Optics Communications 550 (2024) 129841

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

NOTE: Manually corrected from the on-line text by the author 2023-Oct-04

# Solitons in birefringent optical fibers and polarization mode dispersion

Curtis R. Menyuk

*Computer Science and Electrical Engineering Department, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA*

A B S T R A C T

The early work on solitons in optical fibers was all done using single-mode fibers, and these fibers remain important for many applications. The modes in single-mode fibers are for all practical purposes weakly confined plane waves and have two polarizations. The basic equation that describes light evolution in these fibers is the coupled nonlinear Schrödinger equation. The soliton robustness hypothesis and its origins, which were the original motivation for studying these equations, is first described. Limits on robustness due to birefringent walkoff and the impact of random birefringence variations in stabilizing solitons is then described. Questions and controversies that arose shortly after the publication of these equations are addressed. These include their relationship to the Maker and Terhune coefficients, the requirements for the validity of the nonlinear Schrödinger equation and the impact of polarization mode dispersion, and the conditions under which the birefringence can be considered linear. It is a consequence of the fluctuation-dissipation theorem that the birefringence must be linear in any homogeneous, low-loss optical medium with a local dielectric response. Any ellipticity is associated with non-locality, as would occur for example in a twisted optical fiber. This important result, which is not well-known in the optics community, is reviewed.

## 1. Introduction

In this year when we celebrate the 50-th anniversary of the optical fiber soliton [1,2], it is appropriate to look retrospectively at the original work on solitons in birefringent optical fibers [3–5]. I will focus on work that was carried out—first by myself and then by others—starting in the mid-1980s and stretching into the early 2000s. It could be argued that work on optical solitons stretches even farther back than the work by Hasegawa and Tappert [1,2] and the first experimental observation by Mollenauer et al. [6] of solitons in optical fibers. In 1964, Chiao et al. [7] used the nonlinear Schrödinger equation (NLSE) to describe the propagation of optical beams in a medium with a Kerr nonlinearity. This work presented the first observation of what are today called spatial solitons, as opposed to the temporal solitons that exist in optical fibers. Additionally, a theoretical study of these spatial solitons by Kelley [8] was later cited by Zakharov and Shabat [9] in their own seminal work in which they demonstrated that the inverse scattering method that had earlier been applied by Gardner et al. [10] to the Korteweg-de Vries equation could also be applied to the NLSE. Nonetheless, the recognition by Hasegawa and Tappert that the NLSE applied to optical fibers [1,2], along with their experimental observation by Mollenauer et al. [6], was the real trigger that led to an outpouring of work on optical solitons that ensued in the decades following and has never ceased.

The 1970s and 1980s were an exciting time in the nonlinear physics community and in the plasma physics community in which Akira Hasegawa and I were both working at the time. On the one hand, it was discovered that the inverse scattering method could be applied to a large group of equations [11], referred to as integrable, and that this transformation corresponds to an action-angle transformation since all of these systems are Hamiltonian. On the other hand, computers that were sufficiently powerful to simulate the evolution in non-integrable systems became available. While it had been known since Poincaré's work at the end of the 19-th century that the evolution in non-integrable systems is inherently chaotic [12], work in the 1950s and 1960s had led to the Kolgomorov-Arnold-Moser (KAM) theory that implied that regular trajectories (referred to as KAM trajectories) would exist in non-integrable systems and could impede or even block the chaotic motion [13]. These concepts had important applications to ionospheric physics and plasma containment devices, and the computer made possible dramatic visualizations of the chaotic motion, as well as the regular trajectories and their impact on the overall system evolution [14]. The KAM theorem and the impact of the regular trajectories was described in simple terms in a highly influential paper by Chirikov [15] that I first read as a graduate student. My own PhD dissertation was on the subject of Langmuir waves propagating obliquely to a magnetic field and the impact of the KAM trajectories in impeding the flow of electrons [16].

In 1981, I began to work as a post-doctoral research associate at the University of Maryland Baltimore County. My post-doctoral mentors, Hsing-Hen Chen and Yee-Chun Lee, who were interested in studying soliton integrability, first introduced me to solitons. In 1983, I heard a talk at a Physics colloquium by Hans Wilhelmsson in which the work by Hasegawa and colleagues and the work by Mollenauer and colleagues, as well as their potential applications to optical communications, was briefly described. I was instantly hooked! I started reading everything

*E-mail address:* menyuk@umbc.edu.

https://doi.org/10.1016/j.optcom.2023.129841
Received 27 June 2023; Accepted 17 August 2023
Available online 22 September 2023

that I could on the subject of optical fibers and the experiments that Linn Mollenauer, Roger Stolen, and their colleagues at Bell Labs had carried out. I was particularly influenced by papers by Stolen and colleagues [17,18] in which they demonstrated that nonlinear polarization rotation occurs in optical fibers because the cross-polarized Kerr effect in optical fibers is 2/3 as strong as the self-polarized effect. I hypothesized that something like KAM trajectories would exist for physically-realistic Hamiltonian systems that are close enough in some sense to the NLSE, and that solitons should continue to exist in these systems, i.e., solitons would be robust in the presence of Hamiltonian deformations. My plan was to test this idea for practically important optical fiber systems. I wrote down all the Hamiltonian deformations that I could think of and then discussed the list with Roger Stolen. With his help, I narrowed the list to include higher-order dispersion and birefringence as the first two items. Higher-order dispersion eventually became the subject of Ping-kong Alexander Wai's PhD dissertation [19]. H. H. Chen, his PhD advisor, had asked me to suggest a dissertation problem to Alex. The work on birefringence is the subject of this paper.

In the remainder of this paper, I will begin by reviewing the early work on birefringence and continue with a review of work on randomly varying birefringence and polarization mode dispersion (PMD). I will close by describing a little-known result that any homogeneous medium with a local dielectric permittivity can only be linearly birefringent. Any ellipticity requires non-locality. Single-mode optical fibers in which the core-cladding index of refraction is less than 1% only support the $HE_{11}$ mode, which is to all intents and purposes a weakly confined plane wave and, like a plane wave, it has two polarizations. Thus, it is to be expected that the birefringence of single-mode optical fibers will be linear in the absence of twisting or other effects that induce non-locality.

## 2. Solitons in birefringent optical fibers

The starting point for my studies of solitons in birefringent optical fibers was to combine the known dispersion in birefringent optical fibers [20] with the nonlinear response of the fibers, which had earlier been published by Stolen and colleagues [17,18] to yield

$$\begin{aligned} i\frac{\partial u}{\partial z}+i\delta\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial^2 u}{\partial t^2}+\left(|u|^2+\frac{2}{3}|v|^2\right)u=0,\\ i\frac{\partial v}{\partial z}-i\delta\frac{\partial v}{\partial t}+\frac{1}{2}\frac{\partial^2 v}{\partial t^2}+\left(\frac{2}{3}|u|^2+|v|^2\right)v=0, \end{aligned} \tag{1}$$

where $u$ and $v$ are the complex envelopes of the two polarizations, $z$ and $t$ are distance along the fiber and retarded time, and $\delta$ parameterizes the birefringence walkoff. All quantities are normalized, and when writing Eq. (1), I was careful to use the same normalizations that Mollenauer and colleagues were using in order to make the closest possible contact with their prior results. When I wrote down this equation, it was immediately apparent to me that something was missing. Circular symmetry of the transverse fiber profile implies that in the limit $\delta\to 0$, the equation becomes invariant under a transformation $u'=u\cos\theta+v\sin\theta$, $v'=-u\sin\theta+v\cos\theta$, where $\theta$ is an arbitrary angle. I called Roger Stolen for help, who suggested that I expand the vector Kerr nonlinearity, which is proportional to $(\mathbf{E}\cdot\mathbf{E})\mathbf{E}$, where $\mathbf{E}$ is the vector electric field. Expanding the vector electric field yielded the missing term, and Eq. (1) becomes Eq. (2) [3],

