General coupled mode theory in non-Hermitian waveguides

Jing Xu[1,2] **and Yuntian Chen**[1,2,*]

[1] *School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China*
[2] *Wuhan National Laboratory of Optoelectronics, Huazhong University of Science and Technology, Wuhan, China*

[*]*yuntian@hust.edu.cn*

Abstract: In the presence of loss and gain, the coupled mode equation on describing the mode hybridization of various waveguides or cavities, or cavities coupled to waveguides becomes intrinsically non-Hermitian. In such non-Hermitian waveguides, the standard coupled mode theory fails. We generalize the coupled mode theory with a properly defined inner product based on reaction conservation. We apply our theory to the non-Hermitian parity-time symmetric waveguides, and obtain excellent agreement with results obtained by finite element fullwave simulations. The theory presented here is typically formulated in space to study coupling between waveguides, which can be transformed into time domain by proper reformulation to study coupling between non-Hermitian resonators. Our theory has the strength of studying non-Hermitian optical systems with inclusion of the full vector fields, thus is useful to study and design non-Hermitian devices that support asymmetric and even nonreciprocal light propagations.

OCIS codes: (240.6680) Surface plasmons; (230.7370) Waveguides; (230.6080) Sources.

References and links

1. S. A. Schelkunoff, "Conversion of Maxwell's equations into generalized telegraphist's equations," Bell Syst. Tech. J. **34**, 995–1043 (1955).
2. H. A. Haus, "Electron beam waves in microwave tubes," Proc. Symp. Electronic Waveguides, Polytechnic Inst. of Brooklyn, NY, 1958.
3. W. H. Louisell, *Coupled-Mode and Parametric Electronics* (Wiley, 1960).
4. D. Marcuse, "The coupling of degenerate modes in two parallel dielectric waveguides," Bell Syst. Tech. J. **50**, 1791–1816 (1971).
5. A. W. Snyder, "Coupled-mode theory for optical fibers," J. Opt. Soc. Am. **62,** 1267–1277 (1972).
6. A. Yariv, "Coupled-mode theory for guided-wave optics," J. Quantum Electron. **9**, 919–933 (1973).
7. A. Hardy and W. Streifer, "Coupled mode theory of parallel waveguides," J. Lightwave Technol. **3**, 1135–1146 (1985).
8. H. A. Haus, W. P. Huang, S. Kawakami, and N. A. Whitaker, "Coupled-mode theory of optical waveguides," J. Lightwave Technol. **5**, 16–23 (1987).
9. S. L. Chuang, "A coupled mode formulation by reciprocity and a variation principle," J. Lightwave Technol. **5**, 5–15 (1987).
10. W. Huang, "Coupled-mode theory for optical waveguides: an overview," J. Opt. Soc. Am. A **11**, 963–983 (1994).
11. S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, M. J. Khan, C. Manolatou, and H. A. Haus, "Theoretical analysis of channel drop tunneling processes," Phys. Rev. B **59**, 15882–15892 (1999).
12. Y. Xu, Y. Li, R. K. Lee, and A. Yariv, "Scattering-theory analysis of waveguide-resonator coupling," Phys. Rev. E **62**, 7389–7404 (2000).

