Quantum Cascade Laser Theory
Waveguide Effects

Dr. Christopher S. Baird, University of Massachusetts Lowell

1.0 Introduction

The internal waveguide of a quantum cascade laser will only support certain modes of the laser field. The internal waveguide typically consists of reflective or semi-reflective materials layers above and below the active region layers, and may include the substrate. The reflective layers are typically metal or heavily-doped semiconducting material. The waveguide code is tasked with receiving any series of layers from the user (material, width, and doping), and calculating the fundamental mode for that structure at a certain frequency. From the mode, we know the loss, wavenumber, and confinement factor that the laser radiation field will experience. Because the waveguide calculations depend only on the fixed waveguide structure and a frequency, we carry them out at the beginning of the code before entering the iterative loops. However, we do not know in advance the lasing frequency. The solution is to run the waveguide calculations for several possible frequencies and generate a lookup table. Then, later in the iterative loops when the waveguide parameters are needed and the frequency known, the parameters can be simply interpolated from the lookup table.

Our code currently employs a simple one-dimensional slab waveguide model. The width and depth of a QCL is typically so much larger than its height that the QCL is approximated to be uniform and infinite in these dimensions, which reduces the problem down to one dimension. Note that individual layers within the active region are so thin compared to the waveguide layers, that we assume their effects to be negligible. We instead model the entire active region as one waveguide layer with a doping equal to the average of the actual layer dopings.

2.0 Complex Permittivity

We use the Drude model for the complex permittivity ε_c of each layer,

$$\epsilon_c(\omega)=\epsilon+\frac{ne^2\tau^2}{m^*(1+(\omega\tau)^2)}\left(\frac{i}{\omega\tau}-1\right)$$

where n is the free carrier density, and τ is the electron momentum relaxation time. For gold, $n_{\text{Gold}}=5.6\times10^{28}$ and $\tau_{\text{Gold}}=5\times10^{-14}$. For GaAs, the free carrier density is just the ionized doping density and the electron relaxation time is found from experiment to be:

$$\tau_{GaAs} = 10^{-13} + \frac{0.71\times10^{10}}{n+2.2\times10^{21}}$$

where the free carrier density is in units of m^{-3} and the relaxation time is in seconds. Similar expressions for other materials can be found in the literature.

3.0 Mirror Losses

Any loss causes the intensity of the electromagnetic wave to attenuate in space:

$$I(z) = I_0 e^{-\alpha z}$$

After one full round trip, the intensity lost out the front mirror (mirror two) diminishes the wave, so that the resultant intensity is the original intensity times the reflectivity R_2 of the second mirror.:

$$I(2L) = R_2 I_0$$

Compare this to the first equation and solve for the loss:

$$R_2 I_0 = I_0 e^{-\alpha_{M2} 2L}$$

$$\boxed{\alpha_{M2} = -\frac{\ln(R_2)}{2L}}$$ (For a GaAs/air interface, $R = 0.32$)

Here L is the length of the QCL cavity in the direction that the radiation is emitted. The mirror loss due to the back mirror has the exact same form.

4.0 General Waveguide Equations

Maxwell's equations state:

$$\nabla\cdot\mathbf{D} = \rho \qquad \nabla\cdot\mathbf{B} = 0 \qquad \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \qquad \nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial\mathbf{D}}{\partial t}$$

Assume the free currents and free charges are negligible so that this becomes:

$$\nabla\cdot\mathbf{D}=0 \qquad \nabla\cdot\mathbf{B}=0 \qquad \nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} \qquad \nabla\times\mathbf{H}=\frac{\partial\mathbf{D}}{\partial t}$$

In any one region let us assume the material is uniform, linear, and isotropic so that $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$, leading to:

$$\nabla\cdot\mathbf{E}=0 \qquad \nabla\cdot\mathbf{H}=0 \qquad \nabla\times\mathbf{E}=-\mu\frac{\partial\mathbf{H}}{\partial t} \qquad \nabla\times\mathbf{H}=\epsilon\frac{\partial\mathbf{E}}{\partial t}$$

