 1996.

[13] C. Jacoboni F. Nava C. Canali and G. Ottaviani. Electron drift velocity and diffusivity in germanium. *Phys. Rev. B*, 24(2):1014–1026, Jul 1981.

[14] V.N. Abakumov, V.I. Perel, and I.N. Yassievich. *Sov. Phys. Semicond.*, 12:1, 1978.

[15] V.N. Abakumov V.I. Perel and I.N. Yassievich. *Nonradiative Recombination in Semiconductors*. Modern Problems in Condensed Matter Sciences, Vol. 33. North-Holland, 1991.

[16] S.A. Kaufman, K.M. Kulikov, and N.P. Likhtman. *Fiz. Tekh. Poluprovodn. [Sov. Phys.-Semicond.]*, 4:129, 1970.

[17] S.H. Koenig, R.D. Brown, and W. Shillinger. *Phys. Rev.*, 128:1668, 1962.

[18] T. Sanado, K. Matsushita, T. Ohyama, and E. Otsuka. Donor to acceptor electron transfer in germanium at low temperatures. *J. Phys. Soc. Jpn.*, 45(2):501–504, 1978.

[19] S.M. Sze and K.K. Ng. *Physics of Semiconductor Devices*. Wiley-Interscience publication. John Wiley & Sons, 3rd edition, 2007.

[20] K. Seeger. *Semiconductor physics: an introduction*. Advanced texts in physics. Springer, 9th edition, 2004.

[21] S. Dodelson. *Modern Cosmology*. Academic Press, 2003.

[22] A.R. Liddle and D.H. Lyth. *Cosmological inflation and large-scale structure*. Cambridge University Press, 2000.

[23] M. Kowalski et al. Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets. *Astrophys. J.*, 686(2):749–778, 2008.

[24] M. Milgrom. A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophys. J.*, 270:365–370, 1983.

[25] J.D. Bekenstein. Relativistic gravitation theory for the MOND paradigm. *Phys. Rev.*, D70:083509, 2004.

[26] S. M. Faber and J. S. Gallagher. Masses and mass-to-light ratios of galaxies. *Annual review of astronomy and astrophysics*, 17:135–187, 1979.

[27] F. Zwicky. Spectral displacement of extra galactic nebulae. *Helv. Phys. Acta*, 6:110–127, 1933.

[28] M. Girardi, G. Giuricin, F. Mardirossian, M. Mezzetti, and W. Boschin. Optical mass estimates of galaxy clusters. *The Astrophysical Journal*, 505(1):74, 1998.

[29] R.V. Wagoner. Big-bang nucleosynthesis revisited. *Astrophys. J.*, 179:343–360, 1973.

[30] C. Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[31] T. Padmanabhan. *Structure Formation in the Universe*. Cambridge University Press, 1993.

[32] C. Alcock et al. The MACHO project: Microlensing results from 5.7 years of LMC observations. *Astrophys. J.*, 542:281–307, 2000.

[33] J.P. Fillipini. *A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search*. PhD thesis, University of California, Berkeley, 2008.

[34] G. Bertone. *Particle dark matter: observations, models and searches*. Cambridge University Press, 2010.

[35] R. Bradley, J. Clarke, D. Kinion, L.J. Rosenberg, K. van Bibber, S. Matsuki, M. Mück, and P. Sikivie. Microwave cavity searches for dark-matter axions. *Rev. Mod. Phys.*, 75:777–817, 2003.

[36] J. D. Lewin and P. F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astropart. Phys.*, 6:87–112, 1996.

[37] M. Kamionkowski, K. Griest, G. Jungman, and B. Sadoulet. Model independent comparison of direct versus indirect detection of supersymmetric dark matter. *Phys. Rev. Lett.*, 74:5174–5177, 1995.

[38] T. Shutt. *A dark matter detector based on the simultaneous measurement of phonons and ionization at 20 mK*. PhD thesis, University of California, Berkeley, 1993.

[39] P. N. Luke. Voltage-assisted calorimetric ionization detector. *Journal of Applied Physics*, 64(12):6858–6860, 1988.

[40] B. S. Neganov and V. N. Trofimov. *Otkrytia i izobreteniya*, 146(215), 1985.

[41] N. E. Booth. Quasiparticle trapping and the quasiparticle multiplier. *Applied Physics Letters*, 50(5):293–295, 1987.

[42] C.A. Klein. Bandgap dependence and related features of radiation ionization in semiconductors. *J. Applied Physics*, 39:2029–2038, 1968.

[43] J. Lindhard, M. Scharff V. Nielsen, and P.V. Thomsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk.*, 33(10):1–42, 1963.

[44] T. Shutt et al. A solution to the dead-layer problem in ionization and phonon-based dark matter detectors. *Nucl. Instrum. Meth.*, A444:340–344, 2000.

[45] M. Lundstrom. *Fundamentals of Carrier Transport*. Cambridge University Press, 2nd edition, 2000.

[46] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Holt-Saunders International Editions: Science : Physics. Holt, Rinehart and Winston, 1976.

[47] P.Y. Yu and M. Cardona. *Fundamentals of Semiconductors: Physics and Materials Properties*. Graduate Texts in Physics. Springer, 4th edition, 2010.

[48] D. Vasileska and S.M. Goodnick. *Computational electronics*. Synthesis lectures on computational electromagnetics. Morgan & Claypool Publishers, 2006.

[49] R.F. Pierret. *Advanced semiconductor fundamentals*. Modular series on solid state devices. Prentice Hall, 2003.

[50] D.K. Ferry. *Semiconductor Transport*. Taylor and Francis, 2000.

[51] J.P. McKelvey. *Solid state and semiconductor physics*. Harper's physics series. Krieger Pub. Co., 1982.

[52] G.E. Pikus and G.L Bir. Effect of deformation on the hole energy spectrum of germanium and silicon. *Soviet Physics Solid State*, 1(11):1502–1517, 1960.

[53] C. Jacoboni and P. Lugli. *The Monte Carlo method for semiconductor device simulation*. Computational microelectronics. Springer-Verlag, 1989.

[54] S. Rodriguez-Bolivar, F.M. Gomez-Campos, F. Gamiz, and J.E. Carceller. Implications of nonparabolicity, warping, and inelastic phonon scattering on hole transport in pure si and ge within the effective mass framework. *Journal of Applied Physics*, 97:013702, 2005.

[55] H. Brooks. *Adv. Electr. Electron Phys.*, 8:85, 1955.

[56] H. Brooks. Scattering by ionized impurities in semiconductors. *Phys. Rev.*, 83:879, 1951.

[57] E. Conwell and V. F. Weisskopf. Theory of impurity scattering in semiconductors. *Phys. Rev.*, 77(3):388–390, Feb 1950.

[58] B.K. Ridley. Reconciliation of the conwell-weisskopf and brooks-herring formulae for charged-impurity scattering in semiconductors: Third body interference. *J. Phys. C. Solid State Phys.*, 10:1589–1593, 1977.

[59] C. Erginsoy. Neutral impurity scattering in semiconductors. *Phys. Rev.*, 79(6):1013–1014, Sep 1950.

[60] N. Sclar. Neutral impurity scattering in semiconductors. *Phys. Rev.*, 104(6):1559–1561, Dec 1956.

[61] M. Balkanski and R.F. Wallis. *Semiconductor Physics and Applications*. Oxford University Press, 2000.

[62] J.J. Sakurai and San Fu Tuan. *Modern Quantum Mechanics*. Addison-Wesley, 1994.

[63] J.M. Ziman. *Principles of the Theory of Solids*. Cambridge University Press, 2nd edition, 1979.

[64] C. Herring and E. Vogt. Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering. *Phys. Rev.*, 101:944–961, 1955.

[65] E.M. Conwell. *High field transport in semiconductors*. Solid State Physics: Advances in Research and Applications. Supplement. Academic Press, 1967.

[66] P. Lawaetz. Low-field mobility and galvanomagnetic properties of holes in germanium with phonon scattering. *Phys. Rev.*, 174(3):867–880, Oct 1968.

[67] G.L. Bir and G.E. Pikus. *Symmetry and Strain Induced Effects in Semiconductors*. Wiley, New York, 1974.

[68] M. Tiersten. Acoustic-mode scattering of holes. *IBM J. Res. Dev.*, 5:122–131, Apr 1961.

[69] F.L. Madarasz and F. Szumlowicz. Transition rates for acoustic-phonon—hole scattering in p-type silicon with nonparabolic bands. *Phys. Rev. B*, 24(8):4611–4622, Oct 1981.

[70] F. Szumlowicz. Calculation of optical- and acoustic-phonon–limited conductivity and hall mobilities for p-type silicon and germanium. *Phys. Rev. B*, 28(10):5943–5963, May 1983.

[71] J. M. Hinckley and J. Singh. Monte carlo studies of ohmic hole mobility in silicon and germanium: Examination of the optical phonon deformation potential. *Journal of Applied Physics*, 76(7):4192–4200, 1994.

[72] J.P. Wolfe. *Imaging phonons: acoustic wave propagation in solids*. Cambridge University Press, 1998.

[73] H. Ehrenreich and A.W. Overhauser. Scattering of holes by phonons in germanium. *Phys. Rev.*, 104(2):331–342, Oct 1956.

[74] L. Reggiani and M. Asche. *Hot-electron transport in semiconductors*. Topics in applied physics. Springer, 1985.

[75] C. Moglestue. *Monte Carlo simulation of semiconductor devices*. Chapman & Hall, 1993.

[76] M. V. Fischetti. Monte carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures-part i: Homogeneous transport. *IEEE Transactions on Electron Devices*, 38(3):634–649, March 1991.

[77] D. Vasileska and S.M. Goodnick. *Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling*. Springer, 2011.

[78] H.D. Rees. Calculation of steady state distribution function by exploiting stability. *Phys. Rev. Lett.*, 26A:416–7, 1968.

[79] V. Aubry-Fortuna, A. Broniatowski, and P. Dollfus. Field dependence of electron velocity in high-purity germanium at cryogenic temperatures. *AIP Conference Proceedings*, 1185:639–642, 2009.

[80] B. Cabrera, M. Pyle, R. Moffatt, K. Sundqvist, and B. Sadoulet. Oblique propagation of electrons in crystals of germanium and silicon at sub-kelvin temperature in low electric fields. *arXiv*, 1004.1233, 2010.

[81] F. Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill series in fundamentals of physics. McGraw-Hill, 1965.

[82] M.S. Shur. Influence of non-uniform field distribution and frequency limits of gaas field-effect transistors. *Electron Letters*, 12(23):615–616, 1976.

[83] M.S. Shur. *Physics of Semiconductor Devices*. Prentice Hall series in solid state physical electronics. Prentice Hall, 1990.

[84] G.H. Wannier. On the motion of gaseous ions in a strong electric field. i. *Phys. Rev.*, 2:281–289, 1951.

[85] R.E. Robson. *Introductory transport theory for charged particles in gases*. World Scientific, 2006.

[86] M. Shibuya. Hot electron problem in semiconductors with spheroidal energy surfaces. *Phys. Rev.*, 99:1189–1191, 1955.

[87] S.H. Koenig, M.I. Nathan, W. Paul, and A.C. Smith. Effect of high pressure on some hot electron phenomena in n-type germanium. *Phys. Rev.*, 118:1217–1221, 1960.

[88] M.I. Nathan, W. Paul, and H. Brooks. Interband scattering in n-type germanium. *Phys. Rev.*, 124(2):391–407, Oct 1961.

[89] M.I. Nathan. Anisotropy of the conductivity of n-type germanium at high electric fields. *Phys. Rev.*, 130(6):2201–2204, Jun 1963.

[90] H.A. Bethe. *Z Phys.*, 57:815, 1929.

[91] N. M. Haegel, J. W. Beeman, P. N. Luke, and E. E. Haller. Transient photoconductivity in ge:be due to be^+ formation. *Phys. Rev. B*, 39(6):3677–3682, Feb 1989.

[92] E.M. Gershenzon, A.P. Mel'nikov, R.I. Rabinovich, and N.A. Serebryakova. H-like impurity centers and molecular complexes created by them in semiconductors. *Sov. Phys. Usp.*, 23(10):684–698, Oct 1980.

[93] Yu.N. Demkov and V.I. Ostrovskii. *Method potenicalov nulevo radiusa v atomnoi fizike*. Izd. LGU, Leningrad, 1975.

[94] P.T. Landsberg. *Recombination in Semiconductors*. Cambridge University Press, 1991.

[95] S.H. Koenig. Rate processes and low-temperature electrical conduction in n-type germanium. *Phys. Rev.*, 110(4):986–988, May 1958.

[96] M. Lax. Cascade capture of electrons in solids. *Phys. Rev.*, 119(5):1502–1523, 1960.

[97] J.J. Thomson. *Philos. Mag*, 47:337, 1924.

[98] A.M. Stoneham. *Theory of Defects in Solids*. Oxford: Clarendon Press, 1975.

[99] P. Lal and P.T. Landsberg. Radiative cascade theory. *Phys. Rev.*, 140(A46), 1965.

[100] L. Reggiani and V. Mitin. Recombination and ionization processes at impurity centres in hot-electron semiconductor transport. *La Rivista del Nuovo Cimento (1978-1999)*, 12(11):1–90, 1989.

[101] L.P. Pitaevskii. *Sov. Phys. JETP*, 15(919), 1962.

[102] B.K. Ridley. *Quantum Processes in Semiconductors*. Oxford: Clarendon Press, 4th edition, 1999.

[103] D.J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 1999.

[104] A.L Efros, B.I. Shklowskii, and J.Y. Yanchev. *Phys. Status Solidi B*, 50:45, 1972.

[105] E. Otsuka, K. Murase, and J. Iseki. Electron scattering by neutralized acceptors in germanium i.gallium and indium. *Journal of the Physical Society of Japan*, 21(6):1104–1111, 1966.

[106] J. F. Palmier. Impurity ionization in n-type germanium. *Phys. Rev. B*, 6:4557–4571, 1972.

[107] E.A. Gutiérrez-D., M.J. Deen, and C.L. Claeys. *Low Temperature Electronics: Physics, Devices, Circuits, and Applications*. Electronics & Electrical. Academic Press, 2001.

[108] J.S. Lilley. *Nuclear physics: principles and applications*. The Manchester physics series. J. Wiley, 2001.

[109] C.R. Crowell. The Richardson constant for thermionic emission in Schottky barrier diodes. *Solid-State Electronics*, 8(4):395 – 399, 1965.

[110] G. Gamow. Zur Quantentheorie des Atomkernes. *Zeitschrift für Physik A Hadrons and Nuclei*, 51:204–212, 1928. 10.1007/BF01343196.

[111] A.F.J. Levi. *Applied Quantum Mechanics*. Cambridge University Press, 2006.

[112] S. Gasiorowicz. *Quantum Physics*. Wiley, 2003.

[113] D. Bohm. *Quantum Theory*. Prentice-Hall Physics Series. Dover Publications, 1951.

[114] B. Cabrera. Notes on electron-phonon scattering. CDMS internal note, 1993.

[115] P.J. Price. The theory of hot electrons. *IBM J. Res. Dev.*, 14(1):12–24, Jan 1970.

[116] G. Wang. *The Cryogenic Dark Matter Search and Background Rejection with Event Position Information*. PhD thesis, Case Western Reserve University, January 2005.

[117] B.G. Streetman and S.K. Banerjee. *Solid state electronic devices*. Prentice Hall series in solid state physical electronics. Pearson Prentice-Hall, 6 edition, 2009.

[118] J. Clarke and A.I. Braginski. Introduction. In J. Clarke and A.I. Braginski, editors, *The SQUID Handbook*, volume 1. Wiley, 2004.

[119] D. Drung and M. Mück. Squid electronics. In J. Clarke and A.I. Braginski, editors, *The SQUID Handbook*, volume 1, chapter 4, pages 127–170. Wiley, 2004.

[120] R.P. Feynman, R.B. Leighton, and M.L. Sands. *The Feynman Lectures on Physics: Commemorative Issue*. Number v. 1-3 in The Feynman Lectures on Physics. Addison-Wesley, 1989.

[121] M. Tinkham. *Introduction to superconductivity*. Dover books on physics and chemistry. Dover Publications, 2nd edition, 2004.

[122] T. Van Duzer and C.W. Turner. *Principles of Superconductive Devices and Circuits*. Prentice Hall, 2nd edition, 1999.

[123] A.M. Kadin. *Introduction to superconducting circuits*. A Wiley Interscience publication. Wiley, 1999.

[124] M. Tinkham. *Introduction to Superconductivity*. Dover, 2 edition, 1996.

[125] B.A. Hines, K.M. Sundqvist, D. Seitz, and M.E. Huber. Flux-coupled direct feedback in a squid amplifier. *IEEE Trans. Appl. Supercond.*, 21(1):262–266, 2011.

[126] R.A. Webb, R.P. Giffard, and J.C. Wheatley. Noise thermometry at ultralow temperatures. *Journal of Low Temperature Physics*, 13(3):383–429, 1973.

[127] J. Clarke, W.M. Goubau, and M.B. Ketchen. Tunnel junction dc squid: Fabrication, operation, and performance. *Journal of Low Temperature Physics*, 25(1):99–144, 1976.

Appendix A

The "Semi-Ballistic Approximation" for Capture Processes

In this appendix, we present results primarily from what we call our "semi-ballisitic approximation." It was a large attempt to incorporate a simple (step function) sticking probability within a cascade capture model similar to that of Lax. The intent was that $T_L \approx 40\ mK$ lattice temperature would represent an adequately low range of nearly-excited states, such that this would match Abakumov's work in the similar limit. This model provided the possibility to include any capture or generation mechanism, for arbitrary fields. It used the bulk steady-state distributions exported from Monte Carlo transport simulations. This approach was appealing in that it always decoupled "carrier temperature" from "lattice temperature."

At the end of this appendix, we show that we had some discrepancy in the field dependence of the Coulomb cross section derived from this method versus the Abakumov method of chapter 5. We believe our "Lax-like" sticking probability and implementation of energy distributions produce assumptions which are too strict to allow capture to occur as it should.

We make three points about the merits of this appendix: *1)* The exposition of the constituent mechanisms is educational for understanding the myriad of capture kinematics, *2)* the impact ionization predictions presented here are still those calculated by only with the underpinnings described in chapter 6, and *3)* some relative comparisons can be made between capture processes calculated under this method.

To point out a tempting future direction, the Abakumov-based methods described in sections 5.3.4 and 5.3.4 can be invoked numerically, using transport Monte Carlo results directly.

The driving topic: extending the Lax treatment to *non-equilibrium* cases
The critical aspect of our treatment lies in the fact that we are looking at processes

out of equilibrium. Abakumov [15] points out the subtle fact that a mean-free-path, λ should be replaced by $v\tau_E$ where v is the (total) velocity and τ_E is the energy relaxation time. This is the right idea, since an energy relaxation time need not imply equilibrium with the lattice. Analytical approximations for these quantities may still be difficult for extremely hot carriers, especially since the relaxation time appropriate while in the continuum steady state may be altered while in the proximity to a local capture potential. So we wish to construct the framework to predict processes without the reliance of thermal equilibrium between carriers and lattice. As carriers become increasingly "hot" with applied field, the energy dependence results in a field-dependence.

A.1 Introducing the general approach and macroscopic quantities

Although our treatment resorts to numerical techniques, we present a simple model to first introduce the general approach.

Recall Thomson's capture cross section,

$$\sigma = \frac{4\pi}{3}\frac{r_c^3}{\lambda} = \frac{4\pi}{3}\frac{r_c^3}{v\tau} \tag{A.1}$$

In the upcoming sections, we perform integrals that provide an *energy-dependent* cross section, taken as an expectation integral over the potential change of an impurity center. In order to take into account the possible dependence of velocity and relaxation time on the radius, while maintaining cylindrical symmetry, we propose to generalize equation A.1 as

$$\sigma(E) = \int_0^{r_c} \frac{4\pi r^2}{v(E,r)\tau(E,r)} dr \tag{A.2}$$

Here, E is the total energy of the carrier, representing kinetic energy in the continuum steady state. We will explore what are the appropriate limits ("r_c") to incorporate for these integrals, as well as the appropriate $\frac{1}{v\tau}$. Note that, in presenting this integral, one of our implicit approximations is that of a *uniform probability* for incident carriers to occupy any part of the trapping volume.

A.2 The real-space volume of a recombination center: the critical radius

If we considered the simple equilibrium Thomson cross section for a charged center, the critical radius which would depict the capture "sphere of influence" is found by

inverting the Coulomb potential.

$$r_c = \frac{Ze^2/(4\pi\varepsilon\varepsilon_0)}{3/2 kT} \tag{A.3}$$

Considering a temperature of $T = 40\ mK$, this corresponds to a radius of 17 um and a cross section of $10^{-9}\ cm^2$. Another consideration is to ensure that a carrier of zero total energy has enough kinetic energy ($E_k \geq \frac{1}{2}mv_s^2$) while over an attractive (negative) potential to emit an acoustic phonon. Recall figure 5.5 for a reminder of energy conventions over an attractive center. Setting this threshold energy equal to the Coulomb potential also results in a critical radius of about 17 um. While this is not out of the realm of possibility, but it does represent a large cross section. A key insight is to realize that, for our CDMS case, *the thermal contribution to the capture volume limit is typically matched or exceeded by influences of the external potential itself.* This insight was inspired by the Poole-Frankel effect, where the barrier-lowering by an external field of a neutral trap enhances thermionic emission of the occupying carrier.

A.2.1 Using a barrier lowering, corrected by kT_L

Recalling chapter 5, the external field modifies the local potential. We account for the effects of the external electric field by a correction factor to the local potential. Note, that we define this *correction* to the potential as a *positive* term.

$$\delta U_0 = -U(r_b) = A(n+1)\left[\left(\frac{\mathrm{eF}}{An}\right)^{\frac{-1}{1+n}}\right]^{-n} \tag{A.4}$$

We see that we can now express the classically bound states within the potential as contained up to an energy of $\approx |U(r)| - \delta U_0$.

We do not consider the specific functional forms of different "sticking probability" functions proposed through the years. As a reminder, we take the oft-used approximation that a carrier will be assumed "stuck" into the local potential as long as it has found itself in a bound state lower than the continuum levels by an energy kT_L, where T_L is the *lattice* temperature.

We account for the possibility of re-absorbing a phonon and ionizing out of the bound state region by adding a term onto the local-potential correction.

$$\delta U = \delta U_0 + kT_L \tag{A.5}$$

A.2.2 A volume defined by energy

Now, we are at a point to find the real-space limiting distance to the effective trapping volume. Once again, our understanding before was that this volume is

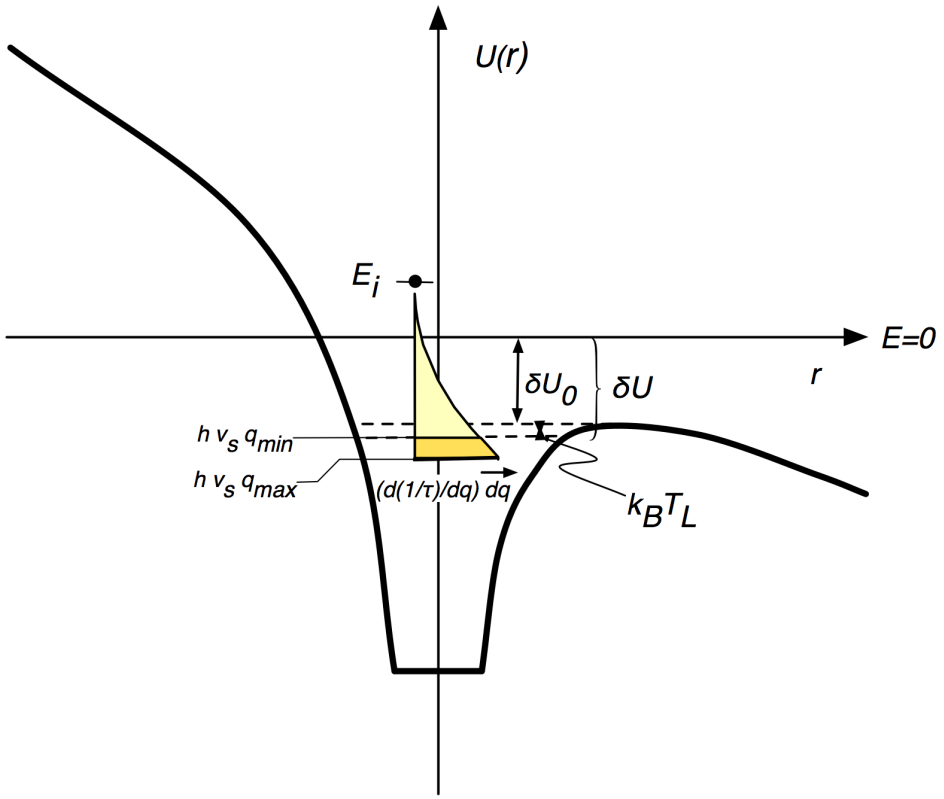

Figure A.1: A schematic of a hydrogenic potential for capture, including barrier lowering and the range of final-state carrier energies contributing to carrier capture. The differential scattering rate is also depicted as a function of energy.

determined by a critical radius where the potential is deep enough to permanently bind a carrier. Assuming a spherical case just for now,

$$volume = \frac{4}{3}\pi r_c^3 \tag{A.6}$$

We now define a critical capture radius by potential energy arguments as follows.

$$|U_{tot}(r_c)| = \delta U = \delta U_0 + kT_L \tag{A.7}$$

Provided $r_c > r_{min}$, this finally gives a useful expression for our effective trapping critical radius as

$$r_c = \left\{ \left[\left(\frac{\mathrm{eF}}{An} \right)^{-\frac{1}{n+1}} \right]^{-n} + \frac{\mathrm{kT}_L}{A} \right\}^{-1/n} \tag{A.8}$$

We have determined the "critical radius" for our volume estimate. Although our integration techniques in the end are numerical, with the limits of integration

determined automatically by energy conservation, the principle of this critical radius and its corresponding "sphere of influence" is conceptually an important one.

A.3 Accounting for an appropriate relaxation length

Here, we account for the fact the the relaxation length is dependent on energy and therefore position.

Within the sphere of influence, we need to know whether the carrier will energetically relax such that it can become bound. In a naive treatment, we might think that the typical bulk scattering rate would be appropriate here. This is an approximation appropriate for an order of magnitude calculation, and it is how the Thomson scattering cross section was found. However, we perform a slightly more sophisticated treatment in that we will consider the scattering relaxation time including only those final states which result in a permanently bound carrier.

We need to find the energy relaxation length,

$$\lambda_E = v\tau_E \tag{A.9}$$

Here, v is the velocity, and τ_E is the energy relaxation time. Recall, however, that we are in the "Thomson limit" where a single scatter accounts for the entire energy relaxation.

A.3.1 Accounting for the relaxation time

Here we account for the relaxation time, taken as an inverse phonon emission rate. For these energy and position dependent integrals, we maintain an "isotropic spirit" to reduce the complexity of this problem to manageable levels. To this end, we group all acoustic emission processes to separate phonon branches into a single coupling rate, separate for electrons and for holes. This rate is based on the isotropic deformation potential and average speed of sound available in the literature [7, 12, 5]. Similarly, we also group the optical branches together for a single optical emission rate [7], separate for electrons and for holes.

Relaxation into bound states by *acoustic* phonon emission

With these fully isotropic assumptions, the relaxation rate is the phase-space integral across the range of possible phonon wavevectors [45].

$$\frac{1}{\tau} = \frac{1}{4\pi^2} \int_{q\,\min}^{q\,\max} I^2(N_q + 1) \left(\frac{\pi m_{dos} \Xi^2}{\hbar \rho v_s p} \right) q^2 dq \tag{A.10}$$

Here, m_{dos} is the density-of-states mass, N_q is the phonon occupation, Ξ is the deformation potential, ρ is the density of germanium, p the carrier momentum. We will revisit q_{min} and q_{max} momentarily, but in the bulk case they are $q_{min} = 0$ and $q_{max} = 2k(1 - \frac{v_s}{\hbar k/m_c})$. As we know, phonon occupation is nearly zero in our temperature range, for quite nearly all phonon wavevectors of interest. We keep this factor around as a correction to acoustic emission processes, as we know such capture is dominated by carriers energies within a magnitude of the speed of sound energy, $\frac{1}{2}m_c v_s^2$. In our condition, $kT_L \approx \frac{1}{2}m_c v_s^2$, so it doesn't hurt to retain an approximation for N_q. We remove N_q from within the integral, but make an approximation for it by evaluating it at a expectation value for q.

$$N_q \approx \left\{ \exp\left[\hbar v_s \left(\frac{3}{4} \frac{q_{\max}^3}{(q_{\max}^2 + q_{\max} q_{\min} + q_{\min}^2)} + \frac{3}{4} q_{\min} \right) \frac{1}{kT} \right] - 1 \right\}^{-1} \tag{A.11}$$

At the threshold carrier energy able to trap, the use of this approximation compared to the actual, numerical integral is good to within 10%, given our working temperature. Setting $N_q = 0$ gives an error of over 30%.

We also make the common assumption that the electron wavefunction overlap factors are $I^2 = 1$ [5]. Considering holes, the incident hole is assumed to occupy the heavy (top) band, and since the density of states for *inter*-band transfer is small, the overlap factor is taken to be $I^2 = \left\langle \frac{1}{4}\left[1 + 3\cos^2(\theta)\right] \right\rangle = 1/2$.

We re-express the incident momentum, p, as a kinetic energy term, $\sqrt{2mE_k}$. We find our simplified rate to be the following.

$$\frac{1}{\tau} = I^2(N_q + 1) \frac{\Xi^2 m_{dos} (q_{max}^3 - q_{min}^3)}{12\sqrt{2}\pi v_s \rho \hbar \sqrt{E_k m_c}} \tag{A.12}$$

So we have determined a useful expression for capture rate that we can combine with the velocity of equation 5.18 for an appropriate expression for relaxation length as a function of initial carrier energy and strength potential.

Next, we will see that there is a lot to learn by inspecting the phonon limits,q_{min} and q_{max}, present in this special capture rate.

A.4 Phonon limits in the case of acoustic phonon emission

In our consideration of a capture rate, we consider the emission rate relevant *only* to permanently bound states. The form of the integral rate does not change, but we revisit the limiting phonon amplitudes that may contribute to carrier capture.

To enter a permanently bound state, a passing carrier of total energy E must emit *at least* a phonon such that δU amount of energy, by equation A.5, is lost. This sets the minimum phonon wavelength for this rate.

$$q_{min} = \frac{E + \delta U}{\hbar v_s} \tag{A.13}$$

Momentum and energy conservation constrain, as always, the maximum phonon wavevector. This corresponds to complete backscatter and is given by

$$q_{max} = 2k(1 - \frac{v_s}{\hbar k/m_c}) = 2\frac{\sqrt{2m_c E_k}}{\hbar}(1 - \frac{m_c v_s}{\sqrt{2m_c E_k}}) \tag{A.14}$$

which is approximately

$$q_{max} \approx 2k \tag{A.15}$$

as long as the carrier is traveling considerably faster than the speed of sound threshold. These phonon limits are important as they introduce limits in the incident carrier energy allowed for capture.

A constraint on the incident carrier energy, using the acoustic phonon limits

For any phonon emission rate to contribute to carrier capture, we must observe $q_{min} < q_{max}$. We inspect the limiting case by setting wavevectors equal to each other. For this purpose, we approximate the maximum allowed acoustic phonon as $q_{max} \approx 2k$.

Considering the relation for capture-causing emission of an acoustic phonon,

$$q_{min} \leq q_{max} \tag{A.16}$$

we make the approximation

$$\frac{E + \delta U}{\hbar v_s} \leq \frac{2}{\hbar}\sqrt{2m_c E_k} = \frac{2}{\hbar}\sqrt{2m_c(E - U)} \tag{A.17}$$

Recall that U is a negative quantity for attractive potentials.

Note figure A.1 as an example. An initial, free carrier has some probability to emit into a bound state between energy ranges determined by our definitions of phonon wavevectors q_{min} and q_{max}.

We now determine *limiting* values for the *energy of the incident carrier*. Consider the limiting case where $q_{min} = q_{max}$. Solving, we find the following quadratic roots for total energy of the incident carrier.

$$\begin{aligned} E_{\min} &= (4m_c v_s^2 - \delta U) - 2\sqrt{2}\left(2\left(m_c v_s^2\right)^2 + m_c v_s^2(|U| - \delta U)\right)^{1/2} \\ E_{\max} &= (4m_c v_s^2 - \delta U) + 2\sqrt{2}\left(2\left(m_c v_s^2\right)^2 + m_c v_s^2(|U| - \delta U)\right)^{1/2} \end{aligned} \tag{A.18}$$

Since we used the approximation of $q_{max} \approx 2k_i$, we loose the constraint that carriers must be traveling above the *speed of sound energy*, $\frac{1}{2}m_c v_s^2$. This is satisfactory as the lower root, $E_{\min}$ is generally unphysical in the majority of physical cases anyhow. Most often, there is no *lower* constraint to the allowed incident carrier energy.

When the potential $|U|$ and potential lowering δU are not large, the upper energy bound becomes

$$E_{max} \approx 8m_c v_s^2 = 16(\frac{1}{2}m_c v_s^2) \tag{A.19}$$

which is about a magnitude above the speed of sound energy.

What we conclude in this subsection is that energy range contributing to carrier capture by acoustic phonon emission is proportionate to the speed of sound energy. Carriers within this span of energy *dominate* the trapping processes. This a range of extremely low energy, well into the "low energy tail" for carrier distributions under standard operational biasing conditions. While acoustic phonon emission may be considered as a near-elastic process for average carrier energies, this energy range demonstrates that only low-energy carriers are kinematically allowed to emit a phonon of energy adequate to become bound.

A constraint on the amount of barrier lowering

We can also determine a limiting relation between the barrier lowering and attractive potential, independent of the carrier energy. If we consider where the maximum allowed energy is equal to zero, we find the following.

$$\frac{\delta U}{|U|} \leq 4\sqrt{\frac{\left(\frac{1}{2}m_c v_s^2\right)}{|U|}} \tag{A.20}$$

Where this limit is surpassed, the incident carrier cannot loose adequate energy to become bound. Therefore, the barrier lowering cannot become too large relative to the depth of the potential. As the barrier lowering grows with the applied field, this effect increasingly forbids capture. This is because the energy width of available trapping states near the continuum is being evacuated out of the capture center.

The "resonant" nature of scattering rate solutions

Another notable aspect of capture rates is their general resonant nature. We saw there are limiting values initial carrier energy, E_{min} and E_{max}, which can contribute to capture. Inspecting these roots reveals the generally parabolic nature of these solutions. In figure A.2, we qualitatively see that the resulting cross sections tend to have a resonant maximum response between E_{min} and E_{max}, with the width determined by the radical term in their solutions. By examining the trend of the radical term present in the solution of both roots, we note the dependence on δU. A larger energy span and larger resonance generally occurs when δU is smaller.

Understanding that these capture rates have resonant features based on minimum and maximum phonon limits gives some insight in interpreting the features in plots of energy-dependent cross sections.

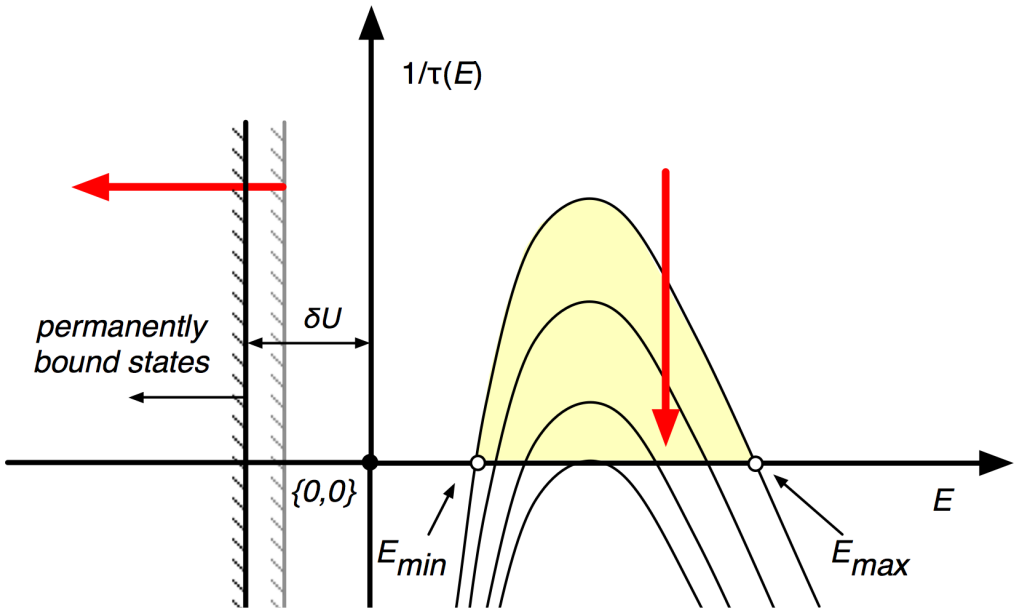

Figure A.2: As the magnitude of the potential lowering (δU) increases (red arrow moving left), the energy span and peak amplitude of carrier trapping decreases (red arrow moving down). We find that the trapping rate is naturally (at least) "2nd-order," and often has resonant peaking at an energy in between maximum and minimum allowed values.

A.5 Cascade capture by optical phonon emission

In the previous section, we considered the possibility for a carrier to emit an acoustic phonon and subsequently find a permanent, bound state in a localized, attractive center. We found that threshold effects due to potential lowering of the external field are important. Here, we consider a different scenario in the form of the large energy loss by way of optical phonon emission.

Optical phonon emission is an inelastic process. Typical optical phonons (or intervalley phonons for electrons) are 10's of milli-electronvolts in energy, while the ground state of a hydrogenic center in germanium is perhaps 12 meV. So it is important to recognize that carriers must have adequate energy to emit an optical phonon while still finding a final state that is within the potential well. The carrier needs to at least be able to "bottom out" to the ground state of the trap.

At the other extreme, perhaps during de-excitation from initial particle interactions, it could be that carriers are so very energetic that a single optical phonon does not represent enough energy loss to force a carrier to de-excite into a bound state.

Luckily, unlike the acoustic phonon case, the optical phonon dispersion relation can be considered to be perfectly flat, thereby having no q dependence in the energy relation. This simplifies these sorts of considerations.

An optical phonon has energy

$$E_{optical} = \hbar\omega_0 \tag{A.21}$$

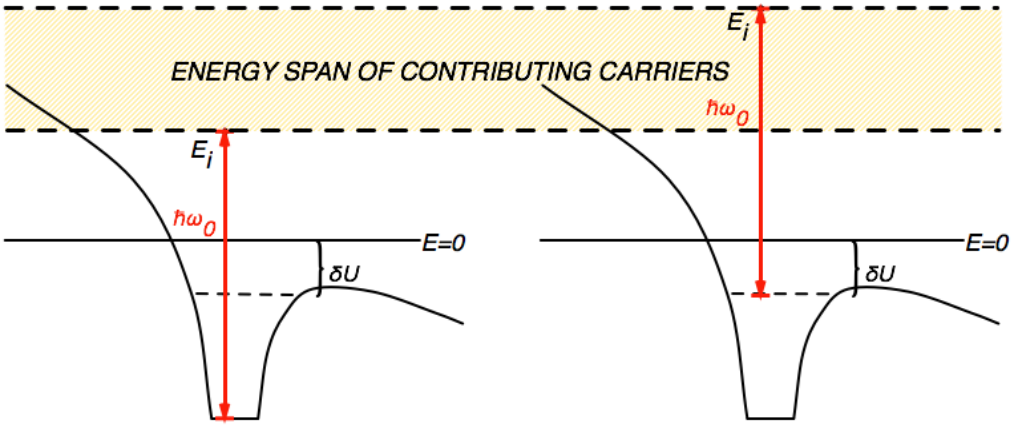

Figure A.3: Optical phonon emission leading to trapping provides a span in initial carrier energy from $E = \hbar\omega_0 + U(r)$ up to $E = \hbar\omega_0 - \delta U$.