where $\mp R\delta z$ is the phase slip between the two polarizations due to the birefringence. For the 5-ps pulses that Mollenauer et al. [6] studied in their original work and a fiber birefringence $\Delta n/n$ between $10^{-7}$ and $10^{-6}$, which was typical [20], the parameter $R\delta$ was on the order of $10^4$–$10^5$, while the parameter $\delta$ was on the order of 0.3–3.0. I concluded that the final terms, shown in red, were rapidly varying and could be dropped in numerical studies. This term does become important when pulses are short and can lead to an instability that was first observed by Trillo et al. [21]; however, this instability was not relevant for the limit that I was trying to model. In subsequent numerical work [4,5], I found that solitons were indeed robust if $\delta$ is not too large. If the birefringent walkoff length is smaller than the nonlinear scale length, then the two polarizations self-trap and a single elliptically-polarized soliton is created. In the opposite limit, the two polarizations split apart and two pulses walk off from one another. I note parenthetically that the (unsymmetrized) Hamiltonian $\mathscr{H}$ for Eq. (1) can be written

$$\begin{aligned} \mathscr{H}=\int dt\Bigg[&-\delta\left(u^*\frac{\partial u}{\partial t}-v^*\frac{\partial v}{\partial t}\right)\\ &+\frac{i}{2}\left(u^*\frac{\partial^2 u}{\partial t^2}+v^*\frac{\partial^2 v}{\partial t^2}+|u|^4+|v|^4+\frac{4}{3}|u|^2|v|^2\right)\Bigg], \end{aligned} \tag{3}$$

where $[u, u^*]$ and $[v, v^*]$ are coordinate and momentum pairs.

I submitted the second of the two numerical papers [5] to a special issue that Roger Stolen edited. The reviewers complained that the paper was not sufficiently different from the earlier paper [4] to merit publication. I pointed out that the paper contained details of the numerical algorithms that I used, which I claimed would be useful to others. Roger agreed and, after I added additional information that he suggested, he accepted the paper for publication. In fact, I received requests for a decade following publication for more details on the algorithms and for the original FORTRAN code that I wrote. However, the difficulty in publishing innovative and useful algorithms persists in the field of optics and photonics.

When I described my results to Linn Mollenauer, I noted that the 5-ps pulses that he had used were in the low-$\delta$ limit and perhaps the birefringent walkoff would not matter since both polarizations would be trapped. Linn explained to me that he was currently doing work with 100-ps pulses so that this effect did potentially matter. He thought for a while and then suggested that perhaps the randomly varying birefringence that had already been studied by Poole and Wagner [22] would stabilize the solitons. Using a version of my numerical code and implementing a randomly varying birefringence by randomly shifting the signs of $\delta$ in Eq. (1), he and his co-authors demonstrated the validity of this suggestion [23]. Linn publicized these equations at Bell Labs and invited me to give a talk in which I described my robustness hypothesis, which I only published years later at the urging of Ray Hawkins [24].

Mohammed Islam was at my talk and suggested that the transition between the trapped and untrapped polarization states could be used to make an optical switch. He first focused on demonstrating the validity of Eq. (1) [25], after which we collaborated on a study of soliton dragging and trapping gates [26]. It became important to understand the impact of the Raman effect on these gates. That led to collaborative work with Jim Gordon in which we modified Eq. (1) to include the Raman effect. I had difficulty publishing this work at first due to overlap with earlier work by Hellwarth [27]. Jim advised me on how to rewrite our work to distinguish it from the work by Hellwarth, and it was published [28].

While it was gratifying that the validity of Eq. (1) was experimentally established [25[ and that randomly varying birefringence could explain the observed soliton stability [23], it was apparent by 1991 that a better

$$\begin{aligned} i\frac{\partial u}{\partial z}+i\delta\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial^2 u}{\partial t^2}+\left(|u|^2+\frac{2}{3}|v|^2\right)u+\frac{1}{3}v^2u^*\exp(-iR\delta z)=0,\\ i\frac{\partial v}{\partial z}-i\delta\frac{\partial v}{\partial t}+\frac{1}{2}\frac{\partial^2 v}{\partial t^2}+\left(\frac{2}{3}|u|^2+|v|^2\right)v+\frac{1}{3}v^2u^*\exp(iR\delta z)=0, \end{aligned} \tag{2}$$

model was needed that unified a physics-based model of the randomly varying birefringence with the Kerr effect. Aside from the lack of physicality, I was intrigued by the success of the NLSE in modeling optical fiber transmission. Indeed, until the development of polarization-division multiplexed systems in the 2000s forced the use of Eq. (1) and its improvements, the NLSE was consistently used. I was convinced that it should be possible to derive the NLSE as a limit of the coupled NLSE with randomly birefringence and thereby determine the limits of its validity. I had joined the University of Maryland Baltimore County (UMBC) in Fall 1986 as the first-ever professor of Electrical Engineering in a newly-created College of Engineering (later the College of Engineering and Information Technology). Shortly after 1990, Alex Wai joined me at UMBC as a research professor, and we began to tackle this problem. Based on the earlier work of Foschini and Poole [29], we derived an analytical expression for the walkoff length and verified that as long as the walkoff length is long compared to the nonlinear scale length, solitons will still be robust [30]. As a practical matter, the lengths over which the randomly varying birefringence remains correlated is very small—typically on the order of meters—when compared to the nonlinear scale length in long-distance telecommunications systems, which is typically on the order of hundreds of kilometers. Modeling the random variations on the length scale of the random variations is computationally inefficient. Working with Dieter Marcuse, who joined Alex and I at UMBC (working remotedly) after retiring from Bell Labs and using two different, physically realistic models of the randomly varying birefringence, we investigated this issue. We demonstrated that artificially increasing the birefringence by a factor $(\Delta z/z_{\mathrm{corr}})^{1/2}$, where $\Delta z$ is the step size and $z_{\mathrm{corr}}$ is the birefringence correlation length, it is possible to take computational steps that are long compared to $z_{\mathrm{corr}}$ and the birefringent beat length, although $\Delta z$ must remain short compared to the dispersive and nonlinear scale lengths [31].

The impact of the coupled NLSE, Eq. (1) has been large. Aside from direct references to Refs. [3–5], these equations appeared in the first edition of *Nonlinear Fiber Optics* [32] by Govind Agrawal, where they have been referenced many times more. The references have become so common that I now often see these equations with no citation at all. Beyond their application to solitons, they are the base equations that govern all nonlinear Kerr interactions in single-mode fibers, and have been extended to include higher-order dispersion, the Raman and Brillouin effects, Rayleigh scattering and noise. They play an important role in fiber lasers that are modelocked using nonlinear polarization rotation or for that matter any fiber laser that does not use polarization-preserving fibers [33].

The field of nonlinear fiber optics has progressed far beyond far beyond single solitons and single-mode fibers. Even before my own work was published, Hasegawa [34] and Crosignani and Di Porto [35] had proposed the use of a multi-component NLSE to describe light propagation in multi-mode fibers, although I was not aware of this work until after [3] was published. At the time, inter-modal dispersion was too large to produce solitons. However, that is no longer the case, and it has recently become possible to create modelocked pulses in which many transverse modes as well as the usual longitudinal modes are locked together [36,37]. Orbital angular modes in structured fibers with large index differences play an important role in modern nonlinear fiber optics [38]. Complex soliton structures, including molecules [39], crystals [40], and supra-molecules [41] have all been studied in fiber lasers. My own contribution to these exciting developments in nonlinear fiber optics has been slight. For the past 15 years, most of my effort has gone into frequency combs, first in lasers and more recently in microresonators. Solitons are important in these applications, and somewhat like Russell chasing a soliton along a canal in Scotland [42], I have spent my career chasing solitons into new domains and applications.

