BEAM-WAVE INTERACTION FOR A TERAHERTZ SOLID-STATE AMPLIFIER

By

Matthew Steven Hodek

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering – Doctor of Philosophy

2021

ABSTRACT

BEAM-WAVE INTERACTION FOR A TERAHERTZ SOLID-STATE AMPLIFIER

By

Matthew Steven Hodek

The push of conventional electronic amplifier technologies into the deep submillimeter wavelength and THz frequency ranges of the electromagnetic (EM) spectrum has been limited by constraints on their fundamental physics of operation and fabrication limitations. At the same time, optical amplifier technologies can only access this spectral region using inefficient frequency down-conversion. This struggle for practical power amplifiers in the THz band will likely require a new type of amplifier and has led to a desire for a solid-state beam-wave style amplifier using semiconductor fabrication techniques.

While there has been considerable progress in creating transistors in the THz region [1], the small size required to achieve the needed transit times and gate capacitances generally precludes them from producing power above 1 mW. Vacuum electronic devices (VEDs), such as traveling wave amplifiers (TWAs), have also shown great progress into this band [2]. A TWA is an example of a beam-wave style device where gain is achieved by transferring energy from an electron beam to an EM wave at electrically large length scales. However, as traditional TWAs are scaled to higher frequencies, the shrinking wavelength makes fabrication of the corresponding interaction circuit structures and miniscule beam tunnels increasingly difficult through micro-machining or other subtractive metal shaping. Thus, combining the strengths of both these systems into a single device has some merit. Solid-state TWAs have been attempted over many years without success largely due to slow electron drift velocities resulting in beam equivalents that are unsuitable for synchronization with EM slow-wave structures. One possible path towards a beam-wave style THz solid-state amplifier is to couple to a plasma wave characterized by phase propagation much faster than the electron velocity limited by scattering in a material, but this requires a substantial redevelopment of the fundamental beam-wave interaction analysis.

Presented here is a novel analysis built upon the prior work on solid-state and VED TWAs with a primary difference in the nature of the charge carrier behavior. In this work the electron beam, which was previously described as bulk carriers in a semiconductor, is now formed with an un-gated 2D electron gas (2DEG). A freely propagating plasma wave is present in the dense 2DEG and takes the place of the typical space charge wave present in VED devices. Example calculations are compared to a generic VED TWA behavior and the basic performance of a realizable device is analyzed through the use of a gallium nitride heterostructure material system and achievable fabrication strategies. It is shown that the concept of a TWA using a 2DEG plasma wave is not practical at best, and fundamentally flawed at worst. However, the understanding gained lays some of the groundwork for other possible beam-wave interaction style amplifiers using a fast 2DEG plasma wave.

To my dad, who taught me to never stop learning.

ACKNOWLEDGMENTS

Thank you to the Defense Advanced Research Projects Administration for supporting this work under contract numbers HR0011-21-9-0058 and W31P4Q-17-C-0037.

Thank you to my collaborators and colleagues I have worked with in the course of this work. Thank you to David Smithe for countless hours of discussion and advice. Thank you to Patrick Wong for helping me to better understand the theory. Thank you to Peng Zhang, Marco Saraniti, and Jung-Hun Seo for sharing their expertise. Thank you to Nicholas Miller at AFRL for helping with the wafer measurements.

Last, but certainly not least, a sincere thank you to my advisor John Albrecht. You may have gotten me into this mess, but you definitely guided me out of it too.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PRACTICAL THz AMPLIFIERS: PHYSICAL CHALLENGES AND THE THz GAP

The use of the electromagnetic (EM) spectrum is ubiquitous in our current society. For over a hundred years people have been experimenting with and subsequently using EM waves at different frequencies to accomplish many technological achievements. RF and microwave frequencies (roughly 1 kHz – 300 GHz) are commonly used for the communication of data and for sensing applications like radar. Optical frequencies (roughly 3 THz - 100 PHz) spanning the ranges of infrared, visible, and ultraviolet light have been used for centuries in too many ways to enumerate here. In-between those two well-known regions of the EM spectrum is the terahertz (THz) Band. The difficulties faced in working within this band, and thus its lack of development, has led to it being called the THz Gap.

Many of the fundamental issues in working in the THz band come from the range of wavelengths and wave energies involved. The free space wavelength at 1 THz is 300 μm, in silicon (Si) it is closer to 25 μm. This is approaching the dimensions of common solid-state RF electronic devices which pushes the limits of how they physically operate. At the same time, the wave energy of about 4 meV is sufficient to start interacting with low frequency crystal vibrational modes or molecular states in materials. While these interactions provide a way to learn more about the materials (through spectroscopy), they can also render commonly used semiconductor and dielectric materials to be too lossy to be effectively used in the THz band.

There have been numerous attempts to create measurement systems in the THz range that have leveraged the established expertise in electronic and optical hardware, and both have faced major challenges. While the THz band is difficult to work within, there are many different applications that are looking to it for answers. In many cases, the same reasons that make it difficult to work with are in fact what make its use interesting. THz imaging is of growing interest in the medical community. THz radiation is non-ionizing which means it doesn't have the health impact that exists with X-rays. By using THz radiation to image soft tissue, it has been shown to reveal detailed information that is

inaccessible with other imaging methods. One modality is to use a 500 GHz reflected signal image to differentiate the water content in soft tissue to highlight damaged or diseased areas [6].

As the vast field of semiconductor-based electronics continues to expand, more detailed information is needed about how the materials involved behave as EM media when extended into the THz Gap. THz-based spectroscopy has been shown as a way to probe certain material properties and transient behaviors. It is possible to measure directly the frequency-dependent complex valued transient conductivity using time resolved fast pulse THz spectroscopy [6]. Detailed information about the basic physical properties of most semiconductors and dielectrics is sparse, unreliable, or absent at THz frequencies.

Communications applications dominate the crowed RF spectrum and wireless communication systems are requiring an increasing amount of bandwidth as their use expands. Whether it is a short range terrestrial system, or space-based system, even a modest fractional bandwidth ($< 1\%$) at 1 THz can be a significant improvement over current technology. A 40 Gbit/s link has already been shown [7], and systems with over 100 Gbit/s capabilities are being planned.

So there is certainly interest in THz systems for various applications. Several options exist for uncontrolled, bright power sources in the THz band, but there is an absence of amplifiers of intentional excitations. For many of these applications a THz amplifier that allows coherent measurements in a small and affordable package is critical for their development. A coherent measurement is one where both the amplitude and phase are quantified. The combination of these two pieces of information allow the community to bring to bear advanced measurement techniques that have been refined over several decades in the wireless communications field. By using these techniques, the signal-to-noise ratio (SNR) can be reduced significantly to allow more sensitive measurements [8]. In the case of wireless communications, the use of coherent systems also can lead to using more complicated modulation schemes. The 40 Gbit/s data link mentioned above could only use a basic on-off keying (OOK) modulation scheme since the THz source was non-coherent. That data rate could be much higher with a coherent system that can use contemporary waveform engineering, such as the popular quadrature amplitude modulation [7].

CHAPTER 2

HIGH FREQUENCY AMPLIFIERS

2.1 Types of Amplifiers

Having established the need for a THz amplifier, the research problem will now be clearly defined. In the most basic and ideal sense, an amplifier is a device that takes an electrical input signal and outputs a congruent signal with a larger amplitude. In the RF/mmW field, typically the signal power is increased, but this can be done via an increase in the signal voltage and/or current depending on the needs of the system. Regardless of the implementation, the amplifier must be provided a source of energy to increase the signal amplitude.

The earliest forms of electronic amplifiers were the ubiquitous vacuum tubes. These devices take a DC stream of electrons boiled off a metal cathode and accelerates them through the vacuum towards an anode with an applied DC voltage between these electrodes. A metal grid is positioned in the middle that can be electrically charged with the time varying input signal. When the grid has a negative charge it will repel the vacuum electrons, and prevent them from reaching the anode where the output signal connection is made. In this fashion a simple vacuum tube acts much like a valve turning on and off a stream of electrons with the application of an input voltage. There are many variations of this basic concept, and they all fall into a category known as vacuum electronic devices (VEDs). Solid-state electronics use transistors to implement the same principle of an electron valve, but in a semiconducting material instead of a vacuum envelope. In a field effect transistor (FET) the gate electrode repels or attracts a conducting channel between the source and drain contacts.

In both the vacuum tube or the transistor, the input signal modulates a stream of electrons. At low frequencies, these devices can easily respond to input signal. As the frequency increases, more parasitic effects come into play, and start to affect the performance of the device. It is generally the case that these devices should be small compared to the wavelength of the signal being amplified. This ensures that the entire grid of a tube or gate of a transistor is at the same voltage at any given

time. This is known as the quasi-static approximation. The charging capacitance of the grid or gate, i.e. how quickly the needed charges can be gathered to affect the electron stream, will also limit how quickly the electron stream can be modulated. Both issues are generally dealt with by making smaller devices (both vertically from gate to channel and horizontally from source to drain) as the frequency of the signal is increased. This continual shrinking the device sizes has at least two major issues. First, smaller devices become increasingly difficult to fabricate. The second is that drain breakdown effects will limit the power handling capabilities of smaller devices and the dynamic range of voltage waveforms.

Given the narrow-band nature of typical valve style vacuum tubes, or even klystrons, and an increasing demand for bandwidth, a new type of vacuum electron device was envisioned in the first half of the 20th century. Instead of having a single grid modulating the electron beam, this device has a specially designed waveguide that is constructed to propagate the input signal along the same axis as and in close proximity to the electron beam. A key feature of this waveguide is that the propagation velocity of the RF wave is much slower than free space propagation to match the electron beam velocity. As such, they are known as slow-wave structures (SWSs).While the RF wave and electron beam are co-propagating down this structure, the wave and beam interact through mutual electromagnetic forces. If this interaction structure is designed correctly, the kinetic energy of the electron beam is transferred into the co-propagating RF signal. Haeff first published about the possible interaction between an electron beam and a traveling RF wave in 1933. His initial idea was for an electronic detector, but theorized that it could also be used as an amplifier. Later, in 1940 Lindenblad made a VED that modulated the electron beam via a grid near the source, and then generated an output signal by propagating that beam down the axis of a helix. The first traveling wave amplifier (TWA) as they are known today, with both the electron bunching and the wave growth due to propagation along the helix, was created by Kompfner in the early 1940's [9].

2.2 Pushing to Higher Frequencies

As with most RF devices, scaling a TWA to higher frequencies requires that the device dimensions shrink proportionally to the wavelength in various fashions. The most obvious is the scaling of the slow wave structure. If a helix waveguide is used, it is important that the slow wave wavelength extends across several turns of the helix. For higher frequencies the helix will need to have a higher number of turns per length and the transverse dimensions will also have to shrink to get the desired phase velocity. Similar arguments can be made for the size of a coupled-cavity SWS. In the THz range, these dimensions are pushing past what can be done with traditional machining techniques. To deal with these small dimensions, new processes that have been largely developed by the micro-electro-mechanical systems (MEMS) have been brought to bear on the problem. Even these processes struggle to fabricate features sufficiently small to support operation above 0.8 THz [6]. Worse, the practical issue of avoiding beam scrape in the vanishingly small beam tunnels of these circuits is often too much to overcome for practical devices.

2.3 Comparison With Other Technologies

It is instructive to compare the physical challenges for extending to 1 THz across technical approaches. Currently there are multiple options for reaching 1 THz, some are based on lower frequency electronics approaches, and some are based on optical techniques from higher frequencies. In most cases the solutions can't produce significant power with reasonable efficiencies in a compact package. The following examples will quantify some of those aspects.

THz transistors have been a major goal of the DARPA THz Electronics program for more than a decade. To maximize the performance of these devices, they use the semiconducting material indium phosphide (InP) which has a very high mobility. With this material they have developed scaled process to shrink the dimensions and boost high frequency performance. A record setting transistor shown in [10] has 25nm gate width and can achieve an Fmax of 1.5 THz. The thing to keep in mind with transistors, and amplifiers in general, is that to accurately reproduce modulated signals the frequency limit of the device (Ft/Fmax for transistors) should be several times higher in

frequency from the carrier to account for the harmonic content of the signal. The devices developed by these programs have been successfully used in a terahertz monolithic integrated circuits (TMICs) with operating frequencies in the 0.3-1 THz range. These TMICs are very compact with dimensions on the mm scale for the entire chip, and have shown good integration into standard waveguide packages [10]. The output power of these heavily scaled devices is not high enough for a transmitter application. These TMICs will likely still play a critical role as low noise amplifiers and driver amplifiers in coherent THz systems.

Lasers and photoconductive materials are currently the easiest way to generate THz signals with benchtop systems for THz time domain spectroscopy (TDS) being commercially available. These systems rely on the laser exciting a photo current in a semiconductor, typically low temperature grown gallium arsenide (LT-GaAs), and use that current to radiate a THz signal. The THz generation typically happens via two different methods. In a THz frequency-tunable system, two laser beams that are offset by the THz frequency desired are directed to a wide bandwidth antenna pattern, such as a spiral, on the LT-GaAs. A single frequency system can be implemented with a single laser shining on a structure with a pattern matching the desired wavelength such as a grating. In both cases, the laser generates pulses of electrons underneath the patterns and excites them to radiate [6]. By carefully tuning the photo-conducting substrate, such systems have achieved up to 90 dB of dynamic range [11]. This is significant for a non-coherent system, but it comes at the expense of a large system and high cost with the use of a fiber laser.

Quantum Cascade Lasers (QCLs) are a source of THz radiation in a small package that have garnered a lot of attention over the past 20 years. Based on a cascade of tuned quantum wells, a QCL can produce watts of power in the 1-5 THz range. However, it still just a THz generator with a very narrow bandwidth [6]. A QCL could potentially be used as local oscillator in a THz system, but THz amplifiers would still be needed.

2.4 Wave Style Amplifiers Using Solid-State Fabrication

As previously stated, it is becoming increasingly difficult to scale the traditional VEDs to higher frequencies. It is also clear that traditional semiconductor devices are not going to be able to produce the amounts of power needed in advanced systems. Hence, a merger of the two technologies is proposed. The concept is to utilize semiconductor fabrication techniques and materials to produce a device that has an extended beam-wave interaction producing amplification on an electrically large length scale.

The general idea of a solid-state TWA was introduced for very similar reasons in the early 1960's, but for lower frequencies than we would presently like to consider. Using semiconductor fabrication techniques allows for much smaller EM SWSs and electron beams than in the vacuum case without the use of hot cathodes or bulky external magnets for beam control. Inherently being an electrically large device, a solid-state TWA would not be limited to the small dimensions of transistors and thus more gain and higher powers can be achieved. It is important to note that within this context, a TWA refers specifically to a class of amplifiers that transfer DC kinetic energy of an electron beam continuously to a co-propagating EM wave within a multi-wavelength interaction region. The term TWA has been used loosely in the literature to describe other types of solid-state amplifiers that use discrete parametric type gain devices. We consider parametric gain devices as distinct from the physics associated with TWAs.

Early work on solid-state TWAs by Sumi suggested simply replacing the vacuum beam with the conducting channel in semiconductor [12]. These concepts did not accurately include loss mechanisms in the form of thermal velocity and electron collisions. Later Solymar introduced a theory for a bulk semiconductor TWA device that included the previously missing semiconductor transport effects [13]. Gover then used this derivation with the addition of a dielectric waveguide for two further studies. One was in the collisionless regime such that $\omega\tau >> 1$ [14], and the other was done in a collision dominated case $\omega\tau < 1$ [15]. For purposes of this discussion $1/\tau$ is the average collision frequency limiting the free electron mobility in the semiconductor and ω is the angular frequency of the THz wave.

The primary problems with the semiconductor TWA concept are the SWS feature sizes and inadequate coupling between the SWS and the beam. One of the main points of Solymar's paper is that a strong coupling constant is needed to overcome the collisions and thermal velocity. Gover tries to achieve this using his proposed dielectric waveguide, but what is not discussed are the relative speeds of the EM wave in the waveguide and the electron velocity. The example Gover gives uses a free space wavelength of 100 μm or around 3 THz. In his proposed waveguide of GaAs ($\epsilon_r \approx 13$) the TM mode wavelength will be on the order of 40 μm, depending on the particular geometry. The grating on the surface of the waveguide that forms the SWS matched to the electron velocity has a pitch of only 0.06 μm. With the SWS pitch being much smaller than the EM wavelength, the grating will only weakly interact with the propagating EM wave. In this thesis, the development of a useful SWS is a key topic that is discussed in Chapter 5.

One possible path towards a beam-wave style THz solid-state amplifier is to couple to a wave that is much faster than the electron velocity in the material, but this requires significant change in the approach to deriving the interacting wave solution compared with traditional TWA theory dating back to the 1950s. The first step is to use a different kind of electron beam. Based on theory by Stern [16], and confirmed by experiment [17], it is known that a plasma wave in a dense gas of electrons confined to two dimensions has a propagating plasma wave that may be appropriate for this use. A key insight of this thesis is recognizing that the common electron channels in electron devices that are confined in the third dimension provide these beam options with many advantages. The following chapter will elaborate this concept in detail.

What is presented in Chapter 4 is a continuation of the prior work of Solymar with a primary difference in the nature of the charge carrier behavior. In this work the electron beam, which was previously described as bulk carriers in a semiconductor, is now formed with an un-gated 2D electron gas (2DEG). A freely propagating plasma wave is present in the dense 2DEG and takes the place of the typical space charge wave present in VED devices as shown in Fig. 2.1. An example calculation is compared to a generic VED TWA behavior. It is shown that the 2DEG plasma wave limits any realization of the hypothetical device in the typical TWA mode of operation,

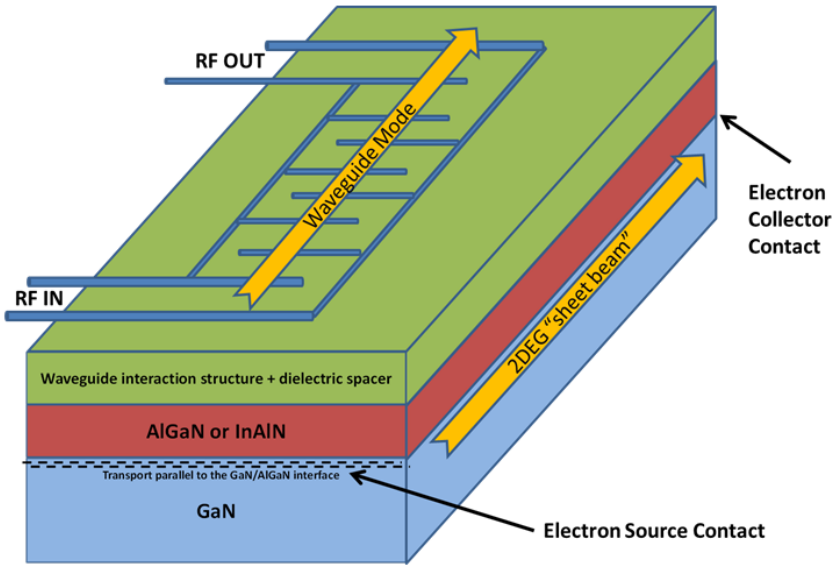

Figure 2.1: A conceptual diagram of a solid-state beam-wave amplifier utilizing a 2DEG beam.

but in the elaboration of the design space it is also shown that it opens up new possibilities for other beam-wave interaction style amplifiers.

CHAPTER 3

TWO-DIMENSIONAL ELECTRON GASSES

3.1 A Brief Introduction to Semiconductor Transport

A little must be said on the properties of the 2DEG before going on to describe a theory for a beam-wave style amplifier that would employ it. This begins with a basic understanding of electron transport within semiconductor materials.

For a solid material to conduct electricity it needs electrons that are in states where they can freely move among the atoms that form the solid and the crystal potential field seen by the conduction electrons. As atoms are taken from isolation and forced closer together, their electron orbitals start interacting and hybridizing into distinct energy bands. These energy bands represent the collection of possible states within which electrons can exist. The tightly-bound core electrons of the atoms preserve the chemical nature of the elements and contribute to the background potential, but the physical extent of some of these extended bands extends between adjacent atoms, and allow for electrons to freely move around the solid. These partially filled states are known as the conduction bands. The lower energy bands of these extended states that are typically filled are known as valance bands and do not contribute to conductivity. The forbidden region between these two energy bands is known as the band gap [18].

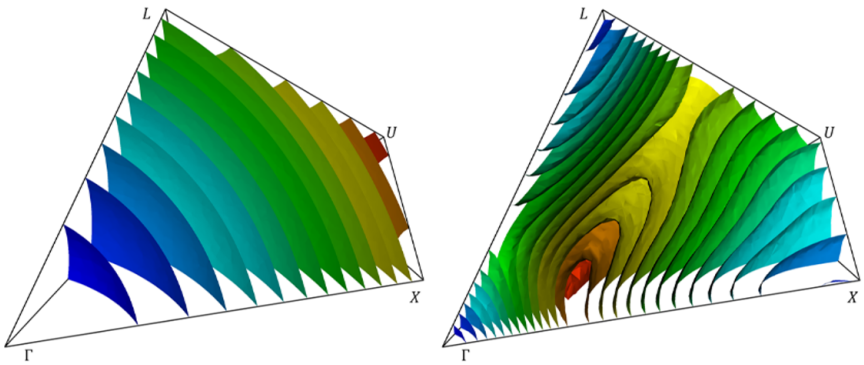

Figure 3.1: Energy Contours for the parabolic energy approximation (left) and the first conduction band of GaAs from EPM (right) within the irreducible wedge.

The arrangement of the energy bands for a solid, collectively referred to as the band structure, is calculated within the momentum space of the crystal. The regular crystalline structure of semiconductors in real space causes a reciprocal structure in the electron momentum space, and the smallest unit cell that can tile the entire momentum space is called the Brillouin zone. The Brillouin zone can be further reduced in size using symmetry arguments into the irreducible wedge. Shown in Fig. 3.1 are two numerical representations of the GaAs band structure in the irreducible wedge for a zinc-blend crystal. The calculated band structure on the right was created using the empirical pseudo-potential method, and is intended to reproduce many features of measured band structures [19]. It is often useful to have a closed form representation of the band structure that can be used in analytical models. Shown on the left in Fig. 3.1 is the band structure using the parabolic bands approximation where the energy above the conduction band minimum is approximated as $E = \frac{\hbar^2 k^2}{2m^*}$. This approximation uses the electron effective mass m^* that allows the electron within a crystal to be described as a pseudo-particle that obeys classical mechanics [20]. The parabolic bands approximation can be used when the electron energy is low enough that most electrons are within the lower curvature of the first conduction band. The theory derived in Chapter 4 does make use of the parabolic approximation which allows the use of force equations and the effective mass instead of including more complicated functions of the band structure.

As its name suggests, the density of states $g(E)$ is a function that describes how many states are available for electrons to occupy at a given energy within the band structure. The density of states for the parabolic bands approximation can be written as $g(E) = \frac{4\pi(2m^*)^{3/2}}{h^3}\sqrt{E}$. The ability of a solid to conduct electrons is determined by which of these bands have electrons in them. The probability of an electron to have the energy E at a given temperature is given by the Fermi-Dirac Distribution function $f_F(E) = 1/\left(1 - exp\left(\frac{E-E_f}{K_B T}\right)\right)$. At T = 0 K, the Fermi function is a step function at the Fermi Energy, E_f. In general the Fermi Energy is the point where there is a 50% probability of being occupied by and electron. As the temperature rises, the tails of the Fermi Function extend in both directions allowing for electrons with more energy. If this tail extends high enough in energy, the electrons can transition to the conduction band and move about the solid. The

density of electrons at a given energy is then given by $n(e) = g(E)f_F(E)$ [18].

What has been described thus far is the band structure and electron states at equilibrium. When an external stimulus is applied, typically by applying a voltage or injecting carriers, the net effect is that it moves the Fermi Energy relative to the energy bands and this changes the probable distribution of electrons. This is the process by which a voltage applied to the gate of a transistor can turn on or off the conduction channel below it. One of the primary differences between the electrical properties of solids is the size of the band gap between the valance and conduction bands. In a metal the gap can be non-existent with the conduction band overlapping the valence band. This ensures that there will always be electrons in the conduction band. At the other end of the spectrum, the band gap is very large in good insulators. A very large voltage is needed to excite an electron into the conduction band. Between these two limits are the semiconductor materials. In these solids the band gap exists, but is small enough that electrons can be promoted to the conduction band by an external stimulus. This is why their conductivity can be controlled.

As an electron in the conduction band moves through a solid there are many interactions which can affect its momentum. Simply put, the more an electron scatters or interacts, the lower its average velocity will be. Most commonly the electrons see a scattering potential field from the charged ion cores of the lattice. In these types of scattering events, the electron momentum can be transferred to a quanta of lattice vibrations called a phonon. Another dominant scattering process is from the ionized impurity elements introduced to change the number of electrons available. Most relevant to this work, electrons can also scatter off other electrons. Since this is generally considered to be an elastic process the electron energy is conserved, and the velocity is randomized. This will be important later when the plasma wave is discussed. One way to quantify how much scattering is occurring in a material is by its mobility. $\mu = \dfrac{e_0\tau}{m^*}$, where e_o is the unsigned unit electron charge, τ is the average time between scattering events, and m^* is the electron's effective mass in the material.

When an electric field is applied along a conducting channel, the force will cause the charge carriers to accelerate. The net force from the applied field and the randomization of the velocity from scattering will result in a steady state current to flow through the material. This drift current

($\mathbf{J}_{drift}$) is a product of the electron mobility (μ), the electron number density (n), and the force on the electron charges in an electric field ($e_0\mathbf{E}$), $\mathbf{J}_{drift} = e_0\mu n\mathbf{E}$ [18].

3.2 Formation and Properties of 2DEGs

For the present purposes, the direct problem with the normal conduction channel created by a gate voltage on a semiconductor is that the electron density is not high enough. Applying a higher voltage to collect more electrons, would exceed the breakdown voltage of the gate dielectric before the desired density is reached. The other way to get higher densities is to use a higher doping concentration, but this greatly decreases the mobility through impurity scattering. To get the higher densities needed, a two-dimensional electron gas (2DEG) which achieved high densities through electrostatic confinement in the third dimension will be utilized.

A 2DEG is a collection of electrons in the conduction band that is confined to a very thin surface in the semiconductor by a strong potential. This confining potential is created by manipulating the band structure in some fashion as shown in Figure 3.2. The important quality of this potential well is that it is deep enough to create quantum energy levels that confine the electrons into a sheet at the interface. In practical terms, the depth of the narrow potential well must be extremely large compared with the thermal energy at practical operating temperature, or K_BT where K_B is the Boltzmann constant and T is the lattice temperature. For strong confining potentials, e.g, $>> 3K_BT$ deep ($K_BT \sim 25$ meV at room temperature), the 2DEG is effectively confined quantum mechanically to two dimensions of free motion. As alluded, two of the important properties of the 2DEG we will later try to exploit are the increased electron mobility, and the higher electron densities compared to regular semiconductor channels with three-dimensional diffusive scattering.