We can say, then, that the initial carrier energy simply has to be within the following range to be able to trap by optical phonon emission.

$$\hbar\omega_0 + U(r) \leq E \leq \hbar\omega_0 - \delta U \tag{A.22}$$

This is illustrated by fig. A.3, where we see how the optical phonon energy "fits in" to our problem.

A.5.1 Optical phonon emission rate

We need to now make use of the optical emission rate.

For optical phonons, we have a slightly weaker dependence on phonon wavevector.

$$\frac{1}{\tau} = \frac{D_0^2 m_{dos}}{4\pi\hbar\rho p\omega_0} \int_{q\,\min}^{q\,\max} (N+1) q dq \tag{A.23}$$

We are certainly in the $T = 0$ limit for ambient optical phonons, so $N \approx 0$. The minimum and maximum wavevectors are as follows.

$$\begin{aligned} q_{\min} &= \tfrac{p}{m_c}\left(1 - \sqrt{1 - \tfrac{\hbar\omega_0}{E - U(r)}}\right) \\ q_{\max} &= \tfrac{p}{m_c}\left(1 + \sqrt{1 - \tfrac{\hbar\omega_0}{E - U(r)}}\right) \end{aligned} \tag{A.24}$$

We have used the relation for kinetic energy, $\varepsilon_k = E - U(r)$. So through these limits of integration, the position dependence in $U(r)$ enters our problem. We find,

$$\frac{1}{\tau} = \frac{D_0^2 m_{dos}}{\sqrt{2}\pi\rho\hbar^3\omega_0}\sqrt{m_c\left(E + \frac{Ze^2}{4\pi\varepsilon\varepsilon_0 r}\right)}\sqrt{\left(1 - \frac{\hbar\omega_0}{E + \frac{Ze^2}{4\pi\varepsilon\varepsilon_0 r}}\right)} \tag{A.25}$$

The factor of $\sqrt{\left(1 - \frac{\hbar\omega_0}{E + \frac{Ze^2}{4\pi\varepsilon\varepsilon_0 r}}\right)}$ depicts the lower limit of allowed energies, provided that we must only use this expression when it is positive definite. We can enforce the upper energy bound by enforcing a unit step function in energy, $\Theta[\varepsilon]$.

$$\frac{1}{\tau} = \frac{D_0^2 m_{dos}}{\sqrt{2}\pi\rho\hbar^3\omega_0}\sqrt{m_c\left(E + \frac{Ze^2}{4\pi\varepsilon\varepsilon_0 r}\right)}\sqrt{\left(1 - \frac{\hbar\omega_0}{E + \frac{Ze^2}{4\pi\varepsilon\varepsilon_0 r}}\right)}\Theta[(\hbar\omega_0 - \delta U) - E] \tag{A.26}$$

Equation A.26 along with equation 5.18 determine the relaxation quantity $\lambda = v\tau$ for the case of carrier capture by optical phonon emission.

So we have now established the constraints needed to include into our capture integral a rate based on either acoustic or optical phonons.

A.5.2 Accounting for the non-ballistic nature of carrier propagation in the capture region

Further extensions to our trapping model account both for the full cylindrical lowering of the potential, in addition to the fact that our carrier distributions are not ballistic and continuously adjust their distribution to the band edge as a function of position along the electric field.

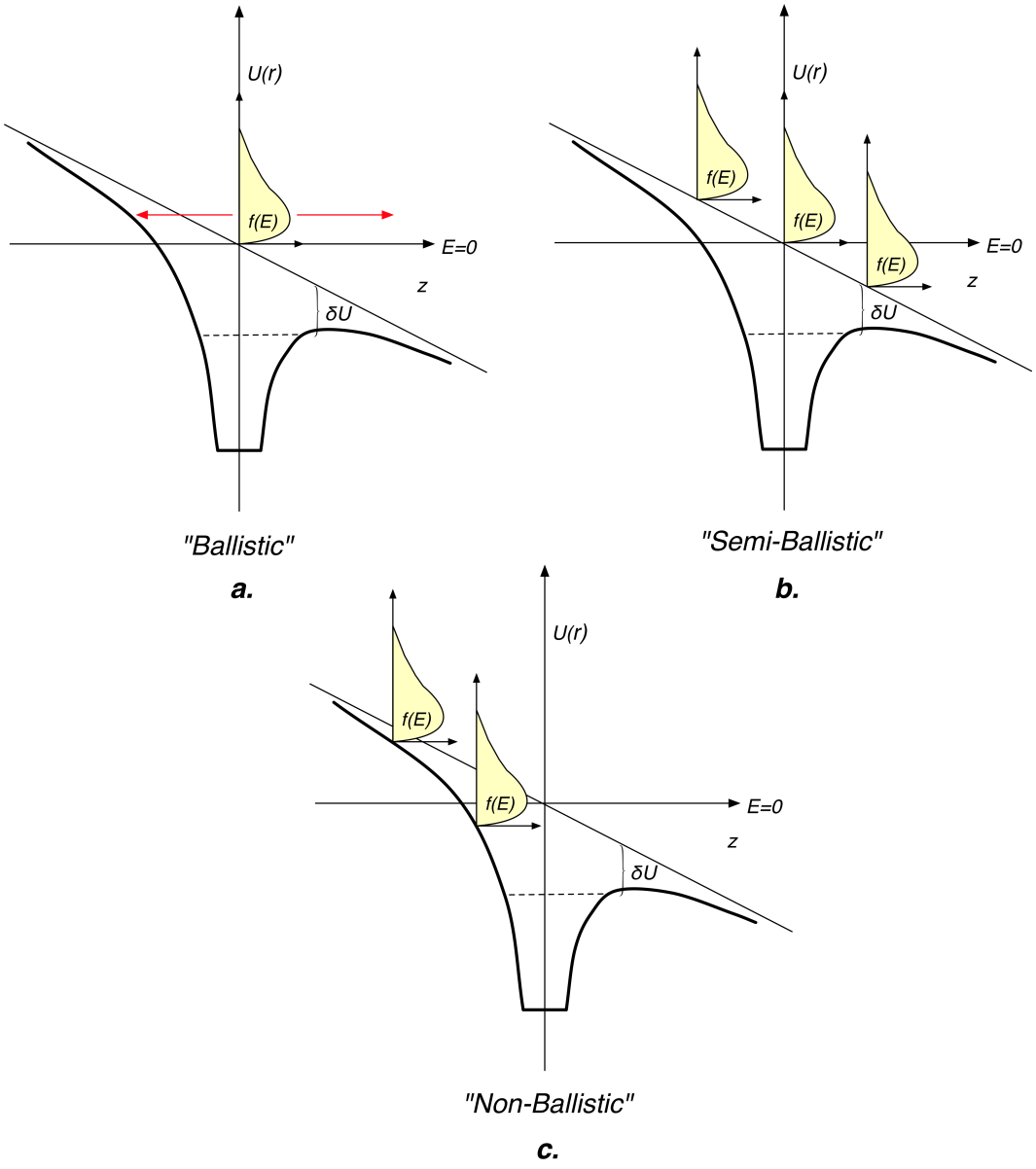

Figure A.4: *a.*) So far, the carrier energy has been referenced from the center of the local potential. Due to the external potential, this underestimates the kinetic energy at negative z coordinates, and overestimates the kinetic energy at positive z coordinates. *b.*) A correction to the total energy accounts for the assumed energy loss by phonons in a constant external field. Carrier energy distributions adopt this correction as a function of position. *c.*) A limiting case is a fully non-ballistic treatment, where the distribution function is assumed aligned with the density of states starting from zero kinetic energy.

Carriers undergo inelastic scattering while drifting in the external electric field. As depicted in previous chapters, a detailed accounting of scattering processes enables us to compute carrier energy distributions. Here, we point out that inelastic scattering also leads us to account for the external potential drop across the local capture site. Although our general approach still assumes capture is dominated by a single phonon scattering event, we introduce a correction to account for the non-ballistic nature of carrier scattering due to the external potential.

This correction is introduced by adjusting the definition of total carrier energy as referenced from the center of the capture site. As demonstrated in figure A.4, we adjust the total energy along the field direction (-z) as

$$E(\rho, z) = E(\rho, 0) - eFz \tag{A.27}$$

A note on cylindrical symmetry and numerical limits

Note that, by equation A.27, the symmetry of the problem can no longer be reduced to a purely central (radial) potential. It is now expressed in cylindrical coordinates. The simple form for energy-dependent cross section as in equation A.2 also changes. As our calculations are numerical, we do not explicitly integrate to a critical radius, r_c, but instead use numerical limits of integration that maintain the relaxation rate as positive definite. We show in figure A.5 that the effect of the external field makes the capture volume anisotropic, which is accounted for numerically.

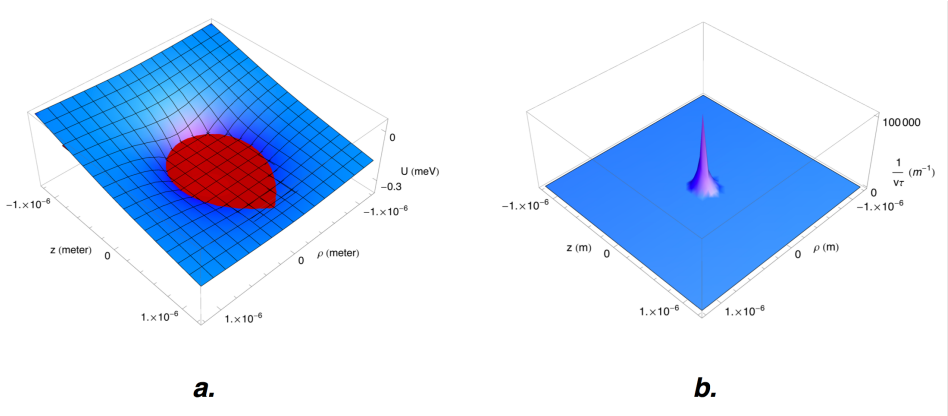

Figure A.5: *a.*) A hydrogenic capture potential in an applied field is plotted in cylindrical coordinates. The red plane represents the lowest energy allowed by acoustic phonon emission for a carrier of a particular initial energy. *b.*) For the same carrier energy, the position dependent property $\lambda^{-1} = \frac{1}{v\tau}$ is plotted.

A.6 A mid-appendix summary, and assembly of the energy-dependent cross section

Let's summarize and review what we have established in this appendix so far.

We took some effort to depict a framework for solving this problem, as applied to our low-temperature case. We also know that we will utilize the non-equilibrium, steady-state energy distributions established from prior simulation work. Therefore, we do not need to make a priori assumptions regarding an incoming carrier's energy distribution.

We incorporated a simplified version of Lax's "sticking probabilty" in a simple way in that we neglect capture to bound states within an energy width kT of the lowest accessible continuum state. The lattice temperature energy is quite small in our case, anyhow.

We also assume that the trapping volume can be traversed uniformly by any incoming carriers. While this method neglects angular momentum, it allows us to solve the problem using a volumetric integral rather than tracking arclengths over the potential trap states. As an aside, an arclength method was established as in appendix E, but it still requires work to perform adequately.

What our method establishes is the volume of real space where it is possible for an incident carrier to emit a phonon (acoustic or optical), thereby de-exciting into a bound impurity state. It turns out this volume is affected by the external potential even at reasonably low fields. Having accounted for the applied field, the lattice temperature, and functional form of the attractive potential, we can calculate radial bounds (ultimately numerically) to where capture to a bound state can no longer take place. These bounds depend on the incident carrier's energy, as well as conserving energy and momentum to only those phonons which would assist in adequate energy loss for a carrier to become bound.

To establish a cross section as a function of incident carrier energy, we weight our volumetric real-space integral by the inverse of an effective trapping length. This trapping length is simply the velocity of carrier multiplied by the isotropic emission rate of phonons appropriate to de-excite the carrier into a bound state.

$$\sigma(E) = \int_{\mathrm{V}} \frac{d^3\mathbf{r}}{\lambda(E(r))} = \int_{\mathrm{V}} \left(\frac{1}{v(E(r))} \frac{1}{\tau(E(r))} \right) d^3\mathbf{r} \tag{A.28}$$

We will then use our statistical energy distributions determined by Monte Carlo to solve for what is actually the expected *rate* per trap density.

$$C(E) = \langle v(E)\sigma(E) \rangle \tag{A.29}$$

What we cover next in this chapter are the specific kinds of attractive potentials considered, and we will also introduce probabilities for impact ionization of these states. Then, we will introduce the different kinds of impurity we consider. We will

present in the next chapter the results for the many specific kinds of capture and emission rates we have established.

A.6.1 Other processes to note, not included for CDMS conditions

Other processes which are not treated in this work include the following.

Excitonic recombination

This is the process where free electrons and holes can directly recombine, across the band gap. Since the carriers are charged oppositely, this is a Coulomb capture problem. Perhaps this could be solved by a translation to the center of mass frame. It is a joint energy distribution problem, with probabilities that either the electron or the hole would emit a phonon to bind the pair to form an exciton which would eventually decay.

Impact ionization across the gap

As mentioned briefly before, a carrier with energy that exceeds the band gap energy (actually $\approx 3/2$ of the band gap energy) can also impact ionize and generate carrier pairs out of the continuum. In a localized state translational wavevector k is no longer a good quantum number, and momentum need not be conserved. In contrast, ionizing carriers from continuum states across the band gap does require momentum conservation.

A.7 Results for specific processes

We have introduced the general method for evaluating capture cross sections, which we call "the semi-ballistic approximation." We now give results for specific cases. For each case, we have determined the cross section numerically, as a function of both carrier energy and applied electric field. Then, in a numerical implementation of equations 5.1 and 5.3, we use our simulated energy ensembles to find the expectation of this cross section times velocity. In other words, we use the appropriate carrier distributions to find the energy-independent rates and cross sections. To illustrate this idea, in figure A.6 we show some typical cross sections for $F = 1\ V/cm$ as a function of energy.

In what follows, we see a number of plots for capture and impact ionization processes, of many specific varieties. We have a few data points for low-field, equilibrium capture cross sections of electrons and holes in the $T \sim 1 - 3\ K$ range. These points match well, considering such data are not quite representative of the same temperature range. What we present is a strong field dependence, unique to the type of

process. Charged/Coulomb capture decreases rapidly with the field, owing to its $1/r$ type potential and associated barrier lowering. We see a pronounced "resonance" with the optical phonon contribution to Coulomb capture, associated with the specific range of carrier energies able to contribute to capture as in figure A.3. Neutral capture cross sections change less so. Impact ionization processes turn on with an energy threshold associated with the effective "Rydberg" of the impurity in question. In figure A.17, we compare processes for electrons and for holes for shallow impurities. We find the general trend that, since holes are "hotter" for a given field, that similar processes occur for both species, but are shifted to lower fields for holes.

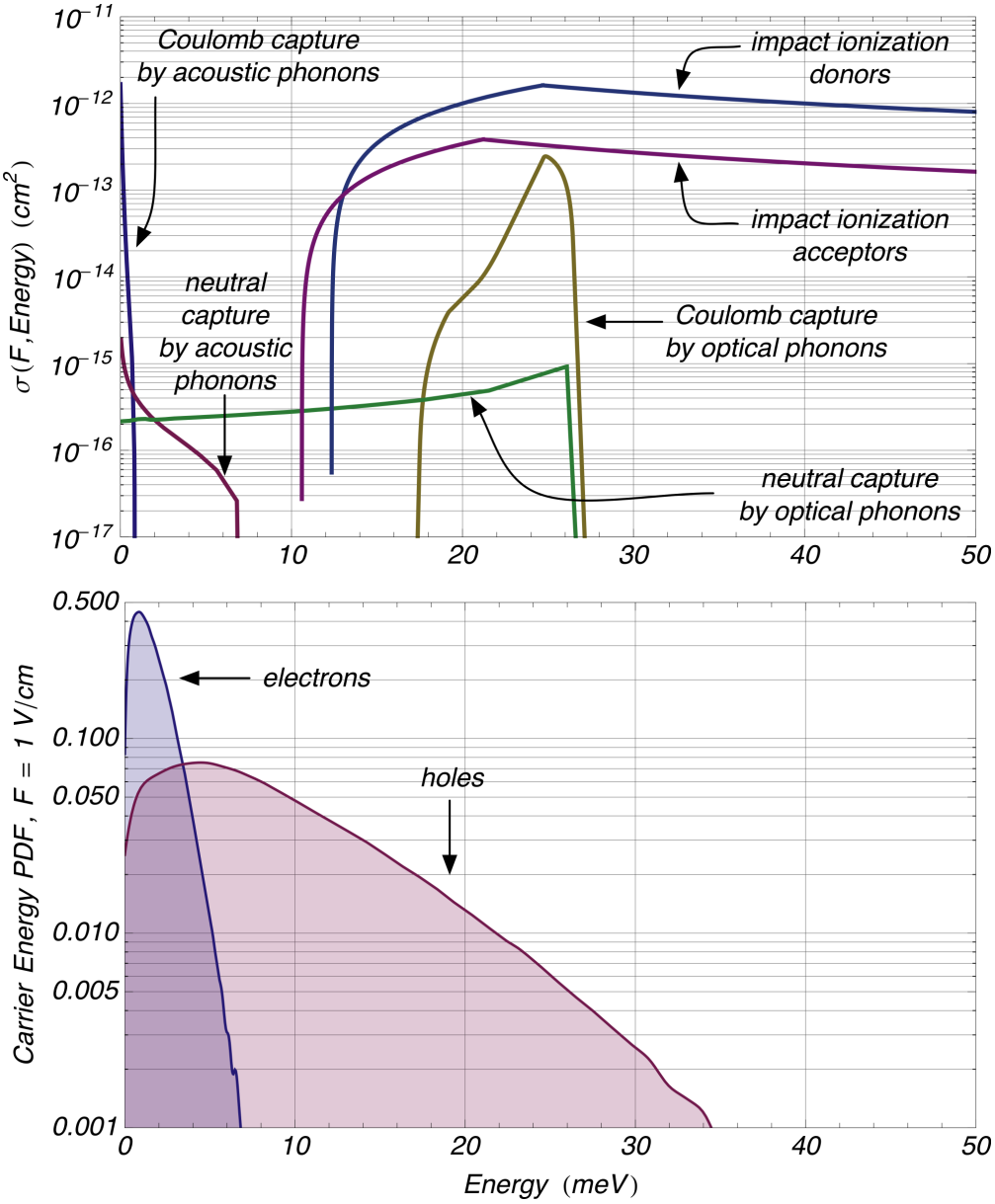

Figure A.6: Energy-dependent cross sections. **Top:** Cross sections for representative processes involving shallow hydrogenic impurities with incident electrons at $F = 1\ V/cm$. The equivalent processes for holes are either similar or identical. Note the threshold effects and resonant peaking determined by energy and momentum conservation. **Bottom:** Energy distributions for electrons at holes at $F = 1\ V/cm$, set to the same energy scale. Kernel smoothing of the statistical distributions was performed here, only for illustrative purposes.

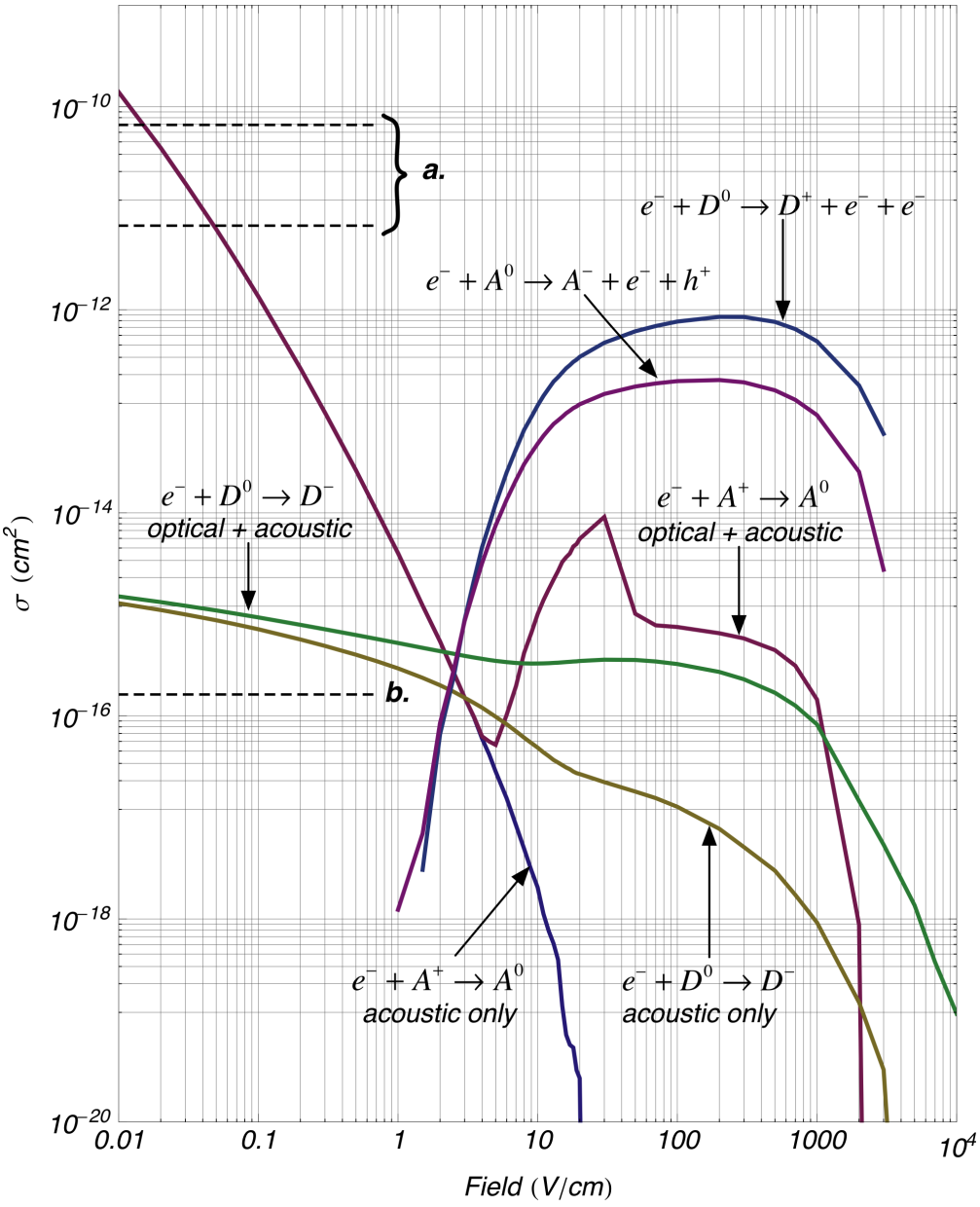

Figure A.7: Predicted cross sections for processes involving electrons and shallow impurities. Note that charged capture, neutral capture, and impact ionization have roughly the same cross section between fields of about $2.1 - 2.5\ V/cm$, where the electron mean energy is a sizable fraction ($\sim 20\%$) of the hydrogenic Rydberg. Dashed black bars indicate published cross section data, presumably at zero field. **a.** electron capture ($T = 3.1K$) on Sb^+ in Ge [16, 17], **b.** electron capture ($T = 1.5K$) on neutral acceptors [18].

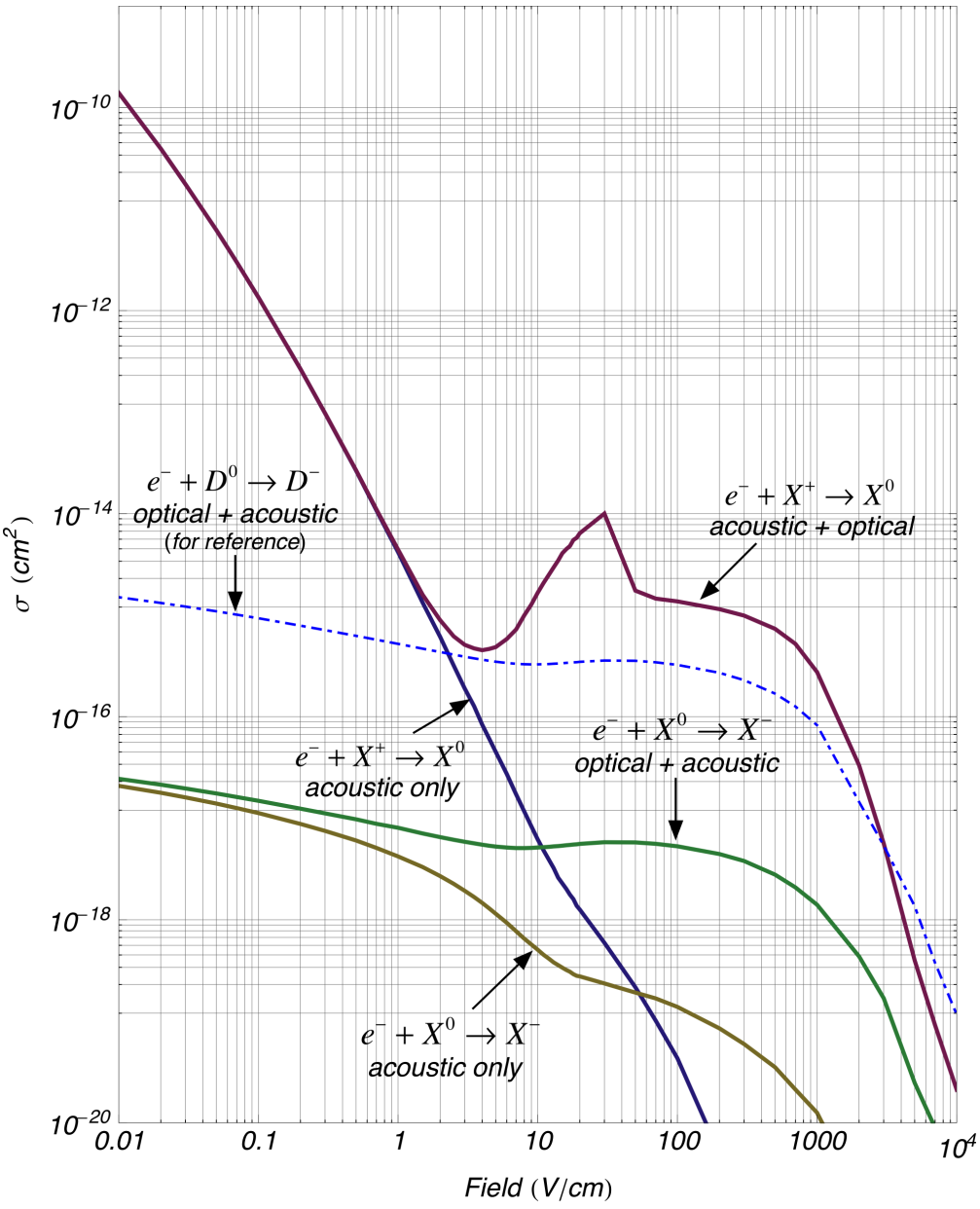

Figure A.8: Predicted cross sections for processes involving electrons and our benchmark model for deep levels. We also present the capture cross section for electrons on neutral shallow acceptors, as a reference.

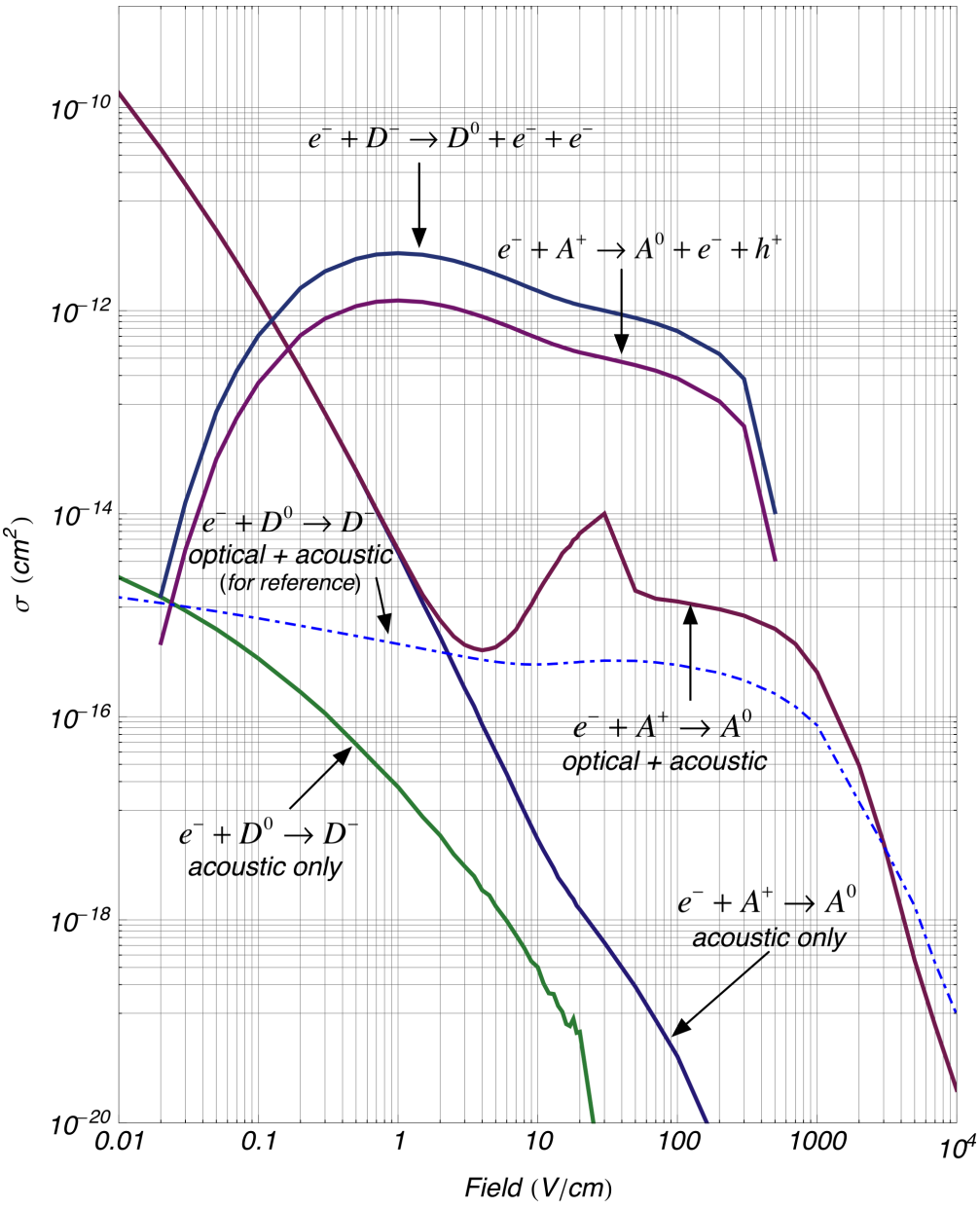

Figure A.9: Predicted cross sections for processes involving electrons and anion states. Intervalley ("optical") phonons are a negligible contribution to neutral capture in this case.

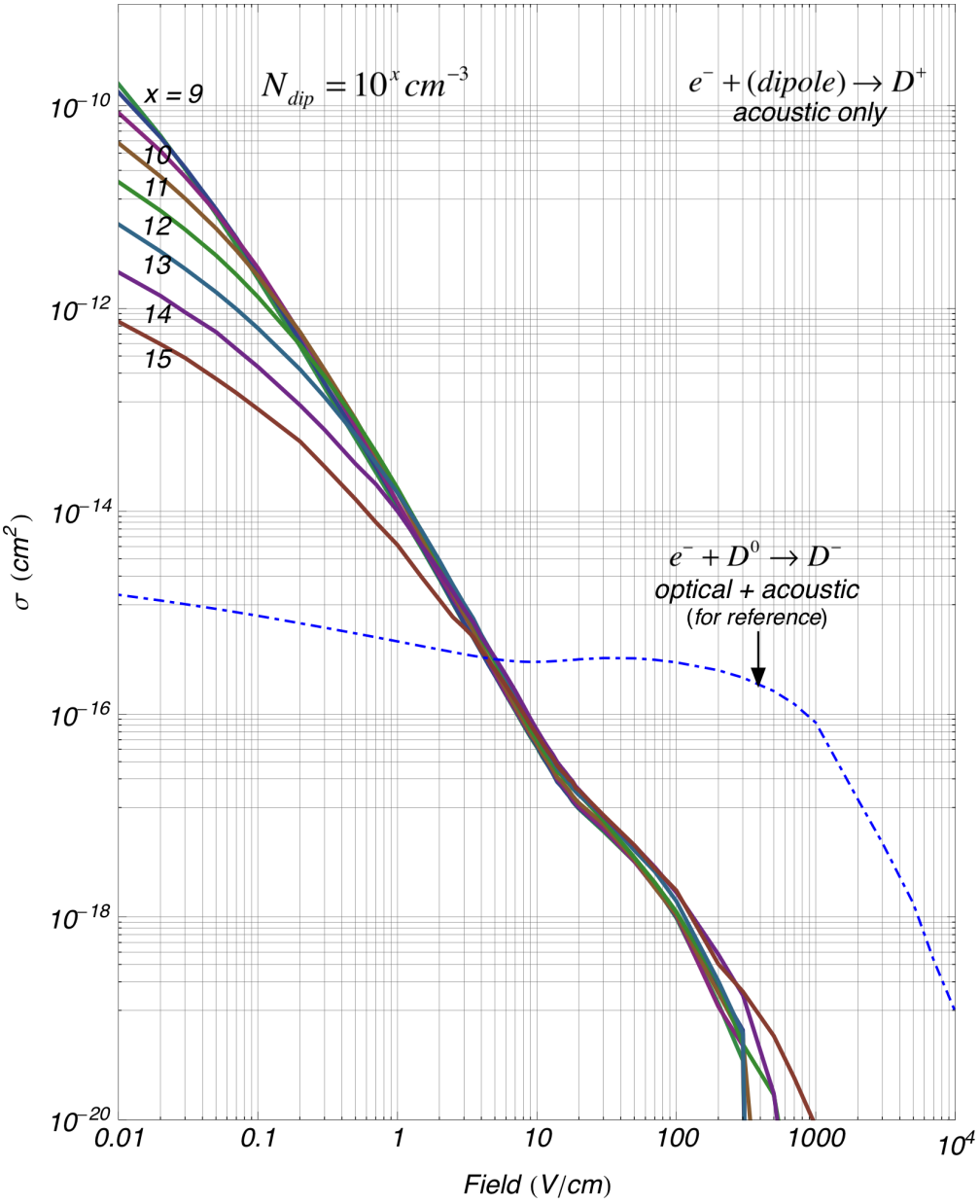

Figure A.10: Predicted cross sections (due to acoustic phonon emission) for electrons on dipoles. The dipole cross section has a dependence on the dipole concentration. The three dimensional integration used for dipoles supplies some numerical noise to these results.

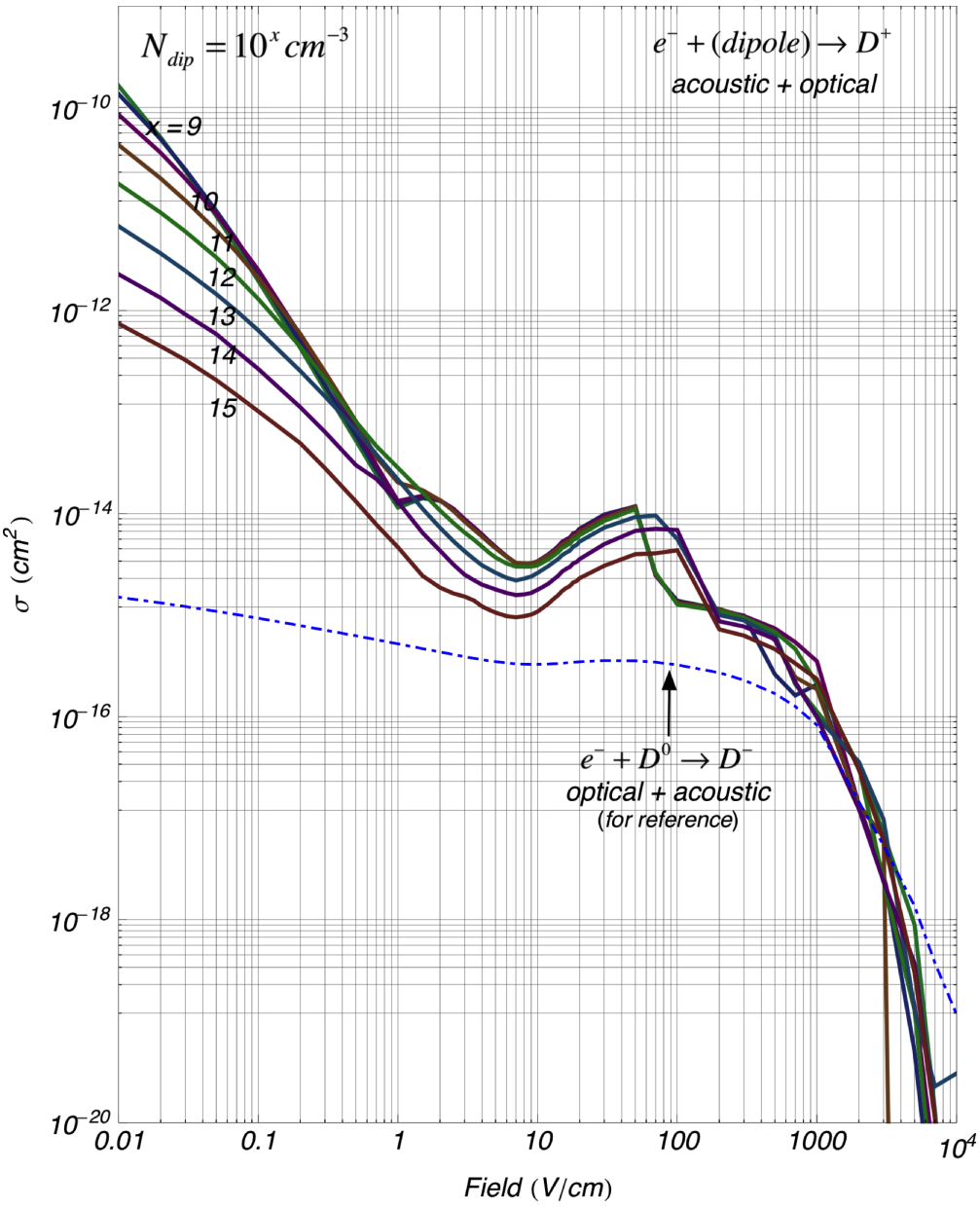

Figure A.11: Predicted cross sections (due to both optical and acoustic phonon emission) for electrons on dipoles. The three dimensional integration used for dipoles supplies some numerical noise to these results.

