Pulse propagation in the slow and stopped light regime

TAL A. WEISS[1] AND YONATAN SIVAN[2,*]

[1]*Department of Physics, Ben-Gurion University of the Negev, Israel*
[2]*Department of Engineering Sciences, Ben-Gurion University of the Negev, Israel*
*sivanyon@bgu.ac.il

Abstract: Electromagnetic pulse propagation in the slow light regime and near a zero group velocity point is relevant to a plethora of potential applications, and has analogies in numerous other wave systems. Unfortunately, the standard *frequency*-based formulation for pulse propagation is unsuitable for describing the dynamics in such regimes, due to the divergence of the dispersion coefficients. Moreover, in the presence of absorption, it is not clear how to interpret the propagation dynamics due to the drastic change induced by absorption upon the dispersion curves. As a remedy, we present an alternative *momentum*-based formulation, which is rapidly converging in these regimes, and naturally suitable for lossy and nonlinear media. It is specialized to a waveguide geometry which provides a significant simplification with respect to existing momentum-based schemes. Doing so, we provide a somewhat alternative, yet intuitive picture of the seeming enhanced absorption and nonlinear response in these regimes, and show that light-matter interactions are *not* enhanced in the slow/stopped light regimes.

References and links

1. J. R. Pierce, *Traveling-Wave Tubes* (D. Van Nostrand Company, Inc., 1950).
2. M. Fleischhauer, A. Imamoglu, and J. P. Marangos, "Electromagnetically induced transparency: optics in coherent media," Rev. Mod. Phys. **77**, 633–673 (2005).
3. P. Siddons, "Light propagation through atomic vapours," J. Phys. B: At. Mol. Opt. Phys **47**, 093001 (2014).
4. J. B. Khurgin, "Slow light in various media: a tutorial," Adv. Opt. Photonics **2**, 287–318 (2010).
5. R. W. Boyd, "Material slow light and structural slow light: similarities and differences for nonlinear optics," J. Opt. Soc. Am. B **28**, A38–A44 (2010).
6. T. Baba, "Slow light in photonic crystals," Nat. Photonics **2**, 465–473 (2008).
7. T. F. Krauss, "Slow light in photonic crystal waveguides," J. Phys. D **40**, 2666–2670 (2007).
8. G. Heinze, C. Hubrich, and T. Halfmann, "Stopped light and image storage by electromagnetically induced transparency up to the regime of one minute," Phys. Rev. Lett. **111**, 033601 (2013).
9. R. Won, "Quantum memory: Extended storage times," Nat. Photonics **7**, 677 (2013).
10. G. Brennen, E. Giacobino, and C. Simon, "Focus on quantum memory," New J. Phys. **17**, 050201 (2007).
11. T. Tanabe, M. Notomi, E. Kuramochi, A. Shinya, and H. Taniyama, "Trapping and delaying photons for one nanosecond in an ultrasmall high-Q photonic-crystal nanocavity," Nat. Photonics **1**, 49–52 (2007).
12. M. Notomi, E. Kuramochi, and T. Tanabe, "Large-scale arrays of ultrahigh-q coupled nanocavities," Nat. Photonics **2**, 741–747 (2008).
13. H. Takesue, N. Matsuda, E. Kuramochi, W. J. Munro, and M. Notomi, "An on-chip coupled resonator optical waveguide single-photon buffer," Nat. Commun. **4**, 2725 (2013).
14. C. Monat, B. Corcoran, D. Pudo, M. E. Heidari, C. Grillet, M. D. Pelusi, D. J. Moss, B. J. Eggleton, T. P. White, L. O'Faolain, and T. F. Krauss, "Slow light enhanced nonlinear optics in silicon photonic crystal waveguides," IEEE J. Sel. Top. Quantum Electron. **16**, 344–356 (2010).
15. S. Lavdas and N. Panoiu, "Theory of pulsed four-wave mixing in one-dimensional silicon photonic crystal slab waveguides," Phys. Rev. B **93**, 115435 (2016).
16. T. Kamalakis, "Derivation of coupled envelope propagation equations in silicon-based photonic crystal waveguides," J. Opt. Soc. Am. B **33**, 2339–2349 (2016).
17. X. Checoury, Z. Han, and P. Boucaud, "Stimulated Raman scattering in silicon photonic crystal waveguides under continuous excitation," Phys. Rev. B **82**, 041308 (2010).
18. J. F. McMillan, M. Yu, D.-L. Kwong, and C. W. Wong, "Observation of spontaneous raman scattering in silicon slow-light photonic crystal waveguides," Appl. Phys. Lett. **93**, 251105 (2008).

19. B. Corcoran, C. Monat, C. Grillet, D. J. Moss, B. J. Eggleton, T. P. White, L. O'Faolain, and T. F. Krauss, "Green light emission in silicon through slow-light enhanced third-harmonic generation in photonic-crystal waveguides," Nat. Photonics **3**, 206–210 (2009).

20. C. Monat, B. Corcoran, M. Ebnali-Heidari, C. Grillet, B. J. Eggleton, T. P. White, L. O'Faolain, and T. F. Krauss, "Slow light enhancement of nonlinear effects in silicon engineered photonic crystal waveguides," Opt. Express **17**, 2944–2953 (2009).

21. Y. Hamachi, S. Kubo, and T. Baba, "Slow light with low dispersion and nonlinear enhancement in a lattice-shifted photonic crystal waveguide," Opt. Lett. **34**, 1072–1074 (2009).

22. Y. Okawachi, M. S. Bigelow, J. E. Sharping, Z. Zhu, A. Schweinsberg, D. J. Gauthier, R. W. Boyd, and A. L. Gaeta, "Tunable all-optical delays via brillouin slow light in an optical fiber," Phys. Rev. Lett. **94**, 153902 (2005).

23. A. Zadok, A. Eyal, and M. Tur, "Stimulated brillouin scattering slow light in optical fibers," Appl. Opt. **50**, E38–E49 (2011).

24. M. Afzelius, N. Gisin, and H. de Riedmatten, "Quantum memory for photons," Phys. Today **68**, 42–47 (2015).

25. H. de Riedmatten, "A long-term memory for light," Physics **80**, 6 (2013).

26. F. Blatt, L. S. Simeonov, T. Halfmann, and T. Peters, "Stationary light pulses and narrowband light storage in a laser-cooled ensemble loaded into a hollow-core fiber," Phys. Rev. A **94**, 043833 (2016).

27. K. V. Rajitha, T. N. Dey, J. Evers, and M. Kiffner, "Pulse splitting in light propagation through N-type atomic media due to an interplay of Kerr nonlinearity and group-velocity dispersion," Phys. Rev. A **92**, 023840 (2015).

28. G. Grigoryan, V. Chaltykyan, E. Gazazyan, O. Tikhova, and V. Paturyan, "Pulse propagation, population transfer, and light storage in five-level media," Phys. Rev. A **91**, 023802 (2015).

29. Y. Chen, Z. Chen, and G. Huang, "Storage and retrieval of vector optical solitons via double electromagnetically induced transparency," Phys. Rev. A **91**, 023820 (2015).

30. L. V. Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 metres per second in an ultracold atomic gas," Nature **397**, 594–598 (1999).

31. M. M. Kash, V. A. Sautenkov, A. S. Zibrov, L. Hollberg, G. R. Welch, M. D. Lukin, Y. Rostovtsev, E. S. Fry, and M. O. Scully, "Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas," Phys. Rev. Lett. **82**, 5229–5232 (1999).

32. P. C. Kuan, C. Huang, W. S. Chan, S. Kosen, and S. Y. Lan, "Large fizeau's light-dragging effect in a moving electromagnetically induced transparent medium," Nat. Commun. **7**, 13030 (2016).

33. H. Zhang, M. Sabooni, L. Rippe, C. Kim, S. Kröll, L. V. Wang, and P. R. Hemmer, "Slow light for deep tissue imaging with ultrasound modulation," Appl. Phys. Lett. **100**, 131102 (2012).

34. M. S. Shahriar, G. S. Pati, R. Tripathi, V. Gopal, M. Messall, and K. Salit, "Ultrahigh enhancement in absolute and relative rotation sensing using fast and slow light," Phys. Rev. A **75**, 053807 (2007).

35. S. Wuestner, T. W. Pickering, J. M. Hamm, A. F. Page, A. Pusch, and O. Hess, "Ultrafast dynamics of nanoplasmonic stopped-light lasing," Faraday Discuss. **178**, 307 (2015).

36. A. F. Page, "Phd thesis," Ph.D. thesis, Imperial College London (2015).

37. K. L. Tsakmakidis, T. W. Pickering, J. M. Hamm, A. F. Page, and O. Hess, "Completely stopped and dispersionless light in plasmonic waveguides," Phys. Rev. Lett. **112**, 167401 (2014).

38. K. L. Tsakmakidis, T. W. Pickering, J. M. Hamm, A. F. Page, and O. Hess, "Cavity-free plasmonic nanolasing enabled by dispersionless stopped light," Nat. Commun. **5**, 4972–4979 (2014).

39. N. Sun, J. Chen, and D. Tang, "Stopping light in an active medium," The Eur. Phys. J. D **69**, 219 (2015).

40. J. I. L. Chen, G. von Freymann, S. Y. Choi, V. Kitaev, , and G. A. Ozin, "Amplified photochemistry with slow photons," Adv. Mater. **18**, 1915–1919 (2006).

41. J. I. L. Chen, G. von Freymann, S. Y. Choi, V. Kitaev, , and G. A. Ozin, "Slow photons in the fast lane in chemistry," J. Mater. Chem. **18**, 369–373 (2008).

42. U. Leonhardt, "A laboratory analogue of the event horizon using slow light in an atomic medium," Nature. **415**, 406–409 (2002).

43. M. Fink, "Acoustic time-reversal mirrors: Imaging of complex media with acoustic and seismic waves," Top. Appl. Phys. **84**, 17–43 (2002).

44. A. Hayrapetyan, K. Grigoryan, R. Petrosyan, and S. Fritzsche, "Propagation of sound waves through a spatially homogeneous but smoothly timedependent medium," Ann. Phys. **333**, 47–65 (2013).

45. H. Cui, W. Lin, H. Zhang, and X. Wang, "Backward waves with double zero-group-velocity points in a liquid-filled pipe," The J. Acoust. Soc. Am. **139**, 1179–1194 (2016).

46. A. Chumak, A. Karenowska, A. Serga, and B. Hillebrands, *in Topics in Applied Physics. Magnonics: from fundamentals to applications* (Springer, 2013).

47. V. Bacot, M. Labousse, A. Eddi, M. Fink, and E. Fort, "Revisiting time reversal and holography with spacetime transformations," Nat. Phys. **12**, 972–977 (2016).

48. C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases* (Cambridge, 2008), 2nd ed.

49. J. Mendonca and A. Guerreiro, "Time refraction and the quantum properties of vacuum," Phys. Rev. A **72**, 063805 (2005).

50. A. V. Turukhin, V. S. Sudarshanam, M. S. Shahriar, J. Musser, B. S. Ham, and P. R. Hemmer, "Observation of ultraslow and stored light pulses in a solid," Phys. Rev. Lett. **88**, 023602 (2002).

51. A. Figotin and I. Vitebskiy, "Slow light in photonic crystals," Waves Random Complex Media **16**, 293–382 (2006).

52. S. P. Shipman and A. T. Welters, "Pathological scattering by a defect in a slow-light periodic layered medium," J. Math. Phys. **57**, 022902 (2016).

53. G. P. Agrawal, *Nonlinear Fiber Optics* (Academics, San Diego, 2011).

54. G. Fibich, *The nonlinear Schrödinger equation, singular solutions and optical collapse* (Springer, Heidelberg, 2015).

55. F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," J. Opt. Soc. Am. B **25**, 1645–1654 (2008).

56. Y. Sivan, S. Rosenberg, and A. Halstuch, "Coupled-mode theory for electromagnetic pulse propagation in dispersive media undergoing a spatiotemporal perturbation: Exact derivation, numerical validation and peculiar wave mixing," Phys. Rev. B **93**, 144303 (2016).

57. P. Tran, "Photonic-band-structure calculation of material possessing kerr nonlinearity," Phys. Rev. B **52**, 10673–10676 (1995).

58. V. Lousse and J. P. Vigneron, "Self-consistent photonic band structure of dielectric superlattices containing nonlinear optical materials," Phys. Rev. E **63**, 027602 (2001).

59. B. Maes, P. Bienstman, and R. Baets, "Bloch modes and self-localized waveguides in nonlinear photonic crystals," J. Opt. Soc. Am. B **22**, 613–619 (2005).

60. C. de Sterke and J. Sipe, "Envelope-function approach for the electrodynamics of nonlinear periodic structures," Phys. Rev. A **38**, 5149–5165 (1988).

61. C. de Sterke and J. Sipe, "Switching dynamics of finite periodic nonlinear media: A numerical study," Phys. Rev. A **42**, 2858–2869 (1990).

62. J. Sipe, L. Poladian, and C. M. de Sterke, "Propagation through non-uniform grating structures," J. Opt. Soc. Am. B **11**, 1307–1320 (1994).