Assume the waveguide has a uniform shape along its axis, and its axis is in the z direction. All of the fields therefore have a harmonic free-wave solution in this dimension with wave number k_z. Also assume that all of the fields are oscillating harmonically in time at the same frequency ω:

$$\mathbf{E}=\mathbf{E}(x,y)e^{ik_z z-i\omega t} \qquad \mathbf{H}=\mathbf{H}(x,y)e^{ik_z z-i\omega t}$$

Using these forms, Maxwell's equations become:

$$\boxed{E_z=\frac{i}{k_z}\left[\frac{\partial E_x}{\partial x}+\frac{\partial E_y}{\partial y}\right]} \qquad \boxed{H_z=\frac{i}{k_z}\left[\frac{\partial H_x}{\partial x}+\frac{\partial H_y}{\partial y}\right]} \qquad \nabla\times\mathbf{H}=-i\omega\epsilon\,\mathbf{E} \qquad \nabla\times\mathbf{E}=i\omega\mu\,\mathbf{H}$$

The first two equations give us the parallel fields if we know the transverse fields. The last two equations can do the opposite. If we manipulate the last two equations, we can get them in a form where we can calculate the transverse fields if we know the parallel fields. Expand the fields and curls into parallel and transverse components:

$$\left(\nabla_t+\hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\times(\mathbf{H}_t+H_z\hat{\mathbf{z}})=-i\omega\epsilon(\mathbf{E}_t+E_z\hat{\mathbf{z}}) \qquad \left(\nabla_t+\hat{\mathbf{z}}\frac{\partial}{\partial z}\right)\times(\mathbf{E}_t+E_z\hat{\mathbf{z}})=i\omega\mu(\mathbf{H}_t+H_z\hat{\mathbf{z}})$$

Distribute through:

$$\nabla_t\times\mathbf{H}_t+\nabla_t\times H_z\hat{\mathbf{z}}+ik_z\hat{\mathbf{z}}\times\mathbf{H}_t=-i\omega\epsilon(\mathbf{E}_t+E_z\hat{\mathbf{z}}) \qquad \nabla_t\times\mathbf{E}_t+\nabla_t\times E_z\hat{\mathbf{z}}+ik_z\hat{\mathbf{z}}\times\mathbf{E}_t=i\omega\mu(\mathbf{H}_t+H_z\hat{\mathbf{z}})$$

Cross both equations on both sides with the unit vector in the z direction to single out the transverse component of the fields on the right:

$$\hat{\mathbf{z}}\times\nabla_t\times H_z\hat{\mathbf{z}}+ik_z\hat{\mathbf{z}}\times\hat{\mathbf{z}}\times\mathbf{H}_t=-i\omega\epsilon\,\hat{\mathbf{z}}\times\mathbf{E}_t \qquad \hat{\mathbf{z}}\times\nabla_t\times E_z\hat{\mathbf{z}}+ik_z\hat{\mathbf{z}}\times\hat{\mathbf{z}}\times\mathbf{E}_t=i\omega\mu\,\hat{\mathbf{z}}\times\mathbf{H}_t$$

The unit vector in the z-direction is perpendicular to the transverse field vectors so that we know $\hat{\mathbf{z}}\times\hat{\mathbf{z}}\times\mathbf{H}_t=-\mathbf{H}_t$ and $\hat{\mathbf{z}}\times\hat{\mathbf{z}}\times\mathbf{E}_t=-\mathbf{E}_t$. We can also expand into components to show that $\hat{\mathbf{z}}\times\nabla_t\times H_z\hat{\mathbf{z}}=\nabla_t H_z$ and $\hat{\mathbf{z}}\times\nabla_t\times E_z\hat{\mathbf{z}}=\nabla_t E_z$. Using these relations, we find:

$$\nabla_t H_z-ik_z\mathbf{H}_t=-i\omega\epsilon\hat{\mathbf{z}}\times\mathbf{E}_t \qquad \nabla_t E_z-ik_z\mathbf{E}_t=i\omega\mu\hat{\mathbf{z}}\times\mathbf{H}_t$$