13. S. Fan, W. Suh, and J. D. Joannopoulos, "Temporal coupled-mode theory for the Fano resonance in optical resonators," J. Opt. Soc. Am. B **20**, 569–572 (2003).
14. A. D. Berk, "Variational principles for electromagnetic resonators and waveguides," IRE IEEE Trans. Antennas Propag. **4**, 104–111 (1956).
15. R. EI-Ganainy, K. G. Makris, D. N. Christodoulides, and Z. H. Musslimani, "Theory of coupled optical $\mathscr{PT}$-symmetric structures," Opt. Lett. , **32**, 2632–2634 (2007).
16. M. Golshani, S. Weimann, Kh. Jafari, M. Khazaei Nezhad, A. Langari, A. R. Bahrampour, T. Eichelkraut, S. M. Mahdavi, and A. Szameit, "Impact of loss on the wave dynamics in photonic waveguide lattices," Phys. Rev. Lett. , **113**, 123903 (2014).
17. S. Klaiman, U. Gunther, and N. Moiseyev, "Visualization of branch points in $\mathscr{PT}$-symmetric waveguides," Phys. Rev. Lett. **101**, 080402 (2008).
18. L. Feng, Y.-L. Xu, W. S. Fegadolli, M.-H. Lu, J. E. B. Oliveira, V. R. Almeida, Y.-F. Chen and A. Scherer, "Experimental demonstration of a unidirectional reflectionless parity-time metamaterial at optical frequencies," Nature Mater. **12**, 108–113 (2013).
19. X. Zhu, H. Ramezani, C. Shi, J. Zhu, and X. Zhang, "$\mathscr{PT}$-symmetric acoustics," Phys. Rev. X **4**, 031042 (2014).
20. Z. Lin, H. Ramezani, T. Eichelkraut, T. Kottos, H. Cao, and D. N. Christodoulides, "Unidirectional invisibility induced by $\mathscr{PT}$-symmetric periodic structures," Phys. Rev. Lett. **106**, 213901 (2011).
21. Y. Shi, Z. Yu and S. Fan, "Limitations of nonlinear optical isolators due to dynamic reciprocity," Nature Photon. **9**, 388–392 (2015).
22. C. E. Ruter, K. G. Makris, R. El-Ganainy, D. N. Christodoulides, M. Segev and D. Kip, "Observation of parity-time symmetry in optics," Nature Phys. **6**, 192–195 (2010).
23. H. Hodaei, M.-A. Miri, M. Heinrich, D. N. Christodoulides, and M. Khajavikhan, "Parity-time -symmetric microring lasers," Science **346**, 975–978 (2014).
24. H. Alaeian and J. A. Dionne, "Non-Hermitian nanophotonic and plasmonic waveguides," Phys. Rev. B **89**, 075136 (2014).
25. H. Alaeian and J. A. Dionne, "Parity-time-symmetric plasmonic metamaterials," Phys. Rev. A **89**, 033829 (2014).
26. Y. Shen, X. H. Deng, and L. Chen, "Unidirectional invisibility in a two-layer non-PT-symmetric slab," Opt. Express **22**, 19440–19447 (2014).
27. V. H. Rumsey, "Reaction concept in electromagnetic theory," Phys. Rev. **94**, 1483–1491 (1954).
28. B. Friedman, *Principles and Techniques of Applied Mathematics* (John Wiley & Sons, 1962).
29. C. H. Chen and C.-D. Lien, "The variational principles for non-self-adjoint electromagnetic problems," IEEE Trans. Microwave Theory Tech. **28**, 878–886 (1980).
30. A. E. Siegman, *Lasers* (University Science Books, 1986).
31. P. Pintus, "Accurate vectorial finite element mode solver for magneto-opti and anisotropic waveguides," Opt. Express **22,** 15737-15756 (2014).
32. R. F. Harrington, *Time-Harmonic Electromagnetic Fields, 2rd* (Wiley-IEEE, 2001).
33. J. Mu and W. P. Huang, "Complex coupled-mode theory for tapered optical waveguides," Opt. Lett. **36,** 1026-1028 (2011).
34. G. Zhu, "Pseudo-hermitian hamltonian formalism of electromagnetic wave propagation in a dielectric medium-application to nonorthogonal coupled-mode theory," J. Lightw. Technol. **29**, 905-911 (2011).
35. J. Xu, B. Wu and Y. Chen, "Elimination of polarization degeneracy in circularly symmetric bianisotropic waveguides: a decoupled case," Opt. Express **23,** 11566-11575 (2015).

1. Introduction

Coupled mode theory (CMT) can be traced back to 3 decades ago [1], and has been an indispensable tool of analyzing and designing waveguides, resonators, couplers or many other devices from microwave frequency [2–4] to optical frequency [5–13], in both time and space domain. CMT has its mathematical root of variational principle [14], which yields a stationary value for target physical quantities of the coupled optical systems, such as propagation constant, eigen-frequencies, or impedance and so on. CMT was invented to study parametric amplifiers, oscillators, and frequency converters in microwaves [3], with a rigorous derivation given by Schelkunoff using mode expansion [1], and Haus using a variational principle [2]. The CMT for optical waveguides was developed by many authors [4–6], and further refined by Hardy et al., using non-orthogonal coupled mode formulation (NCMT) [7–9], in contrast to orthogonal coupled mode theory (OCMT). Subsequently, temporal coupled mode theory was applied to study coupling among optical resonators, or resonators coupled to waveguides [11–13].

It can be proved that CMT is exactly equivalent to Maxwell's equations, as long as a complete set of modes is taken in the mode expansion in constructing the coupled mode equations. In most cases, we only need a few modes (usually two) in the mode expansion, because all the other modes, for instance, the continuum modes in waveguides, do not couple to the target mode that we are interested in. Even though CMT remains approximate in the truncated mode set, yet insightful and often accurate description of target quantities of a coupled system can be obtained. By excluding those irrelevant modes, CMT for a two-modes structure can be reduced into a form of 2×2 matrix, which is a classic model and has been extensively used in describing many different coupled physical systems, such as two coupled mechanical oscillators, light (single-mode) matter (single atom) interaction in Jaynes-Cummings model, two coupled resonators in electric circuits and so on. Mostly, for a lossless system, the energy flow from one mode/oscillator/atom to another mode/oscillator/photon is reciprocal, namely, the energy will flow back in a 'reciprocal' way as 'time' is reversed, which renders the coupled system Hermitian, or the transition matrix of the coupled system as a Hermitian Matrix.

The aforementioned formulations of CMT rely on a definite and conserved optical power of the whole coupled system for either OCMT or NCMT, which essentially by default selects a scheme of complex inner product. However, for coupled optical parity-time ($\mathscr{PT}$)-symmetric structures [15–26] the total integrated power is not a conserved quantity, especially in the broken phase after the exceptional point. Therefore, the standard CMT fails in such non-Hermitian systems [15]. In [15], the authors formulate a CMT to studying $\mathscr{PT}$-symmetric structures through a Lagrangian treatment, in which the vector optical field is approximate as a scalar field envelope. In this work, we provide a conceptually simple model to construct general coupled mode theory (GCMT) using reaction conservation, a well-known concept in electromagnetism. The GCMT is capable of capturing the vectorial nature of the optical fields. Particularly, using a scalar inner product, we prove that the Maxwell equation remains self-adjoint, even in the presence of loss and gain. We further construct GCMT based on perturbation. In the application of GCMT, we study $\mathscr{PT}$-symmetric waveguides with balanced losses and gain. The prediction of our theory shows excellent agreement with results obtained by the fullwave simulation (COMSOL).