63. C. M. de Sterke, K. R. Jackson, and B. D. Robert, "Nonlinear coupled-mode equations on a finite interval: a numerical procedure," J. Opt. Soc. Am. B **8**, 403–412 (1991).

64. J. T. Mok, C. M. de Sterke, I. C. M. Littler, and B. J. Eggleton, "Dispersionless slow light using gap solitons," Nat. Phys. **2**, 775–780 (2006).

65. C. de Sterke, D. Salinas, and J. Sipe, "Coupled-mode theory for light propagation through deep nonlinear gratings," Phys. Rev. E **54**, 1969–1989 (1996).

66. N. Aközbek and S. John, "Optical solitary waves in two- and three-dimensional nonlinear photonic band-gap structures," Phys. Rev. E **57**, 2287–2319 (1998).

67. N. Bhat and J. Sipe, "Optical pulse propagation in nonlinear photonic crystals," Phys. Rev. E **64**, 056604 (2001).

68. J. E. Sipe, "Vector $k \cdot p$ approach for photonic band structures," Phys. Rev. E **62**, 5672–5677 (2000).

69. Y. Sivan and J. B. Pendry, "Time-reversal in dynamically-tuned zero-gap periodic systems," Phys. Rev. Lett. **106**, 193902 (2011).

70. Y. Sivan and J. B. Pendry, "Theory of wave-front reversal of short pulses in dynamically-tuned zero-gap periodic systems," Phys. Rev. A **84**, 033822 (2011).

71. M. Reza, M. M. Dignam, and S. Hughes, "Can light be stopped in realistic metamaterials?" Nature **455**, E10–E11 (2008).

72. J. Joannopoulos, S. Johnson, J. Winn, and R. Meade, *Photonic Crystals - Molding the Flow of Light* (Princeton University Press, 2008), 2nd ed.

73. W. Cai and V. Shalaev, *Optical metamaterials, fundamentals and applications* (Springer, Heidelberg, 2010).

74. P. Yeh, *Optical Waves in Layered Media* (Wiley-Interscience, 2005), 2nd ed.

75. P. Lalanne, W. Yan, K. Vynck, C. Sauvan, , and J.-P.Hugonin, "Light interaction with photonic and plasmonic resonances," ArXiv: 170502433 Phys. (2017).

76. L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamic of Continuous Media* (Pergamon Press, Kidlington, 2008).

77. S. G. Johnson, *Notes on the algebraic structure of wave equations* (http://math.mit.edu/ stevenj/18.369/wave-equations.pdf, 2010).

78. L. Thévenaz, S. Chin, I. Dicaire, J.-C. Beugnot, S. F. Mafang, and M. G. Herraez, "Experimental verification of the effect of slow light on molecular absorption," Proceeding SPIE **7503**, 75034W–1 (2009).

79. I. Dicaire, A. D. Rossi, S. Combrié, and L. Thévenaz, "Probing molecular absorption under slow-light propagation using a photonic crystal waveguide," Opt. Lett. **37**, 4934–4936 (2012).

80. A. Archambault, T. V. Teperik, F. Marquier, and J. J. Greffet, "Surface plasmon fourier optics," Phys. Rev. B **79**, 195414 (2009).

81. Z. I. Samson, P. Horak, K. F. MacDonald, and N. I. Zheludev, "Femtosecond surface plasmon pulse propagation," Opt. Lett. **36**, 250–252 (2011).

82. Q. Lin, O. J. Painter, and G. P. Agrawal, "Nonlinear optical phenomena in silicon waveguides: modeling and applications," Opt. Express **15**, 16604–16644 (2007).

83. C. M. Dissanayake, M. Premaratne, I. D. Rukhlenko, and G. P. Agrawal, "FDTD modeling of anisotropic nonlinear optical phenomena in silicon waveguides," Opt. Express **18**, 21427–21448 (2010).

84. R. Atkinson, W. R. Hendren, G. A. Wurtz, W. Dickson, A. V. Zayats, P. Evans, and R. J. Pollard, "Anisotropic optical properties of arrays of gold nanorods embedded in alumina," Phys. Rev. B **73**, 235402 (2006).

85. G. A. Wurtz, R. J. Pollard, W. R. Hendren, G. P. Wiederrecht, D. J. Gosztola, V. A. Podolskiy, and A. V. Zayats, "Designed ultrafast optical nonlinearity in a plasmonic nanorod metamaterial enhanced by nonlocality," Nat.

Nanotechnol. **6**, 107–111 (2011).

1. Introduction

Electromagnetic waves with slow propagation velocities (slow light) were observed and used in the microwave range as early as the 1940s [1]. In recent decades, slow light was demonstrated also at optical frequencies [2–7], and it was even shown that the light can nearly be brought to a standstill ("stopped light") [8–10]. As a result, various potential applications emerged, including in particular optical memory in optical communication systems [6, 11–13], nonlinear optics [7, 14–21], Brillouin scattering [22, 23], and more recently, quantum computing [10, 24–29], e.g., based on electromagnetic induced transparency [2, 3, 8, 9, 30–32]; Additional potential applications are tissue imaging [33], sensing [34], low threshold lasing [35–39], photochemistry [40, 41] and even tabletop cosmology [42]. Slow and stopped light regimes occur also in other wave systems like acoustic waves [43–45], spin waves [46], water waves [47], quantum/matter waves [48, 49] etc..

Slow light has been demonstrated in various platforms, including atomic vapours [2, 3, 30, 31], solids [50], and various photonic structures, such as Bragg gratings, photonic crystal fibers and waveguides etc., see [4–7, 51, 52].

In many studies (especially, in the context of atomic vapours), pulse propagation in the slow light regime was studied via a somewhat phenomenological transport equation for the pulse envelope [27–29], i.e., neglecting second- and higher-order dispersion terms. More accurate models (usually, for pulse propagation in waveguides or photonic crystals) were based on envelope equations like the nonlinear Schrödinger equation, see Section 2 and [53]. However, the validity of these models is limited to moderate levels of slow light. Otherwise (especially close to a zero group velocity point (ZGVP)), an exceeding number of dispersion terms (in the form of high-order *time* derivatives) has to be taken into account, because the various dispersion coefficients are proportional to the inverse of the group velocity (and its powers) and/or because of the slowly convergence of the series of higher-order derivatives of the propagation constant with respect to the frequency [54]. As a result, standard pulse propagation schemes (see e.g., [53, 54]), as well as coupled mode models (see e.g., [55, 56], which are all *frequency*-based expansions) become inefficient. In these cases, one is forced to resort to numerical solutions of the full Maxwell equations (e.g., using the Finite Differences Time Domain (FDTD) approach) which is far more demanding in terms of computational resources and run times [11–13, 37, 38, 57–59]; such an approach also hampers physical insights.

This computational burden could have been circumvented by using the efficient envelope formulations developed in the context of Bragg gratings. Initially assuming a shallow grating contrast, a scheme has been derived for pulse propagation near the ZGVP on the band edge, and most consequent work focused on the dynamics of nonlinear phenomena such as gap solitons [60–63] and their applications [64]. This formulation was then extended for deeper grating contrasts by de Sterke *et al.* [65] (and then, further extended to higher dimensions [66–68] and pump-probe configurations [69, 70]) by employing a $k \cdot p$ (i.e., *momentum*-based) expansion of the (diffraction [67] and) structural dispersion (sometimes referred to as waveguide dispersion) based on non-paraxial directional fluxes and Bloch carrier waves; the resulting expansion consists of a series of *spatial (longitudinal)* derivatives, instead of the standard *time* derivatives. In particular, in this derivation, the expansion coefficients are proportional to the group velocity (or its powers), rather than to its inverse. As a result, this formulation is especially efficient in the slow light and stopped light regimes, where the derivative series converges rapidly.

Unfortunately, despite the generality and accuracy of these rigorous models [65, 67], they are far from being in wide use. This is related, most likely, to their relative complexity. In the same vein, the (momentum-based) envelope formulations of [65, 67, 69, 70] were limited to configurations involving only a small number of not-so-short interacting pulses, in order to

avoid further complexity in the formulation arising from the need to account for adjacent bands. Finally, these formulations were limited to material constituents of weak dispersion. However, material dispersion can have an important role for semiconductor and plasmonic structures, and can become comparable to the structural dispersion. Such systems are also frequently lossy, so that there is also an ambiguity regarding the convenient way to include the absorption in the formulation [71]. Thus, overall, there are still no efficient (slowly-varying) envelope-based formulations that enable the study of pulse propagation in strongly dispersive and absorptive media, and specifically, in the slow and stopped light regimes.

In this paper, we aim to bridge this gap in the existing pulse propagation schemes. Specifically, in Section 3 we develop a (momentum-based) envelope formulation for pulse propagation in a system with translation invariance along one direction, i.e., a general waveguide system. We, thus, exploit the translation invariance along z to avoid the need for a Bloch carrier wave, hence, we work with the more convenient (and popular) Fourier carrier wave basis. Our approach shares the advantages of previous (momentum-based) formulations, but is more general than previous works [65, 67, 69, 70] as it is suitable also for non-periodic transverse dielectric structure, e.g., the abundant layered structures, photonic crystal fibers [72] etc.; these typically do not involve additional spectrally-adjacent bands.

In Section 4, we demonstrate the convergence of the derived model for the two generic cases - stopped and slow light. In order to emphasize the suitability of our approach to the deep slow light regime and for strongly dispersive materials, we model possibly the simplest geometry that supports a ZGVP and a slow light regime, namely, a plasmonic slot waveguide (metal-dielectric-metal (MDM) layer structure), see Fig. 1(a). The dispersion of the metallic constituent is indeed strong, and the metal constituent involves significant absorption. We demonstrate convergence of the derivative series upon the analytical solution. We also show that unlike the standard approach, where the nonlinearity and loss seem to diverge near the ZGVP, in our formulation, they are regular. This is explained by noting that the relevant criterion for the loss and nonlinear effect is the time rather than distance of propagation, such that the propagation time through a finite length waveguide segment diverges near a ZGVP [7, 14, 15]. More generally, we identify effects known from weakly dispersive systems, and discuss the interplay of dispersion, nonlinearity and absorption. We summarize the main findings and discuss future steps in Section 5.

2. Standard formulation - frequency-based expansion

The starting point of most derivations of pulse propagation schemes is the wave equation for a scalar electric field

$$\nabla^2 \mathcal{E}(\vec{r}, t) - \mu_0 \frac{\partial^2}{\partial t^2} (R(\vec{r}_\perp, t) * \mathcal{E}(\vec{r}, t)) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} (R_{NL}(\vec{r}, t) \mathcal{E}(\vec{r}, t)) = 0. \tag{1}$$

Here, $\mathcal{E}$ is the electric field, $\nabla^2 = \partial_z^2 + \partial_\perp^2$, μ_0 is the vacuum permeability, t is time, $\vec{r}$ is the position vector.

In this work, we focus on a waveguide geometry, so that the position vector, $\vec{r}$, is split into z, the propagation direction, and $\vec{r}_\perp$, the transverse direction; thus, the response function of the media $R(\vec{r}_\perp, t)$ and its Fourier transform, namely, the permittivity $\epsilon(\vec{r}_\perp, \omega)$, are functions of the transverse coordinate only. Lastly, $R_{NL}(\vec{r}, t)$ is a nonlinear response function (see e.g. [53] for explicit expressions) and the asterisk, '$*$', specifies a convolution. Eq. (1) is valid for slab waveguides (including free space) for TE polarization only. In more general cases, the field has more than one component and the first term of the left hand side of Eq. (1) has to be replaced by $\vec{\nabla} \times \vec{\nabla} \times \vec{\mathcal{E}}(\vec{r}, t)$. At this stage, we simplify the derivation by ignoring absorption; this effect will be fully accounted for when we present our alternative expansion.