## 3. Controversies, questions, and answers

The coupled NLSE was not without controversy when it was first published. The first time that I presented these equations at a scientific meeting, I was asked what happened to the Maker and Terhune $A$ and $B$ coefficients [43]. At the time, I didn't even know what they were! A more serious question was how polarization mode dispersion (PMD) would affect these equations. A related question that I had was: Given that the coupled NLSE is a more fundamental equation than the NLSE in single-mode optical fibers, why is the NLSE as successful as it is, and what are the limitations? A final question that I was asked on several occasions is: Why did I assume that the birefringence is linear so that the cross-phase-modulation term in Eq. (1) has a factor 2/3 in front? Before answering these questions, we first recall that the fundamental mode in a step-index fiber is the $HE_{11}$ mode, which is the only mode that exists when the index difference becomes small enough for the fiber to be single-mode. In this limit, the electric field is effectively a weakly confined plane wave [44,45]. So, it is unsurprising that it has two polarizations. We can thus study the nonlinear propagation of light in single-mode fibers using the plane-wave approximation once we appropriately average over the transmission mode amplitude to obtain the effective index and determine the strength of the nonlinearity. This observation can be made rigorous [45,46].

*Question 1: What happened to the Maker and Terhune coefficients?*

To answer this question, we must first examine more closely the nonlinear response of the optical fiber, which we will do in the plane wave approximation. We then find that the polarization density of the medium **P** is given by [45,46].

$$\mathbf{P}(z,t) = \int_{-\infty}^{t} dt_1\, \mathrm{X}_{\mathrm{L}}(t-t_1)\cdot\mathbf{E}(z,t_1) + \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t} dt_3\, \chi_{\mathrm{NL}}(t-t_1,t-t_2,t-t_3)[\mathbf{E}(z,t_1)\cdot\mathbf{E}(z,t_2)]\mathbf{E}(z,t_3), \tag{4}$$

where **E** is the electric field, the susceptibility $\mathrm{X}_{\mathrm{L}}$ is a second-order tensor, since the optical fiber is birefringent, and $\chi_{\mathrm{NL}}$ is a scalar. We make three basic assumptions in writing Eq. (1) in addition to the plane wave approximation. The first is that the change in the index of refraction due to both the birefringence and nonlinearity is small compared to one. We can parameterize this smallness by comparing the length scales for the birefringence and the nonlinearity to the optical wavelength. The length scale for the birefringence, given by the beat length between the two polarizations, is on the order of a meter, while the length scale for the nonlinearity, given by the length over which the nonlinearity rotates the phase by $2\pi$, can be many kilometers. By comparison, the wavelength of light is about a micrometer. One consequence is that the effect of birefringence on the nonlinear response is negligible. Second, we are assuming that the carrier frequency $\omega_0$ is large compared to the bandwidth of the signal so that the slowly-varying envelope approximation applies and the lowest-order in-band nonlinear contributions are cubic. Third, we are assuming that the response of the polarization density to the electric field is local in space and homogeneous in time. We may then write [45,46].

$$\begin{aligned}\mathbf{E}(z,t) &= \mathbf{E}^{+}(z,t)\exp(ik_0 z - i\omega_0 t) + \mathbf{E}^{-}(z,t)\exp(-ik_0 z + i\omega_0 t),\\ \mathbf{P}(z,t) &= \mathbf{P}^{+}(z,t)\exp(ik_0 z - i\omega_0 t) + \mathbf{P}^{-}(z,t)\exp(-ik_0 z + i\omega_0 t),\end{aligned} \tag{5}$$

where $\mathbf{E}^{-} = \mathbf{E}^{+*}$, $\mathbf{P}^{-} = \mathbf{P}^{+*}$ and all four quantities are slowly varying. We have $k_0 = k(\omega_0)$ is the wavenumber at the central frequency. Keeping only the in-band components, we have

$$\begin{aligned}\mathbf{P}^{+}(z,t) &= \int_{-\infty}^{t} dt_1\, \mathbf{X}_{\mathrm{L}}(t-t_1)\cdot\mathbf{E}^{+}(t_1)\exp[i\omega_0(t-t_1)]\\ &+ \int_{\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t} dt_3\, \chi_{\mathrm{NL}}(t-t_1, t-t_2, t-t_3)\\ &\{2[\mathbf{E}^{+}(z,t-t_1)\cdot\mathbf{E}^{-}(z,t-t_2)]\mathbf{E}^{+}(z,t-t_3)\\ &+[\mathbf{E}^{+}(z,t-t_1)\cdot\mathbf{E}^{+}(z,t-t_2)]\mathbf{E}^{-}(z,t-t_3)\,\}\\ &\exp[i\omega_0(t-t_1+t_2-t_4)].\end{aligned} \tag{6}$$

Writing $\mathbf{P}^{+} = \mathbf{P}^{+}_{\mathrm{L}} + \mathbf{P}^{+}_{\mathrm{NL}}$, which correspond to the linear and nonlinear contributions to $\mathbf{P}^{+}$ and taking advantage of the slowly varying envelope approximation, we find that the nonlinear response becomes

$$\begin{aligned}P^{+}_{\mathrm{NL}}(z,t) &= 2\widetilde{\chi}_{\mathrm{NL}}(\omega_0,-\omega_0,\omega_0)[\mathbf{E}^{+}(z,t)\cdot\mathbf{E}^{-}(z,t)]\mathbf{E}^{+}(z,t)\\ &+\widetilde{\chi}_{\mathrm{NL}}(\omega_0,\omega_0,-\omega_0)[\mathbf{E}^{+}(z,t)\cdot\mathbf{E}^{+}(z,t)]\mathbf{E}^{-}(z,t),\end{aligned} \tag{7}$$

where $\widetilde{\chi}_{\mathrm{NL}}$ is the three-dimensional Fourier transform of $\chi_{\mathrm{NL}}$. We see that the nonlinear response has two components that are distinct when the response time of the medium is long compared to the period of the optical field, albeit short compared to the time scale on which the field amplitudes change, which is necessary for the slowly-varying envelope approximation to be valid. That is the limit in which Maker et al. [43] worked, and the two components of the nonlinear susceptibility are proportional to the $A$ and $B$ coefficients. However, the response time of the Kerr effect in optical fibers is short compared to the period of the optical field, and in this limit, we have $\widetilde{\chi}_{\mathrm{NL}}(\omega_0,-\omega_0,\omega_0) = \widetilde{\chi}_{\mathrm{NL}}(\omega_0,\omega_0,-\omega_0) = \widetilde{\chi}_{\mathrm{NL}}(0,0,0)$, and the two Maker and Terhune coefficients collapse into one.