Substitute in back and forth to decouple these equations:

$$\mathbf{H}_t=\frac{i}{\mu\epsilon\omega^2-k_z^2}\left[k_z\nabla_t H_z+\omega\epsilon\hat{\mathbf{z}}\times\nabla_t E_z\right] \qquad \mathbf{E}_t=\frac{i}{\mu\epsilon\omega^2-k_z^2}\left[k_z\nabla_t E_z-\omega\mu\hat{\mathbf{z}}\times\nabla_t H_z\right]$$

Define the variable $\kappa^2=\epsilon\mu\omega^2-k_z^2$. This definition will become clear a few steps later. This variable will be determined by the boundary conditions in terms of the waveguide's material and geometry. Using this definition, the equations become:

$$\boxed{\mathbf{H}_t=\frac{i}{\kappa^2}\left[k_z\nabla_t H_z+\omega\epsilon\hat{\mathbf{z}}\times\nabla_t E_z\right]} \qquad \boxed{\mathbf{E}_t=\frac{i}{\kappa^2}\left[k_z\nabla_t E_z-\omega\mu\hat{\mathbf{z}}\times\nabla_t H_z\right]}$$

We now have equations that let us calculate directly the transverse fields if we know the parallel fields. The problem is therefore reduced to only needing to solve for the transverse or the parallel fields.

Now take the curl of Faraday's law and substitute in the Maxwell-Ampere Law without sources. Also take the curl of the Maxwell-Ampere Law and insert Faraday's law into it to find:

$$\nabla\times\nabla\times\mathbf{E}=-\epsilon\mu\frac{\partial^2}{\partial t^2}\mathbf{E} \qquad \nabla\times\nabla\times\mathbf{H}=-\epsilon\mu\frac{\partial^2}{\partial t^2}\mathbf{H}$$

Use the vector identity $\nabla\times(\nabla\times\mathbf{A})=\nabla(\nabla\cdot\mathbf{A})-\nabla^2\mathbf{A}$ and realize that both the electric and magnetic field have no divergence in the absence of sources, so that the first term in the expansion drops out. This leaves us with the wave equations:

$$\nabla^2\mathbf{E}=\epsilon\mu\frac{\partial^2}{\partial t^2}\mathbf{E} \qquad \nabla^2\mathbf{H}=\epsilon\mu\frac{\partial^2}{\partial t^2}\mathbf{H}$$

Inside the waveguide, we have assumed harmonic time and *z* dependence, which makes the wave equations look like:

$$\left[\nabla_t^2 + \epsilon\mu\omega^2 - k_z^2\right]\mathbf{E} = 0 \qquad \left[\nabla_t^2 + \epsilon\mu\omega^2 - k_z^2\right]\mathbf{H} = 0$$

Now we see in these equations the variable which was defined above as $\kappa^2 = \epsilon\mu\omega^2 - k_z^2$ and recognize it as the transverse wave number $\kappa^2 = k_x^2 + k_y^2$. Solving for the frequency we recognize $\epsilon\mu\omega^2 = \kappa^2 + k_z^2$ and $\epsilon\mu\omega^2 = k_x^2 + k_y^2 + k_z^2$ as we would expect. With this definition, the wave equations become:

$$\left[\nabla_t^2 + \kappa^2\right]\mathbf{E} = 0 \qquad \left[\nabla_t^2 + \kappa^2\right]\mathbf{H} = 0$$

The wave equations apply separately to each component of the vector fields. If we are solving for the parallel components they become

$$\left[\nabla_t^2 + \kappa^2\right]E_z = 0 \qquad \left[\nabla_t^2 + \kappa^2\right]H_z = 0$$

whereas for the transverse components they become

$$\left[\nabla_t^2 + \kappa^2\right]\mathbf{E}_t = 0 \qquad \left[\nabla_t^2 + \kappa^2\right]\mathbf{H}_t = 0$$