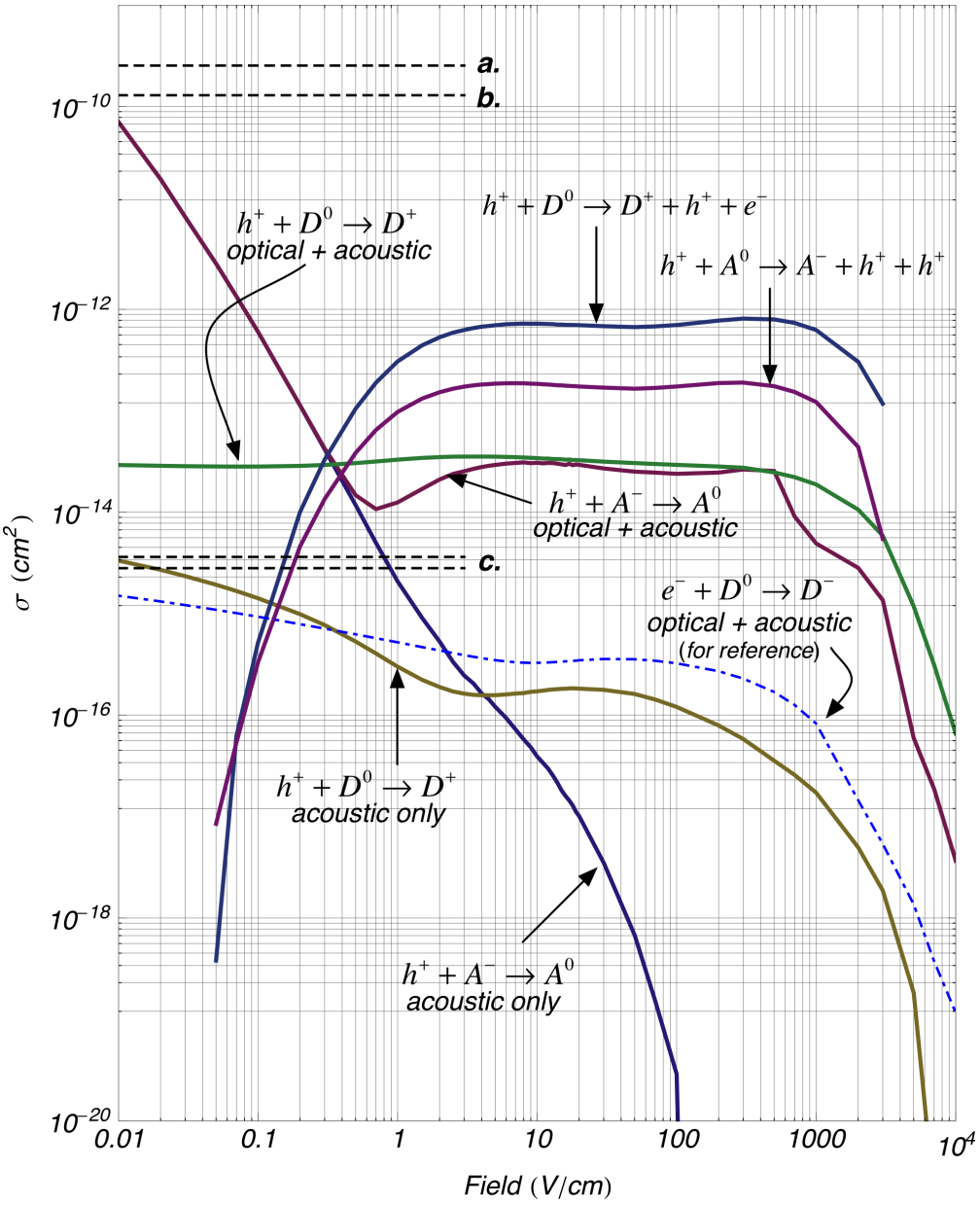

Figure A.12: Predicted cross sections for processes involving holes and shallow impurities. Note that charged capture, neutral capture, and impact ionization have roughly the same cross section between fields of about $0.3 - 0.4\ V/cm$, where the hole mean energy is a sizable fraction of the hydrogenic Rydberg. Dashed black bars indicate published cross section data, presumably at zero field. **a.** hole capture ($T = 1.6K$) on B^- in Ge [16], **b.** hole capture ($T = 2.2K$) on Al^- in Ge [16], **c.** hole capture ($T = 1.7K$) by neutral donors [18].

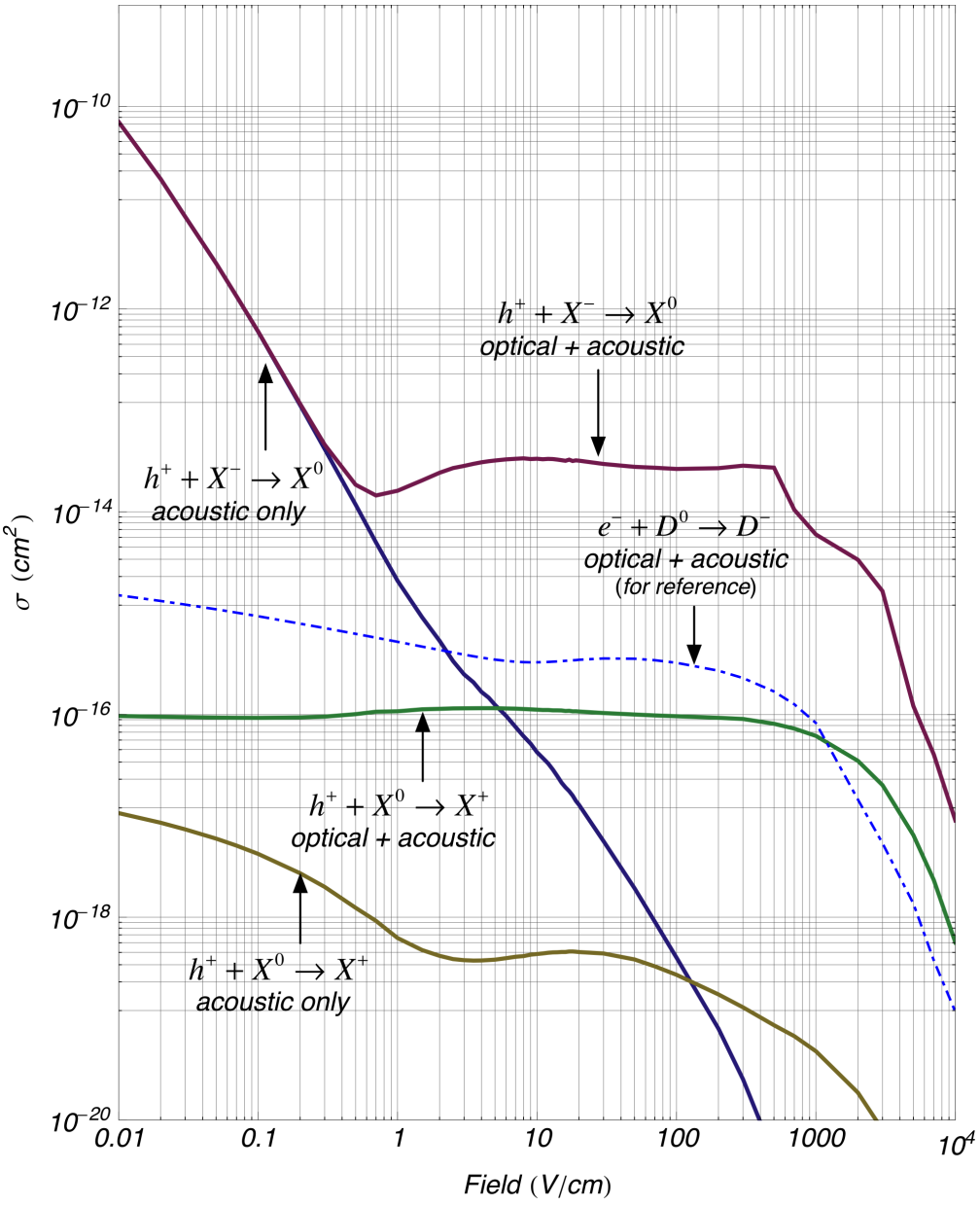

Figure A.13: Predicted cross sections for processes involving holes and our benchmark model for deep levels.

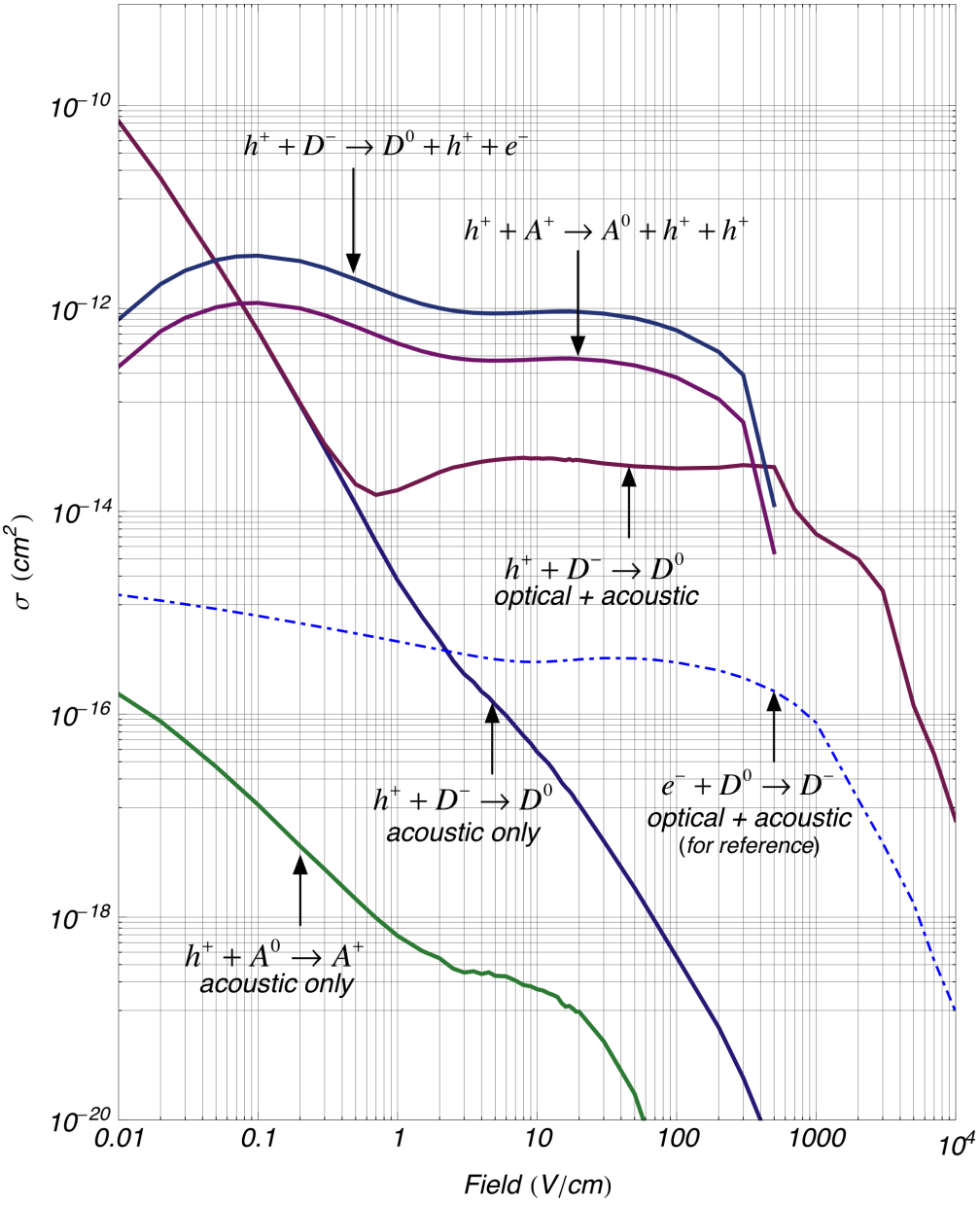

Figure A.14: Predicted cross sections for processes involving holes and anion states. Optical phonons are a negligible contribution to neutral capture in this case.

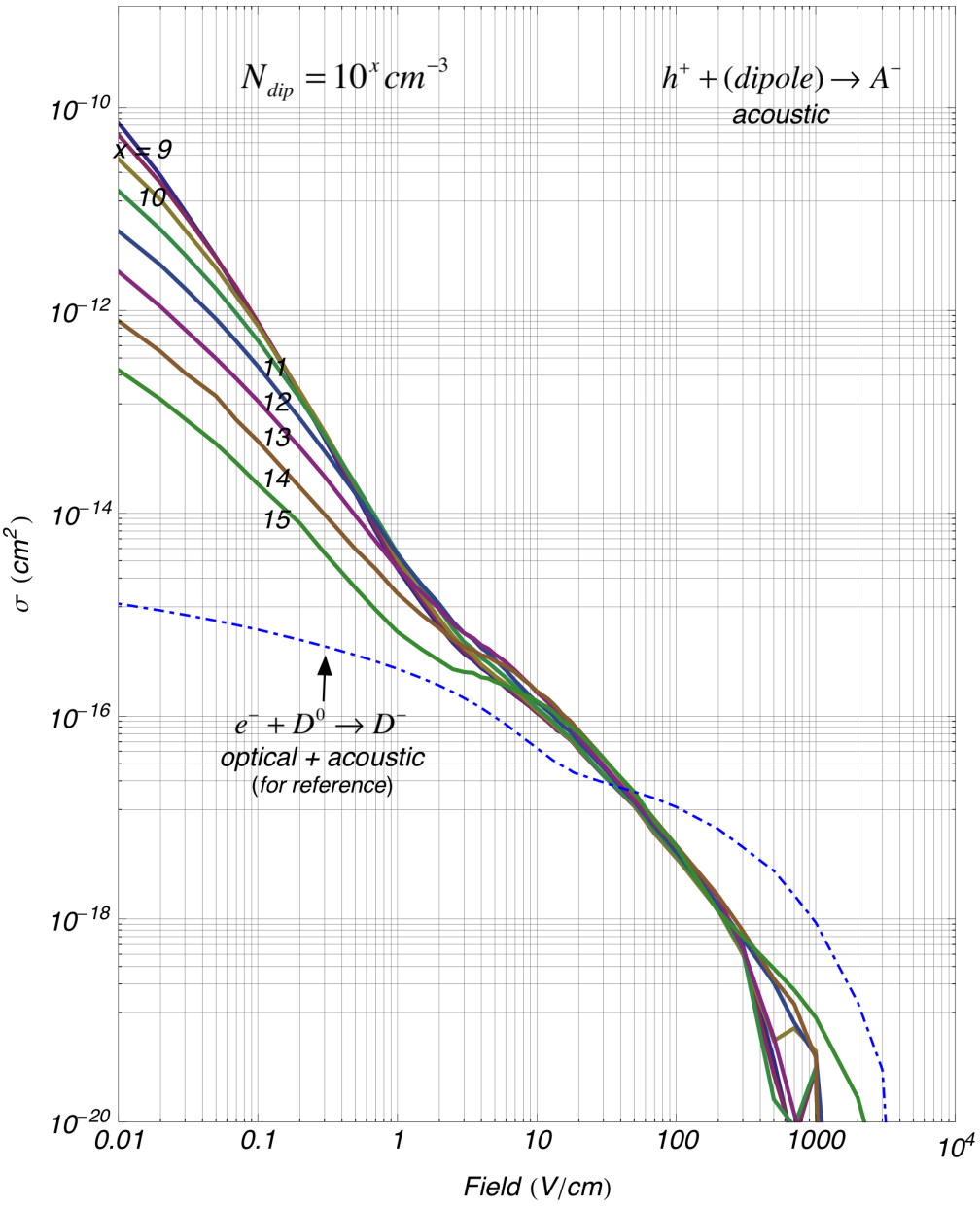

Figure A.15: Predicted cross sections (due to acoustic phonon emission only) for holes on dipoles. The three dimensional integration used for dipoles supplies some numerical noise to these results.

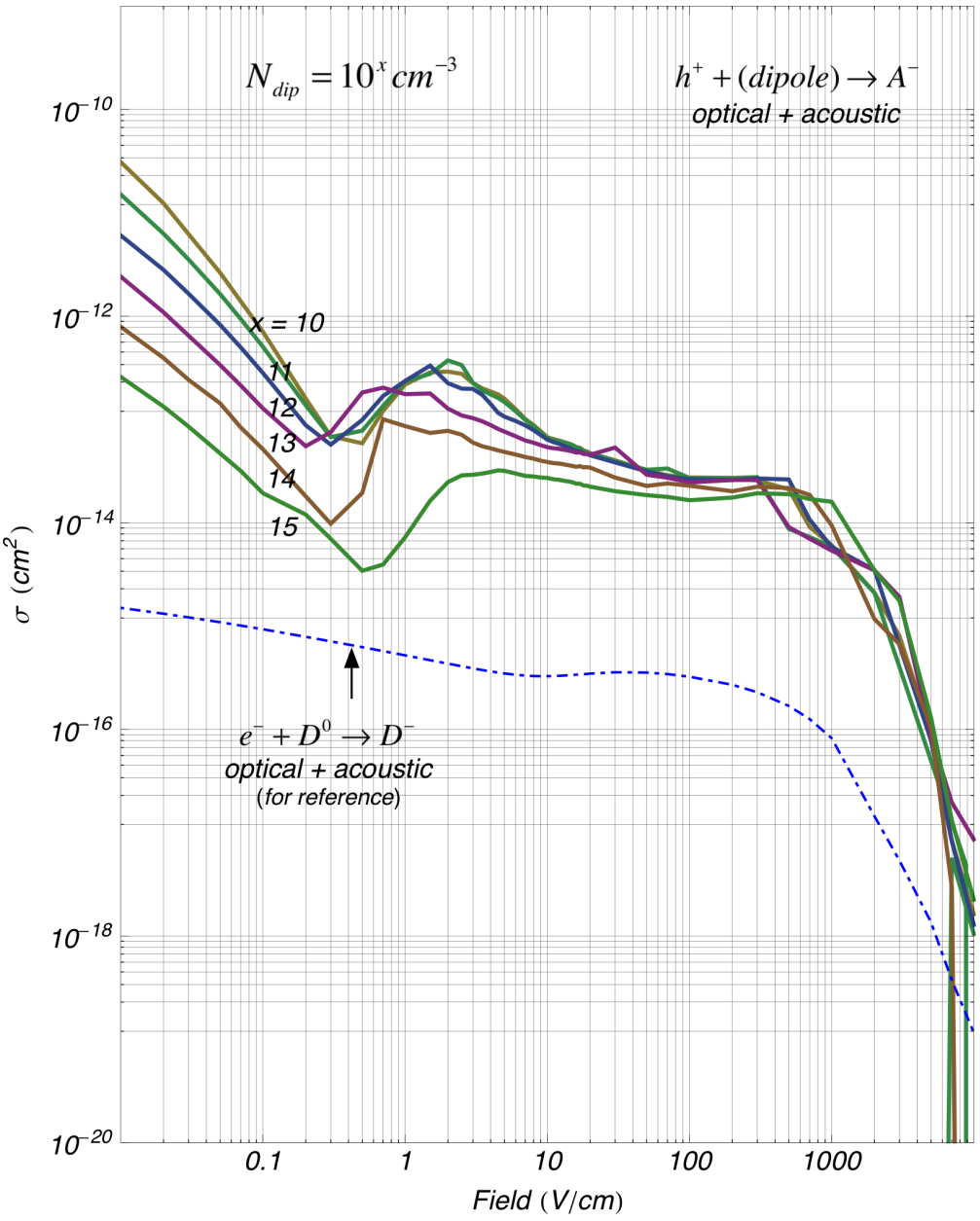

Figure A.16: Predicted cross sections (due to both optical and acoustic phonon emission) for holes on dipoles. The three dimensional integration used for dipoles supplies some numerical noise to these results.

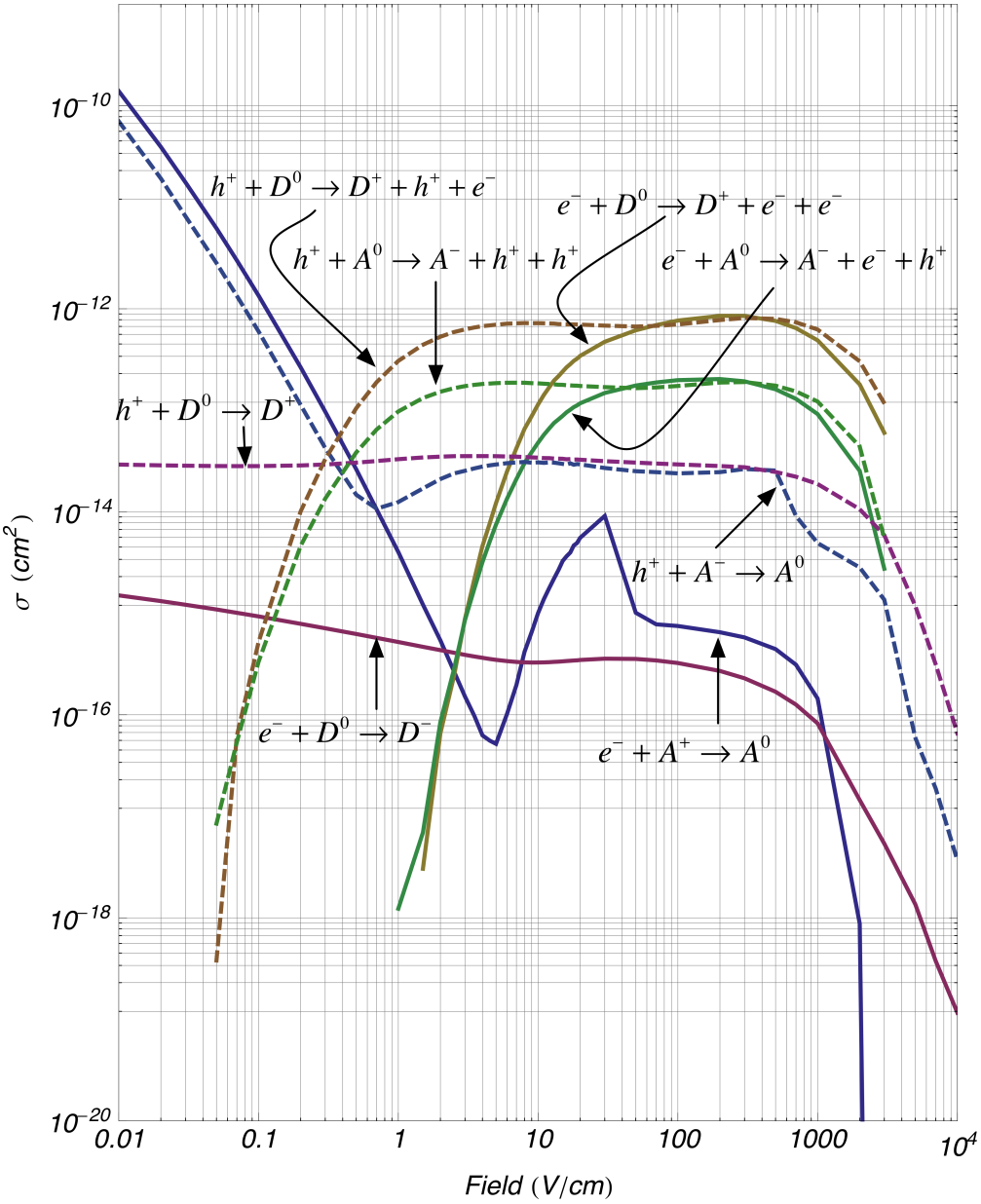

Figure A.17: A direct comparison between electrons and holes for processes involving shallow impurities. Electron (hole) processes are solid (dashed). Displayed capture rates are the total of optical and acoustic phonon contributions. Note that, as holes are generally more energetic than electrons for a given field, predicted threshold effects occur at lower fields.

A.8 Comparing to Abakumov's formulation for the case of Coulomb capture

We can make a comparison of the "semi-ballistic approximation" of this appendix, to the formulation of Abakumov. In addition to the cross section as a function of field for the results of this appendix, we can plot the extended Abakumov formalism of Chapter 5. In addition, for comparison, we can also plot the *equilibrium temperature* expressions from Abakumov for the high and low temperature limits, and the interpolation function [15] frequently used between the two limits.

In figures A.18 and A.19 we plot, for electrons and holes respectively, the pertinent cross sections we have discussed as a function of applied field. We use Monte Carlo output to compute an effective carrier temperature – here based on the mean energy – as a function of applied field. Note our figure 4.10 from our simulation results chapter for these values.

The high- and low-T limits and their interpolation function do fall off with field as we might expect from the underlying effective temperature. What is critical in figures A.18 and A.19 is that we find our "semi-ballistic approximation" falls much more sharply with field compared to Abakumov's formula. Our assumed conditions for capture as treated in Chapter 5 are more strict than Abakumov's, notably in the use of a step-function sticking probability.

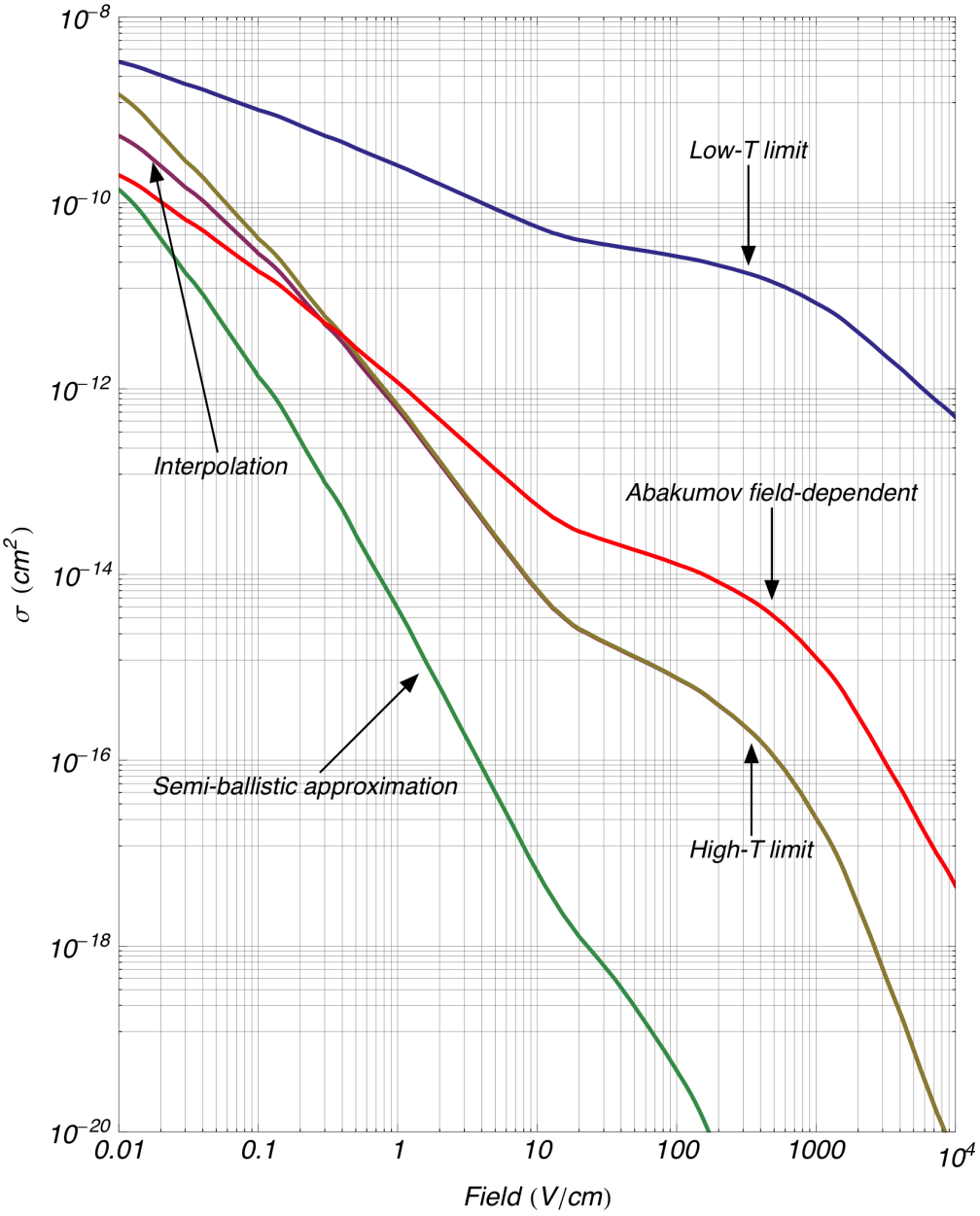

Figure A.18: Various temperature- and field-dependent formulae for the Coulomb capture of **electrons** by acoustic emission. Our numerical treatment of the semi-ballistic approximation of chapter 5 is given, as well as the Abakumov's analytical, field-dependent treatment of equation 5.63. We plot as a function of field, using Monte Carlo output to interpolate a steady-state carrier temperature.

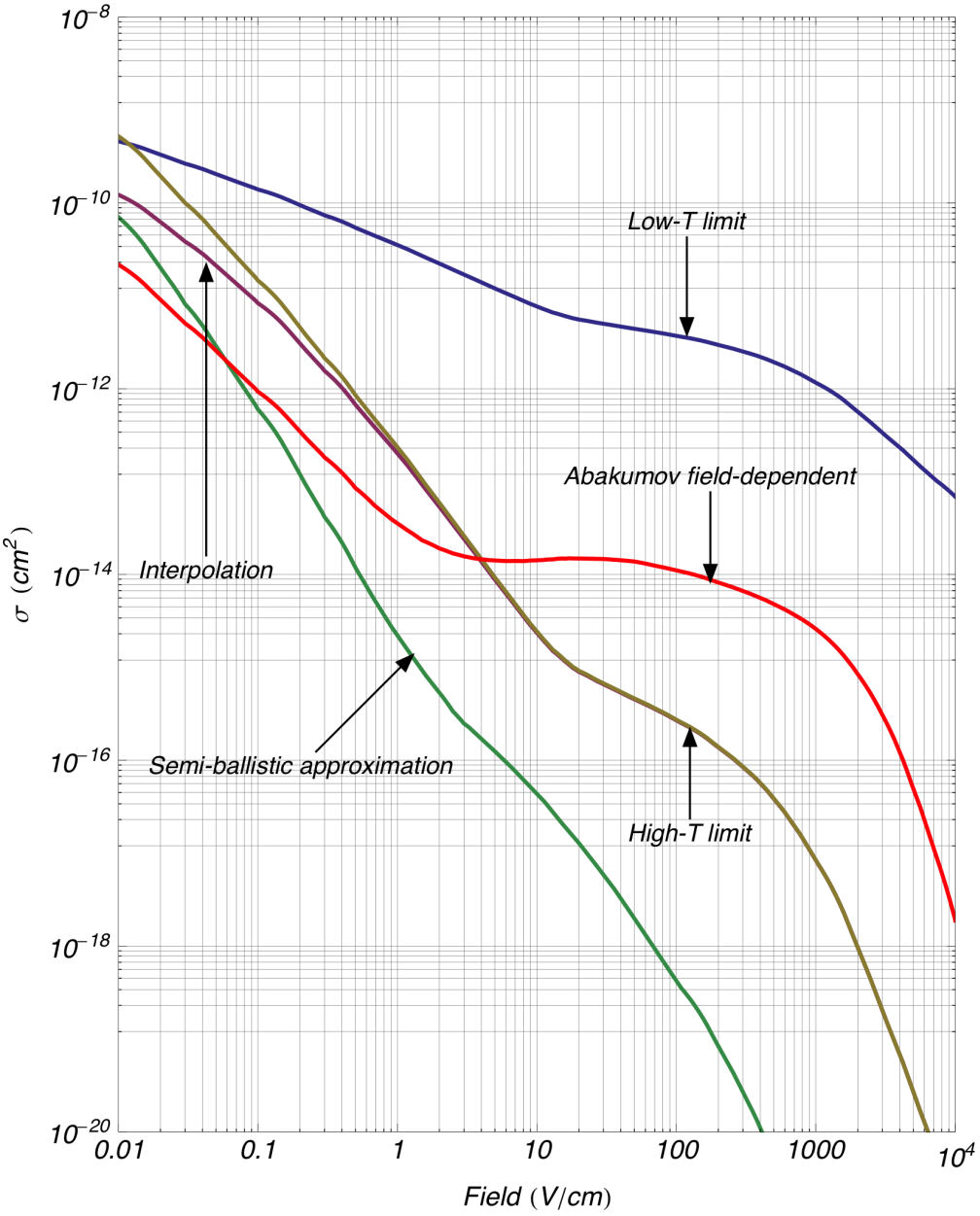

Figure A.19: Various temperature- and field-dependent formulae for the Coulomb capture of **holes** by acoustic emission. Our numerical treatment of the semi-ballistic approximation of chapter 5 is given, as well as the Abakumov's analytical, field-dependent treatment of equation 5.63. We plot as a function of field, using Monte Carlo output to interpolate a steady-state carrier temperature.

Appendix B

An Analytical, Isotropic Treatment of "Hot" Carriers

Charge carriers under CDMS conditions are always "hot." Here we treat the field dependence of carriers far from thermal equilibrium using simple assumptions. This model involves solving a coupled system of equations depicting momentum and energy balance. These two equations are coupled by way of an idealized carrier scattering rate. A single, isotropic band is assumed.

This model is a worthwhile exercise in that analytical expressions are useful for developing insight. Namely, it provides a few simple and ideal relations how some pertinent quantities should change as a function of applied electric field. It also highlights the difference between *momentum* and *energy* relaxation times, which can be powerful concepts once understood.

A treatment for balancing kinematic equations was performed by Blas Cabrera [114, 80], establishing the field dependence pertinent for our low-temperature CDMS limits. This work repeats a balance of equations in the same vein, while borrowing readily from a generalized presentation in the latest edition of B.K. Ridley's textbook [102].

To note, this is *not* a treatment of the Boltzmann Transport Equation. We are manipulating mean values. We are not self-consistently solving for an ensemble distribution. The result is that, while power-law dependencies on coefficients are typically correct, there may be discrepancies with prefactors on the order of unity. Such differences arise due to the difference between ensemble-weighted values versus manipulation of a representative mean value. For a thorough treatment of hot carriers via the Boltzmann Transport Equation, see [115]. A virtue of this simple approach is that it is easy to follow and interpret.

B.1 The steady-state balance of momentum and energy

We proceed in this section by establishing Ridley's relations [102] for momentum and energy balance, using expressions for generalized relaxation times. We will treat specific cases in the upcoming sections.

The equations for momentum and energy rates of change are as follows.

$$\dot{p} = m^* \frac{dv}{dt} = eF - \frac{m^* v}{\tau_m(E)} \tag{B.1}$$

$$\dot{E} = eFv - \frac{E - E_0}{\tau_E(E)} \tag{B.2}$$

Above, m^* is the effective mass, v is the velocity, e the electric charge, F the electric field, E the carrier energy, E_0 the energy with no applied field, $\tau_m(E)$ is the energy-dependent momentum relaxation time, and $\tau_E(E)$ is the energy-dependent energy relaxation time.

We should be careful. To which "velocity" quantity does v correspond? Perhaps you, the reader, are familiar with the *drift* versus *thermal* definitions of velocity. In this present case, those distinctions are not pertinent. Equations B.1 and B.2 are relating the *instantaneous* momentum and energy for a carrier at any point in time. In other words, we are not discussing ensemble properties at this point. Therefore, the quantity v is simply the total and instantaneous velocity.

The definition of the zero-field energy, E_0, is basically the carrier energy when thermalized. One should know that E_0 properly should be equated to the thermal energy, $\frac{3}{2}k_B T_L$, of the lattice. However, our simple model assumes carriers are *quite* hot relative to the lattice. This is a completely reasonable assumption under almost any conceivable CDMS scenario. So, we would ultimately set

$$E_0 = 0 \tag{B.3}$$

although we will retain it in our equations as we proceed.

Finding the steady state, equation B.1 recovers an expression for velocity. We can therefore also incorporate it into the energy-balance equation.

$$\begin{aligned} v &= \tfrac{e\tau_m(E)}{m^*} F \\ E &= E_0 + (eF)^2 \, \tfrac{\tau_m(E)\tau_E(E)}{m^*} \end{aligned} \tag{B.4}$$

Ridley makes the point that, via equation B.4, hot effects should relevant at a critical field where $E \approx E_0$

$$F_c = \sqrt{\frac{E_0}{(\frac{e^2}{m^*})\tau_m(E_0)\tau_E(E_0)}} \tag{B.5}$$

Next, we need to incorporate the energy dependence of the relaxation times. The scattering rates' dependence on energy can be parameterized as power laws [45, 102]. The following forms for relaxation times

$$\begin{aligned} \tau_m(E) &= AE^p \\ \tau_E(E) &= BE^q \end{aligned} \tag{B.6}$$

yield

$$\begin{aligned} v &= \frac{eFA}{m^*}\left(\frac{(eF)^2AB}{m^*}\right)^{\frac{p}{1-p-q}} \\ E &= \left(\frac{(eF)^2AB}{m^*}\right)^{\frac{1}{1-p-q}} \end{aligned} \tag{B.7}$$

So we see that an energy dependence in the *scattering* rates (thereby in the *relaxation* rates) results directly in a dependence on the applied field.

B.2 The steady state under acoustic phonon emission

In this section, we evaluate the steady-state quantities as laid out in the previous section. We first must determine appropriate relaxation times.

B.2.1 Pertinent rates for acoustic phonon emission

Now we treat the case for acoustic phonon emission. We assume an isotropic emission rate, as we defined earlier in this thesis regarding capture cross sections. To maintain simplicity, we assume the typical carrier energy is far above the speed of sound energy. We therefore approximate the minimum allowed phonon wavevector to be zero. We also assume zero ambient phonon occupation of the lattice. The scattering rate is then,

$$\Gamma_{ac}(E) = \frac{4I^2 m_{dos} m_c \Xi^2}{3\pi\hbar^4\rho v_s} E \tag{B.8}$$

where these quantities have been described in our chapter on scattering processes.

Generally, the momentum relaxation time can be related to the scattering rate by a weighting factor [45].

$$\frac{1}{\tau_m(E)} = \left(1 - \frac{k_f}{k_i}\cos\theta_{if}\right)\Gamma(E) \tag{B.9}$$

So the momentum relaxation rate varies from the scattering rate, as we can see from the weighting factor. The weighting factor departs from unity when the emission of phonons is not isotropic. This happens to be the case for acoustic phonon emission, as larger phonon wavevectors are preferred in the emission rate.