*Question 2: When is the NLSE valid?*

The validity of the NLSE, which was used almost universally to model single-mode optical fibers in the first two decades after the work of Hasegawa and Tappert [1,2] and Mollenauer et al. [6] continually intrigued me, as I reported in the prior section. Mollenauer et al. [23] showed that the NLSE continues to hold when the slow and fast axes of the fiber are randomly interchanged. I wanted to show that this conclusion continues to hold for more physically realistic models and to find its limits of validity. Working with Alex Wai and Dieter Marcuse, I tackled this question, and we resolved it [30,31,45]. The first step is to re-write Eq. (2) in vector form and to slightly extend it to yield [45].

$$\begin{aligned}& i\frac{\partial\mathbf{U}}{\partial z} - ig(z)\mathbf{U} + [\cos\theta(z)\boldsymbol{\sigma}_3 + \sin\theta(z)\boldsymbol{\sigma}_1]\left[\Delta\beta(z)\mathbf{U} + i\Delta\beta'(z)\frac{\partial\mathbf{U}}{\partial t}\right]\\ & -\frac{1}{2}\beta''\frac{\partial^2\mathbf{U}}{\partial t^2} - \frac{1}{6}i\beta'''\frac{\partial^3\mathbf{U}}{\partial t^3} + \gamma\left[\left|\mathbf{U}\right|^2\mathbf{U} - \frac{1}{3}(\mathbf{U}^{\dagger}\boldsymbol{\sigma}_2\mathbf{U})\boldsymbol{\sigma}_2\mathbf{U}\right] = 0,\end{aligned} \tag{8}$$

where

$$\mathbf{U} = \begin{bmatrix} u(z,t)\\ v(z,t)\end{bmatrix}, \qquad \mathbf{U}^{\dagger} = [u^{*}(z,t)\;\; v^{*}(z,t)], \tag{9}$$

the $\boldsymbol{\sigma}_j$ are the standard Pauli matrices,

$$\boldsymbol{\sigma}_1 = \begin{bmatrix} 0 & 1\\ 1 & 0\end{bmatrix}, \qquad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & -i\\ i & 0\end{bmatrix}, \qquad \boldsymbol{\sigma}_3 = \begin{bmatrix} 1 & 0\\ 0 & -1\end{bmatrix}, \tag{10}$$

$\Delta\beta$ and $\Delta\beta'$ are the difference in the wavenumber and its frequency derivative evaluated at the central frequency $\omega_0$, $\theta(z)$ is the orientation angle of the birefringence, $\beta''$ and $\beta'''$ are the second- and third-order contributions to chromatic dispersion, $g(z)$ is the position-dependent gain or loss, and $\gamma$ is the Kerr coefficient. Next, we transform Eq. (7) in a way that would diagonalize its evolution were it not for the $z$-variation of $\theta(z)$. We may do that by letting $\mathbf{V} = \mathsf{R}^{-1}\mathbf{U}$, where $\mathsf{R} = \cos(\theta/2)\mathsf{I} + i\sin(\theta/2)\boldsymbol{\sigma}_2$. We then obtain

$$\begin{aligned}& i\frac{\partial\mathbf{V}}{\partial z} - ig\mathbf{V} + \boldsymbol{\sigma}_3\left[\Delta\beta\mathbf{V} + i\Delta\beta'\frac{\partial\mathbf{V}}{\partial t}\right] + \frac{1}{2}\theta_z\boldsymbol{\sigma}_2\mathbf{V}\\ & -\frac{1}{2}\beta''\frac{\partial^2\mathbf{V}}{\partial t^2} - \frac{1}{6}i\beta'''\frac{\partial^3\mathbf{V}}{\partial t^3} + \gamma\left[|\mathbf{V}|^2\mathbf{V} - \frac{1}{3}(\mathbf{V}^{\dagger}\boldsymbol{\sigma}_2\mathbf{V})\boldsymbol{\sigma}_2\mathbf{V}\right] = 0,\end{aligned} \tag{11}$$

where $\theta_z = d\theta/dz$. In this step, we have effectively frozen the rapid motion of the state of birefringence on the equator of the Poincaré sphere. However, the term $(\theta_z/2)\boldsymbol{\sigma}_2\mathbf{V}$ leads to off-diagonal coupling, whose random variation leads to the polarization state of the light uniformly filling the Poincaré sphere. It is worth noting that the linear portion of Eq. (2) with the off-diagonal coupling between the two components of the state of polarization is consistent with Poole and Wagner's phenomenological model [22], and we can identify their coupling coefficient with $\theta_z/2$ in a physically realistic model.

The next step is a bit trickier. The goal is to freeze the motion of the carrier frequency on the Poincaré sphere. That will allow us to obtain the residual spreading of the polarization state as a function of frequency and distance. To accomplish this next transformation, we let $\mathbf{W} = \mathsf{T}^{-1}(z)\mathbf{V}$, where $\mathsf{T}$ satisfies the equation

$$i\frac{\partial\mathsf{T}}{\partial z} + [\Delta\beta\boldsymbol{\sigma}_3 + (\theta_z/2)\boldsymbol{\sigma}_2]\mathsf{T} = 0 \tag{12}$$

with the initial condition $\mathsf{T}(z=0) = \mathsf{I}$. After making this transformation, we obtain

$$\begin{aligned}& i\frac{\partial\mathbf{W}}{\partial z} - ig\mathbf{W} - \frac{1}{2}\beta''\frac{\partial^2\mathbf{W}}{\partial t^2} - \frac{1}{6}i\beta'''\frac{\partial^3\mathbf{W}}{\partial t^3} + \frac{8}{9}\gamma|\mathbf{W}|^2\mathbf{W}\\ & = -i\Delta\beta'\overline{\boldsymbol{\sigma}}_3\frac{\partial\mathbf{W}}{\partial t} + \frac{1}{3}\gamma\left[(\mathbf{W}^{\dagger}\overline{\boldsymbol{\sigma}}_2\mathbf{W})\overline{\boldsymbol{\sigma}}_2\mathbf{W} - \frac{1}{3}|\mathbf{W}|^2\mathbf{W}\right] = 0,\end{aligned} \tag{13}$$

where $\overline{\boldsymbol{\sigma}}_j = \mathsf{T}^{-1}\boldsymbol{\sigma}_j\mathsf{T}$.

The beauty of this expression is that all the effects of the rapidly varying birefringence have been isolated on the right-hand side of this equation. We see immediately that there is no coupling between the components of $\mathbf{W}$ when the right-hand side of Eq. (13) is negligible. The first term on the right-hand side corresponds to the usual polarization mode dispersion (PMD) [45]. The second term on the right-hand side is a nonlinear PMD. Marcuse et al. [31] showed that this term is negligible unless the pulse duration is a small fraction of a picosecond for typical pulse powers at the time.

The conditions for the NLSE to hold are now clear. The initial pulse must be in a single polarization state, and the effects of PMD must be negligible so that the polarization state remains the same at every point in the signal as the signal propagates, although this polarization state will typically be rapidly and randomly changing. I will add parenthetically that answering this question allowed my colleagues and I to start exploring the interaction of nonlinearity and PMD before this interaction became a topic of great interest in the telecommunications industry—pointing in my view to the importance of curiousity-driven research.

*Question 3: Is the birefringence really linear?*

My response to this question for many years was, "Of course it is; Roger Stolen showed it." More precisely, Roger Stolen and his colleagues at Bell Labs demonstrated that the cross-coupling term has 2/3 the strength of the self-coupling term [17,18], and that is only possible if the birefringence is linear. Moreover, it seems physically reasonable since elliptical or circular birefringence is associated with helicity in the optical medium, and none is present unless the optical fiber is twisted, as can happen intentionally or when the fiber is wound on a bobbin [47]. I quickly became aware that twisting a fiber would induce an effective ellipticity [45,46,48]. That effect is apparent in Eq. (11), where if $\theta_z$ is constant, then a constant off-diagonal coupling is induced between the components of **V**. This polarization coupling leads to an elliptical birefringence relative to the rotating axes and changes the nonlinear coupling. The ratio changes from 3:2 to 1:2 in the case of circular birefringence and at a special angle of ellipticity of about 35°, the ratio becomes 1:1, and the Manakov equation is obtained [46,48]. However, this ellipticity is with respect to the rotating axes and applies when the nonlinear scale length is long compared to the beat length, as is the case in telecommunications applications, but not when nonlinear polarization rotation is used as a discriminator, as is the case in many passively modelocked fiber lasers [33].

A further complexity is that modern telecommunication fibers are spun as they are drawn to reduce the PMD. In the late 1990s and early 2000s, my research group at UMBC collaborated with the research group of Andrea Galtarossa at the Università di Padova to model PMD. Andrea's group had pioneered polarization optical time-domain reflectometry (P-OTDR) techniques that could be used to study the evolution of the birefringence along the fibers [49]. While P-OTDR is insensitive to the circular component of the birefringence, the results are consistent with models that assume that this component is negligible for straight untwisted optical fibers.