The paper is organized as follows. In Section 2, we give the foundation of our general coupled theory based on reaction concept. Secondly, we discuss the dimensional reduction and the procedures of constructing GCMT. In Section 3, we study the mode dispersion of $\mathscr{PT}$-symmetric waveguides using GCMT, and compare it with conventional coupled mode theory. Finally, Section 4 concludes the paper.

2. General couple mode theory

2.1. Reaction concept and self-adjointness

The reaction is a physical observable introduced by Rumsey [27] to measure the reaction between sources $\boldsymbol{S}_a$ and $\boldsymbol{S}_b$, defined as follows,

$$(\boldsymbol{F}_b, \boldsymbol{S}_a) = \langle \boldsymbol{F}_b, \sigma \boldsymbol{S}_a \rangle \tag{1}$$

where $\boldsymbol{F}_b$ is the field generated by source $\boldsymbol{S}_b$. As seen from Eq. (1), the inner-product $(\cdot|\cdot)$ is defined for the vector fields, and equals $\langle \cdot \sigma \cdot \rangle$, where $\langle \boldsymbol{\phi}(\boldsymbol{r})|\boldsymbol{\psi}(\boldsymbol{r}) \rangle = \int d\boldsymbol{r} \boldsymbol{\phi}^T(\boldsymbol{r}) \boldsymbol{\psi}(\boldsymbol{r})$ for two column vector fields $\boldsymbol{\phi}(\boldsymbol{r})$, $\boldsymbol{\psi}(\boldsymbol{r})$. It is important to point out that there is no complex conjugate operation over the fields, therefore the relation $\langle \boldsymbol{\phi}(\boldsymbol{r})|\boldsymbol{\psi}(\boldsymbol{r}) \rangle = \langle \boldsymbol{\psi}(\boldsymbol{r})|\boldsymbol{\phi}(\boldsymbol{r}) \rangle$ holds. Denoting the field $\boldsymbol{F} = [\boldsymbol{E}, \boldsymbol{H}]^T$, and $\boldsymbol{S} = [\boldsymbol{J}, \boldsymbol{M}]^T$, $\boldsymbol{J}$ ($\boldsymbol{M}$) electric (magnetic) current density. The reaction given in Eq. (1) can be explicitly given by

$$(\boldsymbol{F}_b, \boldsymbol{S}_a) = \int d\boldsymbol{r} \begin{bmatrix} \boldsymbol{E}_b & \boldsymbol{H}_b \end{bmatrix} \sigma \begin{bmatrix} \boldsymbol{J}_a \\ \boldsymbol{M}_a \end{bmatrix} \tag{2}$$

where the metric tensor $\sigma = \begin{pmatrix} \bar{\mathbf{1}} & \bar{\mathbf{0}} \\ \bar{\mathbf{0}} & -\bar{\mathbf{1}} \end{pmatrix}$, $\bar{\mathbf{1}}$ denotes the identity matrix, $\bar{\mathbf{0}}$ the zero matrix. The link between the source $\mathbf{S}$ and the associated vector fields $\mathbf{F}$ is given by the Maxwell's equations,

$$\bar{\mathbf{L}}\mathbf{F} = \bar{\mathbf{W}}\mathbf{S}, \tag{3}$$

where $\bar{\mathbf{L}} = \begin{pmatrix} \nabla\times & ik_0\bar{\boldsymbol{\mu}}_r \\ -ik_0\bar{\boldsymbol{\varepsilon}}_r & \nabla\times \end{pmatrix}$, $\bar{\mathbf{W}} = \begin{pmatrix} \bar{\mathbf{0}} & -\bar{\mathbf{1}} \\ \bar{\mathbf{1}} & \bar{\mathbf{0}} \end{pmatrix}$.

Based on the reaction concept, the reciprocity theorem can be given

$$(\mathbf{F}_a, \mathbf{S}_b) = (\mathbf{S}_a, \mathbf{F}_b), \tag{4}$$

which states that the response of one source to the external field induced by another source equals to the response of the second source to the field given by the first source. Equation (4) imposes certain constraints on the material parameters as given by $\bar{\boldsymbol{\varepsilon}}_r(\mathbf{r}) = \bar{\boldsymbol{\varepsilon}}_r^T(\mathbf{r})$ and $\bar{\boldsymbol{\mu}}_r(\mathbf{r}) = \bar{\boldsymbol{\mu}}_r^T(\mathbf{r})$. Such medium is called reciprocal medium in the literature, and could be lossy or active, as long as reciprocal conditions are fulfilled. We can reformulate Eq. (3) in a matrix form as follows,

$$\bar{\mathbf{H}}\mathbf{F} = \mathbf{S}, \tag{5}$$

where $\bar{\mathbf{H}} = \bar{\mathbf{W}}^{-1}\bar{\mathbf{L}}$. For any reciprocal medium, we find that for any two vector fields $\boldsymbol{\psi}$, $\boldsymbol{\phi}$, the following relation holds:

$$(\boldsymbol{\psi}, \bar{\mathbf{H}}\boldsymbol{\phi}) = (\bar{\mathbf{H}}\boldsymbol{\psi}, \boldsymbol{\phi}). \tag{6}$$

