Modeling Photon Generation

Problem Presenter

C. J. McKinstrie
Alcatel-Lucent

Problem Participants

Shuchi Agrawal, New Jersey Institute of Technology
David A. Edwards, University of Delaware
Joseph D. Fehribach, Worcester Polytechnic Institute
John Gounley, Old Dominion University
Isaac Harris, University of Delaware
Richard Moore, New Jersey Institute of Technology
Takeshi Takahashi, University of Massachusetts
Jacek Wrobel, New Jersey Institute of Technology
and others...

Twenty-Seventy Annual Workshop on Mathematical Problems in Industry
June 13–17, 2011
New Jersey Institute of Technology

Section 1: Introduction

The two most celebrated developments leading to today's high-speed optical communication networks are the invention of the laser in 1960 and the subsequent decade of development of extremely transparent fused silica for use in optical fiber. Although optical fibers transmit light efficiently, their response to the incident electric fields is nonlinear. This nonlinearity provides a natural counterbalance to chromatic dispersion, producing robust pulses that are very efficient carriers of digital information. A nonlinear optical susceptibility also provides a means to transfer energy between different frequencies of light, a mechanism that can be exploited for signal amplification at large powers or for quantum information experiments at extremely low powers.

Parametric devices based on four-wave mixing (FWM) in fibers can amplify, frequency-convert, phase-conjugate, regenerate, and sample optical signals in classical communication systems (McKinstrie *et al.*, 2007). They can also generate photon pairs for quantum information (communication and computation) experiments (Fan *et al.*, 2007).

As was first suggested by Hasegawa and Tappert in 1973, when a beam of light is formed by slowly modulating the amplitude $\psi(t, z)$ of a laser's output (or the concatenated fields from several lasers), the appropriate model for the amplitude's evolution as it advances through an optical fiber is the nonlinear Schrödinger equation:

$$\frac{\partial \psi}{\partial z} = -\alpha\psi + i\beta(i\partial/\partial t)\psi + i\gamma|\psi|^2\psi. \tag{1.1}$$

Here z represents distance along the fiber axis, α represents loss due to absorption, β represents the propagation constant (with formal dependence on frequency to denote chromatic dispersion), and γ is the nonlinear constant of the fiber. This equation can either be interpreted with ψ a scalar quantity, as in the case of a singly polarized beam of light propagating through an isotropic, polarization-preserving fiber, or it can be interpreted with ψ a vector quantity, giving rise to a large number of potential interactions mediated by the nonlinear tensor. For instance, in the case of FWM ψ is a superposition of four optical pulses, each with a different carrier frequency.

To understand the behavior of solutions of the nonlinear partial differential equation given by (1.1), one must first understand the simplified case of constant-wavelength fields with amplitudes that depend only on z, such that their interaction takes the form of coupled ordinary differential equations referred to as coupled-mode equations (CMEs), as discussed in section 2. Many applications, from amplification in optical communications networks to photon generation in quantum experiments, involve one or more pump fields at relatively large amplitude interacting with one or more signal fields at relatively small amplitude. This scale disparity suggests an approximate linear form of the CMEs in the undepleted-pump regime, where the CMEs can formally be solved to yield input-output equations (IOEs) of the form

$$\mathbf{x}(z) = M(z)\mathbf{x}(0) + N(z)\mathbf{x}^*(0),$$

where $\mathbf{x} \in \mathcal{C}^n$ is a vector of complex field amplitudes, and $M, N \in \mathcal{C}^{n \times n}$ are transfer (Green) matrices found by solving the CMEs.

For simple one- and two-mode interactions, it is easy to solve the CMEs and interpret the IOEs. However, in some systems several modes interact simultaneously, or several two-mode interactions occur sequentially. For such systems, the CMEs and IOEs are complicated and two related questions arise: Under what conditions can we solve the CMEs explicitly and how can we interpret the (explicit or implicit) IOEs?

This report explores the relationship between the linear approximate form of the CMEs and the corresponding IOEs. In particular, the spectral decomposition of the linear operator in the CMEs has implications for the singular value decomposition of the IOEs. This relationship is explored in general to the extent that linear algebraic tools will allow; specific cases of physical interest are explored in some more depth.

Section 2: Coupled-Mode Equations

Parametric interactions of weak sidebands driven by strong pumps are governed by coupled-mode equations (CMEs) of the following form:

$$\frac{d\mathbf{x}}{dz} = A\mathbf{x} + B\mathbf{x}^*, \tag{2.1}$$

where $\mathbf{x} \in \mathcal{C}^n$ is the amplitude vector, and * represents the complex conjugate. The entries of the amplitude vector could be the amplitudes of distinct monochromatic sidebands (continuous waves), or different frequency components of multichromatic sidebands (pulses), with one or two polarization components. For uniform media the coupling coefficients that form the entries of $A, B \in \mathcal{C}^{n\times n}$ are constant, whereas for nonuniform media they vary with position. In this manuscript we focus on the uniform case.

The quantum mechanical properties of parametric processes are not part of the project. However, the laws of quantum mechanics impose constraints on the coefficient matrices (McKinstrie, 2009), namely that

$$A = -A^\dagger, \qquad B = B^t, \tag{2.2}$$

where † denotes complex transpose (Hermitian conjugate) and ^{t} denotes regular transpose.

Since (2.1) is a linear problem, the solution $\mathbf{x}(z)$ at any time can be derived from multiplying two *transfer matrices* by the initial vectors for $\mathbf{x}$ and $\mathbf{x}^*$:

$$\mathbf{x}(z) = M(z)\mathbf{x}(0) + N(z)\mathbf{x}^*(0). \tag{2.3}$$

Note that if B is the zero matrix 0 (the notation used throughout this report), then $M = e^{zA}$. If one can construct $M(z)$ and $N(z)$, then the problem is solved for all z.

One way to solve the problem is to note that the complex conjugate of (2.1) is given by

$$\frac{d\mathbf{x}^*}{dz} = A^*\mathbf{x}^* + B^*\mathbf{x}. \tag{2.4}$$

Combining (2.1) and (2.4), we have

$$\frac{d}{dz}\begin{pmatrix}\mathbf{x}\\ \mathbf{x}^*\end{pmatrix} = \begin{pmatrix}A\mathbf{x} + B\mathbf{x}^*\\ A^*\mathbf{x}^* + B^*\mathbf{x}\end{pmatrix} = \begin{pmatrix}A & B\\ B^* & A^*\end{pmatrix}\begin{pmatrix}\mathbf{x}\\ \mathbf{x}^*\end{pmatrix}.$$

From (2.2), we may write

$$A = iJ, \qquad J = J^\dagger, \qquad B = iK, \qquad K = K^t, \tag{2.5}$$

which allows us to rewrite our system as

$$
\begin{aligned}
\frac{d}{dz}\begin{pmatrix} \mathbf{x} \\ \mathbf{x}^* \end{pmatrix} &= \begin{pmatrix} iJ & iK \\ -iK^* & -iJ^* \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{x}^* \end{pmatrix} \\
\frac{d\mathbf{y}}{dz} &= iL\mathbf{y}, \qquad (2.6) \\
\mathbf{y} &= \begin{pmatrix} \mathbf{x} \\ \mathbf{x}^* \end{pmatrix}, \qquad (2.7a) \\
L &= \begin{pmatrix} J & K \\ -K^* & -J^* \end{pmatrix}. \qquad (2.7b)
\end{aligned}
$$

Note that we have expanded our system from an n-dimensional one to a $2n$-dimensional one. Therefore, we would seem to have introduced some additional degrees of freedom into the problem, but the initial condition

$$
\mathbf{y}(0) = \begin{pmatrix} \mathbf{x}(0) \\ \mathbf{x}^*(0) \end{pmatrix} \qquad (2.8)
$$

takes care of that. Given that (2.6) was generated from the complex-conjugate equations (2.1) and (2.4), the form of the flow will force $\mathbf{y}$ to be of the form (2.7a) for all z as long as it starts out that way. In particular, requiring that the last n elements of $\mathbf{y}$ are the complex conjugates of the first n elements provides the n additional conditions we need to close our $2n$-dimensional system.