I was then surprised to discover in summer 2009 when perusing *Electrodynamics of Continuous Media* by Landau and Lifshitz [50] that the assumption that the birefringence is linear has a rigorous basis. The birefringence is necessarily linear in any homogenous medium in which the permittivity is local [50], as is assumed in Eq. (15) below. Equation (15) and equivalently Eq. (20) below is the starting point for modeling a wide variety of optical resonators and waveguides, not just optical fibers. This result, which is a consequence of the fluctuation-dissipation theorem, implies that elliptical or circular birefringence can only occur due to non-locality, as can happen when the molecules in the medium are helical and large, leading to optical activity, or when the medium is subject to a helical stress, as when an optical fiber is twisted. Landau and Lifshitz report this result as a special case of a more general result that they obtain in their volume *Statistical Physics* [51] and where in the middle of their demonstration, they refer to their volume *Non-Relativistic Quantum Mechanics* for a crucial step [52]. The relevance to single-mode optical fibers is evident since, as stated earlier, the modes in these fibers can be treated as weakly confined plane waves [45,46]. This result was new to me, and I immediately did a deep dive into the textbooks on my shelf, and discovered that they either did not contain this result or mis-stated it.

Given the importance of this result and the difficulty of tracing it through three volumes of Landau and Lifshitz's *Course of Theoretical Physics*, a demonstration of this result that is specialized to electromagnetic media is appropriate. In 2010, I had the opportunity to present this result in a tutorial talk at the IEEE Photonics Society Summer Topicals Meeting [53]. However, I have never until now published the details.

The result that we will use is [50].

$$\tilde{\epsilon}_{ij}(\omega,\mathbf{H}) = \tilde{\epsilon}_{ji}(\omega,-\mathbf{H}), \tag{14}$$

where we write the electric displacement **D** as

$$\mathbf{D}(\mathbf{r},t) = \int_{-\infty}^{t} \boldsymbol{\epsilon}(t-t')\cdot\mathbf{E}(\mathbf{r},t')\,dt', \tag{15}$$

with permittivity $\boldsymbol{\epsilon}$. We use tildes to represent the Fourier (frequency) transforms of all quantities, so that

$$\tilde{\boldsymbol{\epsilon}}(\omega) = \int_{-\infty}^{\infty} \boldsymbol{\epsilon}(t)\exp(-i\omega t)\,dt. \tag{16}$$

Equation (14) is a consequence of the fluctuation-dissipation theorem and holds regardless of the loss. However, Eq. (15) has a number of assumptions that are worth enumerating explicitly.

The most general form of the electric displacement is

$$\mathbf{D}(\mathbf{r},t) = \int_{-\infty}^{\infty} dt' \int_V d^3\mathbf{r}'\,\boldsymbol{\epsilon}(\mathbf{r},\mathbf{r}',t,t')\cdot\mathbf{E}(\mathbf{r}',t'), \tag{17}$$

where we integrate over the volume $V$ of the medium. If we assume *homogeneity* in time and space, then Eq. (17) becomes

$$\mathbf{D}(\mathbf{r},t) = \int_{-\infty}^{\infty} dt' \int_V d^3\mathbf{r}'\,\boldsymbol{\epsilon}(\mathbf{r}-\mathbf{r}',t-t')\cdot\mathbf{E}(\mathbf{r}',t'). \tag{18}$$

Time invariance is almost always a reasonable assumption for times that are short compared to the environmental fluctuations. Homogeneity is also a reasonable assumption for lengths of optical fiber that are subject to the same environment, such as straight lengths of optical fibers or fibers that are wound on a bobbin with a fixed radius. If we now assume *locality* then Eq. (18) collapses to

$$\mathbf{D}(\mathbf{r},t) = \int_{-\infty}^{\infty} dt'\,\boldsymbol{\epsilon}(\mathbf{r},t-t')\cdot\mathbf{E}(\mathbf{r},t'). \tag{19}$$

The locality assumption is separate from homogeneity and would apply for example to a straight fiber, but not, as previously noted, to a twisted fiber. Finally, if we assume *causality*, we obtain Eq. (15). Causality, of course, must always hold in any physical system. Equation (15) becomes in the frequency domain

$$\tilde{\mathbf{D}}(\mathbf{r},\omega) = \tilde{\boldsymbol{\epsilon}}(\omega)\cdot\tilde{\mathbf{E}}(\mathbf{r},\omega). \tag{20}$$

We now specialize to a system in which the loss is low and consider a signal at a single frequency $\omega$ so that

$$\mathbf{E}(\mathbf{r},t) = \frac{1}{2}\mathbf{E}_0\exp(i\omega t) + c.c., \qquad \mathbf{D}(\mathbf{r},t) = \frac{1}{2}\mathbf{D}_0\exp(i\omega t) + c.c. \tag{21}$$

The time rate of change of the energy is given by [50].

$$\frac{dQ}{dt} = \mathbf{E}(\mathbf{r},t)\cdot\frac{d\mathbf{D}(\mathbf{r},t)}{dt}, \tag{22}$$

which when averaged over one period becomes

$$\frac{d\overline{Q}}{dt} = \frac{i\omega}{4}\left[\mathbf{E}_0^*\cdot\tilde{\boldsymbol{\epsilon}}(\omega)\cdot\mathbf{E}_0 - \mathbf{E}_0\cdot\tilde{\boldsymbol{\epsilon}}^*(\omega)\cdot\mathbf{E}_0^*\right] = \frac{i\omega}{4}\mathbf{E}_0^*\cdot[\tilde{\boldsymbol{\epsilon}}(\omega) - \tilde{\boldsymbol{\epsilon}}^\dagger(\omega)]\cdot\mathbf{E}_0^*, \tag{23}$$

where $\overline{Q} = (\omega/2\pi)\int_0^{2\pi/\omega} dt'\,Q(t)$. If the loss is negligible then $\tilde{\boldsymbol{\epsilon}}(\omega)$ must be Hermitian. Using Eq. (14) when no magnetic field is present, we conclude that $\tilde{\boldsymbol{\epsilon}}(\omega)$ is a real symmetric tensor and so the birefringence must be linear.

Equation (14) is a consequence of the fluctuation-dissipation theorem, which in the classical limit may be written

$$\langle E_j(t)E_k(t+\tau)\rangle = \frac{k_B T}{V}\int_\tau^\infty \epsilon_{jk}^{-1}\,d\tau', \tag{24}$$

where $k_B$ is Boltzmann's constant, $T$ is the temperature, and $V$ is the volume being considered. In Appendix A, we include a derivation of this result, which specializes a derivation of Lenk [54] to electrodynamics. The power spectral density then becomes (Wiener-Khintchine theorem)

$$\Phi_{jk}(\omega) = \int_{-\infty}^{\infty}\langle E_j(t)E_k(t+\tau)\rangle\exp(-i\omega\tau)\,d\tau = i\frac{k_B T}{\omega V}\left[\tilde{\epsilon}_{jk}^{-1}(\omega) - \tilde{\epsilon}_{kj}^{-1}(\omega)\right]. \tag{25}$$

An exact quantum-mechanical derivation yields [52].

$$\Phi jk(\omega) = \frac{i\hbar}{2V}\coth\left(\frac{\hbar\omega}{2k_B T}\right)\left[\tilde{\epsilon}_{jk}^{-1}(\omega) - \tilde{\epsilon}_{kj}^{-1}(\omega)\right]. \tag{26}$$

At optical frequencies and room temperatures, the quantum-mechanical form should be used. Microscopic time reversibility implies that the autocorrelation function in Eq. (24) must be the same in both directions of time as long as the sign of the magnetic field is reversed, so that $\Phi_{jk}(\omega, \mathbf{H}) = \Phi_{kj}(\omega, -\mathbf{H})$. It then follows that $\Im\left[\tilde{\epsilon}_{jk}^{-1}(\omega,\mathbf{H})\right] = \Im\left[\tilde{\epsilon}_{kj}^{-1}(\omega,-\mathbf{H})\right]$. From the Kramers-Kronig relations, which are themselves a consequence of causality, we find $\Re\left[\tilde{\epsilon}_{jk}^{-1}(\omega,\mathbf{H})\right] = \Re\left[\tilde{\epsilon}_{kj}^{-1}(\omega,-\mathbf{H})\right]$. It then follows $\tilde{\epsilon}_{jk}^{-1}(\omega,\mathbf{H}) = \tilde{\epsilon}_{kj}^{-1}(\omega,-\mathbf{H})$ and finally $\tilde{\epsilon}_{jk}(\omega,\mathbf{H}) = \tilde{\epsilon}_{kj}(\omega, -\mathbf{H})$.