Equation (6) means that the operator $\bar{\mathbf{H}}$ is self-adjoint in the scheme of the inner-product given by Eq. (1), which has relevant consequences and implications that we intend to discuss. Firstly, the self-adjointness of the operator $\bar{\mathbf{H}}$ and the reaction for defining reciprocity share the same definition on the inner-product. Moreover, the self-adjointness of $\bar{\mathbf{H}}$ is equivalent to reciprocity theorem. In other words, the reciprocal conditions on the material parameters are necessary and sufficient condition both to self-adjointness of $\bar{\mathbf{H}}$, and to reciprocity. Secondly, the self-adjointness of operator $\bar{\mathbf{H}}$ guarantees that the underlying space associated with the operator $\bar{\mathbf{H}}$ is complete, despite the existence of losses or gain in the material parameters. Lastly, the left eigenvector space (bra space $\langle\cdot|$) and the right eigenvector space (ket space $|\cdot\rangle$) are closely related with each other for self-adjoint operator [28], namely, one determines the other, and vice versa. For complex inner-product, the self-adjoint operator is conventionally called as Hermitian operator. The matrix form of right and left eigenvectors are complex conjugate and transpose to each other, i.e., $|\cdot\rangle = [(\langle\cdot|)^T]^*$. As for a scalar inner-product, the matrix form of self-adjoint operator is symmetric, hence the left and right eigenvectors has the relation of $|\cdot\rangle = (\langle\cdot|)^T$.

2.2. Dimension reduction: a non self-adjoint formulation for waveguide problems based on variational principles

It is important to realize that the self-adjointness of the Maxwell's equations in reciprocal medium, as given in Eq. (6), is valid for the inner-product defined in 3D space. As for the development of GCMT or the modal solver based on variational principle for waveguide problems, it is necessary to reduce the 3D formulation into its 2D counterpart.

Considering an infinitely long waveguide, the waveguide modes of the original problem are given by $\boldsymbol{\phi} = [\mathbf{e}(\mathbf{r}), \mathbf{h}(\mathbf{r})]^T$, where $\mathbf{e}(\mathbf{r}) = \mathbf{e}(x, y, \beta)e^{i(\omega t - \beta z)}$, $\mathbf{h}(\mathbf{r}) = \mathbf{h}(x, y, \beta)e^{i(\omega t - \beta z)}$. The adjoint fields are given by $\boldsymbol{\psi} = [\mathbf{e}^a(\mathbf{r}), \mathbf{h}^a(\mathbf{r})]^T$, where $\mathbf{e}^a(\mathbf{r}) = \mathbf{e}^a(x, y)e^{i(\omega t + \beta z)}$, $\mathbf{h}^a(\mathbf{r}) = \mathbf{h}^a(x, y)e^{i(\omega t + \beta z)}$. We choose counter-propagating modes with the same β as our mode sets $\boldsymbol{\phi} = [\mathbf{e}^+, \mathbf{h}^+]^T e^{i(\omega t - \beta z)}$, and $\boldsymbol{\psi} = [\mathbf{e}^-, \mathbf{h}^-]^T e^{i(\omega t + \beta z)}$ respectively. The relation between the fields propagating in $+z$ and $-z$ direction is given by $\mathbf{e}^+ = \{e_x, e_y, e_z\}$, $\mathbf{h}^+ = \{h_x, h_y, h_z\}$,

$e^- = \{e_x, e_y, -e_z\}$, $h^- = \{-h_x, -h_y, h_z\}$. The particular choice of the mode sets [29–33] is relevant: (1) it is necessary to get a functional of coupled modes that is independent of z, meaning the terms $e^{\pm i\beta z}$ in the inner-product between the modes sets ϕ and ψ are canceled; (2) the mode pair of counter-propagating modes has a definite relation, hence ϕ can be deduced from ψ, and vice versa. It is easy to get the z-independent and source-free wave-equation for the mode profiles $\phi_{2d} = [e^+, h^+]$ of the original problem as follows

$$\bar{H}_{2d}\phi_{2d} = 0 \tag{7}$$

where $\bar{H}_{2d} = \begin{pmatrix} \nabla_t \times -i\beta z \times & ik_0\bar{\mu}_r \\ -ik_0\bar{\varepsilon}_r & \nabla_t \times -i\beta z \times \end{pmatrix}$, and $\bar{\varepsilon}_r^a = \begin{pmatrix} \varepsilon_r^{tt} & \varepsilon_r^{tz} \\ \varepsilon_r^{zt} & \varepsilon_r^{zz} \end{pmatrix}$, and $\bar{\mu}_r^a = \begin{pmatrix} \mu_r^{tt} & \mu_r^{tz} \\ \mu_r^{zt} & \mu_r^{zz} \end{pmatrix}$, and ε_r^{tt} (μ_r^{tt}) denotes the 2×2 in-plane components of electric (magnetic) dielectric function, $\nabla_t = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. From Eq. (7), and the predefined modes of ϕ and ψ, one can also obtain the adjoint system from the original ones ($\bar{H}_{2d}$), as given by