The general solution to any component i in *rectangular* coordinates in any one particular region of uniform linear material is:

$$E_i = \sum_{k_x, k_y} (A e^{i k_x x} + B e^{-i k_x x})(C e^{i k_y y} + D e^{-i k_y y}) \quad \text{where} \quad \kappa^2 = k_x^2 + k_y^2$$

$$H_i = \sum_{k_x, k_y} (E e^{i k_x x} + F e^{-i k_x x})(G e^{i k_y y} + H e^{-i k_y y}) \quad \text{where} \quad \kappa^2 = k_x^2 + k_y^2$$

Boundary conditions must be applied at each boundary that connects two regions of uniform linear material in order to determine the coefficients and the wavenumbers. The boundary conditions will cause the span of possible wave numbers k_x and k_y to form a discrete set called modes. The lowest-order modes are typically the ones excited first and are the ones of most interest.
The parallel electric and magnetic field components are typically separated and treated as separate modes. By setting $B_z = 0$ in all of the above equations, we get the Transverse Magnetic (TM) modes and by setting $E_z = 0$ we get the Transverse Electric (TE) modes.

TM: $\mathbf{E}_t = \frac{i k_z}{\kappa^2}\nabla_t E_z \quad \mathbf{H}_t = \frac{i\omega\epsilon}{\kappa^2}\hat{\mathbf{z}}\times\nabla_t E_z$ TE: $\mathbf{H}_t = \frac{i k_z}{\kappa^2}\nabla_t H_z \quad \mathbf{E}_t = -\frac{i\omega\mu}{\kappa^2}\hat{\mathbf{z}}\times\nabla_t H_z$

We can combine each set of equations to relate the transverse fields:

TM: $\boxed{\mathbf{H}_t = \frac{\omega \epsilon}{k_z} \hat{\mathbf{z}} \times \mathbf{E}_t}$ TE: $\boxed{\mathbf{E}_t = -\frac{\omega \mu}{k_z} \hat{\mathbf{z}} \times \mathbf{H}_t}$

In summary, the general rectangular waveguide equations are:

General Rectangular Waveguide Equation Summary: TM Modes ($H_z = 0$)

Transverse to Parallel: $\boxed{E_z = \frac{i}{k_z}\left[\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y}\right]}$ $\boxed{\frac{\partial H_x}{\partial x} = -\frac{\partial H_y}{\partial y}}$ Parallel to Transverse: $\boxed{\mathbf{E}_t = \frac{i k_z}{\kappa^2} \nabla_t E_z}$ $\boxed{\mathbf{H}_t = \frac{i \omega \epsilon}{\kappa^2} \hat{\mathbf{z}} \times \nabla_t E_z}$

Wave Equations: Solve $\boxed{E_i = \sum_{k_x, k_y} (A e^{i k_x x} + B e^{-i k_x x})(C e^{i k_y y} + D e^{-i k_y y})}$ or $\boxed{H_i = \sum_{k_x, k_y} (E e^{i k_x x} + F e^{-i k_x x})(G e^{i k_y y} + H e^{-i k_y y})}$

where $\kappa^2 = \epsilon \mu \omega^2 - k_z^2$ and $\kappa^2 = k_x^2 + k_y^2$ and i denotes any component x, y, z

Magnetic-Electric Field Relations $\boxed{\mathbf{H}_t = \frac{\omega \epsilon}{k_z} \hat{\mathbf{z}} \times \mathbf{E}_t}$

Note that all of these equations apply only in regions of uniform linear materials. Separate regions must be linked by boundary conditions.

General Rectangular Waveguide Equation Summary: TE Modes ($E_z = 0$)

Transverse to Parallel: $\boxed{H_z = \frac{i}{k_z}\left[\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y}\right]}$ $\boxed{\frac{\partial E_x}{\partial x} = -\frac{\partial E_y}{\partial y}}$ Parallel to Transverse: $\boxed{\mathbf{H}_t = \frac{i k_z}{\kappa^2} \nabla_t H_z}$ $\boxed{\mathbf{E}_t = -\frac{i \omega \mu}{\kappa^2} \hat{\mathbf{z}} \times \nabla_t H_z}$