## 4. Conclusions

The 1980s were an extraordinarily fruitful time in nonlinear optics with the advent low-loss optical fibers and laser sources that could operate at 1.5 μm. These two developments made possible the first observation of optical fiber solitons by Linn Mollenauer, Roger Stolen, and Jim Gordon at the Holmdel laboratory of AT&T Bell Laboratories, based on the earlier predictions of Akira Hasegawa and Fred Tappert. It was also an extraordinarily fruitful period in nonlinear dynamics. Powerful computers became available that made it possible to observe the interplay between regularity and chaos that exists in dynamical systems and that had previously only been accessible through difficult analytical calculations. The convergence of these two strands of thought led me to propose that solitons would function like KAM trajectories that remain regular in the presence of chaos and that therefore solitons would remain robust in the presence of Hamiltonian deformations of the NLSE. Applying this hypothesis to birefringent optical fibers, I proposed that solitons would remain robust in the presence of optical fiber birefringence. This hypothesis was correct with the important caveat that the birefringent walkoff length must be smaller than the nonlinear scale length. Since that condition was not satisfied (and still is not) in the regimes of interest for optical fiber communications, it was necessary to invoke the random variation of the birefringence in optical fibers to explain the observed stability of solitons in optical fibers where the nonlinear scale length exceeded the beat length by several orders of magnitude and solitons can propagate stably over many times the nonlinear scale length. The ramifications of these discoveries occupied much of my time and the time of a research group that I established at UMBC in the 1980s and continue to direct.

The basic equations that govern pulse propagation in birefringent single-mode optical fibers with nonlinearity and dispersion are now well-established as the fundamental equations that govern propagation in single-mode fibers. From the time that they were first written down in 1987 to the present time, they have attracted controversy and questions. In this article, I have addressed three of them: (1) Why don't the two Maker and Terhune coefficients appear independently? (2) Why is the NLSE so successful, given that the coupled NLSE is more fundamental, and what are the limits beyond which the NLSE no longer applies? (3) Is the birefringence really linear as was assumed in the original coupled NLSE? The answer to (1) is that the instantaneous response of the Kerr nonlinearity leads the two coefficients to collapse into one. The answer to (2) is that the signal that is injected into the optical fiber must be in a single polarization state, and PMD must be negligible over the propagation length. That was the case for most optical fiber experiments up to about 1995. In modern telecommunication systems in which polarization division multiplexing is commonly used, the coupled NLSE and its improvements must be used if the Kerr nonlinearity cannot be ignored. The answer to (3) is more nuanced. The assumption of linear birefringence was clearly valid for the original experiments of Stolen and colleagues and remains valid for straight lengths of fiber that are used as discriminators in fiber lasers that are passively modelocked using nonlinear polarization rotation. At the same time, twisting a single-mode optical fiber will induce ellipticity in the birefringence of a single-mode optical fiber. In cases where the beat length is smaller than the nonlinear scale length, this effect will lead to a change in the coefficients of the coupled NLSE. If the birefringence is rapidly and randomly varying, even when locally linear, the nonlinear coefficients for both self- and cross-phase modulation become equal.

In any optical system or, for that matter, any electrodynamic system to which Eq. (15) applies, and assuming that the damping rate is small compared to the frequency as is always the case in practical optical systems, the birefringence must be linear. Equation (15) is a common starting point for analyzing both optical resonators and optical waveguides. The basic idea is that the damping of the optical modes that occurs due to material loss must be compensated by thermodynamic fluctuations. That leads to a correlation response that must be an even function of time with a reversal of the magnetic field if it is present, due to microscopic reversibility. From that the symmetry of the imaginary part of the permittivity tensor follows, and the Kramers-Kronig relations then imply that the real part of the permittivity tensor is also symmetric. Assuming that the permittivity tensor is to good approximation Hermitian, the permittivity tensor must be symmetric and real, and so the birefringence must be linear. We have described the mathematical theory in some detail because this important result is not as well known in the optics community as it should be.

I am grateful to the organizers of this special issue, Sonia Boscolo, John Dudley, and Christophe Finot for this opportunity to review my own contributions to the exciting developments in optical solitons. It gives me an opportunity to thank my senior colleagues whose encouragement, guidance, and collaboration made it possible for me to carry out this work. In addition to the opportunity to work with great optical scientists, including Linn Mollenauer, Roger Stolen, Jim Gordon, and Dieter Marcuse, I had the opportunity to work with great applied mathematicians, including Mark Ablowitz, Thanasis Fokas, David Kaup, and Pavel Winternitz. Finally, I am grateful to many colleagues, as well as my students and co-workers at UMBC, too many to name here, without which my work could not have been carried out. Special thanks are due to Mr. Ishraq Md Anjum for his careful reading of this manuscript and his help in typesetting it. Funding for the work reported in this paper, which was carried out over many years, has been provided by the US Department of Energy, the National Science Foundation, the Defense Advanced Research Projects Agency, the Department of Defense, and the Air Force Office of Scientific Research. I am grateful to the program managers who found ways to support my work since the 1980s up to today.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Acknowledgment

Two recent funding sources that are appropriate to explicitly acknowlege are: National Science Foundation, ECCS-1807272, and Air Force Office of Scientific Research, FA9550-20-1-0357.

## Appendix

In this appendix, we derive Eq. (14) in the classical limit, following the approach of Lenk [54] and specializing to electrodynamic systems. While the classical limit does not apply at optical frequencies, the fully quantum description of Callen and Welton [52,55] obscures somewhat the role of microscopic reversibility.

In this discussion, we assume that the electrodynamic field and the material are in thermal equilibrium so that the intrinsic dissipation in the medium, determined by the imaginary part of the permittivity, must be compensated by fluctuations to ensure that there is a fluctuating energy of $k_B T$ per mode. The basic notion is that the relaxation that a medium experiences after the sudden removal of a force makes it possible to probe the magnitude of the fluctuations that are needed to compensate for the dissipation.