In general, the solution of (2.6) subject to (2.8) is given by

$$
\mathbf{y}(z) = e^{izL}\mathbf{y}(0). \qquad (2.9)
$$

Let L be diagonalizable. Then by the spectral decomposition we have

$$
e^{izL} = Se^{iz\Lambda}S^{-1}, \qquad (2.10)
$$

where S is the matrix of eigenvectors of L and Λ is the diagonal matrix of corresponding eigenvalues. Therefore, facts about the eigenvalues of L will tell us about the stability of the system.

Section 3: Eigenvalues of L

Now we wish to determine facts about the eigenvalues and eigenvectors of L.

Lemma 3.1. Let

$$L\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ \mathbf{y}_{\mathrm{b}}\end{pmatrix}=\lambda\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ \mathbf{y}_{\mathrm{b}}\end{pmatrix}.$$

Then

$$L\begin{pmatrix}\mathbf{y}_{\mathrm{b}}^*\\ \mathbf{y}_{\mathrm{a}}^*\end{pmatrix}=-\lambda^*\begin{pmatrix}\mathbf{y}_{\mathrm{b}}^*\\ \mathbf{y}_{\mathrm{a}}^*\end{pmatrix},\qquad L^\dagger\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ -\mathbf{y}_{\mathrm{b}}\end{pmatrix}=\lambda\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ -\mathbf{y}_{\mathrm{b}}\end{pmatrix}$$

Proof.

$$L\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ \mathbf{y}_{\mathrm{b}}\end{pmatrix}=\lambda\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ \mathbf{y}_{\mathrm{b}}\end{pmatrix}\qquad\Longrightarrow\qquad\begin{aligned}J\mathbf{y}_{\mathrm{a}}+K\mathbf{y}_{\mathrm{b}}&=\lambda\mathbf{y}_{\mathrm{a}},\\ -K^*\mathbf{y}_{\mathrm{a}}-J^*\mathbf{y}_{\mathrm{b}}&=\lambda\mathbf{y}_{\mathrm{b}}.\end{aligned}\tag{3.1}$$

Taking the negative conjugate of each equations and reordering, we have

$$\begin{aligned}-J^*\mathbf{y}_{\mathrm{a}}^*-K^*\mathbf{y}_{\mathrm{b}}^*&=-\lambda^*\mathbf{y}_{\mathrm{a}}^*\\ K\mathbf{y}_{\mathrm{a}}^*+J\mathbf{y}_{\mathrm{b}}^*&=-\lambda^*\mathbf{y}_{\mathrm{b}}^*\end{aligned}\qquad\Longrightarrow\qquad\begin{aligned}J\mathbf{y}_{\mathrm{b}}^*+K\mathbf{y}_{\mathrm{a}}^*&=-\lambda^*\mathbf{y}_{\mathrm{b}}^*,\\ -K^*\mathbf{y}_{\mathrm{b}}^*-J^*\mathbf{y}_{\mathrm{a}}^*&=-\lambda^*\mathbf{y}_{\mathrm{a}}^*.\end{aligned}$$

Rewriting the latter system in matrix-vector form, the first result is proved. Similarly,

$$L^\dagger\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ -\mathbf{y}_{\mathrm{b}}\end{pmatrix}=\begin{pmatrix}J^\dagger & -K^t\\ K^\dagger & -J^t\end{pmatrix}\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ -\mathbf{y}_{\mathrm{b}}\end{pmatrix}=\begin{pmatrix}J\mathbf{y}_{\mathrm{a}}+K\mathbf{y}_{\mathrm{b}}\\ K^*\mathbf{y}_{\mathrm{a}}+J^*\mathbf{y}_{\mathrm{b}}\end{pmatrix}=\begin{pmatrix}\lambda\mathbf{y}_{\mathrm{a}}\\ \lambda(-\mathbf{y}_{\mathrm{b}})\end{pmatrix}=\lambda\begin{pmatrix}\mathbf{y}_{\mathrm{a}}\\ -\mathbf{y}_{\mathrm{b}}\end{pmatrix},$$

where we have used (2.5) and (3.1). Hence the proof is complete. ✔

Note that the lemma says nothing about how to determine $\mathbf{y}_{\mathrm{a}}$ and $\mathbf{y}_{\mathrm{b}}$. However, we may use the lemma to determine additional important facts:

Theorem 3.2. The eigenvalues of L come in quartets: $\{\lambda, -\lambda^*, \lambda^*, -\lambda\}$.

Proof. The first two elements of the quartet follow directly from Lemma 3.1. Moreover, we know that if λ is an eigenvalue for $L^\dagger$, then λ^* is an eigenvalue for L. The last member of the quartet results from applying the first part of Lemma 3.1 to λ^*. ✔

Remarks.

1. Note that neither the lemma nor the theorem establish a relationship between the eigenvector *for* L corresponding to λ^* and the eigenvector given in (3.1). We were unable to determine such a relationship.
2. Note that when we substitute our quartet elements into the spectral decomposition, we have that

$$e^{i(-\lambda^*)z}=e^{(i\lambda)^*z}=\left(e^{i\lambda z}\right)^*.$$

Hence the two eigenvalues related directly through Lemma 3.1 correspond to complex conjugate pairs. The other two members of the quartet are also complex conjugates, but correspond to real growth in z (if λ corresponds to decay) or vice versa.

3. These properties can also be inferred from the properties of eigenvalues of the IOEs, as discussed in Section 4.

Corollary 3.2.1. If n is odd, there exist at least two purely real or two purely imaginary eigenvalues in $\pm$ pairs.

Proof. If n is odd, then $2n$ is not divisible by four. Therefore the set $\{\lambda, -\lambda, \lambda^*, -\lambda^*\}$ must be degenerate. Hence either $\lambda = \lambda^*$ (so λ is real, and $-\lambda$ is the other member of the pair) or $\lambda = -\lambda^*$ (so λ is imaginary, in which case $\lambda^* = -\lambda$ is still the other member of the pair). ✔

Physical examples of the case $n = 1$ are discussed in McKinstrie (2009). In subsequent sections we will show that even if n is even, it is possible for L to have purely real or purely imaginary eigenvalues.

We also make some remarks about the normality of L:

Lemma 3.3. L is normal (and hence can be unitarily diagonalized) if and only if $KJ^* = -JK$.

Proof.

$$
\begin{aligned}
LL^\dagger &= \begin{pmatrix} J & K \\ -K^* & -J^* \end{pmatrix} \begin{pmatrix} J^\dagger & -K^t \\ K^\dagger & -J^t \end{pmatrix} = \begin{pmatrix} JJ^\dagger + KK^\dagger & -JK^t - KJ^t \\ -K^*J^\dagger - J^*K^\dagger & K^*K^t + J^*J^t \end{pmatrix} \\
L^\dagger L &= \begin{pmatrix} J^\dagger & -K^t \\ K^\dagger & -J^t \end{pmatrix} \begin{pmatrix} J & K \\ -K^* & -J^* \end{pmatrix} = \begin{pmatrix} J^\dagger J + K^tK^* & J^\dagger K + K^tJ^* \\ K^\dagger J + J^tK^* & K^\dagger K + J^tJ^* \end{pmatrix},
\end{aligned}
$$

If L is normal, these two matrices must be equal. The diagonal entries are equal by (2.5), but matching the upper-right entry yields

$$
\begin{aligned}
-JK^t - KJ^t &= J^\dagger K + K^tJ^* \\
-2KJ^* &= 2JK,
\end{aligned}
$$

as required. The lower-left entry is simply the conjugate transpose of the upper-right entry. ✔

Clearly the requirement in Lemma 3.3 is very restrictive, and hence in general L is not normal (and hence not Hermitian). Even so, L may still be diagonalizable, but not by a unitary matrix. Diagonalizability is not guaranteed, however, as illustrated by the example in Section 6.