In order to eliminate the fast oscillations on the scale of the optical period and wavelength, the electric field is written as a convolution of the transverse field profile and a slowly-varying
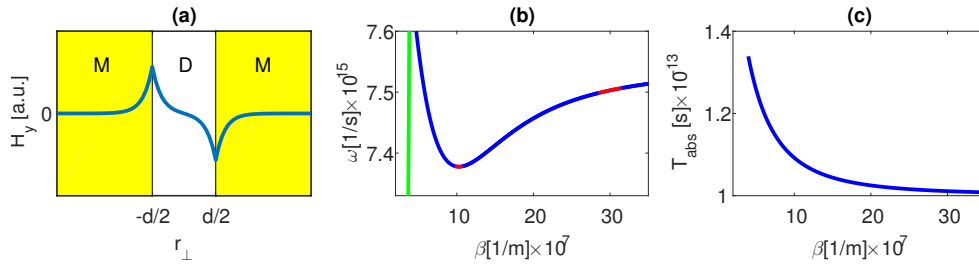
Fig. 1. (a) Cross section of a one dimensional plasmonic MDM waveguide of width d. The solid line features a transverse profile of the magnetic field propagating in z direction, corresponding to the antisymmetric branch. (b) Dispersion curve of antisymmetric plasmonic modes in a MDM slab waveguide where the core width is 35 nm, the permittivity of the dielectric is $\epsilon_D = 2.5$ and the parameters of the permittivity of the silver (see Eq. (45)) are $\epsilon_\infty = 5$, $\omega_P = 1.4 \times 10^{16} s^{-1}$ and $\gamma = 3.2 \times 10^{13} s^{-1}$ [73]. In cases where we ignore absorption, we set $\gamma = 0$. The red segment on the left is a stopped light regime where forward and backward modes coexist. Here, $\beta_0 = 1.51 \times 10^8 m^{-1}$ ($\lambda_0 \approx 363.6$nm) and $\Delta\beta = 0.15 \times 10^8 m^{-1}$, where β_0 lies exactly at the ZGVP. The second red segment lies in a slow light regime where $\beta_0 = 3 \times 10^8 m^{-1}$ ($\lambda_0 \approx 362.9$nm) and $\Delta\beta = 0.3 \times 10^8 m^{-1}$. Here β_0 lies at a point where the group velocity is $5.51 \times 10^{-4} c$. The green line on the left represents the light line, corresponding to $c/\sqrt{\epsilon_D}$. In the presence of absorption, the dispersion curve (and the ZGVP) is shifted slightly (not shown). (c) The values of $T_{abs} = 1/\Im(\omega_0)$ corresponding to the mode shown in (b).

envelope [56], both written as sums of (strictly real) frequency components, i.e.,

$$\mathcal{E}(\vec{r}, t) = e^{i\beta_0 z} \mathcal{F}_\omega^t \left[\left(\int dt' A(z, t') e^{i(\omega - \omega_0) t'} \right) \left(\int dt'' \mathcal{F}_\omega^{t''} \left[\hat{e}(\vec{r}_\perp, \omega) \right] e^{i\omega t''} \right) \right], \tag{2}$$

where $A(z, t')$ is a dimensionless amplitude, and $\hat{e}(\vec{r}_\perp, \omega)$ is the transverse mode profile (customarily, associated with units of electric field), i.e., the solution of the Helmholtz equation,

$$\nabla_\perp^2 \hat{e}(\vec{r}_\perp, \omega) - \beta^2(\omega) \hat{e}(\vec{r}_\perp, \omega) + \mu_0 \epsilon(\vec{r}_\perp, \omega) \omega^2 \hat{e}(\vec{r}_\perp, \omega) = 0. \tag{3}$$

Further, ω is the angular frequency, ω_0 is the center frequency of the pulse, $\beta_0 = \beta(\omega_0) = \omega_0 n_{eff}/c$ is the corresponding (potentially complex) propagation constant and $\mathcal{F}_\omega^t$ is a Fourier transform from the frequency to the time domain. (Note that in some works, $\hat{e}(\vec{r}_\perp, \omega)$ is replaced by $\hat{e}(\vec{r}_\perp, \omega_0)$. While for single mode propagation, there is just insignificant differences between the two approaches, the latter approach leads to some spurious mode coupling [15].) Eq. (1) is now transformed into the frequency domain, yielding

$$\hat{a}(z, \omega - \omega_0) \nabla_\perp^2 \hat{e}(\vec{r}_\perp, \omega) + 2i\beta_0 \mu_0 \epsilon(\vec{r}_\perp, \omega) \hat{e}(\vec{r}_\perp, \omega) \frac{\partial \hat{a}(z, \omega - \omega_0)}{\partial z}$$
$$- \beta_0^2 \mu_0 \epsilon(\vec{r}_\perp, \omega) \hat{e}(\vec{r}_\perp, \omega) \hat{a}(z, \omega - \omega_0) + \omega^2 \mu_0 \epsilon(\vec{r}_\perp, \omega) \hat{e}(\vec{r}_\perp, \omega) \hat{a}(z, \omega - \omega_0)$$
$$- \frac{1}{c} (\omega - \omega_0)^2 \mathcal{F}_t^\omega (R_{NL}(\vec{r}, t) \mathcal{E}(\vec{r}, t)) = 0, \tag{4}$$

where $\mathcal{F}_t^\omega$ represents the inverse Fourier transform from the time domain to the frequency domain, and the convolution is transformed to frequency domain as follows

$$R(\vec{r}_\perp, t) * \mathcal{E}(\vec{r}, t) = \mathcal{F}_\omega^t \left[\epsilon(\vec{r}_\perp, \omega) \hat{e}(\vec{r}_\perp, \omega) \hat{a}(z, \omega - \omega_0) e^{i\beta_0 z} \right]. \tag{5}$$

Finally, the spectral envelope, $\hat{a}(z, \omega - \omega_0)$ is related to the time domain envelope via [53, 56],

$$\mathcal{F}_t^\omega \left[A(z, t) e^{-i\omega_0 t} \right] = \hat{a}(z, \omega - \omega_0). \tag{6}$$

Now, one Taylor expands $\beta^2(\omega)$ around ω_0 as follows

$$\beta^2(\omega) = \sum_{m=0}^{\infty} \frac{1}{m!} \frac{d^m(\beta^2)}{d\omega^m} \bigg|_{\omega_0} (\omega - \omega_0)^m. \tag{7}$$

Substituting Eqs. (3) and (7) in Eq. (4), considering instantaneous Kerr nonlinear response, i.e., $R_{NL} = \epsilon_{NL} |\mathcal{E}|^2$, where ϵ_{NL} is the nonlinear part of the permittivity, and transforming back to time domain gives

$$\frac{i}{2\beta_0} \frac{\partial^2 A(z, t)}{\partial z^2} + \frac{\partial A(z, t)}{\partial z} + \sum_{m=1}^{\infty} \hat{D}_m \frac{\partial^n A(z, t)}{\partial t^m} + \frac{i}{L_{NL}} |A(z, t)|^2 A(z, t) = 0, \tag{8}$$

where the dispersion coefficients $\hat{D}_m$ are defined as

$$\hat{D}_m = \frac{1}{2\beta_0} \frac{(-i)^{m-1}}{m!} \frac{d^m(\beta^2)}{d\omega^m} \bigg|_{\omega_0}, \tag{9}$$

and where

$$L_{NL} = \frac{2P_0}{\epsilon_{NL}\omega_0 \int |\hat{e}(\vec{r}_\perp, \omega_0)|^4 d\vec{r}_\perp}, \tag{10}$$

is the characteristic length scale for the manifestation of the nonlinear effect, where

$$P_0 = \frac{\beta_0}{\mu_0\omega_0} \int |\hat{e}(\vec{r}_\perp, \omega_0)|^2 d\vec{r}_\perp, \tag{11}$$

is the power of the mode at ω_0 [74].

Further simplification is achieved by transforming to a frame of reference co-moving with the pulse such that

$$z^* = z, \qquad t^* = t - z/v_g. \tag{12}$$

Eq. (8) then becomes

$$\frac{i}{2\beta_0} \frac{\partial^2 A(z^*, t^*)}{\partial z^{*2}} + \frac{\partial A(z^*, t^*)}{\partial z^*} - \frac{i}{2\beta_0 v_g} \frac{\partial^2 A(z^*, t^*)}{\partial z^* \partial t^*} - \frac{i}{2} \beta_0'' \frac{\partial^2 A(z^*, t^*)}{\partial t^{*2}}$$
$$+ \sum_{m=0}^{\infty} \hat{D}_m \frac{\partial^m A(z^*, t^*)}{\partial t^{*m}} A(z^*, t^*) + \frac{i}{L_{NL}} |A(z^*, t^*)|^2 A(z^*, t^*) = 0, \tag{13}$$

where $\beta_0'' \equiv \frac{d^2\beta}{d\omega^2}\big|_{\omega_0}$ (not to be confused with the popular notation for an imaginary value; the latter is denoted below explicitly.), and

$$v_g \equiv \left(\frac{d\beta}{d\omega} \right)^{-1} \bigg|_{\omega_0} = \beta_0'^{-1}, \tag{14}$$

is the group velocity of the mode at ω_0.

In order to understand the relative magnitude of the various terms in Eq. (13) we transform once again the coordinates as follows

$$\hat{z} = \frac{z^*}{L_{disp}}, \tag{15a}$$

$$\hat{t} = \frac{t^*}{T_P}, \tag{15b}$$

where T_P is the pulse duration and $L_{disp} = T_P^2/\beta_0''$ is the dispersion length. The new coordinates are now dimensionless, and the derivatives by these coordinates are $O(1)$. Eq. (13) now reads

$$\frac{i}{2\beta_0 L_{disp}}\frac{\partial^2 A(\hat{z},\hat{t})}{\partial \hat{z}^2} + \frac{\partial A(\hat{z},\hat{t})}{\partial \hat{z}} - \frac{i}{2\beta_0 v_g T_P}\frac{\partial^2 A(\hat{z},\hat{t})}{\partial \hat{z}\partial \hat{t}} - i\frac{\partial^2 A(\hat{z},\hat{t})}{\partial \hat{t}^2} + L_{disp}\sum_{m=3}^{\infty}\hat{D}_m\frac{\partial^m A(\hat{z},\hat{t})}{\partial t^{*m}}$$

$$+i\frac{L_{disp}}{L_{NL}}|A(\hat{z},\hat{t})|^2 A(\hat{z},\hat{t}) = 0. \tag{16}$$

Since the wavelength of the center mode is $\lambda_0 = \frac{2\pi n_{eff}}{\beta_0}$ and the pulse spatial width is $Z \sim v_g T_P$, the first (non-paraxial) term and third (mixed derivative) term in Eq. (16) scale as $\frac{\lambda_0}{L_{disp}}$ and $\frac{\lambda_0}{Z}$, respectively. When comparing the magnitudes of all terms, the non-paraxial and the mixed derivative terms are negligible in the standard scenario in which $\lambda_0 \ll Z \ll L_{disp}$. The dispersion terms are proportional to $\frac{\hat{D}_m}{T_P^m}$, such that the high-order terms in Eq. (16) can be neglected if

$$\hat{D}_m \ll T_P \hat{D}_{m-1}. \tag{17}$$

In general, this condition is satisfied for sufficiently long pulses. Then, if dispersion is accounted for only up to second-order, Eq. (13) reduces to

$$\frac{\partial A(z^*,t^*)}{\partial z^*} - \frac{i}{2}\beta_0''\frac{\partial^2 A(z^*,t^*)}{\partial t^{*2}} + \frac{i}{L_{NL}}|A(z^*,t^*)|^2 A(z^*,t^*) = 0, \tag{18}$$

which is the well-known nonlinear Schrödinger equation [53, 54].

Although this approach is well-accepted and efficient in most cases, it becomes inefficient in the slow light regime and completely breaks down at a zero group velocity point. The reason for the failure of Eq. (18) lies in the neglect of the mixed derivative term in Eq. (13), as it is inversely proportional to the group velocity. This term grows in magnitude in the slow light regime, and becomes singular at the ZGVP. Additional singularities arise at the high-order derivatives $\hat{D}_m$ since the m'th term consist of terms which are proportional to v_g^{-m}. Therefore, when approaching the slow light regime or a ZGVP, the derivative series cannot be truncated at the second-order time derivative, and higher-order terms have to be considered, making the standard expansion inefficient.

In order to circumvent these problems, we present below an alternative derivation.

3. Alternative formulation - momentum-based expansion

3.1. Derivation

The starting point of our derivation is the vectorial wave function

$$\vec{\nabla}\times\vec{\nabla}\times\vec{\mathcal{E}}(\vec{r},t) - \mu_0\frac{\partial^2}{\partial t^2}\left(R(\vec{r}_\perp,t) * \vec{\mathcal{E}}(\vec{r},t)\right) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(R_{NL}(\vec{r},t)\vec{\mathcal{E}}(\vec{r},t)\right) = 0. \tag{19}$$

where $\vec{\mathcal{E}} = (\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z)$ is the electric field vector. Eq. (19) is the vectorial version of Eq. (1); thus, it is valid for any waveguide cross section and any polarization. For lossy media, all coefficients in the following derivation attain complex values.

As in former approaches, the convolution is transformed to the frequency domain. However, unlike the standard approach, we do not proceed by separation of variables, replacement of the permittivity with the propagation constant, and Taylor expanding the latter around the central frequency. Instead, inspired by the momentum-based formulations [65, 67], we aim to decompose the fields into a sum of spatial frequencies, i.e., in momentum space. First, however, we expand

the permittivity around the central frequency, ω_0, and then transform back to the time domain, giving

$$R(\vec{r}_\perp, t) * \vec{\mathcal{E}}(\vec{r}_\perp, z, t) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \frac{d^n \epsilon(\vec{r}_\perp, \omega)}{d\omega^n}\bigg|_{\omega_0} \frac{\partial^n \vec{\mathcal{E}}(\vec{r}_\perp, z, t)}{\partial t^n}. \tag{20}$$

Now, we can rewrite the electric field and related physical quantities in terms of their *spatial* Fourier transforms, namely,

$$\vec{\mathcal{E}}(\vec{r}_\perp, z, t) = \mathcal{F}_\beta^z[\tilde{\vec{\mathcal{E}}}(\vec{r}_\perp, \beta, t)] = \mathcal{F}_\beta^z\left[\tilde{a}(\beta - \beta_0, t)\tilde{e}(r_\perp, \beta)\tilde{\vec{e}}^{-i\omega(\beta)t} \right], \tag{21}$$

where as in the standard derivation (Section 2), $\tilde{a}(\beta - \beta_0, t)$ is a dimensionless amplitude and $\tilde{\vec{e}}(\vec{r}_\perp, \omega)$ is the transverse mode profile, associated with units of electric field. Here, and in what follows, we denote the spatial Fourier transforms by a tilde symbol (), in order to distinguish them from the (more standard) temporal Fourier transforms used in the standard derivation (Section 2), which are denoted by a hat symbol ().

