coupled mode theory for weakly-guiding, paraxial, and slowly varying waveguides

Jonathan Lin

January 2024

1 derivation of the coupled mode equations

1.1 background

We apply coupled-mode theory to approximately solve the propagation equation for slowly-varying waveguides. The principle of coupled-mode theory is as follows: suppose in some ideal scenario our physical system supports a basis of ideal eigenmodes which take the form of harmonic oscillators. When a slowly-varying perturbation is added, we assume that perturbed solutions will be "near harmonic", and thus attempt to find new solutions formed from a linear combination of the "instantaneous", perturbed eigenmode basis (which, for waveguides like directional couplers, is often called the "supermode" basis). To account for the perturbation, we allow the modes to cross-couple through mode coefficients that change with longitudinal propagation distance. Neglecting these cross-coupling coefficients is equivalent to applying the adiabatic approximation (e.g. WKB from QM); the adiabatic approximation decouples the evolution equations for all waveguide eigenmode amplitudes and requires that the waveguide modes are never degenerate.

1.2 review: ideal, z-invariant waveguides

We define z to be the propagating direction. We also use bra-ket notation, defining a "transverse" inner product space with

$$\langle u|v\rangle = \iint \bar{u}v\, dxdy \tag{1}$$

where $\bar{u}$ is the complex conjugate of u. Beam propagation under the approximation of weak guidance obeys the Helmholtz equation: for a field function $\psi(x, y, z)$ or equivalently a field ket $|\psi\rangle$ in the above inner product space, we have

$$\left[\nabla^2 + k^2 n_0^2\right]|\psi\rangle = 0 \tag{2}$$

where k is the free space wavenumber and $n_0 = n_0(x, y)$ is the refractive index profile. Under the paraxial approximation, we assume solutions of the form

$$|\psi_j\rangle = |\phi_j\rangle e^{i\beta_j z} \tag{3}$$

where $|\phi\rangle$ describes a transverse wavefront that is assumed to be "slowly-varying" with z, and β_j is some "propagation constant" (which will turn out to be the square root of the eigenvalue for eigenmode j). Making the paraxial approx., higher-order derivatives w.r.t. to z can be ignored and we are left with the so-called "paraxial Helmholtz equation":

$$\left[\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + k^2 n_0^2 - \beta_j^2\right]|\phi_j\rangle = 2i\beta_j \frac{\partial}{\partial z}|\phi_j\rangle. \tag{4}$$

For eigenmodes whose shape are maintained throughout propagation, the RHS is 0 and thus the mode shapes solve the following eigenvalue problem:

$$\left[\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + k^2 n_0^2(x, y)\right]|\phi_j\rangle = \beta_j^2|\phi_j\rangle. \tag{5}$$

1.3 slowly-varying waveguides

We now consider the following perturbation:

$$n_0^2(x,y) \to k^2 n^2(x,y,z) = n_0^2(x,y) + \delta\epsilon(x,y,z) \tag{6}$$

where $\delta\epsilon(x,y,z)$ is a weak function of x,y (in accordance with weak guidance) and z. We will establish what this latter condition on z means precisely later on.

Denote the "instantaneous" eigenmodes $|\xi_j^z\rangle$ such that

$$\left[\nabla_\perp^2 + k^2 n^2(x,y,z)\right] |\xi_j^z\rangle = \beta_j^2(z)\, |\xi_j^z\rangle. \tag{7}$$

where the z superscripts denote evaluation at z and $\nabla_\perp$is the transverse Laplacian. We propose new solutions expressed in the basis of

$$|\Psi_j\rangle \equiv e^{i\int_0^z \beta_j(z')dz'} |\xi_j^z\rangle. \tag{8}$$

A general ansatz for the perturbed Helmholtz equation is:

$$|\Psi\rangle = \sum_j a_j(z)|\Psi_j\rangle = \sum_j a_j(z) e^{i\int_0^z \beta_j(z')dz'} |\xi_j^z\rangle \tag{9}$$

where the mode coefficients $a_j(z)$ are for now unconstrained. Importantly, we allow the coefficients $a_j(z)$ to vary; we are *not* claiming that the states $|\Psi_j\rangle$ solve the Helmholtz equation, but rather that some linear combination might approximate the true solution. Plugging back into the Helmholtz equation gives