Wave Equations: Solve $\boxed{E_i = \sum_{k_x, k_y} (A e^{i k_x x} + B e^{-i k_x x})(C e^{i k_y y} + D e^{-i k_y y})}$ or $\boxed{H_i = \sum_{k_x, k_y} (E e^{i k_x x} + F e^{-i k_x x})(G e^{i k_y y} + H e^{-i k_y y})}$

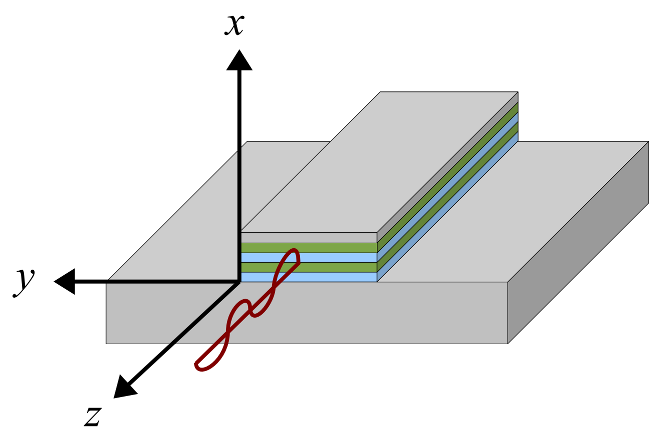
where $\kappa^2 = \epsilon \mu \omega^2 - k_z^2$ and $\kappa^2 = k_x^2 + k_y^2$ and i denotes any component x, y, z

Magnetic-Electric Field Relations $\boxed{\mathbf{E}_t = -\frac{\omega\mu}{k_z}\hat{\mathbf{z}}\times\mathbf{H}_t}$

Note that all of these equations apply only in regions of uniform linear materials. Separate regions must be linked by boundary conditions.

5.0 QCL Slab Waveguide Equations

A QCL's active region consists of a sequence of planar epitaxial layers grown on top of each other in the *x* direction. A quantum selection rule dictates that all coherent radiation generated by any QCL is polarized such that the electric field points in the *x* direction, normal to the epitaxial layers. The active region structure thus automatically dictates that $E_y = 0$ and $H_x = 0$. Plugging these in the general equations yields:



QCL TM Modes

Transverse to Parallel: $H_z = 0$ $\quad E_z(x) = \frac{i}{k_z}\frac{\partial E_x(x)}{\partial x}$

Parallel to Transverse: $E_x(x) = \frac{i k_z}{\kappa^2}\frac{\partial E_z(x)}{\partial x}$ $\quad E_y = 0$ $\quad H_x = 0$ $\quad H_y(x) = \frac{i\omega\epsilon}{\kappa^2}\frac{\partial E_z(x)}{\partial x}$

Wave Equations: $E_x = \sum_{k_x}(A e^{i k_x x} + B e^{-i k_x x})$ or $E_z = \sum_{k_x}(C e^{i k_x x} + D e^{-i k_x x})$ or $H_y = \sum_{k_x}(E e^{i k_x x} + F e^{-i k_x x})$ where $k_x^2 = \epsilon\mu\omega^2 - k_z^2$

Magnetic-Electric Field Relations $H_y(x) = \frac{\omega\epsilon}{k_z} E_x(x)$

QCL TE Modes

Transverse to Parallel: $E_z=0 \quad H_z(y)=\frac{i}{k_z}\frac{\partial H_y(y)}{\partial y}$

Parallel to Transverse: $E_x(y)=\frac{i\omega\mu}{\kappa^2}\frac{\partial H_z(y)}{\partial y} \quad E_y=0 \quad H_x=0 \quad H_y(y)=\frac{ik_z}{\kappa^2}\frac{\partial H_z(y)}{\partial y}$

Wave Equations: $E_x=\sum_{k_y}(Ae^{ik_y y}+Be^{-ik_y y})$ or $H_y=\sum_{k_y}(Ce^{ik_y y}+De^{-ik_y y})$ or $H_z=\sum_{k_y}(Ge^{ik_y y}+He^{-ik_y y})$ where $k_y^2=\epsilon\mu\omega^2-k_z^2$

Magnetic-Electric Field Relations $E_x(y)=\frac{\omega\mu}{k_z}H_y(y)$

5.1 Slab Model

Often the width of the QCL is much larger than the height of the QCL. As an approximation we can thus assume that the waveguide is infinite and uniform in the *y* direction. We therefore assume that none of the fields have a *y*-dependence. This automatically forbids TE modes. The fields in the TM modes are already all independent of *y*.