We emphasize that as we consider a *spatial* (rather than a *temporal*) Fourier transform, we inherently work with real β values. Thus, in the presence of absorption (i.e., when the permittivity of the constituents is complex), the frequency ω obtains complex values. In this case, the dispersion curve remains similar to the curve in the loss-free case (see Fig. 1) and therefore, the ZGVP and slow light regime are preserved [37, 38, 71]. This means that the evaluation of the modal frequencies and various other coefficients is done with complex arguments. This is an unavoidable complication of the use of complex frequencies [75].

Once the amplitude $\tilde{a}(\beta - \beta_0, t)$ is found, the electric field can be written as

$$\begin{aligned}
\vec{\mathcal{E}}(\vec{r}_\perp, z, t) &= e^{-i\omega_0 t}\mathcal{F}_\beta^z\left[\left(\int dz' A(z', t)e^{-i(\beta-\beta_0)z'} \right) \left(\int dz'' \mathcal{F}_\beta^{z''}\left[\tilde{\vec{e}}(\vec{r}_\perp, \beta) \right] e^{-i\beta z''} \right) \right] \\
&= e^{-i\omega_0 t} \int d\zeta A(z - \zeta, t)e^{i\beta_0 \zeta}\mathcal{F}_\beta^\zeta\left[\tilde{\vec{e}}(r_\perp, \beta) \right],
\end{aligned} \tag{22}$$

where $\omega_0 = \omega(\beta_0)$, where

$$A(z, t) = 2\pi e^{-i\beta_0 z}\mathcal{F}_\beta^z\left[a(\beta - \beta_0, t) \right], \tag{23}$$

$$a(\beta - \beta_0, t) = \tilde{a}(\beta - \beta_0, t)e^{-i[\omega(\beta) - \omega_0]t}, \tag{24}$$

and where $\tilde{\vec{e}}(\vec{r}_\perp, \beta) = (\tilde{e}_x, \tilde{e}_y, \tilde{e}_z)$ satisfies the vector Helmholtz equation

$$\begin{aligned}
&\left[\left(\frac{\partial^2 \tilde{e}_y}{\partial x \partial y} - \frac{\partial^2 \tilde{e}_x}{\partial y^2} + \beta^2 \tilde{e}_x + i\beta\frac{\partial^2 \tilde{e}_z}{\partial x} \right)\hat{x} + \left(\frac{\partial^2 \tilde{e}_x}{\partial x \partial y} - \frac{\partial^2 \tilde{e}_y}{\partial x^2} + \beta^2 \tilde{e}_y + i\beta\frac{\partial^2 \tilde{e}_z}{\partial y} \right)\hat{y} \right. \\
&\left. + \left(i\beta\left(\frac{\partial \tilde{e}_x}{\partial x} + \frac{\partial \tilde{e}_y}{\partial y} \right) - \frac{\partial^2 \tilde{e}_x}{\partial x^2} - \frac{\partial^2 \tilde{e}_y}{\partial y^2} \right)\hat{z} \right] + \mu_0\epsilon(\vec{r}_\perp, \beta)\omega^2\tilde{\vec{e}}(\vec{r}_\perp, \beta) = 0.
\end{aligned} \tag{25}$$

For simplicity, we omitted in Eq. (25) the arguments in the components of $\tilde{\vec{e}}(\vec{r}_\perp, \beta)$.

Substituting Eqs. (20) and (24) in Eq. (19), and transforming into momentum space gives

$$
\left[\left(\frac{\partial^2 \tilde{e}_y}{\partial x \partial y} - \frac{\partial^2 \tilde{e}_x}{\partial y^2} + \beta^2 \tilde{e}_x + i\beta \frac{\partial^2 \tilde{e}_z}{\partial x} \right) \hat{x} + \left(\frac{\partial^2 \tilde{e}_x}{\partial x \partial y} - \frac{\partial^2 \tilde{e}_y}{\partial x^2} + \beta^2 \tilde{e}_y + i\beta \frac{\partial^2 \tilde{e}_z}{\partial y} \right) \hat{y} \right.
$$

$$
\left. + \left(i\beta \left(\frac{\partial \tilde{e}_x}{\partial x} + \frac{\partial \tilde{e}_y}{\partial y} \right) - \frac{\partial^2 \tilde{e}_x}{\partial x^2} - \frac{\partial^2 \tilde{e}_y}{\partial y^2} \right) \hat{z} \right] a(\beta - \beta_0, t) + \mu_0 \omega_0^2 \epsilon(\vec{r}_\perp, \omega_0) \tilde{\vec{e}}(\vec{r}_\perp, \beta) a(\beta - \beta_0, t)
$$

$$
+ i\mu_0 \tilde{\vec{e}}(\vec{r}_\perp, \beta) \left(2\omega_0 \epsilon(\vec{r}_\perp) + \omega_0^2 \frac{d\epsilon(\vec{r}_\perp, \omega)}{d\omega} \right) \frac{\partial a(\beta - \beta_0, t)}{\partial t}
$$

$$
+ \mu_0 \tilde{\vec{e}}(\vec{r}_\perp, \beta) \sum_{n=2}^{\infty} i^n \left(\frac{1}{(n-2)!} \frac{d^{n-2}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-2}} \bigg|_{\omega_0} + \frac{2}{(n-1)!} \omega_0 \frac{d^{n-1}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-1}} \bigg|_{\omega_0} \right.
$$

$$
\left. + \frac{1}{n!} \omega_0^2 \frac{d^n \epsilon(\vec{r}_\perp, \omega)}{d\omega^n} \bigg|_{\omega_0} \right) \frac{\partial^n a(\beta - \beta_0, t)}{\partial t^n} + e^{i\omega_0 t} \frac{\partial^2}{\partial t^2} \mathcal{F}_z^\beta \left[R_{NL}(\vec{r}_\perp, z, t) \vec{\mathcal{E}}(\vec{r}_\perp, z, t) \right] = 0. \qquad (26)
$$

Substituting Eq. (25) in Eq. (26) allows us to replace the first two terms by $\epsilon(\vec{r}_\perp, \omega)\omega^2 a(\beta - \beta_0, t)\tilde{\vec{e}}(\vec{r}_\perp, \beta)$. We now get

$$
\left(\epsilon(\vec{r}_\perp, \omega)\omega^2 - \epsilon(\vec{r}_\perp, \omega_0)\omega_0^2 \right) a(\beta - \beta_0, t) \tilde{\vec{e}}(\vec{r}_\perp, \beta)
$$

$$
- i\tilde{\vec{e}}(\vec{r}_\perp, \beta) \left(2\omega_0 \epsilon(\vec{r}_\perp, \omega) + \omega_0^2 \frac{d\epsilon(\vec{r}_\perp, \omega)}{d\omega} \bigg|_{} \right) \frac{\partial a(\beta - \beta_0, t)}{\partial t}
$$

$$
- \tilde{\vec{e}}(\vec{r}_\perp, \beta) \sum_{n=2}^{\infty} i^n \left(\frac{1}{(n-2)!} \frac{d^{n-2}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-2}} \bigg|_{\omega_0} + \frac{2}{(n-1)!} \omega_0 \frac{d^{n-1}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-1}} \bigg|_{\omega_0} \right.
$$

$$
\left. + \frac{1}{n!} \omega_0^2 \frac{d^n \epsilon(\vec{r}_\perp, \omega)}{d\omega^n} \bigg|_{\omega_0} \right) \frac{\partial^n a(\beta - \beta_0, t)}{\partial t^n} - e^{i\omega_0 t} \frac{\partial^2}{\partial t^2} \mathcal{F}_z^\beta \left[R_{NL}(\vec{r}_\perp, z, t) \vec{\mathcal{E}}(\vec{r}_\perp, z, t) \right] = 0. \qquad (27)
$$

In order to remove the dependence on the transverse coordinate, we now multiply by $\tilde{\vec{e}}^*(\vec{r}_\perp, \beta)$ and integrate over $\vec{r}_\perp$. Then, Eq. (27) becomes

$$
\frac{\partial a(\beta - \beta_0, t)}{\partial t} + \frac{i}{\omega_0 U(\beta, \beta_0)} \left[\int \left(\epsilon(\vec{r}_\perp, \omega)\omega^2 - \epsilon(\vec{r}_\perp, \omega_0)\omega_0^2 \right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp \right] (\beta - \beta_0, t)
$$

$$
+ \frac{i}{\omega_0 U(\beta, \beta_0)} \sum_{n=2}^{\infty} i^n \left[\int \left(\frac{1}{(n-2)!} \frac{d^{n-2}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-2}} \bigg|_{\omega_0} + \frac{2}{(n-1)!} \omega_0 \frac{d^{n-1}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-1}} \bigg|_{\omega_0} \right. \right.
$$

$$
\left. \left. + \frac{1}{n!} \omega_0^2 \frac{d^n \epsilon(\vec{r}_\perp, \omega)}{d\omega^n} \bigg|_{\omega_0} \right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp \right] \frac{\partial^n a(\beta - \beta_0, t)}{\partial t^n}
$$

$$
+ \frac{i e^{i\omega_0 t}}{2\omega_0 U(\beta, \beta_0)} \frac{\partial^2}{\partial t^2} \int \mathcal{F}_z^\beta \left[R_{NL}(\vec{r}_\perp, z, t) \vec{\mathcal{E}}(\vec{r}_\perp, z, t) \right] \tilde{\vec{e}}^*(\vec{r}_\perp, \beta) d\vec{r}_\perp = 0, \qquad (28)
$$

where

$$
U(\beta, \beta_0) = \int \left(\epsilon(\vec{r}_\perp, \omega) + \frac{1}{2}\omega \frac{\epsilon(\vec{r}_\perp, \omega)}{d\omega} \right) \bigg|_{\omega(\beta_0)} |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp. \qquad (29)
$$

Note that the terms in parentheses are evaluated at $\omega_0 \equiv \omega(\beta_0)$ (rather than the other way around in the standard approach), but $\tilde{\vec{e}}(\vec{r}_\perp, \beta)$ is evaluated at a general β. Thus, one can see that (in the absence of absorption, Otherwise, U [76] becomes a non-trivial complex function.) $U(\beta_0, \beta_0)$ is related to the energy density of the carrier mode β_0 (see Appendix A), but $U(\beta, \beta_0)$ equals this

quantity only to leading-order. As noted many times in the past, the energy represents the most convenient (and natural) norm for solutions of the vector Helmholtz equation for the electric field (see e.g., discussion in [77]).

Since $\omega = \omega(\beta)$, we can Taylor expand the ω-dependent coefficients in terms of β, in analogy to the expansion of β^2 in terms of ω in the standard approach (7). Thus, Eq. (28) becomes

$$\frac{\partial a(\beta - \beta_0, t)}{\partial t} + i \sum_{m=1}^{\infty} \tilde{D}_m (\beta - \beta_0)^m a(\beta - \beta_0, t) + \sum_{n=2}^{\infty} \sum_{m=0}^{\infty} i^{n+1} \tilde{D}_{n,m} (\beta - \beta_0)^m \frac{\partial^n a(\beta - \beta_0, t)}{\partial t^n}$$

$$+ \frac{i e^{i\omega_0 t}}{2\omega_0 U(\beta, \beta_0)} \frac{\partial^2}{\partial t^2} \int \mathcal{F}_z^\beta \left[R_{NL}(\vec{r}_\perp, z, t) \vec{\mathcal{E}}(\vec{r}_\perp, z, t) \right] \tilde{\vec{e}}^*(\vec{r}_\perp, \beta) d\vec{r}_\perp = 0, \qquad (30)$$

where

$$\tilde{D}_m = \frac{1}{m!} \frac{1}{2\omega_0} \frac{d^m}{d\beta^m} \left[\frac{\int \left[\omega^2(\beta)\epsilon(\vec{r}_\perp, \omega(\beta)) - \omega_0^2 \epsilon(\vec{r}_\perp, \omega_0) \right] |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp}{U(\beta, \beta_0)} \right]\Bigg|_{\beta_0}, \qquad (31)$$

and

$$\tilde{D}_{n,m} = \frac{1}{m!} \frac{1}{2\omega_0} \frac{d^m}{d\beta^m} \left[\frac{1}{U(\beta, \beta_0)} \int \left(\frac{1}{(n-2)!} \frac{d^{n-2}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-2}} + \frac{2}{(n-1)!} \omega_0 \frac{d^{n-1}\epsilon(\vec{r}_\perp, \omega)}{d\omega^{n-1}} \right. \right.$$

$$\left. \left. + \frac{1}{n!} \omega_0^2 \frac{d^n \epsilon(\vec{r}_\perp, \omega)}{d\omega^n} \right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp \right]\Bigg|_{\omega_0}. \qquad (32)$$

Note that both $\tilde{D}_m$ and $\tilde{D}_{m,n}$ account for the dispersion of the material parameters, the mode and the "energy density" U; The index n represents the order of material dispersion, and the index m represents the order of the combined effect of material and structural dispersion. Note that our formulation involves a differentiation of the mode profiles themselves ($\tilde{\vec{e}}(\vec{r}_\perp, \beta)$), a complication which is avoided in the standard formulation. Such complication arises because the initial temporal Fourier transform (Eq. (20)) does not allow one to get rid of the mode profiles, as in the standard derivation.