$$\begin{aligned} \left[\nabla_\perp^2 + \frac{\partial^2}{\partial z^2} + k^2 n^2\right] |\Psi\rangle = 0 \\ \left[\nabla_\perp^2 + \frac{\partial^2}{\partial z^2} + k^2 n^2\right] \sum_j a_j(z)|\Psi_j\rangle = 0 \\ \sum_j \left[\beta_j^2 + \frac{\partial^2}{\partial z^2}\right] a_j(z)|\Psi_j\rangle = 0. \end{aligned} \tag{10}$$

I'd like to point out briefly that the first line can be written as

$$\frac{\partial^2}{\partial z^2}|\Psi\rangle = -A|\Psi\rangle \tag{11}$$

for $A \equiv \nabla_\perp^2 + k^2 n^2$, which *almost* looks like the usual form of the equation assumed in coupled-mode theory (e.g. the time-dependent Schrodinger equation), except we have a second-order derivative instead of first-order (so, more reminiscent of the type of differential equation solved by the WKB approximation). Expanding out fully and applying the product rule, we have

$$\begin{aligned} 0 = \sum_j & 2e^{i\int_0^z \beta_j(z')dz'} \frac{\partial a_j}{\partial z} \left[i\beta_j + \frac{\partial}{\partial z}\right] |\xi_j^z\rangle \\ & + i a_j e^{i\int_0^z \beta_j dz'} \left[\frac{\partial \beta_j}{\partial z}|\xi_j^z\rangle + 2\beta_j(z)\frac{\partial}{\partial z}|\xi_j^z\rangle\right] \\ & + \frac{\partial^2 a_j}{\partial z^2}|\xi_j^z\rangle + a_j e^{i\int_0^z \beta_j(z')dz'} \frac{\partial^2}{\partial z^2}|\xi_j^z\rangle. \end{aligned} \tag{12}$$

Multiply through by $\langle\Psi_i|$ to get:

$$\begin{aligned} 0 = {} & 2i\beta_i \frac{\partial a_i}{\partial z} + i a_i \frac{\partial \beta_i}{\partial z} \\ & + \sum_j e^{i\int_0^z (\beta_j - \beta_i)dz'} \left[2\frac{\partial a_j}{\partial z} + 2i a_j \beta_j\right] \langle\xi_i^z|\frac{\partial}{\partial z}|\xi_j^z\rangle \\ & + \frac{\partial^2 a_i}{\partial z^2} + \sum_j a_j e^{i\int_0^z (\beta_j - \beta_i)dz'} \langle\xi_i^z|\frac{\partial^2}{\partial z^2}|\xi_j^z\rangle \end{aligned} \tag{13}$$

We now successively apply approximations to the above "coupled-mode" equations.

1.3.1 approx. 1: paraxial

The paraxial equation is usually stated as:

$$\left|\frac{\partial^2 \xi_j}{\partial z^2}\right| \ll \beta_j \left|\frac{\partial \xi_j}{\partial z}\right|. \tag{14}$$

This approximation removes all second-order derivatives of $|\xi\rangle$, giving:

$$\begin{aligned} 0 &= 2i\beta_i \frac{\partial a_i}{\partial z} + i a_i \frac{\partial \beta_i}{\partial z} \\ &+ \sum_j e^{i\int_0^z (\beta_j - \beta_i) dz'} \left[2\frac{\partial a_j}{\partial z} + 2i a_j \beta_j\right] \langle \xi_i^z | \frac{\partial}{\partial z} | \xi_j^z \rangle + \frac{\partial^2 a_i}{\partial z^2} \end{aligned} \tag{15}$$

1.3.2 approx. 2: slow envelope

In the "slowly varying envelope approximation", we claim:

$$\left|\frac{\partial^2 a_j}{\partial z^2}\right| \ll \beta_j \left|\frac{\partial a_j}{\partial z}\right| \ll \beta_j^2 |a_j|. \tag{16}$$

This condition is the same as claiming that the cross-coupling between instantaneous eigenmodes is weak and as such the modal coefficients a_j vary slowly compared to $\exp(i\beta_j z)$. This approximation removes removes the first term of the middle line in equation 15, and the remaining second-order derivative term. In the following, I also divide out i.