For this work we have chosen a gallium nitride (GaN) based 2DEG (this is explained later in detail). In group III-nitride based systems there is a spontaneous polarization charge due to the bond arrangements in the wurtzite crystal. A second larger band gap III-nitride is used as the insulating material to form a barrier above the 2DEG, and the resulting potential notch is due to the difference in polarization electric fields in the two materials that can be represented by an

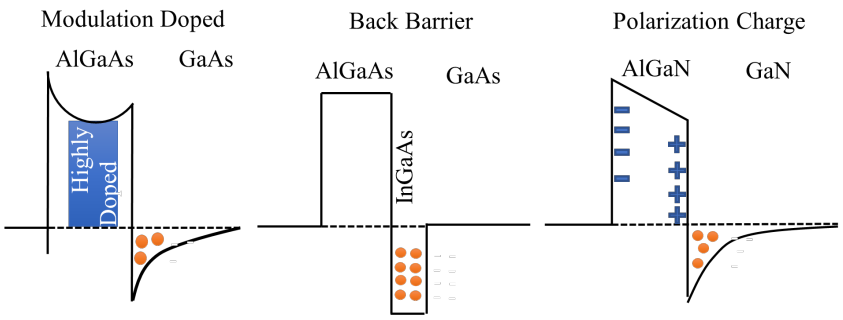

Figure 3.2: Conceptual diagrams of band structures that produce a 2DEG

interface sheet of polarization charges. It is common in GaN HEMTs to additionally produce a piezoelectric response with a strained AlGaN barrier that further increases the confining potential. The piezoelectric contribution is a natural result of the barrier material being epitaxial to the GaN lattice, thus inducing pseudomorphic strain fields.

One of the most exploited properties of the 2DEG is its reduced scattering rate which leads to an increase in the electron mobility. The reduction of scattering for all 2DEGs is mainly due to the inability to scatter out of the 2DEG plane. This effectively cuts the normal scattering rate by a third through phase space blocking of final states. In some systems there is also a reduction of scattering because there are few defects in the 2D channel compared to a doped 3D channel in a silicon based transistor. It is this property that forms the basis of the high electron mobility transistors (HEMTs) which are the active devices around which almost all GaN electronics are implemented.

3.3 Plasma Waves in 2DEGs

The electrons in a 2DEG can be thought of as a plasma just like the electrons in a metal. As such, their collective motion can be described as a plasma wave. We are specifically interested in what is sometimes called the Langmuir wave in the plasma community. This wave is a result of a restoring electric field being created by a displaced charge from the equilibrium distribution. Because this electric field is in the direction of the propagation, and thus no magnetic field is generated, this is why it is known as an electrostatic wave [21]. The wave can also be considered an oscillation between the kinetic energy of a current and the potential energy of an increased electron density.

Thus, in other contexts these same waves can also be referred to as electrokinetic waves [22].

In an infinite 3D plasma of interacting free electrons it can be shown that the dielectric tensor only has components on the diagonal.

$$\epsilon_{ii} = 1 - \frac{\omega_p^2}{\omega^2} \tag{3.1}$$

where ω_p is defined as the plasma frequency given by

$$\omega_p^2 = \frac{ne_0^2}{\epsilon m}. \tag{3.2}$$

Considering that the electric field only has a component along the propagation direction, the plasma wave equation simplifies to the following:

$$\frac{\omega^2}{c^2}\epsilon(\omega)\mathbf{E} = 0. \tag{3.3}$$

Given that the field is non-zero, this implies that $\epsilon\left(\omega\right) = 0$, and therefore the Langmuir wave can only occur at the plasma frequency [23]. In a 2D system, the plasma dispersion changes. In section 4.4 it is shown that it no longer has a single defined plasma frequency but rather results in a dispersion of the following form:

$$\omega^2 = \frac{e_0^2 n_0^{2D}}{2\epsilon m^*} k_z. \tag{3.4}$$

3.4 Simulation of 2DEG Plasma Waves

One difficulty in designing a 2DEG plasma wave interaction is the present inability to simulate them in the context of a useful device. There are several methods that can simulate certain aspects of the problem, but none are able to provide a complete picture. The main difficulty is that these devices will have two physical mechanisms with different length scales that need to interact in a single simulation domain. The description of electromagnetic waves on the slow-wave circuit has a discretization length scale set by the Courant-Friedrichs-Lewy condition $C = a\Delta t/\Delta x \leq C_m ax$.

This sets a ratio of the discretization of the physical space and the time scale that prevents a signal from traveling past the nearest mesh point in a single time step [24]. In practice this favors a spatial grid around a small fraction of the wavelength or on the order of a μm at 1 THz. The electron transport on the other hand requires a small spatial grid to resolve important electron interactions. Ideally the confining potential that forms the 2DEG is resolved by the spatial grid to properly account for the collisional environment. The Debye length of the 2DEG should also be resolved to properly model the screened coulomb interaction that forms the plasma wave. Both of these semiconductor effects would require spatial grids on the order of a nm. Satisfying all of these conditions with a single discretization for a realistic device would require a large spatial grid and a small time step that would require excessive computational resources. Another approach to the discretization is to use two different spatial grids and time steps, but keeping those synchronized and numerically stable is also a challenge.

One family of tools to consider are the codes currently used to simulate HEMTs. Common commercial semiconductor device codes like Sentaurus [25] use fluid type transport models that don't include advanced material properties and scattering in a way that can capture the electron-electron interaction needed for the plasma wave. Even more advanced codes such as Fermi Kinetic Theory being developed by the Air Force Research Laboratory, that does include realistic material effects, doesn't time resolve the electron-electron scattering. Instead the scattering rates are calculated by integrating over all the electron energies [26]. Monte Carlo codes can include these scattering effects, and there have been simulations that show plasma wave behaviors [27][28] . However, because they require very fine spatial grids (usually with dimensions in Angstroms) these codes have difficulty simulating full transistors that are already small compared to the electrically large TWA-style devices with RF circuits. It is also difficult to incorporate EM waves in Monte Carlo codes due to the reasons outlined above so they typically capture only electrostatic field distributions through Poisson's equation.

From a plasma physics approach there are both fluid style codes and particle-in-cell (PIC) codes. The simulation of choice for this problem would be a PIC model. While these codes can handle

electron-electron collisions and full-wave EM adequately, they lack a natural way to reproduce the other effects involved in semiconductor transport such as band structure and phonon collisions. Simulations have been ongoing with the fluid style plasma code Vsim from Tech-X, and might provide a starting point for a device design [29][30]. However, the simulations represent something of an idealized situation. The boundaries on the 2DEG are discrete and perfectly reflecting, which is not realistic. The electrokinetic oscillation of the plasma waves is also presupposed as more of an EM boundary condition than a natural response of the collective electrons. While there is evidence to support this type of behavior, it cannot represent how it may interact with other effects in the solid. Therefore, designs for prototype devices and experimental test plans will need to take these differences into account. This is discussed more in section 7.2

CHAPTER 4

COUPLED-MODE THEORY OF A SEMICONDUCTOR 2DEG INTERACTION

4.1 Coupled-Mode Theory of TWAs

The basic operation of a TWA hinges on the phase velocity of an axial electric field component propagating on a SWS to be near the electron velocity. Thus the EM wave and a perturbed wave on the electron beam can interact to produce gain.

Pierce first derived an encompassing theory for VED TWAs that uses a system of equations describing the mutual interactions between the electron beam transport and an EM wave propagating on the SWS. This coupled mode theory uses these equations to form a dispersion relation for the modes existing in the TWA interaction region. Pierce theory, as it is often referred to, is a first-order (small-signal) linear response theory intended to describe the amplitude and phase velocity of the coupled modes as they propagate along the beam-wave interaction structure [31].

Sumi [12] and then Solymar [13] extended this derivation beyond VEDs to instead include the electron transport behavior in a semiconductor in place of the beam. The devices they and Gover describe attempt to operate in a direct solid-state analog to the VED devices. These formulations include a simplified form of the electron transport present in solid-state materials. There are some notable differences from the transport in the VED case. The effective electron mass is used to approximate the average electron interaction with the solid lattice as a quantum mechanical quasi-particle that obeys classical mechanics (also known as the semi-classical approximation). A thermal diffusion term is added that is sometimes represented as an average thermal velocity with random direction. Electron collisions, mostly with the lattice via phonons, are also included by an average collision time τ. These are rather basic approximations of these effects, and Solymar includes a discussion about their applicability.

In traditional Pierce theory, the space charge term is often used to describe higher order perturbations to the vacuum beam dispersion and takes the form of a constant QC term calculated

for a particular SWS [32]. Here we will replace the bulk semiconductor carriers with a dense 2DEG. Within this 2DEG a plasma oscillation more accurately describes the AC carrier motion rather than the traditional space charge wave of VEDs. It is the 2D analog of solid-state plasmons [33] in 3D systems. The 2DEG plasma wave can be described as oscillations between electrostatic potential energy among the electrons in higher density regions and the AC kinetic energy of electrons in higher current regions, which is similar to the space charge wave. There is some evidence to suggest that the plasma wave and the space charge wave are the same wave in opposite limits of electron density [34], but for this discussion it is preferred to differentiate between them. The important difference is that the space charge wave phase velocity is near that of the DC electron velocity, whereas in the 2DEG plasma wave the phase velocity of the collective oscillation can be very different to that of the individual electrons. Further discussion of the relationship of these waves is certainly warranted, but beyond the present scope.

To include this plasma wave mode to the traditional Pierce theory, the space-charge contribution is not assumed to be a static perturbation, but instead is the main mode of coupling via the electrostatic field of an oscillating sheet charge. This relationship for the coupled wave can then be used to predict the small-signal behavior of the amplifier. This style of calculation using the VED space-charge term to predict gain has been shown in [35] to be equivalent to the traditional use of Pierce theory. In general, this implies that the EM wave phase velocity does not necessarily need to be near the DC electron velocity. However, we will show why there are still some practical limits to this in the operation of a TWA-type device.

4.2 Electron Transport Equation

One part of the coupled system is to describe the motion of the electrons under the influence of an external force. This starts with the first two moments of Boltzmann Equation (BE) as are typically used in semiconductor transport simulations within linear response theory. The derivation of moments starting from the distribution function in phase space $f(\mathbf{r}, \mathbf{k}, t)$ has been published many times and can be found in [20], so it is not reproduced in full here. The salient feature of

working in the moment expansion is the connection of the physical quantities density and velocity are approximated by their averaged values $\langle n \rangle$ and $\langle \mathbf{v} \rangle$ through

$$n(\mathbf{r}, t) = \int f(\mathbf{r}, \mathbf{k}, t) d\mathbf{k} \tag{4.1}$$

and

$$\langle \mathbf{v} \rangle = \frac{\int \mathbf{v} f(\mathbf{r}, \mathbf{k}, t) d\mathbf{k}}{\int f(\mathbf{r}, \mathbf{k}, t) d\mathbf{k}} = \frac{\mathbf{j}}{n(\mathbf{r})}. \tag{4.2}$$

The first two moments obtained by multiplying the Boltzmann transport equation by the zeroth and first powers of momentum and integrating over all momentum space can be written as

$$\frac{\partial}{\partial t} n + \nabla \cdot \mathbf{j} = 0 \tag{4.3}$$

and

$$\frac{\partial}{\partial t} \mathbf{j} + \nabla \cdot \mathbf{W} - n \left\langle \frac{\mathbf{F}}{m^*} \right\rangle = -\frac{1}{\tau} \mathbf{j}. \tag{4.4}$$

Here, n is the number density inside of a differential volume, and $\mathbf{j}$ is the particle flux across the surface defining that volume. $\mathbf{W}$ is defined as $W_{ij} \equiv n \langle v_i v_j \rangle = n \langle v_i \rangle \langle v_j \rangle + n \langle \delta v_i \delta v_j \rangle$, and m^* is the effective mass of the electron. For this small-signal case we will assume the mass is valuated at the bottom of the lowest conduction band.

As shown in [20] the second term on the left of Eq. (4.4) can be expanded as the following

$$\nabla \cdot \mathbf{W} = \nabla \cdot (n \langle \mathbf{v} \rangle \langle \mathbf{v} \rangle) + \nabla n \cdot \mathbf{w} + n \nabla \cdot \mathbf{w} \tag{4.5}$$

Using $\mathbf{F} = -e_0 \mathbf{E}$ (no magnetic fields), and the previous expansion bring the first moment equation to the form

$$\frac{\partial}{\partial t} \mathbf{j} + \nabla \cdot (n \langle \mathbf{v} \rangle \langle \mathbf{v} \rangle) + \nabla n \cdot \mathbf{w} + n \nabla \cdot \mathbf{w} - n \left\langle \frac{-e_0}{m^*} \mathbf{E} \right\rangle = -\frac{1}{\tau} \mathbf{j}. \tag{4.6}$$

Here it is assumed that $\nabla \cdot \mathbf{w} = 0$ based on diffusion being isotropic, and as stated above we define total particle current $\mathbf{j} = n \langle \mathbf{v} \rangle$. Using common definitions for mobility $\mu = \tau \frac{e_0}{m^*}$, and the diffusion coefficient tensor $D = \tau \mathbf{w}$ results in the form of the drift-diffusion equation within which we will apply wave perturbations:

$$\frac{\partial}{\partial t}\mathbf{j} + \frac{1}{\tau}\mathbf{j} + \nabla \cdot \left(\frac{\mathbf{jj}}{n}\right) + \frac{D}{\tau}\nabla n = -\frac{ne_0}{m^*}\left\langle \mathbf{E} \right\rangle . \tag{4.7}$$

Equations (4.3) and (4.7) are split into unperturbed (n_0, $\mathbf{j}_0$, $\mathbf{E}_0$), and perturbed (n_1, $\mathbf{j}_1$,$\mathbf{E}_1$) quantities with dependences described in the following equations:

$$\frac{\partial}{\partial t}n_0 + \nabla \cdot \mathbf{j}_0 = 0 \tag{4.8}$$

$$\frac{\partial}{\partial t}\mathbf{j}_0 + \frac{1}{\tau}\mathbf{j}_0 + \nabla \cdot \left(\frac{\mathbf{j}_0\mathbf{j}_0}{n_0}\right) + \frac{D}{\tau}\nabla n_0 = -\frac{n_0 e_0}{m^*}\left\langle \mathbf{E}_0 \right\rangle \tag{4.9}$$

and

$$\frac{\partial}{\partial t}n_1 + \nabla \cdot \mathbf{j}_1 = 0 \tag{4.10}$$

$$\frac{\partial}{\partial t}\mathbf{j}_1 + \frac{1}{\tau}\mathbf{j}_1 + \nabla \cdot \left(\frac{\mathbf{j}_1\mathbf{j}_0}{n_0} + \frac{\mathbf{j}_0\mathbf{j}_1}{n_0} - n_1\frac{\mathbf{j}_0\mathbf{j}_0}{n_0^2}\right) + \frac{D}{\tau}\nabla n_1 = -\frac{n_1 e_0}{m^*}\left\langle \mathbf{E}_0 \right\rangle - \frac{n_0 e_0}{m^*}\left\langle \mathbf{E}_1 \right\rangle . \tag{4.11}$$

The unperturbed quantities are in steady-state $\frac{\partial}{\partial t}n_0 = \frac{\partial}{\partial t}\mathbf{j}_0 = 0$, this means from Eq. (4.8) that $\nabla \cdot \mathbf{j}_1 = 0$. Eq. (4.9) simplifies to the drift-diffusion equation given by

$$\frac{1}{\tau}\mathbf{j}_0 + \frac{D}{\tau}\nabla n_0 = -\frac{n_0 e_0}{m^*}\left\langle \mathbf{E}_0 \right\rangle . \tag{4.12}$$

The third term in Eq. (4.11) can be expanded as follows:

$$\begin{aligned}
&\nabla \cdot \left(\frac{\mathbf{j}_1\mathbf{j}_0}{n_0} + \frac{\mathbf{j}_0\mathbf{j}_1}{n_0} - n_1\frac{\mathbf{j}_0\mathbf{j}_0}{n_0^2}\right) = \nabla \cdot (\mathbf{j}_1\mathbf{v}_0 + \mathbf{v}_0\mathbf{j}_1 - n_1\mathbf{v}_0\mathbf{v}_0) \\
&= \mathbf{j}_1 (\nabla \cdot \mathbf{v}_0) + \mathbf{v}_0 \cdot \nabla\mathbf{j}_1 + \mathbf{v}_0 (\nabla \cdot \mathbf{j}_1) + \mathbf{j}_1 \cdot \nabla\mathbf{v}_0 - n_1\nabla \cdot \mathbf{v}_0\mathbf{v}_0 - \mathbf{v}_0\mathbf{v}_0 \cdot \nabla n_1 \\
&= \mathbf{v}_0 \cdot \nabla\mathbf{j}_1 + \mathbf{v}_0 (\nabla \cdot \mathbf{j}_1) - \mathbf{v}_0\mathbf{v}_0 \cdot \nabla n_1.
\end{aligned} \tag{4.13}$$

This uses the expansion $\nabla \cdot (\mathbf{jv}) = (\nabla \cdot \mathbf{v})\mathbf{j} + (\mathbf{v} \cdot \nabla)\mathbf{j}$ found in [36]. The perturbed equations then can be written as

$$\frac{\partial}{\partial t}n_1 + \nabla \cdot \mathbf{j}_1 = 0 \tag{4.14}$$

and

$$\frac{\partial}{\partial t}\mathbf{j}_1 + \frac{1}{\tau}\mathbf{j}_1 + \mathbf{v}_0 \cdot \nabla\mathbf{j}_1 + \mathbf{v}_0 (\nabla \cdot \mathbf{j}_1) - \mathbf{v}_0\mathbf{v}_0 \cdot \nabla n_1 + \frac{D}{\tau}\nabla n_1 = -\frac{n_1 e_0}{m^*}\left\langle \mathbf{E}_0 \right\rangle - \frac{n_0 e_0}{m^*}\left\langle \mathbf{E}_1 \right\rangle . \tag{4.15}$$

The desired solution of this system would have propagating waves of the form $e^{i(\omega t-\mathbf{k}\cdot\mathbf{r})}$ and this wave will ultimately be oriented as propagation along the z-direction. Therefore, the time derivatives and spatial derivatives can be replaced with $i\omega$ and $-i\mathbf{k}$, respectively. Eq. (4.14) can then be rearranged to a useful form that is then substituted into Eq. (4.15).

$$n_1 = \frac{1}{\omega}\mathbf{k}\cdot\mathbf{j}_1 \tag{4.16}$$

$$i\omega\mathbf{j}_1 + \frac{1}{\tau}\mathbf{j}_1 + \mathbf{v}_0\cdot(-i\mathbf{k})\mathbf{j}_1 + \mathbf{v}_0\left((-i\mathbf{k})\cdot\mathbf{j}_1\right) - \mathbf{v}_0\mathbf{v}_0\cdot(-i\mathbf{k})n_1 + \frac{D}{\tau}(-i\mathbf{k})(\frac{1}{\omega}\mathbf{k}\cdot\mathbf{j}_1)$$
$$= -(\frac{1}{\omega}\mathbf{k}\cdot\mathbf{j}_1)\frac{e_0}{m^*}\left\langle\mathbf{E}_0\right\rangle - \frac{n_0e_0}{m^*}\left\langle\mathbf{E}_1\right\rangle \tag{4.17}$$

Dividing the unperturbed Eq. (4.12) by n_0, and assuming there is no spatial variation in the unperturbed density such that $\nabla n_0 = 0$ leads to an expression for $\langle\mathbf{E}_0\rangle$:

$$\left\langle\mathbf{E}_0\right\rangle = -\frac{m*}{\tau e_0}\mathbf{v}_0. \tag{4.18}$$

Substituting this result into (4.17), and simplifying results in

$$\left(\omega - i\frac{1}{\tau}\right)\mathbf{j}_1 - \mathbf{v}_0\cdot\mathbf{k}\mathbf{j}_1 + \left(i\frac{1}{\omega\tau} - 1\right)\mathbf{v}_0\left(\mathbf{k}\cdot\mathbf{j}_1\right) + \frac{1}{\omega}\mathbf{v}_0\mathbf{v}_0\cdot\mathbf{k}\mathbf{k}\cdot\mathbf{j}_1 - \frac{D}{\tau\omega}\mathbf{k}\mathbf{k}\cdot\mathbf{j}_1 = i\frac{n_0e_0}{m^*}\left\langle\mathbf{E}_1\right\rangle. \tag{4.19}$$

The electrons under consideration will be strongly confined to a single plane by the effective polarization sheet charge that created the potential well for the 2DEG. They can further be confined in another direction by etching the barrier sufficiently narrow to restrict their effective motion in all but one direction. The practical channel width will be discussed later. The transport equation can continue to be worked with only the z component with wave propagation of the form $e^{i(\omega t-k_z z)}$ which results in

$$\left[\omega^2 - i\frac{\omega}{\tau} - 2\omega v_{z0}k_z + i\frac{1}{\tau}v_{z0}k_z + v_{z0}^2k_z^2 - \frac{D}{\tau}k_z^2\right]j_{z1} = i\omega n_0^{2D}\frac{e_0}{m^*}\left\langle E_{z1}\right\rangle\ . \tag{4.20}$$

A slight rearrangement puts it into a form very similar to common VED TWA transport equations. Since this is a 2D system, it can it can be integrated over the perpendicular direction to reflect that. To make the meaning clearer the relation $\frac{D}{\tau} = \frac{K_B T}{m^*}$ is also substituted.

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2\right] j_{z1}^{2D} = i\omega n_0^{2D} \frac{e_0}{m^*} \langle E_{z1} \rangle \tag{4.21}$$

The first term within the brackets is the normal coupling to the DC beam velocity. The middle complex term is a loss term due to collisions. The last term in the brackets represents the thermal velocity. In plasma physics this can be referred to as the pressure term.

4.3 Slow Wave Structure Equation

The SWS is approximated as a lumped element model of a transmission line as Pierce has done. This might seem overly simplified at first, but [37] shows that a lumped element model is directly analogous to the field descriptions of actual RF waveguides. In reality, only a few basic properties of the EM structure need to be known, and in many cases those can be calculated directly from independent EM simulations of the circuit.

This part of the derivation starts with the discrete telegrapher's equations with an added impressed current from the beam. It is assumed that the impressed current is equal in magnitude to the perturbed 2DEG current. In this case, we will not assume the transmission line to be lossless and include the loss terms found in [38]. We are also guided by [38] in how to appropriately add the extra current such that the end result agrees with previous derivations [31][39].

The telegrapher equations can be found by starting at Kirchhoff's voltage and current laws applied to an infinitesimal section of a transmission line represented by lumped elements, as shown in Figure 4.1, which are given by

$$\begin{aligned} V(z,t) - R\Delta z\, I(z,t) - L\,\Delta z \frac{\partial I(z,t)}{\partial t} - V(z+\Delta z, t) &= 0 \\ \frac{V(z,t) - V(z+\Delta z, t)}{\Delta z} &= R\, I(z.t) + L \frac{\partial I(z,t)}{\partial z} \\ -\frac{\partial V(z,t)}{\partial z} &= R\, I(z,t) + L \frac{\partial I(z,t)}{\partial t} \end{aligned} \tag{4.22}$$

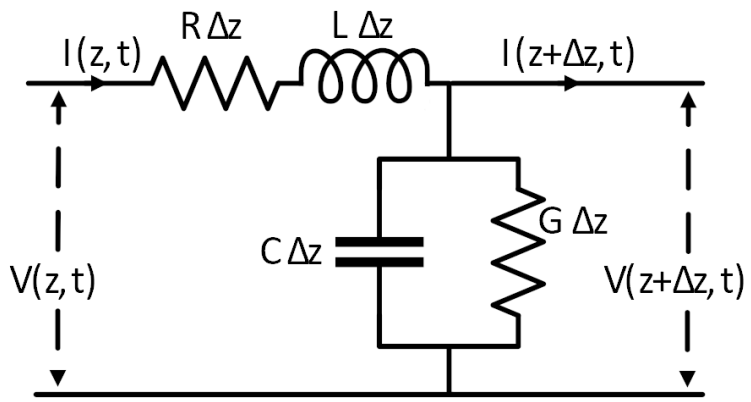

Figure 4.1: Lumped element model of a transmission line.

and

$$I(z,t) - G\Delta z V(z+\Delta z,t) - C\Delta z \frac{\partial V(z+\Delta z,t)}{\partial t} - I(z+\Delta z,t) = 0$$

$$-\frac{\partial I(z,t)}{\partial z} = GV(z,t) + C\frac{\partial V(z,t)}{\partial t}. \tag{4.23}$$

Letting I and V vary as $e^{i(\omega t - k_c z)}$ gives the traditional telegrapher's equations:

$$\frac{\partial V(z,t)}{\partial z} = -(R + i\omega L) I(z,t) \tag{4.24}$$

$$\frac{\partial I(z,t)}{\partial z} = -(G + i\omega C) V(z,t). \tag{4.25}$$

Subsequent differentiation and substitution gives

$$\frac{\partial^2 V(z,t)}{\partial z^2} = (-ik_c)^2 V(z,t) = (R + i\omega L)(G + i\omega C) V(z,t) \tag{4.26}$$

$$\frac{\partial^2 I(z,t)}{\partial z^2} = (-ik_c)^2 I(z,t) = (R + i\omega L)(G + i\omega C) I(z,t) \tag{4.27}$$

whereby wave propagation is characterized by

$$(-ik_c)^2 = (R + i\omega L)(G + i\omega C)$$

$$k_c = -i\sqrt{(R + i\omega L)(G + i\omega C)}. \tag{4.28}$$

The choice of sign for the square root is taken such that if $L = G = C = 0$, $k_c = -i\sqrt{R}$, and $e^{-ikz} = e^{-i\left(-i\sqrt{R}\right)z} = e^{-\sqrt{R}z}$. So in the limit of the circuit being purely resistive, the signal decays as it should. Substituting the assumed solution form into Eq. (4.24) results in

$$-ik_c V_0 = -\left(R + i\omega L\right) I_o \, . \tag{4.29}$$

The line impedance can be than be defined as

$$Z_c = \frac{V_0}{I_0} = \frac{\left(R + i\omega L\right)}{ik_c} = \frac{\left(R + i\omega L\right)}{\sqrt{\left(R + i\omega L\right)\left(G + i\omega C\right)}} = \frac{\sqrt{\left(R + i\omega L\right)}}{\sqrt{\left(G + i\omega C\right)}} \, . \tag{4.30}$$

These characteristic properties of the transmission line (k_c and Z_c)s generally depend on frequency. In later chapters, EM simulations will be used to calculate the properties for particular SWSs of interest and used in this coupled mode calculation.