We solve for an expectation value of momentum relaxation time.

$$\frac{1}{\langle\tau_m(E)\rangle} = \left\langle 1 - \frac{k_f}{k_i}\cos\theta_{if}\right\rangle\Gamma(E) \tag{B.10}$$

For this process of acoustic phonon emission, we consider the interaction as near elastic for modestly high carrier energies. So we approximate that the final and initial carrier momenta are about the same ($k_i \approx k_f$). The law of cosines to determine the angle between initial and final electrons states is,

$$\begin{aligned} q^2 &= k_i^2 + k_f^2 + 2k_i k_f\cos\theta_{if} \\ &\approx 2k_i^2\left(1+\cos\theta_{if}\right) \end{aligned} \tag{B.11}$$

The expectation value regarding the final carrier state angle relative to the original is

$$\left\langle 1 - \frac{k_f}{k_i}\cos\theta_{if}\right\rangle \approx \langle 1-\cos\theta_{if}\rangle = \frac{\int_0^{2k_i}\left(1-\frac{2k_i^2-q^2}{2k_i^2}\right)q^2dq}{\int_0^{2k_i}q^2dq} = \frac{6}{5} \tag{B.12}$$

So we say the expected momentum relaxation time for acoustic phonon emission is

$$\langle\tau_m^{ac}(E)\rangle = \frac{5}{6}\frac{1}{\Gamma_{ac}(E)} \tag{B.13}$$

$$\boxed{\langle\tau_m^{ac}(E)\rangle = \frac{5}{8}\frac{\pi v_s\rho\hbar^4}{I^2 m_c m_{dos}\Xi^2}\frac{1}{E}} \tag{B.14}$$

In general, we define the energy relaxation time by the following [45].

$$\frac{1}{\tau_E(E_i)} = \left(1 - \frac{E_f}{E_i}\right)\Gamma(E_i) \tag{B.15}$$

Since we know the final state is initial state missing a phonon quantum, $E_f = E_i - \hbar\omega$, we have

$$\frac{1}{\tau_E(E_i)} = \left(\frac{\hbar\omega}{E_i}\right)\Gamma(E_i). \tag{B.16}$$

For acoustic phonon emission, we take an expectation value of $\hbar\omega = \hbar v_s q$ as

$$\left\langle \frac{\hbar\omega}{E_i} \right\rangle = \left\langle \frac{\hbar v_s q}{E_i} \right\rangle = \frac{\hbar v_s \left(3k_i/2\right)}{E_i} = \frac{3}{2}\frac{v_s\sqrt{2m^*E_i}}{E_i} \tag{B.17}$$

such that the energy relaxation time in this case is

$$\langle \tau_E^{ac}(E) \rangle = \frac{2E}{3v_s\sqrt{2m^*E}}\frac{1}{\Gamma_{ac}(E)} \tag{B.18}$$

$$\boxed{\langle \tau_E^{ac}(E) \rangle = \frac{\pi\rho\hbar^4}{2\sqrt{2m_c}m_c m_{dos} I^2 \Xi^2}\frac{1}{\sqrt{E}}} \tag{B.19}$$

B.2.2 Field dependence for the case of acoustic phonon emission

By our power-law prescription of Ridley, we see that the momentum relaxation time goes as E^{-1} while the energy relaxation time goes as $E^{-1/2}$. Employing equations B.7, we find average values of velocity and energy. We now may think of these as the drift velocity and the mean carrier energy.

$$\boxed{v_d^{ac} = \frac{5^{3/5}\pi^{1/5}}{2^{6/5}}\frac{v_s^{3/5}\rho^{1/5}\hbar^{4/5}}{(I^2)^{1/5}m_c^{3/5}m_{dos}^{1/5}\Xi^{2/5}}(eF)^{1/5}} \tag{B.20}$$

$$\boxed{\langle E^{ac} \rangle = \frac{5^{2/5}\pi^{4/5}}{2^{9/5}}\frac{v_s^{2/5}\rho^{4/5}\hbar^{16/5}}{(I^2)^{4/5}\, m_c^{14/5}m_{dos}^{4/5}\Xi^{8/5}}(eF)^{4/5}} \tag{B.21}$$

To note in particular, drift velocity goes as $\propto F^{1/5}$ and mean energy $\propto F^{4/5}$.

We should be cautious of the expression for the "threshold field" of equation B.5. At a CDMS working temperature of $T_L = 40\ mK$, $\frac{3}{2}k_B T_L \approx (0.3 - 1.0) \times \frac{1}{2}m_c v_s^2$, depending on the effective mass of the carrier. Elastic approximations are not valid at this energy, so equation B.5 is not particularly useful. In any case, we know that carriers cannot be in thermal equilibrium as they cannot emit phonons to reach equilibrium with a temperature below the speed of sound energy.

B.3 The steady state under optical phonon emission

This framework also allows us to consider steady-state conditions when scattering is dominated by other processes. Here, we consider the non-polar optical phonon emission in the isotropic, $T = 0$ limit.

$$\Gamma_{op}(E) = \frac{I^2 \Xi_0^2 m_{dos} \sqrt{m_c}}{\sqrt{2}\pi\rho\omega_0\hbar^3} \sqrt{E - \hbar\omega_0} \tag{B.22}$$

where these quantities have been introduced in our chapter on scattering mechanisms. Note that the scattering rate must be either zero or positive, so emission with a carrier energy below the $\hbar\omega_0$ threshold is energetically not allowed.

The process of optical phonon emission is highly inelastic. The final momentum state is randomized and much smaller in amplitude than the initial momentum. Considering the momentum relaxation time of equation B.9, we say that the momentum relaxation time is simply the inverse scattering rate,

$$\boxed{\langle \tau_m^{op}(E) \rangle = \frac{\sqrt{2}\pi\rho\omega_0\hbar^3}{I^2 \Xi_0^2 m_{dos} \sqrt{m_c}} \frac{1}{\sqrt{E - \hbar\omega_0}}} \tag{B.23}$$

where we know that the scattering rate remains zero below threshold, so that the momentum relaxation time must be infinite.

For energy relaxation, the phonon quanta are always set by the optical phonon frequency, ω_0. We have

$$\langle \tau_E^{op}(E) \rangle = \left(\frac{E}{\hbar\omega_0} \right) \frac{1}{\Gamma_{op}(E)} \tag{B.24}$$

$$\boxed{\langle \tau_E^{op}(E) \rangle = \frac{\sqrt{2}\pi\rho\hbar^2}{I^2 \Xi_0^2 m_{dos} \sqrt{m_c}} \frac{E}{\sqrt{E - \hbar\omega_0}}} \tag{B.25}$$

We have to implement our velocity and energy expressions carefully, as we have a threshold effect in place for this scenario. It still works out under the framework we have established above, however.

$$\boxed{\langle E_{op} \rangle = \hbar\omega_0 + \frac{2\pi^2\rho^2\omega_0\hbar^5}{\left(I^2\right)^2 m_c^2 m_{dos}^2 \Xi_0^4} \left(eF\right)^2} \tag{B.26}$$

The mean energy offset of $\hbar\omega_0$ is just the obvious outcome of a non-trivial energy threshold. Otherwise, the field dependence goes as $\propto F^2$.

We find that the field dependence of the energy in the momentum relaxation time return a drift velocity of

$$\boxed{v_d^{op} = \sqrt{\frac{\hbar\omega_0}{m_c}}} \tag{B.27}$$

The solution achieved via Boltzmann Transport Equation [115] is

$$v_d^{op} = \sqrt{\frac{2^3\ \hbar\omega_0}{3\pi\ m_c}} \tag{B.28}$$

which represents the same limit, but is 15% smaller.

So the drift velocity, when dominated by optical phonon emission, is saturated at this value. If there is an adequate "plateau" in future drift velocity measurements at higher fields, comparing to this value may make for a good double-check of systematics.

In reality, perfect saturation does not take place. As a function of increasing field, equation B.23 begins to become a poor approximation. As carrier energy increases with field, it becomes substantially larger than $\hbar\omega_0$. Emission of a single optical phonon does not completely randomize the momentum. The momentum relaxation time starts to become longer than simply the scatter time. As a result, drift velocity begins to again increase with field. Furthermore, other non-ideal effects such as non-parabolicity and interaction with the higher conduction (valence) minima also change the average properties of the system as an ensemble.

B.4 The lack of a steady state under impurity scattering

Impurity scattering is considered elastic. There is no energy lost in a scattering event, so the energy relaxation time is infinite. There can be no steady state as there is no dissipation to balance the power gained from the electric field.

A momentum relaxation time can be a valid concept under ionized impurity scattering, however. To note, the momentum relaxation time takes considerably longer than a scatter time–at least for ionized impurities–because small-angle scattering is favored [57].

Appendix C

Displaced Maxwellian Distribution, as a Function of Energy

Here, we examine the form of the displaced Maxwellian. We find this defined in terms of momentum and not energy. It is useful to convert this to a function of energy. Although that may seem trivial at first, we can actually use this to derive the Wannier relation [84] for expected energy.

C.1 A multivariate normal distribution

Even accounting for mass anisotropy and other effects, carriers undergoing randomizing, near-elastic phonon scattering can be depicted by a multivariate normal distribution.

A general form for the multivariate normal distribution (tong) is

$$f(\boldsymbol{k}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-Q(\boldsymbol{k})/2} \tag{C.1}$$

where $n = 3$ is the dimensionality, and

$$Q(\boldsymbol{k}) = (\boldsymbol{k} - \boldsymbol{k_0})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{k} - \boldsymbol{k_0}) \tag{C.2}$$

In this case, $\boldsymbol{k_0}$ is the average momentum wavevector, corresponding to "drift" momentum.

To further connect to physical quantities, we define an element of $\boldsymbol{\Sigma}$ as

$$\Sigma_{ij} = \frac{k_B T_{ij}}{\hbar^2} (\frac{1}{m_{ij}})^{-1} \tag{C.3}$$

To formulate an isotropic model, we can consider this covariance matrix as a diagonal matrix with equal components. So for the sake of a simplified model, we assert

$$\Sigma_{ij} = \frac{k_B m_c T}{\hbar^2}\delta_{ii} \tag{C.4}$$

where m_c is the conductivity effective mass.

C.1.1 An effectively isotropic average momentum wavevector

In dealing with crystal anisotropies, we have several different mass terms for which we must account. We of course have not been assuming the conductivity mass, m_c, as the effective mass in most of our work. The exercise before us is to determine an adequate isotropic approximation for general use. We can "normalize" the average momentum wavevector amplitude such that the drift energy term is preserved and the "average" conductivity mass can be used.

In other words, we want to ensure that the energy held in the average wavevector under the full anisotropy treatment will be equivalent to a similar term in a simple isotropic approximation.

$$\frac{\hbar^2}{2}\boldsymbol{k_0}^T(\boldsymbol{m}^{-1})\boldsymbol{k_0} = \frac{\hbar^2 k_0^{*2}}{2m_c} \tag{C.5}$$

where the quantity k_0^* is an isotropically weighted vector,

$$k_0^* = |\boldsymbol{k_0}|\left(\frac{< m^{-1} >}{m_c}\right)^{1/2} \tag{C.6}$$

and this mass expectation value is determined by

$$< m^{-1} >= \frac{\boldsymbol{k_0}^T\boldsymbol{m}^{-1}\boldsymbol{k_0}}{\boldsymbol{k_0}\cdot\boldsymbol{k_0}} \tag{C.7}$$

So in this way, we are able to identify an isotropically balanced average wavevector, k_0^*, for use with an isotropic conducitvity mass, m_c, for use in an isotropically approximate distribution.

C.1.2 The displaced Maxwellian distribution

We are therefore able to make a simplification of a *displaced Maxwellian* distribution, of the normalized form

$$f(k,\vartheta) = \frac{1}{\left(2\pi k_B m_c T/\hbar^2\right)^{3/2}}\mathrm{e}^{-\left(\frac{k^2+k_0^{*2}-2kk_0^*\cos\vartheta}{2(k_B m_c T/\hbar^2)}\right)} \tag{C.8}$$

where ϑ is the angle between momentum vector $\boldsymbol{k}$ and $\boldsymbol{k_0^*}$, the drift wavevector. We further reduce this function by taking an angular expectation value over polar

coordinates in k-space. We multiply by 2π for the azimuthal angle integration, we perform the ϑ expectation integral and weight by the k-vector volume element. We do not yet do the integration over wavevector amplitude.

$$g(k)dk = 2\pi \left(\int_0^\pi f(k,\vartheta)\sin\vartheta d\vartheta \right) k^2 dk = \frac{\left(e^{-\frac{\hbar^2(k-k_0^*)^2}{2k_B m_c T}} - e^{-\frac{\hbar^2(k+k_0^*)^2}{2k_B m_c T}} \right)}{\sqrt{2\pi}\sqrt{\frac{k_B m_c T}{\hbar^2}}} \frac{k}{k_0^*} dk \quad \text{(C.9)}$$

C.1.3 Normallzation and Wannier relation

To check, we can see that we are properly normalized, and

$$\int_0^\infty g(k)dk = 1 \quad \text{(C.10)}$$

Furthermore, we find an important relation by taking the expectation value for typical energy.

$$< \varepsilon(k) >= \int_0^\infty \varepsilon(k)g(k)dk = \int_0^\infty \frac{\hbar^2 k^2}{2m_c} g(k)dk \quad \text{(C.11)}$$

Evaluating, we find

$$< \varepsilon(k) >= \underbrace{\frac{3}{2}k_B T}_{\varepsilon_{diffusion}} + \underbrace{\frac{\hbar^2 k_0^{*2}}{2m_c}}_{\varepsilon_{drift}} \quad \text{(C.12)}$$

What we have established is the *Wannier relation*, which states that carrier energy is approximated as the energy held in stochastic, thermal, or *diffusion energy*, $\frac{3}{2}k_B T$, is summed with the *drift energy* which is held in the average momentum, $\frac{\hbar^2 k_0^{*2}}{2m_c}$.

C.1.4 Expressing the displaced Maxwellian as a function of energy

Most processes are *energy* dependent, so *energy* is the quantity we are most interested in for distribution functions.

We re-evaluate the integrated displaced Maxwellian, substituting in for energy. We utilize

$$k = \sqrt{2m_c\varepsilon}/\hbar \quad \text{(C.13)}$$

and therefore

$$dk = \frac{dk}{d\varepsilon}d\varepsilon = \frac{m_c}{\hbar^2 k}d\varepsilon \quad \text{(C.14)}$$

to find

$$g(k)dk = g(\varepsilon)d\varepsilon = \frac{m_c \left(e^{-\frac{\hbar^2\left(\frac{\sqrt{2}\sqrt{\varepsilon m_c}}{\hbar} - k_0^*\right)^2}{2Tk_B m_c}} - e^{-\frac{\hbar^2\left(\frac{\sqrt{2}\sqrt{\varepsilon m_c}}{\hbar} + k_0^*\right)^2}{2Tk_B m_c}} \right)}{\hbar k_0^* \sqrt{2\pi}\sqrt{k_B T m_c}} d\varepsilon \tag{C.15}$$

Multiplying out the arguments of the exponentials, we find the following

$$\begin{aligned} g(\varepsilon) \;=\; & \frac{m_c}{k_0\hbar\sqrt{2\pi}\sqrt{k_B T m_c}} \left[\exp\left(-\frac{\varepsilon}{k_B T} + \frac{\sqrt{2}k_0\hbar\sqrt{\varepsilon m_c}}{k_B T m_c} - \frac{k_0^2\hbar^2}{2m_c}\frac{1}{k_B T} \right) \right. \\ & \left. - \exp\left(-\frac{\varepsilon}{k_B T} - \frac{\sqrt{2}k_0\hbar\sqrt{\varepsilon m_c}}{k_B T m_c} - \frac{k_0^2\hbar^2}{2m_c}\frac{1}{k_B T} \right) \right] \end{aligned} \tag{C.16}$$

To interpret, this distribution is the difference of two exponential factors. At zero energy, the distribution becomes identically zero. About zero, a lowest-order series expansion in energy gives a $\sqrt{\varepsilon}$ dependence. At highest energies, the distribution falls off like a Boltzmann factor, as $exp(-\varepsilon/k_B T)$.

The energy at the distribution maximum is difficult to solve for, analytically. However, in most situations the drift energy tends to be smaller than the thermal/diffusive energy,

$$\frac{\hbar^2 k_0^{*2}}{2m_c} << \frac{3}{2}k_B T \tag{C.17}$$

Therefore, the distribution maximum is within the range,

$$\frac{\hbar^2 k_0^{*2}}{2m_c} < \varepsilon_{max} < (\frac{\hbar^2 k_0^{*2}}{2m_c} + \frac{3}{2}k_B T) \tag{C.18}$$

To summarize, we have formulated an approximate distribution function, possibly useful for describing the energy of carriers in an isotropic case.

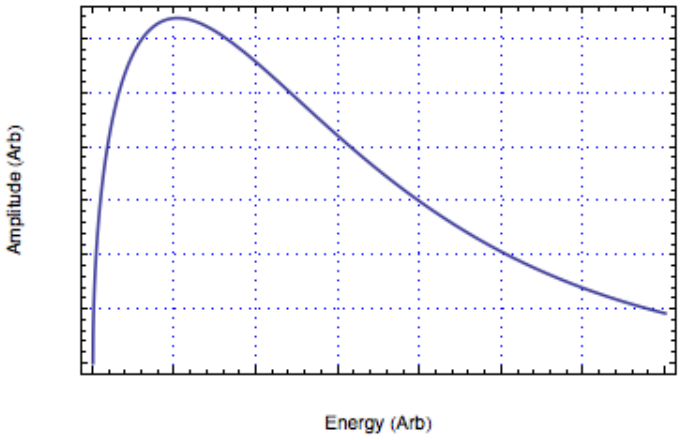

Figure C.1: A displaced Maxwellian distribution, which is similar to a Maxwell-Boltzmann distribution but with a drift component.

Appendix D

In Search of the Non-Equilibrium Chemical Potential

The Fermi level in semiconductors is not defined when equilibrium conditions do not exist. In this appendix, we continue to piece together all that we can make of this situation by looking at generation and recombination rates in bulk. For purposes of finding a steady-state condition when dealing with interfaces, we find a representation of a bulk mean energy when charge-neutral conditions exist. This "bulk neutrality level," "proxy Fermi level," or simply "non-equilibrium chemical potential" may be an adequate substitution for the Fermi level under some non-equilibrium conditions.

D.1 Balancing non-thermal rates deep in bulk (no gradients)

We balance generation and recombination rates of electrons, holes, donors, and acceptors. To start, we first treat the free carriers.

D.1.1 Balancing rates of electrons and holes in deep bulk

From previous work we know that electron and hole concentrations, deep in the bulk of a semiconductor, should asymptote to a steady state. This state is where the rates of generation and recombination are balanced.

In this case, the continuity equation for electrons is

$$\dot{n} = g - \frac{n}{\tau_n} \tag{D.1}$$

Quite simply, we find a static value where the rate change is zero. This is true also for holes.

$$\begin{aligned} n &= g\tau_n \\ p &= g\tau_p \end{aligned} \tag{D.2}$$

In this treatment, we continue to leave out recombination by excitons (electrons recombining to holes directly). Ignoring excitonic recombination not only simplifies the equations, but it also represents an interesting departure from typical assumptions in device literature. We believe ignoring excitions is representative of our typical device conditions. However, it may prove beneficial to consider them later.

For this note, we further assume no probability for impact ionization.

The recombination rate in our simple model of shallow impurities will include terms for both the ionized and neutral capture.

For this note, we commit a common crime in that we pretend capture rates are uncorrelated and energy-independent. We can then simply add rates for ionized and neutral capture by way of Mathieson's rule. For electrons, these rates are given by products of thermal velocities, capture cross-sections, and trap densities of ionized donors and neutral acceptors. For convenience, the products of thermal velocities and capture-cross sections are grouped into "c_n..." capture coefficients. The following rate is the inverse relaxation time for electrons to recombine both to ionized donors and to neutral acceptors.

$$\begin{aligned} 1/\tau_n &= 1/\tau_{nD+} + 1/\tau_{nA0} \\ 1/\tau_n &= \sigma_{nD+} v_{th_n} N_D^+ + \sigma_{nA0} v_{th_n} N_A^0 \\ 1/\tau_n &= c_{nD+} N_D^+ + c_{nA0} N_A^0 \end{aligned} \tag{D.3}$$

We also express the corresponding capture rate of holes to ionized acceptors and neutral donors.

$$\begin{aligned} 1/\tau_p &= 1/\tau_{pA-} + 1/\tau_{pD0} \\ 1/\tau_p &= \sigma_{pA-} v_{th_p} N_A^- + \sigma_{pD0} v_{th_p} N_D^0 \\ 1/\tau_p &= c_{pA-} \quad N_A^- + c_{pD0} \quad N_D^0 \end{aligned} \tag{D.4}$$

So in deep bulk, carrier concentrations could be said to be the following.

$$\begin{aligned} n &= g\left(\frac{1}{c_{nD+} N_D^+ + c_{nA0} N_A^0}\right) \\ p &= g\left(\frac{1}{c_{pA-} N_A^- + c_{pD0} N_D^0}\right) \end{aligned} \tag{D.5}$$

D.1.2 Balance of shallow trap rates

Now we consider the case for steady-state populations of donors and acceptors.

We begin with donors. By capture of carriers, we see that the evolution of donor concentration is the following.

$$\dot{N}_D^+ = c_{pD0} N_D^0 p - c_{nD+} N_D^+ n \tag{D.6}$$

For a steady state, we consider that

$$c_{pD0} N_D^0 p = c_{nD+} N_D^+ n \tag{D.7}$$

The corresponding evolution of acceptors is the following.

$$\dot{N}_A^- = c_{nA0} N_A^0 n - c_{pA-} N_A^- p \tag{D.8}$$

For this acceptor steady state, we consider that

$$c_{nA0} N_A^0 n = c_{pA-} N_A^- p \tag{D.9}$$

In order to express the variables N_D^+ and N_A^-, we will make the substitution relating neutral traps to the total number of traps.

$$\begin{aligned} N_D^0 &= N_D - N_D^+ \\ N_A^0 &= N_A - N_A^- \end{aligned} \tag{D.10}$$

Note that this excludes the possibility of "overcharged" states in this model. As mentioned in previous work, we ignorantly claim the capture cross-section for overcharged capture to be negligible by argument of extremely small potential well energies.

D.1.3 Combining steady-state rates for carriers and shallow traps

Plugging in the steady-state rates of carrier recombination into the equations for traps, we find the following two equations to solve. For donors,

$$\frac{g c_{\mathrm{pD0}} \left(N_D - N_D^+\right)}{c_{\mathrm{pA\,-}}\, N_A^- + c_{\mathrm{pD0}} \left(N_D - N_D^+\right)} = \frac{g c_{\mathrm{nD\,+}}\, N_D^+}{c_{\mathrm{nA0}} \left(N_A - N_A^-\right) + c_{\mathrm{nD\,+}}\, N_D^+} \tag{D.11}$$

and for acceptors,

$$\frac{g c_{\mathrm{nA0}} \left(N_A - N_A^-\right)}{c_{\mathrm{nA0}} \left(N_A - N_A^-\right) + c_{\mathrm{nD\,+}}\, N_D^+} = \frac{g c_{\mathrm{pA\,-}}\, N_A^-}{c_{\mathrm{pA\,-}}\, N_A^- + c_{pD0} \left(N_D - N_D^+\right)} \tag{D.12}$$

To keep the entropy in our expressions to a minimum, define now a *fraction of ionized traps*, which we will call ζ.

For the fraction of ionized donors, we will use the following.

$$N_D^+ = \zeta_D \; N_D \tag{D.13}$$

The fraction of ionized acceptors will likewise be defined as

$$N_A^- = \zeta_A \; N_A \tag{D.14}$$

Using "ζ" terms in our expressions and solving, we find the following expression for the fraction of ionized acceptors.

$$\zeta_A = \frac{N_A^-}{N_A} = \left(1 - \frac{\zeta_D c_{\mathrm{nD}} + c_{\mathrm{pA}}}{(\zeta_D - 1) c_{\mathrm{nA0}} c_{\mathrm{pD0}}}\right)^{-1} \tag{D.15}$$

We now need to interpret this expression.

D.1.4 Interpreting the ionization fraction of traps.

Let's examine this expression more thoroughly.

$$\zeta_A = \frac{N_A^-}{N_A} = \left(1 - \frac{\zeta_D c_{\mathrm{nD}} + c_{\mathrm{pA}}}{(\zeta_D - 1) c_{\mathrm{nA0}} c_{\mathrm{pD0}}}\right)^{-1} \tag{D.16}$$

As ζ_D becomes identically zero, the fraction of ionized donors becomes unity.

$$\lim_{\varsigma_D \to 0} \zeta_A = 1 \tag{D.17}$$

Conversely, as ζ_D becomes one we have

$$\lim_{\varsigma_D \to 1} \zeta_A = 0 \tag{D.18}$$

So quite obviously we have a pronounced inverse relationship. Now it happens that extremely small but positive definite values of ζ_D give ζ_A values constant near unity until ζ_D becomes much larger than a value of

$$\zeta_D \leq \frac{1}{\frac{c_{\mathrm{nD+}} c_{\mathrm{pA-}}}{c_{\mathrm{nA0}} c_{\mathrm{pD0}}} - 1} \approx \frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{nD+}} c_{\mathrm{pA-}}} \approx 10^{-8} \quad \text{(in order for } \zeta_A \approx 1\text{)} \tag{D.19}$$

We are free to neglect the one in the denominator as it becomes trivial compared to the ratio of capture rates. The rates for ionized impurity trapping are several magnitudes higher than those for neutral traps.

As ζ_D approaches unity, ζ_A inversely follows until ζ_A becomes close to $\frac{c_{\mathrm{nA0}} \; c_{\mathrm{pD0}}}{c_{\mathrm{nD+}} \;\; c_{\mathrm{pA-}}}$. As ζ_D becomes identically one, ζ_A abruptly falls to zero.

So as long as both ζ_D and ζ_A are somewhat larger than $\frac{c_{\mathrm{nA0}} \; c_{\mathrm{pD0}}}{c_{\mathrm{nD+}} \;\; c_{\mathrm{pA-}}}$, we find the ionization ratios follow a simple inverse relation.

$$\zeta_A \approx \frac{1}{\zeta_D} \frac{c_{\text{nA0}}\ c_{\text{pD0}}}{c_{\text{nD+}}\ c_{\text{pA-}}} \tag{D.20}$$

To summarize our work up to this point, we can posit this as

$$\boxed{\zeta_D \zeta_A = (\frac{N_D^+}{N_D})(\frac{N_A^-}{N_A}) \approx \frac{c_{\text{nA0}}\ c_{\text{pD0}}}{c_{\text{nD+}}\ c_{\text{pA-}}} \approx 10^{-8}} \tag{D.21}$$

D.2 Using charge neutrality in the bulk

What else can we say? We could assume that in deep bulk that charge neutrality would hold. In this case, we would say

$$n + N_A^- = p + N_D^+ \tag{D.22}$$

Inserting our steady-state expressions from the beginning, we find that solving for N_D^+ is still difficult. Solving for N_D^+ is not analytically approachable as it becomes a 5th degree polynomial. Our nice approximation from the previous section does not help, as the polynomial becomes then degree six.

Perhaps we can take this in approximation. We re-write.

$$n + N_A^- - p - N_D^+ = 0 \tag{D.23}$$

This becomes

$$\frac{g}{c_{\text{nA0}}\left(N_A - N_A^-\right) + c_{\text{nD+}}\ N_D^+} + N_A^- - \frac{g}{c_{\text{pA-}}\ N_A^- + c_{\text{pD0}}\left(N_D - N_D^+\right)} - N_D^+ = 0 \tag{D.24}$$

In the above expression, we take advantage of our ionization fraction approximation.

$$N_A^- = \frac{c_{\text{nA0}}\ c_{\text{pD0}}}{c_{\text{nD+}}\ c_{\text{pA-}}} \frac{N_A N_D}{N_D^+} \tag{D.25}$$

So we should solve for N_D^+ from the following equation.

$$\frac{g}{c_{\text{nA0}}\left(N_A - N_A^-\right) + c_{\text{nD+}}\ N_D^+} + \frac{c_{\text{nA0}}\ c_{\text{pD0}}}{c_{\text{nD+}}\ c_{\text{pA-}}} \frac{N_A N_D}{N_D^+} - \frac{g}{c_{\text{pA-}}\ N_A^- + c_{\text{pD0}}\left(N_D - N_D^+\right)} - N_D^+ = 0 \tag{D.26}$$

Now, if we *series expand* this expression in N_D^+ to order one, we find the following.

$$\left(\frac{c_{\text{nA0}} c_{\text{pD0}} N_A N_D}{c_{\text{nD +}}\ c_{\text{pA -}}}\right) \frac{1}{N_D^+} - \left(\frac{c_{\text{pD0}} N_A N_D c_{\text{nA0}}^2 + g c_{\text{nD -}}\ c_{\text{nA0}} + g c_{\text{nD +}}\ c_{\text{pA +}}}{c_{\text{nA0}}^2 c_{\text{pD0}} N_A N_D}\right) N_D^+ = 0 \tag{D.27}$$

Grouping N_D^+, we see that

$$\left(N_D^+\right)^2 = \frac{\left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}} N_A N_D}{c_{\mathrm{nD+}}\ c_{\mathrm{pA\text{-}}}}\right)}{\left(\frac{c_{\mathrm{pD0}} N_A N_D c_{\mathrm{nA0}}^2 + g c_{\mathrm{nD\text{-}}}\ c_{\mathrm{nA0}} + g c_{\mathrm{nD+}}\ c_{\mathrm{pA+}}}{c_{\mathrm{nA0}}^2 c_{\mathrm{pD0}} N_A N_D}\right)} \tag{D.28}$$

We regroup this expression.

$$\left(N_D^+\right)^2 = \left(\frac{c_{\mathrm{nD}\ +}\ c_{\mathrm{pA}\ \text{-}}}{c_{\mathrm{nA0}} c_{\mathrm{pD0}} N_A N_D} + \frac{g c_{\mathrm{pA}\ \text{-}}^2\ c_{\mathrm{nD}\ +}^2\ + g c_{\mathrm{nA0}} c_{\mathrm{pA}\ \text{-}}\ c_{\mathrm{nD}\ +}^2}{c_{\mathrm{nA0}}^3 c_{\mathrm{pD0}}^2 N_A^2 N_D^2}\right)^{-1} \tag{D.29}$$

For reasonable values, the predominance of neutral trapping probabilities makes the second term in parentheses several orders of magnitude smaller than the first term. It is reasonable to say,

$$\left(N_D^+\right)^2 = \left(\frac{c_{\mathrm{nD+}} c_{\mathrm{pA-}}}{c_{\mathrm{nA0}} c_{\mathrm{pD0}} N_A N_D} + \underbrace{\frac{g c_{\mathrm{pA-}}^2 c_{\mathrm{nD+}}^2 + g c_{\mathrm{nA0}} c_{\mathrm{pA-}} c_{\mathrm{nD+}}^2}{c_{\mathrm{nA0}}^3 c_{\mathrm{pD0}}^2 N_A^2 N_D^2}}_{\approx 0}\right)^{-1} \approx \left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{nD+}}\ c_{\mathrm{pA-}}}\right) N_A N_D \tag{D.30}$$

With these approximations, we find the rather balanced result

$$\boxed{N_D^+ = N_A^- = \sqrt{\left(\frac{c_{\mathrm{nA0}}\ c_{\mathrm{pD0}}}{c_{\mathrm{nD+}}\ c_{\mathrm{pA-}}}\right) N_A N_D}} \tag{D.31}$$

As we are claiming bulk neutrality, it should not be all that surprising that the population of ionized donors and ionized acceptors is balanced. What is different about this general scenario, versus true thermal equilibrium, is that both populations of traps remain highly neutralized. Unlike a thermal equilibrium situation, the population of electrons and holes were *generated from the intrinsic bulk*, yet they *recombine only to impurities*.

D.3 So what, then, of this Fermi level substitution?

The conditions of the previous section indicate that the bulk is rather symmetric in its charged constituency. What, then, can we say about some singular, characteristic energy? From other work, we know that such a concept would be quite useful when considering the interface to an external system.

Equilibration to some external system would be performed by the mobile carriers. In other words, free electrons and holes are the agents of interaction. The kinetic energy distributions of the electrons and holes span small energies in comparison to the germanium band gap. We therefore take electrons to have an energy of the conduction band, and holes to have energy of the valance band. Our Fermi-level proxy, E_F^*, is a weighted mean of the carrier energy.

$$E_F^* = \frac{n\ E_C + p\ E_V}{n+p} \tag{D.32}$$

We know that steady-state values in the bulk are determined by a generation-recombination balance.

$$E_F^* = \frac{(g\tau_n)\ E_C + (g\tau_p)\ E_V}{(g\tau_n) + (g\tau_p)} = \frac{\tau_n\ E_C + \tau_p\ E_V}{\tau_n + \tau_p} \tag{D.33}$$

From the results of the previous section, we use the following.

$$\begin{aligned}
\tau_n &= \left\{ c_{\mathrm{nD+}} \sqrt{\left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{pA-}} c_{\mathrm{nD+}}}\right) N_A N_D} + c_{\mathrm{nA0}} \left(N_A - \sqrt{\left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{pA-}} c_{\mathrm{nD+}}}\right) N_A N_D} \right) \right\}^{-1} \\
\tau_p &= \left\{ c_{\mathrm{pA-}} \sqrt{\left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{pA-}} c_{\mathrm{nD+}}}\right) N_A N_D} + c_{\mathrm{pD0}} \left(N_D - \sqrt{\left(\frac{c_{\mathrm{nA0}} c_{\mathrm{pD0}}}{c_{\mathrm{pA-}} c_{\mathrm{nD+}}}\right) N_A N_D} \right) \right\}^{-1}
\end{aligned} \tag{D.34}$$

This is a little difficult to deal with unless we again take some approximations.

D.3.1 Assume symmetry between electrons and holes.

We've been doing this all along, more or less. This will be interesting to re-question later, with expectation values from microscopic transport plugged in. in my retirement, in other words. For now, we just say neutralization capture probabilities are similar for electrons and holes. Likewise, ionized trap recombination is similar for electrons and holes.

$$\begin{aligned}
c_0 &= c_{\mathrm{nA0}} = c_{\mathrm{pD0}} \\
c_I &= c_{\mathrm{nD+}} = c_{\mathrm{pA-}}
\end{aligned} \tag{D.35}$$

D.3.2 Recognize that the capture probability by a neutral impurity concentration is small.

Knowing that c_0 is small, it gives us reasonable variable with which to series expand. Expanding in terms of c_0, the first two terms of our expression become

$$E_F^* \approx \frac{\left(N_D + \sqrt{N_A N_D}\right) E_C + \left(N_A + \sqrt{N_A N_D}\right) E_V}{N_D + N_A + 2\sqrt{N_A N_D}} + \underbrace{\frac{\sqrt{N_A N_D}\,(N_D - N_A)\,(E_C - E_V)}{\left(N_D + N_A + 2\sqrt{N_A N_D}\right)^2}\left(\frac{c_0}{c_I}\right)}_{\approx 0} \tag{D.36}$$

The first term dominates the expression. We find the rewarding result that

$$\boxed{E_F^* \approx \frac{\left(N_D + \sqrt{N_A N_D}\right) E_C + \left(N_A + \sqrt{N_A N_D}\right) E_V}{N_D + N_A + 2\sqrt{N_A N_D}}} \tag{D.37}$$

For balanced impurity concentrations, this expression returns a proxy Fermi level around mid-gap ($E_F^* \approx \frac{E_C + E_V}{2}$). However, unbalanced impurity concentrations may allow the proxy Fermi level to veer substantially towards one or the other band-edge. This occurs even though *the vast majority of impurities remain neutralized*. One might ask how this energy level can veer to one band edge or the other if there is overall charge neutrality. What matters is that the populations of *electrons and holes* can be substantially unbalanced. This is because there are *so many* impurities – even counting just the ionized ones – that the burden of charge neutrality is taken up predominantly by impurities. This is true even when the electron/hole generation rate is on the high side of reasonable.

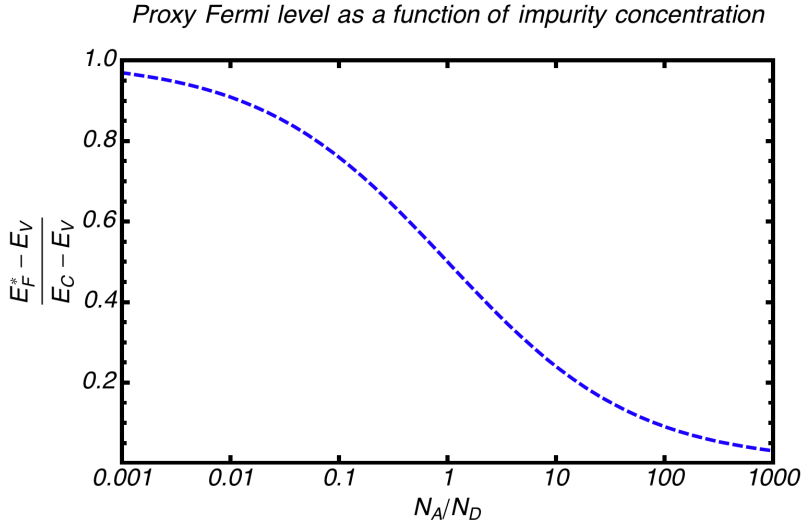

Figure D.1: The proxy Fermi level, plotted normalized.

Considering a straightforward *metal-germanium* interface, a typical Schottky-induced surface level is pinned around 0.1 eV above the valence band, E_V. At ultra-low temperature, this represents about $0.14 \times (E_C - E_V)$, as we have a 0.74 eV gap.

As per our proxy Fermi level prediction, we would need $N_A/N_D > 40$ to find the bulk of such a sample veer to p-type behavior near this contact.

D.4 Conclusion

We predict a steady-state condition based on a non-equilibrium chemical potential, which depends primarily on impurity concentrations. Knowledge of surface state equilibrium energy is also required for full-out "band-bending" predictions.

Nothing in this short note is intended to be conclusive. It is expected that the conservation of recombination current at the contact should result in appropriate band-bending to meet this proxy Fermi level. To be theoretically conclusive, showing this requires the simultaneous solution of Poisson's equation, as well as continuity equations for electrons and holes. This work is underway, but the solution need not lend itself to a general closed form and will to some degree rely on numerical solutions.

The degree to which the proxy Fermi level holds valid may also depend upon bias conditions. Some caveats mentioned above would need to be reconsidered under bias-dependent conditions. The strict charge neutrality condition in the bulk should be scrutinized. In addition, gradient terms in the continuity equations, due to drift or diffusion currents, could complicate this derivation.

To conclude, it is a good guess that a considerable amount of detector phenomenology likely originates from band-alignment of a given contact's work function to this level. A lot of charge collection behavior – dynamic over position and in time – could possibly be interpreted as due to the regions of accumulated charges required for this band-alignment.

Appendix E

Appendix: Simple 1D Drift-Diffusion Equations, for Simple Insight

The purpose of this section is to demonstrate a simple understanding from a simple treatment of drift - diffusion equations. By understanding how electron and hole concentrations are distributed within a simple one - dimensional slab, we can infer how bound space charge distributions may accrue.

We examine a few simple relations, adding slightly more sophistication along the way. What we find is that, for uniform generation throughout the bulk, we tend to find that there are only a few parameters which describe how carriers are distributed across the crystal. The nature of a "drift velocity" may pull carriers to one side, but the constant bulk generation rate sees to it that carriers are populous throughout the crystal in the steady state.

Carriers are generated in the bulk with uniform illumination. Carrier-carrier interaction is neglected. The surfaces at $z = 0$ and $z = L$ are assumed to be strongly absorbing, so carrier concentrations are assumed to be zero there. Figure E.1 is illustrative of the system. The dynamics of carriers in the bulk are described by drift and diffusion currents. While carriers transit the detector bulk, they may also trap and recombine into bound states.

We treat the the drift velocity and diffusion constants with asserted values, appropriate for $\sim V/cm$ sort of fields. Here, nothing is modeled self-consistantly with an electric field. However, if we consider that the effect of space charge upon the total field in the crystal is initially a perturbation, the distributions we solve for here still bring us insight as to how built-in fields develop while detectors are in operation. So, solving for the distribution of free carriers tells us where charge bound states are initially occurring.