$$\bar{H}_{2d}^a \psi_{2d} = 0 \tag{8}$$

where $\psi_{2d} = [e^-, h^-]$, $\bar{H}_{2d}^a = \begin{pmatrix} \nabla_t \times +i\beta z \times & ik_0\bar{\mu}_r^a \\ -ik_0\bar{\varepsilon}_r^a & \nabla_t \times +i\beta z \times \end{pmatrix}$, $\bar{\varepsilon}_r^a = \begin{pmatrix} \varepsilon_r^{tt} & -\varepsilon_r^{tz} \\ -\varepsilon_r^{zt} & \varepsilon_r^{zz} \end{pmatrix}$, and $\bar{\mu}_r^a = \begin{pmatrix} \mu_r^{tt} & -\mu_r^{tz} \\ -\mu_r^{zt} & \mu_r^{zz} \end{pmatrix}$. It is easy to find that the following relation holds

$$(\psi_{2d}, \bar{H}_{2d}\phi_{2d}) = 0 = (\bar{H}_{2d}^a\psi_{2d}, \phi_{2d}), \tag{9}$$

where the inner-product is carried out over 2D computational domain, i.e., transverse plane of the waveguides. As regards to the comparison between 2D formula (Eq. (9)) and 3D formula (Eq. (6)), a few remarks may deserve attentions. Firstly, the operator H_{2d} with inner product defined over 2D domain is not self-adjoint [31,32,34], e.g., $\bar{H}_{2d}^a \neq \bar{H}_{2d}$, for reciprocal medium. Secondly, for self-adjoint system, the adjoint system and the original system can be treated separately. As for a non self-adjoint electromagnetic problem [29], one need to solve original problem, as well as its adjoint problem simultaneously to provide a complete but biorthogonal mode sets to construct the coupled mode equations, or any other modal solver based on variational principles, e.g., method of moments (MoM) and finite element method (FEM). As for Eq. (9), it is necessary to obtain a unified variational form that contains contribution from both H_{2d}^a and H_{2d}. The explicit unified variational form for eigen-mode problem of waveguides is the following,

$$Y = (\delta\psi_{2d}, \bar{H}_{2d}\phi_{2d}) + (\bar{H}_{2d}^a\psi_{2d}, \delta\phi_{2d}) = 0. \tag{10}$$

which is the first variation, e.g., $\delta\psi_{2d}$ and $\delta\phi_{2d}$, to the functional $I = (\psi_{2d}, \bar{H}_{2d}\phi_{2d})$. According to variational principle, Eq. (10) indicates simultaneously the optimal solution of ψ_{2d} and ϕ_{2d} to Eq. (7) and Eq. (8), as long as the functional I is stationary for any ψ_{2d} and ϕ_{2d}. Thirdly, the mode set (ψ_{2d}) of the adjoint system is selected as the countering propagating modes of the original system (ϕ_{2d}). As such, the total degree of freedom of unknows is halved [31]. We note that GCMT can also be applied to bianisotropic waveguides [35].

2.3. Procedures of constructing GCMT for waveguide problem based on perturbation

Given a perturbation to the adjoint system $\bar{H}_{2d}^a$ by $\Delta\bar{H}$, i.e., $\bar{H}_{2d}^\# = \bar{H}_{2d}^a + \Delta\bar{H}$, we shall have the relation according to linear response of Maxwell's equations, $(\psi_{2d}, \bar{H}_{2d}\phi_{2d}) = ([\bar{H}^\# - \Delta\bar{H}]\psi_{2d}, \phi_{2d})$, which can be reformulated as

$$(\psi_{2d}, \bar{H}_{2d}\phi_{2d}) - (\bar{H}_{2d}^\#\psi_{2d}, \phi_{2d}) = (-\Delta\bar{H}\psi_{2d}, \phi_{2d}). \tag{11}$$

Perturbation implies that $\Delta\bar{H}$ is small, hence $(-\Delta\bar{H}\boldsymbol{\psi}, \boldsymbol{\phi})$ can be approximately taken as 0, which leads to

$$(\boldsymbol{\psi}_{2d}, \bar{H}_{2d}\boldsymbol{\phi}_{2d}) - (\bar{H}_{2d}^{\#}\boldsymbol{\psi}_{2d}, \boldsymbol{\phi}_{2d}) = 0. \tag{12}$$

Equation (12) gives the connection between the original system and the perturbed adjoint system via the reaction conversation under a small perturbation $\Delta\bar{H}$, and can be transcribed into a set of coupled mode equations, which is essentially the GCMT proposed in this paper. Firstly, we use normalized fields $\boldsymbol{e}(\boldsymbol{r}) = \boldsymbol{e}^{+}e^{i(\omega t - \beta z)}$, $\boldsymbol{h}(\boldsymbol{r}) = \boldsymbol{h}^{+}e^{i(\omega t - \beta z)}$ propagating in the $+z$ direction, satisfying Maxwell equations

$$\nabla_t \times \boldsymbol{e}_{0,i}^{+} - i\beta_{0,i}\boldsymbol{z} \times \boldsymbol{e}_{0,i}^{+} = -ik_0\bar{\boldsymbol{\mu}}_r^0\boldsymbol{h}_{0,i}^{+}, \tag{13a}$$

$$\nabla_t \times \boldsymbol{h}_{0,i}^{+} - i\beta_{0,i}\boldsymbol{z} \times \boldsymbol{h}_{0,i}^{+} = ik_0\bar{\boldsymbol{\varepsilon}}_r^0\boldsymbol{e}_{0,i}^{+}, \tag{13b}$$

where the subscripts 0 and i stand for no perturbation case and mode labels, respectively.