$$\begin{aligned} 0 &= 2\beta_i \frac{\partial a_i}{\partial z} + a_i \frac{\partial \beta_i}{\partial z} \\ &+ 2\sum_j e^{i\int_0^z (\beta_j - \beta_i) dz'} a_j \beta_j \langle \xi_i^z | \frac{\partial}{\partial z} | \xi_j^z \rangle. \end{aligned} \tag{17}$$

I believe the above form is the most general under the paraxial and slow envelope approximations. Note that the second term on the first line (the "WKB" term which we discuss in the next section) is diagonal. Cross-coupling between instantaneous eigenmodes occurs via the term on the second line. We often define the "cross-coupling matrix" κ_{ij} as something like

$$\kappa_{ij} \equiv \langle \xi_i^z | \frac{\partial}{\partial z} | \xi_j^z \rangle \tag{18}$$

though the exact definition from text to text might vary. As defined above, and assuming real-valued eigenmodes (which we can do without loss of generality, at least under prior approximations), the cross-coupling matrix is antisymmetric. Proof:

$$\begin{aligned} \kappa_{ij} &= \int \xi_i(x,y,z) \frac{\partial}{\partial z} \xi_j(x,y,z)\, dxdy \\ &= \int \frac{\partial}{\partial z} [\xi_i \xi_j]\, dxdy - \int \xi_j \frac{\partial}{\partial z} \xi_i\, dxdy \\ &= \frac{\partial}{\partial z} \int \xi_i \xi_j\, dxdy - \int \xi_j \frac{\partial}{\partial z} \xi_i\, dxdy \\ &= \frac{\partial}{\partial z} \delta_{ij} - \int \xi_j \frac{\partial}{\partial z} \xi_i\, dxdy \\ &= -\int \xi_j \frac{\partial}{\partial z} \xi_i\, dxdy = -\kappa_{ji}. \end{aligned} \tag{19}$$

This also means that the cross-coupling matrix has a 0 diagonal; $\kappa_{ii} = 0$. In the case of complete degeneracy and constant β, it can be shown that the solution will be of the form $\vec{a} \propto U(z)\vec{a}_0$ for a unitary matrix U

which is a function of z, since the matrix exponential of an anti-Hermitian matrix is unitary. In this case, power is clearly preserved.

For cross-coupling occur, we need both κ_{ij} to be non-zero and $\beta_j - \beta_i$ to be small (otherwise, the sign of the cross-coupling will rapidly flip with z and very little net power will be transferred).

1.3.3 approx 3: adiabatic approximation

In the full adiabatic approximation, we assume that the cross-coupling term is 0; thus, the modes are decoupled as

$$0 = 2\beta_i \frac{\partial a_i}{\partial z} + a_i \frac{\partial \beta_i}{\partial z}. \tag{20}$$

If the β_i are constant, as in a straight fiber, we have

$$\frac{\partial a_i}{\partial z} = 0 \tag{21}$$

as expected. More generally, the solution for the mode amplitudes a_i is of the form

$$a_i(z) = a_i(0)\sqrt{\frac{\beta_i(0)}{\beta_i(z)}} \tag{22}$$

and the full solution for a propagating wavefront is

$$|\Psi(z)\rangle = \sum_j a_j(0)\sqrt{\frac{\beta_j(0)}{\beta_j(z)}} e^{i\int_0^z \beta_j(z')dz'} |\xi_j^z\rangle \tag{23}$$

which is reminiscent of the 1st order WKB approximation for a particle in a slowly-varying potential. For an overview of the WKB approximation, see appendix A.