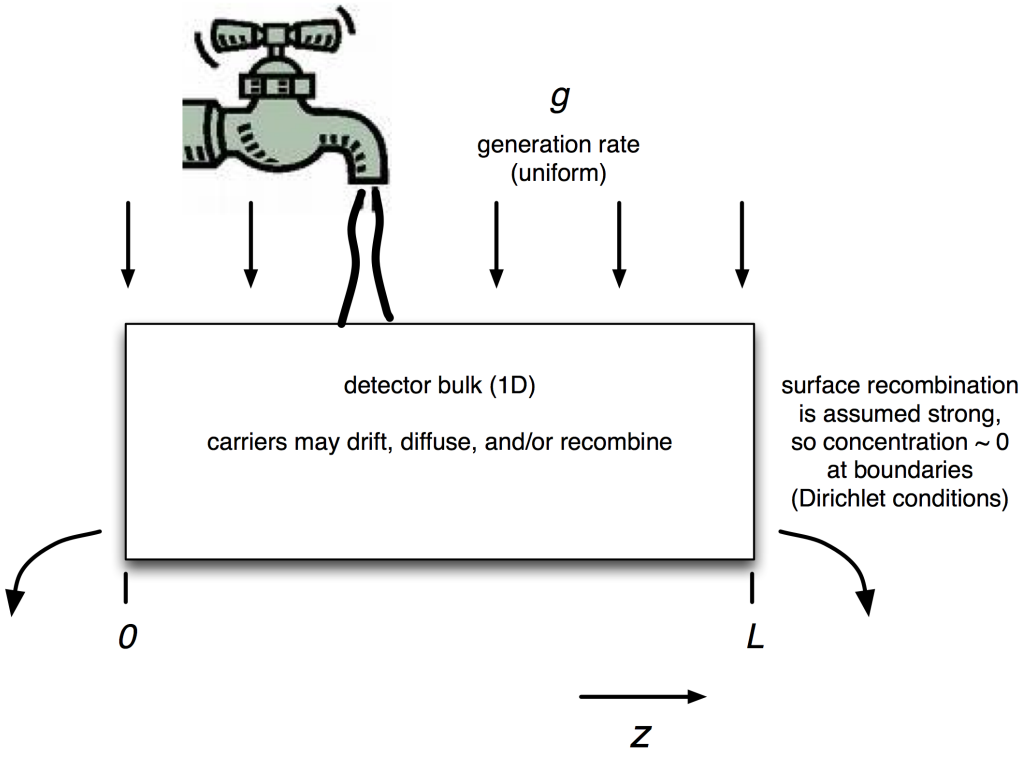

Figure E.1: A depiction of the drift-diffusion equations, including generation/recombination rates, and boundary conditions

The 1D drift-diffusion equations for electrons (n) and holes (p) are the following.

$$\begin{aligned} \dot{n} &= v_{dn}\frac{\partial n}{\partial z} + D_n\frac{\partial^2 n}{\partial z^2} - \frac{n}{\tau_n} + g \\ \dot{p} &= -v_{dp}\frac{\partial p}{\partial z} + D_p\frac{\partial^2 p}{\partial z^2} - \frac{p}{\tau_p} + g \end{aligned} \tag{E.1}$$

We also make use of Poisson's equation throughout, using bound charge densities of positive (N_p) and negative (N_n) states.

$$\frac{\partial^2 V}{\partial z^2} = \frac{-\rho}{\varepsilon\varepsilon_0} = \frac{-|e|}{\varepsilon\varepsilon_0}\left(N_p - N_n\right) \tag{E.2}$$

The integration of bound charge is found by the simple recombination rates ($\frac{1}{\tau_n}$ & $\frac{1}{\tau_p}$) and integration time, τ_{int}.

$$\begin{aligned} N_n &= \int_0^{\tau_{\rm int}} \frac{dN_n}{dt}dt = \frac{n}{\tau_n}\tau_{\rm int} \\ N_p &= \int_0^{\tau_{\rm int}} \frac{dN_p}{dt}dt = \frac{p}{\tau_p}\tau_{\rm int} \end{aligned} \tag{E.3}$$

E.1 Parameters for a toy model

Here, we introduce some parameters for making some rough estimates for carriers under drift at the scale of $\sim\ V/cm$ fields.

Parameter	Value	Description
g	$10^4\ cm^{-3}s^{-1}$	uniform generation rate
v_{dn}	$2.6\times10^6\ \ cm\ s^{-1}$	electron drift velocity
D_n	$3,900\ cm^2\ s^{-1}$	electron longitudinal diffusion
τ_n	$20\ \mu s$	electron recombination time
s_n	$10^5\ cm\ s^{-1}$	electron surface recombination velocity
v_{dp}	$1.6\times10^6\ \ cm\ s^{-1}$	hole drift velocity
D_p	$38,000\ cm^2\ s^{-1}$	hole longitudinal diffusion
τ_p	$50\ \mu s$	hole recombination time
s_p	$10^5\ cm\ s^{-1}$	hole surface recombination velocity
L	$2.54\ cm$	detector length
$\epsilon\epsilon_0$	$1.417\times10^{-12}\ A\ s\ cm^{-1}\ V^{-1}$	germanium dielectric factor
τ_{int}	$600\ s$	an integrated exposure time of 10 minutes

E.2 Considering only diffusion and generation terms

Solving for carrier concentrations in the steady state ($\dot{n} = \dot{p} = 0$), we solve

$$\begin{aligned} D_n \frac{\partial^2 n}{\partial z^2} &= -g \\ D_p \frac{\partial^2 p}{\partial z^2} &= -g \end{aligned} \tag{E.4}$$

with the Dirichlet boundary conditions $n(0) = n(L) = p(0) = p(L) = 0$.

The solutions are

$$\begin{aligned} n(z) &= \frac{g}{2D_n}\left(Lz - z^2\right) \\ p(z) &= \frac{g}{2D_p}\left(Lz - z^2\right) \end{aligned} \tag{E.5}$$

which are parabolic in nature.

From these distributions, we can use Poisson's equation as in equation E.2 to immediately solve for the potential energy *due to only the carrier distributions*. Recall that *potential energy* is defined as $U = eV = -|e|V$, and retains the negative sign of the electron charge. We can then solve for the *integrated* space charge distribution as per equation set E.3. As this space charge is simply scaled by the integration time, the distribution of bound charge for this toy model replicates the same spatial distributions as the free carrier distributions. Figure E.2 depicts the free carrier distributions, as well as the resultant potential energy distributions.

The integrated potential terms go as

$$\begin{aligned} U_{Nn} &= \frac{e^2 g z(L^3 - 2Lz^2 + z^3)}{24 D_n \varepsilon \varepsilon_0}\left(\frac{\tau_{\text{int}}}{\tau_n}\right) \\ U_{Np} &= -\frac{e^2 g z(L^3 - 2Lz^2 + z^3)}{24 D_p \varepsilon \varepsilon_0}\left(\frac{\tau_{\text{int}}}{\tau_p}\right) \end{aligned} \tag{E.6}$$

As we would expect, these solutions are symmetric as are the carrier distributions.

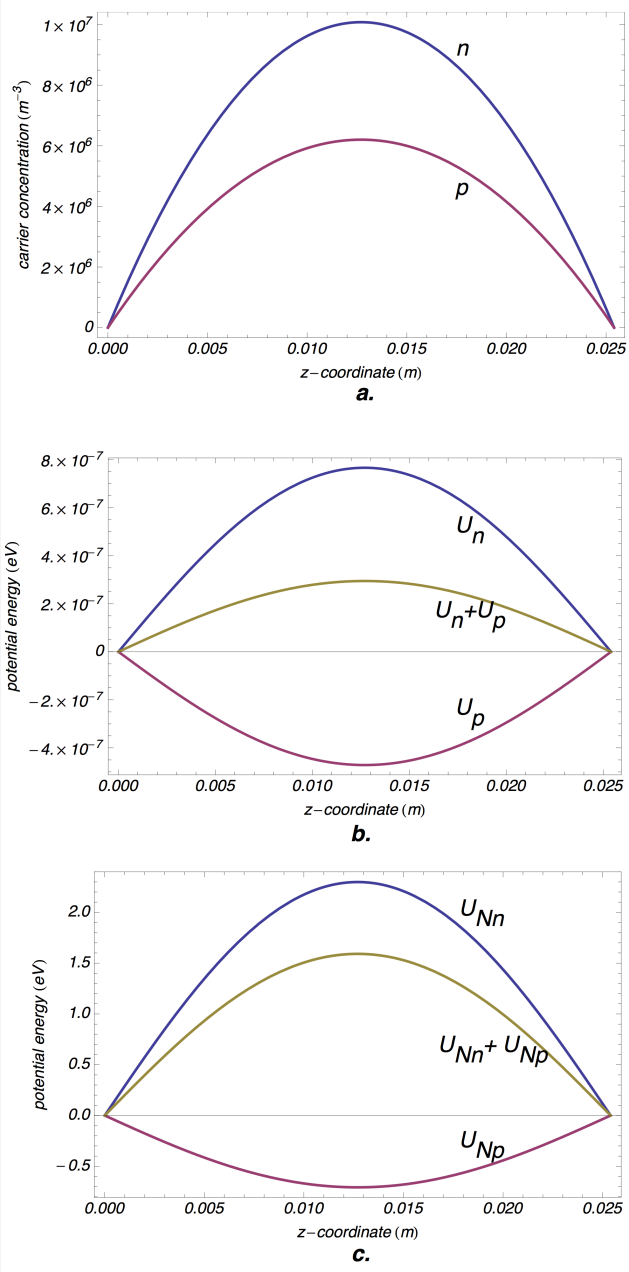

Figure E.2: The case for only diffusion and generation terms. **a.** electron and hole free carrier distributions in the steady state **b.** the potential energy terms due to the free carrier distributions **c.** the potential energy terms due to the space charge integrated onto bound bulk states

E.3 Considering drift, diffusion, and generation terms

Now we introduce the *drift* term back into our solutions.

$$\begin{aligned} \dot{n} &= v_{dn}\frac{\partial n}{\partial z} + D_n\frac{\partial^2 n}{\partial z^2} + g \\ \dot{p} &= -v_{dp}\frac{\partial p}{\partial z} + D_p\frac{\partial^2 p}{\partial z^2} + g \end{aligned} \tag{E.7}$$

For the steady state, we solve the following.

$$\begin{aligned} v_{dn}\frac{\partial n}{\partial z} + D_n\frac{\partial^2 n}{\partial z^2} &= -g \\ -v_{dp}\frac{\partial p}{\partial z} + D_p\frac{\partial^2 p}{\partial z^2} &= -g \end{aligned} \tag{E.8}$$

The general form for the free carrier distributions looks like the following.

$$\begin{aligned} n(z) &= C_{n1}\frac{D_n}{v_{dn}}e^{-\frac{v_{dn}z}{D_n}} - \frac{gz}{v_{dn}} + C_{n2} \\ p(z) &= C_{p1}\frac{D_p}{v_{dp}}e^{+\frac{v_{dp}z}{D_p}} + \frac{gz}{v_{dp}} + C_{p2} \end{aligned} \tag{E.9}$$

The constants "C_x" are determined by boundary conditions. Notice that the included terms go either proportionately with z, or by $exp(\mp v_d z/D)$.

Including our $n(0) = p(0) = n(L) = p(L) = 0$ boundary conditions, we find the following specific free carrier distributions.

$$\begin{aligned} n(z) &= -\frac{gL}{v_{dn}}\frac{e^{v_{dn}(L-z)/D_n}}{(e^{v_{dn}(L)/D_n}-1)} + \frac{g}{v_{dn}}\left(\frac{Le^{Lv_{dn}/D_n}+z-ze^{Lv_{dn}/D_n}}{e^{v_{dn}(L)/D_n}-1}\right) \\ p(z) &= -\frac{gL}{v_{dp}}\frac{e^{v_{dp}(+z)/D_p}}{(e^{v_{dp}(L)/D_p}-1)} + \frac{g}{v_{dp}}\left(\frac{L-z+ze^{Lv_{dp}/D_p}}{e^{v_{dp}(L)/D_p}-1}\right) \end{aligned} \tag{E.10}$$

So overall, we find a $\sim\ z$ "ramp-like" term which traverses the majority of the detector bulk and signifies that carriers are drifting to one side as they are being generated. This distribution is combined with the Euler-characteristic $exp(\mp v_d z/D)$ response very close to the boundary condition, where drift current into the contact balances with diffusive currents back into the bulk. Note that the characteristic length $D/v_d \approx 10\ \mu m$ for electrons, ($\approx 100 + \mu m$ for holes!) arises completely due to the balance between diffusion and drift currents. This is a unique length constant; in literature a "diffusion length" is typically defined in conjunction with the relaxation time, τ, as $\lambda_{diff} = \sqrt{D\tau}$. A translational drift length might be defined by $\lambda_{drift} = v_d\tau$. Here, where bulk relaxation occurs slowly, enough that it's contribution is negligible to the dynamical balance between drift and diffusion.

The resulting potential terms become more cumbersome than the case of diffusion+generation only, but they still retain a surprisingly symmetric – nearly parabolic – nature. Figure E.3 illustrates this.

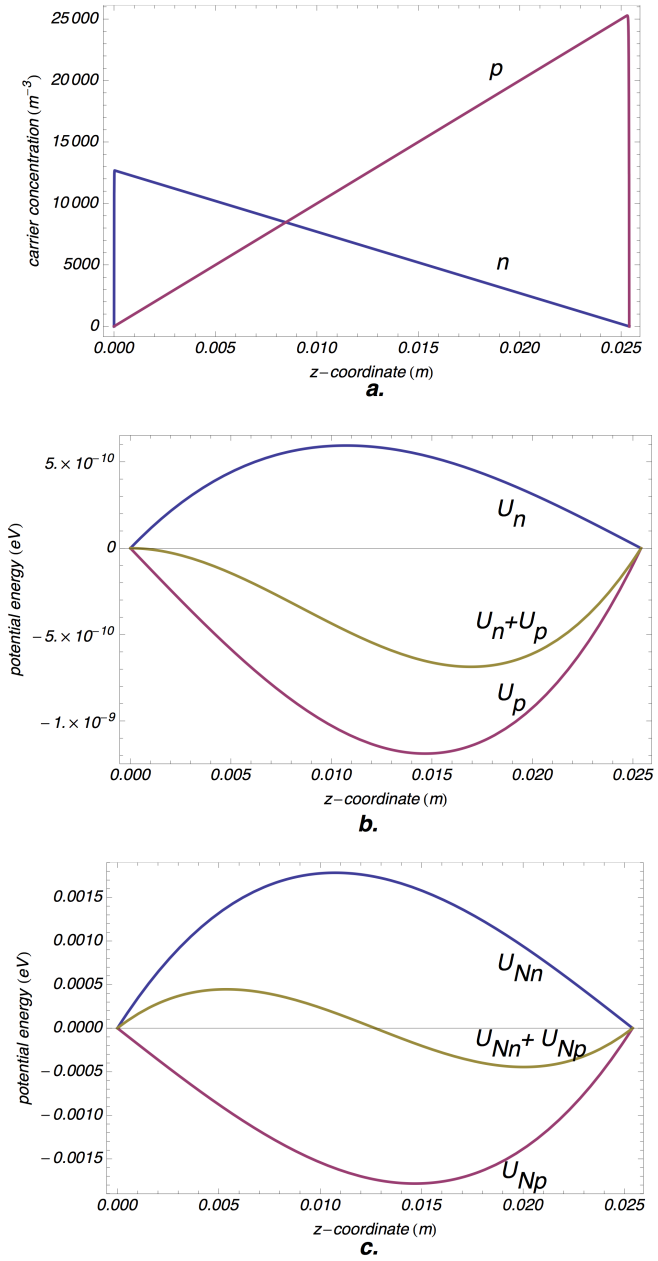

Figure E.3: The case for drift, diffusion, and generation terms. **a.** electron and hole free carrier distributions in the steady state **b.** the potential energy terms due to the free carrier distributions **c.** the potential energy terms due to the space charge integrated onto bound bulk states

E.4 Considering drift, diffusion, and generation as well as recombination terms

When we consider the effect of the recombination times τ_n and τ_p, we find that they have little impact on the shape and amplitude of the free carrier distributions unless these constants approach times about as small as the drift transit times, L/v_{dn} or L/v_{dp}, respectively. Of course, the relaxation times do impact the space charge amplitudes only because the integration of free carriers onto bound bulk states is directly proportional to the relaxation times.

As these functions are becoming cumbersome, we only point out the free carrier distributions without regard to boundary conditions.

$$\begin{aligned} n(z) &= g\tau_n + C_{n1}\exp\left(-\frac{z\left(v_{dn}+\sqrt{\frac{4D_n+v_{dn}^2\tau_n}{\tau_n}}\right)}{2D_n}\right) + C_{n2}\exp\left(-\frac{z\left(v_{dn}-\sqrt{\frac{4D_n+v_{dn}^2\tau_n}{\tau_n}}\right)}{2D_n}\right) \\ p(z) &= g\tau_p + C_{p1}\exp\left(-\frac{z\left(v_{dp}-\sqrt{\frac{4D_p+v_{dp}^2\tau_p}{\tau_p}}\right)}{2D_p}\right) + C_{p2}\exp\left(-\frac{z\left(v_{dp}+\sqrt{\frac{4D_p+v_{dp}^2\tau_p}{\tau_p}}\right)}{2D_n}\right) \end{aligned} \tag{E.11}$$

Considering a finite τ, note that the arguments in these exponentials will vary between characteristic lengths of a traditional *diffusion length*, $\sqrt{D\tau}$, to a *drift length*, $v_d\tau$. We are typically in a "drift-like" versus a "diffusion-like" scenario as

$$v_d\tau \gg \sqrt{D\tau} \tag{E.12}$$

under most circumstances for large τ.

Once again, for any impact on the shape of free carrier distributions, we must consider *very short* relaxation times. In figure E.4, we tune down constants to $\tau_n = 2\ \mu s$ and $\tau_p = 4\ \mu s$

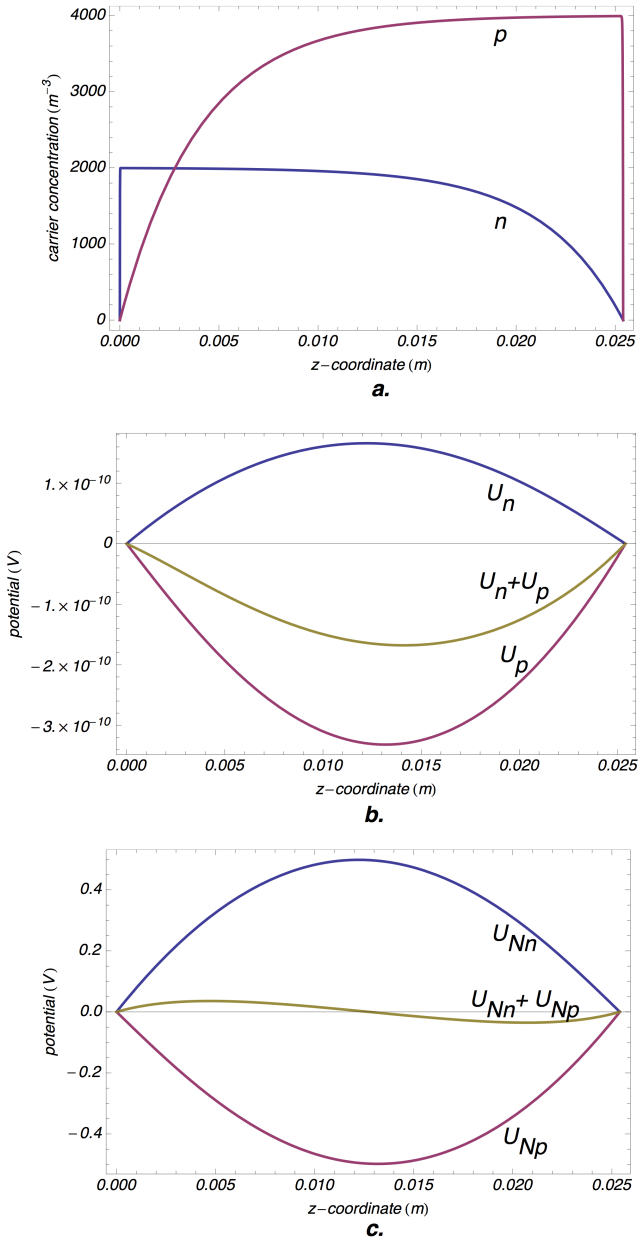

Figure E.4: The case for drift, diffusion, generation, as well as recombination terms. For a pronounced difference in the relative *shape* of the distributions, we had to substantially shorten the recombination times τ_n and τ_p. **a.** electron and hole free carrier distributions in the steady state **b.** the potential energy terms due to the free carrier distributions **c.** the potential energy terms due to the space charge integrated onto bound bulk states

E.5 Reconsideration of the boundary conditions

Rather than *Dirichlet* boundary conditions, it also important to consider where absorption of carriers into the boundaries is not *infinite*. In the case of *finite* absorption into the boundaries, we specify *Neumann* boundary conditions which is a constraint on the *carrier flux*. The current into the contact is specified by a *surface recombination velocity*. We denote s_n for electrons, and s_p for holes.

Considering only holes for illustrative purposes, the concentration current balance at $z = L$ and at $z = 0$ are the following.

$$\begin{aligned} v_{dp}p(L) - D_p\frac{\partial p(L)}{\partial z} &= s_p p(L) \\ -v_{dp}p(0) + D_p\frac{\partial p(0)}{\partial z} &= s_p p(0) \end{aligned} \tag{E.13}$$

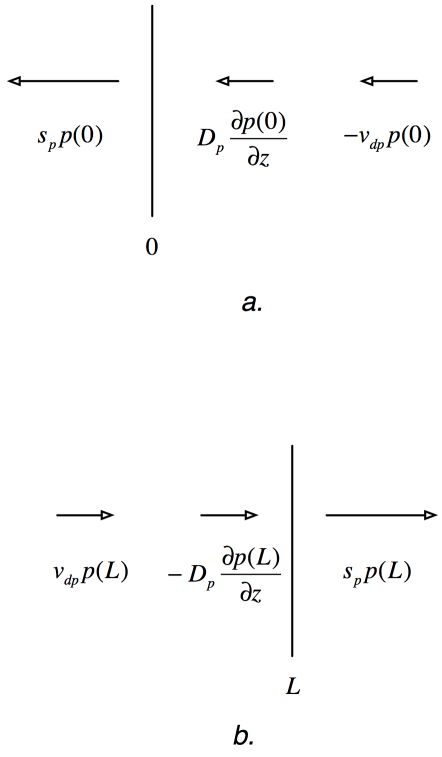

Figure E.5: Surface recombination boundary conditions, illustrated by considering hole current into the boundaries **a.** The current of holes balanced by surface recombination at $z = 0$. **b.** The current of holes balanced at $z = L$.

It is interesting to note when the *derivative* of the carrier concentration flips between decreasing at the contact (representing a local depletion to the boundary), versus an increasing derivative, representing *accumulation*.

For instance, at $z = L$ we can simply solve for $\frac{\partial p(L)}{\partial z} = 0$. In this case, we find that we have a *flat distribution* (no gradient) when $v_{dp} = s_p$. So when the drift velocity is exactly balanced by the surface recombination velocity the contact can accept, the free carrier distribution has no derivative at the contact.

For illustrative purposes, we show in figure the free carrier distribution in the extreme case where $s_n = s_p = 10^5\ cm/s$. This shows that free carriers are accumulating in a shock front immediately at the contact, subtending a width on order of length $\sim\ D/v_d$.

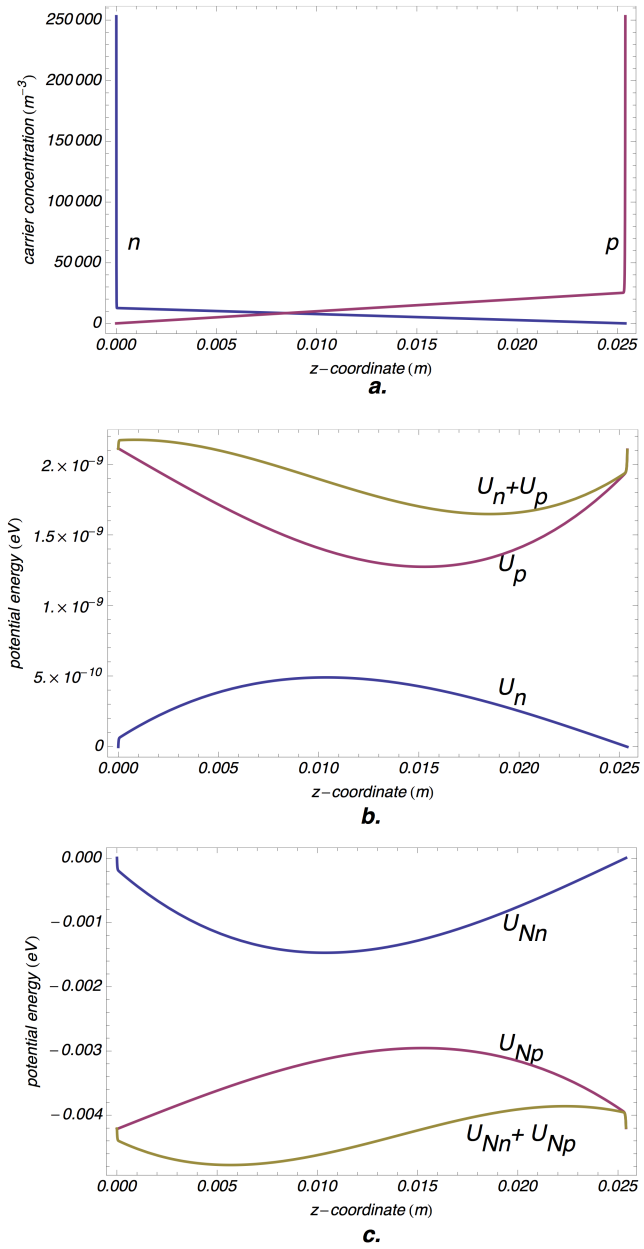

Figure E.6: The case full drift, diffusion, generation, and recombination, but with finite surface recombination velocities for both electrons and holes. **a.** electron and hole free carrier distributions in the steady state **b.** the potential energy terms due to the free carrier distributions **c.** the potential energy terms due to the space charge integrated onto bound bulk states

Appendix F

A Simple One-Dimensional Continuity Equation for Holes at a Contact

The purpose of this appendix is to explore analytical solutions to the continuity (or "drift-diffusion") equation at a boundary, for a single species. We wish to explore the form of the solutions under different assumptions, to build intuition. Arbitrarily, we chose to work with holes.

We have an interest in contacts to germanium with an assumed negative charge. To that end, we disregard electrons for now, and consider how a hole density may accumulate as a function of depth into the substrate. We solve the partial differential equation for holes, assuming no coupling to Poisson's equation. The electric field enters this treatment only as a drift velocity. The purpose is to establish an analytical solution which may offer some easily interpretable insight. We take insight into solving the continuity equation in this manner, following Balkanski and Wallis [61].

F.1 The continuity equation, in one dimension

The continuity equation for holes, in one dimension, is written as

$$\frac{dp}{dt} = -\mu\frac{dp}{dx}E - \mu p\frac{dE}{dx} + D\frac{d^2p}{dx^2} + g - r \tag{F.1}$$

Let's solve for equilibrium by setting $\frac{dp}{dt} \to 0$. To simplify our work, we also assume the electric field gradient is small, so $\frac{dE}{dx} \to 0$.

$$D\frac{d^2p}{dx^2} = \mu\frac{dp}{dx}E - g + r \tag{F.2}$$

F.1.1 Recombination

We could go on and on about recombination (capture) rates. There is a wealth of scattering-level input to be utilized here. For now, a simple recombination rate is assumed. Naively, we assume holes are trapping to neutral sites. As we disregard self-consitancy with Poisson's equation and ionized centers, the actual trapping mechanism is not important for this treatment.

We express the recombination rate in terms of a lifetime.

$$r = \frac{p}{\tau} \tag{F.3}$$

F.1.2 Drift velocity

We assume carriers are always in equilibrium with the electric field. The electric field will enter our idealized continuity equation only in this way.

We make the replacement,

$$v_d = \mu E \tag{F.4}$$

With this, we adjust our continuity equation.

F.2 Solving for the hole distribution

We now solve the following expression.

$$D\frac{d^2p}{dx^2} = (v_d)\frac{dp}{dx} + (\frac{1}{\tau})p - g \tag{F.5}$$

This has solutions of the form

$$p(x) = g\tau + p_1 e^{-x/L_{eff}} + p_2 e^{x/L_{eff}} \tag{F.6}$$

where p_1 and p_2 are constants and

$$L_{eff} = \frac{1}{2}\{v\tau + (4D\tau + v^2\tau^2)^{1/2}\}. \tag{F.7}$$

This effective length, L_{eff}, reduces to the *diffusion length*, $L_D = \sqrt{D\tau}$, when the drift velocity is zero.

To be physical, we throw out solutions of the growing exponential form. Therefore we set p_2 to zero. To finish, we must solve for the constant p_1. We can do this by incorporating a proper boundary condition.

F.3 Boundary conditions, and surface recombination velocity

We use the device-level concept of a *surface recombination velocity*, s, to aid us in our understanding the boundary conditions of a contact.

F.3.1 Balancing current density at the contact

Consider that the carrier flux, or the *current density*, out of the substrate is only partially absorbed on incidence.

The current density of outbound holes – meaning they are leaving the substrate and are incident at the contact – can be expressed by

$$j_{out} = (e)\frac{1}{4}v_{th}\ p. \tag{F.8}$$

Technically, we speak of an electric current, so we keep an electron charge e associated with this for now. Note it is the thermal velocity, v_{th}, that determines this stochastic flux. A factor of 1/4 arises through an expectation integral over solid angle (show this!).

Some fraction of this flux is scattered back into the substrate. We introduce a reflection coefficient, r_r, which is some fraction of unity. We use "r_r" with a subscript to keep it separate from "r," our recombination rate.

We also consider that the amount of carriers injected from the contact is a flux, S.

The current density of these *inbound* holes – entering from the contact – is given as

$$j_{in} = r_r(e\frac{1}{4}v_{th}\ p) + eS \tag{F.9}$$

We utilize equilibrium, solving for the *net flux* equal to zero,

$$j_{net} = j_{out} - j_{in} = (e)\{(\frac{1}{4}(1-r_r)v_{th}\ p\} - eS = 0 \tag{F.10}$$

or

$$S = \frac{1}{4}(1-r_r)v_{th}\ p \tag{F.11}$$

So to maintain the steady-state across the boundary, we need to observe this net flux, S.

We simplify this condition by introducing the *surface recombination velocity* as

$$s = \frac{1}{4}(1-r_r)v_{th}. \tag{F.12}$$

This quantity, by inspection, is some fraction of the thermal velocity. The steady-state flux becomes simply

$$S = sp. \tag{F.13}$$

F.4 Solving for the boundary condition

In this section, we make use of the current determined by drift and diffusion, and equate it to the flux equilibrium we deduced above.

The current density by drift and diffusion can be written as

$$\mathbf{j_p} = ep\mathbf{v_d} - eD\nabla p. \tag{F.14}$$

Define our boundary at $x = 0$. We consider the amplitude of flux at $x = 0$ in one dimension. Note that we express the flux moving *into* the contact, i.e., in the negative x direction.

$$(\frac{-j_p}{e})|_{x=0} = (-v_d p + D\frac{dp}{dx})|_{x=0} = -v_d p(0) + D\frac{(-p_1)}{L_{eff}} = -v_d(g\tau + p_1) - \frac{D}{L_{eff}}p_1 \tag{F.15}$$

The steady-state incident flux must be equal to the flux which is recombining. This was determined in the previous section.

$$S|_{x=0} = s\ p(0) = s\ (g\tau + p_1) \tag{F.16}$$

Equating, we find

$$v_d(g\tau + p_1) + \frac{D}{L_{eff}}p_1 = -s\ (g\tau + p_1) \tag{F.17}$$

which provides

$$p_1 = -g\tau\ (\frac{1}{1 + \frac{D}{L_{eff}(s+v_d)}}) \tag{F.18}$$

So we replace "p_1" in the expression for hole density as a function of depth, and find

$$p(x) = g\tau\ \{1 - (\frac{1}{1 + \frac{D}{L_{eff}(s+v_d)}})e^{-x/L_{eff}}\} \tag{F.19}$$

The density far off in the bulk asymptotes to

$$p(\infty) = g\tau \tag{F.20}$$

which is where generation balances recombination, and local drift/diffusion has no impact.

At $x = 0$, the density expression becomes

$$p(0) = g\tau(\frac{D}{D + L_{eff}(s + v_d)}) \tag{F.21}$$

which signifies a density deficit compared to the bulk, provided all constants are positive. This demonstrates that recombination at the contact is a lossy mechanism and is pulling carriers out of the ambient population at the substrate surface.

We were wondering about an *accumulation* layer of holes, which would screen the presence of negative charge at the contact. How does this work, then, if carriers are being *captured*, and their density *is thinned* near the contact? Perhaps the answer is that the electric field due to negative surface charge enters this expression by way of the drift velocity? To point out, a drift velocity due to negative charges at $x < 0$ requires a minus sign in our convention. This is because drift current of holes will be flowing backwards in this case. Can we reach an enhanced density over bulk at $x = 0$? Well, at $v_d = -s$ we see we at least break even. Is this really the solution we were looking for, though?

F.5 Equating the full current for the boundary condition

In the previous section, we assumed only the drift and diffusion terms mattered at the contact. In the literature, we often see that even the drift term is ignored. Can we call this well enough, then? Well, in the case of large CDMS detectors, we have a centimeter or more of bulk generating holes and electrons. All these carriers, modulo the ones that recombine, pile up at the contacts. So above, we have forgotten to integrate the carriers pulled from out of the deep bulk.

Let's go back to the continuity equation to define our current.

$$-\frac{dp}{dt} = \frac{1}{e}\nabla \cdot \mathbf{j_p} \tag{F.22}$$

What does this say about the negative-going, contact-incident current at $x = 0$?

Perhaps we say

$$-\frac{j_p(0)}{e} = \int_{L_{det}}^{0} (\frac{dp}{dt})dx \tag{F.23}$$

As we have a spatial integral now, we can define L_{det} is the length of the detector.

The negative-going flux at $x = 0$ is then

$$-\frac{j_p(0)}{e} = \{-v_d p + D\frac{dp}{dx} - g(x - L_{det}) + \frac{p}{\tau}(x - L_{det})\}|_{x=0} \tag{F.24}$$

giving

$$-\frac{j_p(0)}{e} = -v_d\ (g\tau + p_1) + D\frac{-p_1}{L_{eff}} + gL_{det} - \frac{(g\tau + p_1)}{\tau}L_{det} \tag{F.25}$$

We should balance this expression against the steady-state surface recombination to satisfy our boundary condition.

$$-v_d\ (g\tau + p_1) + D\frac{-p_1}{L_{eff}} + gL_{det} - \frac{(g\tau + p_1)}{\tau}L_{det} = s\ (g\tau + p_1) \tag{F.26}$$

So we find

$$p_1 = -g\tau(\frac{1}{1 + \frac{L_{\rm det}L_{\rm eff} + D\tau}{L_{\rm eff}(s\tau + v_d\tau)}}) \tag{F.27}$$

Using this coefficient back in our expression for $p(x)$,

$$p(x) = g\tau(1 - \frac{1}{1 + \frac{L_{\rm det}L_{\rm eff} + D\tau}{L_{\rm eff}(s\tau + v_d\tau)}}e^{-x/L_{\rm eff}}) \tag{F.28}$$

We see that at $x = 0$,

$$p(0) = \frac{g\tau\ (D\tau + L_{\rm eff}L_{\rm det})}{(D\tau + L_{\rm eff}L_{\rm det}) + L_{\rm eff}s\tau + L_{\rm eff}v_d\tau} \tag{F.29}$$

Which tells us that, again, the hole density at the contact is suppressed compared to the $p(\infty) = g\tau$ of the bulk, unless the condition that $v_d < -s$.

Is this as it should be? Is it only through the strong field (and position) dependence on v_d that we may get a strong hole distribution near the contact?

Strong recombination to the contact will still make the hole population tend to zero at $x = 0$.

This work has been an interesting aside, but not as informative as hoped. To end, we also point out possibilities for a simple numerical method.

F.6 An Euler-step method to numerically find the field & density distribution

We know the form of the density distribution, given a drift velocity from the electric field. If we have a known (negative) surface charge at $x = 0$, we know the field and drift velocity ($v_d = \mu\frac{-\sigma}{\epsilon}$) immediately at the surface. Therefore, we also have the numerical value for $p(0) = g\tau + p_1$. We can integrate in depth for the electric field using the density distribution.

$$E(x) = -\frac{\sigma}{\epsilon} + e\int_0^x p(E(x'), x')dx' \tag{F.30}$$

If we take infinitesmal steps, Δx, we might use a form such as

$$E(n+1) = E(n) + ep(E(n), n)\Delta x \tag{F.31}$$

If the desired accuracy is not met by this approximation, using a stock Runge-Kutta routine is also a possibility.

F.7 End remarks regarding this single-species model at a boundary

This toy model is very basic. Although we may use non-equilibrium values for the appropriate drift and diffusion constants, a fundamental flaw is that we assume a "semi-infinite bulk." In our very mobile and completely depleted (frozen-out) CDMS detectors, a depth of several centimeters into a substrate still does not constitute finding a steady-state bulk condition. The behavior of concentration dynamics on one side of the detector may still influence behavior all the way across. Nonetheless, this appendix still represents a healthy exercise for understanding the behavior of free carriers near an interface.

Appendix G

Hypothetical Space Charge Layers, and a Simple Dead Layer Model Based on Drift and Diffusion

This is an overview of a space charge layer understanding at contacts, assuming completely charged impurity regions. This approximation is appropriate where impurity charge states can be completely saturated by adequately high concentrations and diffusion profiles can be ignored. A treatment of 1D drift / diffusion dynamics then gives a response profile of carriers near a contact, leading to a formulation of an ideal dead layer.

G.1 The appearance of a space charge layer between a metal and semiconductor

Here is an explanation of space charge layer formation. This is common to many device physics references.

Consider the interface between two materials: a metal and a semiconductor. We could have instead considered the interface between two metals or two semiconductors. Why study the metal/semiconductor interface? The metal/semiconductor system is a good way to understand the energetic dynamics of one semiconductor to external conditions. A metal represents a sea of free electrons, with an energy level that is extremely hard to displace by steady-state processes. As it is complicated enough to understand one semiconductor's steady state while imposing external boundary conditions, a metal/semiconductor interface is a good system to study. In figure G.1 a., we draw as a function of position the band diagrams of the metal (left) and semiconductor (right). These materials we first draw as spatially separated. The potential energy of the metal, relative to the energy of the vacuum, is called the *work function*, and is denoted ϕ_m. A semiconductor has a forbidden energy gap between

it's valence and conduction bands, so referencing these two levels is just slightly more complicated. It is useful to define a work function also for a semiconductor, ϕ_s, which is the energy difference from vacuum to what is typically the equilibrium Fermi level. As an aside, the issue of what to do when we have NO Fermi level, (i.e., no thermal equilibrium) is the subject of another work. We also define the energy difference between the vacuum level and the *conduction band*, as a quantity referred to as the electron *affinity*, χ.