Section 4: Relating the Decompositions

Physicists often like to work with the transfer matrices on the right-hand side of (2.3), expressing them in a *Schmidt* (singular value) decomposition

$$M = U_M \Sigma_M V_M^\dagger, \tag{4.1}$$

where U_M and V_M are unitary matrices and Σ_M is a diagonal positive semidefinite matrix consisting of the non-negative square roots of the eigenvalues of $M^\dagger M$.

One might suspect that there would be some interesting relationships between the singular value (Schmidt) and eigenvalue (adjoint) decompositions; for example that the diagonal elements would have the same modulus. There is one general relationship that does hold:

Lemma 4.1. Let C be a general diagonalizable matrix, $C = U\Sigma V^\dagger$, $C = S\Lambda S^{-1}$. Then

$$|\det C| = |\det \Sigma| = |\det \Lambda|.$$

Proof. Since U and V are unitary, we have that $|\det U| = |\det V^\dagger| = 1$. Similarly, $\det S^{-1} = (\det S)^{-1}$. ✔

Beyond this general result, the key issue is symmetry: If the matrix $C = C^\dagger$, then its eigenvectors are orthogonal, implying that $S^{-1} = S^\dagger$. So in this case, S is unitary and $C = S\Lambda S^\dagger$.

Lemma 4.2. If C is Hermitian, then $\sigma_j = |\lambda_j|$ for some appropriate ordering of eigenvalues. If C is also positive semidefinite, then $\sigma_j = \lambda_j$. If C is skew-Hermitian ($C = -C^\dagger$), then its eigenvalues are pure imaginary and $\sigma_j = |\lambda_j|$ where here $|\cdot|$ is modulus.

Proof. Since C is Hermitian, the entries of Σ are the non-negative square roots of the eigenvalues of $C^\dagger C = C^2$. So $\sigma = \sqrt{\lambda(C^2)} = \sqrt{[\lambda(C)]^2} = |\lambda|$. If C is also positive semidefinite, $\lambda \geq 0$, so $|\lambda| = \lambda$. Similarly, if C is skew-Hermitian, then $C^\dagger C = -C^2$, so $\sigma = \sqrt{\lambda(-C^2)} = \sqrt{-[\lambda(C)]^2} = |\lambda|$. ✔

But if C is not real symmetric or Hermitian, no results beyond the first proposition seem possible, as the following counter examples indicate:

Example 1: Suppose that

$$U_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad V_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Then

$$C_1 = U_1 \Sigma V_1^\dagger = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}.$$

But also $C_1 = S_1 \Lambda_1 S_1^{-1}$ where

$$S_1 = \begin{pmatrix} \sqrt{2}i & \sqrt{2}i \\ -1 & 1 \end{pmatrix}, \qquad S_1^{-1} = \frac{-i}{2\sqrt{2}} \begin{pmatrix} 1 & -\sqrt{2}i \\ 1 & \sqrt{2}i \end{pmatrix}, \qquad \Lambda_1 = \begin{pmatrix} -\sqrt{2}i & 0 \\ 0 & \sqrt{2}i \end{pmatrix}.$$

For this C_1, the elements of Σ and Λ_1 are unique (up to order), so clearly the elements of these two matrices do not have the same modulus, are not all real, and seem related only in that they have the same product (the same determinant).

A second example may well be helpful; here the matrix C is symmetric, but not Hermitian:

Example 2: Suppose that

$$U_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 & 2-i \\ 2+i & 0 \end{pmatrix}, \qquad V_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 & 2+i \\ 2-i & 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

So Σ is the same as in the first example. Then

$$C_2 = U_2 \Sigma V_2^\dagger = \frac{1}{5} \begin{pmatrix} 6-8i & 0 \\ 0 & 3+4i \end{pmatrix}$$

and the eigenvalues and eigenvectors are clear:

$$S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \Lambda_2 = \frac{1}{5} \begin{pmatrix} 6-8i & 0 \\ 0 & 3+4i \end{pmatrix}.$$

So again there is no easy relationship between the singular values and eigenvalues and as a result, one will not easily be able to replace the singular value decomposition by a diagonalization.

We can relate e^{izL} to M and N by combining (2.3) and its complex conjugate. Rewriting in terms of $\mathbf{x}$, we have

$$\begin{aligned} \begin{pmatrix} \mathbf{x}(z) \\ \mathbf{x}^*(z) \end{pmatrix} &= \begin{pmatrix} M(z) & N(z) \\ N^*(z) & M^*(z) \end{pmatrix} \begin{pmatrix} \mathbf{x}(0) \\ \mathbf{x}^*(0) \end{pmatrix} \\ e^{izL} &= \begin{pmatrix} M(z) & N(z) \\ N^*(z) & M^*(z) \end{pmatrix}. \end{aligned} \tag{4.2}$$

Noting that $e^{i(-z)L} = (e^{izL})^{-1}$, we have

$$\begin{pmatrix} M(z) & N(z) \\ N^*(z) & M^*(z) \end{pmatrix} \begin{pmatrix} M(-z) & N(-z) \\ N^*(-z) & M^*(-z) \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \tag{4.3}$$

which gives us relations between M and N of positive and negative argument.

However, from the laws of quantum mechanics we have that (McKinstrie, 2009)

$$MM^\dagger - NN^\dagger = I, \tag{4.4a}$$

$$MN^t - NM^t = 0. \tag{4.4b}$$

(Note that if $N = 0$, this implies that M is Hermitian.) Equations (4.4) yield the relation in the following lemma.

Lemma 4.1.

$$M(-z) = M^\dagger(z), \qquad N^*(-z) = -N^\dagger(z) \qquad \Longrightarrow \qquad N(-z) = -N^t(z). \tag{4.5}$$

Proof. We use (4.5) as an ansatz. Substituting (4.5) into (4.3) and dropping the arguments, we have

$$\begin{aligned}
\begin{pmatrix} M & N \\ N^* & M^* \end{pmatrix} \begin{pmatrix} M^\dagger & -N^t \\ (-N^t)^* & (M^\dagger)^* \end{pmatrix} &= \begin{pmatrix} MM^\dagger - NN^\dagger & -MN^t + NM^t \\ N^*M^\dagger - M^*N^\dagger & -N^*N^t + M^*M^t \end{pmatrix} \\
&= \begin{pmatrix} I & 0 \\ (NM^t - MN^t)^* & (-NN^\dagger + MM^\dagger)^* \end{pmatrix} \\
&= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},
\end{aligned}$$

where we have used (4.4). Since the system is linear (and hence the transfer matrices are unique), the result is proved. ✔

The quantum mechanical relations (4.4) also provide a simplifying relation between the Schmidt decompositions of M and N. First we have that

$$N = U_M \Sigma_N V_M^t. \tag{4.6}$$

Though the proof of this is somewhat involved (McKinstrie, 2009), it can be verified to be true by substituting (4.1) and (4.6) into (4.4b):

$$\begin{aligned}
(U_M \Sigma_M V_M^\dagger)(U_M \Sigma_N V_M^t)^t - (U_M \Sigma_N V_M^t)(U_M \Sigma_M V_M^\dagger)^t &= 0 \\
U_M \Sigma_M \Sigma_N U_M^t - U_M \Sigma_N \Sigma_M U_M^t &= 0,
\end{aligned}$$

where the result holds because diagonal matrices commute.