To solve for the coupled mode, the beam current I_1 is added in Kirchhoff's current law for a node on the SWS. This approximation is that the current of the beam is assumed to be so close to the conductor in can be thought of as part of it (i.e. $I + I_1$ are inducing the voltage on the SWS). This is similar to multi-strand wire. This is also congruent to assuming that the beam induces a current in the wire as Pierce does by adding the induced current to the opposite side of the equation, and saying the induced current is the negative of the beam current. Within this framework, we have the following:

$$I\left(z,t\right) + I_1\left(z,t\right) - G\Delta z V\left(z + \Delta z, t\right) - C\Delta z \frac{\partial V\left(z + \Delta z, t\right)}{\partial t} - I\left(z + \Delta z, t\right) - I_1\left(z + \Delta z, t\right) = 0$$

and

$$-\frac{\partial I\left(z,t\right)}{\partial z} - \frac{\partial I_1\left(z,t\right)}{\partial z} = GV\left(z,t\right) + C\frac{\partial V\left(z,t\right)}{\partial t} \, . \tag{4.31}$$

Using this new current equation with Eq. (4.24), and letting quantities vary as $e^{i(\omega t - k_z z)}$ results in the following equations for the perturbed quantities. It is useful to note that this uses k_z since this is now the coupled mode in contrast the pure circuit mode above that uses k_c.

$$-ik_Z V_1 = -\left(R + i\omega L\right) I \tag{4.32}$$

and

$$-ik_z I = -\left(G + i\omega C\right) V_1 + ik_z I_1. \tag{4.33}$$

Substituting the former into the latter for I and rearranging leads to the result that

$$\left(k_z^2 + \left(R + i\omega L\right)\left(G + i\omega C\right)\right) V_1 = \; i\left(R + i\omega L\right) k_z I_1 \; . \tag{4.34}$$

From the previous definitions for our transmission line we have

$$\left(R + i\omega L\right)\left(G + i\omega C\right) = -k_c^2 \tag{4.35}$$

and

$$k_c Z_c = -i\sqrt{\left(R + i\omega L\right)\left(G + i\omega C\right)} \; \left[\frac{\sqrt{\left(R + i\omega L\right)}}{\sqrt{\left(G + i\omega C\right)}}\right] \; = -i\left(R + i\omega L\right) \tag{4.36}$$

which can be substituted into Eq. (4.34) to give

$$\begin{aligned} \left(k_z^2 - k_c^2\right) V_1 = \; & -k_c Z_c k_z I_1 \\ V_1 = \; & -\frac{k_c Z_c k_z}{\left(k_z^2 - k_c^2\right)} I_1. \end{aligned} \tag{4.37}$$

The relation to the electric field component can then be found by differentiation as

$$E_{z1,EM} = \; -\frac{\partial V_1}{\partial z} = \; \frac{k_c Z_c k_z}{\left(k_z^2 - k_c^2\right)}\left(-ik_z\right) I_1 = -i\frac{k_c Z_c k_z^2}{\left(k_z^2 - k_c^2\right)} I_1 \; . \tag{4.38}$$

To put the current into the terms of the transport equation, define $I_1 = -e_0 w_b j_{z1}^{2D}$. Since it uses the 2D number density, w_b is the width of the 2DEG beam with the result that

$$E_{z1,EM} = i\frac{k_c Z_c k_z^2}{\left(k_z^2 - k_c^2\right)} e_0 w_b j_{z1}^{2D} \; . \tag{4.39}$$

As mentioned, in this formulation transmission line loss is included which makes k_c a complex function of frequency. A more suitable form of this would be the commonly used

$$k_c = \frac{\omega}{v_c(\omega)} - i\alpha_c(\omega) = \frac{2\pi}{\lambda_c(f)} - i\alpha_c(\omega) \tag{4.40}$$

where $\omega = 2\pi f$ is the angular frequency of interest, $v_c = \lambda_c f$ is the phase velocity of a wave on the transmission line, and α_c is the loss coefficient of the transmission line. Using this loss coefficient

means that the field will decrease in magnitude as $e^{-\alpha_c z}$. To calculate α_c from EM simulations of a SWS it is useful to relate it to the EM power loss. From the Poynting vector, the square of the electric field integrated over a plane transverse to the waveguide is proportional to the power flowing through that plane. If the waveguide does not change in cross-section between two points along it, the integral of the fields will also be proportional at those two points.

$$\frac{\int \mathbf{E}^2(z_2) \cdot ds}{\int \mathbf{E}^2(z_1) \cdot ds} = \frac{P(z_2)}{P(z_1)} \tag{4.41}$$

$$\frac{E_0^2 e^{-2\alpha_c z_2}}{E_0^2 e^{-2\alpha_c z_1}} = \frac{P(z_2)}{P(z_1)} \tag{4.42}$$

Setting $z_1 = 0, z_2 = l, P(z_1) = P_{in}, P(z_2) = P_{in} - P_{loss}$

$$e^{-2\alpha_c l} = \frac{P_{in} - P_{loss}}{P_{in}} \tag{4.43}$$

$$\alpha_c = \frac{-1}{2l} \ln\left(1 - \frac{P_{loss}}{P_{in}}\right) \tag{4.44}$$

P_{loss} includes losses in the dielectrics, and finite conductivity losses in the metals used. P_{in} is the RF power that initially starts propagating on the SWS, and takes into account any mismatch reflections from launching the slow EM wave. From these definitions, $v_c(\omega)$ and $\alpha_c(\omega)$ for a particular SWS can be calculated for EM simulations as shown in section 5.3.

4.4 2DEG Plasma Wave Equation

This section derives a mathematical description of the electrostatic field generated by the perturbed 2DEG, which is the electron beam in the proposed interaction structure. This will take the place of the space-charge term in the traditional VED derivation. Let us consider the perturbed 2DEG sheet charge density in the $y - z$ plane: $\sigma_1(y, z) = \sigma_1' \, e^{-ik_z z}$. This assumes the perturbation is along the z-direction, or beam drift direction, and that the magnitude of that perturbation σ_1' is constant and independent of the transverse y-direction. From Poisson's equation, the resulting electric field from this charge distribution is

$$\mathbf{E}_1(\mathbf{r}') = \frac{1}{4\pi\epsilon} \int \frac{\rho}{r^2} dV \; \hat{r} \tag{4.45}$$

where $\hat{r} = \frac{\mathbf{r}}{r}$. Evaluating the electric field at a position $\mathbf{r}' = (x', 0, 0)$ above 2DEG at time t=0 produces

$$E_1(x', 0, 0) =$$
$$\frac{\sigma_1'}{4\pi\epsilon} \iiint_{-\infty}^{\infty} \delta(x) \frac{e^{-i(k_z z+\theta)}}{((x'-x)^2+y^2+z^2)^{\frac{3}{2}}} \left((x'-x)\hat{x} - y\hat{y} - z\hat{z}\right) dxdydz \quad (4.46)$$

where θ accounts for a phase difference along the z-direction between the observation point, and the density modulation.

$$E_1(x', 0, 0) = \frac{\sigma_1' e^{-i\theta}}{4\pi\epsilon} \int_{-\infty}^{\infty} e^{-ik_z z} \left\{ \int_{-\infty}^{\infty} \frac{(x'\hat{x} - y\hat{y} - z\hat{z})}{(x'^2+y^2+z^2)^{\frac{3}{2}}} \, dy \right\} dz \quad (4.47)$$

$$E_1(x', 0, 0) = \frac{\sigma_1' e^{-i\theta}}{4\pi\epsilon} \int_{-\infty}^{\infty} e^{-ik_z z} \left\{ \frac{2x'}{x'^2+z^2}\hat{x} - \frac{2z}{x'^2+z^2}\,\hat{z} \right\} dz \quad (4.48)$$

$$E_1(x', 0, 0) = \frac{\sigma_1' e^{-i\theta}}{2\pi\epsilon} \int_{-\infty}^{\infty} \frac{x' e^{-ik_z z}}{x'^2+z^2}\hat{x} - \frac{z e^{-ik_z z}}{x'^2+z^2}\,\hat{z}\, dz \quad (4.49)$$

$$E_1(x', 0, 0) = \frac{\sigma_1' e^{-i\theta}}{2\pi\epsilon} \left[\pi e^{-|k_z||x'|} \frac{x'}{|x'|}\hat{x} + i\pi e^{-|k_z||x'|} \frac{k_z}{|k_z|}\,\hat{z} \right] \quad (4.50)$$

$$E_1(x', 0, 0) = \frac{\sigma_1' e^{-i\theta}}{2\epsilon} e^{-|k_z||x'|} \left[sign(x')\,\hat{x} + i\; sign(k_z)\;\hat{z} \right] \quad (4.51)$$

This analytic solution of the integral does assume k_z to be real, but for this portion of the derivation k_z is explicitly real and the imaginary part is added in the full dispersion when collisions are added. We are only interested at this time in the z component, and we can take the limit as $x' \to 0$. The phase constant θ also needs to be equivalent to the original variation in z of the density, so

$$\sigma_1' e^{-i\theta} = \sigma_1'\, e^{-ik_z z} = \sigma_1 = -\frac{e_0 k_z}{\omega} j_{z1}^{2D} \quad (4.52)$$

$$E_{z1,2DEG} = -i\frac{e_0\; sign(k_z)\; k_z}{2\epsilon\omega} j_{z1}^{2D} \quad (4.53)$$

If this field term is then substituted into Eq. (4.21), the dispersion of the 2D plasma wave itself can be found. To do this, remove the drift velocity component $v_{z0} = 0$, take it in the collisionless regime such that $1/\tau = 0$, and remove thermal effects by setting $T = 0$.

$$\left[(\omega - kv_{z0})^2 + \frac{i}{\tau}(kv_{z0} - \omega) - \frac{K_B T}{m^*} k^2 \right] j_{z1}^{2D} = i\omega n_0^{2D} \frac{e_0}{m^*} \left\langle -i\frac{e_0\; sign(k)\; k}{2\epsilon\omega} j_{z1}^{2D} \right\rangle \quad (4.54)$$

For this specific case, the variable k is used instead of k_z because k_z specifically refers to the coupled mode solution. It should also be noted that k is strictly real for this derivation of the plasma wave dispersion. The imaginary part of k would be included if $1/\tau > 0$. In a VED TWA the corresponding space-charge term would result in a plasma frequency term ω_p that is a constant. In this case, the corresponding term is a function of k.

$$\omega_{p,2D}^2 = \frac{e_0^2 n_0^{2D}}{2\epsilon m^*} k \; sign(k) \tag{4.55}$$

This result matches the dispersion relation derived in [16]. Since the coupled-mode dispersion being derived uses k as a solution variable, $\omega_{p,2D}$ cannot be used in the coupled-mode dispersion. Instead, it is useful to reformulate the plasma wave dispersion into k_p as a function of ω. In this form ω is an input to the plasma dispersion just as it will be for the coupled-mode dispersion. This can then be simplified by using $v_{p,2D} \equiv \frac{\omega}{k_p}$, since again k_p is strictly real. The $sign(k)$ term also explicitly refers to the real part of k which will become important when in the coupled-mode dispersion k_z is complex.

$$\frac{\omega^2}{k_p} = \omega v_{p,2D} = \frac{e_0^2 n_0^{2D}}{2\epsilon m^*} sign(Re(k)) \tag{4.56}$$

4.5 Assembling the Coupled-Mode Dispersion Equation

The three key pieces of the system that are needed have been derived, and can now be assembled to give the coupled-mode solutions of the beam-wave interaction structure. They are repeated below to give a clear starting point for this derivation of a mode propagating with k_z in terms of the perturbed quantities indicated by the subscripts $z1$, the static beam velocity v_{z0}, and the 2DEG carrier density n_0^{2D}:

Transport Equation (4.21) confined to a 2D sheet:

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2\right] j_{z1}^{2D} = i\omega n_0^{2D} \frac{e_0}{m^*} \langle E_{z1} \rangle$$

Transmission Line (Circuit) Equation (4.39):

$$E_{z1,EM} = i \frac{k_c Z_c k_z^2}{(k_z^2 - k_c^2)} e_0 w_b j_{z1}^{2D}$$

2DEG Plasma Wave Equation (4.53):

$$E_{z1,2DEG} = -i\frac{e_0\ sign\,(k_z)\ k_z}{2\epsilon\omega} j_{z1}^{2D}$$

The electric field in the transport equation will be comprised of the superposition of the 2DEG plasma wave charge-induced field, and the EM circuit field. After substitution,

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2\right] j_{z1}^{2D} = i\omega n_0^{2D} \frac{e_0}{m^*} \langle E_{z1,EM} + E_{z1,2DEG} \rangle \tag{4.57}$$

which can be simplified in the following steps by

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2\right] j_{z1}^{2D}$$
$$= i\omega n_0^{2D} \frac{e_0}{m^*} \left[i\frac{k_c Z_c k_z^2}{(k_z^2 - k_c^2)} e_0 w_b j_{z1}^{2D} - i\frac{e_0\ sign(Re(k_z))\ k_z}{2\epsilon\omega} j_{z1}^{2D}\right]$$

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2\right]$$
$$= -\omega n_0^{2D} \frac{e_0}{m^*} \frac{k_c Z_c k_z^2}{(k_z^2 - k_c^2)} e_0 w_b + n_0^{2D} \frac{e_0}{m^*} \frac{e_0\ sign(Re(k_z))\ k_z}{2\epsilon}$$

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2 - \frac{e_0^2 n_0^{2D}}{2\epsilon m^*} sign(Re(k_z)))k_z\right] \left(k_z^2 - k_c^2\right)$$
$$+ n_0^{2D} \frac{e_0}{m^*} Z_c e_0 w_b k_c k_z^2 \omega = 0 \tag{4.58}$$

This is further simplified by introducing the coupling C^3 and phase velocity $v_{p,2D}$ defined in (4.56) to provide the final coupled mode dispersion:

$$C^3 = \frac{1}{2} n_0^{2D} \frac{e_0}{m^*} Z_c e_0 w_b \tag{4.59}$$

$$\left[(\omega - k_z v_{z0})^2 + \frac{i}{\tau}(k_z v_{z0} - \omega) - \frac{K_B T}{m^*} k_z^2 - \omega v_{p,2D} k_z\right] \left(k_z^2 - k_c^2\right) + 2C^3 k_c k_z^2 \omega = 0. \tag{4.60}$$

The roots of Eq. (4.60) are the coupled-mode dispersion relations of the beam-wave interaction system within Pierce theory.

CHAPTER 5

CREATING THE INTERACTION STRUCTURE

5.1 Overview of the Key Elements Required

The dispersion derived in the previous chapter intentionally uses abstract circuit equations and material properties to represent the SWS and the interaction region behaviors. This makes it easier to explore multiple configurations, as opposed to a full field solution that would need to be re-derived for each SWS type. The goal of this chapter is to provide more physical details of the intended interaction structure that can be included in the dispersion calculation. This not only increases the general accuracy of the results, but also provides context for whether or not a particular set of parameters is realistic. While these concepts are developed with a TWA in mind, they are also applicable to other styles of coupled-mode beam-wave style amplifiers.

The purpose of the interaction region in general is to put the EM wave to be amplified into close proximity to the beam such that they can strongly interact with each other. There are a variety of ways to create this interaction region, so they need to be quantified in some manner.

One of the most important goals of the interaction structure is guiding the slow EM wave to phase match with the beam charge perturbation. These two waves need to travel at approximately the same speed in order for there to be a meaningful transfer of energy, i.e. power gain. One of the main struggles of the solid-state TWA is indeed getting the EM wave slow enough for this interaction to be useful and realistic.

The strength of the interaction can be quantified using what Pierce calls the interaction impedance, defined in Eq. (5.3), that replaces the previous generic circuit definition of Z_c in the SWS field equation. Achieving an adequate coupling is certainly another main priority in amplifier design. The strength of the coupling in the final analysis is directly controlling the gain of the device, but it is also competing with losses from both the electron transport (collisions, thermal velocity, etc.) and the EM propagation (dielectric, metal, and radiation losses). In this chapter, the

EM loss will also be discussed, but the inclusion of transport loss and the determination of whether a parameter set will produce gain will be covered in the following chapter.

There is also the matter of being able to measure and use the device once it is fabricated. This primarily entails including well-matched input and output ports at each end of the SWS that can be connected in some fashion to test equipment. As will be shown, some SWS types are more amenable to this than others.

5.2 Semiconductor Material Properties and Choices

There are two different groups of materials that need to be considered in building the interaction structure. One type or group of materials will serve as the conductor and dielectrics of the SWS. The main interest in these is low loss and a somewhat suitable dielectric constant. These will be discussed in the context of the SWS in a following sections. Considered here are the other group of materials, semiconductor heterostructures, that are responsible for forming the 2DEG beam. The general concept of how 2DEGs are formed was discussed in Chapter 3, but now a suitable material system can be chosen with properties needed for a beam-wave amplifier. At present the focus will be on a TWA-style device with the EM wave synchronized to the plasma wave, and changes to these requirements will be discussed as the analysis progresses.

This research is focused on designing a power amplifier in the THz band and as such GaN has been chosen to provide the 2DEG. This is not to say other materials cannot be used to make amplifiers of this style, but the reasons outlined below indicate GaN being the best material for this application.

One crucial requirement is that the plasma wave phase velocity needs to be an appreciable fraction of the speed of light so that a matching SWS can be realistically constructed. From Eq. (4.56) the phase velocity of a 2DEG plasma wave is

$$v_{p,2D} = \frac{e_0^2 n_0^{2D}}{2\epsilon m^* \omega} sign(Re(k)) \tag{5.1}$$

As shown in [16] and [40], the phase velocity of a capacitive 2DEG due to a metal layer (such as a

gate) being nearby, is

$$v_{p,2D,C} = \frac{e_0^2 n_0^{2D}}{2\epsilon m^*} d_{cap} \tag{5.2}$$

where d_{cap} is the distance of the metal from the 2DEG. In order for a gate to maintain electrostatic control of a 2DEG, d_{cap} would have to be small, thus resulting in a slower phase velocity. For this reason, an ungated (normally on) 2DEG is chosen. This concept also happens to be the fundamental feature of depletion-mode GaN HEMTs which form the basis of nearly all modern GaN transistor amplifiers. Another key reason to use an ungated 2DEG is to remove the metal gate electrodes from the exact region where the circuit must interact with the beam.

Another key influence on the phase velocity is the 2DEG density. A GaN 2DEG number density of free electrons can be in excess of $2 \times 10^{13} cm^{-2}$, which equates to a phase velocity of around 1/10th the speed of light (c) in the THz band. Ungated InP and GaAs systems cannot support such large densities, and the resulting slower phase velocity would make the SWS design more difficult. While $c/10$ is slow compared to the phase velocity in VEDs, it is orders of magnitude faster than the average electron velocity in any semiconductor. As stated above, this is why the 2DEG plasma wave is the dominant charge effect desired in the 2DEG TWA concept.

Now that a suitable plasma wave can be formed, it needs to propagate far enough along the interaction structure to adequately couple with the SWS. For this to happen, the frequency of oscillation needs to be larger than the electron collision frequency. Thus it requires a material with a high electron mobility due to modest scattering rates. While InP and GaAs have a larger mobility than GaN (due in large part to their comparatively small electron effective masses), their lower densities would only work at lower frequencies where traditional VEDs are likely to perform better. It is this combination of high density and reasonably high mobility that makes the GaN 2DEG attractive for this device.

The ideal material system will also have an appropriate velocity saturation curve to achieve maximum gain and output power. As part of the previous remark on the collision frequency, the device must operate in the linear region of the electron velocity versus applied electric field curve. In most materials, this curve saturates when phonon collisions begin to be dominated by spontaneous

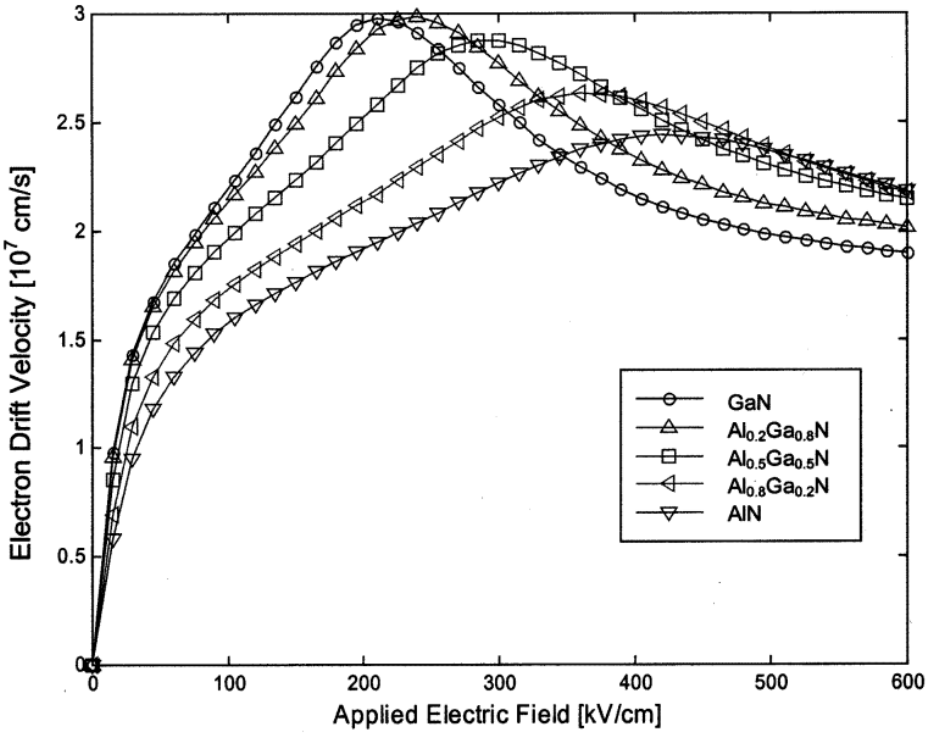

Figure 5.1: Velocity versus field curves for different semiconductors [3]. ©2001 IEEE

phonon emission scattering (energy transfer from the electrons to the lattice), and the associated collision frequency drastically rises. While staying in the linear region, it is desirable to operate the proposed device at a higher DC bias to increase the power available to produce gain. Fig 5.1 shows that GaN has a high saturation velocity and the linear operation region in excess of 200 kV/cm.

At these higher DC bias conditions, the 2DEG system will need to be capable of supporting the high current density that will be created. Typically the confinement of a 2DEG in the third dimension is about 2-4 nm, but with a large current they can spread out to behave more three dimensional and the mobility will drop with the increased scattering. GaN HEMTs have already shown support for current densities of $\sim 1A/mm$, and we believe we can achieve meaningful gain within that limit.

For this work commercially fabricated GaN on SiC wafers were procured from IQE. Leveraging their expertise, a heterostructure was design to produce a nominal 2DEG density of $2 \times 10^{13}\ \mathrm{cm}^{-2}$. The cross-section of the wafer stack is shown in Fig. 5.2. Table 5.1 shows the electron density

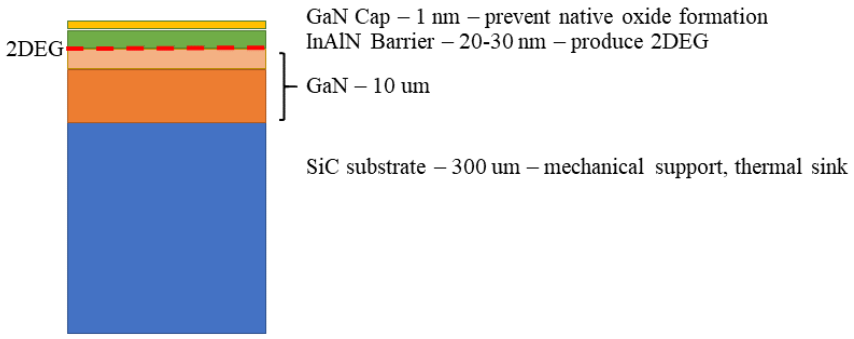

Figure 5.2: A diagram of the wafer stack procured for device fabrication.

Table 5.1: Measured Electron Density [$\times 10^{13}$ cm^{-2}]

Wafer ID	Avg	Max	Min	Std Dev
DH0143-06-EV	2.032	2.133	1.973	0.05180
DU0156-04-EV	2.019	2.067	1.973	0.03456
GP0247-02-EV	2.078	2.131	2.013	0.03784
HK0157-06-EV	2.055	2.111	1.983	0.04048
HK0157-16-EV	1.990	2.065	1.931	0.03948

for these wafers measured using a contactless Hall-effect Lehighton system confirming a highly uniform 2DEG.

5.3 Interaction Structure Details

In order to add more realistic detail to the dispersion calculation in the previous chapter, a few EM quantities can be calculated for a specific SWS configuration. In a way, this is similar to the extraction process that is done in RF circuit design. In both cases, scalar quantities as a function of frequency are generated from 3D geometry that are then used in lower dimensional calculations. The primary values to quantify for the SWS are the wavelength of the resulting axial field, the interaction impedance, and the EM power loss. More details of particular structures will be compared in a following section, here the quantities to compare are established. For all of the simulations in this section, the EM simulator HFSS which is a part of Ansys Electronics Desktop is used.

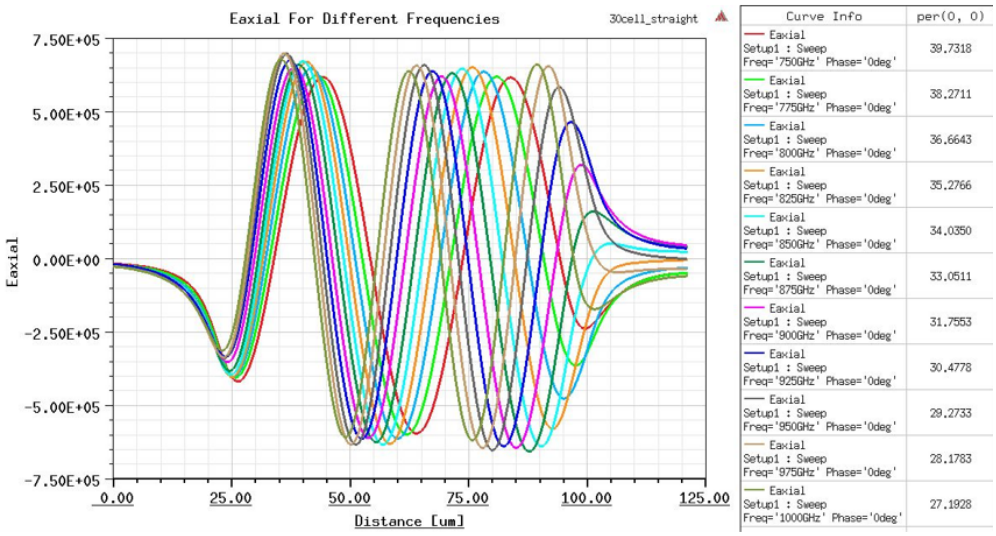

Figure 5.3: The axial field is plotted for a range of frequencies, and the wavelength is tabulated.