We bring the metal and semiconductor in proximity to each other. As the charge carriers in both materials are able to interact, we find generally a net difference in population density will occur. Either electrons or holes from the semiconductor will migrate to the metal in order for electrostatic charging to occur, producing an energy difference that brings both materials into steady-state. In other words, electrostatic charging aligns the Fermi levels of both materials. In figure G.1 b., we notice that the semiconductor has a Fermi level, denoted by the dashed line, which is rather high in the bandgap representing an n-type semiconductor. To lower this fermi level and match the metal's work function, electrons from the semiconductor must migrate into the metal. Once again, the metal represents a sea of free carriers, so this excess carrier concentration makes no appreciable change in the metal's band alignment. What used to be a net neutral semiconductor now has some region of net positive charge. A space charge region has now been produced inside of the semiconductor. If we solve Poisson's equation, we find that this boundary dipole has created the potential energy difference to align the Fermi level of the metal to that of the net neutral semiconductor. So we see that this electric displacement is a necessary feature of the steady-state, and we should expect it to occur at any metal/semiconductor interface. Our CDMS detectors are no exception. Figure G.1 c. demonstrates that an n-type semiconductor, with its positive space charge region, bends upward toward the interface with a metal or other material of a larger work function. The case is opposite for a p-type semiconductor.

The specific case of a junction between a metal and a semiconductor is called a *Schottky barrier*. "Space charge layer" can refer generally to the charge region in a semiconductor due to any surface or contact interface. So this is a good term to use. Actually, the case of a metal on an *intrinsic* semiconductor can be called a *Mott barrier*, but no one is going to know what you are talking about. CDMS germanium detectors are of course an epitome of intrinsic operation, so it is a shining opportunity for some sort of paper with "Mott barrier" in the title.

G.2 The space charge layer

In the last section, we found that a space charge layer produces band bending. We also found that each material on its own was net neutral, so we must have formed a balanced, electrostatic dipole. We can relate the energy difference overcome by

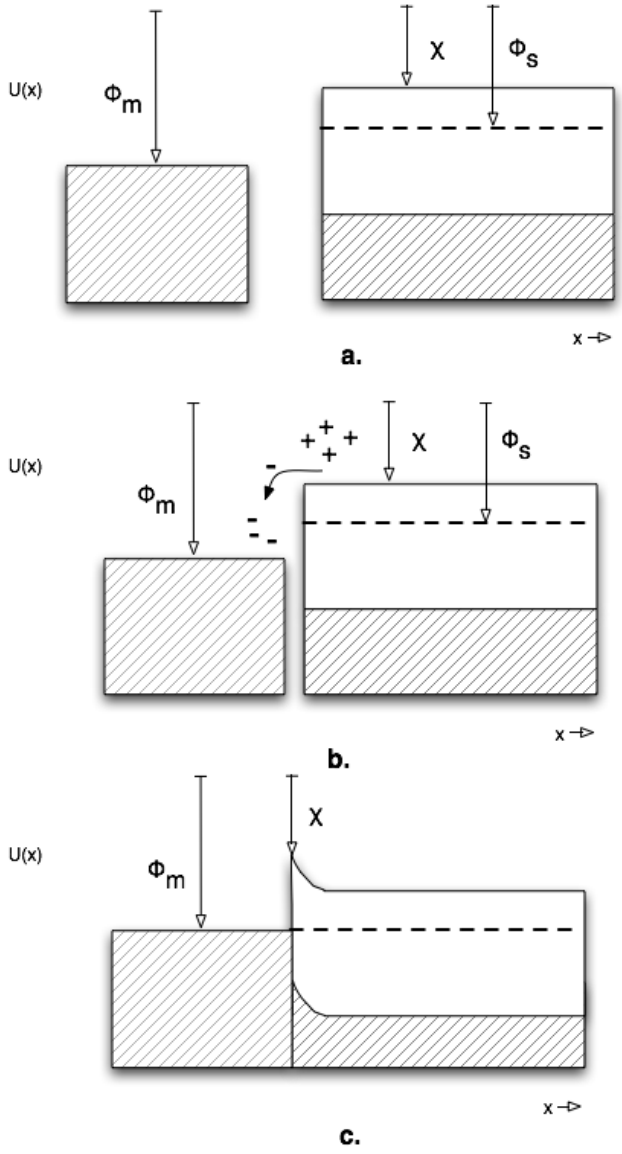

Figure G.1: a. metal and a semiconductor have work functions relative to vacuum energy, b. in proximity, carriers from the semiconductor will migrate to the metal and leave behind an imbalanced charge forming the dipole necessary for energy equalization. c. for the case of this n-type semiconductor, band-bending is upward toward the metal as the metal has a larger work function.

carrier displacement to find the charge needed.

Other works will describe charge accumulation under different conditions. Diffusion lengths are an important and interesting topic, and are likely important in describing charge accumulation near detector bare sidewalls. Here, we assume that some impurity, either donors or acceptors, will become completely charged over some rectangular spatial width to meet steady state conditions. This is a good assumption when microscopic drift fields are large, localized carriers to the interface before recombining and charging impurities. Therefore, a population of completely charged impurities will grow from the interface back into the semiconductor bulk until some equilibrium is reached.

Assume planar contacts and a constant charged density, N. Using Poisson's equation in 1D,

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon} \tag{G.1}$$

Assume that we have a charge density of completely charged donors, $\rho = N_D$. We can integrate from the metal boundary into the semiconductor to find

$$E = -\frac{d\phi}{dx} = \frac{eN_D}{\epsilon}x \tag{G.2}$$

In turn, if we solve for the full potential drop, we find

$$(\phi_m - \phi_s) = \phi_{ms} = -\frac{eN_D W^2}{2\epsilon} \tag{G.3}$$

where W is the width of the charge layer.

So,

$$W = \sqrt{\frac{2\epsilon\phi_{ms}}{eN_D}} \tag{G.4}$$

The expression is similar for p-type layers, as the quantity ϕ_{ms} becomes negative, as does the charge density.

We see we have an integrated charge at the interface of

$$Q_m = -Q_s = -eN_D W \tag{G.5}$$

Plugging in some numbers, we claim that a typical potential energy across the gap edge would be $\phi_{ms} = 0.1\ eV$. In our intrinsic detectors, assume $N_D = 10^{11}\ cm^{-}3$. We find an expected charge layer of $\approx 40\ \mu m$! This is huge. We *should* be prepared for values of ϕ_{ms} of this size, but it goes to show that our very pure crystals may force a large charge width to form, as it takes some depth to integrate enough charge density to screen the semiconductor interface.

While we are at it, consider also the *strength* of the electric field produced. We can solve for the maximum field, at the interface, and find for our example,

$$E_{max} = \frac{eN_D}{\epsilon} W = 2\phi_{ms}/W \tag{G.6}$$

which comes about to be $\approx 50\ V/cm$. This is an order of magnitude larger than our typical drift fields. So in fact, we would expect a zero-field point and a complete *reversal* of the drift field somewhere a distance W into the bulk of our detector, near one of the two contacts. Near which contact would depend on the polarity of the applied field and whether the crystal is n or p-type.

G.3 A simple device with both contacts

We illustrate these scenarios in figure G.2. We suggest a symmetric device as a detector, noting that we apply a voltage bias across it. The resulting conditions for either an n-type or p-type crystal show how carriers would react in the bulk. It is tempting to speculate if the reversed field region is responsible, above and beyond diffusion dynamics, for low-charge events or even reversed charge collection to some extent (at the percent level). Near-surface events could be affected by the field-reversal regions that would act as "collection gutters." To note, this naturally occurring space charge is a benefit to charge collection on the side where it enhances the field.

The evolution of space charge layers with time and with bias will definitely be important. The placement of the germanium non-equilibrium chemical potential (defining ϕ_s) can change with drift field amplitude. Integrated charge from the bulk can also "eat" away or contribute to these charge layers. Likely, though, a return to grounded conditions will re-enable these charge layers to evolve back into this steady state. Keep in mind that the space charge layer accumulated in the first place for the sake of energy equilibrium, for material interface reasons.

G.4 Peculiarities of germanium: level pinning, damage, and hydrogen

There are specific propensities of germanium that add to this device understanding.

G.4.1 Level pinning

Fermi level pinning is a property that comes up a lot in regards to Schottky barriers on particular semiconductors. Material properties of the metal/semiconductor interface can allow a large density of surface states to dominate the apparent work function of the interface as seen by the semiconductor. So absolute values of the work function of a metal, as measured for instance by photoemission techniques, may

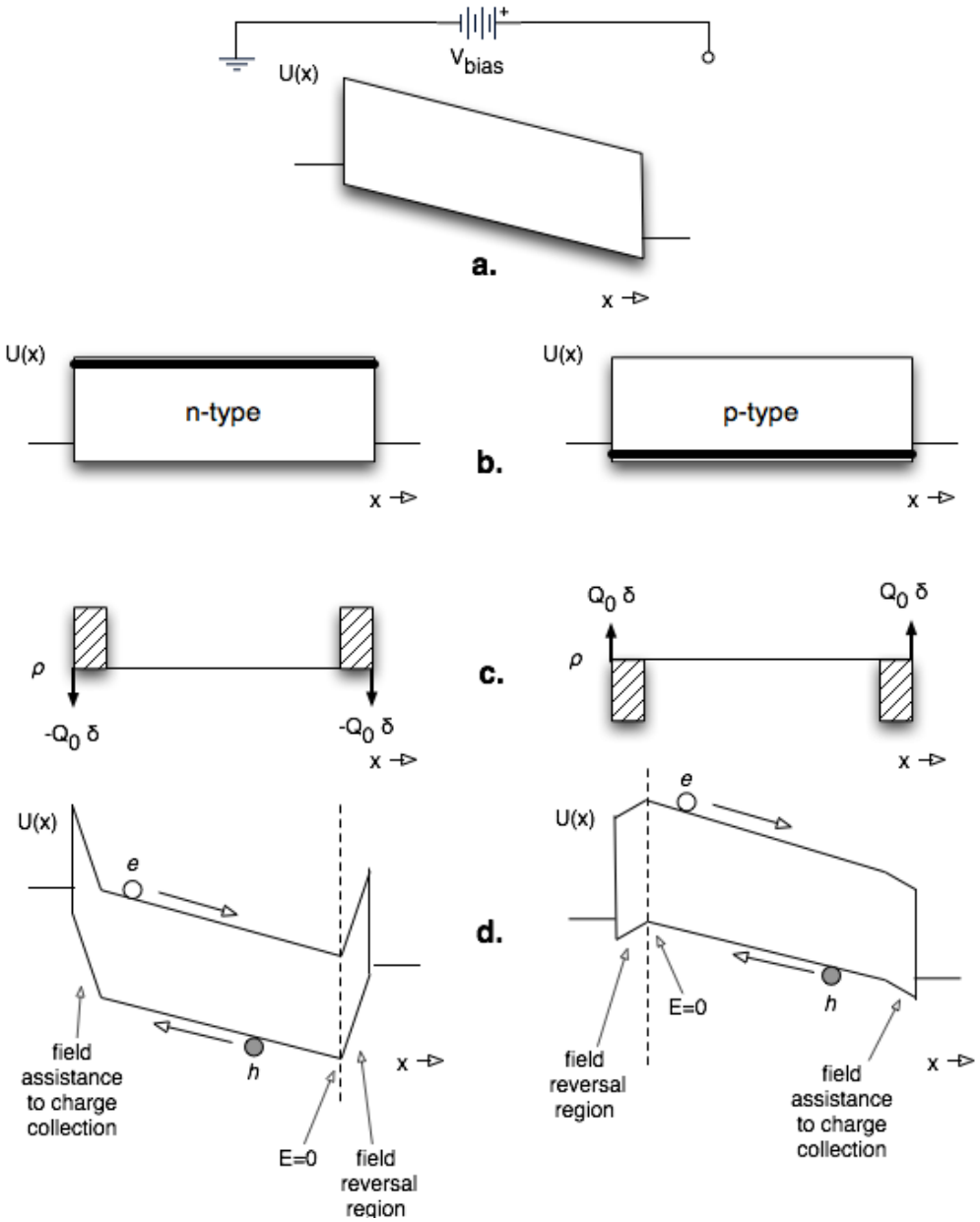

Figure G.2: a. this model suggests a simple, symmetric device as a detector with applied voltage bias, b. examples are in columns for n-type and p-type behavior, c. the charge density mapping out the space charge regions at both contacts, d. note how drifting electrons and holes interact with the resulting net field, including field-assisted and reversed-field regions.

no longer completely determine the height of the Schottky barrier. A way around this is to specify for a semiconductor a modifier to introduce an effective metal work function for the junction.

In the case of germanium, the surface state physics to a metal is so pronounced that pinning more or less takes over the energy level placement. This is the subject of a lot of condensed matter theory of interfaces. Surface states responsible for pinning are not defects or "dirty" interfaces, but are quantum mechanically allowed states that arise from the proximity to the metallic states in the adjacent metal. These are called Metal Induced Gap States (MIGS).

Metals tend to demonstrate an effectively pinned work function that is low along the germanium gap, close to the valence band edge.

I propose a benchmark model, where any metal (or generally any contact, for now) is simply given a work function pinned to 0.1 eV above the valence band.

G.4.2 Damage and hydrogen

Germanium is looking for an excuse to be p-type at surfaces, it seems. It is known from traditional nuclear physics detectors that damage in germanium detectors tends to act p-type. Some theories suggest that both hydrogen and lattice damage in germanium may allow states that exist even *below* the valence band edge. This is not something that comes up a lot. Since hydrogen and damage both have a propensity to act p-type, it suggests that trying to passivate in germanium may be a bad idea as one tends to just add acceptor-like states when hydrogenating. This is not the understanding gained in silicon, where hydrogen passivation works because it is ambivalent and will couple with bonds that need repair.

So add to a benchmark model that lattice damage (or atomically diffused hydrogen) act as acceptors, with states near or at the valence band edge.

G.4.3 A specific case of implantation damage

Here is a SRIM simulation result of ion implanting 50 keV Fe^+ into germanium with a 400 angstrom a-Si layer, figure G.3. About half of the ions make it into the germanium. Of those ions, the typical number of damage sites created was 150. So say a dose of $10^{12}\ cm^{-2}\ Fe^+$ generates $10^{14}\ cm^{-2}$ damage sites into 200 angstroms of the surface. Now this damage is present only where there is a lack of metalized electrode coverage, but it still represents a lot of the surface area of the detector face.

Iron is a deep center. It could act ambivalent. Forget it for now. But the *damage* may be notable. A concentration of damage acceptor sites at $5 \times 10^{19}\ cm^{-3}$ is enough to bend a p-type work function across $0.1\ eV$ in a distance

$$W_{damage} \approx 2\ nm. \tag{G.7}$$

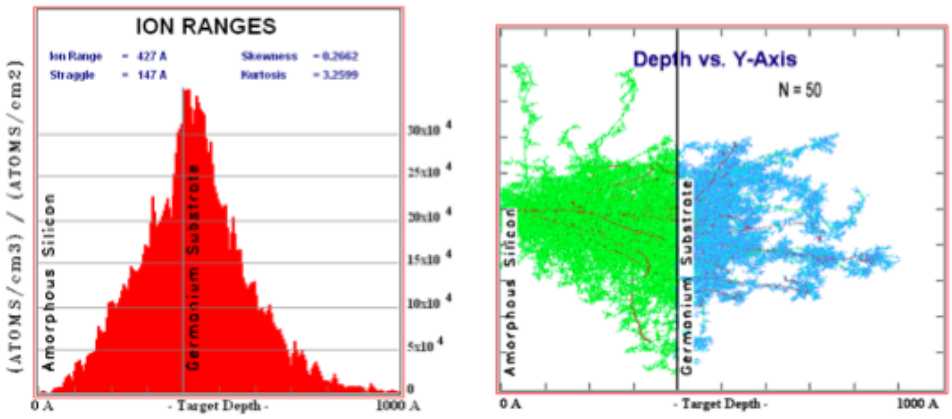

Figure G.3: SRIM ion stopping simulation of implantation damage, where there is a lack of electrode coverage.

This distance happens to be approaching atomic scales. Likely tunneling by field emission becomes probable, but *for p-type* germanium devices.

To the point, germanium damage such as this (or from other surface prep) might act as a p^+ region. So for p-type detectors, such regions could account for band-bending over shorter distances. N-type detectors would see the opposite effect. Germanium n-type devices tend to rectify and require long space-charge regions anyway, but an added p^+ region only adds potential energy to the barrier. The work function difference to overcome may be closer to 0.74 eV (the T=0 full gap) rather than 0.74 - 0.1 = 0.64 eV. Space charge regions in n-type detectors are probably wider, with damage and hydrogen effects acting as a detriment.

Moreover, n-type detectors with their wider space charge regions, generally have a lower electric field due to the space charge. This means the abrupt field model may need to be revisited. Drift-diffusion dynamics may take place over a wider, low-field region. Dead layers on the field-reversal side of n-type detectors may likely be the worst we see, with hydrogenation and damage only contributing to this.

G.5 Summarizing space charge layers

To summarize, both n-type and p-type detectors may have space charge layers near their surfaces. These regions of space charge would superimpose contributions to the electric field. These contributions would assist charge collection with higher fields near one surface, but with a conversely low (and likely even reversed) field direction at the opposite contact. This would decrease the drift length by some margin for all events, and be a detriment to events which originate within the reversed-field region. The spectra of emitted Luke phonons may also be complicated by non-uniform drift fields.

The model presented here may be a worst-case limit, in that the trace impurity concentrations of our detectors make for a very large space-charge widths. It could be that the physical treatments of the actual surfaces in reality provide higher concentrations of states with which to bend the bands. This would result in much shorter distances.

The evolution of space charge under bias is important. It may be these space charge layers "burn off," with charging events. This would be rate-dependent, however, and the regions would return when unbiased and left to naturally find their steady state. Evolution of space charge layers could be predicted, for instance, by solving drift-diffusion equations along with continuity equations (see other work).

Interesting prospects are that doping procedures should likely be able to fix (or preferentially enhance) these effects in both n and p-type crystals.

Knowing some of the peculiarities of germanium, we should be suspect of n-type crystals as having potentially larger space-charge regions.

G.6 Drift-diffusion dynamics for a dead layer model

Now we also introduce a simple dead layer model, appropriate for use with transport quantities for mobility and diffusion. We basically put a literal treatment of "diffusion" into the term we always use, "back-diffusion." We speak of back-diffusion at the contacts regarding the dead layer, so there is no reason not to posit this in a simple model with a proper diffusion constant. We will balance the stochastic nature of diffusion-dominated current, and see where that motion gives way to drift in the electric field.

G.6.1 The abrupt field approximation

We saw in previous sections that the electric field can be large compared to the external drift field when entering a space charge region. Consider the appropriate contact next to the supposed field-reversal region. For an applied field of $\sim V/cm$, the space-charge field rather sharply can invert this field to opposite polarity over a short distance. Note figure G.4. We call this "the abrupt field approximation." Therefore, under this approximation, we treat the field dependence of transport quantities of v_d and D as constant in the drift-field region, and assume that point of entering the field-reversal region serves as a point of no return.

So this will be the simple model assumed in this treatment.

G.6.2 Carrier concentration

The diffusion equation is separable, so we start with a drift-diffusion solution for either electron or hole transport, appropriate in 1D. As a function of time, we call

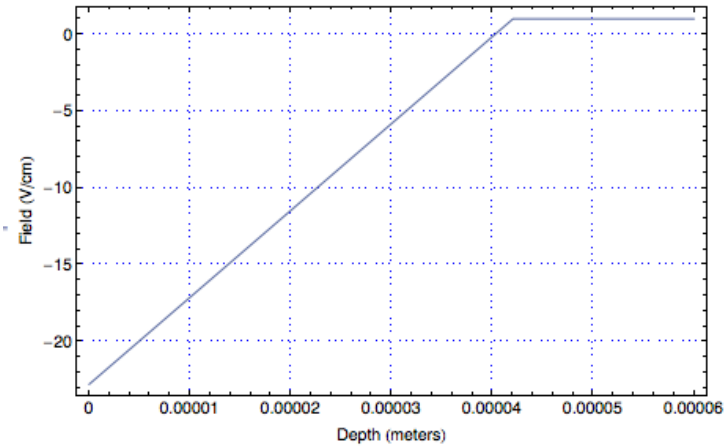

Figure G.4: Here is a hypothetical electric field (V/cm) experienced by a carrier, near the field-reversal contact. We call the "abrupt field approximation" a constant drift field that goes to some extreme space-charge field, discretely, beyond some threshold distance.

the 1D carrier concentration as,

$$N(z, z_0, t) = \frac{N_0}{\sqrt{4\pi Dt}} e^{-\frac{(v_d t + z - z0)^2}{2(2Dt)}} \tag{G.8}$$

where N_0 is the initial carrier concentration per unit depth, z. The initial event location started at z_0. The diffusion constant (implicit that it is field dependent), is given by D. Likewise, the drift velocity is given by v_d. We see that we have a simple gaussian that spreads by diffusion with time, and has a constant velocity in the mean due to the drift velocity. The intuition will be that diffusion lengths go as $l_{diff} = \sqrt{Dt}$, whereas drift velocity increments by $l_{drift} = v_d\ t$. So as a function of time, there will be transition from a diffusion-dominated step size to a drift-limited step size.

We can define drift and diffusion currents.

The electric current density due to a diffusing concentration of carriers is

$$J_{diff} = -(q)D\frac{dN}{dz}. \tag{G.9}$$

Using our expression and ignoring any time derivates of the drift velocity, we find

$$J_{diff} = qD\frac{\mathrm{N_0}(v_d t + z - \mathrm{z_0})e^{-\frac{(v_d t + z - z0)^2}{2(2Dt)}}}{2t\sqrt{4\pi Dt}} \tag{G.10}$$

The current density for a simple drift current is

$$J_{drift} = qv_d N \tag{G.11}$$

giving

$$J_{drift} = qv_d \frac{N_0}{\sqrt{4\pi Dt}} e^{-\frac{(v_d t + z - z0)^2}{2(2Dt)}}. \tag{G.12}$$

Comparing the two currents, as in the bottom of figure G.5, we find that the drift current is always traveling forward. Diffusion, on the other hand, is omnidirectional and gives a bipolar distribution. The negative-going current on this plot represents the current that actually moves against the applied electric field. It is our *back-diffusion*.

What if we ask *when* does the current due to back diffusion equal the current due to forward drift, for a given point in space? Balancing currents in this way tells us when the distribution behind the centroid quit moving backward on average, and began a drift-dominated net motion forward in the electric field. We solve for this particular time of balance, as a function of position between initial event point z_0, and the "point of no return," z'.

$$t_{bal} = \frac{z_0 - z'}{3v_d} \tag{G.13}$$

As a function of position variable z', we take the concentration at this special time of balance. For an event originating at a point z_0, t_bal represents the time for which the gaussian distribution gives a back-diffusion current balanced by a forward drift current in the field. Assume that everything farther than the point z' is lost to the wrong contact, or lost to the space-charge region first.

We can take the cumulative distribution of this concentration, and essentially map out how much of the gaussian will be swept away toward the proper contact, as a function of position.

$$\mathbb{N}_{\mathbb{CDF}}(z', z_0) = \int_{z'}^{\infty} N(z, z_0, t_{bal}) dz \tag{G.14}$$

As long as $z_0 > z'$, then plugging in for the concentration returns

$$\mathbb{N}_{\mathbb{CDF}}(z', z_0) = \frac{N_0}{2}(1 + erf\left[\sqrt{\frac{v_d}{3D}(z_0 - z')}\right]). \tag{G.15}$$

This function, normalized, gives the fraction of charge collected for a given initial position, z_0. Note that a characteristic length depends on the ratio of balancing mobility against diffusion.

$$l_{back-diffusion} \approx \frac{D}{v_d} \tag{G.16}$$

Note also that the CDF at a given value of $(z_0 - z') = \frac{D}{v_d}$ returns a numerical value of $0.794 \approx 0.80$. So a collection level of 80% may be an appropriate benchmark to specify a dead layer in the future.

We plot the CDF in figure G.6. We use the charge layer width from figure G.4 as the value of z'. 1D values of D and v_d from transport simulations have been used, at different electric fields. Note that charge collection in this ideal model begins at 50% rather than at zero. This is because drift and diffusion currents in the forward half of the gaussian concentration are already forward-going, and diffusion complements drift.

G.7 Concluding the simple dead layer model

We used a simple approximation for electric field, which we called the abrupt field approximation. Solving drift and diffusion in a simple 1D model for non-ideal electric fields is possible, but does not keep this simple analytical form. One could also think about initial non-steady-state effects or extra diffusive processes like initial plasma droplet diffusion. A way to account for initial diffusion effects would be to introduce an initial time t_0, adjusted to give diffusion a "head start" over drift, again originating at some initial position z_0.

The merit of what we have produced here is a benchmark formula, in the form of a simple error function. It relates in an ideal way the dominant dynamics of the problem, which is the field-dependent drift velocity balanced against the stochastic random walk of the diffusion current.

We found that a typical dead layer length should go as the ratio of two fundamental quantities, $\frac{D}{v_d}$.

Under the abrupt field approximation, we also suggest the possibility of a "very dead" dead layer, on order of the width of the space-charge region, $\sqrt{\frac{2\epsilon\phi_{ms}}{eN}}$. Here, we would not be sensitive to collection dynamics, unless we can evidence some portion of backward going events at the $\approx$ percent level. Is it possible we have been oblivious to this region? The charge collection dynamics we typically consider would originate from positions outside of this region. Again, this model is not claiming that these space charge regions *must* be this large, as long as there would be some high concentration of appropriate impurities near the surface to "bend the bands" over a much smaller space-charge region. Since damage acts slightly more p-type, it would make sense that p-type detectors, depending on surface treatment, may have less of these effects. Intentional doping would also serve this purpose, for either n or p-type crystals.

On one last note, electric fields in very wide space charge regions may not work well under the abrupt field approximation. If the typical space charge width goes on for 100's of um, then the intensity of the charged back-field is generally smaller. This back-field becomes on order of the applied field. This may be the case for n-type detectors, for instance. In this case, the abrupt field approximation doesn't hold and collection dynamics may depart substantially from this ideal model. Diffusion currents would be allowed to gradually dominate in a different way. Recombination

effects may even be present. It would make sense that the field-reversal regions of n-type detectors would show a response that is substantially different than this ideal model.

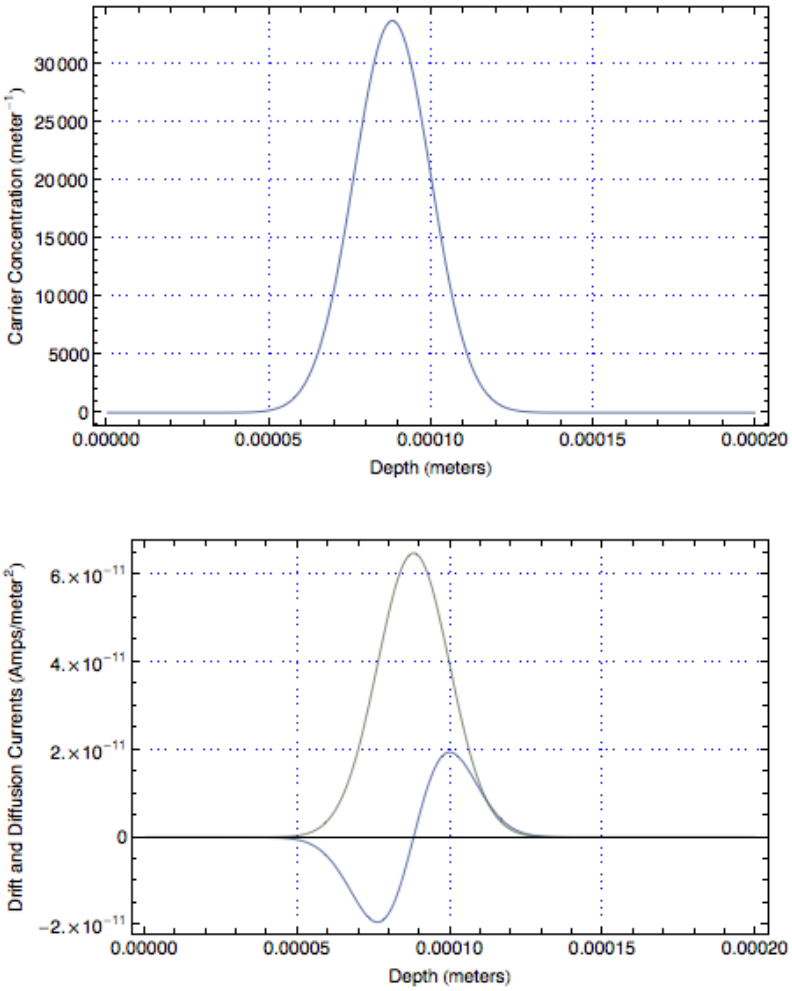

Figure G.5: top; a carrier concentration we see is a gaussian distribution. bottom; the contributions to the total current for drift and diffusion are shown at one point in time.

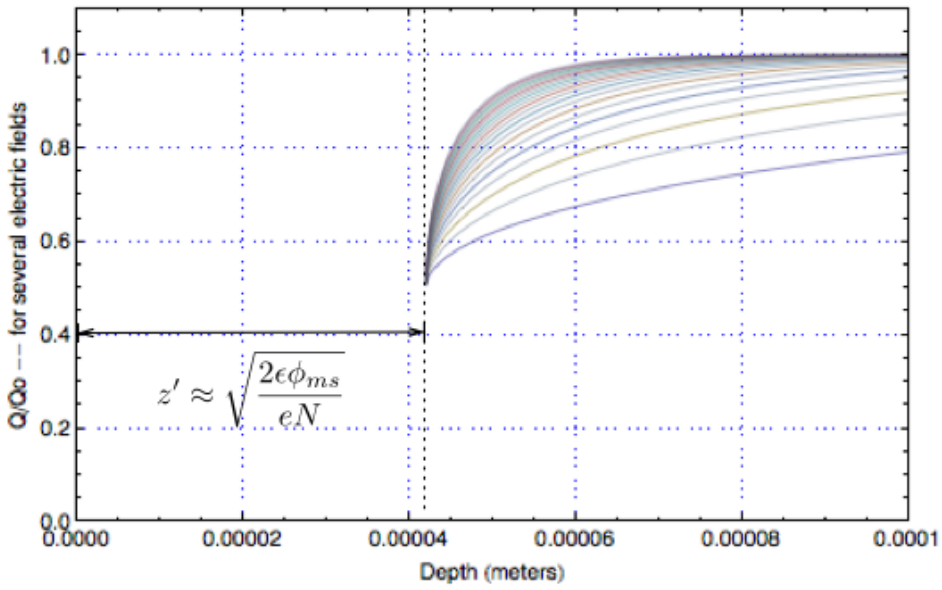

Figure G.6: The dead layer response. The fractional charge collected as a function of initial event location, z_0. The ideal response starts at a value 50%. Under the abrupt field approximation, this occurs at position z' at the edge of the space-charge region.

Appendix H

A Note About Our Existing LED Wavelength for Neutralization

This appendix simply points out a fact about the photon wavelength emitted by the LEDs we use to neutralize our detectors. It remains to be seen what role it may play.

After periods of biased use, CDMS detectors are grounded and "baked" with LED flashes of light. The principle is for photons to generate free carriers within the detector bulk while the detectors are grounded. This generates free charges to nullify possible bound space charge. We call this our "neutralization" process.

An interesting fact is that the wavelength chosen for our LEDs is actually too short. The light from our GaAs LEDs (Photonic Detectors, Inc. SMT Type PDI-E940SM) is peaked at 1.31 eV (940 nm). Values close to this were verified experimentally at cryogenic temperatures despite bandgap changes [116]. From the wavelength distribution presented in manufacturer's spec sheet, the light emitted can be well approximated by a Gaussian distribution with a deviation of 0.04 eV.

In *silicon*, 1.31 eV photons are of an energy just above the lowest gap minimum, able to excite an electron-hole pair and an optical phonon. The generated phonon is necessary, as the lowest gap in silicon is *indirect*. An indirect gap is such that the conduction and valence minima are offset in wavevector, or momentum. Note figure H.1. So for generating electron/hole excitations across the indirect gap, a phonon is a required constituent of every excitation [117]. This is what we were intending to occur in germanium, as well.

In germanium, however, this LED wavelength behaves quite differently. Here, 1.31 eV represents an energy well into the germanium's *direct gap* (Γ). A photon excitation of the direct gap occurs without subtending a change in momentum. Photons alone can excite this gap, without the assistance of a phonon to conserve momentum. This kind of absorption is a simpler interaction, and can occur much more readily [20].

A consequence of exciting the direct gap is that photons absorb over substantially smaller depths. Light is absorbed in materials as $I = I_0\ exp(-\alpha x)$. As per [20],

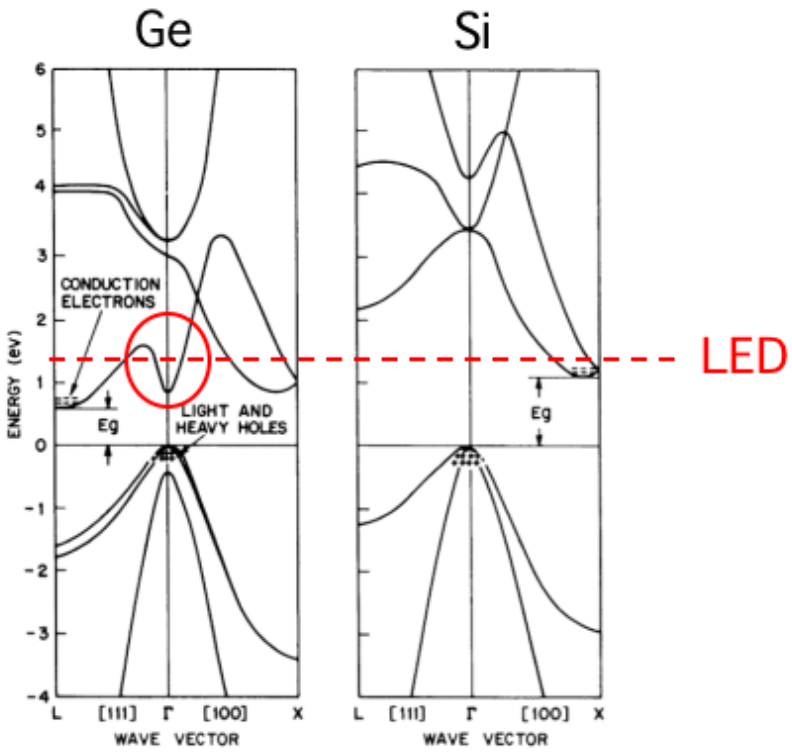

Figure H.1: A comparison of band placement, germanium relative to silicon. Our LED photons excite the indirect gap in silicon as expected. For germanium, they excite well into the *direct* gap at Γ ($k = 0$). Underlying figure taken from Sze [19].

the absorption coefficient for 1.31 eV light is 25 cm^{-1} in the indirect gap of silicon, vs. 10^4 cm^{-1} in the direct gap of germanium. Note figure H.2. This is true despite the slight ($\sim$ 10%) increase in the bandgaps near $T = 0$ K versus 300 K. The characteristic absorption length in silicon of $\sim 1/2$ mm is therefore **400** times longer than the $\sim \mu m$ length in germanium.

Incidentally, a sharp increase in absorption coefficient with increasing energy is measured at the direct gap. This is referred to as the "absorption edge [20]," and is actually how these bands were measured in the first place.

A consequence of this short-distance absorption is that electron/hole pairs are generated only quite near the surfaces of our detectors. They must diffuse deep into the crystal to neutralize bulk space charge. Conceivably, the neutralization process may be less efficient with this photon absorption profile.

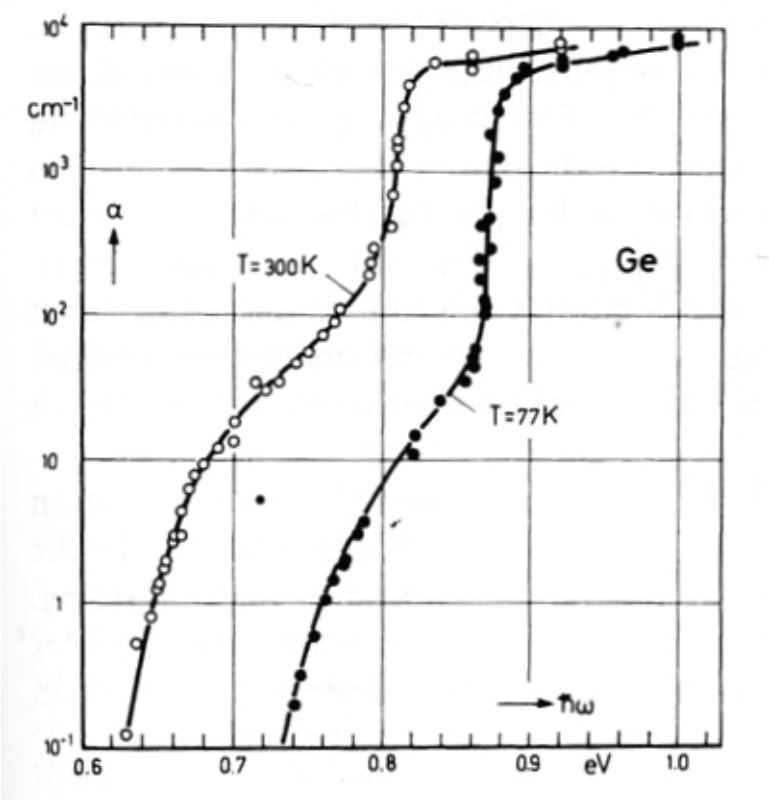

Figure H.2: The absorption coefficient for germanium, as a function of incident photon energy. The "absorption edge" occurs at the direct gap, as phonons can readily absorb by the direct process. Even for $T \sim 0$, our LED photons penetrate only at $\sim \mu m$ scales. From [20].

Appendix I

The CDMS SQUID Amplifier

In this appendix, we introduce the CDMS SQUID amplifier, exploring it using basic, linearized feedback theory. This way, it should be possible to develop a good appreciation for the signal transduction of our TES current pulses into readout signals. This is also a good way to understand the trade-offs we balance when we "tune" the SQUID arrays for optimum performance.

It was a pleasure to work with Dennis Seitz in this effort, putting into practice this knowledge in the lab. We came to appreciate the behavior of the error signal, which allowed us to develop criteria for the SQUID tuning procedure at the Soudan Laboratory. Our SQUID arrays are produced by Martin Huber at U. of Colorado, Denver. He and Bruce Hines contributed a good deal to the later material in this appendix.

For a more general review on SQUID amplifiers, see reviews such as [118, 119].

I.1 The SQUID array

The cornerstone of the open-loop gain is of course the SQUID itself. The SQUID acts as a transresistance pre-amplifier, converting current signals in the input coil into voltage signals for the next-stage amplifier. Although our circuit uses a many-junction SQUID *array*, operationally it can be considered similar to a two-junction DC SQUID. Below, we give the important details of basic SQUID parameterization. For reference, possibly the most popular explanation of a SQUID [120] was written before the acronym was adopted. Other handy references are found in [121, 122, 123].

I.1.1 An ideal two-junction SQUID

In brief, SQUIDs utilize the fact that magnetic flux in a superconducting loop is *quantized* in increments of the flux quantum, Φ_o. The SQUID is itself a superconducting loop, though spliced by Josephson junctions. As magnetic flux is passed

through the loop, the superconducting phase across the junctions must change relative to each other to maintain the flux quantization condition. This high sensitivity to magnetic flux is how the parallel Josephson junctions collectively modulate their summed supercurrent, and is the key aspect of SQUID performance.