Transforming Eq. (30) back to real space gives

$$i\tilde{D}_{2,0} \frac{A(z,t)}{\partial t^2} + \frac{\partial A(z,t)}{\partial t} + \omega_0' \frac{\partial A(z,t)}{\partial z} - i\tilde{D}_2 \frac{\partial^2 A(z,t)}{\partial z^2}$$

$$+ \sum_{m=3}^{\infty} (-i)^{m-1} \tilde{D}_m \frac{\partial^m A(z,t)}{\partial z^m} + \sum_{n=2}^{\infty} \sum_{\substack{m=0 \\ m \neq 0 \text{ if } n=2}}^{\infty} (-1)^m i^{n+m+1} \tilde{D}_{n,m} \frac{\partial^{n+m} A(z,t)}{\partial t^n \partial z^m}$$

$$+ i\frac{e^{i\omega_0 t}}{2\omega_0} \mathcal{F}_\beta^z \left[\frac{1}{U(\beta, \beta_0)} \frac{\partial^2}{\partial t^2} \int \mathcal{F}_z^\beta \left[R_{NL}(\vec{r}_\perp, z, t) \vec{\mathcal{E}}(\vec{r}_\perp, z, t) \right] \tilde{\vec{e}}^*(\vec{r}_\perp, \beta) d\vec{r}_\perp \right] = 0, \qquad (33)$$

where

$$\tilde{D}_2 = \frac{\omega_0''}{2} + \tilde{D}_{2,0} \omega_0'^2, \qquad (34)$$

and

$$\tilde{D}_{2,0} = \frac{1}{2\omega_0 U(\beta_0, \beta_0)} \int \left(2\epsilon(\vec{r}_\perp, \omega_0) + 4\omega_0 \frac{d\epsilon(\vec{r}_\perp, \omega)}{d\omega}\Bigg|_{\omega_0} \right.$$

$$\left. + \omega_0^2 \frac{d^2 \epsilon(r_\perp, \omega)}{d\omega^2}\Bigg|_{\omega_0} \right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta)|^2 d\vec{r}_\perp. \qquad (35)$$

In Eqs. (34)-(35), we used Eqs. (31)-(32), where it follows from the chain rule that $\frac{d\epsilon(\vec{r}_\perp,\omega)}{d\beta}\big|_{\beta_0} = \frac{d\epsilon(\vec{r}_\perp,\omega)}{d\omega}\big|_{\omega_0}\frac{d\omega}{d\beta}\big|_{\beta_0}$. The coefficients ω_0' and ω_0'' stand for $\frac{d\omega}{d\beta}\big|_{\beta_0}$ and $\frac{d^2\omega}{d\beta^2}\big|_{\beta_0}$, respectively.

Eq. (33) is our main result; it is an *exact* reduction of Maxwell's equations, analogous to the standard derivation, Eq. (8), but with the roles of z and t exchanged (as in previous momentum-based expansions [65, 67]). Thus, for example, it accounts for non-paraxiality via a second-order derivative in time rather than in space, i.e., in z, and for dispersion via a series of derivatives in space rather than time. (In a 3D configuration, the spatial derivatives also account for diffraction effects [67].) Another manifestation of the exchange of roles of time and space is seen through $\tilde{D}_2$, which represents the group velocity dispersion (GVD). It consists of two terms. The first is the second derivative of the dispersion curve, whereas the second term is the group velocity at the central wavenumber β_0 in the current approach, similarly to Eq. (14), defined as

$$v_g \equiv \Re\left(\omega_0'\right). \tag{36}$$

For materials with negligible absorption, the second term in Eq. (34) is proportional to v_g^2, in analogy to the GVD coefficient in the standard derivation (where the second term is proportional to v_g^{-2}, see Eq. (9) and [53, 54]). As in the standard (frequency-based) derivation, the relative magnitude of the two terms in Eq. (34) depends on the position of the central mode on the dispersion curve. However, near a ZGVP/in the slow light regime, the second contribution to the GVD *vanishes*, in contrast to the standard derivation, where the second term *diverges*, again, due to the exchanged roles of time and space.

The first two orders of dispersion are analogous to those in the standard derivation (8), however, the higher-order dispersion terms in (33) involve additional mixed derivative terms. In that respect, our derivation yields a more complicated model compared with the standard one.

Eq. (33) generalizes the final results of the previous momentum-based expansions [65, 67] to lossy and strongly dispersive media, including in particular, the slow light regime and near ZGVPs. Importantly, in these regimes, Eq. (33) has a major advantage. Indeed, the series of dispersion terms which diverges in the standard derivation (18) as v_g^{-m} is replaced by a series of terms that scale with v_g^m, hence, is rapidly converging. In addition, the high-order derivatives of ω, are also small in the slow light regime.

3.2. Dimensional analysis

We now assume that the nonlinear response is spatially-local, i.e., that R_{NL} represents a cubic (Kerr-like) response, $R_{NL} = \epsilon_{NL}|\vec{\mathcal{E}}|^2$; Furthermore, U (29) is evaluated at β_0 in the nonlinear term. In order to further simplify the analysis, we also would like to adopt the slowly-varying envelope approximation. However, this has to be done with care. Indeed, the first term in Eq. (33) (the second derivative in time) is not necessarily small, even if the envelope is slowly varying - it consists of slow changes in time ($\sim T_{disp}$, see below), but also a $(v_g/Z)^2$ scale change due to the propagation; this term is comparable to $\tilde{D}_{2,0}\omega_0'^2$ (which is dominant, as explained in Section 3.1, see Eq. (34)), as in the standard derivation [54]. Yet, the latter contribution can be removed after the coordinates are transformed to a frame co-moving with the pulse, namely,

$$z^* = z - v_g t, \tag{37a}$$

$$t^* = t. \tag{37b}$$

Accordingly, the derivatives of the coordinates transform as

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z^*}, \tag{38a}$$

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t^*} - v_g \frac{\partial}{\partial z^*}. \tag{38b}$$

Then, Eq. (33) becomes

$$
i\tilde{D}_{2,0}\frac{\partial^2 A(z^*,t^*)}{\partial t^{*2}} + \frac{\partial A(z^*,t^*)}{\partial t^*} + i\Im\left(\omega_0'\right)\frac{\partial A(z^*,t^*)}{\partial z^*}
$$
$$
-\left[\frac{i}{2}\omega_0'' - \left(2\Im\left(\omega_0'\right)v_g + i\Im\left(\omega_0'\right)^2\right)\tilde{D}_{2,0}\right]\frac{\partial^2 A(z^*,t^*)}{\partial z^{*2}}
$$
$$
-2iv_g\tilde{D}_{2,0}\frac{\partial^2 A(z^*,t^*)}{\partial z^*\partial t^*} + \sum_{m=3}^{\infty}(-i)^{m-1}\tilde{D}_m\frac{\partial^m A(z^*,t^*)}{\partial z^{*m}}
$$
$$
+\sum_{n=2}^{\infty}\sum_{\substack{m=0\\ m\neq 0 \text{ if } n=2}}^{\infty}\sum_{k=0}^{n}\binom{n}{k}i^{n+1}(-i)^m(-v_g)^k\tilde{D}_{n,m}\frac{\partial^{n+m}A(z^*,t^*)}{\partial t^{*n-k}\partial z^{*m+k}}
$$
$$
+\frac{i}{T_{NL}}|A(z^*,t^*)|^2 A(z^*,t^*) = 0, \tag{39}
$$

where

$$
T_{NL} = \frac{2U(\beta_0,\beta_0)}{\epsilon_{NL}\omega_0\int|\tilde{\vec{e}}(\vec{r}_\perp,\beta_0)|^4 d\vec{r}_\perp}, \tag{40}
$$

is the characteristic time for the nonlinear effect to become significant. For lossy media, ω_0' is complex, so that the $\partial/\partial t$ term does not vanish, but rather, leaves behind its imaginary part, $\Im\left(\omega_0'\right)$; note that it does not represent a velocity, but rather the dispersion of the absorption. Similarly, the $\tilde{D}_{2,0}\omega_0'^2$ term does not cancel completely either.

In order to reveal the hierarchy of the terms in Eq. (39), we normalize the coordinates as follows

$$
\tilde{z} = \frac{z^*}{Z}, \tag{41a}
$$

$$
\tilde{t} = \frac{t^*}{T_{disp}}, \tag{41b}
$$

where Z is the initial pulse width and $T_{disp} = Z^2/\omega_0''$ is the dispersion time, the analogue of the dispersion length L_{disp}. Then, Eq. (39) obtains the dimensionless form

$$
if\frac{\partial^2 A(\tilde{z},\tilde{t})}{\partial\tilde{t}^2} + \frac{\partial A(\tilde{z},\tilde{t})}{\partial\tilde{t}} - 2ig\frac{\partial^2 A(\tilde{z},\tilde{t})}{\partial\tilde{z}\partial\tilde{t}} + iq\frac{\partial A(\tilde{z},\tilde{t})}{\partial\tilde{z}} - \frac{i}{2}\frac{\partial^2 A(\tilde{z},\tilde{t})}{\partial\tilde{z}^2}
$$
$$
+s\frac{\partial^2 A(\tilde{z},\tilde{t})}{\partial\tilde{z}^2} + \sum_{m=3}^{\infty}(-i)^{m-1}T_{disp}\frac{\tilde{D}_m}{Z^m}\frac{\partial^m A(\tilde{z},\tilde{t})}{\partial\tilde{z}^m}
$$
$$
+\sum_{n=2}^{\infty}\sum_{\substack{m=0\\ m\neq 0 \text{ if } n=2}}^{\infty}\sum_{k=0}^{n}\binom{n}{k}i^{n+1}(-i)^m\left(\frac{(-v_g)}{Z}\right)^k\frac{\tilde{D}_{n,m}}{T_{disp}^{n-k-1}}\frac{1}{Z^m}\frac{\partial^{n+m}A(\tilde{z},\tilde{t})}{\partial\tilde{t}^{n-k}\partial\tilde{z}^{m+k}}
$$
$$
+\frac{T_{disp}}{T_{NL}}|A(\tilde{z},\tilde{t})|^2 A(\tilde{z},\tilde{t}) = 0, \tag{42}
$$

where the dimensionless coefficients are

$$f \equiv \frac{\tilde{D}_{2,0}}{T_{disp}}, \tag{43a}$$

$$g \equiv \frac{v_g \tilde{D}_{2,0}}{Z}, \tag{43b}$$

$$q \equiv \frac{\Im\left(\omega_0'\right) T_{disp}}{Z} = \frac{\Im\left(\omega_0'\right)}{\omega_0''} Z, \tag{43c}$$

$$s \equiv \frac{\Im\left(\omega_0'\right)}{\omega_0''} \tilde{D}_{2,0} \left(2v_g + i\Im\left(\omega_0'\right)\right). \tag{43d}$$

The second and fifth terms in Eq. (42) are of the order of unity, and typically the dominant.

Now, in order to estimate the dispersion terms, we assume that the waveguide includes highly dispersive constituents (e.g., metal layers as in Fig. 1(a)). Thus, the estimates that follow provide an upper bound for the magnitude of these terms, while for dielectric materials, the dispersion terms would be typically smaller. Thus, $\tilde{D}_{2,0}$ can be estimated as $1/\omega_0$ (see Appendix B). Then, $f \sim T_0/T_{disp}$ and $g \sim T_0/T_P$ where we defined the period $T_0 = 2\pi/\omega_0$, and the pulse duration, T_P, which scales as Z/v_g. Thus, in the typical case where $T_0 \ll T_P \ll T_{disp}$, we can omit the first (non-paraxiality) and the third (mixed derivative) terms in Eq. (42); this is analogous to the standard derivation (Section 2).

The fourth and the sixth terms represent the effect of the complex ω terms on the propagation in the presence of absorption, namely, they account for the first- and second-order dispersion of the absorption. Generically, the absorption dispersion is a weak effect even for strongly absorbing media, like metals; accordingly, it is safe to neglect these terms.