We will focus here on the linear (as opposed to nonlinear) polarization density

$$\mathbf{P}(z,t) = \int_{-\infty}^{t} dt' \,\mathsf{X}(t-t')\cdot\mathbf{E}(z,t'), \tag{A1}$$

which is the linear part of Eq. (4) and where $\mathsf{X}$ is the linear polarizability tensor, so that in the frequency domain $\widetilde{\epsilon}(\omega) = \epsilon_0\left[\mathsf{I} + \widetilde{\mathsf{X}}(\omega)\right]$. Hereafter, we drop the $z$-dependence of $\mathbf{P}$ since we are considering a homogeneous volume $V$. We focus on $\mathbf{P}$, rather than $\mathbf{D}$, since the discussion with this choice is somewhat simpler and is also closer to the quantum-mechanical derivation. We may now write

$$\langle P_j(t)P_k(t+\tau)\rangle = \langle P_j(t)\Xi_{jk}(\tau)\rangle, \tag{A2}$$

where $\Xi_{jk}(\tau) = \langle P_k(t+\tau)|P_j(t)\rangle$ and where the brackets $\langle\cdot\rangle$ denote an ensemble average over the equilibrium distribution. Due to ergodicity, this ensemble average is equivalent to a long time average when the system is in equilibrium. Eq. (A2) states that the joint ensemble average of $P_k(t+\tau)$ and $P_j(t)$ is equivalent to the ensemble average of $P_k(t+\tau)$, given a particular $P_j(t)$ and then averaged over all values of $P_j(t)$. We now define the covariance

$$\Phi_{jk}(\tau) = \langle P_j(t)P_k(t+\tau)\rangle = \langle P_j(t)\Xi_{jk}(\tau)\rangle \tag{A3}$$

and the corresponding power spectral density

$$S(\omega) = \int_{-\infty}^{\infty} \Phi_{jk}(\tau)\exp(-i\omega\tau)\,d\tau, \tag{A4}$$

and we explicitly express the covariance $\Phi_{jk}(\tau)$ in terms of the probability density function for $P_j$ as

$$\Phi_{jk}(\tau) = \int w_0(P_j)P_j\Xi_{jk}(\tau)\,d^3P, \tag{A5}$$

where the integrand is over all three components of $\mathbf{P}$ and $w_0$ is the Boltzmann distribution in the absence of an imposed electric field.

In order to relate $\Phi_{jk}(\tau)$ to the susceptibility, we consider a specific electric field perturbation that has the form

$$E_j(t) = \begin{cases} E_{j0}, \text{when } t<0, \\ 0, \text{when } t>0. \end{cases} \tag{A6}$$

The electric field polarizes the medium, which decreases the energy in the medium by an amount $-\epsilon_0 V P_j(t)E_j(t)$. When $t<0$, the equilibrium probability density function for $P_j$ is given by

$$w(P_j) = w_0(P_j)\frac{\exp(\epsilon_0 V P_j E_{j0}/k_B T)}{\langle\exp(\epsilon_0 V P_j E_{j0}/k_B T)\rangle}. \tag{A7}$$

When $t>0$, the relaxation process for $P_k(t)$ is governed by $\Xi_{jk}(t)$ so that the mean value of $P_k(t)$ is given by

$$\langle P_k(t)\rangle = \int w(P_j)\Xi_{jk}(t)\,d^3P = \frac{\langle\exp(\epsilon_0 V P_j E_{j0}/k_B T)\Xi_{jk}(t)\rangle}{\langle\exp(\epsilon_0 V P_j E_{j0}/k_B T)\rangle}. \tag{A8}$$

We may also write

$$\langle P_k(t)\rangle = \int_0^{\infty} X_{kj}(\tau)E_j(t-\tau)\,d\tau, \tag{A9}$$

which becomes in our case

$$\langle P_k(t)\rangle = E_{j0}\int_t^{\infty} X_{kj}(\tau)\,d\tau. \tag{A10}$$

Assuming that $|\epsilon_0 V P_j(t)E_j(t)| \ll k_B T$, we may expand the exponential functions in Eq. (A8), keeping only first-order contributions in $E_{j0}$. We then find

$$\langle P_k(t)\rangle = \frac{\epsilon_0 V E_{j0}}{k_B T}\langle P_j(0)P_k(t)\rangle = \frac{\epsilon_0 V E_{j0}}{k_B T}\Phi_{jk}(t). \quad \text{(A11)}$$

We conclude

$$\Phi_{jk}(t) = \frac{k_B T}{V}\int_t^\infty X_{jk}(\tau)\,d\tau. \quad \text{(A12)}$$

Alternative forms of Eq. (A12) are

$$\begin{aligned} \langle D_j(t)D_k(t+\tau)\rangle &= \frac{k_B T}{V}\int_t^\infty \epsilon_{jk}(\tau)\,d\tau, \\ \langle E_j(t)E_k(t+\tau)\rangle &= \frac{k_B T}{V}\int_t^\infty \epsilon_{jk}^{-1}(\tau)\,d\tau. \end{aligned} \quad \text{(A13)}$$

The second of these two alternative forms is just Eq. (24). Noting that $\Phi_{jk}(-t) = \Phi_{jk}(t)$, we obtain

$$S_{jk}(\omega) = i\frac{k_B T}{\omega V}\left[\widetilde{X}_{jk}(\omega) - \widetilde{X}_{jk}^*(\omega)\right]. \quad \text{(A14)}$$

Equation (A14) is the classical form of the fluctuation-dissipation theorem.

We now invoke microscopic time-reversibility, changing the sign of the magnetic field if a magnetic field is present, so that

$$\begin{aligned} \Phi_{jk}(\tau,\mathbf{H}) &= \langle P_j(t)P_k(t+\tau),\mathbf{H}\rangle = \langle P_j(t)P_k(t-\tau),-\mathbf{H}\rangle = \langle P_j(t+\tau)P_k(t),-\mathbf{H}\rangle \\ &= \Phi_{kj}(\tau,-\mathbf{H}). \end{aligned} \quad \text{(A15)}$$

It then follows that $\Im[\widetilde{X}_{kj}(\omega,\mathbf{H})] = \Im[\widetilde{X}_{kj}(\omega,-\mathbf{H})]$. From the Kramers-Kronig relations, we now infer $\Re[\widetilde{X}_{kj}(\omega,\mathbf{H})] = \Re[\widetilde{X}_{kj}(\omega,-\mathbf{H})]$. Putting these these two pieces together, we finally conclude

$$\widetilde{\epsilon}_{kj}(\omega,\mathbf{H}) = \widetilde{\epsilon}_{jk}(\omega,-\mathbf{H}). \quad \text{(A16)}$$