The relationship between magnetic flux and the superconducting phase difference between the junctions can be found by solving the Josephson relations, as in Chapter 5 of [122], with a constraint on the superconducting order parameter from the Ginzburg-Landau equations. For the single best reference on this, reference [122] is recommended. What is of immediate practical concern is the functional dependance. Say we have junctions *1* and *2*. The phase for *2* is

$$\phi_2 = \phi_1 + 2n\pi - 2\pi\frac{\Phi}{\Phi_o} \tag{I.1}$$

If we assume each junction contributes $\frac{I_c}{2}sin(\phi)$ in current, then the sum of current through the SQUID is

$$I_{tot} = \frac{I_c}{2}sin\phi_1 + \frac{I_c}{2}sin\phi_2 \tag{I.2}$$

$$I_{tot} = \frac{I_c}{2}(sin\phi_1 + sin(\phi_1 - \frac{2\pi\Phi}{\Phi_o}) \tag{I.3}$$

Maximizing this total current gives us the maximum critical current through the device.

$$I_{tot_C} = I_c|cos\frac{\pi\Phi}{\Phi_o}| \tag{I.4}$$

The above equation is what is useful to us in characterizing the SQUID response. As the SQUID self-inductance is low, we ignore self-induced contributions in the flux.

I.1.2 Connection to a single Josephson junction

We seek the *I-V* curve of the ideal SQUID. To do this, it is useful to first consider a single Josephson junction.

The Stewart-McCumber treatment [122] models a single Josephson junction as components (current source, resistor, capacitor) in parallel. Solving with the Josephson relations for current and voltage, we arrive at what is sometimes known as the pendulum-analogy equation. This equation relates the total current through the junction to the time-dependent phase. Although the *instantaneous* voltage is oscillating at a frequency of order 100 GHz, the Stewart-McCumber treatment allows us to solve for a non-trivial time-averaged voltage. Ignoring the capacitance term,

$$V = 0 \text{ for } I < I_c \tag{I.5}$$

$$V = \frac{I_c}{G}\sqrt{(I/I_c)^2 - 1} \text{ for } I > I_c \tag{I.6}$$

Where G is the conductance of the junction. Notice this function extrapolates to the ohmic regime $V = I/G$ at large currents $I \gg I_c$.

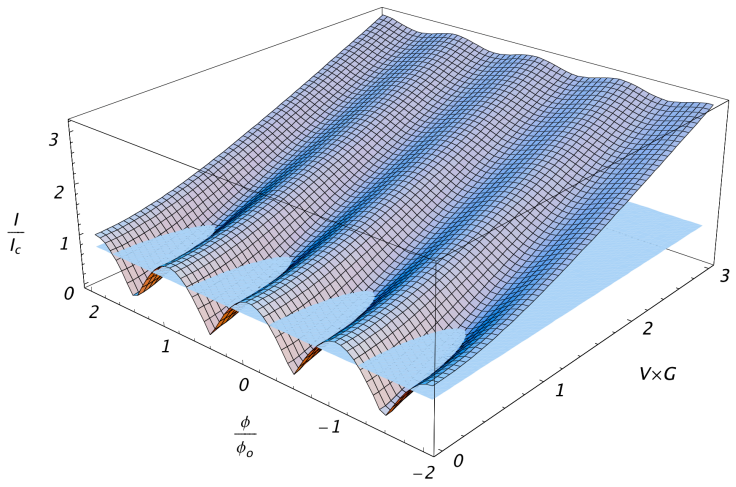

Figure I.1: The idealized relation between $I(\Phi)$, $V(\Phi)$, and Φ. Our SQUID amplifier maintains a constant current bias, represented by the constant plane. Notice the $V - \Phi$ response this creates. A bias current $I > I_c$ will not allow any $V = 0$ points.

It can be shown (as on p. 225 of [124]), that the time-averaged voltage across a SQUID is similar to that across a single junction *except with a flux-dependance*. The natural extension of Eqn. I.4 combined with Eqn. I.6 results in the proven result for the ideal SQUID I-V relation:

$$V(\Phi) = \frac{I_c}{G}\sqrt{(I/I_c)^2 - cos^2(\pi\frac{\Phi}{\Phi_o})} \tag{I.7}$$

or equivalently,

$$I(\Phi) = \sqrt{V^2G^2 + I_c^2 cos^2(\pi\frac{\Phi}{\Phi_o})} \tag{I.8}$$

We can see from the above relations that the utility of a SQUID comes from its flux sensitivity, present in both V or I. In utilization of SQUIDs in circuit, a constant I (V) must bias the SQUID while a transient response in phi is readout as a change in V (I). Either choice is valid. Most often, biasing a SQUID with constant *current*

and measuring a transient *voltage* is useful in coupling to high-impedance amplifiers. This is the situation in our amplifier.

I.1.3 SQUID equivalent circuit and internal impedance

We have established the voltage relation for our current-biased SQUID. To develop this further as a circuit model, let's now consider what internal impedance the SQUID line should have. Shunt resistors bridge the Josephson junctions on every SQUID in the 100-SQUID array. Their purpose is to reduce hysteretic effects in the I-V curve not discussed here. These additional shunt resistors supply greater conductance, so less voltage is induced across each SQUID element. Yet in a series array of 100, an adequately large voltage can be established to amplify an input signal above the voltage noise inherent in the SQUID line. Each SQUID element receives approximately the same flux, so the array acts coherently as an ensemble to produce the output voltage. Although the array flux-dependance may not precisely resemble $cos^2(\pi\frac{\Phi}{\Phi_o})$, it is clearly has a flux-period of Φ_o. This is how we may treat the array as an equivalent single-SQUID. Two resistors in parallel in each element, with 100 elements in series, gives a conductance in this case $G \approx (50\ \Omega)^{-1}$. The SQUID coils themselves produce a net self-inductance that is perhaps $\sim 40 - 45\ pH$.

I.1.4 Simplification for the small-signal model: SQUID responsivity

We think in terms now of a model representing small-signal deviations from quiescent levels. Our SQUID will be in a closed-loop circuit, which means transient SQUID voltages (should) never be far from the equilibrium "lock point."

It is intuitive to think of the SQUID as a device which drives a (change in) current through its own impedance as a function of applied flux. The equivalent voltage model can be represented as a source with the impedance now in series. We should point out that the stage following the SQUID (the preamp) in our amplifier has an enormous input impedance compared to the SQUID itself. For most purposes, the SQUID line impedance (G_s and L_s) can actually be disregarded.

Furthermore, we simplify characterization by introducing the *responsivity*, Z_{SQ}. Responsivity is a *transresistance*. It is practical to measure, and denotes the voltage change across the SQUID for a given change in coil current: $(\frac{dv_{SQ}}{di_{coil}})$. The AC input coil current, i_{in}, passes through an input coil which couples magnetic flux to the SQUID via a mutual inductance M_{in}. All we require from the $V - \Phi$ curve (the empirical equivalent of Eqn. I.7) is the *differential* voltage response, $dv_{SQ}/d\phi$, calculated at the lock point.

When considering Z_{SQ} values, one must take care to distinguish whether the quoted value is referred to the *input* coil or the *feedback* coil. For the remainder of this analysis, Z_{SQ} will always be referenced to the *input coil*.

$$Z_{SQ} = \frac{dv_{SQ}}{di_{in}} = (\frac{dv_{SQ}}{d\phi})(\frac{d\phi}{di_{in}}) = \eta\ M_{in} \tag{I.9}$$

where

$$\eta = \frac{dv_{SQ}}{d\phi} \tag{I.10}$$

$$M_{in} = \frac{d\phi}{di_{in}} \tag{I.11}$$

Responsivity is chosen by "tuning" the equilibrium voltage (lock point) of the circuit to the appropriate point on the $V - \Phi$ curve.

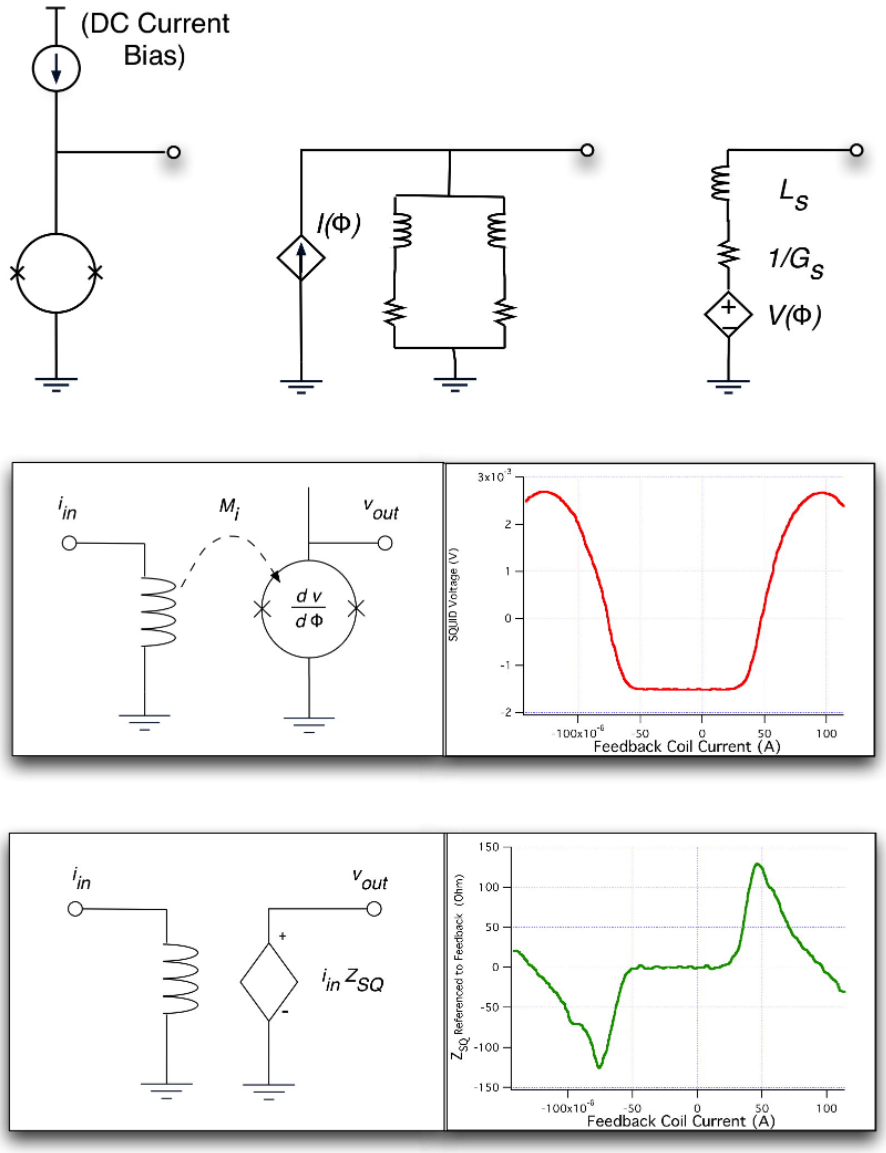

Figure I.2: The top row demonstrates the reduction of the SQUID into a simple (controlled) voltage source and series impedance. This impedance is small and can be neglected most of the time. The middle row shows the actual SQUID voltage relative to DC input current. The derivative of this relation is shown in the bottom row, which is used to deduce the value Z_{SQ} for a given current bias. In this small-signal model, all we need to consider is that the SQUID acts as a voltage source $v_{SQ} = i_{in} Z_{SQ}$. Be it known that, for ease of measurement, these plots were actually *measured* relative to the *feedback* coil.

I.2 Transfer function and bandwidth

I.2.1 Closed-loop amplifier with simple feedback

In this treatment, a traditional closed-loop transfer function for a linear feedback amplifier is used. In this introduction, a quick summary of the method is explained for review.

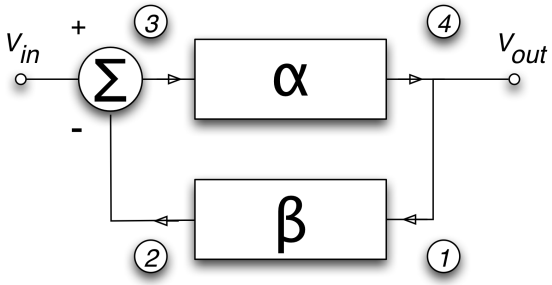

Figure I.3: A signal-flow diagram of an amplifier, demonstrating feedback.

Consider Figure I.3. In this example amplifier, we have an input voltage and an output voltage. Let's start first at the *output* voltage. Points *1* and *4* are common to the output, and must share the same value. Starting at point *1*, the output voltage signal is passed through the feedback network, passing the value βV_{out} to point *2*. At Σ, the summing node, point *2* is *subtracted* from the input voltage (since we have stable, *negative* feedback). As we follow along to point *3*, the voltage must therefore be equal to $V_{in} - \beta V_{out}$. This voltage is now amplified through the α, the open-loop gain network. Arriving back at the output, point *4* must be

$$V_{out} = \alpha(V_{in} - \beta V_{out}) \tag{I.12}$$

We use this to solve for the *transfer function*, which is the ratio of output to input,

$$H = \frac{V_{out}}{V_{in}} = \frac{\alpha}{1 + \alpha\beta} \qquad \text{(general example)} \tag{I.13}$$

The above equation demonstrates the basic relationship between closed- and open-loop gain. Again, α is the open-loop gain and would be the transfer function of the amplifier if the feedback loop were disconnected/opened. Eqn. I.13 is the *closed-loop gain* transfer function where we see the effect of the feedback network, β. For the most part, the open-loop gain of an amplifier is made to be far greater than 1. This effectively cancels α out of Eqn. I.13, which then becomes $\frac{1}{\beta}$.

The transfer function and its constituent parts will usually have a frequency dependance. Frequency dependancies will be dealt with in Laplace formalism, where we

consider the complex angular frequency $s = \sigma + j\omega$. If you are unaccustomed to this, be assured that it is not too cumbersome. Angular frequency ω is the same frequency variable one would see in Fourier formalism. In fact, Laplace formalism is simply the natural extension of Fourier formalism in that frequency has been generalized to have both these real and imaginary components. The purpose of this is to account for both the oscillatory response of the system (solutions $\propto \ e^{j\omega t}$) as well as growing/dying transient responses (solutions $\propto \ e^{\pm \ \sigma t}$). In a nutshell, using Laplace formalism is advantageous because it retains both oscillatory and transient information of the system response.

In the SQUID circuit we are about to investigate, most of the work will stem from finding the closed-loop gain similar to Eqn. I.13. A cosmetic detail is that we are concerned with a SQUID's output voltage relative to its input *current*. The SQUID transfer function will therefore describe an overall impedance (a *transimpedance*) between input and output.

I.2.2 SQUID transfer function and bandwidth

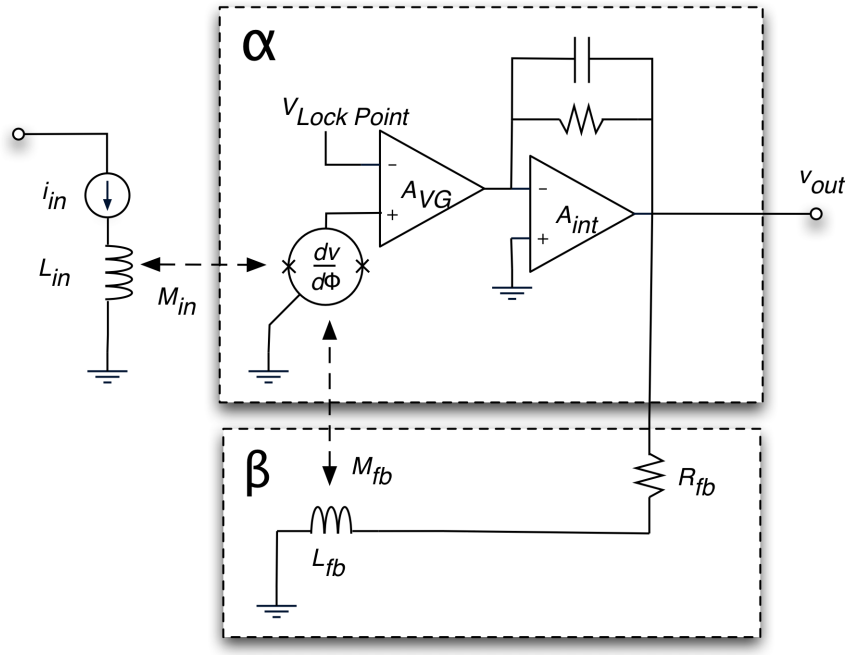

Figure I.4: A signal-flow schematic of the SQUID amplifier.

The SQUID amplifier can be decomposed into amplifier subsets α and β as per the above signal-flow model. If we consider the geometrical layout, the SQUID itself serves as the summing node as it couples the input line and feedback line. However,

the SQUID voltage response ($dv_{SQ}/d\phi$) is operationally part of the open-loop gain, so the SQUID itself is best represented as a component within α. See Figure I.4.

As we begin to identify parameters of the amplifier circuit, please note that numerical values of the parameters are also summarized in a table at the end of this note.

It is helpful to describe the amplifier chain in sequence. The circuit remains flux-locked, so Z_{SQ} is valid to describe the coupling of input current to the SQUID. The resulting voltage output of the SQUID is amplified by subsequent amplifiers within the open-loop subsystem α. The amplifier chain is functionally idealized as two amplifiers: the *variable gain* amplifier, and the *integrator*. The variable gain A_{VG} is adjustable to values between 20 and 100. The integrator has a DC gain component A_{int}. The important role of the integrator is to provide the open-loop frequency roll-off, set by an RC pole at angular frequency ω_{int}. The total gain contribution of the integrator is therefore $A_{int}\frac{\omega_{int}}{\omega_{int}+s}$. The output signal v_{out} is a voltage built up across the feedback impedance $R_{fb} + sL_{fb}$. The inductance L_{fb} is around $4.5 \times 10^{-9}H$, and can be ignored. The current in this feedback line is magnetically coupled back to the SQUID via a mutual inductance M_{fb}.

Putting these terms together, we find

$$\alpha(s) = \frac{dv_{SQ}}{d\phi} A_{VG} A_{int} \frac{\omega_{int}}{\omega_{int} + s} \tag{I.14}$$

$$\beta(s) = \beta = \frac{M_{fb}}{R_{fb}} \tag{I.15}$$

Representing α in terms of Z_{SQ},

$$\alpha(s) = \frac{Z_{SQ}}{M_{in}} A_{VG} A_{int} \frac{\omega_{int}}{\omega_{int} + s} \tag{I.16}$$

Furthermore, ω_{int} is a small frequency in relation to the frequency intercepts we will generally be concerned with in this circuit. It is an adequate approximation to adjust the denominator in $\alpha(s)$:

$$\alpha(s) \approx \frac{Z_{SQ}}{M_{in}} A_{VG} A_{int} \frac{\omega_{int}}{s} \tag{I.17}$$

or finally, define the DC open-loop gain

$$\alpha_o = \frac{Z_{SQ}}{M_{in}} A_{VG} A_{int} \tag{I.18}$$

so we can use

$$\alpha(s) = \alpha_o \frac{\omega_{int}}{s} \tag{I.19}$$

With α and β settled, we now consider the closed-loop transfer function. The mutual inductance of the input coil comes *before* the summing node, so we must amend the closed loop gain to read

$$H(s) = M_{in}\frac{\alpha(s)}{1+\alpha(s)\beta} \tag{I.20}$$

We now have

$$\begin{aligned} H(s) &= \frac{Z_{SQ}A_{VG}A_{int}\frac{\omega_{int}}{s}}{1+(\frac{Z_{SQ}}{M_{in}}A_{VG}A_{int}\frac{\omega_{int}}{s})(\frac{M_{fb}}{R_{fb}+sL_{fb}})} \\ &= \frac{Z_{SQ}A_{VG}A_{int}\frac{\omega_{int}}{s}}{1+\frac{M_{fb}}{M_{in}}\frac{Z_{SQ}A_{VG}A_{int}\frac{\omega_{int}}{s}}{R_{fb}+sL_{fb}}} \\ &= \frac{Z_{SQ}A_{VG}A_{int}\omega_{int}}{s+\frac{M_{fb}}{M_{in}}\frac{Z_{SQ}A_{VG}A_{int}\omega_{int}}{R_{fb}+sL_{fb}}} \end{aligned} \tag{I.21}$$

From here on out, it also simplifies the situation to set $L_{fb} \equiv 0$. This is justified as $L_{fb} \sim 25\ nH$, and the time constant associated with the feedback network, L_{fb}/R_{fb}, is small ($\sim 25\ ps$).

Rearranging to show the single-pole nature of this function,

$$H(s) = R_{fb}\frac{M_{in}}{M_{fb}}\Big(\frac{1}{1+s\times\frac{M_{in}R_{fb}}{M_{fb}Z_{SQ}A_{VG}A_{int}\omega_{int}}}\Big) \tag{I.22}$$

Eqn. I.22 is the transfer function, which we will utilize periodically. Already, consider the following observations:

- the output voltage signal is developed across R_{fb}, with a factor $\frac{M_{in}}{M_{fb}}$ more current than the input current.
- the term in parentheses is $\sim$ unity for modest values of s.
- the term in parentheses acts as a single-pole filter, passing frequency components under a particular cut-off frequency.

I.2.3 Bandwidth

From Eqn. I.22, we can characterize amplifier bandwidth by the -3dB point. Finding the f_{3dB} frequency is a matter of finding where

$$|H(j2\pi f)| = \frac{|H(0)|}{\sqrt{2}} \tag{I.23}$$

The parentheses term from Eqn. I.22 has the only frequency dependance. By inspection, it has the functional form of a single-pole low-pass filter.

$$\Big(\frac{1}{1+s\times\frac{M_{in}R_{fb}}{M_{fb}Z_{SQ}A_{VG}A_{int}\omega_{int}}}\Big) \tag{I.24}$$

The pole frequency is synonymous with the -3dB point in a single-pole circuit, so we can expect a -3dB frequency where $\alpha(s)$ meets the overall DC gain, $\approx 1/\beta$. Notice this is internal to the feedback loop, so we neglect M_{in}.

$$|\alpha_o\frac{\omega_{int}}{j\omega_{3dB}}| = 1/\beta \tag{I.25}$$

$$\omega_{3dB} = \alpha_o\beta\ \omega_{int} \tag{I.26}$$

So then

$$f_{3dB} = \frac{\omega_{3dB}}{2\pi} = \frac{\omega_{int}}{2\pi}(\frac{\frac{M_{fb}}{M_{in}}Z_{SQ}A_{VG}A_{int}}{R_{fb}}) \tag{I.27}$$

I.2.4 Final form of the transfer function & gain-bandwidth product

Having defined the -3dB frequency, note that possibly the most elegant way to express the transfer function, Eqn. I.22, is,

$$H(s) = \frac{H_o\omega_{3dB}}{\omega_{3dB}+s} \tag{I.28}$$

where

$$H_o = M_{in}/\beta = R_{fb}\frac{M_{in}}{M_{fb}} \tag{I.29}$$

Notice we have the equality

$$H_o\omega_{3dB} = \frac{M_{in}}{\beta}\alpha_o\beta\omega_{int} = M_{in}\alpha_o\omega_{int} \tag{I.30}$$

We find that the *gain-bandwidth product* is a conserved quantity defined solely by the open-loop gain,

$$GBP = \frac{\alpha_o}{R_{fb}}\omega_{int} = \frac{H_o}{R_{fb}M_{in}}\omega_{3dB} \tag{I.31}$$

Above, we must divide by a factor of R_{fb} as the transfer function as defined considers *output voltage* relative to *input current*. This way, we have a *current-current* relation. We also drop the prefactor M_{in} because we only consider how the closed loop behaves and the mutual inductance comes before the summing node.

I.2.5 Numerical values for transfer function and bandwidth

In this section, we simply put numerical values to the equations in this section. Numerical for individual parameters are tabulated at the end of this document. Here is the transfer function, with $Z_{SQ} = 500\ \Omega$, plotted for several values of A_{VG}.

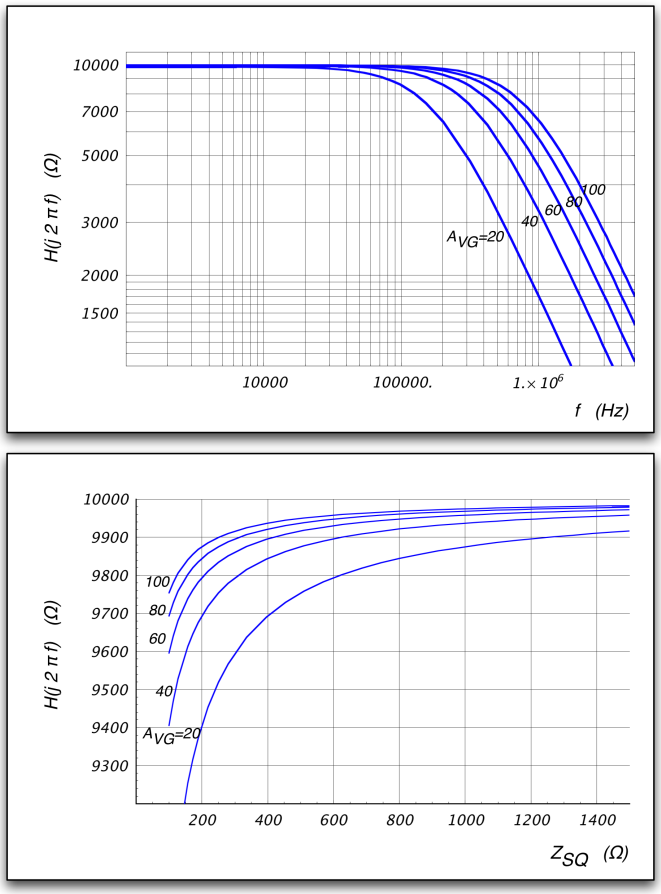

Figure I.5: The transfer function, for $A_{VG} = 20, 40, 60, 80, 100$

I.2.6 Introduction to impulse response

We now examine how the output signal performs, given a known input current. In this section, we determine the time-domain output from the transfer function.

A small detour: convolutions in the time-domain

This small subsection is to demonstrate that the Laplace/Transfer-Function method is more convenient than a direct convolution. This topic generally falls under the category "Signals and Systems," which is the title of many texts and undergraduate

engineering courses. This may be too elementary for the reader. If so, feel free to skip ahead.

If we wanted to know the time-domain response of the output given an input current as a function of time, the direct way would be to perform a *convolution* of the current and impulse response of the system...

$$v_{out}(t) = i_{in}(t) \otimes h(t) = \int_0^t i_{in}(t)h(t-\tau)d\tau \tag{I.32}$$

In general, this is awkward to perform.

Notice this *impulse response*, $h(t)$, is an interesting function. It represents the time-dependent effect of the amplifier. It is sometimes called *the instrument response*. We would like our amplifier to preserve as much information of the signal as possible, so we would like to understand this $h(t)$ function well.

If we take the Inverse Laplace Transform of the transfer function, we automatically find the *impulse response*, $h(t)$. The impulse response represents how the system *reacts to a delta-function stimulus*.

$$H(s) = \mathfrak{L}[h(t)] \tag{I.33}$$

$$h(t) = \mathfrak{L}^{-1}[H(s)] \tag{I.34}$$

Where the Laplace Transform is defined as in

$$H(s) = \int_0^\infty e^{-st}h(t)dt \tag{I.35}$$

What is nice about having transfer function already in hand, is that we do not need to perform the convolution as presented in equation I.32.

Provided we have knowledge of $i_{in}(t)$, we can take its transform

$$I_{in}(s) = \mathfrak{L}[i_{in}(t)] \tag{I.36}$$

Now a simple product in the s-domain:

$$V_{out}(s) = H(s)\ I_{in}(s) \tag{I.37}$$

The output is now represented in time-domain as

$$v_{out}(t) = \mathfrak{L}^{-1}[V_{out}(s)] = \mathfrak{L}^{-1}[I_{in}(s)\ H(s)] \tag{I.38}$$

These past few steps may seem like a circular way to arrive at $v_{out}(t)$, but in actuality it is generally less work than performing the direct convolution of Eqn. I.32.

I.2.7 SQUID impulse response and step function

So again consider Eqn. I.22:

$$H(s) = \frac{H_o}{1 + \frac{s+\omega_{int}}{s_0}} \tag{I.39}$$

$$H_o = R_{fb}\frac{M_{in}}{M_{fb}} \tag{I.40}$$

$$s_0 = \frac{M_{fb} Z_{SQ} A_{VG} A_{int} \omega_{int}}{M_{in} R_{fb}} \tag{I.41}$$

Eqn. I.39 gives us a clear-cut function to transform. It resembles the form of a single-pole filter with a gain factor. From it, we get the impulse response

$$h(t) = H_o s_0 \; e^{-(s_0+\omega_{int})t} \tag{I.42}$$

$$= Z_{SQ} A_{VG} A_{int} \omega_{int} \; e^{-2\pi f_{3dB} t} \tag{I.43}$$

So what we should notice in Eqn. I.43 is that $f_{3dB} = \frac{s_0+\omega_{int}}{2\pi}$ is in the exponential: it is the characteristic *falltime* of the feedback signal that the SQUID amplifier supplies.

While we are at it, let's do the response to an input step with some amplitude. The Laplace transform of a unit step is $1/s$.

$$\mathfrak{L}[I_o \; u(t)] = \frac{I_o}{s} \tag{I.44}$$

in which case

$$V_{out}(s) = H(s)\frac{I_o}{s} \tag{I.45}$$

so

$$v_{out}(t) = H_o I_o s_o \frac{(1 - e^{-t(s_o+\omega_{int})})}{s_o + \omega_{int}} \tag{I.46}$$

So, just as with a simple filter, we have a finite risetime due to the lack of response at higher frequencies. It is worthwhile to point out that a current, voltage, or any test signal has extra units of Hz^{-1} while in Laplace space.

I.3 The effect of finite bandwidth on risetime

This section is largely an aside from the particulars of the SQUID amplifier. Here, we generalize risetime effects for a simple amplifier. The risetime of an input signal is effected whenever it is passed through a circuit with any sort of frequency response. Amplifiers most often act as a linear, time-invariant filter of some kind. Here we treat the effect of a stable, 1st-order low-pass amplifier on an input signal.

I.3.1 Simple input

Say we have a generic input function with risetime τ_o. Give it a simple exponential dependance:

$$input(t) = 1 - e^{-t/\tau_o} \tag{I.47}$$

This has the Laplace transform

$$Input(s) = \frac{1}{s} - \frac{1}{s + \frac{1}{\tau_o}} \tag{I.48}$$

I.3.2 A simple transfer function

The effect of the finite-bandwidth amplifier is represented as a function acting as a filter with some amount of gain. This "transfer function" is sometimes also called the "instrumentation function." We give this amplifier unity gain, and define it directly in Laplace formalism:

$$Amplifier(s) = \frac{\frac{1}{\tau_B}}{s + \frac{1}{\tau_B}} \tag{I.49}$$

In this section, τ_B is the time constant of the amplifier's bandwidth.

I.3.3 Output

We find the output to be

$$Output(s) = Input(s) \times Amplifier(s) \tag{I.50}$$

Taking the inverse Laplace transform,

$$output(t) = \frac{\tau_o(1 - e^{-t/\tau o}) - \tau_B(1 - e^{-t/\tau_B})}{\tau_o - \tau_B} \tag{I.51}$$

We see that the time-domain output is truly a convolution of time-constants. Assume, though, that the effect of the amplifier is subtle. We make the approximation that the output has the same functional dependance as the input signal, only with a new effective time-constant slightly modified by the filtering of the amplifier.

$$output(t) = 1 - e^{-t/\tau_{out}} \tag{I.52}$$

We now equate the output expressions to determine the effective time constant τ_{out}.

$$\frac{\tau_o(1 - e^{-t/\tau o}) - \tau_B(1 - e^{-t/\tau_B})}{\tau_o - \tau_B} = 1 - e^{-t/\tau_{out}} \tag{I.53}$$

We want to solve this expression for t at $t = \tau_{out}$, but it is cumbersome. We make a reasonable approximation; the bandwidth time constant is much shorter than τ_{out}. We say

$$e^{-\tau_{out}/\tau_B} \approx 0 \tag{I.54}$$

So

$$\frac{\tau_o(1 - e^{-\tau_{out}/\tau_O}) - \tau_B}{\tau_o - \tau_B} \approx 1 - e^{-1} \tag{I.55}$$

Therefore,

$$\tau_{out} \quad = \quad \tau_o \; ln[\frac{\tau_o}{(\tau_o - \tau_B)(e^{-1} - 1 + \frac{\tau_o}{\tau_o - \tau_B} - \frac{\tau_B}{\tau_o - \tau_B})}] \tag{I.56}$$

which reduces to

$$\tau_{out} = \tau_0 \; log[(\frac{\tau_o}{\tau_o - \tau_B})e] \tag{I.57}$$

This is a good approximation to the actual output risetime around the 63% mark.

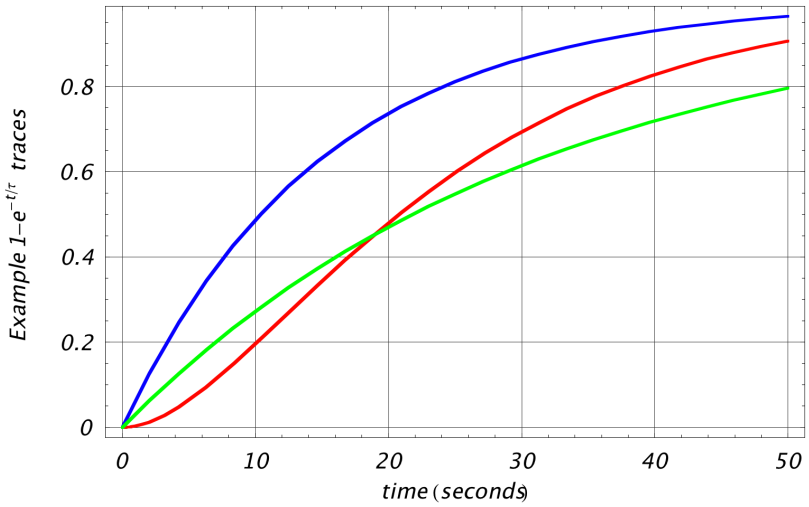

Figure I.6: Here we demonstrate a convolution of a simple risestime with finite bandwidth of $\tau_B = \tau_0/5$. Blue is the original pulse. Red is the convolved pulse. Green is the single-time constant approximation from Eqn. I.57.

We have additional filtering in the phonon-signal chain for the purpose of anti-aliasing. If we *approximate* the effect of the anti-aliasing filter as *another* simple pole, we have now *two* limiting bandwidths. Call this new bandwidth τ_{B2}.

Making similar assumptions $e^{-t/\tau_B} = e^{-t/\tau_{B2}} \approx 0$, we get the approximate formula *for two bandwidths*,

$$\tau_{out} = \tau_0 \; log[(\frac{\tau_o^2}{(\tau_o - \tau_B)(\tau_o - \tau_{B2})})e] \tag{I.58}$$

I.3.4 Application to SQUID amp: numerical values for timing

We have found empirically that TES physics occurs with perhaps about a $\sim$ 100 kHz bandwidth. This is primarily dependent on the L_{in}/R_{TES} timing quiescent in the input coil.

Anti-aliasing filters are also in the read-out electronics and give a bandwidth of 500 kHz.

Assuming single poles for each, we can plot Eqn. I.58 as a function of SQUID bandwidth frequency.

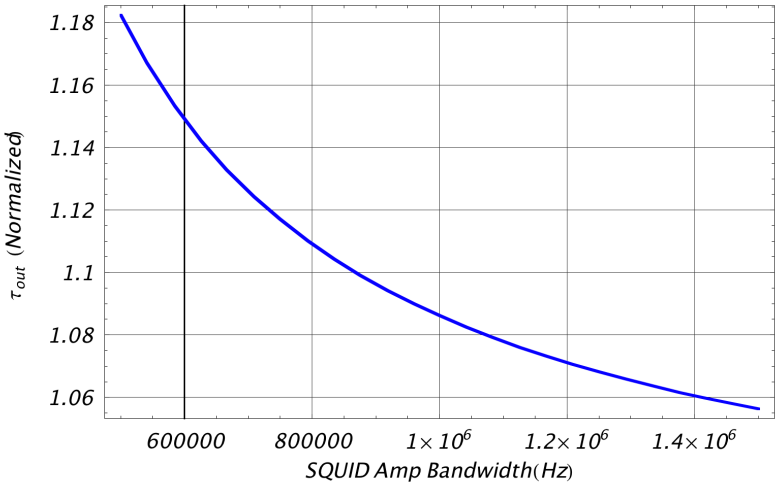

Figure I.7: This depicts effective output risetime for an input pulse of bandwidth 100 kHz, and single-pole filtering already in place at 500 kHz. It is normalized to the τ_{out} with infinite SQUID-amp bandwidth.

We see from Fig. I.3.4 that at $1MHz$, SQUID Amplifier Bandwidth will contribute less than 10% additional error in the risetime.

From Eqn. I.27 we find the constraint for bandwidth to be $f_{3dB} > \; 1 \; MHz$,

$$Z_{SQ} \times A_{VG} \geq \frac{(2\pi(1 \; MHz) - \omega_{int})R_{fb}M_{in}}{M_{fb}A_{int}\omega_{int}} \tag{I.59}$$

which is to say

$$Z_{SQ} \times A_{VG} \geq 57.5\ k\Omega \tag{I.60}$$

In practice, we find we have a resonance in the SQUID amp around $\sim 700\ kHz$. It has been seen at Soudan and in multiple test setups. Tuning the SQUIDs above this level only makes the poles interact and become more resonant. Empirically, we have set a lower threshold for required bandwidth. Using a ballpark $\sim 600\ kHz$, we set a lesser constraint,

$$Z_{SQ} \times A_{VG} \geq 35\ k\Omega \tag{I.61}$$

I.4 Noise

Noise in the SQUID amplifier is modeled as three uncorrelated sources, as figure I.8.

I.4.1 Defining noise sources

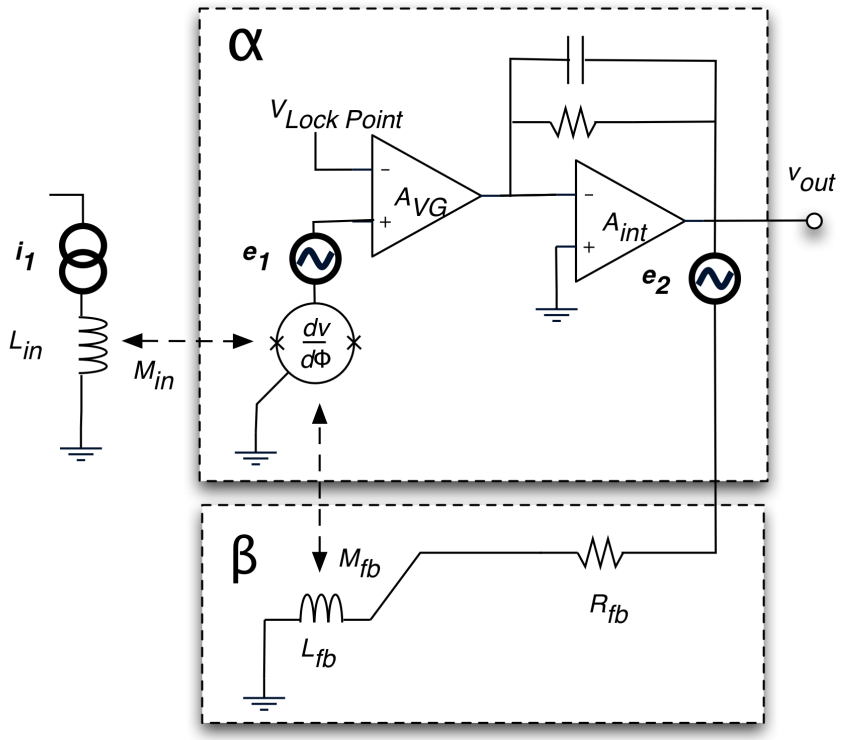

Figure I.8: Diagraming sources of noise.