Wavelength of Slow EM Wave

The primary purpose of the SWS is to create an axial field that has a phase velocity slower than would normally exist on a typical waveguide. While HFSS could simulate a time domain velocity, it is easier (and perhaps cleaner) to determine the wavelength in a steady-state frequency domain simulation. The axial electric field is plotted for a range of frequencies in Fig. 5.3. HFSS can then internally calculate the period in space, i.e. the wavelength λ_c. This can then be included in k_c as shown in Eq. 4.40.

Interaction Impedance

In Section 4.3, the term Z_c is originally given as the impedance of the transmission line used in the SWS. In the perturbed case with the beam, this term can be adapted to quantify the strength of the interaction between the beam and the SWS. Pierce called it the interaction impedance [31]. It is defined as

$$Z_c = \frac{v_c^2 |E_{z1}|^2}{8\pi^2 f^2 P_{RF}}, \tag{5.3}$$

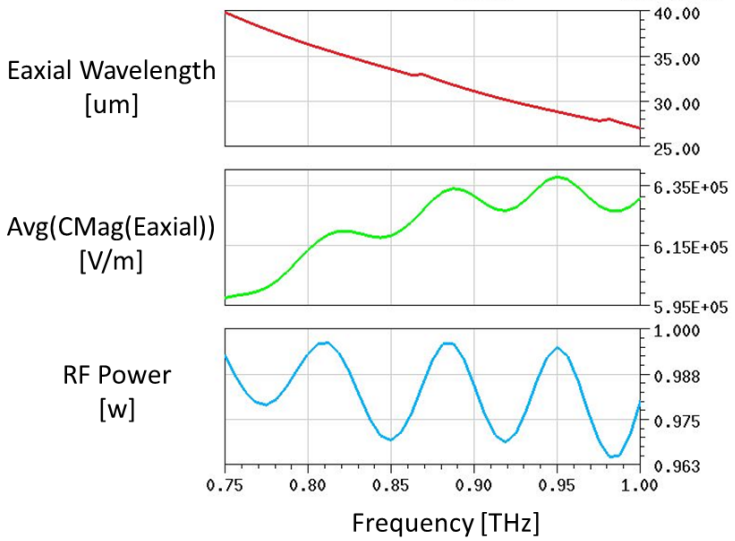

Figure 5.4: The quantities used to calculate the interaction impedance as a function of frequency.

where $|E_{z1}|$ is the magnitude of the axial electric field that interacts with the 2DEG, and v_c is the phase velocity of the axial field propagating down the SWS. P_{RF} is the RF power used to excite the SWS. The interaction impedance is then a measure of how much axial field the electron beam will experience per wavelength normalized to the power on the SWS. Each of these quantities used is calculated in HFSS as a function of frequency as shown in Fig. 5.4.

In most cases for a modal type simulation within HFSS, the input power is given as a way to normalize the fields. This input power could be used directly in Z_c, but it might not be very accurate. Using the input power ignores and possible mismatches in the way the SWS is fed. For this reason, the simulations are done using the best match possible. For instance, with the V-meander described in the next section the relative width of the microstrip lines that feed and form the SWS determines their impedance match as shown in Fig. 5.5. The power is also calculated using the Poynting vector on the transverse mid-plane of the SWS. As shown in Fig. 5.4, there is a slight variation in power, and this is used to more accurately normalize the field quantities also used in Z_c.

Extracting an accurate and meaningful axial field strength is the most difficult part of calculating Z_c. If care is not taken the quality of the mesh can play a role, but the mesh refinement process in HFSS generally does a sufficiently good job of this. The difficulty is that of impedance mismatches and wave reflections. Measuring the power as was described previously only represents an average

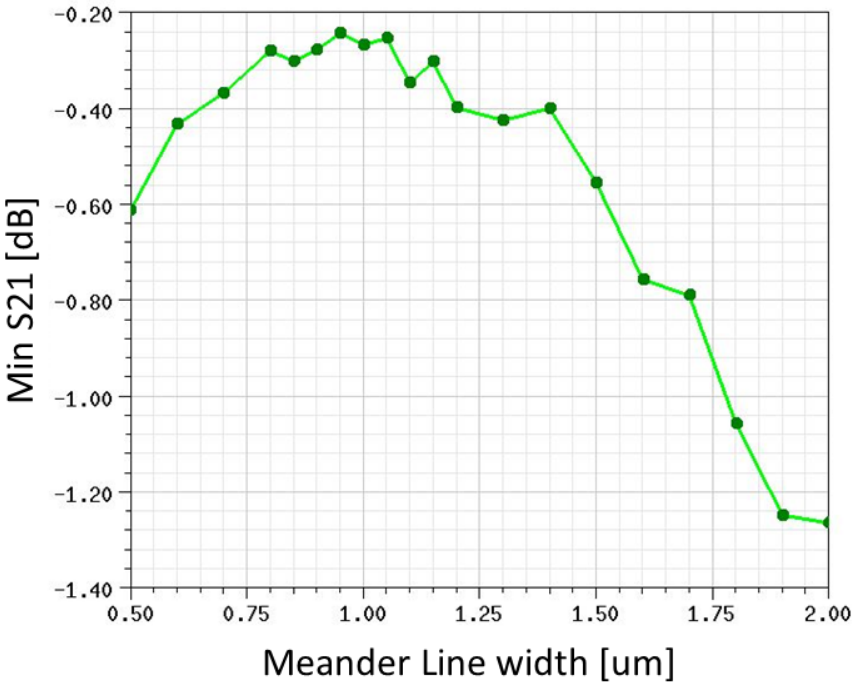

Figure 5.5: Parametric sweep of the line width for a V-meander to maximize the transmitted power and provide the best match.

effect of the mismatch. The fields themselves vary in space and phase within the SWS, so a way of averaging them is needed. More precisely, the average of the maximum field strength is desired. This quantity within HFSS is referred to as the complex magnitude of the field that is shown in Fig. 5.6. The complex magnitude of the quantity is simply the maximum field value at a position in space over all phases. Even though the mismatch was minimized, there is still some variation of the complex magnitude of the axial field due to reflected waves beating with the forward wave. The complex magnitude for each frequency is averaged over the length of the SWS to help smooth out these mismatch reflections.

EM Power Loss

In section 4.4 a complex propagation constant was used for the SWS to allow the inclusion of EM losses within the structure. To do this the imaginary loss coefficient α_c needs to be calculated for realistic geometries and material properties. HFSS does support finite conductivity metals and dielectric loss tangents within the EM simulation. The values used within these simulations are

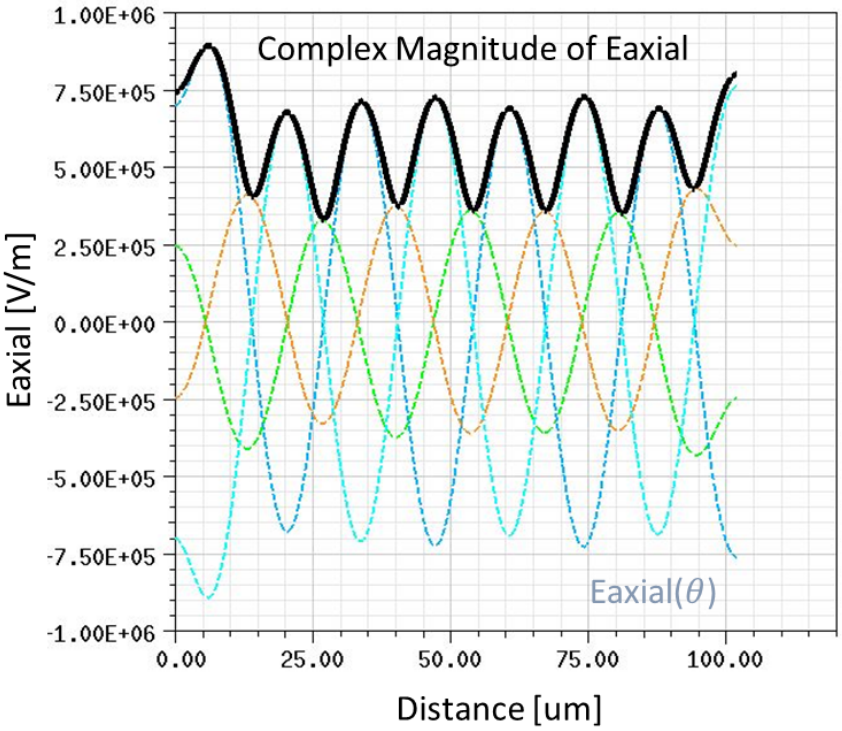

Figure 5.6: For a single frequency, the axial electric field for various phases is plotted with dashed lines. The complex magnitude of the axial field is shown and is the maximum field value over all phases.

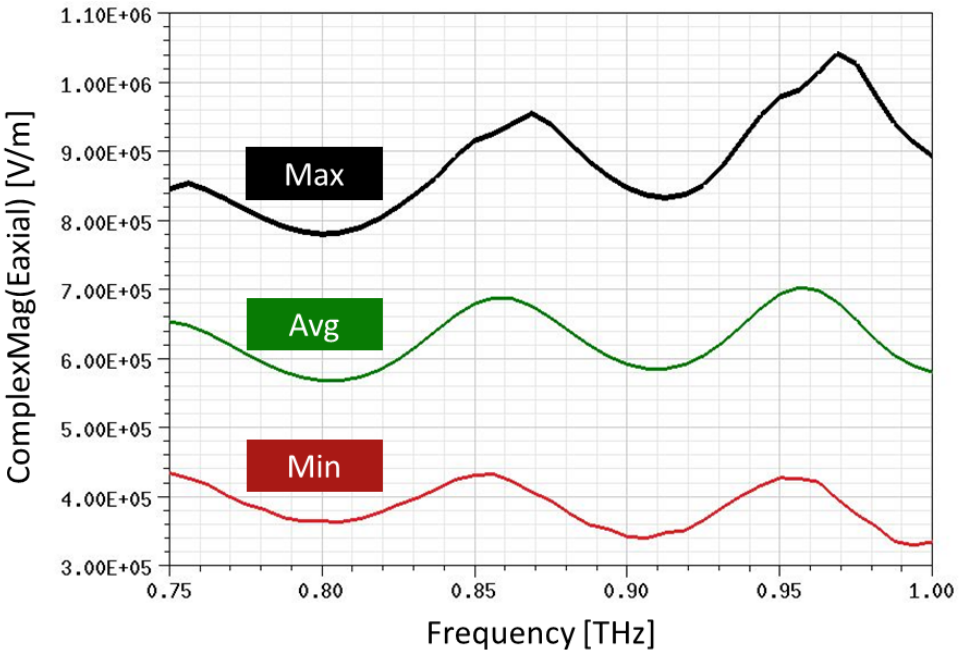

Figure 5.7: The maximum, average, and minimum values of the complex magnitude of the axial field is computed over the length of the interaction region for each frequency.

shown in Table 5.2. The direct method for finding α_c would be to fit a curve to the field decay, as this is what the term explicitly represents. However, it is more convenient to calculate it from the power loss of the SWS through Eq. (4.43). The P_{loss} should be calculated by an integral over the surface and volume loss densities, and P_{in} is the integral of the Poynting vector normal to a plane near the start of the SWS. These calculated values should be used instead of the s-parameters provided by HFSS because the loss can be calculated specifically for the SWS section and not have it include the feed structures or any mismatch losses. The loss within the feed lines is still important in the measurement context, but it should not be included in the dispersion calculation where it is in direct competition with gain production.

The exponential fit to the field decay shown in Fig. 5.8 gives a value for α_c of 46 cm^{-1}. The transmitted power loss for the same 180 $\mu\mathrm{m}$ SWS was 0.78 W, resulting in $\alpha_c = 42\ \mathrm{cm}^{-1}$. While the error in those numbers does include some simple numerical noise, there is actually a subtle difference between the two calculations. Measuring the dissipated power as a surface loss integral is at best a representation of the meaning of α_c that can be extracted from this type of simulation. The field decay, as well as calculating the Poynting power decay, will be close to the exponential loss but are a summation of all waves traveling on the SWS and not just the main forward wave of interest. Certainly any reflected wave will also go into the surface loss calculation, but the effect there should be much smaller.

While it is known that metal loss becomes significant in the THz band, there is not yet enough information to conclude whether it will overwhelm any gain present in the interaction. In the following section, more detailed simulations regarding loss are given for particular SWS geometries. In the following chapter, that will also be used to compute the overall gain of a structure.

HFSS Model Details

The aforementioned quantities have been calculated in two separate HFSS models for the same geometry. The axial wavelength, the axial field strength, and the RF power are calculated first in one lossless model, and then the α_c coefficient is calculated in a second model with material losses.

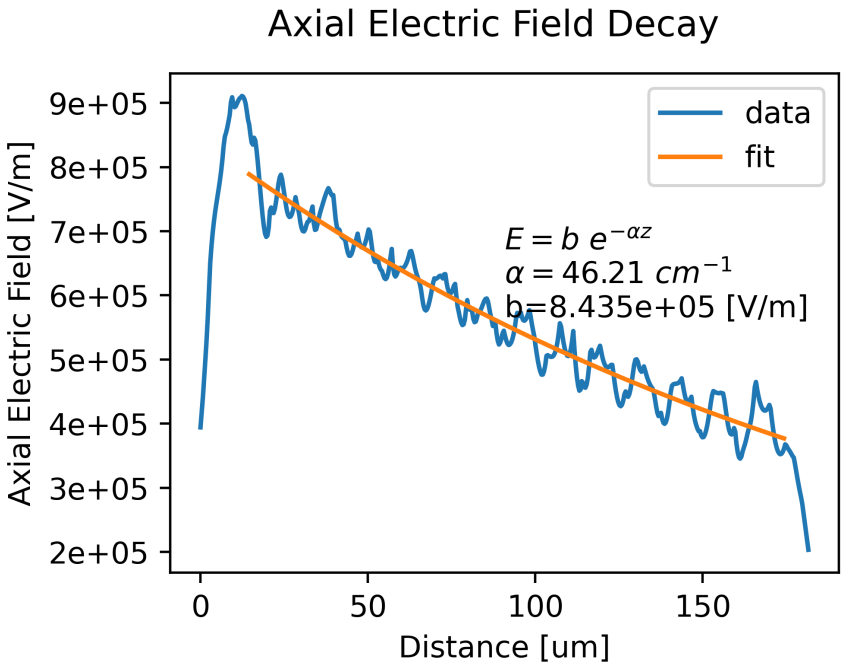

Figure 5.8: Axial field as a function of distance on a SWS with a conductivity of 5×10^8 S/m showing the field decay.

Table 5.2: Dielectric Material Properties

Material	Relative Dielectric	Loss Tangent	Ref.
Si	11.7	1.5e-5	[41]
GaN	9.6	5.2e-4	[42]
SiC	9.6	1.5e-3	[41]
SiO2	3.8	4.87e-3	[41]
BCB	2.45	1e-2	[43]

Table 5.3: High Frequency Metal Conductivity

Metal	Meas Freq [THz]	Conductivity [S/m]	Ref.
Gold	1.0	1.5e7-2e7	[44] [45]
Copper	0.85	2e7-5e7	[46]
Brass	0.6	1.5e6-6e6	[47]
Aluminum	1.0	3.8e7	[48]

Calculating the wavelength requires several periods of the axial field be present on the SWS at one time, and so it requires more cells of the SWS geometry. This also gives a larger region to average the field strength to minimize the impact of numerical discretization noise and mismatch beat patterns. When calculating the loss coefficient, a long SWS can dissipate almost all of the available power which will introduce noise into the data. Thus the second lossy model typically has a fewer number of SWS cells. Ideally, a length that dissipates somewhere between 30-70% of the input power would be best. It is also known that the metal loss accounts for at least 95% of the power dissipation in a given SWS, so often dielectric losses are not included.

5.4 Types of Slow Wave Structures

Now that the important quantities for a SWS have been defined, several different geometries can be explored and compared. All of the geometries presented here are possible to fabricate using current semiconductor fabrication techniques without stressing them unrealistically. For the most part these techniques favor stacked planer components with metal vias connecting them between layers. More details about the actually techniques and the physical limits they impose on the SWS geometry will be discussed in the following section on fabrication challenges.

Meander Microstrip

The primary SWS geometry that has been used within this study has been the microstrip meander line. This was previously published as a possible SWS for a VED device [49]. The microstrip meander is almost an ideal SWS for a solid-state beam-wave style amplifier due to it being completely planar and because the microstrip is established in RFIC design. Simply put, it is a microstrip waveguide that has connected lines going back and forth within the plane and transverse to the direction of propagation. If the lines are phased properly the fringe fields of successive lines couple to create a wave that travels much slower than modes of the normal microstrip. The meander line could be considered a planar analogy to the helix SWS used in VEDs. Figs. 5.9 and 5.10 show the basic geometry of the design. Because the microstrip lines were originally intended to be applied on top of

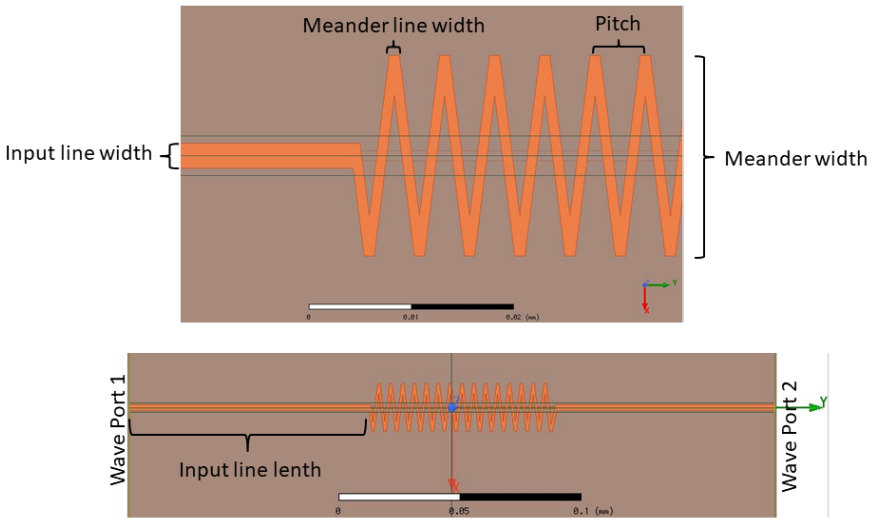

Figure 5.9: Top view of the microstrip meander highlighting some of the key geometry parameters.

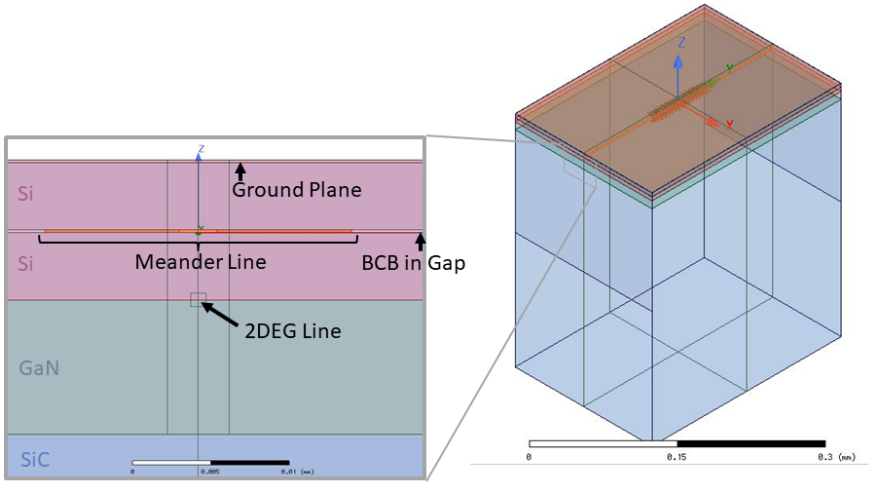

Figure 5.10: An end-view and isolated view of a microstrip meander showing the materials used in the simulation.

a GaN wafer, the microstrips are shown inverted with the ground plane on top of the material stack. In the section on fabrication recent techniques will be discussed that could alter this arrangement, but it generally does not change the fundamental analysis.

An example SWS dispersion is shown in Fig. 5.12 along with the dispersion of a 2DEG calculated using Eq. (4.56). The SWS dispersion was calculated using two simulation types within HFSS. The driven modal solution is as described previously using a single tone frequency domain solution with waveports. The eigenmode solution for the SWS uses a single cell of the meander,

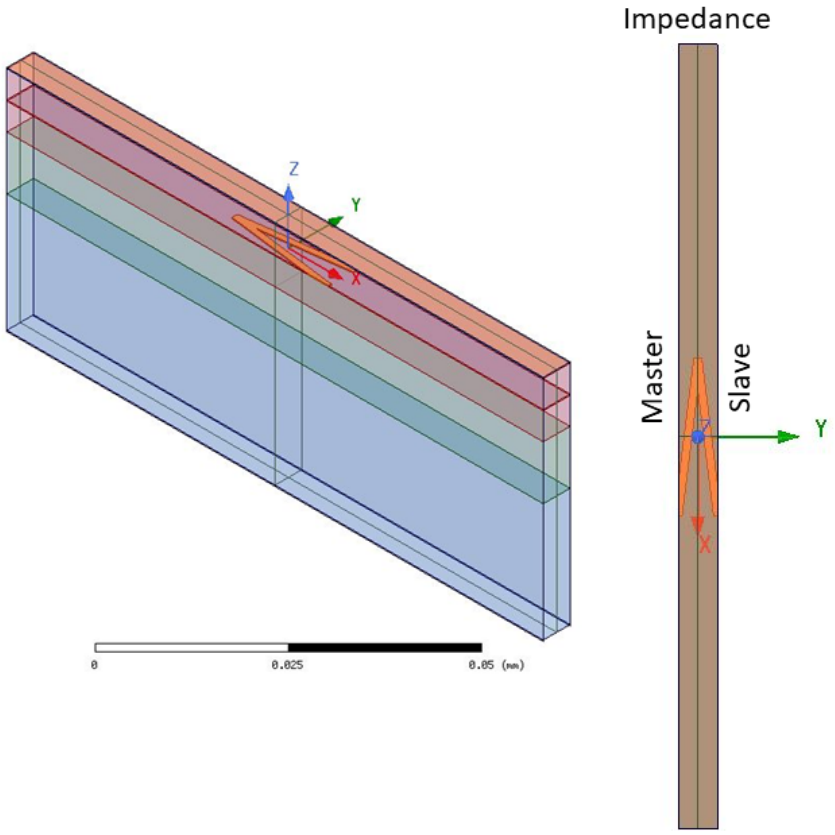

Figure 5.11: A single cell of a meander with periodic master/slave boundaries used for an eigen solution to find higher order modes in the SWS dispersion.

as shown in Fig. 5.11. For that solution, periodic master/slave boundaries set the wavelength of the mode through a prescribed phase shift, and the resulting mode frequencies are calculated. In the dispersion plot, the fundamental interaction range is marked, as well as two other possible interactions with the backwards wave, and a higher order mode. While these other interactions are not currently used in the derived theory, it is important to consider them for fabricated devices. The higher order modes of the forward-going wave can contribute to perturbations from the basic theory, and are sometimes included in an equivalent QC term for the SWS [32]. The point at which the beam dispersion intersects the backwards wave is generally more concerning. This interaction generally produces a resonant feedback oscillation that is sometimes used intentionally as a backwards wave oscillator (BWO).

In practice the microstrip line feeding the meander section is designed to be matched to 50 Ω.

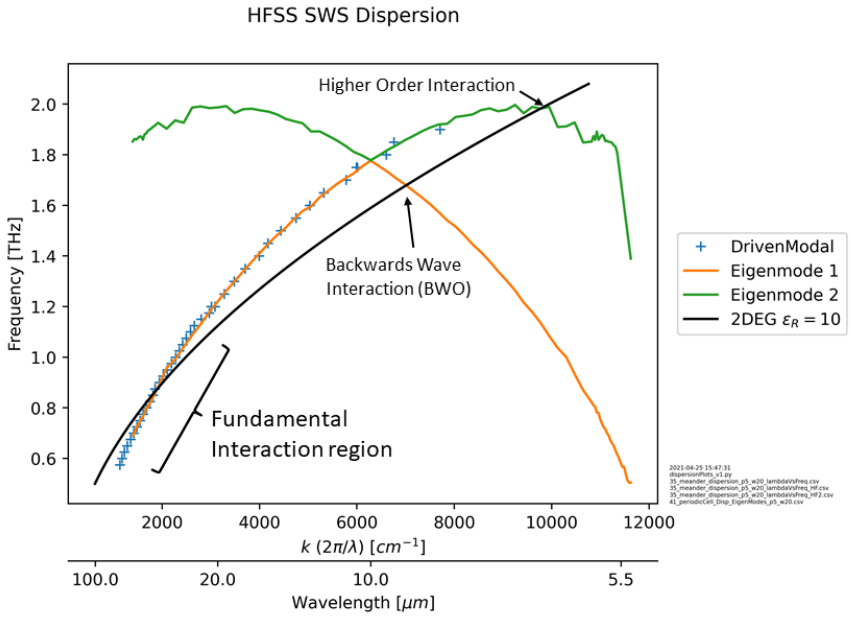

Figure 5.12: Simulated dispersion of a meander SWS with both driven and eigenmode solutions. A theoretical 2DEG dispersion is overlaid to show the region of operation.

The line width of the meander can then be adjusted to minimize an impedance mismatch as shown in Fig. 5.5. It is important to do this with no material losses in the EM model, as those will obscure any mismatches. Fig. 5.13 and Fig. 5.14 show that a reasonable match over the intended band of operation is feasible regardless of the length of the meander.