We also obtain a relationship between the singular values of M and N:

Lemma 4.2. Let M and N be transfer matrices with the Schmidt decomposition given in (4.1) and (4.6). Then

$$\Sigma_M^2 - \Sigma_N^2 = I. \tag{4.7}$$

Proof. The result follows directly from substituting (4.1) and (4.6) into (4.4a):

$$
\begin{aligned}
(U_M\Sigma_M V_M^\dagger)(U_M\Sigma_M V_M^\dagger)^\dagger - (U_M\Sigma_N V_M^t)(U_M\Sigma_N V_M^t)^\dagger &= I \\
U_M\Sigma_M^2 U_M^\dagger - U_M\Sigma_N^2 U_M^\dagger &= U_M U_M^\dagger \\
U_M(\Sigma_M^2 - \Sigma_N^2 - I)U_M^\dagger &= 0. \qquad ✔
\end{aligned}
$$

Note that the spectral decomposition (2.10) has the convenient form that all the z-dependence of the transfer matrix is confined to the diagonal matrix $e^{iz\Lambda}$. In contrast, the Schmidt decompositions in (4.1) and (4.6) generally have z-dependence in all three matrices (McKinstrie, 2009).

We can relate the Schmidt decompositions of the n- and $2n$-dimensional complex systems by noting that

$$
\begin{aligned}
e^{izL} &= \begin{pmatrix} M & N \\ N^* & M^* \end{pmatrix} = \begin{pmatrix} U_M\Sigma_M V_M^\dagger & U_M\Sigma_N V_M^t \\ (U_M\Sigma_N V_M^t)^* & (U_M\Sigma_M V_M^\dagger)^* \end{pmatrix} = \begin{pmatrix} U_M\Sigma_M V_M^\dagger & U_M\Sigma_N V_M^t \\ U_M^*\Sigma_N V_M^\dagger & U_M^*\Sigma_M V_M^t \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} U_M & 0 \\ 0 & U_M^* \end{pmatrix}\begin{pmatrix} (\Sigma_M+\Sigma_N)+(\Sigma_M-\Sigma_N) & (\Sigma_M+\Sigma_N)-(\Sigma_M-\Sigma_N) \\ (\Sigma_M+\Sigma_N)-(\Sigma_M-\Sigma_N) & (\Sigma_M+\Sigma_N)+(\Sigma_M-\Sigma_N) \end{pmatrix} \times \\
&\qquad \begin{pmatrix} V_M^\dagger & 0 \\ 0 & V_M^t \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} U_M & 0 \\ 0 & U_M^* \end{pmatrix}\begin{pmatrix} \Sigma_M+\Sigma_N & \Sigma_M-\Sigma_N \\ \Sigma_M+\Sigma_N & -(\Sigma_M-\Sigma_N) \end{pmatrix}\begin{pmatrix} I & I \\ I & -I \end{pmatrix}\begin{pmatrix} V_M^\dagger & 0 \\ 0 & V_M^t \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} U_M & 0 \\ 0 & U_M^* \end{pmatrix}\begin{pmatrix} I & I \\ I & -I \end{pmatrix}\begin{pmatrix} \Sigma_M+\Sigma_N & 0 \\ 0 & \Sigma_M-\Sigma_N \end{pmatrix}\begin{pmatrix} V_M^\dagger & V_M^t \\ V_M^\dagger & -V_M^t \end{pmatrix} \\
&= \frac{1}{\sqrt{2}}\begin{pmatrix} U_M & U_M \\ U_M^* & -U_M^* \end{pmatrix}\Sigma V^\dagger, \\
\Sigma &= \begin{pmatrix} \Sigma_M+\Sigma_N & 0 \\ 0 & (\Sigma_M+\Sigma_N)^{-1} \end{pmatrix}, \qquad V = \frac{1}{\sqrt{2}}\begin{pmatrix} V_M & V_M \\ V_M^* & -V_M^* \end{pmatrix}, \qquad (4.8a)
\end{aligned}
$$

where we have used (4.6). Continuing to simplify, we have

$$
e^{izL} = U\Sigma V^\dagger, \qquad U = \frac{1}{\sqrt{2}}\begin{pmatrix} U_M & U_M \\ U_M^* & -U_M^* \end{pmatrix}. \qquad (4.8b)
$$

Note that this is a singular value decomposition since

$$
\begin{aligned}
U^\dagger U &= \frac{1}{2}\begin{pmatrix} U_M^\dagger & -U_M^t \\ U_M^\dagger & U_M^t \end{pmatrix}\begin{pmatrix} U_M & U_M \\ -U_M^* & U_M^* \end{pmatrix} = \frac{1}{2}\begin{pmatrix} I+(U_M^\dagger U_M)^* & I-(U_M^\dagger U_M)^* \\ I-(U_M^\dagger U_M)^* & I+(U_M^\dagger U_M)^* \end{pmatrix} = I, \\
V^\dagger V &= \frac{1}{2}\begin{pmatrix} V_M^\dagger & V_M^t \\ V_M^\dagger & -V_M^t \end{pmatrix}\begin{pmatrix} V_M & V_M \\ V_M^* & -V_M^* \end{pmatrix} = \frac{1}{2}\begin{pmatrix} I+(V_M^\dagger V_M)^* & I-(V_M^\dagger V_M)^* \\ I-(V_M^\dagger V_M)^* & I+(V_M^\dagger V_M)^* \end{pmatrix} = I,
\end{aligned}
$$

where we have used the fact that U_M and V_M are unitary, and the entries in Σ are all positive. The relationship between the Schmidt decompositions of the n-dimensional complex system and the $2n$-dimensional real system is explored in McKinstrie and Alic (to appear).

An alternative proof of a part of Theorem 3.2 may be discovered by deducing that e^{izL} is *symplectic*. A matrix $T \in \mathcal{C}^{2n\times 2n}$ is called symplectic if (Mackey and Mackey, 2003)

$$T\Omega T^t = \Omega, \qquad \Omega = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \qquad I \in \mathcal{C}^{n\times n}. \tag{4.9}$$

Lemma 4.3. e^{izL} is symplectic.

Proof.

$$\begin{aligned} e^{izL}\Omega(e^{izL})^t &= \begin{pmatrix} M & N \\ N^* & M^* \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} M^t & N^\dagger \\ N^t & M^\dagger \end{pmatrix} \\ &= \begin{pmatrix} -N & M \\ -M^* & N^* \end{pmatrix} \begin{pmatrix} M^t & N^\dagger \\ N^t & M^\dagger \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \Omega, \end{aligned}$$

where we have used (4.2) and (4.4).

Corollary 4.3.1. If λ is an eigenvalue for L, so is $-\lambda$.

Proof. Let T be a symplectic matrix, and let μ be an eigenvalue of T^t (and hence T) corresponding to $\mathbf{z}$. Then

$$\Omega\mathbf{z} = T\Omega T^t\mathbf{z} = \mu T(\Omega\mathbf{z}).$$

Hence μ^{-1} is an eigenvalue of T. In our case

$$T = e^{izL},$$

so $\mu = e^{i\lambda z}$ for some eigenvalue λ of L. Hence $\mu^{-1} = e^{i(-\lambda)z}$, and corresponds to a second eigenvalue $-\lambda$ of L. ✔

Section 5: Two-Mode Case, Isotropic Propagation

We now specialize to the case of two-mode interaction. First, we consider the case of linear, uncoupled wave propagation (*e.g.*, two polarization components in a non-birefringent fiber, or many frequency components in a non-dispersive fiber). In that case, all the components of the signal amplitude $\mathbf{A}_{\mathrm{s}}$ and idler amplitude $\mathbf{A}_{\mathrm{i}}$ experience the same interaction with themselves. Hence we have

$$\frac{d}{dz}\begin{pmatrix}\mathbf{A}_{\mathrm{s}}\\ \mathbf{A}_{\mathrm{i}}^*\end{pmatrix} = iL\begin{pmatrix}\mathbf{A}_{\mathrm{s}}\\ \mathbf{A}_{\mathrm{i}}^*\end{pmatrix}, \qquad L = \begin{pmatrix}\delta I & K\\ -K^* & -\delta I\end{pmatrix}. \tag{5.1}$$

Here the δI term models the isotropic self-interaction, and K models the interaction between $\mathbf{A}_{\mathrm{s}}$ and $\mathbf{A}_{\mathrm{i}}$. Note that (5.1) is in exactly the same form as (2.6) and (2.7b), and so the results from the previous sections hold.