We consider there is a small perturbation of ε. The idea is that the fields under perturbed system can be approximated by a linear combination of the unperturbated fields of the adjoint systems. Therefore, the fields of perturbed system could be written as $\boldsymbol{e}'(\boldsymbol{r}) = (\Sigma_j a_j \boldsymbol{e}_{0,j}^{-})e^{i(\omega t + \beta z)}$, $\boldsymbol{h}'(\boldsymbol{r}) = (\Sigma_j a_j \boldsymbol{h}_{0,j}^{-})e^{i(\omega t + \beta z)}$ and satisfy

$$\nabla_t \times \Sigma_j a_j \boldsymbol{e}_{0,j}^{-} + i\beta\boldsymbol{z} \times \Sigma_j a_j \boldsymbol{e}_{0,j}^{-} = -ik_0\bar{\boldsymbol{\mu}}_r\Sigma_j a_j \boldsymbol{h}_{0,j}^{-}, \tag{14a}$$

$$\nabla_t \times \Sigma_j a_j \boldsymbol{h}_{0,j}^{-} + i\beta\boldsymbol{z} \times \Sigma_j a_j \boldsymbol{h}_{0,j}^{-} = ik_0\bar{\boldsymbol{\varepsilon}}_r\Sigma_j a_j \boldsymbol{e}_{0,j}^{-}, \tag{14b}$$

In equivalence with Eq. (12), we derive GCMT from perturbation as follows,

$$\iint \{Eq.\ (13a) \cdot \Sigma_j a_j \boldsymbol{h}_{0,j}^{-} - Eq.\ (14b) \cdot \boldsymbol{e}_{0,i}^{+} + Eq.\ (13b) \cdot \Sigma_j a_j \boldsymbol{e}_{0,j}^{-} - Eq.\ (14a) \cdot \boldsymbol{h}_{0,i}^{+}\}dxdy \tag{15}$$

In case that a small perturbation is present in the imaginary part of $\bar{\boldsymbol{\varepsilon}}_r$, i.e. $\bar{\boldsymbol{\varepsilon}}_r = \bar{\boldsymbol{\varepsilon}}_r^0 + i\Delta\varepsilon(x,y)$, the formula resulted from Eq. (15) can be simplified as follows,

$$\Sigma_j a_j[k_{ij} + b_{ij} - i(\beta - \beta_{0,i})p_{ij}] = 0 \tag{16}$$

where $b_{ij} = \iint\{\nabla_t \cdot (\boldsymbol{h}_{0,i}^{+} \times \boldsymbol{e}_{0,j}^{-}) - \nabla_t \cdot (\boldsymbol{h}_{0,j}^{-} \times \boldsymbol{e}_{0,i}^{+})\}dxdy$, $p_{ij} = \iint\{\boldsymbol{z} \cdot (\boldsymbol{e}_{0,j}^{-} \times \boldsymbol{h}_{0,i}^{+}) - \boldsymbol{z} \cdot (\boldsymbol{e}_{0,i}^{+} \times \boldsymbol{h}_{0,j}^{-})\}dxdy$, $k_{ij} = ik_0 \iint \Delta\varepsilon(x,y)\boldsymbol{e}_{0,j}^{-} \cdot \boldsymbol{e}_{0,i}^{+}dxdy$. For $\Delta\varepsilon(x,y) = 0$, we shall have the following relation

$$b_{ij} = i\left(\beta_{0,j} - \beta_{0,i}\right)p_{ij}. \tag{17}$$

Inserting b_{ij} back into Eq. (16) yields

$$\Sigma_j a_j(\beta - \beta_{0,j})p_{ij} = \Sigma_j a_j k_{ij}. \tag{18}$$

Equation (18) is the matrix form of GCMT proposed in this paper, which will be used to study Hermitian and non-Hermitian waveguide with some concrete examples in the following section.

It is worthy to write down the CMT derived from conventional CMT (CCMT) [8] for comparison. To this end, Eq. (12) shall be reformulated as

$$(\boldsymbol{\psi}_{2d}^{*}, \bar{H}\boldsymbol{\phi}_{2d}) - ((\bar{H}^{\#}\boldsymbol{\psi}_{2d})^{*}, \boldsymbol{\phi}_{2d}) = 0, \tag{19}$$

where $*$ indicates the operation of complex conjugation, and the metric tensor is modified as $\sigma = \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}$ accordingly. Following the same procedure, we have