The largest drawback to the meander SWS is the metallic loss in the THz band. Fig. 5.15 shows the surface loss density, which is proportional to the current density, is greatest along the inner edges of the V-meander. The geometry of the meander can be altered in an attempt to reduce the current density in two main ways. The effect of using a thicker metal is shown in Fig. 5.16. It does reduce the loss by about 4% for doubling in thickness, however, this is not significant for the thicknesses that would be possible to fabricate. Another modification was to round out the inner corner of the V-meander with a fillet. This initially showed even larger reduction in loss with no fabrication difficulty as seen in Fig. 5.17. However, the fillet effectively shortens the path length of the meander, and thus reduces the axial field wavelength as shown in Fig. 5.18. When the width of the meander was made larger to compensate, the net loss returned to roughly the original value. This seems to imply that the sharp corner was not the main contribution to the metal loss, and that

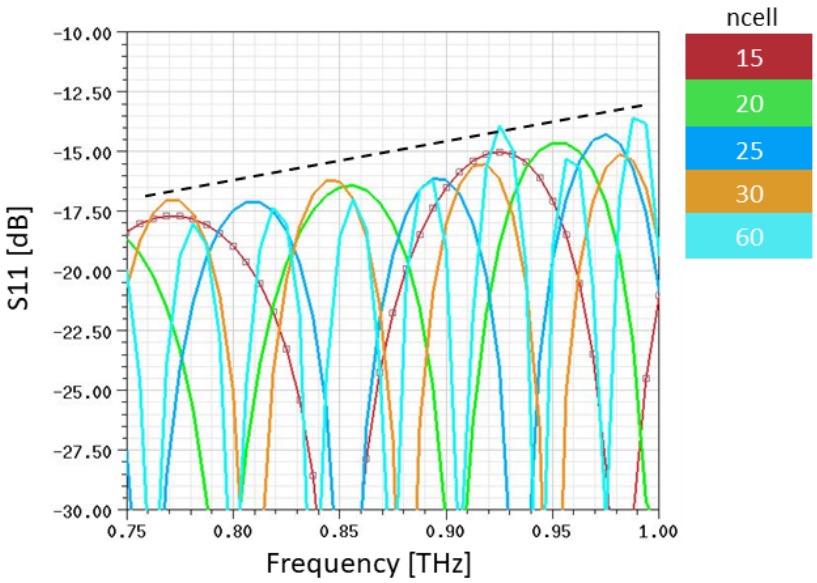

Figure 5.13: S11 of a microstrip meander with varying number of cells showing a reasonable match to minimize any resonant effects within the SWS.

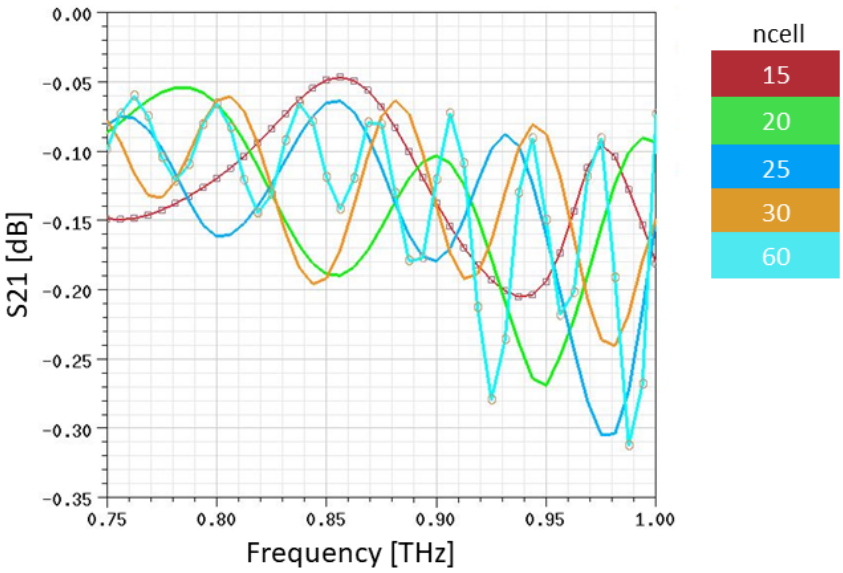

Figure 5.14: S21 of meander lines of different number of cells without material losses.

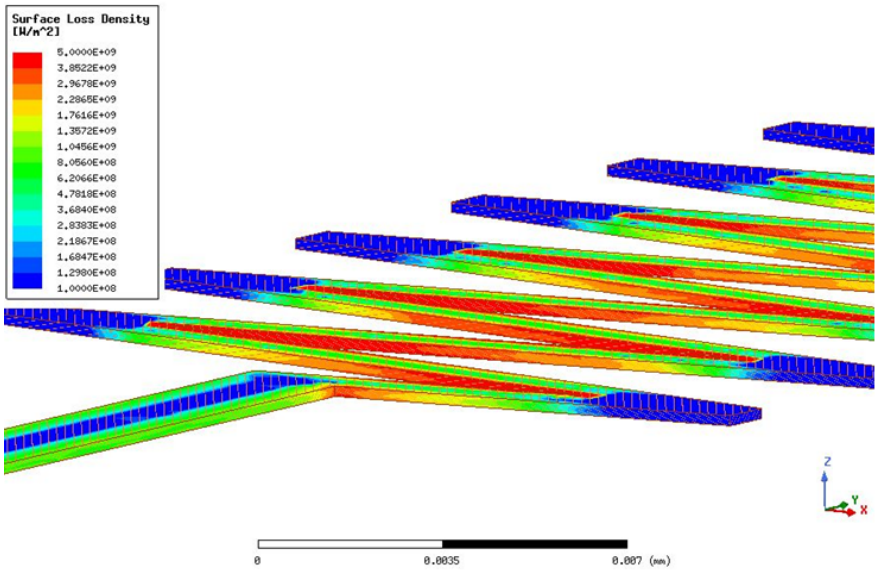

Figure 5.15: The surface loss density for a V-meander microstrip. Much of the field loss is at the edges of the metal lines that couple to the adjacent lines.

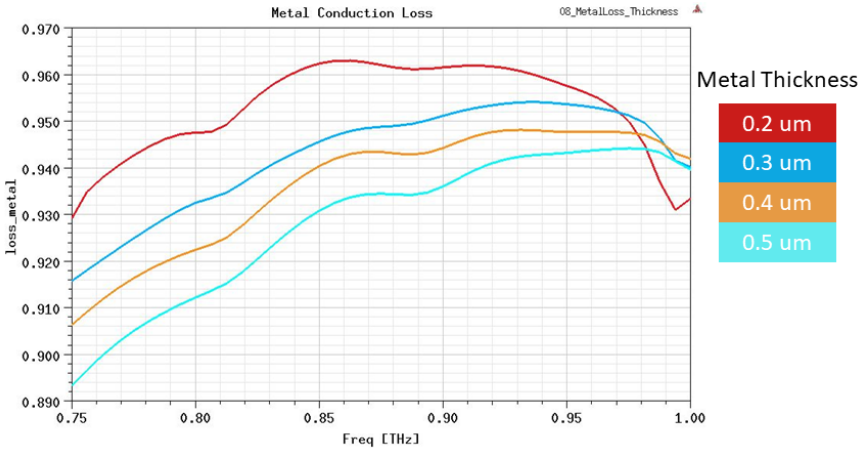

Figure 5.16: Metal surface loss for various metal thicknesses. Thick lines are unsurprisingly less lossy, but are harder to fabricate reliably.

the edges themselves being in close proximity might play a role. Based on this hypothesis, another variation called the U-meander was simulated shown in Fig. 5.19. This design has the central lines of the meander parallel to each other to separate the lines as much as possible while still maintaining the required pitch. This will spread out the loss on the metal surfaces. This improvement can be seen in Fig. 5.24. The U-meander also provides a more uniform field over the width of the 2DEG channel and less transverse fields. Both of these factors are desirable for optimizing performance of devices.

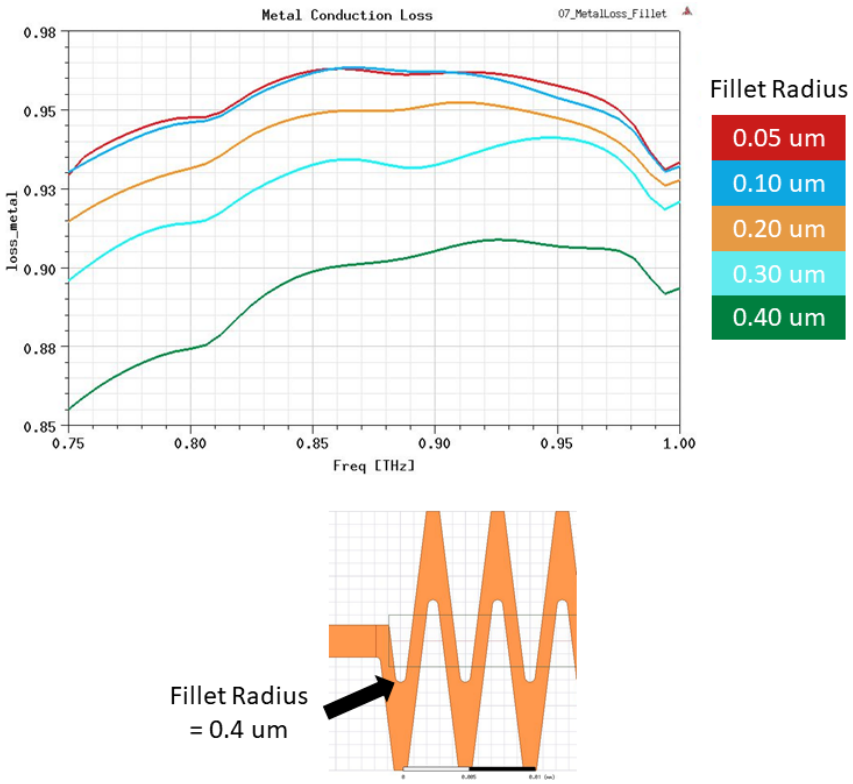

Figure 5.17: A fillet added to the inside edge of the meander to reduce current density at that location. The plot shows loss as a function of fillet radius.

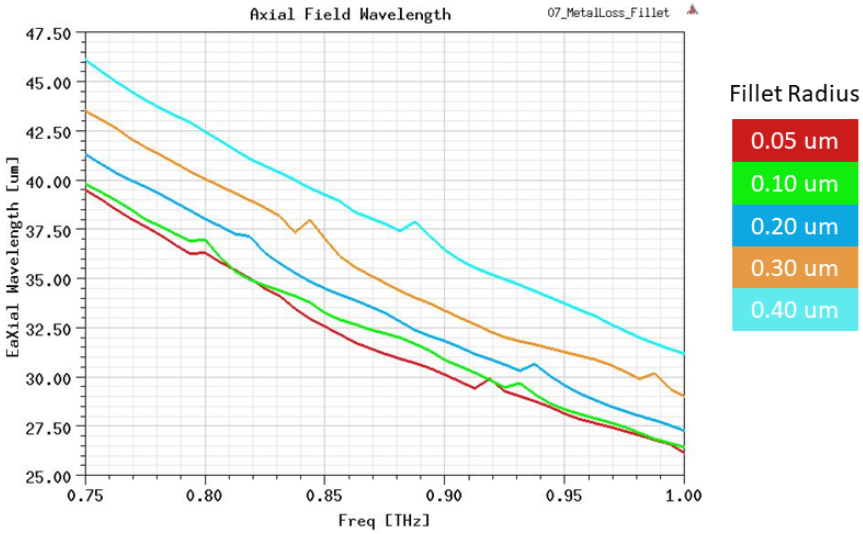

Figure 5.18: The fillet radius also effects the path length along the SWS, and therefore effects the axial field wavelength.

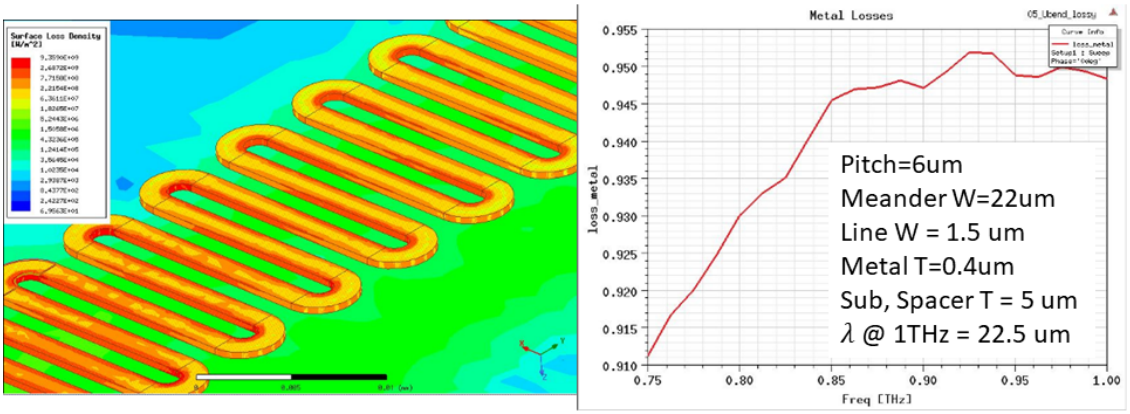

Figure 5.19: Another variation of the microstrip meander, called the U-meander, is aimed at reducing current density and thus the metal losses in the meander.

Rectangular Helix

The helix is one of the most common SWSs used in TWAs. A true cylindrical symmetric helix would be very challenging to fabricate on the desired scale, but an adaptation of it could be possible. Shown in Fig. 5.20 and Fig. 5.21 is a model of such a rectangular helix using planar metal traces and plated through vias. Since the metal traces are spread out slightly more, the metal losses shown in Fig. 5.23 are less than those of the meander SWS. The axial electric field pattern seen in Fig. 5.22 show how the field is distributed along the helix. The higher dielectric constant of Si was useful to manage the size of the meander line SWS, but with the rectangular helix, the transverse dimensions are a bit smaller than might be useful. Ideally the SWS metals are far enough away from the 2DEG such that they do not affect the dispersion of the plasma wave. In the example case, the Si rectangular helix has a height of 3 μm and a width of 9 μm to achieve the desired 24 μm axial field wavelength at 1 THz. This puts the metal traces only 1.5 μm away from the 2DEG. Another helix was modeled using quartz crystal as the dielectric material that allows more space between the metal and the 2DEG.

It should also quickly noted that the oscillating nature of the field strength, and thus Z_c, seen in Fig. 5.24 for both helix variations are due to mismatch reflections on the ports. A suitable coupling structure to the rectangular helix has yet to be designed, and these simulations are done with a rectangular wave port directly feeding the helix mode.

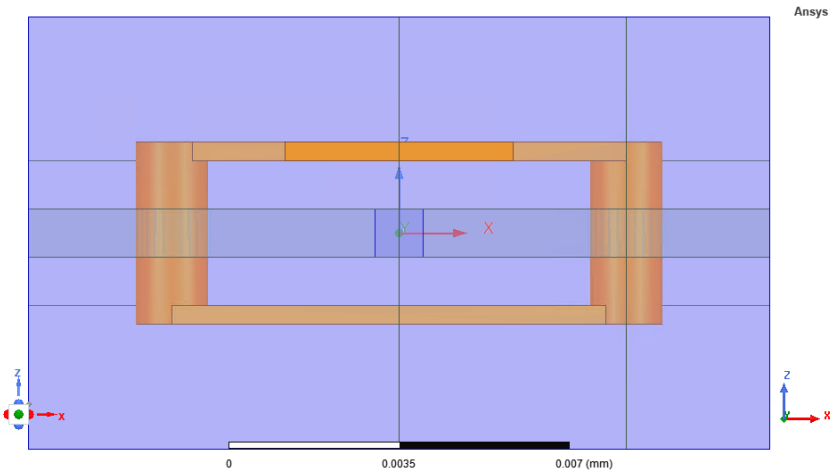

Figure 5.20: End view down the axis of a rectangular helix in Si with a central layer of GaN with the 2DEG.

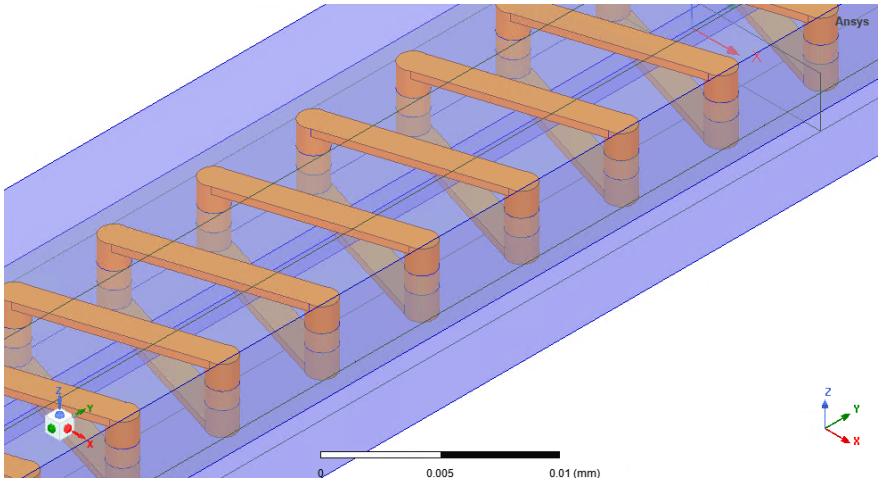

Figure 5.21: Isometric view of a rectangular helix in Si.

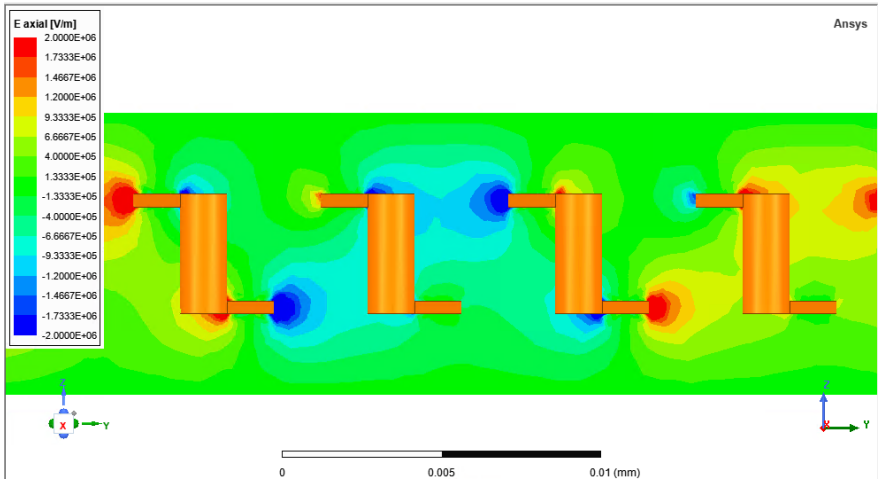

Figure 5.22: Axial electric field pattern at 1THz for the example rectangular helix.

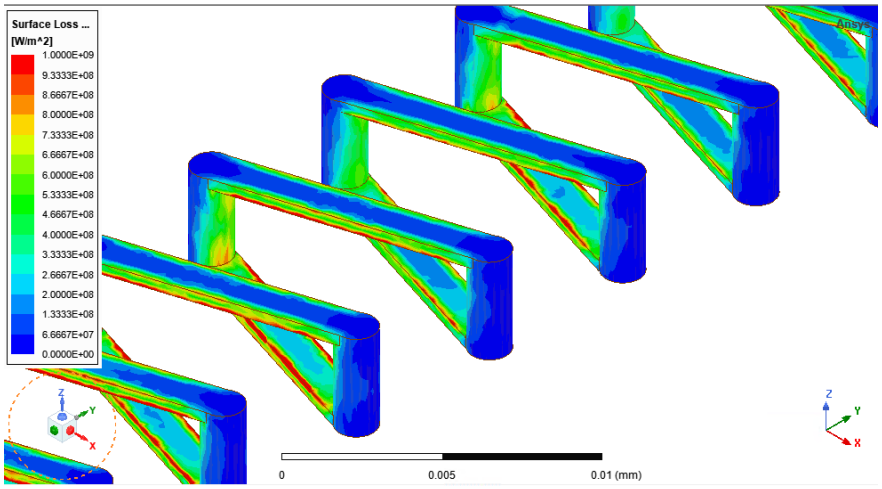

Figure 5.23: Surface loss density on the metal lines and vias in the rectangular helix.

Dielectric Waveguide Grating

Comparison of SWS Types

Now that a few different types of SWSs have been presented, a comparison of their performance can be made. Fig. 5.24 shows four of the calculated quantities as a function of frequency for 5 different configurations. An effort was made to make the comparison between types as fair as possible.

The first quantity of comparison is the normalized phase velocity. For each case, a nominal plasma wave velocity was calculated using Eq. (4.56). The relative permittivity used was an average of GaN and that of the substrate (7 for quartz and 10.5 for Si). Each SWS was tuned to approximately match the plasma velocity at 1 THz to help minimize the roll off of Z_c, and thus gain, at higher frequencies. It is desirable to have as flat of a dispersion as possible to increase the bandwidth over which the resulting device has gain. In that regard the U-meander appears to be the best.

There are a few observations to make about the axial field strength. The value indicated in the legend of Fig. 5.24 is the between the SWS metal traces and the 2DEG layer. As would be expected, the higher field values are closer to the metal, and so putting the 2DEG closer to them would increase their effect. The V-meander with a 1.5 μm spacing does have a very similar field strength as the Si helix with the same spacing. There are downsides to having the metal so close

to the 2DEG. One is that the field is much more likely to have higher harmonic content that could excite other plasma modes or simply reduce its efficiency. Another issue is that the presence of metal around the beam is well-known to perturb the solution in VED devices, and is expected to do the same here.

Another observation of the field strength can be made by comparing the 5 μm V-meander and U-meander. The geometry and design of these two meanders is very similar except for how their metal traces are arranged. Originally, the V-meander was thought to be superior since it seems to concentrate the higher electric field towards the center of the meander where the 2DEG would be placed. Even though the U-meander might not do that, it still consistently shows a higher field at the 2DEG location. Another advantage to the U-meander in this comparison is that the field will be much more consistent over the width of the 2DEG giving a much more well-behaved and predictable behavior.

The interaction impedance is a combination of the field strength and the phase velocity by Eq. (5.3). Since the quartz-based helix has a longer wavelength its interaction impedance is close to that of the similar Si helix, even though it has a lower field strength.

The loss coefficient α_c shows a large variation among the different geometries. It is expected that the two V-meander geometries would have the same loss since the waveguide itself does not change. The U-meander does indeed show a reduction of loss with the metal traces being uniformly spaced to reduce current crowding at the edges. The helix configurations additionally furthers this trend since in that geometry the nearest neighbor to a trace is a full pitch away compared to a half pitch in the meanders, not including the trace width.

Choosing the best SWS to use from this information is not entirely obvious or trivial to quantify. Ideally, a geometry with a high Z_c and a low α_c is desirable. By those merits alone, the two helix configurations would be the likely design choice. However, those geometries are harder to couple the RF signal into, and a design solution for that input coupling structure has not yet been attempted. As such, any plan to fabricate these devices would likely involve both a helix and the U-meander.

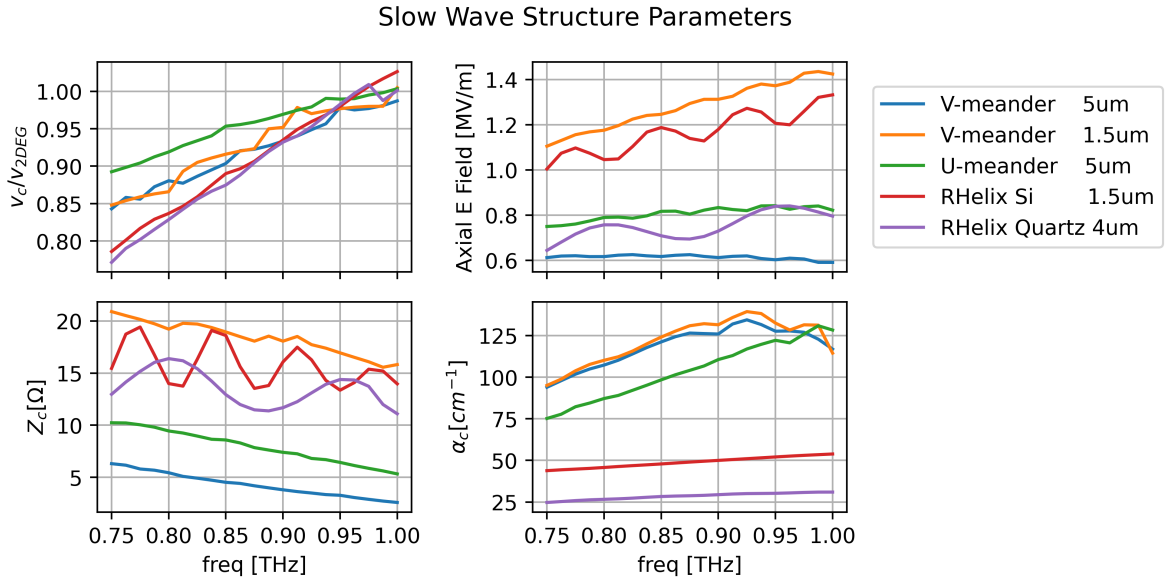

Figure 5.24: A comparison of the SWS properties. In this case, only the metal losses were included in the calculation of α_c as they are responsible for almost all the power dissipation. The field level is normalized to 1 watt of input power.

5.5 Fabricating the Interaction Region

While a detailed analysis of the device fabrication is beyond the present scope, a high level discussion is included as it will ultimately dictate what device can be realistically fabricated. Traditionally, semiconductor devices are mostly composed of different layers of materials being added to or subtracted from a stack (of materials) one layer at a time. Depending on the desired properties of the layers, they are most often either grown epitaxially or deposited in bulk on top of the stack. Material is removed by either chemically etching or mechanically abrading the material. These processes tend to be very good at adding or removing materials in very slow controlled ways, or in faster more coarse ways [50]. In several of the designs of the previous section, there are crystal silicon or quartz layers on the order of 5-10 μm that would be difficult to achieve with the common processes. It would take a long time to grow such a thick crystalline layer using epitaxy, faster deposition techniques would lead to an amorphous solids with higher RF losses, and the thickness tolerance for mechanically polishing a thicker layer is a sizable fraction of the desired value. The best way these layers can be created is to use a membrane transfer process developed by a collaborator as

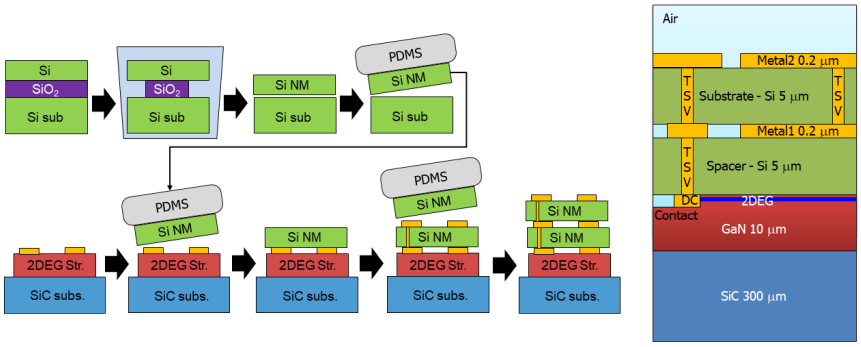

Figure 5.25: An example process for transferring Si layers to form the dielectric stack.[4]

seen in [51].