For the TM modes, we have a choice of three approaches to solving the problem, either solve for E_x, E_z or H_y. Since the magnetic field only has one component, the math should be simplest if we choose to solve for H_y. All of the relevant equations are then:

Solve $H_y=\sum_{k_x}(Ae^{ik_x x}+Be^{-ik_x x})$ by applying boundary conditions.

Then use $E_x(x)=\frac{k_z}{\omega\epsilon}H_y(x)$ and $E_z(x)=\frac{i}{\omega\epsilon}\frac{\partial}{\partial x}H_y(x)$ remembering $E_y=0$, $H_x=0$, and $H_z=0$.

We can carry these out into explicit form because we already know the solution:

$$E_x(x)=\sum_{k_x}\frac{k_z}{\omega\epsilon}(Ae^{ik_x x}+Be^{-ik_x x}) \qquad E_z(x)=\sum_{k_x}\frac{k_x}{\omega\epsilon}(-Ae^{ik_x x}+Be^{-ik_x x})$$

The final solutions are then $\mathbf{H}=\hat{\mathbf{y}}\, H_y(x)\, e^{ik_z z-i\omega t}$ and $\mathbf{E}=(\hat{\mathbf{x}}\, E_x(x)+\hat{\mathbf{z}}\, E_z(x))e^{ik_z z-i\omega t}$ where $k_x^2=\epsilon\mu\omega^2-k_z^2$.

In the one-dimensional QCL waveguide approach, all of the material boundaries are planes parallel to the *y-z* plane. The applicable boundary conditions then are that the tangential components of the magnetic field **H** must be continuous and the tangential components of the electric field **E** field must be continuous across the boundary. Let us state this more formally. Each region of linear uniform material has its solutions with its own **H** and **E** as a function of its own k_x, A, B, ε, and μ. Let us use the index i to denote the ith region so that all these properties in the ith region are denoted $\mathbf{H}_i$ and $\mathbf{E}_i$, $k_{x,i}$, A_i, B_i, ε_i, and μ_i. Note that the boundary conditions require that all the regions have the same z-directional wave number k_z and frequecy ω. Let the i=0th region be the semi-infinite substrate at the bottom of the stack, the i=1st region be the one directly above the substrate, and so on. Also denote the known location of the boundaries between regions as x_i where $x_0 = 0$ is the origin of the x coordinate and is also the location of the zeroth boundary, the one between the substrate and the next layer. With these definitions, the boundary conditions become:

$\hat{\mathbf{x}}\times\mathbf{H}_i(x_i)=\hat{\mathbf{x}}\times\mathbf{H}_{i+1}(x_i)$ where i = 0,1... N-1 where N is the number of layers not including the substrate

$$H_{y,i}(x_i)=H_{y,i+1}(x_i)$$

$$\boxed{A_i e^{ik_{x,i}x_i}+B_i e^{-ik_{x,i}x_i}=A_{i+1}e^{ik_{x,i+1}x_i}+B_{i+1}e^{-ik_{x,i+1}x_i}}$$

and the other boundary condition is:

$$\hat{\mathbf{x}}\times\mathbf{E}_i(x_i)=\hat{\mathbf{x}}\times\mathbf{E}_{i+1}(x_i)$$

$$E_{z,i}(x_i)=E_{z,i+1}(x_i)$$

$$\boxed{\frac{k_{x,i}}{\epsilon_i}(A_i e^{ik_{x,i}x_i}-B_i e^{-ik_{x,i}x_i})=\frac{k_{x,i+1}}{\epsilon_{i+1}}(A_{i+1}e^{ik_{x,i+1}x_i}-B_{i+1}e^{-ik_{x,i+1}x_i})}$$

6.0 Formulating for Numerical Analysis

The problem is to solve for A_i and B_i in each region in terms of the known permittivities ε_i, boundary locations x_i, frequency ω and the guessed wave number k_z. The wave number k_z is guessed and refined until the outermost boundary conditions are met, that of fields approaching zero at positive and negative infinity for x.