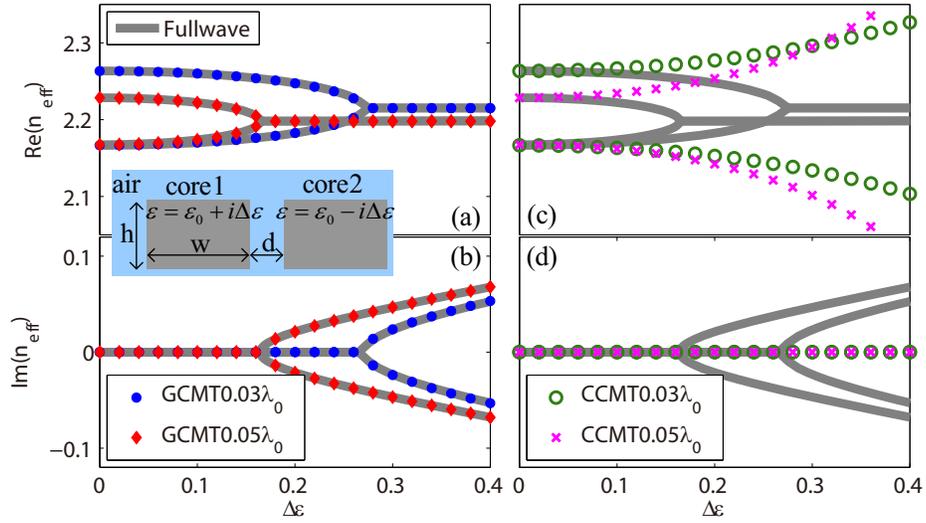$$\Sigma_j a_j(\beta^{*} - \beta_{0,j}^{*})p_{ij} = \Sigma_j a_j k_{ij}. \tag{20}$$

Fig. 1. Real part and imaginary part of effective mode indices (n_{eff}) versus $\Delta\varepsilon$ using GCMT (a,b) and CCMT (c,d). Gray solid lines are calculated from fullwave simulation. Inset shows the schematic diagram of two coupled core layers with loss and gain, surrounded by air. Dimensions are $h = 0.2\lambda_0$, $w = 0.3\lambda_0$, $d = 0.03\lambda_0$ (blue solid circles, green open circles) or $0.05\lambda_0$ (red diamonds, magenta crosses), $\varepsilon_0 = 10$. λ_0 is vacuum wavelength.

where $b_{ij} = \iint \{\nabla_t \cdot (\boldsymbol{h}_{0,i}^+ \times \boldsymbol{e}_{0,j}^*) + \nabla_t \cdot (\boldsymbol{h}_{0,j}^* \times \boldsymbol{e}_{0,i}^+)\}dxdy$, $p_{ij} = \iint \{\boldsymbol{z} \cdot (\boldsymbol{e}_{0,j}^* \times \boldsymbol{h}_{0,i}^+) + \boldsymbol{z} \cdot (\boldsymbol{e}_{0,i}^+ \times \boldsymbol{h}_{0,j}^*)\}dxdy$, $k_{ij} = ik_0 \iint \Delta\varepsilon(x,y)\boldsymbol{e}_{0,j}^* \cdot \boldsymbol{e}_{0,i}^+ dxdy$. In this case, for $\Delta\varepsilon(x,y) = 0$, $b_{ij} + f_{ij} = i\left(\beta_{0,j}^* - \beta_{0,i}\right)p_{ij}$ where $f_{ij} = -ik_0 \iint \{(\bar{\boldsymbol{\mu}}_{r,0} + \bar{\boldsymbol{\mu}}_{r,0}^{T,*})h_j^* \cdot h_i^+ + (\bar{\boldsymbol{\varepsilon}}_{r,0} + \bar{\boldsymbol{\varepsilon}}_{r,0}^{T,*})e_j^* \cdot e_i^+\}dxdy$. We will show in the next section that, CCMT works fine for Hermitian waveguides, but not for the case where non-Hermitian waveguides are considered.

3. Results and discussions

In the following, we use Eq. (18) to analyze dispersion relations in $\mathscr{PT}$-symmetric waveguides as discussed in [15, 17]. The structure of $\mathscr{PT}$-symmetric waveguides is shown by the inset of Fig. 1. It is composed of two waveguides with identical geometry dimensions placed close to each other. It is well known that the pair of an even and odd super mode is formed in this case. Bifurcation occurs as the magnitude of the imaginary part of $\bar{\boldsymbol{\varepsilon}}_r$ of two waveguides, i.e., $\bar{\boldsymbol{\varepsilon}}_r = \bar{\boldsymbol{\varepsilon}}_{r,0} + i\Delta\varepsilon$ in core layer 1 and $\bar{\boldsymbol{\varepsilon}}_r = \bar{\boldsymbol{\varepsilon}}_{r,0} - i\Delta\varepsilon$ in core layer 2, crosses a critical value as shown by the inset. This is used to create a symmetric index guiding profile and an anti-symmetric gain-loss profile. The gain/loss perturbation creates coupling between the odd and even mode pair so that the effective index of the two supermodes becomes closer and closer until an exceptional point where they become identical. Beyond the exceptional point, the real part of n_{eff} remains the same, but the imaginary part of n_{eff} of two modes break into two branches. When two waveguides are put more closer, the modes of two waveguides coupled more intensely so the even and odd supermodes have larger separation in the effective mode index. Therefore, it needs larger gain/loss parameter to get to the exceptional point. These are confirmed by COMSOL where two different gap size between two core layers are considered, as shown by the gray lines of Fig. 1.