An example of this process for the inverted meander line SWS is shown in Fig. 5.25. The Si layers used for the waveguide dielectrics are taken from commercially produced silicon on insulator (SOI) wafers and are then applied to the device stack with metal traces in between. This process allows for a much more flexible stack fabrication for the different SWS configurations.

One of the important things that it allows is the possibility of moving the 2DEG off of the wafer on which it was grown. An example of this can be seen in [51] where the 2DEG conduction channel is shown to function normally even after being transferred. This will allow more complex SWS arrangements like the Si or quartz rectangular helix shown previously. The 2DEG layer can be transferred to a Si layer that forms the lower half of the helix interior, and the other Si half is applied on top of it. This could also become critical in future developments discussed in Chapter 7 where multiple SWSs might need to be constructed around a single 2DEG layer.

5.6 SWS Measurement Challenges

As with all simulations of experimental devices, there needs to be measurements to demonstrate that the model results agree with reality. In this case, there are a number of quantities and behaviors in the EM models that would need to be confirmed. These range from the basic measurement of metal loss and dielectric permittivity, to the more difficult confirmation of the SWS dispersion. An example of a full device ready for measurement is shown in Figs. 5.26 and 5.27. Here RF probe pads, DC bias pads, and a microstrip inversion are included for an on-wafer measurement. Given the

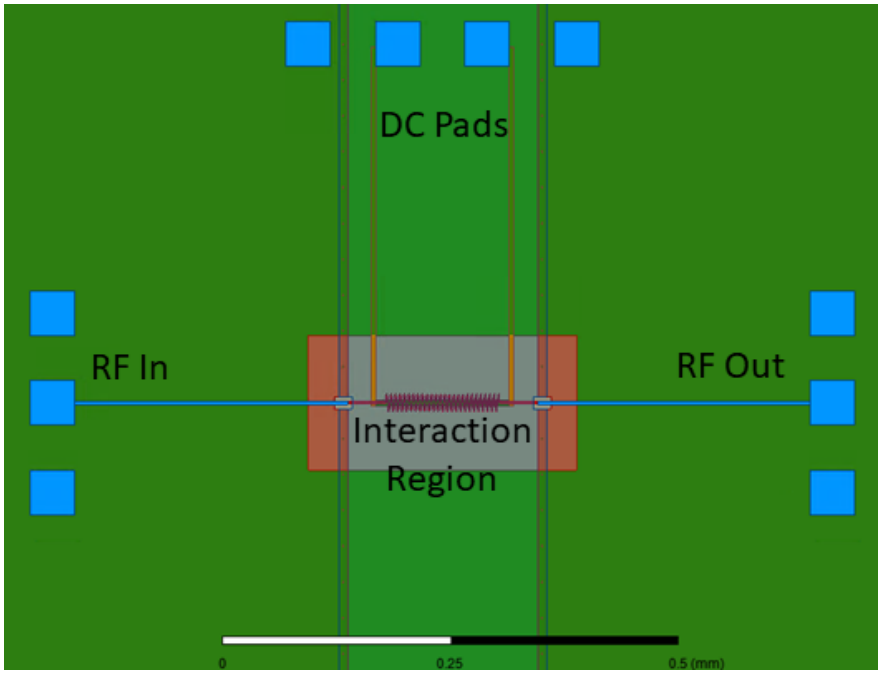

Figure 5.26: A top down view of a hypothetical test device. RF probe pads are on the east and west side, with DC bias lines for the 2DEG on the north side.

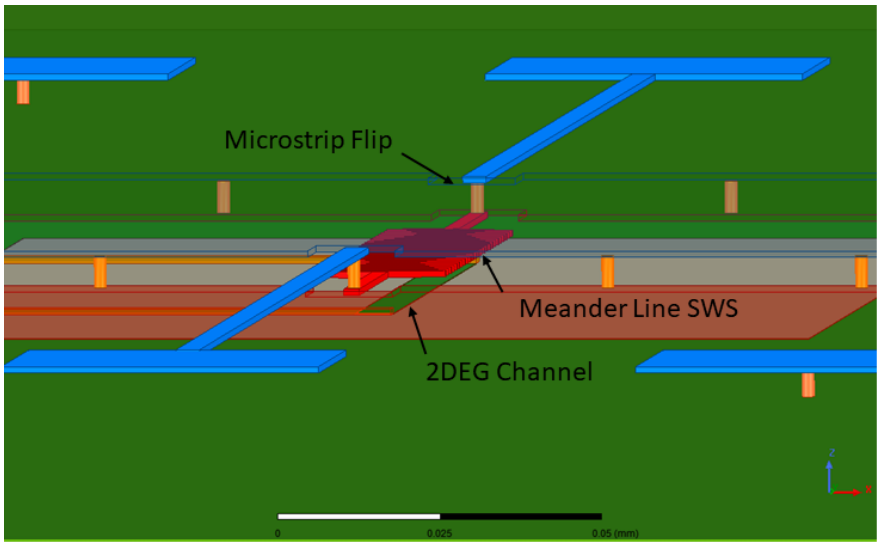

Figure 5.27: An oblique angle view of the test structure showing the microstrip flip, and the meander line above the 2DEG channel.

possibilities provided by the membrane transfer process in fabrication, several other configurations can also be considered with the same types of measurements.

To do these measurements, a set of VNA extender heads would be used to extend the frequency range of a VNA into the THz band. Extenders for the 0.75-1.1 THz band are available from Virginia

Table 5.4: HFSS simulation of insertion loss at 1 THz due to metal conductivity for different microstrip geometries matched to 50 Ω with air above the metal trace.

Configuration	ϵ_r	σ [S/m]	Diel T [μm]	Metal T [μm]	Micro W [μm]	Ins Loss [dB/mm]
High Cond 2D	11.7	5e7	5	0	4.06	8.85
High Cond Thick	11.7	5e7	5	0.2	4.06	4.87
High Cond Thin	11.7	5e7	5	0.1	4.06	5.32
Med Cond Thick	11.7	2e7	5	0.2	4.06	7.49
Low Cond Thick	11.7	8e6	5	0.2	4.06	11.26
Low Diel	2.34	5e7	5	0.2	14.5	3.0
Low Thick Diel	11.7	8e6	10	0.2	29.3	1.61

Diodes Inc and have a dynamic range of at least 75 dB with an output power of -23 dBm [52]. These are paired with wafer probes appropriate for the frequency band such as GGB's model 1100B that have an insertion loss of about 10 dB per probe [53]. Expected losses of microstrip lines are between 5-10 dB / mm at 1 THz, see Table 5.4, and the loss of the meander lines are about twice those of a straight microstrip waveguide. So depending on how closely the probes can be spaced, the insertion loss for testing a SWS geometry can easily be 50-60 dB. While this does reduce the usable range of the extenders, it should still allow for measurements where features in the frequency response are mainly of interest. The primary difficulty is whether we can still resolve the frequency peaks such as those shown in 5.30. The extenders have a quoted magnitude stability of 0.5 dB which is an appreciable fraction of the indicated peaks. It is possible that resolution can be improved with averaging. It may also be possible to use phase information for the SWS dispersion. The extender heads have a phase stability of 6 degrees which should be small enough to measure a difference in phase shift for different length SWSs.

There is not much data available for material properties in the THz band, so a basic attempt at confirming the values used should be made. The metal conductivity could easily be measured via the insertion loss versus length of a simple thru line. In general, the dielectrics also add some loss, but are much smaller than the metal losses as shown in table 5.5. Knowing the metal conductivity more accurately will give a better indication of possible gain for a coupled device. The permittivity of the dielectrics could also be measured using microstrip ring resonators of different sizes. The

Table 5.5: HFSS simulations of insertion loss at 1 THz due to dielectric losses for different substrate materials. Each microstrip geometry is matched to 50 Ω with air above the metal trace.

Material	ϵ_r	loss tangent	Micro W [μm]	Ins Loss [dB/mm]
Si	11.7	1.4e-5	4.06	0.02
GaN (Wurtzite)	9.6	5.23e-4	5.0	0.10
HDPE	2.34	6.24e-4	13.8	0.08
SiC	9.6	1.54e-3	4.94	0.36
Fused Silica	3.84	4.87e-3	10.6	0.71
BCB	2.45	0.01	14.4	1.16

permittivity will not only affect the SWS dispersion, but the 2DEG plasma wave dispersion as well. A small shift in this value will shift both dispersions in the same direction, so it is not expected to be critical. However, it will again play a small role in the overall gain of the device because it will change the interaction impedance.

The primary interest in measuring the SWS is to confirm the axial field dispersion. Because full access to the field information is not possible, some other means is needed to connect the frequency or wavelength response of the axial wave to an S-parameter. One way to do this is with a series of resonant cavities as shown in Figs. 5.28 through 5.33. These use a short section of SWS as a cavity and produces a local minima in the S11 or S21 response. The primary issue with this measurement is that the minima are not very deep compared to the through losses as indicated by the dashed line in Fig. 5.30. For the reflection style cavity, the relative widths of the feed microstrip and the line within the SWS impacts the impedance match at that point. Any variation in this boundary condition of the resonant SWS will shift the dispersion, and that is thought to be the primary reason the mapped modes marked with x's on the right in Fig. 5.31 do not line up perfectly with the long SWS results. In an experiment, there will be some deviation from the nominal line width that will affect this. The resulting error does not seem to be a large correction, and should not preclude a beam-wave interaction in a fabricated device.

Another possible cold-test to measure the SWS dispersion is by measuring the phase shift as a function of distance on the SWS. This would require measuring the two-port phase difference as a

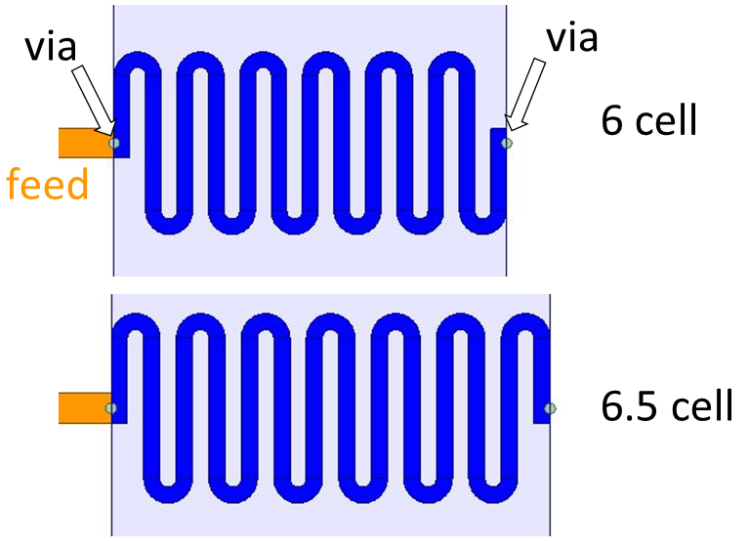

Figure 5.28: Two examples of resonant SWS cavities with different numbers of cells.

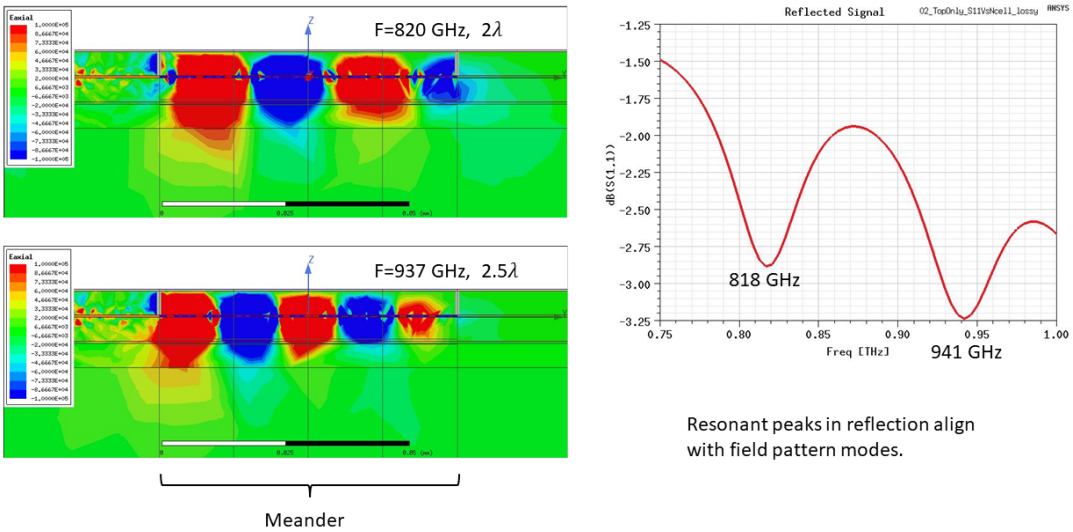

Figure 5.29: On the left the axial field within a cavity SWS is shown for two different resonant modes indicated on the S11 plot shown on the right.

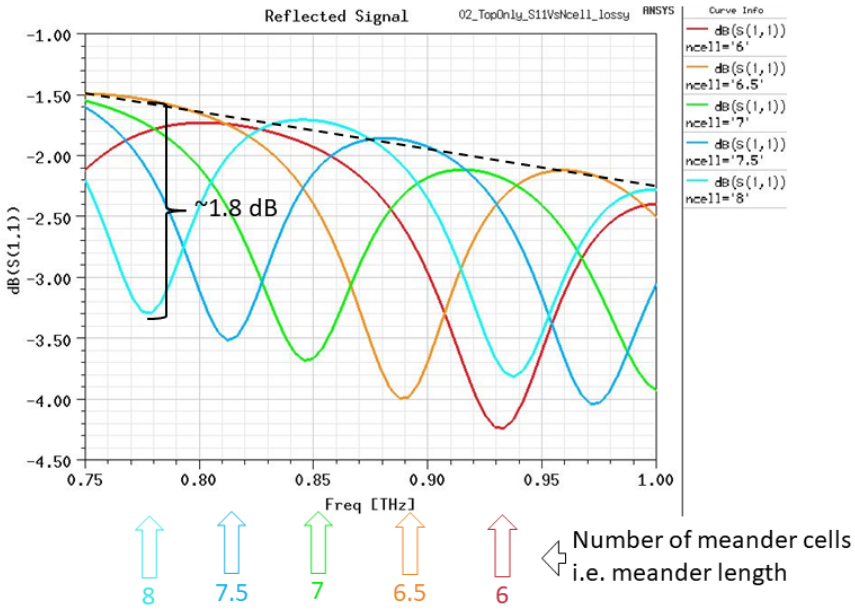

Figure 5.30: Changing the number of cells in the SWS cavity shifts the resonant mode frequency, and this can be used to measure the SWS dispersion. This simulation includes SWS loss, and so the features are limited in magnitude making the measurement more difficult.

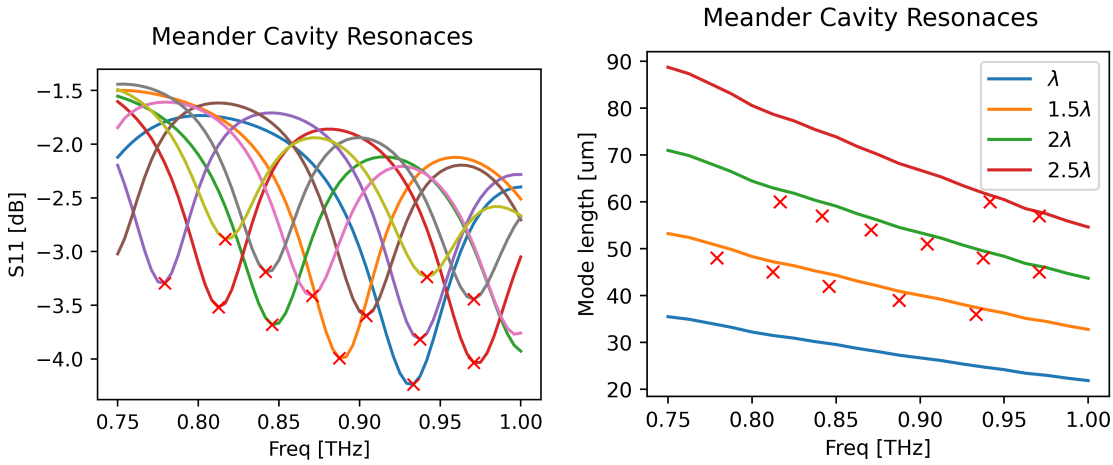

Figure 5.31: All of the S11 resonant features plotted for SWS with different numbers of cells are shown on the right. Plotting those frequencies vs length of the SWS on the right shows the dispersion compared to the predicted dispersion from a 30 cell non-resonant SWS.

Figure 5.32: Using a second smaller section of a SWS shown in red is used as a resonant cavity similar to how a ring resonator is used near a microstrip.

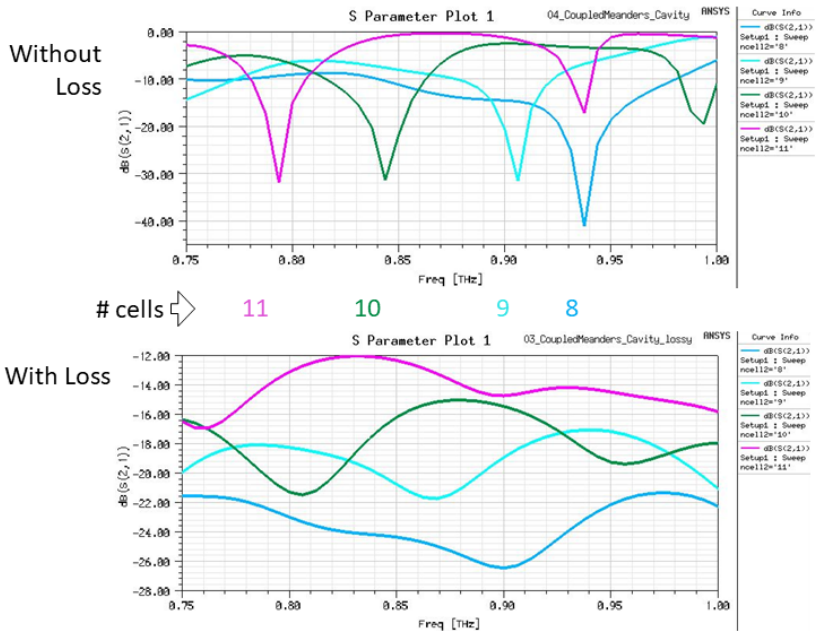

Figure 5.33: The S21 curves for different length resonator cavities with (bottom) and without (top) metal losses.

function of frequency for a series of SWS of different length. Each of the SWSs would have the same geometry, and the included number of cells would change the length. For the simulated data shown in Fig. 5.34, the phase shift for three frequencies is shown versus distance for a SWS with 2-10 cells with a 6 μm pitch. The slope is taken from a linear fit, and the wavelength is calculated using the simple relationship:

$$\frac{-\Delta\theta}{360} = \frac{\Delta L}{\lambda} \rightarrow \lambda = \frac{360}{-slope}. \tag{5.4}$$

The resulting dispersion curve is plotted in Fig. 5.35 both with and without losses. The 'direct length' dispersion curve in those plots is simulated as described in section 5.3 where the zero crossings of the axial electric field along a line are directly calculated. In both the included plots, the direct curve is taken from a lossless SWS simulation. The difference between the two slope derived curves is due to the losses in the SWSs. While it is often ignored as a higher order perturbation, the expansion of the propagation constant for the circuit modeled transmission line shows the affect [37]. For a wave that travels as $V(z) \propto e^{-\gamma z}$

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = j\omega\sqrt{LC}\sqrt{1 - j\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right) - \frac{RG}{\omega^2 LC}}. \tag{5.5}$$

The last term in the expansion is often ignored, but in the THz band losses appear to be significant enough to require it. It is still unclear what the losses will be in the fabricated device, but the change in wavelength of a few microns should not be detrimental to a prototype design. In fact, if the phase noise of the measurement is low enough, this might provide another way to cross-check the loss in the material.

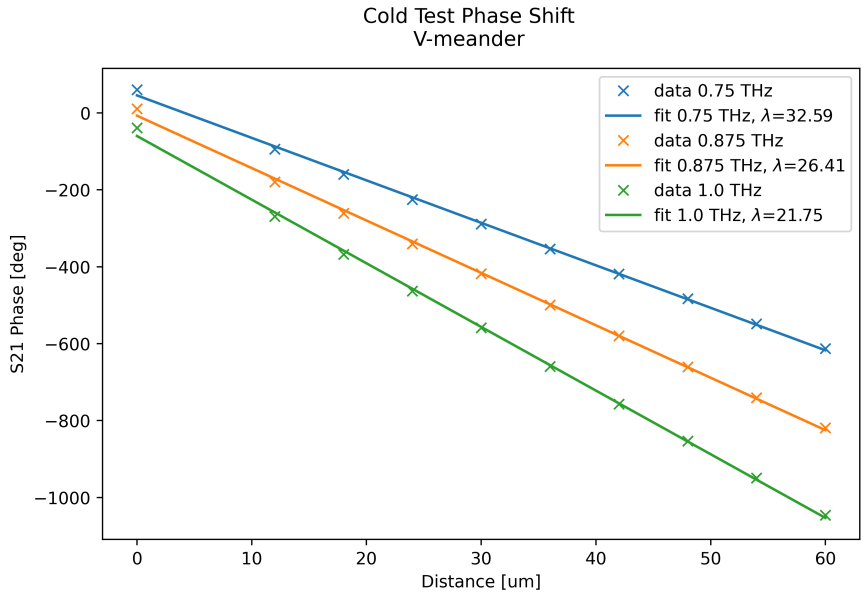

Figure 5.34: Fitting the phase shift as a function of SWS length to calculate the wavelength at different frequencies.

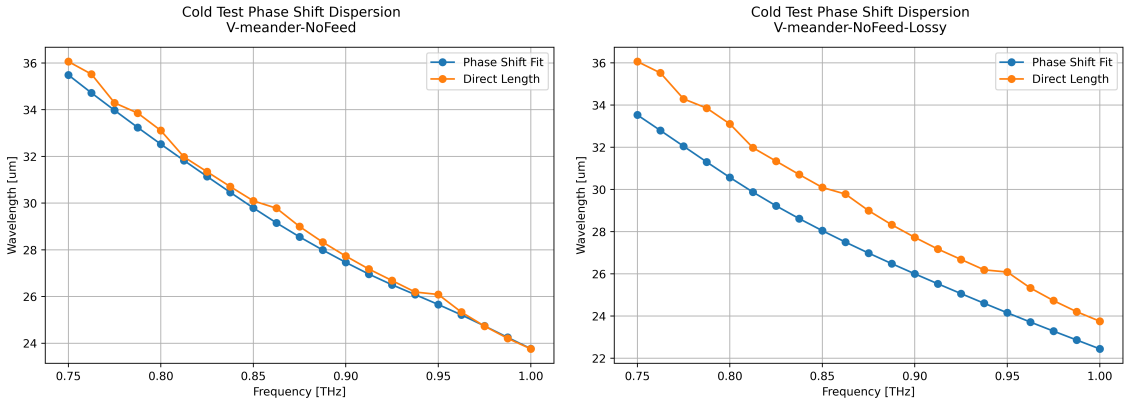

Figure 5.35: The dispersion curves calculated using the slope of the phase shift for the same SWS without loss (left) and with loss (right). For both plots the direct wavelength data is for a lossless SWS.

CHAPTER 6

NUMERICAL SOLUTIONS AND BEHAVIOR OF 2DEG BEAM-WAVE AMPLIFIER

6.1 TWA like behavior

Perhaps the easiest and most complete method to understand the coupled modes of the dispersion equation is to plot the numerical solutions. By solving them numerically, solutions can be visualized for both simplified, and complete dispersion relations. First a traditional vacuum TWA dispersion is shown to introduce the plots that will be used. The following dispersion will be used that is found in [39].

$$(k_z^2 - k_c^2)[(k_z - k_0)^2 - k_p^2] + 2C^3 k_c^3 k_0 = 0 \tag{6.1}$$

Here, $k_0 = \frac{\omega}{v_0}$ is the propagation of the DC electron velocity, and $k_p = \frac{\omega_p}{v_0}$ is due to a 3D plasma frequency. The quantity C in this case is analogous to the definition presented earlier, but a value of $C = 0.1$ that is normal in VED TWAs is used. It is common to use a space-charge term $q = 4QC = \frac{k_p^2}{k_0^2 C^2}$. Following the distinction made earlier, any reference to space charge in this analysis specifically refers to the VED behavior that is distinct from the 2DEG plasma wave behavior even though they are being compared.