To derive when the adiabatic approximation holds, we differentiate the "transverse" eigenmode equation with respect to z and reduce to matrix form:

$$\begin{aligned}
A &\equiv \nabla_\perp^2 + k^2 n^2(x,y,z) \\
A|\xi_j^z\rangle &= \beta_j^2(z)|\xi_j^z\rangle \\
\langle\xi_i^z|\frac{\partial}{\partial z}\left[A|\xi_j^z\rangle\right] &= \langle\xi_i^z|\frac{\partial}{\partial z}\left[\beta_j^2(z)|\xi_j^z\rangle\right] \\
\langle\xi_i^z|\frac{\partial A}{\partial z}|\xi_j^z\rangle + \langle\xi_i^z|A\frac{\partial}{\partial z}|\xi_j^z\rangle &= \beta_j^2\langle\xi_i^z|\frac{\partial}{\partial z}|\xi_j^z\rangle \ ; \ i \neq j \\
\langle\xi_i^z|\frac{\partial A}{\partial z}|\xi_j^z\rangle + \beta_i^2\langle\xi_i^z|\frac{\partial}{\partial z}|\xi_j^z\rangle &= \beta_j^2\langle\xi_i^z|\frac{\partial}{\partial z}|\xi_j^z\rangle \\
\langle\xi_i^z|\frac{\partial}{\partial z}|\xi_j^z\rangle &= \frac{\langle\xi_i^z|\frac{\partial A}{\partial z}|\xi_j^z\rangle}{\beta_j^2 - \beta_i^2}
\end{aligned} \tag{24}$$

so the adiabatic approximation blows up as modes approach degeneracy. In other words, no rate of change is "slow enough" to prevent cross-coupling when considering degenerate modes. We also note

$$\frac{\partial A}{\partial z} = k^2 \frac{\partial}{\partial z} n^2(x,y,z) = k^2 \frac{\partial}{\partial z}\delta\epsilon \tag{25}$$

so the magnitude of the cross-coupling terms go to 0 as the z-derivative of the index perturbation goes to 0, as expected.

1.3.4 approx. 0: weak guidance

The weak guidance approximation asserts that the index contrast of the waveguide is small, and converts Maxwell's equations into a set of decoupled and identical scalar Helmholtz equations, which we assumed outright at the beginning of this document. It can also be shown (I think) that the propagation constants of guided modes obey

$$k \min(n) \leq \beta \leq k \max(n). \tag{26}$$

As a corollary, if the minimum and maximum refractive indices do not change significantly over the the length of the waveguide, we can ignore the WKB-like correction term in equation 17, leaving

$$0 = \beta_i \frac{\partial a_i}{\partial z} + \sum_j e^{i\int_0^z (\beta_j - \beta_i)dz'} a_j \beta_j \langle \xi_i^z | \frac{\partial}{\partial z} | \xi_j^z \rangle. \tag{27}$$

1.4 mode degeneracy

Even when continuously evolving a mesh, modes that are near-degenerate (note that numerically computed modes will typically never be completely degenerate due to asymmetries induced by the mesh) can still "rotate" with z, sometimes rapidly. If we can identify groups of eigenmodes that remain so close to degenerate throughout the entire waveguide that we can treat them as always degenerate, there is a simple solution. When going from $z \to z + \delta z$, we apply a unitary transformation to the degenerate modes at $z + \delta z$ to "match" the modes at z. Mathematically, we search for a unitary transformation Q that minimizes the difference

$$Q|\xi_k^{z+\delta z}\rangle - |\xi_k^z\rangle \tag{28}$$

where k indexes the modes in the degenerate subspace (specifically, we'd like to minimize the square of the residuals, as written above, summed over k). Let's change to matrix form by multiplying through with $\langle \xi_n^{z+\delta z}|$ as follows:

$$\begin{aligned} &\langle \xi_n^{z+\delta z}|Q|\xi_k^{z+\delta z}\rangle - \langle \xi_n^{z+\delta z}|\xi_k^z\rangle \\ &\equiv Q_{nk} - C_{nk}. \end{aligned} \tag{29}$$

Thus, to minimize the above while maintaining unitary of Q, we should set

$$Q = UV^T \tag{30}$$

where U and V are given by the SVD of C:

$$C = USV^T. \tag{31}$$

Denote instantaneous basis of degenerate eigenmodes computed at z as the matrix $M_{jk}(z)$ where j iterates through spatial positions (e.g. FE mesh nodes) and k iterates through the degenerate subspace of modes. At $z + \delta z$, we should choose a new basis of degenerate eigenmodes $M'(z + \delta z)$ such that

$$\begin{aligned} M'_{jk}(z+\delta z) &= \sum_i Q_{ki} M_{ji}(z) = \sum_i M_{ji}(z) Q^T_{ik} \\ M' &= MQ^T \end{aligned} \tag{32}$$

where

$$Q^T = Q^{-1} = VU^T. \tag{33}$$

Even when degeneracy correction may not be strictly necessary for an accurate propagation result, correction can still speed up adaptively-stepped computations, since the cross-coupling matrix will vary less strongly with z.