Taking into account the scaling of $\tilde{D}_m$ and $\tilde{D}_{n,m}$ (see Appendix B), one can see that the seventh and eighth terms scale as $(T_{disp}/T_0)(\beta_0 Z)^{-m}$ and $(T_{disp}/T_0)(\beta_0 Z)^{-m}(v_g T_{disp}/Z)^k (T_{disp}/T_0)^{-n}$, respectively. Typically, T_{disp}/T_0 is much greater than unity; $\beta_0 Z$ and $(v_g T_{disp}/Z)$ are also larger than unity, but not as much. Thus, since $m \geq 3$ in the seventh term, it is typically smaller than unity; further, since $n \geq k$ and T_{disp}/T_0 is larger than $v_g T_{disp}/Z$, it is possible to show that the eighth term is typically smaller than the seventh term in the slow and stopped light regimes. In the specific case of propagation at the ZGVP (slow light regime), all terms proportional to v_g vanish (are small). In all other cases, one has to carefully map the relative magnitudes of the many higher-order terms.

The scaling of the ninth (nonlinear) term turns naturally to be by $U(\beta_0, \beta_0)$ (rather than as P_0, as in the standard derivation, see Eq. (40)). In lossy media, T_{NL} becomes complex as it contains the complex values $U(\beta_0, \beta_0)$ and ω_0. Overall, the nonlinear term becomes important in cases where T_{NL} is comparable or shorter than T_{disp}.

If all mixed terms are indeed small, Eq. (39) becomes

$$\frac{\partial A(z,t)}{\partial t} - i\frac{\omega_0''}{2}\frac{\partial^2 A(z,t)}{\partial z^2} + \sum_{m=3}^{\infty}(-i)^{m-1}\tilde{D}_m\frac{\partial^m A(z,t)}{\partial z^m} + \frac{i}{T_{NL}}|A(z,t)|^2 A(z,t) = 0, \tag{44}$$

where the * symbol was removed for brevity such that z and t are the (dimensional) coordinates of the moving frame (37). In that sense, the additional complexity of the model does not exceed that of the standard model for pulse propagation (Section 2 and [53]).

3.3. Scaling with the group velocity

Many previous studies showed that linear effects such as absorption scale with v_g^{-1} and that cubic nonlinear effects scale with v_g^{-2}, see e.g., [7, 14, 15, 67]. The former effect appears only implicitly in the scaling of the linear loss term with the group velocity, e.g., via the dispersion

relations [71]; the latter is seen from the natural scaling of the cubic nonlinearity with the power P_0 (see Eq. (11)), or explicitly in the momentum-based expansions of de Sterke, Sipe *et al.* (e.g., [65, 67]) where the dynamics is described by directional fluxes.

This inverse proportionality with the group velocity implies that these effects are enhanced in the slow light regime, and diverge near a ZGVP! This behaviour motivated many studies of nonlinear effects in the slow light regime [7, 14–21, 67]. Yet, rather remarkably, our final equation *does not* show any particular enhancement of (linear and) nonlinear effects in the slow and stopped light regimes. The solution to this seeming paradox is that the strength of the various effects is related to the *time* of propagation, rather than to the *distance* of propagation. Thus, if the dynamics is monitored in time, there is no particular sensitivity of the strength of nonlinear effects on the group velocity. In other words, the nonlinear effects are enhanced for low group velocities since a longer time of propagation is required to traverse a given distance in space, but per unit time, the nonlinear effect has the same strength. This is seen implicitly in Fig. 1(c) that shows that T_{abs} is only weakly dependent on β, such that the attenuation scales linearly with t (and is independent of z). Furthermore, the nonlinear term in our formulation scales with the energy density rather than with the power, hence, is independent of v_g.

This result also shows that the enhancement of the nonlinear effects reported before were a result of the experimental configuration - the pulse is injected into a finite length nonlinear media, and the nonlinear effect is measured at the output. Thus, unlike many previous claims, light-matter interactions are not enhanced in the slow/stopped light regimes, and the observed enhancements of the (linear and) nonlinear effects is unrelated to the source of the strong dispersion, be it material or structural dispersion [78, 79]. An alternative experimental configuration where the absorption or nonlinearity are measured in systems with different group velocities for the *same duration* is expected to show the insensitivity to the group velocity we predict here.

The disappearance of the dependence on v_g is a result of two rather subtle mathematical differences our derivation has in comparison to previous studies. First, our formulation is based on momentum modes (rather than frequency modes). Second, unlike [67], we adopt a frame moving in *space* (see Eq. (37)) rather than in *time* (12).

4. Examples

We now demonstrate the validity of our derivation by solving Eq. (44). By its very nature, Eq. (44) is solved as an initial value problem (IVP) in *time*. Specifically, the initial condition is set as $a(\beta, t = 0) = Ce^{-\left(\frac{\beta - \beta_0}{\Delta\beta}\right)^2}$ where β_0 is the center of the spectrum, $\Delta\beta$ is the spectral width of the pulse and C is a normalization constant such that $A(z = 0, t = 0) = 1$. This is similar to how the Schrödinger equation appears in quantum mechanics, but is in contrast to the standard approach in optics, where the equation of propagation is solved as an IVP in *space*, see e.g., [53] and Section 2. Due to the above, it is also not straightforward to compare the two methods directly.

In order to demonstrate the dynamics of short pulses in the slow light regime and near a ZGVP, we consider the simplest geometry that supports such regimes, namely, an MDM waveguide (see Fig. 1(a)). The permittivity of the metal is assumed to be given by the Drude model,

$$\epsilon_M(\omega) = \epsilon_\infty - \frac{\omega_P^2}{\omega^2 + i\gamma\omega}, \tag{45}$$

where ϵ_∞ is a constant which determines the value of the permittivity when $\omega \to \infty$, ω_P is the plasma frequency and γ is a constant real number which relates to the absorption in the metal.

Determination of the coefficients in Eq. (44) requires calculating the dispersion relations, $\omega(\beta)$ and the mode profiles. If absorption is neglected ($\gamma = 0$), one finds a unique set of solutions which satisfy the dispersion relation, where both β and ω attain real values. However, in the presence of absorption, there are two generic sets of solutions to the dispersion relations [80].

The first possibility is to account for the absorption via complex-β values, as in the standard derivation (Section 2). In this case, the dispersion curve splits at the ZGVP, resulting in two branches of fast light [71]; in addition, the dispersion relation curve backbends near plasmon resonance [80]. We note that the standard approach does not totally fail in this case since the terms that include factors such as $1/v_g^m$ are large but do not diverge anymore. However, the slow light regime becomes limited, and it is not a-priori clear how to link the resulting complex modal structure and the original ZGVP that disappeared.

The second possibility is to account for the absorption via complex ω values. In this case, the dispersion curve remains similar to the curve in the loss-free case (see Fig. 1(a)-(b)) and therefore, the ZGVP and slow light regime are preserved [37, 38]. As noted above, our approach is inherently based on complex frequency modes, which thus allows for a convenient description of the pulse dynamics near a ZGVP [37, 38]. Note that while the mode profiles exhibit outgoing energy flow away from the interface, they are also exponentially decaying due to the negative real part of the metal permittivities in the cladding. In that respect, and unlike the structures studied in [35–38], the modes we study are below the light line, hence, they do not suffer from radiation losses. As a result, normalization of the modes is trivial, and free from the complications arising frequently in the normalization of exponentially growing complex frequency modes [75].

Once $\omega(\beta)$ was computed, we ensured that the non-paraxiality and various mixed derivative terms are indeed relatively small, so that Eq. (44) provides an excellent approximation for the exact model, Eq. (42).

When simulating pulse propagation according to Eq. (44) and considering absorption, we have to recall that $e^{-i\omega_0 t}$ in Eq. (22) also describes the attenuation of the pulse in time as ω_0 consists of an imaginary part. Therefore, in the following plots, we will present the value $|A(z,t)e^{-i\omega_0 t}|$ instead of just the amplitude, $A(z,t)$. In that sense, the characteristic time for absorption is $T_{abs} = 1/\Im(\omega_0)$.

4.1. Simulations of pulse propagation in the slow and stopped light regimes; linear media

We first study the linear dynamics by solving Eq. (44) without the nonlinear term. In order to test the convergence of the derivative series in Eq. (44), we set $\Omega^{(M)}(\beta) \equiv \sum_{m=1}^{M} \tilde{D}_m (\beta - \beta_0)^m$, such that the solution of Eq. (44) is

$$A(z,t) = 2\pi e^{-i\beta_0 z} \mathcal{F}_\beta^z \left[a(\beta - \beta_0, t = 0) e^{-i\Omega^M(\beta) t} \right]. \tag{46}$$

Here, M accounts for the order beyond which dispersion terms in Eq. (44) are truncated. For a reference, we use the solution of the non-truncated series, found by repeating the procedure with $\Omega^{(\infty)}(\beta) \equiv \omega(\beta)$. Since there is no source for backward scattering in our case, it can be shown that this approach is equivalent to solving exactly the vector wave equation (19), i.e., taking into account even the non-paraxiality and various mixed-derivative terms; see e.g., [81]. Hence, we can refer to the $\Omega^{(\infty)}$ solution as an exact reference.

Our first example presents pulse propagation in the stopped light regime, where β_0 was chosen to be at the ZGVP, see Fig. 1(b). The spectrum width is chosen to be $\Delta\beta \approx 0.1\beta_0$, to ensure convergence of the dispersion expansion. In Fig. 2(a), we demonstrate the rapid convergence of Eq. (44) upon the exact solution in the absence of absorption. Indeed, the simulation results for the field at $t = T_{disp}$ decently match the exact solution when $M = 2$, and becomes more accurate when $M = 3$. Specifically, the relative difference of the outcome was tested according to

$$R_{M,M_{max}}(t) = \frac{\int |A^{(M)}(z,t) - A^{(M_{max})}(z,t)|^2 dz}{\int |A^{(M_{max})}(z,t)|^2 dz}, \tag{47}$$

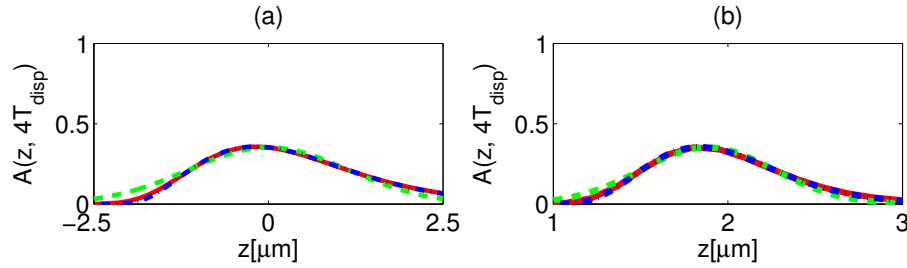where $A^{(M_{max})}$ is the reference (the most accurate simulation). In the linear simulation we set

Fig. 2. Pulse propagation at (a) ZGVP and (b) slow light regime, where the spectral width of the pulse, $\Delta\beta$ and the center value of the propagation constant, β_0 are the corresponding values shown in Fig. 1(b). The green dashed line is $A(z, t = T_{disp})$ where $M = 2$, the blue dotted line corresponds to $M = 3$ and the red solid line represents the exact solution $M = \infty$.

$M_{max} = \infty$. Here, we measured $R_{2,\infty} = 1.1$ and $R_{3,\infty} = 0.035$. This demonstrate the rapid convergence of Eq. (44) with low-order dispersion.

In the following simulations, we chose a narrower spectral regime such that $\Delta\beta = 0.01\beta_0$. The solution of Eq. (44) in the absence of absorption is shown in Fig. 3(a). The pulse is centered at $z = 0$ at all times, as expected, but broadens due to dispersion. In the presence of realistic absorption (see Fig. 3(b)), T_{abs} turns out to be much shorter than T_{disp}, such that the field decays rapidly within $t < T_{disp}$. Figs. 4(a)-(b) show the respective dynamics of the spatial spectrum of the pulse in these two cases. The spectrum does not change over time if we only consider dispersion, however, its intensity is attenuated in the presence of absorption, as expected.

In order to show the validity of Eq. (44) in the slow light regime, we repeated the process above for β_0 centered at the slow light regime where $v_g = 5.5 \times 10^{-4}c$ (see Fig. 1(b)). Here, as earlier, we choose $\Delta\beta \approx 0.1\beta_0$. However, although $\Delta\beta$ here is nominally wider than the spectral width chosen in the case of the ZGVP (since the value of β_0 is higher), the GVD and higher-order dispersion coefficients are much smaller. As a result, satisfactory accuracy is obtained even for $M = 2$, as shown in Fig. 2(b), where $R_{2,\infty} = 0.1$ and $R_{3,\infty} = 0.005$. This is, again, in contrast to the standard approach (Section 2), where due to the large factors $\sim v_g^{-m}$ in the mth-order terms, many more terms are required for convergence.

Fig. 3(c) shows the pulse dynamics in the reference of the moving pulse. The results are remarkably similar to the dynamics in the ZGVP (Fig. 3(a)). In that respect, it is worth emphasizing that one aspect of this similarity is somewhat surprising. Indeed, Fig. 1(c) shows that T_{abs} in the slow light regime is comparable to that at the ZVGP, although the latter is not as close to the SPP resonance. This can be understood by noting that the well-known enhanced absorption near SPP resonance is a result of the longer propagation time through a finite length waveguide segment, i.e., it is a direct result of the fact that the propagation is typically formulated as an IVP in space, an approach that makes the long propagation distance implicit (see also discussion in Section 3.3). Nevertheless, T_{abs} indeed becomes shorter when the pulse spectrum is near resonance. This is a result of the increasing portion of energy residing in the metal side. We will see below a similar behaviour in the context of the nonlinear regime.