In [39], normalized values are introduced as follows. The parameter of non-synchronism between the circuit velocity and electron velocity

$$b = \frac{k_c - k_0}{k_0 C}, \tag{6.2}$$

the normalized change in amplitude

$$x = \frac{Im(k_c)}{k_0 C}, \tag{6.3}$$

and the normalized propagation constant

$$y = \frac{k_0 - Re(k_z)}{k_0 C}. \tag{6.4}$$

In Fig. 6.1, both the raw and normalized solutions can be seen. Both plots are provided because one or the other might be more useful depending on the particular conditions being calculated. In the raw plot, $\beta_1 = Re(k_c) = \frac{\omega}{v_z}$ is equivalent to the reciprocal of the mode velocity. A solution above the red line representing the beam is propagating slower than the beam. In the normalized plot, the values of y that are below 0 are slower than the beam. The change in amplitude of the mode as it propagates is given by β_2 on the raw plot or x on the normalized plot. For $\beta_2 > 0$ and $x > 0$, the mode is growing as it travels down the interaction structure. The gain of a device of length l can be calculated as

$$gain[dB] = 20\log_{10}(e^{\beta_2 l}). \tag{6.5}$$

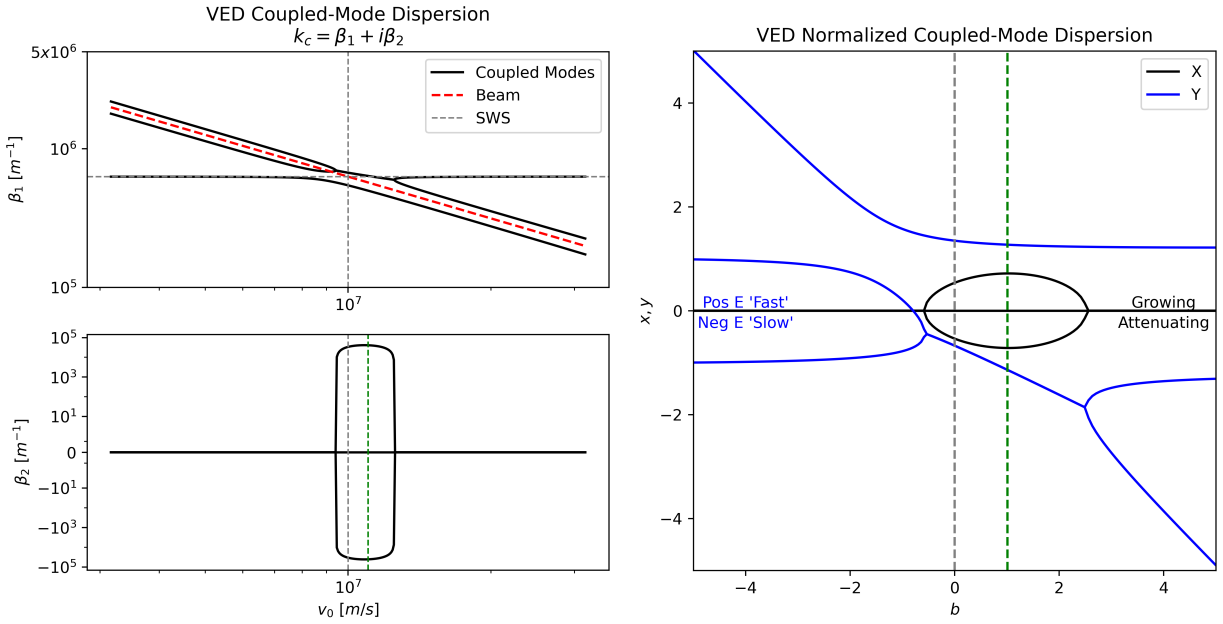

Figure 6.1: The 3 forward coupled mode solutions for a traditional VED TWA. Raw solutions are on the left, and normalized values on the right. The green dashed line indicates the optimal velocity for gain.

In the case shown in Fig. 6.1, there is a small amount of space charge influence. In the traditional Pierce variables $4QC \approx 1$ in this case, and the maximum value of x is at a slightly higher value of b than the synchronous case (b=0). This is congruent with known TWA behavior, and can be interpreted as the RF wave synchronizing with the slow space charge wave on the beam.

Table 6.1: Example Parameters for Numerical Solution

Symbol	Value	Quantity
v_{z0}	$2e6\ cm/s$	2DEG DC beam velocity
v_c	$2e7 - 3e7\ cm/s$	SWS Circuit phase velocity
n_0^{2D}	$2e13\ cm^{-2}$	2DEG density
μ	$1200\ cm^2/(V \cdot s)$	2DEG mobility
w_b	$2 - 6\ \mu m$	Width of 2DEG sheet
Z_c	$1 - 10\ \Omega$	SWS Circuit impedance
L_c	$1000\ \mu m$	SWS Circuit length
ϵ	$10.5\epsilon_0$	Material Permittivity

2DEG TWA Like Behavior

The 2DEG dispersion can be solved by expanding and then gathering Eq. (4.58) into a polynomial of k_z

$$k_z^4\left(v_{z0}^2 - \frac{K_B T}{m^*}\right) + k_z^3\left(-sign(Re(k_z))\frac{e_0^2 n_0^{2D}}{2\epsilon m^*} + i\frac{v_{z0}}{\tau} - 2v_{z0}\omega\right) + k_z^2\left(\frac{e_0^2}{m^*}w_b Z_c n_0^{2D} k_c\omega + \frac{K_B T}{m^*}k_c^2 - v_{z0}^2 k_c^2 - i\frac{\omega}{\tau} + \omega^2\right) + k_z\left(sign(Re(k_z))\frac{e_0^2 n_0^{2D}}{2\epsilon m^*}k_c^2 - i\frac{v_{z0}}{\tau}k_c^2 + 2v_{z0}k_c^2\omega\right) + i\frac{k_c^2\omega}{\tau} - k_c^2\omega^2 = 0 \quad (6.6)$$

This can be solved by a numerical root finding algorithm. For this work a simple python script was used. Example input parameters are given in 6.1, but the values will vary greatly between plots depending on the behavior being shown. Most plots show a list of parameters to the left that are either input parameters or are calculated from them. For instance μ is an input parameter, and τ is calculated. If the plot shows a variable sweep, the parameter list is reflective of the last value in the sweep.

If the DC drift velocity of the 2DEG electrons is allowed to be larger than plasma wave velocity, this system does appear to behave like a traditional TWA as shown in Fig. 6.2. In this case the drift velocity v_0 is much higher than possible in a semiconductor material, but it simply shows a general behavior. Similar behavior can be shown for lower velocities, but those cases also involve highly idealized parameters.

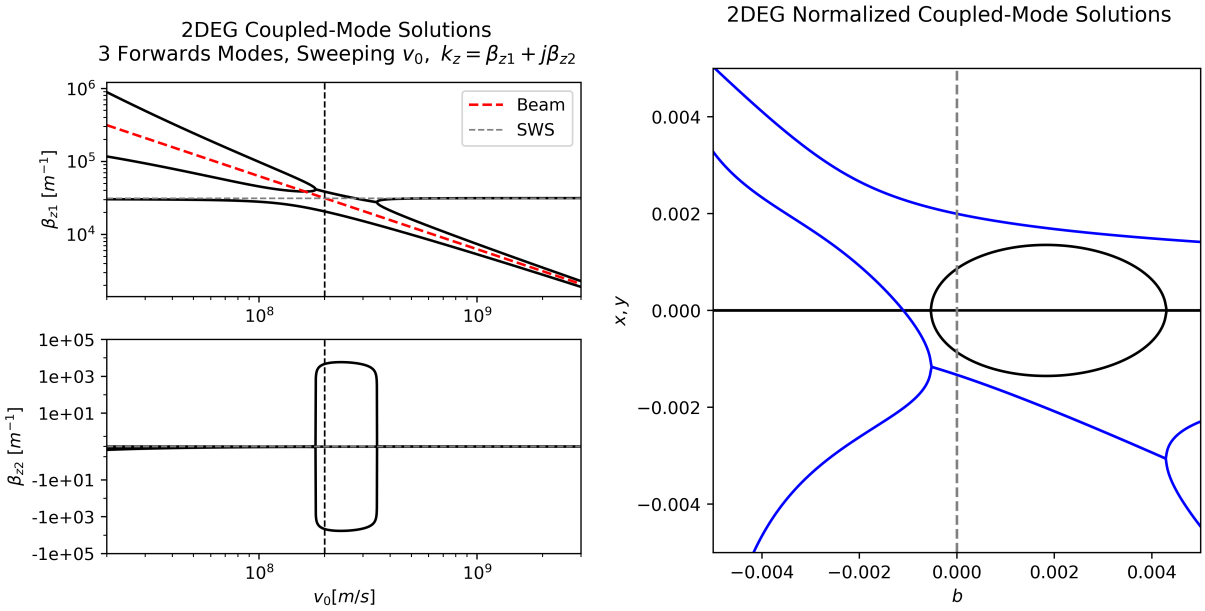

Figure 6.2: The 3 forward coupled mode solutions for a 2DEG TWA. The 2D density is 2×10^{13} cm^{-2} resulting in a v_{p2D} of 2×10^7 m/s. The circuit velocity was set higher at $v_c = 2 \times 10^8$ m/s.

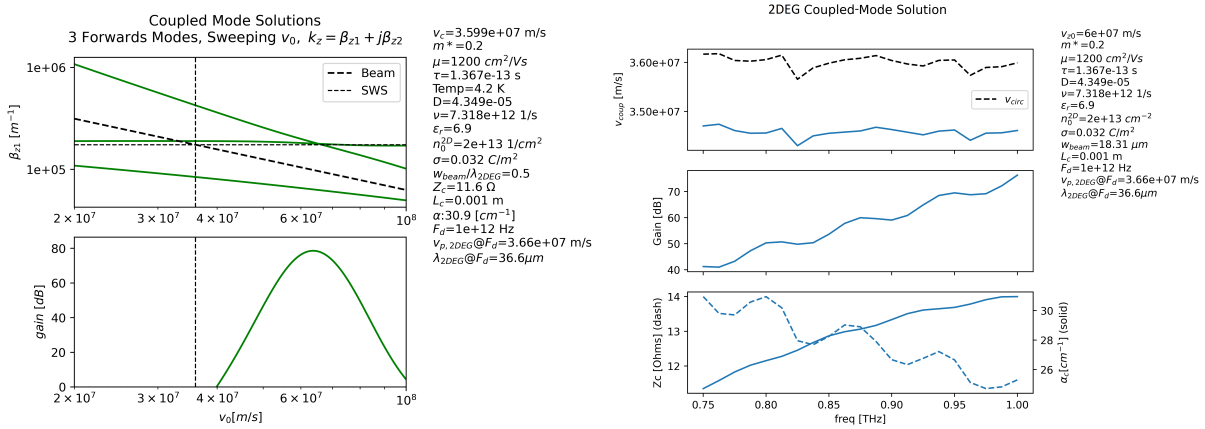

Figure 6.3: A predicted helix response versus electron drift velocity (left), and the corresponding frequency response of a device using EM simulated data for the SWS (right).

If the quantities from the SWS EM simulations discussed in Chapter 5 are used, the frequency response of an idealized device can be calculated as shown in Fig. 6.3.

The gain versus frequency of the SWSs compared in Fig. 5.24 can now be calculated with the 2DEG coupled-mode dispersion using the inputs from the EM simulations. Fig. 6.4 shows that the losses in the two meander line SWSs overcome any energy transfer to the EM wave and therefor have no net gain. As is shown in Fig. 6.8 later, this might be improved by increasing the interaction

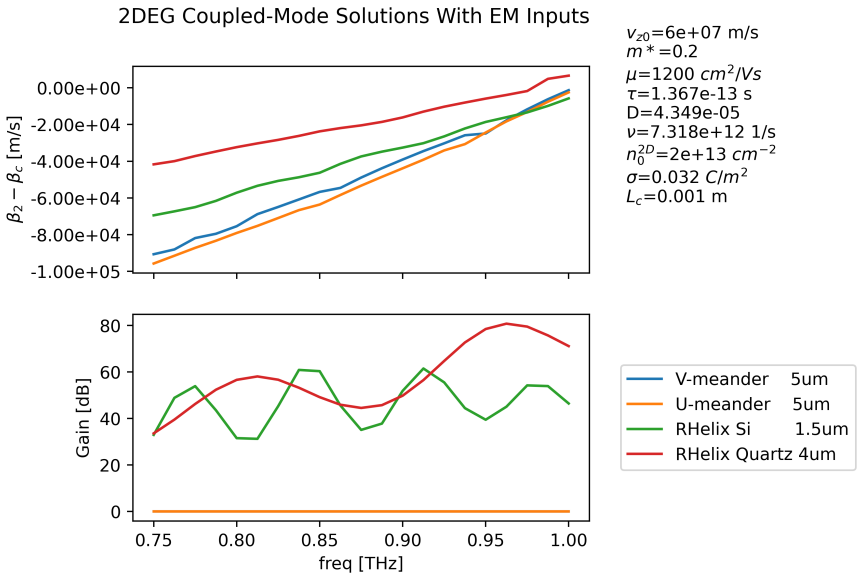

Figure 6.4: The gain for several SWSs is calculated using the coupled-mode theory and input from the EM simulations.

impedance.

One of the important behaviors of the VED TWA is that the maximum gain occurs when the drift velocity is larger than the circuit velocity. For a sufficiently large space charge term that maximum is shown to exist when $k_c = k_0 + k_p$ or in terms of velocities $v_0 = v_c(1 + \omega_p/\omega)$ [39]. This is to say that the circuit wave is synchronizing with the slow space charge wave. Because of the difference in the 2DEG dispersion, namely the inclusion of k_z in the plasma term, it is difficult to derive a similar statement about where the gain peak should be located in terms of the different velocities. It was initially speculated that the max gain would happen when $v_0 = v_c + v_p$, but this is not shown by the numerical solutions in Fig. 6.5

Another suspicious behavior shown in the lower right plot in Fig. 6.5, is that the plasma velocity shown by the dashed line labeled v_p is larger than the drift velocity shown to produce gain. This would seem to be incongruent with only slower plasma waves producing gain, but this particular set of parameters also violates the small signal approximation since the plasma velocity (the AC component) is on the order of the drift velocity (DC component). This particular behavior is discussed further in section 6.3.

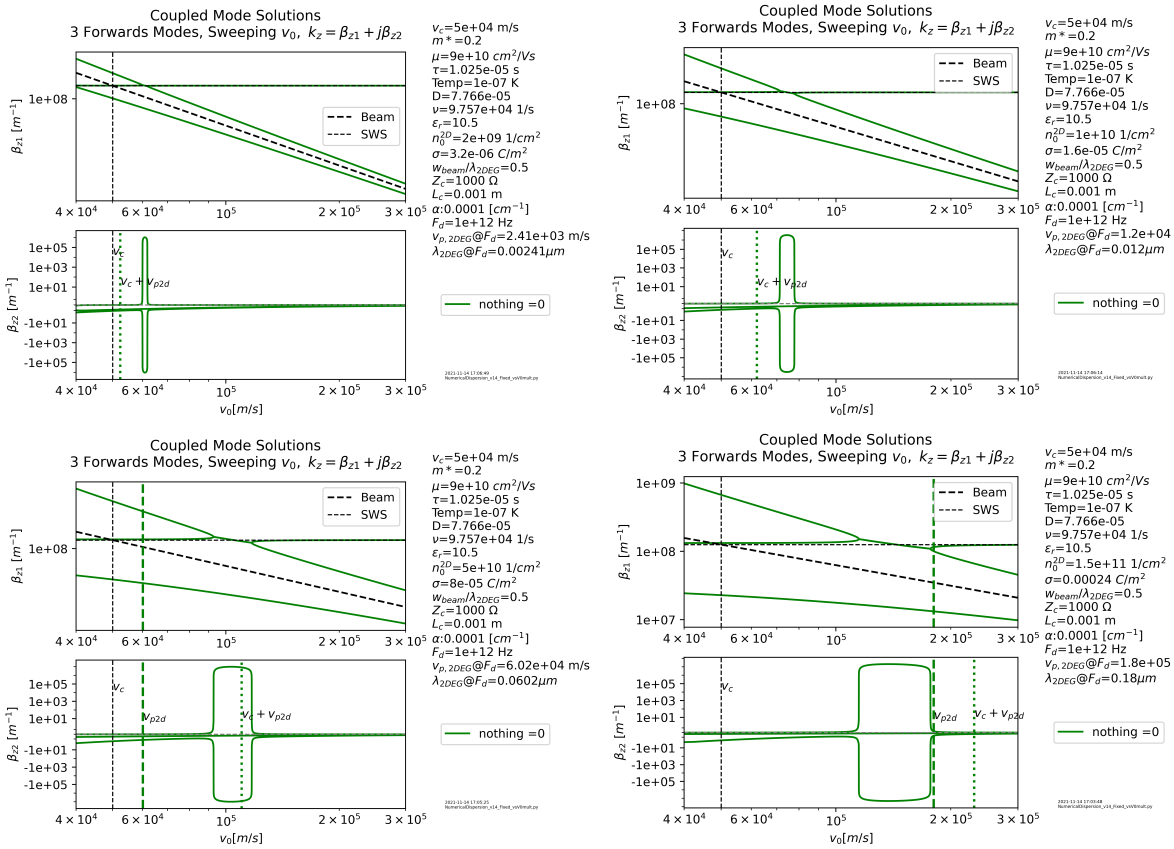

Figure 6.5: A series of dispersion solutions are plotted for different electron densities that show both peak gain, and where the expected peak in gain should be with the dotted line labeled $v_c + v_p$.

6.2 Realistic Device Behavior

The difficulty faced in realistic situations is that the saturation drift velocity is several orders of magnitude lower for known 2DEGs. In GaN for example, the saturation velocity is around 3^5 m/s. If the device were to be operated in the linear region of the velocity vs field curve, this would limit the DC velocity to around 10^5 m/s. For classical TWA style amplifier, the density of the 2DEG would then need to be reduced such that the resulting plasma wave phase velocity was lower than the DC drift velocity. The width of the 2DEG channel would also need to be narrowed to prevent transverse modes. The resulting gain from such a device would be limited by the reduced current.

Other realities of the semiconductor device that need to be considered are the mobility of the electrons in the 2DEG, and their temperature. In this context the primary contribution to the mobility is the acoustic phonon collisions, and it is assumed to be the constant slope of the velocity

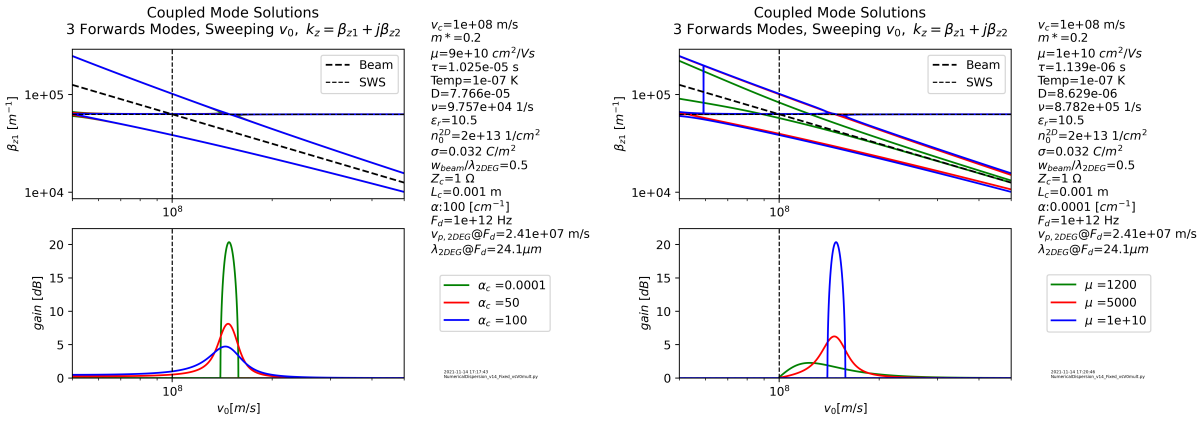

Figure 6.6: The forward coupled-mode solutions for a 2DEG interaction with varying circuit loss α_c on the left, and 2DEG mobility μ on the right.

vs electric field curve. One of the main reasons for operating in the linear velocity range is to avoid exciting optical phonons that are partially responsible for the velocity saturating. The mobility, or the corresponding collision frequency, can be considered a loss term in the context of the plasma wave propagation as it simply randomizes the electron velocity distribution. The effect of increasing mobility can be seen in Fig. 6.6. The gain curve broadens and lowers as the mobility is decreased. Also shown in Fig. 6.6 is the effect of EM loss on the SWS via the α_c term.

The temperature of the electrons, or more directly their thermal velocity, also introduces a similar loss to that mobility. Whereas the acoustic phonon collisions mainly randomize the direction of a given velocity, the thermal energy adds an additional velocity in a random direction. The net result for both is a loss of electrons contributing in phase with the plasma wave. One way in which they differ is that the thermal velocity sets a lower bound on the velocity of the slow plasma wave as seen in Fig. 6.7. The thermal velocity of the electrons at 77 K defined using $\frac{1}{2}m^*v_{th}^2 = K_BT$ gives $v_{th} \approx 1 \times 10^5 m/s$, which is on the order of the drift velocity (regardless of the particular definition for thermal velocity). This would prohibit a coupled plasma wave to grow over any appreciable distance and provide gain.

One of the main concepts argued by Solymar [13], is that most of these losses can be overcome with a large enough coupling impedance. Fig. 6.8 shows that a large value of Z_c is needed to produce gain with a realistic mobility and circuit loss. The quartz helix has a larger impedance-to-loss ratio

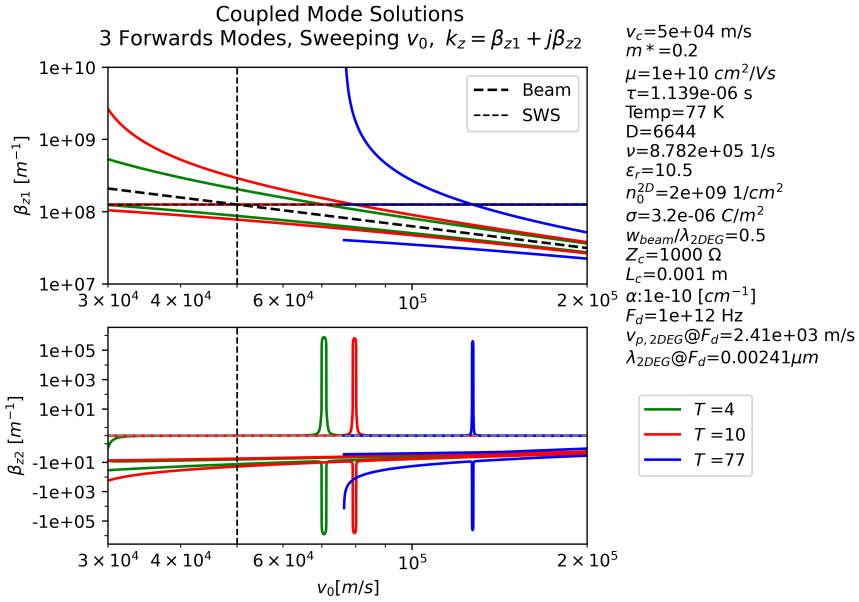

Figure 6.7: Coupled-mode solutions with a realistic drift velocity showing the effect of temperature.

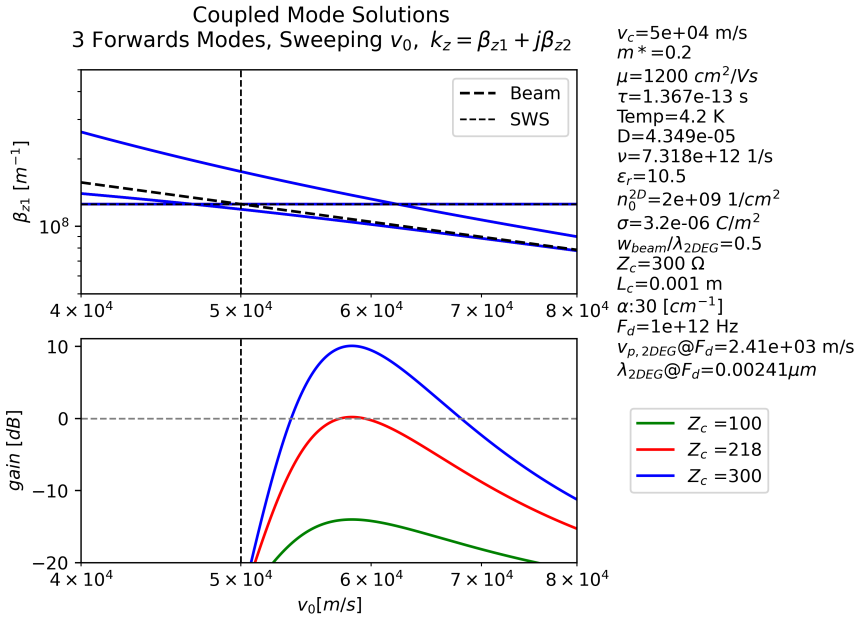

Figure 6.8: Gain for a realistic drift velocity device showing a large coupling impedance, Z_c, is needed to produce gain.

than any of the SWS studied, so it is the most likely to succeed. At a wavelength of around 30 μm, it has an impedance of about 30 Ω at 1 THz. Since Eq. (5.3) shows the impedance is inversely proportional to the wavelength, the impedance should be smaller for the shorter wavelengths associated with lower drift velocities.

To summarize, even if the physical interaction of a 2DEG TWA is accepted, the physical realities of the system would almost certainly preclude it from producing meaningful gain. The SWS would need to have a very small pitch on the order of 10s of nm, and an impossibly high interaction impedance for it to overcome circuit and mobility losses. Even if such a structure could be designed, fabrication would be difficult with such small traces, and even then they would not be able to support very much power leaving it short of the intended goal.

6.3 Energy Flow in a Coupled Systems

Thus far, this analysis of a solid-state TWA has included assumptions and implications that now need to be discussed. Primarily, the assumption that the circuit velocity synchronizes with a perturbed wave slower than the DC electron velocity. This is a basic assumption of TWA operation, and the explanation is not always clearly given. This velocity relation is essential to TWAs because of how they produce gain, or more precisely, how they transfer energy from the DC power source to the RF signal wave. The most basic way to explain the energy transfer in a TWA is that incoming RF signal perturbs the DC electron beam, the electrons that are slowed down by this interaction give up some of their kinetic energy, and start propagating as part of a perturbed AC wave on the beam. The RF wave is given energy by this 'slow space charge wave', and in turn causes an even bigger perturbation. The slow space charge wave can be thought of as a periodic slowing down of the electrons such that the magnitude of the perturbed AC energy on the beam is negative. Thus the RF wave and the slow space charge wave are the two coupled modes that are velocity matched and grow together in the theory described in Chapter 4. The slow space charge wave grows in amplitude by giving up energy. The total beam energy is the sum of the constant DC energy and the negative AC energy, and this allows the transfer of the DC energy to the RF signal through the AC beam perturbation to produce gain [54]. This use of the slow space charge wave to produce gain is the core of all of the O-type VED amplifiers (TWAs, klystrons, resistive wall amplifiers, etc.) [39]. They all rely on the slow space charge wave giving up AC energy from a DC source.

At the same time that the slow space charge wave is formed, there is also energy given from

the RF wave to other electrons in the beam speeding them up. These electrons form the fast space charge wave. However, the faster electrons are only gaining AC energy from the RF wave, they cannot use any of their DC energy for this. These fast electrons cannot give the RF wave any more AC energy than they have already been given from it. Thus the fast space charge wave does not convert the DC kinetic energy into AC energy like the slow space charge wave does and cannot be used in this manner to produce gain [54].