We next wish to examine the eigenvalues of such a system. We begin by considering a much more general case:

Lemma 5.1. Let $JK = KJ^*$, and let $K\mathbf{u}_K^* = \sigma_K \mathbf{u}_K$, $\sigma_K \geq 0$. (In other words, σ_K is one of the singular values of the Schmidt decomposition of K.) Then the following hold:

1. $\mathbf{u}_K$ is an eigenvector for J.
2. There exist constants β, λ_L such that

$$L\begin{pmatrix}\beta\mathbf{u}_K\\ \mathbf{u}_K^*\end{pmatrix} = \lambda_L\begin{pmatrix}\beta\mathbf{u}_K\\ \mathbf{u}_K^*\end{pmatrix}. \tag{5.2}$$

Proof. Consider the following quantities:

$$\begin{aligned}(JK)\mathbf{u}_K^* &= J(\sigma_K\mathbf{u}_K) = \sigma_K J\mathbf{u}_K = \sigma_K\mathbf{w}_K, \qquad \mathbf{w}_K = J\mathbf{u}_K,\\ (KJ^*)\mathbf{u}_K^* &= K(J^*\mathbf{u}_K^*) = K\mathbf{w}_K^*.\end{aligned}$$

By hypothesis, these two quantities are equal. (Indeed, this is how the relationship $JK = KJ^*$ was chosen for the hypothesis.) Hence $\mathbf{w}_K^*$ must be proportional to $\mathbf{u}_K^*$, as long as all the singular values are all distinct. (The proof for the nondistinct case is more complicated, but doable.) Hence

$$\mathbf{w}_K^* = J^*\mathbf{u}_K^* = \lambda_J^*\mathbf{u}_K^* \qquad J\mathbf{u}_K = \lambda_J\mathbf{u}_K,$$

as required. Performing the multiplication, we have

$$L\begin{pmatrix}\beta\mathbf{u}_K\\ \mathbf{u}_K^*\end{pmatrix} = \begin{pmatrix}J & K\\ -K^* & -J^*\end{pmatrix}\begin{pmatrix}\beta\mathbf{u}_K\\ \mathbf{u}_K^*\end{pmatrix} = \begin{pmatrix}(\beta\lambda_J + \sigma_K)\mathbf{u}_K\\ -\beta K^*\mathbf{u}_i - \lambda_J\mathbf{u}_i^*\end{pmatrix}.$$

But $K^* \mathbf{u}_K = (K\mathbf{u}_K^*)^* = (\sigma_K \mathbf{u}_K)^* = \sigma_K \mathbf{u}_K^*$ since σ_K is real. Continuing to simplify, we have

$$L \begin{pmatrix} \beta \mathbf{u}_K \\ \mathbf{u}_K^* \end{pmatrix} = \begin{pmatrix} (\lambda_J + \sigma_i/\beta)(\beta \mathbf{u}_K) \\ -(\beta\sigma_K + \lambda_J)\mathbf{u}_K^* \end{pmatrix}. \tag{5.3}$$

Therefore, our theorem is true iff

$$\lambda_L = \lambda_J + \frac{\sigma_K}{\beta} = -(\beta\sigma_K + \lambda_J)$$

$$\sigma_K \beta^2 + 2\lambda_J \beta + \sigma_K = 0$$

$$\beta = \frac{-2\lambda_J \pm \sqrt{4\lambda_J^2 - 4\sigma_K^2}}{2\sigma_K} = \frac{-\lambda_J \pm \sqrt{\lambda_J^2 - \sigma_K^2}}{\sigma_K}, \tag{5.4a}$$

$$\lambda_L = \mp\sqrt{\lambda_J^2 - \sigma_K^2}. \tag{5.4b}$$

✔

Remarks.

1. Note that in this case, more general than (5.1), the eigenvalues of L will be either real or imaginary.
2. The fact that there are two choices for β and λ_L produce the set of $2n$ eigenvalues and eigenvectors for which we are looking.
3. Note that any eigenvector will be orthogonal to $2(n-1)$ other eigenvectors, because

$$\begin{pmatrix} \beta_i \mathbf{u}_{K,i} \\ \mathbf{u}_{K,i}^* \end{pmatrix}^\dagger \begin{pmatrix} \beta_j \mathbf{u}_{K,j} \\ \mathbf{u}_{K,j}^* \end{pmatrix} = 0$$

as the vectors $\mathbf{u}_K$ are orthogonal by the Schmidt decomposition. It will not be orthogonal only to the second eigenvector corresponding to the same $\mathbf{u}_K$.

Now we consider the specific system (5.1). First, we note from (5.2) that both signal and idler have the same basis vectors $\mathbf{u}$. We can calculate the eigenvalues directly:

Corollary 5.1.1. The eigenvalues of the system (5.1) are given by (5.4) with

$$\beta = \frac{-\delta \pm \sqrt{\delta^2 - \sigma_K^2}}{\sigma_K}, \qquad \lambda_L = \mp\sqrt{\delta^2 - \sigma_K^2}.$$

Proof. The system in (5.1) is in the form (2.7b) with $J = \delta I$. Hence $\lambda_J = \delta$ for any eigenvector and the result follows. ✔

Another way to analyze the system (5.1) is to note that since K is symmetric,

$$K = U\Sigma V^\dagger, \qquad K^t = V^* \Sigma U^t = K,$$

so one possible Schmidt decomposition has $U = V^*$. Hence we may write L in (5.1) as

$$L = \begin{pmatrix} U(\delta I)U^\dagger & U\Sigma U^t \\ -U^*\Sigma U^\dagger & U(-\delta I)U^\dagger \end{pmatrix}. \tag{5.5}$$

The case where K is real is quite similar:

Lemma 5.2. Let K be real in the definition of L in (2.7b), $JK = KJ$, and $K\mathbf{z}_K = \lambda_K \mathbf{z}_K$. Then

1. Both λ_K and $\mathbf{z}_K$ are real, and the eigenvectors are orthogonal.
2. There exist constants β, λ_L such that

$$L \begin{pmatrix} \beta \mathbf{z}_K \\ \mathbf{z}_K \end{pmatrix} = \lambda_L \begin{pmatrix} \beta \mathbf{z}_K \\ \mathbf{z}_K \end{pmatrix}. \tag{5.6}$$

Proof. Since K is now real and symmetric, it has real eigenvalues and a full set of real orthonormal eigenvectors, so item #1 is proved. Since J and K commute, they are simultaneously diagonalizable, so $J\mathbf{z}_K = \lambda_J \mathbf{z}_K$ for λ_J not necessarily equal to λ_K. Performing the multiplication, we have

$$L \begin{pmatrix} \beta \mathbf{z}_K \\ \mathbf{z}_K \end{pmatrix} = \begin{pmatrix} J & K \\ -K & -J^* \end{pmatrix} \begin{pmatrix} \beta \mathbf{z}_K \\ \mathbf{z}_K \end{pmatrix} = \begin{pmatrix} (\beta\lambda_J + \lambda_K)\mathbf{z}_K \\ -\beta\lambda_K \mathbf{z}_K - J^* \mathbf{z}_K \end{pmatrix}.$$