## References


[1] A. Hasegawa, F. Tappert, Transmission of stationary nonlinear pulses in optical fibers. I. Anomalous dispersion, Appl. Phys. Lett. 23 (3) (1973) 142–144.
[2] A. Hasegawa, F. Tappert, Transmission of stationary nonlinear pulses in optical fibers. II. Normal dispersion, Appl. Phys. Lett. 23 (4) (1973) 171–172.
[3] C.R. Menyuk, Nonlinear pulse propagation in birefringent optical fibers, IEEE J. Quant. Electron. 23 (2) (1987) 174–176.
[4] C.R. Menyuk, Stability of solitons in birefringent optical fibers. I: equal propagation amplitudes, Opt. Lett. 12 (8) (1987) 614–616.
[5] C.R. Menyuk, Stability of solitons in birefringent optical fibers. II. Arbitrary amplitudes, J. Opt. Soc. Am. B 5 (2) (1988) 392–402.
[6] L.F. Mollenauer, R.H. Stolen, J.P. Gordon, Experimental observation of picosecond pulse narrowing and solitons in optical fibers, Phys. Rev. Lett. 45 (13) (1980) 1095–1098.
[7] R.Y. Chiao, E. Garmire, C.H. Townes, Self-trapping of optical beams, Phys. Rev. Lett. 15 (13) (1964) 479–482.
[8] P.L. Kelley, Self-focusing of optical beams, Phys. Rev. Lett. 15 (26) (1965) 1005–1008.
[9] V.E. Zakharov, A.B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional modulation of waves in nonlinear media," Sov. Phys. JETP 34:1, 62–69 (1972), [Zh. Eksp. Theor. Fiz. 1 (61) (1971) 118–134.
[10] C.S. Gardner, J.M. Greene, M.D. Kruskal, R.M. Miura, Method for solving the Korteweg-de Vries equation 19 (19) (1967) 1095–1097.
[11] M.J. Ablowitz, H. Segur, Solitons and the Inverse Scattering Transform, 1981. SIAM, Philadelphia.
[12] H. Poincaré, Les méthodes nouvelles de la mécanique céleste, in: Tome, vol. 1, 1882. Gauthier-Villars, Paris.
[13] H. Scott Dumas, The KAM Story: A A Friendly Introduction to the Content, History, and Significance of Classical Kolgomorov-Arnold-Moser Theory, 2014. World Scientific.
[14] A.J. Lichtenberg, M.A. Lieberman, Regular and Chaotic Dynamics, 1992. Springer-Verlag.
[15] B.V. Chirikov, A universal instability of many-dimensional oscillator systems, Phys. Rep. 52 (5) (1979) 263–379.
[16] C.R. Menyuk, Non-linear Evolution of an Obliquely Propagating Langmuir Wave, 1981. PhD Dissertation, U. California Los Angeles.
[17] R.H. Stolen, J. Botineau, A. Ashkin, Intensity discrimination of optical pulses with birefringent fibers, Opt. Lett. 7 (1982) 10.
[18] J. Botineau, R.H. Stolen, Effect of polarization on spectral broadening in optical fibers, J. Opt. Soc. Am. 72 (1982) 12.
[19] P.-K.A. Wai, Solitons Near the Zero Dispersion Point of Optical Fibers, 1988. PhD Dissertation, U. Maryland College Park.
[20] I. Kaminow, Polarization in optical fibers, IEEE J. Quant. Electron. 17 (1) (1981) 15–22.
[21] S. Trillo, S. Wabnitz, R.H. Stolen, G. Assanto, C.T. Seaton, G. I Stegeman, Experimental observation of polarization instability in a birefringent optical fiber, Appl. Phys. Lett. 49 (19) (1986) 1224–1226.
[22] C.D. Poole, R.E. Wagner, Phenomenological approach to polarisation dispersion in long single-mode fibres, Electron. Lett. 22 (19) (1986) 1029–1030.
[23] L.F. Mollenauer, K. Smith, J.P. Gordon, C.R. Menyuk, Resistance of solitons to the effects of polarization dispersion in optical fibers, Opt. Lett. 14 (21) (1989) 1219–1221.
[24] C.R. Menyuk, Soliton robustness in optical fibers, J. Opt. Soc. Am. B 9 (10) (1993) 1585–1591.
[25] M.N. Islam, C.D. Poole, J.P. Gordon, Soliton trapping in birefringent optical fibers, Opt. Lett. 14 (18) (1989) 1011–1013.
[26] M.N. Islam, C.R. Menyuk, C.-J. Chen, C.E. Soccolich, Chirp mechanisms in soliton dragging logic gates, Opt. Lett. 16 (4) (1991) 214–216.
[27] R.W. Hellwarth, Third-order optical susceptabilities of liquids and solids, Prog. Quant. Electron. 5 (1977) 1–68.
[28] C.R. Menyuk, M.N. Islam, J.P. Gordon, Raman effect in birefringent optical fibers, Opt. Lett. 16 (8) (1991) 566–568.
[29] G.J. Foschini, C.D. Poole, Statistical theory of polarization dispersion in single mode fibers, J. Lightwave Technol. 9 (11) (1991) 1439–1456.
[30] P.K.A. Wai, C.R. Menyuk, Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence, J. Lightwave Technol. 14 (2) (1996) 148–157.
[31] D. Marcuse, C.R. Menyuk, P.K.A. Wai, Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence, J. Lightwave Technol. 15 (9) (1997) 1735–1746.
[32] G.P. Agrawal, Nonlinear Fiber Optics, 1989. Academic.
[33] M.E. Fermann, Ultrafast fiber oscillators, in: M. Fermann, A. Galvanauskas, G. Sucha (Eds.), Ultrafast Lasers: Technology and Applications, 2003. Marcel Dekkar.
[34] A. Hasegawa, Self-confinement of multimode optical pulse in a glass fiber, Opt. Lett. 10 (5) (1980) 416–417.
[35] B. Crosignani, P. Di Porto, Soliton propagation in multimode fibers, Opt. Lett. 6 (7) (1981) 329–330.
[36] L.G. Wright, D.N. Christodoulides, F.W. Wise, Spatiotemporal mode-locking in multimode fiber lasers, Science 358 (6359) (2017) 94–97.
[37] L.G. Wright, P. Sidorenko, H. Pourbeyam, Z.M. Ziegler, A. Isichenko, B. A. Malomed, C.R. Menyuk, D.N. Christodoulides, F.W. Wise, Mechanisms of spatotemporal mode-locking, Nat. Phys. 16 (2020) 565–570.

[38] N. Bozinovic, Y. Yue, Y. Ren, M. Tur, P. Kristensen, H. Huang, A.E. Willner, S. Rmachandran, Terabit-scale orbital angular momentum mode division multiplexing, Science 340 (6140) (2013) 1545–1548.
[39] M. Stratmann, T. Pagel, F. Mitschke, Experimental observation of temporal soliton molecules 95 (2005), 143902.
[40] A. Haboucha, H. Leblond, M. Salhi, A. Kamarov, F. Sanchez, Coherent soliton pattern formation in a fiber laser 33 (5) (2008) 524–526.
[41] W. He, M. Pang, D.H. Yeh, C.R. Menyuk, P. St, J. Russell, Formation of optical supramolecular structures in a fibre laser by tailoring long-range soliton interactions, Nat. Commun. 10 (2019) 5756.
[42] J.S. Russell, Report on Waves" in Report of the 14th Meeting of the British Association for the Advancement of Science, 1844, pp. 311–390. John Murray, London.
[43] P.D. Maker, R.W. Terhune, C.M. Savage, Intensity-dependent changes in the refractive index of liquids, Phys. Rev. Lett. 12 (18) (1964) 507–509.
[44] D. Gloge, Weakly guiding fibers, Appl. Opt. 10 (10) (1971) 2252–2258.
[45] C.R. Menyuk, Application of multiple-length-scale methods to the study of optical fiber transmission, J. Eng. Math. 36 (1–2) (1999) 113–136.
[46] C.R. Menyuk, Pulse propagation in an elliptically birefringent Kerr medium, IEEE J. Quant. Electron. 25 (12) (1989) 2674–2682.
[47] A. Simon, R. Ulrich, Evolution of polarization along a single-mode fiber, Appl. Phys. Lett. 31 (8) (1977) 517–520.
[48] C.R. Menyuk, P.K.A. Wai, Elimination of nonlinear polarization rotation in twisted fibers, J. Opt. Soc. Am. B 7 (11) (1994) 1307–1309.
[49] A. Galtarossa, L. Palmieri, Reflectometric measurements of polarization properties in optical-fiber links, in: A. Galtarossa, C.R. Menyuk (Eds.), Polarization Mode Dispersion, 2005, pp. 168–197. Springer.
[50] L. D. Landau and E. M. Lifshitz, Electrodynamics of continuous media (pergamon, 1960). Sec. 96, pp. 313–315. See in particular Eq. (76.3). NOTE: These authors use Gaussian units and the negative frequency convention. This paper uses SI units and the positive frequency convention. See also Sec. 83, pp. 337–342.
[51] L.D. Landau, E.M. Lifshitz, Statistical physics, Pergamon, Secs 125–127 (1969) 384–400.
[52] L.D. Landau, E.M. Lifshitz, Quantum Mechanics — non-relativistic theory, Pergamon, Sec 42 (1958) 146–147.
[53] C. R. Menyuk, "The fluctuation-dissipation theorem and birefringence in optical fibers," 2010 IEEE Photonics Society Summer Topicals Meeting on Polarization Division Multiplexed Transmission Systems, Paper TuA1.
[54] R. Lenk, A simple proof of the classical fluctuation dissipation theorem, Phys. Lett. 3 (25A) (1967) 198–199.
[55] H.B. Callen, T.A. Welton, Reversibility and generalized noise, Phys. Rev. 83 (1) (1951) 34–40.