The preceding paragraphs are explicitly referring to the space charge waves in a VED TWA with a sparse beam. In section 4.1, it was suggested that the 2DEG plasma waves are the same as space charge waves in the limit of a dense beam. This then suggests that the above argument regarding the energy of fast and slow space charge waves will hold true for a fast and slow plasma wave as well. On a simplistic level, the same physical interactions are involved in both situations which implies they should behave similarly in regards to energy transfer. However, there is one crucial piece of evidence that at the very least suggests the exchange of energy does not happen with the slow plasma wave. The VED space charge wave is intimately tied to the DC motion such that the group velocity of the space charge wave is the DC beam velocity. While this is not a sufficient reason to associate energy transfer due to the space charge wave with the DC velocity, it would seem to be a necessary one. On the other hand, the plasma wave regardless of its velocity does not have an association with the DC electron velocity. This is seen from the fact that the group velocity of the plasma wave is not constant, nor equal to the DC velocity. In fact, the plasma wave exists without any DC velocity at all. So it would seem that any energy content the plasma wave has does not depend on the DC energy. So even if the plasma wave velocity is positioned in a way where it should behave like a slow space charge wave, it is doubtful that it will do so since there is not any fundamental change in the underlying behavior of the wave.

This leads to the natural question of why the coupled-mode theory predicted gain where it seems to be nonphysical. It is most likely that one or more of the underlying assumptions in the setup was violated. There has already been a demonstration of the relative velocities violating the small signal approximation, i.e. the perturbed AC velocities are much smaller than the DC velocity. In a

VED device, this does not generally need to be enforced rigorously due to the nature of the space charge waves being an oscillation around the DC velocity. In the 2DEG plasma case however, the plasma wave velocity is independent and therefore can be set to a wide range of values. Another, subtler, violation is the unintentional decoupling of the DC energy from the AC energy as outlined above. An adaptation of Chu's power theorem [55] for the one dimensional linearized small signal TWA system is derived in [54] showing the nature of the power in the VED space charge waves and that the coupled mode TWA equations do in fact conserve energy despite being linearized. That derivation justifies the comments about the space charge wave energy in the previous paragraph. A similar derivation might be possible to show the nature of the 2DEG plasma wave energy in the same type of system and indicate whether or not it can access the DC energy.

At this point, the concept of a TWA style solid-state device seems without a resolution given the unfavorable energetics for producing gain. Even if the plasma wave could access the DC energy, the other considerations of section 6.2 show that real gain is unlikely to be produced in physical devices. However, the motivation for a beam-wave style amplifier still exists, and the TWA was just the most well-known type to first be explored. Since it has been shown that the slow beam waves are not viable options for a solid-state amplifier, a fast wave style device becomes the next avenue of exploration. The fast plasma wave can easily approach velocities of the same magnitude as the RF signals, and that makes coupling to them via a SWS much more tractable. The fast plasma wave cannot use DC energy for amplification, so the problem now is how to configure a system to give the fast plasma wave energy that can be transferred to the RF wave. A few examples of systems that might be adapted to this use are discussed in the next chapter.

CHAPTER 7

THE PATH TOWARDS 2DEG FAST WAVE AMPLIFIERS

This last chapter will summarize some of the ongoing and future work to be done along with the cold SWS measurements discussed in Section 5.6. There are several open questions about solid-state beam-wave interactions and there continues to be a desire for this style of amplifier.

7.1 Types of Fast Wave Amplifiers

For the fast plasma wave to be used in a beam-wave style amplifier as suggested at the conclusion of the last chapter, it needs to have energy that can be transferred in some manner to the RF wave.

One possible style is a distributed parametric amplifier such as described in [54], and demonstrated in [56]. This style uses nonlinear mixing on the beam to couple energy from a higher frequency pumped space charge wave to a signal space charge wave. The main issue with this technique is that there is an energy penalty if the space charge waves have a dispersion such that their phase velocities are unequal [54]. As shown previously, the dispersion for the 2DEG plasma waves would result in a difference in the phase velocity of the signal plasma wave and the higher frequency pump plasma wave. The scale of this energy penalty has not yet been determined for the 2DEG beam device, so it is unknown whether this is feasible based solely on efficiency reasons. Even if the energy transfer is favorable, incorporating a source of higher frequency pump power is also an open topic. Integrated laser like sources would be ideal, but currently they lack the output power likely needed.

Another family of amplifiers to consider are cross-field TWAs, sometimes referred to as linear cross field amplifiers (CFA). In this style of amplifier, the energy used for gain comes from an electron beam that is formed with a high electrostatic potential energy, and the beam is perturbed to release that energy in phase with the incoming EM signal wave [39]. A typical CFA is in a ring configuration like a magnetron, but a linear version such as found in [57] might work better with semiconductor fabrication techniques. It is not yet clear whether a 2DEG plasma wave could be

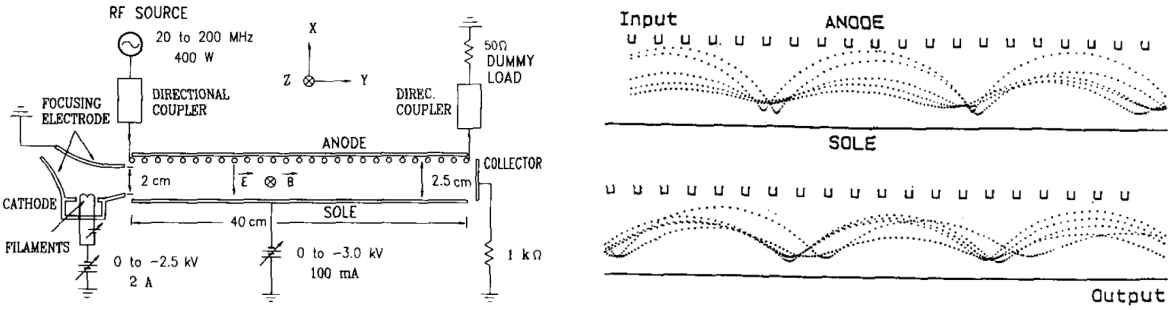

Figure 7.1: A diagram of a VED CFA (left), and the resulting electron motion (right).[5] ©1992 IEEE

generated and used in the exact same manor. However, there are other ways in which a 2DEG plasma can be put into a higher energy state to be perturbed. It should be noted that there is a subtle difference between this concept and the physics used for gain in semiconductor optical amplifiers and lasers. In the optical amplifier case, individual excited electron states are directly producing photons at the operating frequency. Whereas the CFA case does not depend on discrete energy states matching the operating frequency, and there is a continuous exchange of energy from the electron beam to the EM fields as in a TWA.

7.2 Beam-Wave Coupling Experiments

One of the key missing pieces in the development of this type of amplifier is a clear demonstration of the coupling between a SWS and a 2DEG plasma wave. Some of the general difficulties of performing experiments in this frequency band were covered in section 5.6. Presented here is a more specific plan for experiments to show the beam-wave coupling. Both cold-test and coupling experiments will likely proceed in the future as part of the same project that supported this thesis since they share a lot of work in common.

Simulations for the coupled interaction between the SWS and 2DEG plasma wave are currently done in VSim by Tech-X [29]. A collaborator created a VSim input file that includes the 2DEG plasma wave behavior via a field updater in the finite different time domain solution [58]. This setup does create a highly idealized and prescribed plasma response, but more importantly the coupling experiments are intended to show that this is an appropriate starting point. A particle-in-

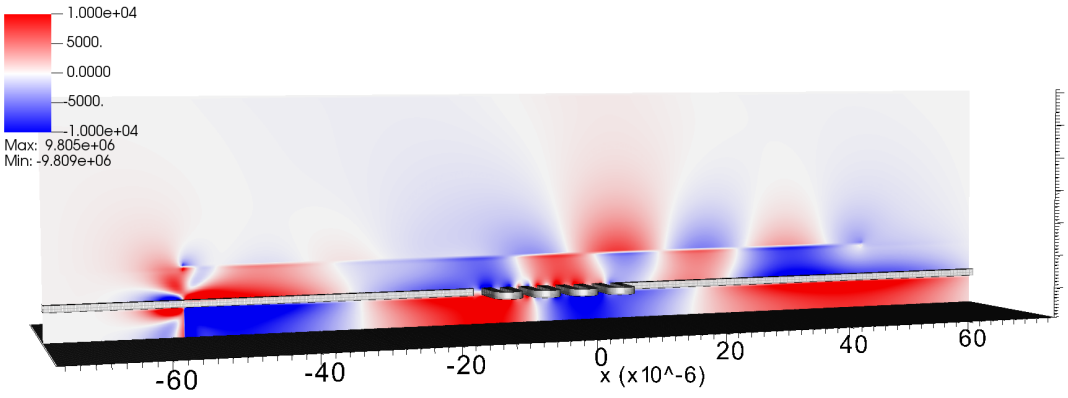

Figure 7.2: An example simulation in VSim showing a 4 cell U-meander SWS exciting a 2DEG above it. Plotted is the vertical component of the electric field.

cell (PIC) style simulation would be better in terms of letting the electrons just respond as they will without assuming a particular reaction, but PIC simulations at high densities become difficult due to super-particle representation issues.

Pictured in Fig. 7.2 is an example of a VSim result. A four cell U-meander geometry was created in HFSS to have a good impedance match, and the geometry was exported. This structured was meshed and imported into VSim as a perfectly conducting metal. The 2DEG layer can be seen from the field discontinuity above the SWS. The SWS is excited on the left side by a sine wave voltage. The longer wavelength of the microstrip can be seen to the left of the SWS, and the shorter wavelength of the SWS is seen to excite a wave in the 2DEG above it.

One possible beam-wave coupling experiment is to look at the transmitted wave of the system described above with a varying number of SWS cells. This would vary the length of interaction and change the amount of power transferred between the 2DEG and the EM wave. This idea is based on the mode coupling between microstrips used in quadrature couplers for instance shown in Fig. 7.3. When the two mode couple, power oscillates between them as they travel. By changing the length of the coupled interaction, the amount of power in each line can be changed. In Fig. 7.4, a proof of concept model is shown to mimic the same behavior with coupled SWS. The SWS on the bottom of that setup has a lower conductivity to mimic the losses in a 2DEG. A clear pattern can be seen between the two curves in the included plot. The blue curve indicates just the normal SWS losses

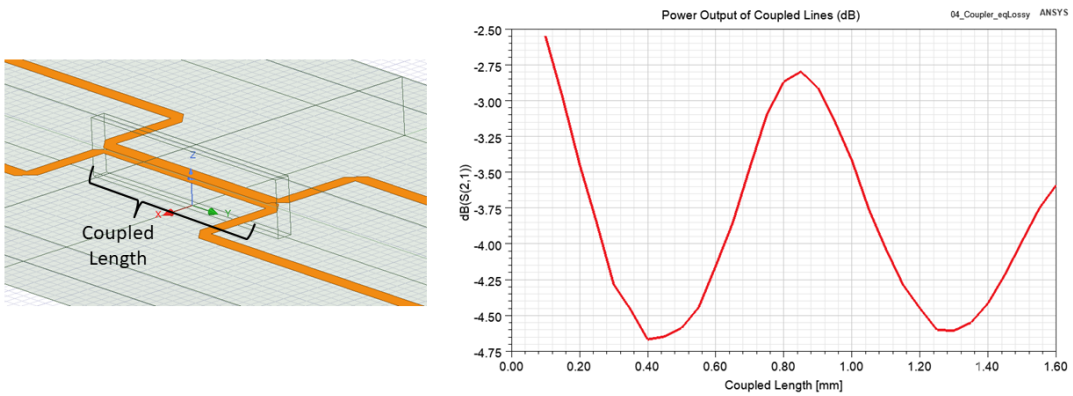

Figure 7.3: A broadside microstrip coupler model (ground planes not shown) in HFSS (left) with the resulting transmitted power as a function of the coupler length (right).

without the 2DEG, and the red curve shows the type of energy exchange with the 2DEG. While these differences are small, they might be measurable.

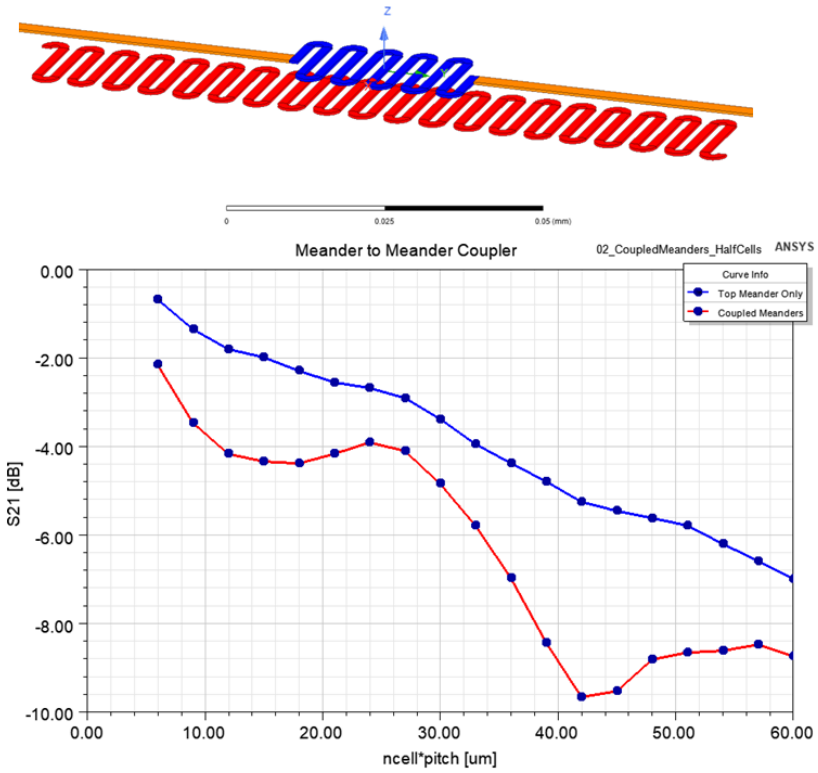

Figure 7.4: A HFSS model for a pair of coupled SWSs. The upper meander (blue) has a typical metal conductivity of 5e7 S/m. The lower (red) meander has a conductivity of 1e6 S/m to mimic the decay constant for a 2DEG. The results show a clear exchange of energy oscillating as a function of distance.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] J. D. Albrecht, M. J. Rosker, H. B. Wallace, and T. Chang, "THz Electronics projects at DARPA: Transistors, TMICs, and amplifiers," in *2010 IEEE MTT-S International Microwave Symposium*, May 2010, pp. 1–1.

[2] M. Field, T. Kimura, J. Atkinson, D. Gamzina, N. C. Luhmann, B. Stockwell, T. J. Grant, Z. Griffith, R. Borwick, C. Hillman, B. Brar, T. Reed, M. Rodwell, Y.-M. Shin, L. R. Barnett, A. Baig, B. Popovic, C. Domier, R. Barchfield, J. Zhao, J. A. Higgins, and Y. Goren, "Development of a 100-W 200-GHz High Bandwidth mm-Wave Amplifier," *IEEE Transactions on Electron Devices*, vol. 65, no. 6, pp. 2122–2128, Jun. 2018.

[3] M. Farahmand, C. Garetto, E. Bellotti, K. Brennan, M. Goano, E. Ghillino, G. Ghione, J. Albrecht, and P. Ruden, "Monte Carlo simulation of electron transport in the III-nitride wurtzite phase materials system: binaries and ternaries," *IEEE Transactions on Electron Devices*, vol. 48, no. 3, pp. 535–542, Mar. 2001.

[4] J.-H. Seo, private communication.

[5] J. Browning, C. Chan, J. Ye, G. Dombrowski, and T. Ruden, "Electron plasma and wave measurements in a crossed-field amplifier and comparison with numerical simulation," *IEEE Transactions on Electron Devices*, vol. 39, no. 10, pp. 2401–2407, Oct. 1992.

[6] S. S. Dhillon, M. S. Vitiello, E. H. Linfield, A. G. Davies, M. C. Hoffmann, J. Booske, C. Paoloni, M. Gensch, P. Weightman, G. P. Williams, E. Castro-Camus, D. R. S. Cumming, F. Simoens, I. Escorcia-Carranza, J. Grant, S. Lucyszyn, M. Kuwata-Gonokami, K. Konishi, M. Koch, C. A. Schmuttenmaer, T. L. Cocker, R. Huber, A. G. Markelz, Z. D. Taylor, V. P. Wallace, J. A. Zeitler, J. Sibik, T. M. Korter, B. Ellison, S. Rea, P. Goldsmith, K. B. Cooper, R. Appleby, D. Pardo, P. G. Huggard, V. Krozer, H. Shams, M. Fice, C. Renaud, A. Seeds, A. Stöhr, M. Naftaly, N. Ridler, R. Clarke, J. E. Cunningham, and M. B. Johnston, "The 2017 terahertz science and technology roadmap," *Journal of Physics D: Applied Physics*, vol. 50, p. 043001, Jan. 2017.

[7] T. Nagatsuma, S. Horiguchi, Y. Minamikata, Y. Yoshimizu, S. Hisatake, S. Kuwano, N. Yoshimoto, J. Terada, and H. Takahashi, "Terahertz wireless communications based on photonics technologies," *Optics Express*, vol. 21, no. 20, pp. 23 736–23 747, Oct. 2013.

[8] M. Rosker, "Terahertz Electronics BAA, DARPA-BAA-08-51." [Online]. Available: https://sam.gov/opp/9078dc01af1dc24b8e5d69617b799516/view

[9] A. S. Gilmour, *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-field Amplifiers, and Gyrotrons*. Boston, MA: Artech House, 2011.

[10] X. Mei, W. Yoshida, M. Lange, J. Lee, J. Zhou, P.-H. Liu, K. Leong, A. Zamora, J. Padilla, S. Sarkozy, R. Lai, and W. R. Deal, "First Demonstration of Amplification at 1 THz Using 25-nm InP High Electron Mobility Transistor Process," *IEEE Electron Device Letters*, vol. 36, no. 4, pp. 327–329, Apr. 2015.

[11] R. J. B. Dietz, B. Globisch, H. Roehle, D. Stanze, T. Göbel, and M. Schell, "Influence and adjustment of carrier lifetimes in InGaAs/InAlAs photoconductive pulsed terahertz detectors: 6 THz bandwidth and 90dB dynamic range," *Optics Express*, vol. 22, no. 16, pp. 19 411–19 422, Aug. 2014.

[12] M. Sumi, "Traveling-wave amplification by drifting carriers in semiconductors," *Applied Physics Letters*, vol. 9, no. 6, pp. 251–253, Sep. 1966.

[13] L. Solymar and E. Ash, "Some Travelling-wave Interactions in Semiconductors Theory and Design Considerations," *International Journal of Electronics*, vol. 20, no. 2, 1966.

[14] A. Gover, K. H. Burrell, and A. Yariv, "Solid-state traveling-wave amplification in the collisionless regime," *Journal of Applied Physics*, vol. 45, no. 11, pp. 4847–4851, Nov. 1974.

[15] A. Gover and A. Yariv, "Monolithic solid-state traveling-wave amplifier," *Journal of Applied Physics*, vol. 45, no. 6, pp. 2596–2600, Jun. 1974.

[16] F. Stern, "Polarizability of a Two-Dimensional Electron Gas," *Physical Review Letters*, vol. 18, no. 14, pp. 546–548, Apr. 1967.

[17] L. C. Súilleabháin, H. P. Hughes, A. C. Churchill, D. A. Ritchie, M. Grimshaw, and G. A. C. Jones, "Raman studies of intrasubband plasmon dispersion for a single 2DEG under a Schottky gate," *Solid State Communications*, vol. 87, no. 6, pp. 517–521, Aug. 1993.

[18] D. A. Neamen, *Semiconductor Physics and Devices: Basic Principles*. New York, NY: McGraw-Hill, 2012.

[19] P. Harrison and A. Valavanis, *Quantum Wells, Wires and Dots: Theoretical and computational physics of semiconductor nanostructures*. John Wiley & Sons, Ltd, 2016.

[20] C. Jacoboni, *Theory of Electron Transport in Semiconductors*. Springer, 2010.

[21] R. J. Goldston and P. H. Rutherford, *Introduction to Plasma Physics*. CRC Press, 2018.

[22] M. Steele and B. Vural, *Wave Interactions in Solid State Plasmas*. New York, NY: McGraw-Hill, 1969.

[23] A. Piel, *Plasma Physics - An introduction to laboratory, space, and fusion plasmas*. Springer International Publishing AG, 2017.

[24] E. W. Weisstein, "Courant-Friedrichs-Lewy condition," https://mathworld.wolfram.com/Courant-Friedrichs-LewyCondition.html.

[25] Y.-C. Wu and Y.-R. Jhan, "Introduction of Synopsys Sentaurus TCAD Simulation," in *3D TCAD Simulation for CMOS Nanoeletronic Devices*. Singapore: Springer Singapore, 2018, pp. 1–17.

[26] M. Grupen, "Energy transport model with full band structure for GaAs electronic devices," *Journal of Computational Electronics*, vol. 10, no. 3, pp. 271–290, Sep. 2011.

[27] J. Wang, G. Du, and X. Liu, "Investigation of terahertz plasma oscillations in nano-scaled double-gate MOSFETs by Monte Carlo method," in *2014 12th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Oct. 2014, pp. 1–3.

[28] L. Reggiani, P. Golinelli, L. Varani, T. González, D. Pardo, E. Starikov, P. Shiktorov, and V. Gružinskis, "Monte Carlo analysis of electronic noise in semiconductor materials and devices," *Microelectronics Journal*, vol. 28, no. 2, pp. 183–198, Feb. 1997.

[29] "VSim - Tech-X," https://txcorp.com/vsim/.

[30] C. Nieter and J. R. Cary, "VORPAL: a versatile plasma simulation code," *Journal of Computational Physics*, vol. 196, no. 2, pp. 448–473, May 2004.

[31] J. R. Pierce, "Traveling-wave Tubes," *The Bell System Technical Journal*, vol. 29, no. 1, pp. 1–59, Jan. 1950.

[32] P. Wong, P. Zhang, and J. Luginsland, "Recent theory of traveling-wave tubes: a tutorial-review," *Plasma Research Express*, vol. 2, no. 2, p. 023001, Jun. 2020.

[33] W. Jones and N. H. March, *Theoretical Solid State Physics*. New York: Dover Publications, 1985.

[34] P. Zhang, J. Albrecht, M. Hodek, D. Smithe, P. Wong, and Y. Zhou, "Space Charge Waves In A 2DEG," in *48th IEEE International Conference On Plasma Science (ICOPS)*, Sep. 2021.

[35] Y. Y. Lau and D. Chernin, "A review of the ac space-charge effect in electron–circuit interactions," *Physics of Fluids B: Plasma Physics*, vol. 4, no. 11, pp. 3473–3497, Nov. 1992.

[36] F. Chen, *Introduction to Plasma Physics and Controlled Fusion*. Springer, 2018.

[37] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, Ltd, 2017.

[38] D. K. Cheng, *Field and Wave Electromagnetics*. Reading, Mass: Addison-Wesley, 2004.

[39] S. Tsimring, *Electron Beams and Microwave Vacuum Electronics*. John Wiley & Sons, Ltd, 2006.

[40] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, no. 2, pp. 437–672, Apr. 1982.

[41] M. Naftaly, *Terahertz Metrology*. Artech House, 2015.

[42] M. T. Hibberd, V. Frey, B. F. Spencer, P. W. Mitchell, P. Dawson, M. J. Kappers, R. A. Oliver, C. J. Humphreys, and D. M. Graham, "Dielectric response of wurtzite gallium nitride in the terahertz frequency range," *Solid State Communications*, vol. 247, pp. 68–71, Dec. 2016.

[43] E. Perret, N. Zerounian, S. David, and F. Aniel, "Complex permittivity characterization of benzocyclobutene for terahertz applications," *Microelectronic Engineering*, vol. 85, no. 11, pp. 2276–2281, Nov. 2008.

[44] M. Walther, D. G. Cooke, C. Sherstan, M. Hajar, M. R. Freeman, and F. A. Hegmann, "Terahertz conductivity of thin gold films at the metal-insulator percolation transition," *Physical Review B*, vol. 76, no. 12, p. 125408, Sep. 2007.

[45] N. Laman and D. Grischkowsky, "Terahertz conductivity of thin metal films," *Applied Physics Letters*, vol. 93, no. 5, p. 051105, Aug. 2008.

[46] M. P. Kirley and J. H. Booske, "Terahertz Conductivity of Copper Surfaces," *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 6, pp. 1012–1020, Nov. 2015.

[47] M. P. Kirley, N. Carlsson, B. B. Yang, and J. H. Booske, "Study of the effect of surface roughness and skin depth on the conductivity of metals at 650 GHz," in *IVEC 2012*, Apr. 2012, pp. 239–240.

[48] N. Laman and D. Grischkowsky, "Reduced conductivity in the terahertz skin-depth layer of metals," *Applied Physics Letters*, vol. 90, no. 12, p. 122115, Mar. 2007.

[49] F. Shen, Y. Wei, H. Yin, Y. Gong, X. Xu, S. Wang, W. Wang, and J. Feng, "A Novel V-Shaped Microstrip Meander-Line Slow-Wave Structure for W-band MMPM," *IEEE Transactions on Plasma Science*, vol. 40, no. 2, pp. 463–469, Feb. 2012.

[50] H. Geng, *Semiconductor Manufacturing Handbook, Second Edition*. McGraw-Hill Education, 2018.

[51] S. J. Cho, D. Liu, A. Hardy, J. Kim, J. Gong, C. J. Herrera-Rodriguez, E. Swinnich, X. Konstantinou, G.-Y. Oh, D. G. Kim, J. C. Shin, J. Papapolymerou, M. Becker, J.-H. Seo, J. D. Albrecht, T. A. Grotjohn, and Z. Ma, "Fabrication of AlGaAs/GaAs/diamond heterojunctions for diamond-collector HBTs," *AIP Advances*, vol. 10, no. 12, p. 125226, Dec. 2020.

[52] "Vector Network Analyzer Extension Modules (VNAX)." [Online]. Available: https://www.vadiodes.com/en/products/vector-network-analyzer-extension-modules

[53] "GGB Industries | Model 1100B." [Online]. Available: https://ggb.com/home/model-1100b/

[54] W. H. Louisell, *Coupled Mode and Parametric Electronics*. Wiley: New York, 1960.

[55] L. Chu, "A kinetic power theorem," in *IRE-PGED Electron Tube Research Conf., Durham, NH*, 1951.

[56] A. Ashkin, "Parametric Amplification of Space Charge Waves," *Journal of Applied Physics*, vol. 29, no. 12, pp. 1646–1651, Dec. 1958.

[57] M. Pearlman and J. Browning, "Simulation of a Distributed Cathode in a Linear-Format Crossed-Field Amplifier," *IEEE Transactions on Plasma Science*, vol. 46, no. 7, pp. 2497–2504, Jul. 2018.

[58] D. Smithe, private communication.