1.5 finite difference estimate of cross-coupling matrix

At some given z, denote v_i^- the eigenmodes at some $z - \Delta z/2$ and v_j^+ the eigenmodes at some $z + \Delta z/2$. The centered finite difference for the eigenmode derivative is

$$\frac{\partial v_i}{\partial z} \approx \frac{v_i^+ - v_i^-}{\Delta z}. \tag{34}$$

We can also estimate the eigenmodes at z via the average:

$$v_i \approx \frac{v_i^+ + v_i^-}{2}. \tag{35}$$

The inner product of the two is

$$\begin{aligned} \langle v_i | \frac{\partial v_j}{\partial z} \rangle &\approx \frac{1}{2\Delta z} \iint \left[v_i^+ + v_i^- \right] \left[v_j^+ - v_j^- \right] \, dxdy \\ &= \frac{1}{2\Delta z} \iint \left[v_i^- v_j^+ - v_i^+ v_j^- \right] \, dxdy. \end{aligned} \tag{36}$$

where terms disappear due to orthonormality relation. This estimate is automatically antisymmetric.

A The WKB approximation; connections to coupled-mode theory

Here, we provide a brief derivation of the WKB approximation, which will hopefully be applicable to both QM and waveguide optics.

The WKB method can be derived through the method of successive approximations. Suppose we have an ordinary differential equation of the form

$$\frac{dy}{dx} = f(x, y). \tag{37}$$

The above can be recast in integral form as

$$\begin{aligned} y(x) &= \int_{x_0}^{x} f(x', y(x'))dx' + C \\ C &= y(x_0) \end{aligned} \tag{38}$$

To solve this integral equation we can provide an initial guess for $y(x)$, denoted $y_0(x)$. Substituting this guess back into the integral equation allows us to determine a more accurate, higher-order guess $y_1(x)$. This process can be repeated:

$$y_n(x) = \int_{x_0}^{x} f(x', y_{n-1}(x'))dx' + C. \tag{39}$$

The resulting sequence y_n should converge to the true solution under certain (boundary and Lipschitz) conditions.

In both QM and optics, we seek to solve a wave equation. For QM this is the time-independent Schrodinger equation. In both cases, we deal with a type of *Helmholtz* equation.

$$\frac{d^2\psi}{dx^2} + k^2(x)\psi = 0. \tag{40}$$

For $k(x) = k$ constant, solutions are of the form $\psi(x) = \exp(\pm ikx)$. If $k(x)$ is allowed to vary, then we might replace kx with $\int k(x)dx$. This is gives us the 0th order approximation

$$\psi_0(x) = e^{\pm i \int k(x)dx}. \tag{41}$$

Plugging back into equation 40, it can be shown that if the 0th order approximation is to be accurate then we must also have

$$\left|\frac{dk}{dx}\right| \ll k^2. \tag{42}$$

Heuristically, $k(x)$ must be a slowly-varying function of x.

We now apply the method of successive approximations to solve the wave equation 40. We first assume the following general form for the field ψ:

$$\psi(x) = e^{iS(x)}. \tag{43}$$

Plugging this into equation 40 yields

$$iS'' - (S')^2 + k^2 = 0. \tag{44}$$

We manipulate the above, derive the series relation for S_n, and insert the 0th order approximation:

$$\begin{aligned}
S'(x) &= \pm\sqrt{iS''(x) + k^2(x)} \\
S_n(x) &= \pm \int \sqrt{iS''_{n-1} + k^2} dx \\
S_0(x) &= \pm \int k(x)\, dx \\
S_1(x) &= \pm \int \sqrt{\pm ik' + k^2} dx \\
&\approx \pm \int k \left[1 \pm \frac{i}{2}\frac{k'}{k^2}\right] dx \\
&= \pm \int k\, dx + \frac{i}{2}\ln(k) + C
\end{aligned} \tag{45}$$