Then we apply our theory to predict the dispersion curves of the above mentioned structures.
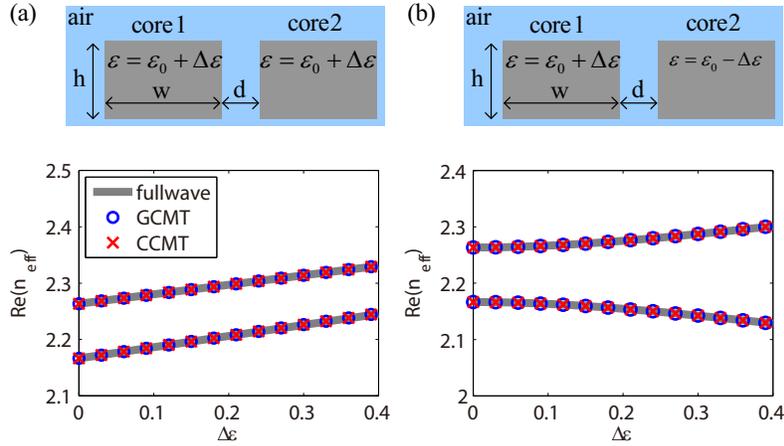
Fig. 2. Real part of n_{eff} versus $\Delta\varepsilon$. Gray solid lines are calculated from fullwave simulations. Blue open circles (red crosses) represent results derived by GCMT(CCMT). $d = 0.03\lambda_0$. Other parameters are the same as Fig. 1.

In this scenario, Eq. (18) is an eigenvalue problem with the following form

$$\left[\begin{array}{cc} \beta_{0,1}p_{11} - ik_{11} & \beta_{0,2}p_{12} - ik_{12} \\ \beta_{0,1}p_{21} - ik_{21} & \beta_{0,2}p_{22} - ik_{22} \end{array} \right] \left[\begin{array}{c} a_1 \\ a_2 \end{array} \right] = \beta \left[\begin{array}{cc} p_{11} & p_{12} \\ p_{21} & p_{22} \end{array} \right] \left[\begin{array}{c} a_1 \\ a_2 \end{array} \right] \qquad (21)$$

where p_{ij} and k_{ij} are defined in previous section. As a starting point, we use the mode fields provided by COMSOL at $\Delta\varepsilon = 0$. Propagation constants as well as eigenvectors are updated according to Eq. (21) using mode fields at this point. Next, in-plane fields are recalculated using updated eigenvectors and used for deriving propagation constants as well as eigenvectors in the next step. By choosing a small step, full dispersion relations as a function of $\Delta\varepsilon$ can be resolved. The solid symbols in Fig. 1(a) and 1(b) show effective mode indices given by $n_{eff} = \beta/k_0$ derived by our method, where they match results from COMSOL remarkably well. For comparison, we also show the results obtained by CCMT in Fig. 1(c) and 1(d). In this case Eq. (21) becomes

$$\left[\begin{array}{cc} \beta_{0,1}^{*}p_{11} - ik_{11} & \beta_{0,2}^{*}p_{12} - ik_{12} \\ \beta_{0,1}^{*}p_{21} - ik_{21} & \beta_{0,2}^{*}p_{22} - ik_{22} \end{array} \right] \left[\begin{array}{c} a_1 \\ a_2 \end{array} \right] = \beta^{*} \left[\begin{array}{cc} p_{11} & p_{12} \\ p_{21} & p_{22} \end{array} \right] \left[\begin{array}{c} a_1 \\ a_2 \end{array} \right]. \qquad (22)$$

It is clear that in this case CCMT fails to capture the major feature of $\mathscr{PT}$-symmetric waveguides. However, instead of perturbing $\bar{\boldsymbol{\varepsilon}}_r$ in the imaginary part, CCMT works fine for the case when $\bar{\boldsymbol{\varepsilon}}_r$ is present the real part. Red crosses in Fig. 2(a) shows the case that $\Delta\varepsilon$ is real and increases with identical sign in both core layers. In this case, the separation between two mode indices remain the same but their absolute values increases as $\Delta\varepsilon$ increases. Red crosses in Fig. 2(b) shows the case that $\Delta\varepsilon$ is real but increases with opposite sign in two core layers. In this case, two mode indices are further separated as $\Delta\varepsilon$ increases, indicating anti-crossing features. In both figures, only real part of n_{eff} is shown since imaginary part of n_{eff} in all cases is zero. Dispersion relations calculated according to GCMT are also shown in Fig. 2 with blue open circles. Clearly, GCMT developed in this work gives the same results as CCMT does, agreeing well with full wave simulations given by gray lines shown in Fig. 2.

4. Conclusions

From reaction conservation, we provide a solid foundation for the construction of general couple mode theory that can handle mode hybridization in non-Hermitian waveguides. Using a scalar inner product, we establish the equivalence between the self-adjointness of Maxwell's equations and reaction conservation. As for waveguide problems, the dimension of the self-adjoint relation need to be reduced from 3D to 2D, in which the formula turns out be non self-adjoint problem. Using coutering-propagating modes as the dual space of 2D non self-adjoint waveguide problem, the eigenmodes can be resolved from variational principles. Importantly, the 2D non self-adjoint relation can be elaborated into a set of coupled mode equation. We give a detailed discussion of the dimensional reduction for waveguide problem that relies on variational principle. We then provide a procedure of constructing GCMT for waveguide problem based on perturbation, which yields a set a coupled mode equations. To illustrate the effectiveness of GCMT developed in this work, it is applied to study the phase transition of coupled $\mathcal{PT}$-symmetric structures and shows excellent agreement with fullwave simulations. For comparison, results derived from CCMT are also shown, which fails to capture the major features of $\mathcal{PT}$-waveguides. Our theory provide direct analysis of eigenvalues of $\mathcal{PT}$-symmetric structures with in-cooperation of full vector fields. Thus it might be useful to study and design non-Hermitian devices that support asymmetric and even nonreciprocal light propagation.