4.2. Simulations of pulse propagation in the slow and stopped light regimes; nonlinear media

We now repeat the simulations shown in Section 4.1 in the presence of a nonlinear term, now employing the split step method [53]. Under these circumstances, one has to consider the interplay between dispersion, absorption and nonlinear dynamics via their characteristic time scales T_{disp},
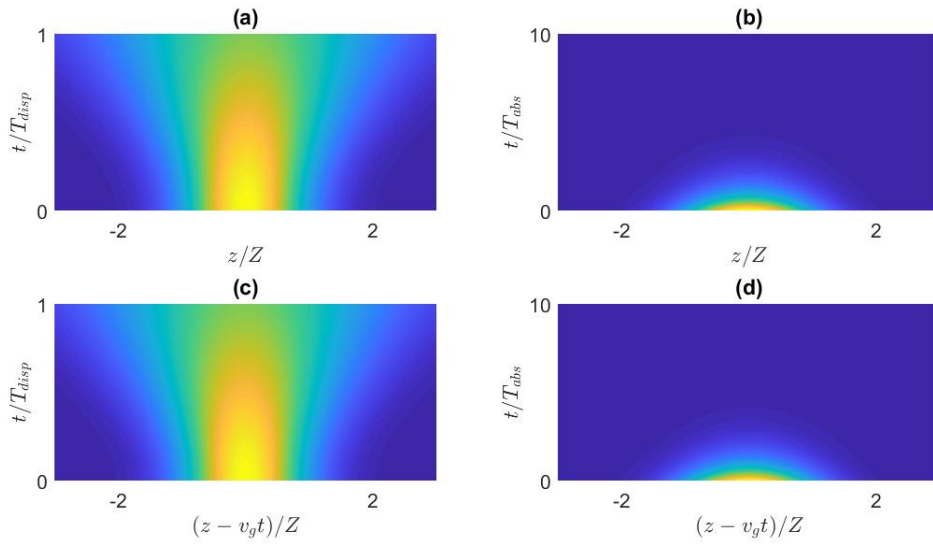
Fig. 3. Spatio-temporal pulse dynamics of the amplitude of the electric field (color represents the value of $|A(z,t)|e^{-\Im(\omega_0)t}$) in the linear case. Plots (a) and (b) represent the ZGVP regime case in the absence and presence of absorption, respectively. Here, $\beta_0 = 1.51 \times 10^8 \frac{1}{m}$, $\Delta\beta = 0.015 \times 10^8 \frac{1}{m}$, $T_{disp} = 7.7 \times 10^{-10}s$; in (b), $T_{abs} = 6.3 \times 10^{-14}s$, such that $T_{abs} \ll T_{disp}$. Plots (c) and (d) represent the slow light regime case in the absence and presence of absorption, respectively. Here, $\beta_0 = 3 \times 10^8 \frac{1}{m}$, $\Delta\beta = 0.03 \times 10^8 \frac{1}{m}$, $T_{disp} = 1.39 \times 10^{-9}s$, $T_{abs} = 6.26 \times 10^{-14}s$ in (d) such that $T_{abs} \ll T_{disp}$. In both cases, the total time of propagation is T_{disp} and $10T_{abs}$, respectively.
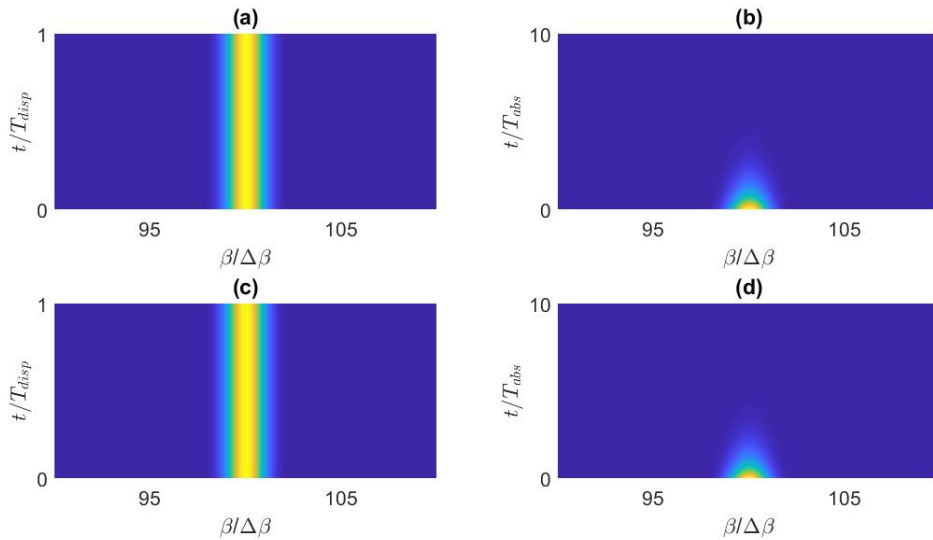

Fig. 4. The spatial spectrum associated with the results shown in Figs. 3(a)-(d).
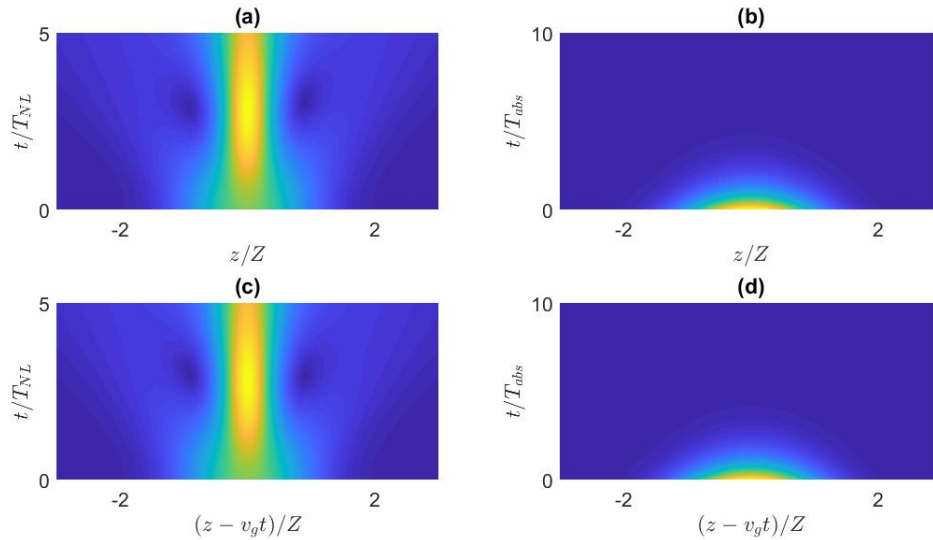
Fig. 5. Spatio-temporal dynamics of the amplitude of the electric field (color represents the value $|A(z,t)|e^{-\Im(\omega_0)t}$) for a pulse in the non-linear case. Plots (a) and (b) represent pulse propagation in the ZGVP regime in the absence and presence of absorption, respectively. Here, the time scales of the dispersion time and the absorption time are exactly as in Figs 3(a)-(b). The nonlinearity time scale is $T_{NL} = 0.2T_{disp}$ in the absence of absorption (a) and $T_{NL} = 0.1T_{abs}$ when absorption is present (b). Plots (c) and (d) represent pulse propagation in the slow light regime in the absence and presence of absorption, respectively. Here, the dispersion and absorption time scales are exactly as in Figs 3(c)-(d). The nonlinearity time scale is $T_{NL} = 0.2T_{disp}$ in the absence of absorption (c) and $T_{NL} = 0.1T_{abs}$ when absorption is present (d). In both cases, the total time of propagation is T_{NL} and $10T_{abs}$, respectively.

T_{abs} and T_{NL}; in the presence of absorption, we replace $T_{NL} \rightarrow \Re[T_{NL}]$.

Overall, the nonlinearity results in self-focusing in space and spectral broadening (due to self-phase modulation). As a result, the higher-order dispersion terms become relatively more important than in the linear cases studied above.

We first study pulse propagation at the stopped light regime. We assume a negative nonlinearity which counters the effects of the normal dispersion occurring in this regime. The outcome of this simulation in the absence of absorption, where we set $T_{NL} \ll T_{disp}$ (T_{abs} is infinite) is presented in Fig. 5(a), and the corresponding spectrum, is presented in Fig. 6(a). In this case, we set $M = 2$, after we found that $R_{2,3} = 0.008$. The corresponding dynamics in the presence of absorption is shown in Fig. 5(b), and Fig. 6(b). A comparison of Fig. 5(a) to Fig. 3(a) shows that since the nonlinear effect is much stronger than the dispersion (for this choice of parameters), the pulse exhibits rapid self-focusing. In momentum space, when comparing Fig. 6(a) to Fig. 4(a), we observe the corresponding behavior where the spectral width broadens, and refocuses again [53].

When considering absorption, in order to observe a significant nonlinear effect, we choose T_{NL} to be much shorter than T_{abs}, thus, also much shorter than T_{NL} in the absence of absorption. If T_{NL} was longer, the absorption would attenuate the pulse before any nonlinear effect can take place. When comparing Fig. 5(b) and Fig. 3(b), it seems that in terms of duration, the nonlinear process does not change the fate of the pulse, as the pulse is attenuated very quickly in both cases. Nevertheless, in momentum space, when comparing Fig. 6(b) to Fig. 4(b), the spectrum is
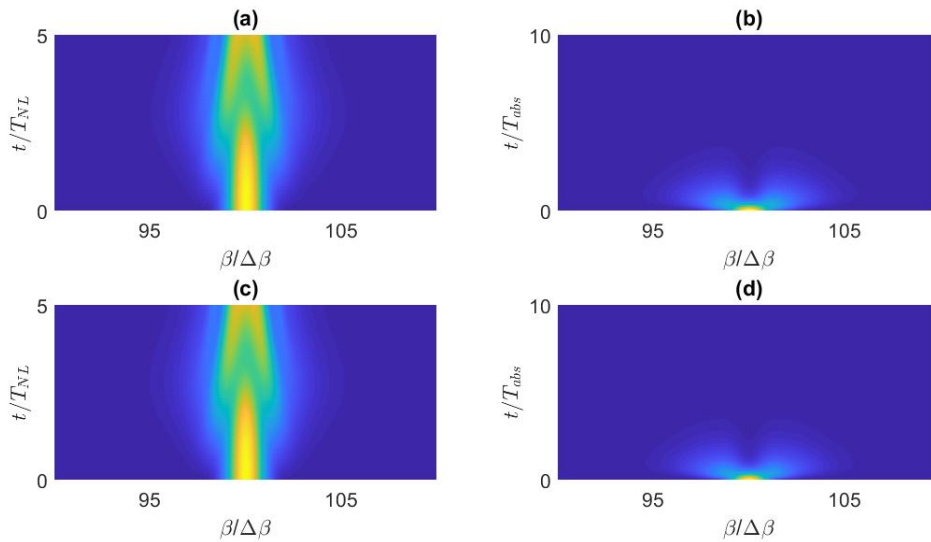
Fig. 6. The spatial spectrum associated with the results shown in Figs. 5(a)-(d).

broadened very quickly due to self-phase modulation.

Next, we study the influence of nonlinearity on pulse propagation in the slow light regime. We choose the same initial pulse as in the slow light regime simulation in Section 4.1. In this regime, the dispersion is anomalous, so we set a positive nonlinear term to counteract the effects of the dispersion. As in the case of the ZGVP, we set T_{NL} such that nonlinear effects are stronger than the dispersion and omit the absorption. Such configuration results in $T_{NL} \ll T_{disp}$ (T_{abs} is infinite). The associated pulse dynamics is shown in Fig. 5(c) and the corresponding spatial spectrum is presented in Fig. 6(c). We note that $M = 2$ is sufficient in this case as $R_{2,3} = 0.001$. Fig. 5(b) and (d) (as well as in Fig. 6(b) and (d)) show the dynamics in the presence of absorption. Overall, as in the linear case, the dynamics in the slow light regime is qualitatively (and even quantitatively) similar to that in the ZGVP. Specifically, T_{NL} in both cases is similar, and also similar to cases where the mode has a "normal" velocity, see discussion in Section 3.3.

5. Discussion

In this paper, we developed an efficient formulation suitable for describing pulse propagation in the slow and stopped light regimes. It enables simple simulations for far more extreme regimes of dispersion than performed before. However, since the linear pulse propagation problem has an analytical solution, the main contribution of this part of our work is in providing a somewhat new point of view on pulse propagation in lossy and nonlinear media, which may in fact be intuitive. Indeed, our formulation shows that the relevant quantity is the propagation time (rather than propagation space) - if measured in time, the rate of phase accumulation is not different than in weakly dispersive systems. Thus, the apparent enhanced phase accumulation (hence, nonlinear effect) and absorption in the slow light regime are a direct result of the longer time spent in the system. This is seen also in the nonlinear simulations (Section 4.2) which do not have an analytic solution.