where going from line 4 to line 5 we made use of the slowly-varying condition 42 to apply a Taylor expansion. Therefore, the first-order field solution is

$$\psi_1(x) \propto \frac{1}{\sqrt{k(x)}} e^{\pm i \int k(x)\, dx}. \tag{46}$$

Notice that $k^{-1/2}$ amplitude dependence! This is exactly the form we derived in 1.2 under the adiabatic approximation. So the coupled-mode solution we derived reduces to first-order WKB when making the adiabatic approximation.

B Connections between the varying basis and static basis approaches

For the moment, we assume mode non-degeneracy. Following equation 8, define the instantaneous eigenmodes as the solution to the eigenmode equation

$$\left[\nabla_\perp^2 + k^2 n^2(x,y,z)\right] |\xi_j^z\rangle \equiv A|\xi_j^z\rangle = \beta_j^2(z)\, |\xi_j^z\rangle. \tag{47}$$

Now, consider two closely values of z separated by some small δz. Define

$$\delta\epsilon = n^2(x,y,z+\delta z) - n^2(x,y,z). \tag{48}$$

Following perturbation theory, we can treat $k^2\delta\epsilon$ as a perturbation to the operator A. The corrected eigenmode is approximated as

$$|\xi_j^{z+\delta z}\rangle \approx |\xi_j^z\rangle + \sum_{i\neq j} \frac{\langle \xi_i | k^2 \delta\epsilon | \xi_j^z\rangle}{\beta_j^2 - \beta_i^2} |\xi_i^z\rangle \tag{49}$$

Therefore, we can approximate the cross-coupling matrix as

$$\begin{aligned}
\langle \xi_k^z | \frac{\partial}{\partial z} | \xi_j^z \rangle &\approx \frac{1}{\delta z} \sum_{i \neq j} \frac{\langle \xi_i^z | k^2 \delta\epsilon | \xi_j^z \rangle}{\beta_j^2 - \beta_i^2} \langle \xi_k^z | \xi_i^z \rangle \\
&= \begin{cases} \dfrac{1}{\delta z} \dfrac{\langle \xi_k^z | k^2 \delta\epsilon | \xi_j^z \rangle}{\beta_j^2 - \beta_k^2}, & k \neq j \\ 0, & k = j. \end{cases}
\end{aligned} \tag{50}$$

This formula connects the cross-coupling matrix of the varying-basis approach in §1 to the that of the fixed basis approach in §2.

Degeneracy can be treated if the degeneracy is listed by the perturbation. If the degeneracy is broken at first order, then perturbation theory tells us the following. Definite $|\xi_{nk}^z\rangle$ as the kth eigenmode in the degenerate subspace admitting the β_n^2 as its shared eigenvalue, and $|\xi_m^z\rangle$ as an eigenmode outside the degenerate subspace. The correction for the degenerate eigenmodes where degeneracy is lifted to 1st order is

$$|\xi_{nk}^{z+\delta z}\rangle \approx |\xi_{nk}^z\rangle + \sum_{m \neq n} \frac{\langle \xi_m^z | k^2 \delta\epsilon | \xi_{nk}^z \rangle}{\beta_m^2 - \beta_n^2} \left(-|\xi_m^z\rangle + \sum_{l \neq k} \frac{\langle \xi_{nl}^z | k^2 \delta\epsilon | \xi_m^z \rangle}{\beta_{nl}^2 - \beta_{nk}^2} |\xi_{nl}^z\rangle \right) \tag{51}$$

where β_{nk}^2 is the first order correction to the degenerate eigenvalues:

$$\beta_{nk}^2 = \langle \xi_{nk}^z | k^2 \delta\epsilon | \xi_{nk}^z \rangle. \tag{52}$$

Note that we have implicitly assumed that our basis in the degenerate subspace is chosen to diagonalize the perturbation $k^2\delta\epsilon$.

However, for certain waveguides (think a circular step-index fiber whose core radius slowly grows with z), the degeneracy within mode groups is not broken by the perturbation. I believe that in these cases, there is no cross-coupling.