Add the two boundary condition equations together to eliminate B_{i+1} and solve for A_{i+1}:

$$\boxed{A_{i+1}=\frac{1}{2}\left(1+\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)A_i e^{i(k_{x,i}-k_{x,i+1})x_i}+\frac{1}{2}\left(1-\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)B_i e^{-i(k_{x,i}+k_{x,i+1})x_i}}$$

Subtract the two equations to eliminate A_{i+1} and solve for B_{i+1}:

$$\boxed{B_{i+1}=\frac{1}{2}\left(1-\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)A_i e^{i(k_{x,i}+k_{x,i+1})x_i}+\frac{1}{2}\left(1+\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)B_i e^{-i(k_{x,i}-k_{x,i+1})x_i}}$$

We now have iteration equations. If we have a trial k_z, then because we know ω, ε_i, and μ_i we also know $k_{x,i}$ using $k_{x,i}=\sqrt{\epsilon_i\mu_i\omega^2-k_z^2}$. Thus if we know A_0 and B_0, we can find all other A_i and B_i using these iteration equations.
We are interested in bound modes, so that the wave must die down to zero at negative and positive infinity. This leads to the conditions:

$$A_0=0 \;\text{ if }\; \Im(k_{x,0})>0 \;\text{ or }\; B_0=0 \;\text{ if }\; \Im(k_{x,0})<0$$

$$A_N=0 \;\text{ if }\; \Im(k_{x,N})<0 \;\text{ or }\; B_N=0 \;\text{ if }\; \Im(k_{x,N})>0$$

There is an overall normalization constant for the mode that can be found and applied after the problem is solved. We take this into account by setting the non-zero coefficient of the zeroth layer equal to one and normalizing it at the end to its proper value.

$$A_0=1 \;\text{ if }\; \Im(\kappa_0)<0 \;\text{ or }\; B_0=1 \;\text{ if }\; \Im(\kappa_0)>0$$

The system is solved iteratively for the propagation constant k_z. First a guess for k_z is made (or a grid of guesses), then the iteration equations are applied to find the fields everywhere corresponding to this k_z. The amount that the coefficients A_N and B_N in the last region differ from what they should be as dictated by the boundary conditions is calculated as the error corresponding to that k_z. The right k_z is found by refining its value until the error is minimized. The lowest-order mode is the one of most importance and it is the one that is found. Note that k_z is complex-valued so that finding the physical modes amounts to finding the minima in a two-dimensional error landscape.

6.1 Final Numerical Recipe

A. Calculate the complex permittivity of each layer using the Drude model (see above) and assume non-magnetic material $\mu_i = \mu_0$.

B. Sweep across a grid of initial possible values for the real and imaginary part of the wave vector k_z

1. For each complex k_z value, calculate the complex x-directional wave number of each waveguide layer using $k_{x,i}=\sqrt{\epsilon_i \mu_i \omega^2 - k_z^2}$
2. Set A_0 and B_0 according to the following rule: if $\Im(k_{x,0})>0$ then $A_0=0$ and $B_0=1$, else $A_0=1$ and $B_0=0$
3. Calculate all remaining A_i and B_i using

$$A_{i+1}=\frac{1}{2}\left(1+\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)A_i e^{i(k_{x,i}-k_{x,i+1})x_i}+\frac{1}{2}\left(1-\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)B_i e^{-i(k_{x,i}+k_{x,i+1})x_i}$$

$$B_{i+1}=\frac{1}{2}\left(1-\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)A_i e^{i(k_{x,i}+k_{x,i+1})x_i}+\frac{1}{2}\left(1+\frac{k_{x,i}\epsilon_{i+1}}{k_{x,i+1}\epsilon_i}\right)B_i e^{-i(k_{x,i}-k_{x,i+1})x_i}$$

4. Find the error as the difference between the calculated A_N or B_N and the required A_N or B_N according to:
   if $\Im(k_{x,N})<0$ then $A_N=0$ else $B_N=0$

C. Find all mimima in the 2D error landscape as a function of k_z. A minima is a grid point where its error value is lower than the error value of its eight nearest neighbors.