But $J^*\mathbf{z}_K = (J\mathbf{z}_K^*)^* = (J\mathbf{z}_K)^*$ since $\mathbf{z}_K$ is real, and $(J\mathbf{z}_K)^* = (\lambda_J \mathbf{z}_K)^* = \lambda_J \mathbf{z}_K$, since J is Hermitian. Continuing to simplify, we have

$$L \begin{pmatrix} \beta \mathbf{z}_K \\ \mathbf{z}_K \end{pmatrix} = \begin{pmatrix} (\lambda_J + \lambda_K/\beta)(\beta \mathbf{z}_K) \\ -(\beta\lambda_K + \lambda_J)\mathbf{z}_K \end{pmatrix},$$

which is exactly of the form (5.3). Hence we have

$$\beta = \frac{-\lambda_J \pm \sqrt{\lambda_J^2 - \lambda_K^2}}{\lambda_K}, \tag{5.7a}$$

$$\lambda_L = \mp\sqrt{\lambda_J^2 - \lambda_K^2}. \tag{5.7b}$$

✔

Corollary 5.2.1. If K is real, the eigenvalues of the system (5.1) are given by (5.6) with

$$\beta = \frac{-\delta \pm \sqrt{\delta^2 - \lambda_K^2}}{\lambda_K}, \qquad \lambda_L = \mp\sqrt{\delta^2 - \lambda_K^2}.$$

Proof. The result trivially follows from Lemma 5.2 and Corollary 5.1.1. ✔

We may use the same sort of technique to derive the factorization in (4.8) in a different way:

Lemma 5.3. Let M and N be transfer matrices with $M\mathbf{v}_M = \sigma_M \mathbf{u}_M$ for some $\sigma_M \geq 0$. (So σ_M is the singular value, and $\mathbf{u}_M$ and $\mathbf{v}_M$ are columns of the matrices U_M, V_M in (4.1).) Then

$$e^{izL} \begin{pmatrix} \mathbf{v}_M \\ \mathbf{v}_M^* \end{pmatrix} = (\sigma_M + \sigma_N) \begin{pmatrix} \mathbf{u}_M \\ \mathbf{u}_M^* \end{pmatrix}, \qquad e^{izL} \begin{pmatrix} \mathbf{v}_M \\ -\mathbf{v}_M^* \end{pmatrix} = (\sigma_M - \sigma_N) \begin{pmatrix} \mathbf{u}_M \\ -\mathbf{u}_M^* \end{pmatrix}. \tag{5.8}$$

Proof. We note that

$$N\mathbf{v}_M^* = U_M \Sigma_N V_M^t \mathbf{v}_M^* = \sigma_N \mathbf{u}_M,$$

where we have used the fact that $V_M^t \mathbf{v}_M^*$ will pick out the same singular value and vector that $M\mathbf{v}_M$ does. Performing the multiplications, we have

$$e^{izL}\begin{pmatrix}\mathbf{v}_M \\ \mathbf{v}_M^*\end{pmatrix} = \begin{pmatrix} M & N \\ N^* & M^* \end{pmatrix}\begin{pmatrix}\mathbf{v}_M \\ \mathbf{v}_M^*\end{pmatrix} = \begin{pmatrix}(\sigma_M+\sigma_N)\mathbf{u}_M \\ (\sigma_M+\sigma_N)\mathbf{u}_M^*\end{pmatrix},$$
$$e^{izL}\begin{pmatrix}\mathbf{v}_M \\ -\mathbf{v}_M^*\end{pmatrix} = \begin{pmatrix} M & N \\ N^* & M^* \end{pmatrix}\begin{pmatrix}\mathbf{v}_M \\ -\mathbf{v}_M^*\end{pmatrix} = \begin{pmatrix}(\sigma_M-\sigma_N)\mathbf{u}_M \\ (-\sigma_M+\sigma_N)\mathbf{u}_M^*\end{pmatrix} = (\sigma_M-\sigma_N)\begin{pmatrix}\mathbf{u}_M \\ -\mathbf{u}_M^*\end{pmatrix},$$

as required.

With this result, we may derive the previous Schmidt decomposition for e^{izL}:

Corollary 5.3.1.

$$e^{izL} = U\Sigma V^\dagger,$$

where the component matrices are defined in (4.8).

Proof. V is made up of $2n$ columns $\mathbf{v}$, each of which must satisfy $e^{izL}\mathbf{v} = \sigma\mathbf{u}$. But n of those columns are given by the first equality in (5.8), and the second are given in the second equality. Hence

$$V \propto \begin{pmatrix} V_M & V_M \\ V_M^* & -V_M^* \end{pmatrix},$$

where the $2^{-1/2}$ in front of (4.8a) assures the proper normalization. Note that the minus sign in front of the final entry assures orthogonality. The corresponding U is described by

$$U \propto \begin{pmatrix} U_M & U_M \\ U_M^* & -U_M^* \end{pmatrix},$$

where the normalization factors and negative sign play the same role. The corresponding entries of Σ are $\sigma_M + \sigma_N$ for the first N columns and $\sigma_M - \sigma_N$ for the second n columns. Once one recalls from (4.7) that $\sigma_M - \sigma_N = (\sigma_M + \sigma_N)^{-1}$, the result is proved.

Section 6: Two-Mode Case, Nonisotropic Propagation

We next consider one particular case of nonisotropic propagation, namely

$$\frac{d}{dz}\begin{pmatrix}\mathbf{A}_{\mathrm{s}}\\ \mathbf{A}_{\mathrm{i}}^*\end{pmatrix} = i\begin{pmatrix}\delta P & K\\ -K^* & -\delta P\end{pmatrix}\begin{pmatrix}\mathbf{A}_{\mathrm{s}}\\ \mathbf{A}_{\mathrm{i}}^*\end{pmatrix}, \qquad P = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{6.1}$$

so $L \in \mathcal{C}^{4\times 4}$. Note that in this case, one of the modes is coupled directly to itself, while the other is coupled negatively. (This type of coupling is representative of birefringence.) In this case we can establish conditions under which the eigenvalues of L are either real or imaginary. But first we prove the following theorem about the eigenvectors.

Theorem 6.1. Let L be defined as in (6.1). Then L^2 has one eigenvector of the form

$$\begin{pmatrix}\mathbf{z}\\ \mathbf{e}_1\end{pmatrix}, \qquad \mathbf{z}, \mathbf{e}_1 \in \mathcal{C}^2.$$

Proof.

$$L^2 = \begin{pmatrix}\delta P & K\\ -K^* & -\delta P\end{pmatrix}\begin{pmatrix}\delta P & K\\ -K^* & -\delta P\end{pmatrix} = \begin{pmatrix}\delta^2 P^2 - KK^* & \delta(PK - KP)\\ \delta(PK^* - K^*P) & \delta^2 P^2 - K^*K\end{pmatrix}.$$

$P^2 = I$, so the lower-right entry is Hermitian, so we define

$$H = \delta^2 I - K^*K. \tag{6.2}$$

We also note that

$$\begin{aligned}
PK - KP &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}k_{11} & k_{12}\\ k_{12} & k_{22}\end{pmatrix} - \begin{pmatrix}k_{11} & k_{12}\\ k_{12} & k_{22}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\\
&= \begin{pmatrix}k_{11} & k_{12}\\ -k_{12} & -k_{22}\end{pmatrix} - \begin{pmatrix}k_{11} & -k_{12}\\ k_{12} & -k_{22}\end{pmatrix} = \begin{pmatrix}0 & 2k_{12}\\ -2k_{12} & 0\end{pmatrix},\\
H^t &= \delta^2 I - K^t K^{\dagger} = \delta^2 I - KK^*,\\
PK^* - K^*P &= (P^*K - KP^*)^* = (PK - KP)^* = \begin{pmatrix}0 & 2k_{12}^*\\ -2k_{12}^* & 0\end{pmatrix}.
\end{aligned}$$

so we rewrite L^2 as

$$L^2 = \begin{pmatrix} H^t & 2\delta k_{12}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\\ 2\delta k_{12}^*\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} & H\end{pmatrix}.$$