The rigorous derivation of our main equations is quite subtle, as the lengthy derivation shows, especially, if one has to go beyond second-order dispersion. However, a-posteriory, the final form is quite intuitive as its first- and second-order terms could have even be guessed due to the close

analogy to the standard formulation. This familiar form allows us to enjoy the wide knowledge accumulated from studies of its standard analogue, including, in particular, the availability of analytical solutions for Gaussian pulses, effects associated with the pulse shape, the understanding of the interplay between dispersion and Kerr nonlinearity (see [53]) and the nonlinear dynamics within the (generalized) nonlinear model, in general. The formulation also shows how to handle material absorption in a convenient manner, via complex frequencies. The advantage of their use was not well understood at the time when the more complex momentum-based formulations were derived. For example, the tedious FDTD calculations of the threshold-less lasing configuration [35–38] and wave mixing and free carrier generation effects [15, 82, 83] could now be replaced with simple, even analytic, solutions.

Our formulation can be generalized for a weak longitudinal grating, by adopting the approach in [60–63]. The formulation will also be trivially valid for pulse propagation in (uniform) atomic media [2, 3, 30, 31] and quantum models of stopped light [26–28], which as noted, were so far usually modelled with rather simplistic models in which second-order and higher-order dispersion effects were neglected.

In addition, the importance of the current study is in performing the first step towards the extension to more peculiar/demanding problems of (linear and especially nonlinear) pulse propagation that do not have an analytical solutions like, e.g., photonic crystals with defects or degenerate band edges [51], or plasmonic photonic crystals, which are absorptive and highly dispersive [84, 85]. Further, our formulation is a starting point for the study of more complex configurations involving the coupled dynamics of various modes. In this case, one would be able to derive coupled mode analysis in the slow and stopped light regime, so far treated only with the full Maxwell equations via FDTD algorithms [6, 11–13, 37, 38, 57–59]. Finally, our results could also be used in many additional contexts such as acoustic waves [43–45], spin waves [46], water waves [47], quantum/matter waves [48, 49] etc..

Appendix A   Finding the energy density of the carrier mode

We start with Eq. (29) for $\beta = \beta_0$, namely,

$$
\begin{aligned}
U(\beta_0, \beta_0) &= s\left(\epsilon(\vec{r}_\perp, \omega_0) + \frac{1}{2}\omega_0 \frac{d\epsilon(\vec{r}_\perp, \omega)}{d\omega}\bigg|_{\omega_0}\right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 d\vec{r}_\perp, \\
&= \frac{1}{2}\int\left(\epsilon(\vec{r}_\perp, \omega_0) + \frac{d\left[\omega\epsilon(\vec{r}_\perp, \omega)\right]}{d\omega}\bigg|_{\omega_0}\right) |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 d\vec{r}_\perp.
\end{aligned}
\tag{48}
$$

We now show that although the constituent materials are dispersive, the relation that stands for the equality of the electric and magnetic energies in dispersionless media, namely, $\int \epsilon(\vec{r}_\perp, \omega_0)|\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 = \int \mu|\tilde{\vec{h}}(\vec{r}_\perp, \beta_0)|^2$, where $\tilde{\vec{h}}(\vec{r}_\perp, \beta_0)$ is the transverse magnetic field of mode β_0, is still valid in the presence of dispersion.

The term $\int \epsilon(\vec{r}_\perp, \omega_0)|\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 dxdy$ in Eq. (48) can be written as

$$
\begin{aligned}
\int \epsilon(\vec{r}_\perp, \omega_0)|\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 dxdy &= \int \epsilon(\vec{r}_\perp, \omega_0)(|\tilde{e}_x|^2 + |\tilde{e}_y|^2 + |\tilde{e}_z|^2)dxdy \\
&= \int \epsilon(\vec{r}_\perp, \omega_0)\left(\frac{1}{\omega_0\epsilon(\vec{r}_\perp, \omega_0)}\left(-\frac{\partial h_z}{\partial y} + i\beta_0 h_y\right)\tilde{e}_x^* + \frac{1}{\omega_0\epsilon(\vec{r}_\perp, \omega_0)}\left(\frac{\partial h_z}{\partial x} - i\beta_0 h_x\right)\tilde{e}_y^* \right. \\
&\quad \left. + \frac{1}{\omega_0\epsilon(\vec{r}_\perp, \omega_0)}\left(\frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y}\right)\tilde{e}_z^*\right)dxdy \\
&= \frac{1}{\omega_0}\int\left(-\frac{\partial h_z}{\partial y}\tilde{e}_x^* + i\beta_0 h_y \tilde{e}_x^* + \frac{\partial h_z}{\partial x}\tilde{e}_y^* - i\beta_0 h_x \tilde{e}_y^* + \frac{\partial h_y}{\partial x}\tilde{e}_z^* - \frac{\partial h_x}{\partial y}\tilde{e}_z^*\right)dxdy,
\end{aligned}
\tag{49}
$$

where we have used Maxwell-Ampéré law,

$$\tilde{e}_x = \frac{i}{\omega_0 \epsilon(\vec{r}_\perp, \omega_0)} \left(i\beta_0 h_y - \frac{\partial h_z}{\partial y} \right), \tag{50a}$$

$$\tilde{e}_y = \frac{i}{\omega_0 \epsilon(\vec{r}_\perp, \omega_0)} \left(\frac{\partial h_x}{\partial y} - i\beta_0 h_y \right), \tag{50b}$$

$$\tilde{e}_z = \frac{i}{\omega_0 \epsilon(\vec{r}_\perp, \omega_0)} \left(\frac{\partial h_x}{\partial y} - \frac{\partial h_y}{\partial x} \right). \tag{50c}$$

Integrating by parts gives

$$\int \epsilon(\vec{r}_\perp, \omega_0) |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 dxdy =$$
$$\frac{1}{\omega_0} \int \left(h_z \frac{\partial \tilde{e}_x^*}{\partial y} + i\beta_0 h_y \tilde{e}_x^* - h_z \frac{\partial \tilde{e}_y^*}{\partial x} + i\beta_0 h_y \tilde{e}_x^* - h_y \frac{\partial \tilde{e}_z^*}{\partial x} + h_x \frac{\partial \tilde{e}_z^*}{\partial y} \right) dxdy, \tag{51}$$

where we used

$$\int h_{r_i} \tilde{e}_{r_j} dr_k \bigg|_{r_i=-\infty}^{r_i=\infty} = 0, \tag{52}$$

where i, j, k represent any spatial coordinate. We now substitute Faraday-Henri Law

$$\frac{\partial \tilde{e}_z}{\partial x} = i\beta_0 \tilde{e}_x + i\omega_0 \mu_0 h_y, \tag{53a}$$

$$\frac{\partial \tilde{e}_z}{\partial x} = i\beta_0 \tilde{e}_x + i\omega_0 \mu_0 h_y, \tag{53b}$$

in Eq. (51) and use Eq. (52) again. Eq. (51) becomes

$$\int \epsilon(\vec{r}_\perp, \omega_0) |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 dxdy =$$
$$\frac{1}{\omega_0} \int \left[h_z \left(\frac{\partial \tilde{e}_x^*}{\partial y} - \frac{\partial \tilde{e}_y^*}{\partial x} \right) + i\omega_0 \mu_0 \left(|h_x|^2 + |h_y|^2 \right) \right] dxdy. \tag{54}$$

In the final stage, we substitute

$$\frac{\partial \tilde{e}_x^*}{\partial y} - \frac{\partial \tilde{e}_y^*}{\partial x} = i\omega_0 \mu_0 h_z^*, \tag{55}$$

and get

$$\int \epsilon(\vec{r}_\perp, \omega_0) |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 dxdy = \mu_0 \int \left(|h_x|^2 + |h_y|^2 + |h_z|^2 \right) dxdy$$
$$= \mu_0 \int |\vec{h}(\vec{r}_\perp, \beta_0)|^2 dxdy. \tag{56}$$

This proves our claim. Substituting Eq. (56) in Eq. (48) gives

$$U(\beta_0, \beta_0) = \frac{1}{2} \int \left(\frac{d(\epsilon(\vec{r}_\perp, \omega_0)\omega)}{d\omega} \bigg|_{\omega_0} |\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|^2 + \mu_0 |\vec{h}(\vec{r}_\perp, \beta_0)|^2 \right) dxdy. \tag{57}$$

Since $|\tilde{\vec{e}}(\vec{r}_\perp, \beta_0)|$ accounts for the total amplitude of the electric field (i.e., $max(A) = 1$), Eq. (57) is the expression for energy of the carrier mode in dispersive, loss-free, non-magnetic media.

Appendix B   Estimate of $\tilde{D}_{n,m}$ and $\tilde{D}_m$

In this Appendix, we estimate the magnitude of the coefficients $\tilde{D}_{n,m}$ (32) and $\tilde{D}_m$ (31).

Starting with index n, when n increases, according to Eq. (32), the relative magnitude of consecutive terms in n, namely, $\tilde{D}_{n+1,m}/\tilde{D}_{n,m}$, can be estimated by

$$R_n \sim \frac{\int \frac{d\epsilon(\vec{r}_\perp,\omega_0)(\vec{r}_\perp,\omega)}{d\omega}\big|_{\omega_0} d\vec{r}_\perp}{\int \epsilon(\vec{r}_\perp,\omega)|_{\omega_0} d\vec{r}_\perp}. \tag{58}$$

For a Drude metal, when taking into account Eq. (45) and ignoring the details of the waveguide width, we can estimate, $R_n \sim \frac{1}{\omega_0}$. For a dispersive dielectric, R_n will be smaller, hence, the following analysis can be regarded as an estimate for an upper limit. It follows that

$$\tilde{D}_{n,m} \sim \frac{1}{\omega_0^{n-1}} \tilde{D}_{1,m}. \tag{59}$$

Continuing with index m, according to Eq. (32), one can evaluate

$$R_m \sim \frac{\int \frac{d|\tilde{\vec{e}}(\vec{r}_\perp,\beta)|^2}{d\beta}\big|_{\beta_0} d\vec{r}_\perp}{\int |\tilde{\vec{e}}(\vec{r}_\perp,\beta)|^2 d\vec{r}_\perp}, \tag{60}$$

where R_m is the ratio of order of magnitude of two adjacent terms of $\tilde{D}_{n,m}$ in m, specifically, $\tilde{D}_{n,1}/\tilde{D}_{n,0}$. The magnitude of R_m may vary when considering different waveguides, however, in this work it will be reasonable to consider a plasmonic waveguide where

$$|\tilde{\vec{e}}(\vec{r}_\perp,\beta)| \sim e^{-|\vec{k}_\perp \cdot \vec{r}_\perp|}, \tag{61a}$$

$$|\vec{k}_\perp| = \sqrt{\beta^2 - \epsilon(\vec{r}_\perp,\beta)(\omega/c)^2}. \tag{61b}$$

Substituting Eqs. (61) in Eq. (60) results in

$$R_m \sim \frac{\int \vec{r}_\perp |\tilde{\vec{e}}(\vec{r}_\perp,\beta)|^2 d\vec{r}_\perp}{\int |\tilde{\vec{e}}(\vec{r}_\perp,\beta)|^2 d\vec{r}_\perp}. \tag{62}$$

Notice that Eq. (62) is not a direct result of the substitution but rather a term which maintains the magnitude in Eq. (60). It is deduced from Eq. (62) that $R_m \sim \langle \vec{r}_\perp(\beta) \rangle$, where $\langle \vec{r}_\perp(\beta) \rangle$ is the characteristic width of the waveguide mode $\tilde{\vec{e}}(\vec{r}_\perp,\beta)$.

In cases where $\beta_0 \sim |\vec{k}_\perp|$ at any of the constituents, $\langle \vec{r}_\perp(\beta) \rangle$ scale as $1/\beta_0$ as β_0 is the dominant value in Eq. (61b). If we consider small values of β_0 with respect to $|\vec{k}_\perp|$, due to the nature of the metal cladding in a MDM waveguide, $\langle \vec{r}_\perp(\beta) \rangle$ cannot exceed a range which is much larger than the core, thus $R_m \sim \max(\langle \vec{r}_\perp(\beta) \rangle, 1/\beta_0)$.

Considering the above, $\tilde{D}_{n,1} \sim \frac{\tilde{D}_{n,0}}{\beta_0}$, where similar analysis in the high-order terms gives,

$$\tilde{D}_{n,m} \sim \frac{\tilde{D}_{n,0}}{\beta_0^m}. \tag{63}$$

Since $\tilde{D}_{1,0} \equiv 1$, according to Eqs. (59) and (63), the magnitude of $\tilde{D}_{n,m}$ is $\frac{1}{\omega_0^{n-1}\beta_0^m}$. For small values of β_0 the scale of $\tilde{D}_{n,m}$ is even smaller as $\beta_0 < |\vec{k}_\perp|$.

The structure of $\tilde{D}_m$ is similar of $\tilde{D}_{n,m}$. Considering the same arguments as above, it can be shown that the magnitude of $\tilde{D}_m$ is ω_0/β_0^m, where it can be smaller with large β_0.

Funding

Acknowledgements