D. Refine every minimum k_z by repeating step B for additional guesses and keeping the ones that give the least error. Use the method of gradient descent: step iteratively in the opposite direction of the gradient of the error landscape until the desired precision is met according to: $[k_z]_{n+1}=[k_z]_n-\gamma_n \nabla f([k_z]_n)$ where $f(k_z)$ is the 2D functional error landscape as a function of k_z.
Due to the complexity of the error landscape, the gradient cannot be solved analytically but the finite difference can be used. The step size γ_n is optimized by finding the first value that minimizes $f([k_z]_{n+1})$ locally.

E. Choose the k_z which corresponds to the lowest-order mode. It is the one with the lowest value for (k_x/k_z) in the largest layer.

F. Once the lowest order k_z has been identified and refined, use step B again to get the final values for A_i and B_i

G. Using the final values of A_i and B_i calculate all the fields across a fine grid of x locations using:

$$H_{y,i}(x)=A_i e^{ik_{x,i}x}+B_i e^{-ik_{x,i}x} \qquad E_{x,i}(x)=\frac{k_z}{\omega\epsilon_i}(A_i e^{ik_{x,i}x}+B_i e^{-ik_{x,i}x}) \qquad E_{z,i}(x)=\frac{k_{x,i}}{\omega\epsilon_i}(-A_i e^{ik_{x,i}x}+B_i e^{-ik_{x,i}x})$$

$$|\mathbf{H}_i(x)|^2=|H_{y,i}(x)|^2 \qquad |\mathbf{E}_i(x)|^2=|E_{x,i}(x)|^2+|E_{z,i}(x)|^2$$

H. Calculate the confinement factor for this waveguide structure at this frequency using the trapezoidal or cubic splines method:

$$\Gamma=\frac{\int_{\text{act. reg.}}|E_{x,i}(x)|^2 dx}{\int_{-\infty}^{\infty}|\mathbf{E}_i(x)|^2 dx}$$

I. Calculate associated parameters:

Waveguide Loss: $\alpha_w=2\Im(k_z)$

Total Cavity Loss: $\alpha=\alpha_w+\alpha_{M1}+\alpha_{M2}$

Threshold Gain: $g_{th}=\alpha/\Gamma$

Velocity: $v=\frac{\omega}{\Re(k_z)}$

Effective Index of Refraction: $n=\frac{c}{v}$

Total Cavity Photon loss rate: $W_p=\alpha v$

Total Cavity Photon lifetime: $\tau_p=1/W_p$

J. Repeat Steps A through I for many possible frequencies in order to generate look-up tables.

6.2 Cutoff Frequency

Upon generating the look-up tables, which essentially establish loss vs. frequency and confinement factor vs. frequency trendlines, a problem arises. Below a certain frequency, the cutoff frequency, electromagnetic modes cannot propagate. In terms of the practical requirements of the rest of the code, the waveguide code should return a confinement factor of zero and a loss of infinity for these frequencies. However, the waveguide code is designed to find modes as minima in an error landscape, and cannot find the lack of modes. As implemented above, the waveguide code will output erroneous values for frequencies below cutoff, and the rest of the code will use these values as if valid, leading to widespread error.

The solution is to have the code determine the cutoff frequency and then hard-set all confinement factors to zero and all losses to infinity below this frequency. Note that if we set the confinement factor to perfectly zero, then we will end up with divide-by-zero errors later on in the rate equations. In practice, we must set each confinement factor below cutoff to a very small non-zero number; so small that it is essentially zero. Similarly, we cannot actually set the losses below cutoff to infinity in a numerical code. We instead set each loss equal to a very high number that essentially behaves as infinity.

The cutoff frequency is found by starting at a high frequency on the confinement factor vs. frequency curve, and asymptotically finding at what frequency the curve approaches zero confinement.