Without loss of generality, for algebraic simplicity we redefine our eigenvector $\mathbf{z}_{L^2}$ for L^2 as

$$\mathbf{z}_{L^2} = \begin{pmatrix} z_1 \\ z_2 \\ 2\delta k_{12}^* \\ 0 \end{pmatrix}.$$

We now compute each row of $L^2\mathbf{z}_{L^2} = \lambda \mathbf{z}_{L^2}$, starting with the third row:

$$\begin{aligned} 2\delta k_{12}^* z_2 + 2\delta k_{12}^* h_{11} &= \lambda(2\delta k_{12}^*) \\ z_2 &= \lambda - h_{11}, \end{aligned} \tag{6.3a}$$

which gives us z_2 in terms of λ. Moving to the fourth row, we have

$$\begin{aligned} -2\delta k_{12}^* z_1 + 2\delta k_{12}^* h_{21} &= 0 \\ z_1 &= h_{21}, \end{aligned} \tag{6.3b}$$

which gives us z_1. Using our results from (6.3) in the first row, we have

$$\begin{aligned} h_{11}z_1 + h_{21}z_2 &= \lambda z_1 \\ h_{11}h_{21} + h_{21}(\lambda - h_{11}) &= \lambda h_{21}, \end{aligned}$$

as required. Using our results from (6.3) in the second row, we have

$$\begin{aligned} h_{12}z_1 + h_{22}z_2 - 2\delta k_{12}(2\delta k_{12}^*) &= \lambda z_2 \\ h_{12}h_{21} + h_{22}(\lambda - h_{11}) - 4\delta^2|k_{12}|^2 &= \lambda(\lambda - h_{11}), \end{aligned} \tag{6.4}$$

which is a quadratic with at least one root λ, and generically has two roots. ✔

Now we have the foundation to prove the following result about the eigenvalues:

Corollary 6.1.1. Let L be defined as in (6.1). Then L has all real and imaginary eigenvalues iff

$$(h_{11} - h_{22})^2 + 4|h_{12}|^2 \geq 16\delta^2|k_{12}|^2. \tag{6.5}$$

Proof. Calculating the discriminant d of (6.4), we have

$$\begin{aligned} 0 &= \lambda^2 - (h_{11} + h_{22})\lambda + (h_{11}h_{22} - h_{12}h_{12} * + 4\delta^2|k_{12}|^2), \\ d &= (h_{11} + h_{22})^2 - 4(h_{11}h_{22} - |h_{12}|^2 + 4\delta^2|k_{12}|^2) \\ &= (h_{11} - h_{22})^2 + 4|h_{12}|^2 - 16\delta^2|k_{12}|^2, \end{aligned}$$

where we have used the fact that H is Hermitian. Since λ corresponds to an eigenvalue of L^2, L^2 will have real eigenvalues iff $d \geq 0$, which is equivalent to the condition in (6.5). Since $\lambda(L^2) = [\lambda(L)]^2$, we have that at least one (but generally two) eigenvalues of L must be either real or imaginary. But by the quartet structure, that forces all of them to be either real or imaginary, since $n = 2$. ✔

In degenerate FWM, a strong pump drives weak signal and idler sidebands. In this case, the linearization of the Schrödinger equation (1.1) yields a coupling term of the form (McKinstrie *et al.*, 2004)

$$K = \begin{pmatrix} \alpha p_x^2 & p_x p_y \\ p_x p_y & \alpha p_y^2 \end{pmatrix}, \tag{6.6}$$

where p_x and p_y are (possibly complex) coupling coefficients in the x- and y-directions (orthogonal to propagation) and α is a real constant. In this case, we have

$$\begin{aligned} H &= \begin{pmatrix} \delta^2 & 0 \\ 0 & \delta^2 \end{pmatrix} - \begin{pmatrix} \alpha(p_x^*)^2 & p_x^* p_y^* \\ p_x^* p_y^* & \alpha(p_y^*)^2 \end{pmatrix} \begin{pmatrix} \alpha p_x^2 & p_x p_y \\ p_x p_y & \alpha p_y^2 \end{pmatrix} \\ &= \begin{pmatrix} \delta^2 - (\alpha^2|p_x|^4 + |p_x|^2|p_y|^2) & -(\alpha|p_x|^2 p_x^* p_y + \alpha|p_y|^2 p_x^* p_y) \\ -(\alpha|p_x|^2 p_x p_y^* + \alpha|p_y|^2 p_x p_y^*) & \delta^2 - (|p_x|^2|p_y|^2 + \alpha^2|p_y|^4) \end{pmatrix} \\ &= \begin{pmatrix} \delta^2 - |p_x|^2(\alpha^2|p_x|^2 + |p_y|^2) & -\alpha p_x^* p_y(|p_x|^2 + |p_y|^2) \\ -\alpha p_x p_y^*(|p_x|^2 + |p_y|^2) & \delta^2 - |p_y|^2(|p_x|^2 + \alpha^2|p_y|^2) \end{pmatrix}, \end{aligned}$$

and hence the discriminant becomes

$$\begin{aligned} d &= \left[|p_x|^2(\alpha^2|p_x|^2 + |p_y|^2) - |p_y|^2(|p_x|^2 + \alpha^2|p_y|^2)\right]^2 + 4\alpha^2|p_x|^2|p_y|^2(|p_x|^2 + |p_y|^2)^2 \\ &\quad - 16\delta^2|p_x|^2|p_y|^2 \\ &= \left[\alpha^2(|p_x|^4 - |p_y|^4)\right]^2 + 4\alpha^2|p_x|^2|p_y|^2(|p_x|^2 + |p_y|^2)^2 - 16\delta^2|p_x|^2|p_y|^2 \\ &= \alpha^2(|p_x|^2 + |p_y|^2)^2\left[\alpha^2(|p_x|^2 - |p_y|^2)^2 + 4|p_x|^2|p_y|^2\right] - 16\delta^2|p_x|^2|p_y|^2. \end{aligned}$$

Note that the form of the matrix K as defined in (6.6) is not guaranteed to be diagonalizable:

Lemma 6.2. Let K take the form indicated in (6.6). Then K is diagonalizable and nontrivial if and only if the complex pump amplitudes p_x and p_y satisfy

$$p_x p_y \neq \pm\frac{1}{2} i\alpha(p_x^2 - p_y^2).$$

Proof. Eigenvalues of K satisfy

$$\lambda^2 - \alpha(p_x^2 + p_y^2)\lambda + (\alpha^2 - 1)p_x^2 p_y^2 = 0.$$

Thus, K has a single eigenvalue with algebraic multiplicity 2 only when the discriminant vanishes, *i.e.*,

$$\alpha^2(p_x^2 + p_y^2)^2 = 4(\alpha^2 - 1)p_x^2 p_y^2,$$

which reduces to the above condition after straightforward manipulations. The condition is therefore necessary.

Now suppose the condition is satisfied. Then solving for the eigenvalue gives

$$\lambda = \frac{1}{2}\alpha(p_x^2 + p_y^2),$$

yielding

$$K - \lambda I = \frac{1}{2}\alpha(p_x^2 - p_y^2)\begin{pmatrix} 1 & \pm i \\ \pm i & -1 \end{pmatrix}.$$

This matrix has eigenvectors $(1, \pm i)^t$ each with geometric multiplicity 1, unless $\alpha = 0$ or $p_x = \pm p_y$. Since both of these cases imply the trivial matrix $K = 0$, the condition is sufficient.