1. The input current noise, i_1, represents white noise currents in the input coil. These terms account for Johnson noise terms from input resistances and excess QET noise. This also accounts for *flux-induced* SQUID noise contributions, which are easiest to model by referring them back to the input coil.

2. The noise at the SQUID, e_1, models the dominant flux-independent voltage noise. SQUID literature elaborates on many possible sources of noise across the SQUID component itself. However, e_1 also takes into account the contribution from the warm electronics. Referenced to the preamp input, the warm electronics noise acts primarily as a *voltage* noise at the SQUID as well. Therefore, it is easy to believe how this combined voltage term dominates at this point in the circuit, and is all that need be considered.

3. The feedback resistor is rather large, and its Johnson noise is addressed as a voltage noise source, e_2, at the amplifier output.

Uncorrelated noise adds in quadrature. From where noise enters in the amplifier chain, we can use superposition to construct proper transfer functions for each term. Consider the combined contributions of noise at the output.

$$V_n(s)^2 = H(s)^2 \times i_1^2 + \frac{H(s)^2}{Z_{SQ}^2} \times e_1^2 + \frac{H(s)^2}{M_{in}^2\ \alpha(s)^2} \times e_2^2 \tag{I.62}$$

or, as a function of frequency, a signal-to-noise ratio

$$SNR(s)^2 = \frac{V_{out}(s)^2}{V_n(s)^2}\ \ (units\ of\ \frac{1}{Hz}) \tag{I.63}$$

$$= \frac{H(s)^2 I_{in}(s)^2}{V_n(s)^2} \tag{I.64}$$

$$= \frac{I_{in}(s)^2}{i_1^2 + \frac{e_1^2}{Z_{SQ}^2} + \frac{e_2^2}{M_{in}^2 \alpha(s)^2}} \tag{I.65}$$

I.4.2 Numerical evaluation of noise

$$i_1 = 4\ \frac{pA}{\sqrt{Hz}} \tag{I.66}$$

$$e_1 = 1\ \frac{nV}{\sqrt{Hz}} \tag{I.67}$$

$$e_2 = \sqrt{4k(4.2\ K)1\ k\Omega} = 0.5\ \frac{nV}{\sqrt{Hz}} \tag{I.68}$$

As it should be, we find that the input current noise dominates under most conditions.

In terms of noise performance, variable gain is not an issue. Variable gain changes only where the roll-off frequency occurs and effects the signal-to-noise ratio only outside the signal bandwidth (*above* the roll-off frequency).

The feedback resistor noise e_2 is negligible. It only turns on at high frequencies, where even then it is magnitudes smaller than the other terms. In terms of operating the circuit, the main consideration is to make sure that the SQUID has enough responsivity to amplify the input signal well above the noise voltage e_1. Output noise falls as a function of SQUID transimpedance. We see that for a value of $Z_{SQ} = 500\ \Omega$, the signal-to-noise ratio is less than 10% from its asymptotic maximum.

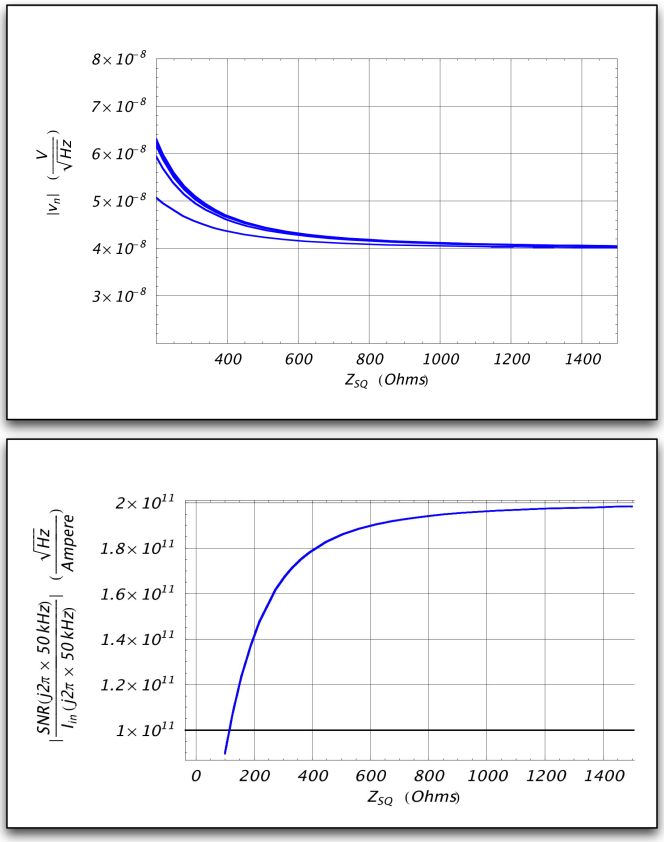

Figure I.9: Noise plots referenced at 50 kHz and plotted vs. Z_{SQ}. Top: absolute magnitude of total noise at output. Bottom: signal-to-noise ratio plotted normalized to $I(s)$.

I.5 SQUID Jumps, and slew rate

The periodic nature of the SQUID $V-\Phi$ curve means there are stable equilibrium points every flux increment of Φ_o. When the circuit is first closed, the amplifier chain ramps from zero until it reaches equilibrium on slope of the first Φ_o cycle. From time to time, events may make it unstable such that it finds a new working point $\Delta\Phi = n\Phi_o$ away.

The extra offset voltage becomes

$$\begin{aligned} \Delta V_{fb} &= \Delta I_{fb} R_{fb} = (n\Phi_o/M_{fb}) R_{fb} && \text{(I.69)} \\ &= n\ 0.23\ Volt && \text{(I.70)} \end{aligned}$$

where n is an integer.

The problem associated with SQUIDs operating at larger feedback offsets is that the overall closed-loop gain can vary by as much as 10%. Possible reasons for this may be changes in Z_{SQ} from offset to offset as the SQUID array becomes more or less coherent, or else R_{fb} thermally changes resistance with larger offset currents.

We are considered with how amplifier performance may accidentally produce a SQUID *jump* to a new voltage offset. To do this, we consider the *slew rate*.

Slew rate is an important quantity in determining stable operation of an amplifier. It defines the maximum rate of change of the amplifier's output. In the case of regular opamps, defining the slew rate is a matter of distinguishing the operation regime where signals are large and/or fast such that a power-bandwidth deficit can not let the amplifier output track the input. In the case of our SQUID amplifier, we are not as concerned about non-linearity or slew-limited pulses. Instead, we define a slew rate that will maintain the quasi-stable lock-point at the SQUID and prevent jumps.

We are concerned with the flux input of the SQUID. We do not want it to *jump* by exceeding the maximum of the $V-\Phi$ curve. We know that the flux error signal at the SQUID has a dynamic range that is some fraction of Φ_o. For now we make the (poor) assumptions that Z_{SQ} is constant over the dynamic range, and furthermore that the gain of the system does not effect the dynamic range. We could probably call ballpark the dynamic range as

$$\delta\Phi = \frac{\Phi_o}{4} \tag{I.71}$$

Referred to the input coil, the above dynamic range represents an amplitude

$$\frac{\delta\Phi}{M_{in}} = 5.34 \times 10^{-6}\ (Amp) \tag{I.72}$$

We define the *gain-bandwidth product* for the amplifier. This the DC gain multiplied by the break frequency, ω_{3dB}. For a single-pole amplifier, this product gives us the *unity gain* frequency, ω_t.

$$GBP = \omega_t = \frac{H_0}{R_{fb}}\ \omega_{3dB} = \frac{Z_{SQ}A_{VG}A_{int}}{R_{fb}}\omega_{int} \tag{I.73}$$

Above, we needed to borrow a factor of R_{fb} to account for the comparison of input current to output current.

So now we define a slew rate, which is the effective dynamic range multiplied by unity frequency. We divide by mutual inductance to refer it to the input coil.

$$S_R = \frac{\delta\Phi}{M_{in}} \times \omega_t = \frac{\Phi_o Z_{SQ}A_{VG}A_{int}}{4M_{in}R_{fb}}\omega_{int} \tag{I.74}$$

$$S_R = 4.9 \times 10^{-3}\ Z_{SQ}A_{VG}\ (\frac{Amp\ Hz}{\Omega}) \tag{I.75}$$

For the SQUID to jump, this slew rate needs to be surpassed by $\frac{di}{dt}$ of the input coil. Under normal conditions, the input coil is limited by R/L bandwidth. Given that is the case, the largest current pulse amplitude the input coil should be able to see is

$$i_{max} = \frac{S_R}{(R_{TES}/L_{in})} \approx 6 \times 10^{-4}\ Amp \tag{I.76}$$

for modest values Z_{SQ} and A_{VG}

In the frequency domain, noise referenced to the input coil can also be determined to meet this criterion if the noise-bandwidth product is comparable with the dynamic range, as from Eqn. I.72.

For instance, the input current noise dominates as we saw in the last section. Given R/L bandwidth once again, the noise-bandwidth product for this contribution is

$$\begin{aligned} NB_i &= |i_1\sqrt{R_{TES}/L_{in}}| && \text{(I.77)} \\ &= 3.2 \times 10^{-9}\ (Amp) && \text{(I.78)} \\ &\approx \frac{1}{1500} \times \frac{\delta\Phi}{M_{in}} && \text{(I.79)} \end{aligned}$$

I.6 Response to a test input pulse

Considering how the SQUID amplifier responds to actual physics signals is a better validation than characterizing with approximations or at a single frequency. To that end, we empirically describe a simple current pulse. We are concerned with *large*, *rapid* pulses, so quantities are approximated on the extreme side to that end.

Let's briefly touch on the input detector circuit (we otherwise do not touch on TES physics in this write-up). A TES bias current, I_b, is applied via a DAQ-controlled

voltage source passing through a $1k\ \Omega$ biasing resistor. $R_{shunt} = 20m\Omega$, and is always a magnitude less in resistance than a typical R_{TES}. Therefore, the majority of I_b will flow down the R_{shunt} branch, thereby locally voltage-biasing the TES. We are concerned with the largest change in current through the TES as it reaches high resistance.

We will *1)* find the largest expected quiessent current in the TES, *2)* assume the resistor reaches a high value during a pulse such that the quiessent current becomes the largest *change* in current, and *3)* use this value to normalize a time-dependent pulse.

We use a function of the form:

$$i_{input}(t) = \kappa(-e^{-t/\tau_{rt}} + e^{-t/\tau_{ft}}) = \kappa(-e^{-\omega_{rt}t} + e^{-\omega_{ft}t}) \tag{I.80}$$

In Laplace formalism, we take the transform

$$i_{input}(s) = \kappa(-\frac{1}{\omega_{rt} + s} + \frac{1}{\omega_{ft} + s}) \tag{I.81}$$

A biased QET channel gives typical bandwidths $\sim\ 100kHz$. Assume 100kHz. Considering an L_{in}/R_{QET} time-constant with $L_{in} = 4.5 \times 10^{-7}H$, we have $R_{QET} = 160m\Omega$. A typical I_b current might be $150 - 200\mu A$. Assume $200\mu A$. A DC current division between R_{shunt} and R_{QET} gives a quiessent I_s

$$I_s = \frac{R_{shunt}}{R_{shunt} + R_{QET}} I_{bias} = 22uA \tag{I.82}$$

Empirically, we see from phonon raw traces that fast QET risetimes may be 10 μs in Ge detectors, and as small as 5 μs in Si. Assume the extreme case in this scenario as 1.6 μs from the L/R time. ETF falltimes are $\sim\ 100\ \mu s$. Assume a relatively fast pulse has a falltime of 75 μs.

There is a phenomenon of *ballistic deficit*, where a finite risetime does not allow the overall pulse to reach the maximum amplitude calibrated for $\tau_{rt} = 0$. We make a correction such that the maximum pulse height is 22 μA.

where κ is normalized to make the maximum 22 μA.

$$\kappa = \frac{22\ \mu A}{0.90} = 24.4\ \mu A \tag{I.83}$$

$$\omega_{rt} = 1/\tau_{rt} = 630\ kHz \tag{I.84}$$

$$\omega_{ft} = 1/\tau_{ft} = 13.3\ kHz \tag{I.85}$$

I.7 Output response to test pulse

Armed with the above test pulse, we can calculate characteristic output parameters,

$$v_{out}(s) = i_{input}(s)H(s) \tag{I.86}$$

taking the inverse Laplace transform,

$$v_{out}(t) = \tag{I.87}$$
$$A_{int}A_{VG}R_{fb}Z_{SQ}\omega_{int}\kappa\{ \tag{I.88}$$
$$\frac{-e^{-\omega_{rt}t}}{A_{int}A_{VG}\frac{M_{fb}}{M_i}Z_{SQ}\omega_{int} + R_{fb}(-\omega_{rt} + \omega_{int})} \tag{I.89}$$
$$+\frac{e^{-\omega_{ft}t}}{A_{int}A_{VG}\frac{M_{fb}}{M_i}Z_{SQ}\omega_{int} + R_{fb}(-\omega_{ft} + \omega_{int})} \tag{I.90}$$
$$-\frac{R_{fb}(\omega_{ft} + \omega_{rt})e^{-\omega_{int}(1+\frac{\frac{M_{fb}}{M_{in}}Z_{SQ}A_{VG}A_{int}}{R_{fb}})t}}{(R_{fb}(\omega_{ft} - \omega_{int}) - A_{int}A_{VG}\frac{M_{fb}}{M_i}Z_{SQ}\omega_{int})(A_{int}A_{VG}\frac{M_{fb}}{M_i}Z_{SQ}\omega_{int} + R_{fb}(-\omega_{rt} + \omega_{int}))} \tag{I.91}$$
$$\} \tag{I.92}$$

So there we have it, in all its glory. The first term is proportionate to the risetime factor, while the second term is proportionate to the falltime factor. So far we just have an amplified pulse. The third term represents the filtering action of the amplifier, slowing down the pulse at the frequency of the amplifier's $\omega_3 dB$ bandwidth.

For the rest of this section, we see what effect amplifier parameters have on this test pulse.

I.7.1 Closed-loop amplitude from test pulse

Using the template time-domain pulse, we can calculate more specifically the error in the gain. The following plot shows the peak amplitude of the output signal, normalized to an ideal gain of $10k\Omega$.

We see that $Z_{SQUID} = 500\ \Omega$ and $A_{VG} = 40$ give a reasonable, consistent response. We would like to have these as minimum settings.

I.8 SQUID error signal and dynamic range

The SQUID's periodic response makes dynamic range a concern. We spoke of dynamic range before, by approximating the amount of flux subtended by the SQUID. We approach it slightly different now. Here, we have a numerical response and ask,

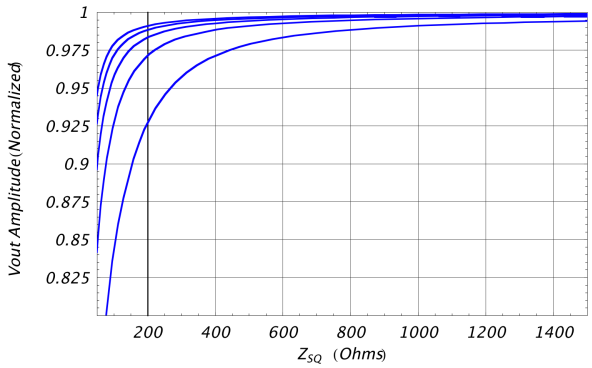

Figure I.10: Here we show the peak amplitude of the amplified pulse, normalized to perfect closed-loop gain. Again, the typical values of A_{VG} are traced.

how large does the SQUID voltage error signal become for a maximum input pulse. We consider the SQUID voltage output directly because we no longer need to ballpark the error signal by what portion of a Φ_o cycle it subtends. Even though the amplifier is closed in a feedback loop, the voltage error signal at the SQUID may be large enough to make the output pulse non-linear, or even cause the amplifier to loose lock (i.e., "jump").

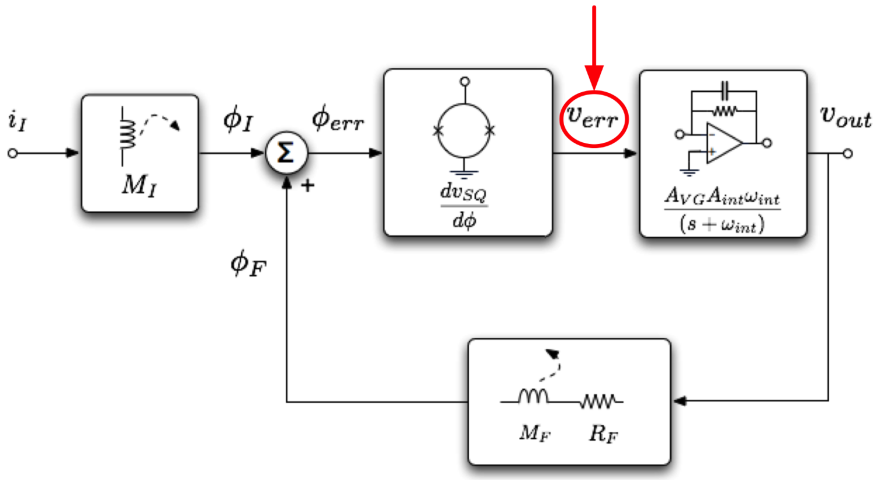

Figure I.11: Recasting the signal flow diagram, we see where the voltage error signal occurs along the readout chain. This helps us make sense of equation I.93.

Assuming the SQUID quiescent gain values stay constant during a pulse, the SQUID voltage error signal was calculated considering the test pulse template as input. The overall closed-loop gain of the amplifier has already been discussed. The

frequency response of the warm electronics is known to roll off at $\omega_{int} = 2.4kHz$. The voltage produced by SQUID within the closed loop must make up the remainder of the amplifier response via the following equation:

$$v_{error}(s) = \frac{i_{input}(s)H(s)}{A_{VG}A_{int}\frac{\omega_{int}}{\omega_{int}+s}} \tag{I.93}$$

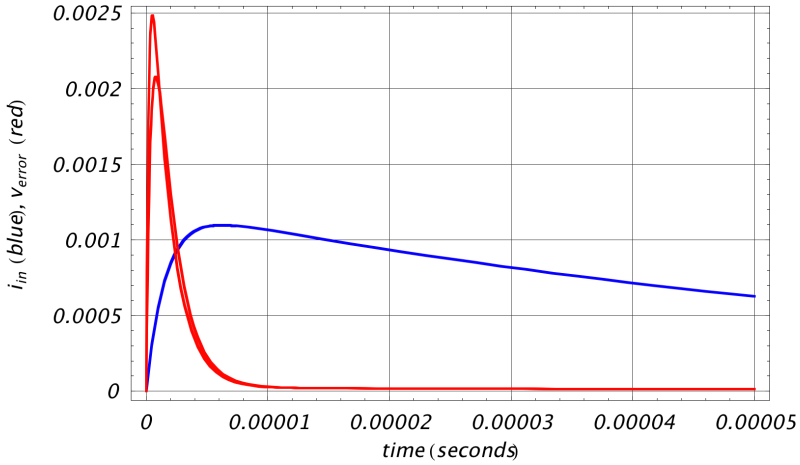

Figure I.12: Here is the SQUID voltage error in the time domain for $A_{VG} = 40$ and for values of $Z_{SQ} = 500, 1000\Omega$. Imposed (in blue) is the input signal, scaled by a multiplier of 50.

We are concerned with the SQUID error signal becoming too large. Empirically, we impose a minimum dynamic range for the SQUID voltage. This requires a $V - \Phi$ modulation curve with usable (relatively linear) range 1.5 times the maximum error signal above the lock point, and 0.5 times that below. This forms a window around the lockpoint. A_{VG} lowers the maximum error signal amplitude, while Z_{SQ} increases it.

This requirement is probably the most demanding in SQUID tuning. We now see how SQUID bias current and choice of voltage lock point become trade-offs. We can adjust current bias to increase dynamic range, which generally broadens out the response and decreases Z_{SQ}. In light of this requirement, we have begun tuning with broad $V - \Phi$ curves, as it is possible to have a Z_{SQ} that is too large. As long as the input coil noise is dominating ($Z_{SQ} \geq 500\ \Omega$), the other requirements are generally met by the $Z_{SQ}A_{VG}$ product. Therefore, it is best to tune for a lower Z_{SQ}, and compensate with A_{VG}.

To emphasize, this examination of the error signal shows that this node of the amplifier is *very susceptible* to high frequencies. Within the amplifier chain, the

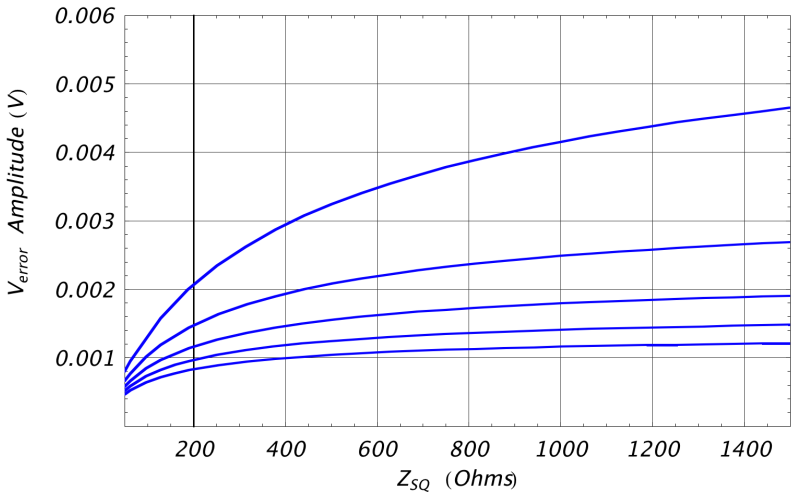

Figure I.13: The pulse maximum for the SQUID voltage error as a function of Z_{SQ}, for the standard values of $A_{VG} = 20, 40, 60, 80, 100$. Smaller A_{VG} results in *larger* error signal.

integrator is rolling off by $1/f$ at frequencies above the integrator pole. Yet, the passband of the amplifier is required by feedback to be flat out to some total amplifier bandwidth. Therefore, to account for the difference, the amplitude of the voltage of the error signal must *increase* by f within the passband. So the error voltage across the SQUID carries the burden of larger amplitudes for faster signals. This leads to greater instability in terms of flux jumps when the SQUID amplifier is tuned for higher overall bandwidth.

I.9 Direct, flux-coupled feedback in the SQUID amplifier

Experimentally, we discovered the oddity that a change in input inductance is seen in SQUID operation which is different between closed and open loop configurations. This bandwidth change is measurable and repeatable. It appears to be due to an additional source of feedback. It was found that the SQUID array washer focuses magnetic flux in a way that accounts for this effect, providing a direct flux-coupling between feedback and input coils.

I.9.1 Transformer equations

If we consider that there exists a direct transformer coupling between the feedback and input branches of the SQUID amplifier, in addition to the typically assumed input-SQUID and feedback-SQUID coupling, then we find an avenue for a parallel feedback channel to occur. Note figure I.14a. Through this mechanism, feedback signal is induced directly back into the input coil, then back again into the SQUID summing node via the the input mutual inductance.

A review of transformer equations is a healthy exercise. There are several good electronics references on the topic. The idea is to represent Kirchhoff's laws for both coil voltages, using a 2x2 matrix formalism of impedances and currents. Solving for the currents returns expressions with somewhat large determinants in their denominators.

A simple and yet general outcome of the transformer equations is the *current ratio*. We will not derive this here, but it is to say that determinants cancel out and it is rather elegant there are few terms we need struggle with. For the transformer circuit in figure I.14b, the current ratio is

$$\frac{i_s}{i_f} = \frac{-sM_{if}}{(R_{sh} + R_s + sL_i)} \tag{I.94}$$

Here, i_s is the sensor (input) current, i_f is the feedback current, M_{if} is the feedback-input mutual inductance (assumed symmetric, incidentally), R_{sh} is the shunt resistor, R_s the sensor resistance, and L_i the input coil inductance.

Let's interpret this. No matter the polarity chosen for SQUID lock in our amplifier, when in stable operation the input coil current is *always opposing* feedback current. Therefore, the minus sign of this term due to Lenz's law always *assists* the input coil. This tells us that impedance is being lowered. It makes sense that the strength of this effect is proportionate to the coupling M_{if}, and the denominator informs us that more impedance in the sensor line loop will logically diminish the induced sensor current. Also, as this effect induces *more* signal current, we can see that it is fighting stability and is a positive feedback mechanism.

I.9.2 Revisiting the closed-loop SQUID amplifier

The transfer function from the beginning sections of this appendix give us the form:

$$H(s) = M_i \frac{\alpha(s)}{1 + \alpha(s)\beta(s)} \tag{I.95}$$

Here, $\alpha(s)$ is the "open loop gain," and $\beta(s)$ is the "feedback factor." The prefactor of M_i in this case is particular to our SQUID amp situation, and serves to convert the input current signal into proper units (of flux) at the SQUID summing node. The

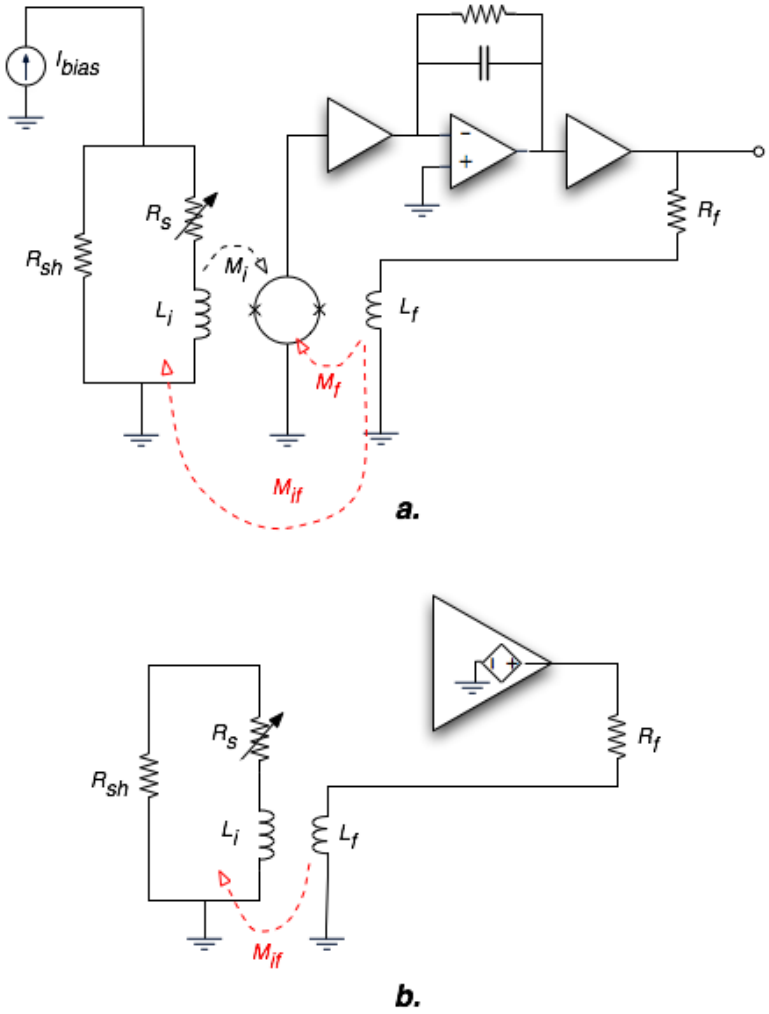

Figure I.14: **a.** A direct mutual inductance between feedback and input allows for another feedback channel. **b.** The ratio of input and feedback current for this mechanism is determined from transformer equations. The active nature of the amplifier output assists input current through the inductor and lowers the impedance.

open loop gain, α is determined from the transducer (SQUID) chain at the input on up through the active electronics.

We have

$$\alpha(s) = \frac{Z_{SQ}}{M_i} A_{VG} \left(\frac{A_{int}\omega_{int}}{s + \omega_{int}}\right) \tag{I.96}$$

Here, Z_{SQ} is the transresistance ("responsivity" sometimes) of the SQUID. This relates a change in input current to a change in SQUID voltage. The rest of the terms represent stock amplifier voltage-to-voltage gains. A_{VG} is the variable gain (including preamp). The last term in parentheses represents the gain and frequency dependence of the integrator, whose frequency dependence will ultimately allow the entire amplifier reach a stable, finite bandwidth.

The standard feedback factor is

$$\beta(s) = M_f / R_f \tag{I.97}$$

This relates an output amplifier voltage back to the SQUID as a flux signal. From now on, we will call this term β_0, to distinguish it as the original, intended feedback term.

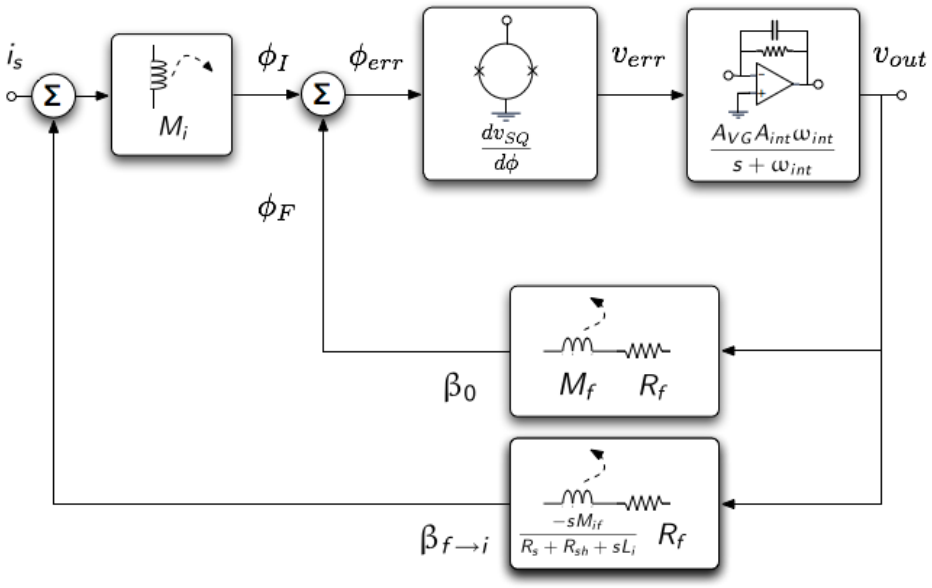

Figure I.15: A signal-flow diagram, including the additional feedback factor. Note that $Z_{SQ} = \phi \frac{dv_{sq}}{d\phi}$ was used to group the elements in a logical way.

I.9.3 The new feedback factor

Looking at figure I.15, we see a signal-flow diagram that puts our new situation into perspective.

For the portion of sensor current determined by voltage output, via the new feedback factor, we have

$$i_s = v_o[\frac{1}{R_f}(\frac{-sM_{if}}{R_{sh}+R_s+sL_i})] \tag{I.98}$$

From here, this induces flux signal back into the SQUID as typical, by way of a factor M_i.

Re-expressed,

$$\beta_{f\to i}(s) = \frac{M_i}{R_f}(\frac{-sM_{if}}{R_{sh}+R_s+sL_i}) \tag{I.99}$$

Some thought will show us that our feedback factors, from the output amplifier *voltage* to the input branch *current* (or flux), are acting as *conductances*. They are also in parallel. Therefore, they add accordingly.

$$\beta(s) = \beta_0(s) + \beta_{f\to i}(s) \tag{I.100}$$

This expression is all that is new in this work. We now make sense of its impact.

I.9.4 Combined transfer function

Again, our transfer function is

$$H(s) = M_i\frac{\alpha(s)}{1+\alpha(s)\beta(s)} \tag{I.101}$$

We find one zero and two poles. Taking ω_{int} to be small, calling $M_{if} = k\sqrt{L_iL_f}$, with k as the coupling,

$$H(s) = \frac{\mathrm{A_{int}A}_{VG}\mathrm{Z}_{SQ}\omega_{int}}{s\left(1+\mathrm{A_{int}A}_{VG}\mathrm{Z}_{SQ}\omega_{\mathrm{int}}\left(\frac{\mathrm{M_f}}{\mathrm{sR_fM}_i}-\frac{k\sqrt{\mathrm{L_fL}_i}}{\mathrm{R}_f(\mathrm{R}_s+\mathrm{R}_{\mathrm{sh}}+s\mathrm{L}_i)}\right)\right)} \tag{I.102}$$

I.9.5 The case of ideal gain

It is informative to take the case where the open-loop gain is infinite. Here,

$$H(s) \approx \frac{M_i}{\beta(s)} = \left(\frac{\frac{M_i}{M_f}R_f}{1-k\sqrt{\frac{L_f}{L_i}}\frac{M_i}{M_f}}\right)\frac{s-\left(\frac{-(R_s+R_{sh})}{L_i}\right)}{s-\left(\frac{-(R_s+R_{\mathrm{sh}})}{L_i-\frac{M_i}{M_f}k\sqrt{L_fL_i}}\right)} \tag{I.103}$$

The expression is put intentionally into zero-pole-gain form. The first term in parentheses is the "gain" term, and we can recognize the factor of $\frac{M_i}{M_f}R_f$, which is the

DC gain in the ideal SQUID amp transfer function. Also, we have a zero at $\frac{-(R_s+R_{sh})}{L_i}$ and a pole at $\frac{-(R_s+R_{\rm sh})}{L_i-\frac{M_i}{M_f}k\sqrt{L_fL_i}}$.

What does this mean? Well, the input coil circuit itself has a natural pole at $\frac{-(R_s+R_{sh})}{L_i}$. Now, the *amplifer* nulls that out with a zero precisely at the same frequency. However, the ideal amplifier gain also inserts its pole at a higher frequency. In effect, this is actively raising the innate bandwidth of the input coil. Perhaps we could call this an active inductance lowering, where the effective input-coil inductance becomes

$$L_{eff} = L_i - \frac{M_i}{M_f}k\sqrt{L_fL_i} \tag{I.104}$$

I.9.6 A sensor impedance with frequency structure

What if we also consider that our sensor resistor is not simply a static resistor, but has internal frequency structure (which is the case).

$$R_s \rightarrow Z_{TES}(s) \tag{I.105}$$

We see that we would have an ideal gain expression of the form,

$$H(s) = \left(\frac{\frac{M_i}{M_f}R_f}{1-k\sqrt{\frac{L_f}{L_i}}\frac{M_i}{M_f}}\right)\frac{s-\left(\frac{-(\mathbf{Z_{TES}(s)}+R_{sh})}{L_i}\right)}{s-\left(\frac{-(\mathbf{Z_{TES}(s)}+R_{\rm sh})}{L_i-\frac{M_i}{M_f}k\sqrt{L_fL_i}}\right)} \tag{I.106}$$

I.9.7 Finding a 2nd-order transfer function due to M_{if}

Bruce Hines followed this through to write out this model in the form of a 2nd-order response. His experimental measurements of the open-loop SQUID chain provided precise values for the coupling quantities based on this model. The interested reader may want to read our reference [125]. Based on referee reports, we were informed how similar treatments had accounted for the effect of a direct mutual coupling, M_{if}, in the past [126, 127].

Below in figure I.16, we show a small-signal SQUID-equivalent circuit that we simulated using NI MultiSim 10. Through the use of analog behavioral modeling, we were able to reproduce the effects depicted by these models.

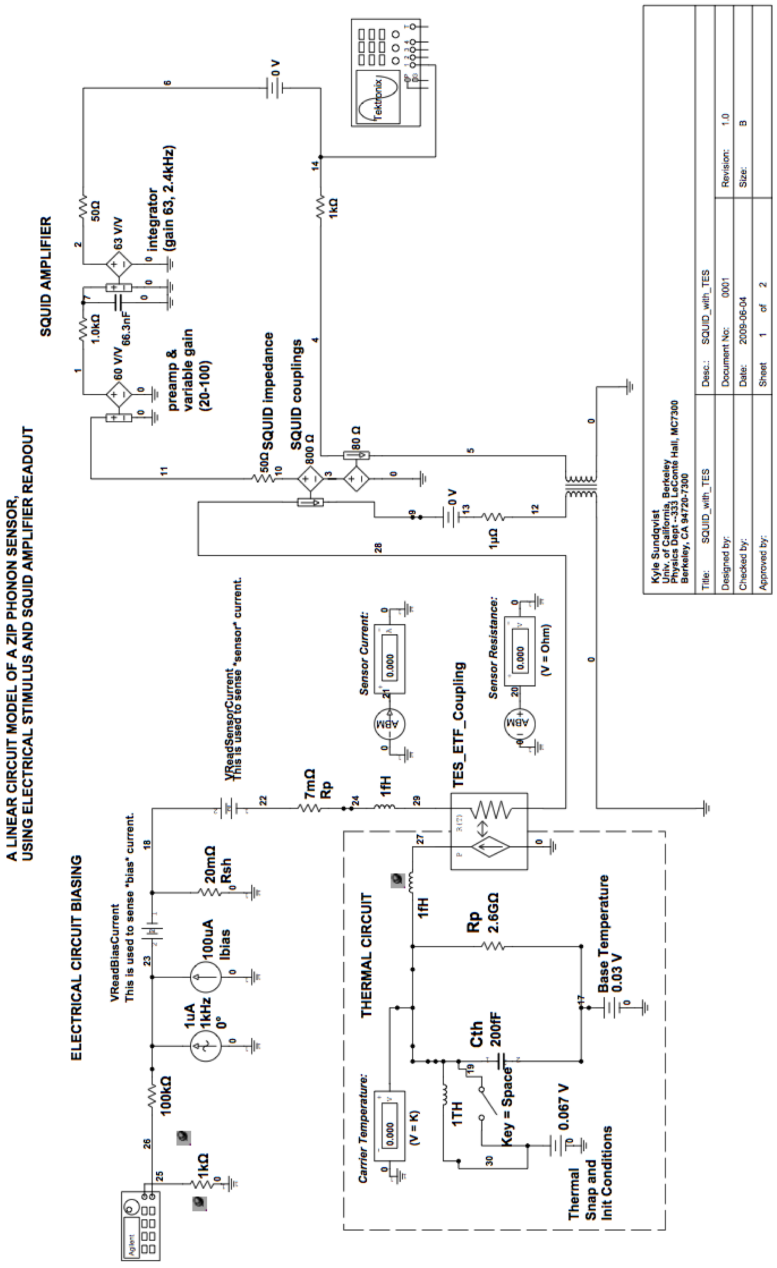

Figure I.16: A circuit simulation of a linearized TES + SQUID system, using analog behavioral modeling in NI MultiSim.

I.10 Conclusion

In this appendix, we have defined a small-signal transfer function for the CDMS SQUID amplifier. From this, we developed an understanding of some key concepts of gain and bandwidth tradeoffs. This gave us insight for developing the SQUID-tuning guidelines used for the "SQUETTest" tuning program at Soudan. A key understanding comes from realizing that high-frequency signals affect the SQUID error signal much more dramatically than lower frequencies. This sets the limits that need to be established for bandwidth and dynamic range.

Another discovery described here is that there has been a significant flux-coupling directly between feedback and input coils. This is responsible for an actively boosted bandwidth (i.e., an actively lowered input inductance) at the input coil.