Although the lack of diagonalizability has no bearing on the analysis included in this section, it precludes the extension of Lemma 5.2 to the more general case where K is complex and symmetric.

The following extensions of Theorem 6.1 were conjectured at the workshop:

1. Consider L to be of arbitrary dimension (instead of 2×2) with P having either ± 1 on the diagonal. Does the result in Theorem 6.1 still hold true? That would (by the corollary) provide conditions on which at least four eigenvalues of L would be real or imaginary.
2. It can be shown that Theorem 6.1 still holds true if one replaces $\mathbf{e}_1$ by $\mathbf{e}_2$ in Theorem 6.1. (This is how you get the other two eigenvectors.) For L of arbitrary dimension, can you replace $\mathbf{e}_1$ by $\mathbf{e}_j$ and still obtain the same result? That would (by the corollary) provide conditions on which all the eigenvalues of L would be real or imaginary.
3. It is easy to show that the $\mathbf{e}_j$ are eigenvectors of P. Could this result be extended somehow to arbitrary P using the form of Theorem 6.1 with eigenvectors of P replacing the $\mathbf{e}_j$?

Section 7: Conclusions and Further Research

The primary objective of this working group was to analyze two complementary descriptions of linearly coupled envelope equations representing optical fields interacting through the Kerr nonlinearity present in optical fiber. The two descriptions are the differential system (2.6) obtained by linearizing the coupled nonlinear Schrödinger equations for resonantly interacting fields, referred to as the coupled-mode equations (CMEs), and the algebraic system (2.3) obtained by solving the differential system, referred to as the input-output equations (IOEs).

The CMEs are characterized by the matrix L in (2.6), which was shown to have a spectrum consisting of quartets $\{\lambda, \lambda^*, -\lambda, -\lambda^*\}$ of eigenvalues. This has the immediate consequence that an odd number of interacting (complex) modes will exhibit at least two purely real or purely imaginary eigenvalues, implying either a growth instability or oscillatory dynamics, respectively. The IOEs are characterized by the matrix e^{izL}, which was shown to be symplectic. This allowed the group to identify a simple singular value decomposition for the matrix and to obtain an alternative proof of the eigenvalue quartets.

Of particular interest was the prospect of relating the spectrum and Schmidt decomposition (singular values) of the solution matrix e^{izL}. Unfortunately, several examples show that there is no discernible relationship between the eigenvalues and singular values apart from the obvious product of their moduli, except in the trivial case of a Hermitian matrix.

The group also considered two particular two-mode cases. In both cases, the signal and idler modes interact only because of the pump-induced fiber nonlinearity. In the undepleted-pump approximation, this coupling is linear in the signal and idler (sideband) amplitudes. In the isotropic (non-birefringent) case, the sideband polarization components have the same wavenumber, so the coupling term is δI. In this case (and, more generally, whenever $JK = KJ^*$), the eigenvalues and eigenvectors of L can be constructed entirely from the singular vectors of the symmetric submatrix K. The eigenvalues are either real or pure imaginary.

In the non-isotropic (birefringent) case, the polarization components have different wavenumbers. If one transforms out the average wavenumber, the differences that remain produce the δP term. In this case the eigenvalues can be shown to be real or pure imaginary if a certain criterion involving the matrix entries is satisfied.

Further work suggested by the working group's results center primarily on extending the two-mode cases considered to an arbitrary number of interacting modes, specifically to determine when sets of eigenvalues would be either purely real (corresponding to oscillations) or purely imaginary (corresponding to a growth instability). The cases considered also assume a very simple linear form for the modal interactions, posing the question of whether the constructive method used to determine the spectrum could be extended to more general linear interaction terms.

Nomenclature

The equation number where a particular quantity first appears or is defined is listed, if appropriate.

$\mathbf{A}$: amplitude vector in 2-mode case (5.1).
A: skew-Hermitian matrix in $\mathbf{x}$ system (2.1).
B: symmetric matrix in $\mathbf{x}$ system (2.1).
C: arbitrary matrix, variously defined.
d: discriminant.
H: Hermitian matrix (6.2).
i: indexing variable, variously defined.
J: Hermitian block of L (2.5).
j: indexing variable, variously defined.
K: symmetric block of L (2.5).
L: matrix in $\mathbf{y}$ system (2.7b).
$M(z)$: transfer matrix in $\mathbf{x}$ system (2.3).
$N(z)$: transfer matrix in $\mathbf{x}$ system (2.3).
n: dimension of original system.
P: matrix in two-mode case (6.1).
p: coupling coefficient in two-mode case (6.6).
S: matrix of eigenvectors of L (2.10).
T: symplectic matrix (4.9).
$U(z)$: unitary matrix in Schmidt factorization of M and N (4.1).
$\mathbf{u}$: column of U.
$V(z)$: unitary matrix in Schmidt factorization of M and N (4.1).
$\mathbf{v}$: column of V.
$\mathbf{w}$: arbitrary vector, variously defined.
$\mathbf{x}$: n-dimensional vector of mode amplitudes (2.1).
$\mathbf{y}$: $2n$-dimensional vector composed of $\mathbf{x}$ and $\mathbf{x}^*$ (2.7a).
$\mathbf{z}$: eigenvector, variously defined.
z: distance (1.1).
α: constant, variously defined.
β: arbitrary constant, variously defined.
γ: nonlinear fiber constant (1.1).
δ: self-coupling constant in two-mode case (5.1).
Λ: diagonal matrix of eigenvalues of L (2.10).
λ: eigenvalue, variously defined.
μ: eigenvalue of T.
$\Sigma(z)$: diagonal positive semidefinite matrix in Schmidt factorization of M and N (4.1).
$\psi(z)$: amplitude in Shrödinger equation (1.1).

Ω: matrix used in definition of symplecticity (4.9).

Other Notation

i: as a subscript on $\mathbf{A}$, used to indicate idler (5.1).
s: as a subscript on $\mathbf{A}$, used to indicate signal (5.1).
t: as a superscript, used to indicate transpose (2.2).
$\dagger$: as a superscript, used to indicate conjugate transpose (2.2).
$*$: as a superscript, used to indicate conjugate (2.1).

References

J. Fan, A. Migdall and L. Wang, "A twin photon source based on optical fiber," Opt. Photon. News, **18** (3) 26–33 (2007).

A. Hasegawa and F. Tappert, "Transmission of Stationary Nonlinear Optical Physics in Dispersive Dielectric Fibers I: Anomalous Dispersion," Appl. Phys. Lett., **23** (3), 142–144 (1973).

A. Hasegawa and F. Tappert, "Transmission of Stationary Nonlinear Optical Physics in Dispersive Dielectric Fibers II: Normal Dispersion," Appl. Phys. Lett., **23** (4), 171–172 (1973).

D. S. Mackey and N. Mackey, "On the Determinant of Symplectic Matrices," Numerical Analysis Report No. 422, Manchester Centre for Computational Mathematics, Manchester, England. 2003.

C. J. McKinstrie, "Unitary and Singular Value Decompositions of Parametric Processes in Fibers," *Opt. Commun.* **282**, 583–593 (2009). (and references therein)

C. J. McKinstrie and N. Alic, "Information Efficiencies of Parametric Devices," to appear in J. Sel. Top. Quantum Electron.

C. J. McKinstrie, H. Kogelnik, R. M. Jopson, S. Radic and A. V. Kanaev, "Four-wave Mixing in Fibers with Random Birefringence," Opt. Express **12**, 2033–2055 (2004).

C. J. McKinstrie, S. Radic and A. H. Gnauck, "All-optical signal processing by fiber-based parametric devices," Opt. Photon. News **18** (3), 34–40 (2007).