Fast and Slow Nonlinearities in Epsilon-Near-Zero Materials

Jacob B. Khurgin,* Matteo Clerici, and Nathaniel Kinsey

Novel materials, with enhanced light–matter interaction capabilities, play an essential role in achieving the lofty goals of nonlinear optics. Recently, epsilon-near-zero (ENZ) media have emerged as a promising candidate to enable the enhancement of several nonlinear processes including refractive index modulation and harmonic generation. Here, the optical nonlinearity of ENZ media is analyzed to clarify the commonalities with other nonlinear media and its unique properties. Transparent conducting oxides as the family of ENZ media with near-zero permittivity in the near-infrared (telecom) band are focused on. The instantaneous and delayed nonlinearities are investigated. By identifying their common origin from the band nonparabolicity, it is shown that their relative strength is entirely determined by a ratio of the energy and momentum relaxation (or dephasing) times. Using this framework, ENZ materials are compared against the many promising nonlinear media that are investigated in literature and show that while ENZ materials do not radically outpace the strength of traditional materials in either the fast or slow nonlinearity, they pack key advantages such as an ideal response time, intrinsic slow light enhancement, and broadband nature in a compact platform making them a valuable tool for ultrafast photonics applications for decades to come.

1. Introduction

Nonlinear optics has been and remains a fascinating field since the 1960s, ignited shortly after the invention of the laser.[1,2] In the most general form, nonlinear optics enables the optical control of light in the temporal and frequency domains, resulting in phenomena such as the intensity-dependent index, nonlinear and multiphoton absorption, self and cross-phase modulation, wave mixing, sum and difference-frequency generation, and multiple harmonic generation. Over the last 60 years, advances made in the study and application of these effects have culminated in notable practical solutions including Kerr mode-locking,[3] optical frequency combs,[4] optical parametric generators and oscillators,[5] high-harmonic generation,[6] and others.

Yet from the inception of the field, it has been understood that these nonlinear photon–photon interactions are mediated by materials. As such, the development of new and high-performance nonlinear optical media has been intertwined with the discovery of new processes and the realization of practical solutions. Despite decades of study and the several potentially transformative developments— think of organics,[7] semiconductor quantum wells and superlattices,[8] plasmonic metals,[9] carbon nanotubes,[10] and 2D materials[11,12]—the list of materials playing a leading role in nonlinear optical applications has not significantly expanded beyond what was available in the 1980s. Indeed, second-order processes are mostly enabled by conventional crystals like LiNbO$_3$,[13] BBO,[14] and KTP[15] from UV to near-IR, and AgGaSe$_2$ and other chalcogenides for mid- and far-IR.[16] Similarly, applications relying on third-order processes are mostly based on silica fiber,[17] SiN,[18] Si,[19] III–V,[20] and II–VI[21] semiconductors as well as chalcogenide glasses.[22] This illustrates the extreme difficulty to find a material that can satisfy the sometimes opposing requirements imposed by applications, such as large nonlinearity, speed, wide optical bandwidth, high damage threshold, thermal stability, and others.

In the last few years, a new class of promising nonlinear materials has emerged, characterized by a refractive index (and hence the real part of dielectric constant) that approaches zero, called epsilon-near-zero or ENZ materials. The real part of ϵ may approach zero in many systems, such as in polar crystals near the longitudinal optical phonon (SiC) and in metals at the plasma frequency. In the visible and near-IR range, ENZ materials are predominantly realized with heavily (degenerately) doped semiconductors such as indium tin oxide (ITO), aluminum zinc oxide (AZO), gallium-doped zinc oxide (GZO), and others in the transparent conducting oxide (TCO) family.[23] We shall focus on TCOs from here on, due to their relevance to ultrafast photonics applications in the visible and telecom spectral ranges.

Prof. J. B. Khurgin
Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218, USA
E-mail: jakek@jhu.edu

Dr. M. Clerici
School of Engineering
University of Glasgow
Glasgow G128QQ, UK

Dr. N. Kinsey
Department of Electrical and Computer Engineering
Virginia Commonwealth University
Richmond VA 23284, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/lpor.202000291

Simply, the concept of ENZ can be understood by considering the contributions to the dielectric constant of the TCO described by the well-known Drude–Lorentz relative permittivity $\varepsilon(\omega) = \varepsilon_\infty - Ne^2/\varepsilon_0 m^*\omega(\omega + i\gamma)$ where m^* is effective mass. When the free carrier density is sufficiently large, the negative contribution due to free electrons, and the positive contribution ε_∞, arising from bound electrons, nearly cancel each other at a particular wavelength and a number of interesting effects ensue such as low group velocity, enhanced diffraction, and in particular, enhanced nonlinear phenomena. In the case of TCOs, when the carrier density reaches 10^{20}–10^{21} cm^{-3}, the cancellation occurs in the telecommunications spectral range, which has clear practical implications. As a result, ENZ materials have experimentally demonstrated enhancement to several nonlinear phenomena including (but not limited to) third[24] and higher[25] harmonic generation and cross-phase modulation.[26,27]

In experiments dealing with switching,[28] phase conjugation,[29] negative refraction,[30] cross-phase modulation,[31] and adiabatic frequency shifts,[32,33] the origin of the nonlinearity is generally explained by an increase of the electron temperature that exhibits a temporal response of a few hundred femtoseconds, commensurate with the rate of thermal relaxation of hot carriers—a *slow* nonlinearity (although it is still fast, on the order of 1 ps, in an absolute sense). At the same time, experiments demonstrating enhanced harmonic generation[24,25,34] cannot be explained by such a slow nonlinearity and require a *fast* or instantaneous temporal effect.

It appears then that ENZ materials possess more than one kind of nonlinearity, and it is not clear if and how they are related to each other. Here, we seek to solve this conundrum, discussing the relation between fast and slow nonlinearities in TCOs. In addition, we utilize this framework to compare the performance of ENZ materials to the wide range of previous nonlinear materials. To do so, in Section 2, we outline the common origin of *fast* and *slow* optical nonlinearities in dielectrics and establish how their strengths are related. In Section 3, we derive equations for carrier transport in a nonparabolic band. Based on that, in Section 4 we derive the expression and estimate the magnitude of *fast* nonlinearity in ENZ and show that it is of the same order of magnitude as fast nonlinearity in conventional materials at the same wavelength. In Section 5, we estimate the *slow* nonlinearity in ENZ material, and show that it has the same physical origin as the fast nonlinearity and that the strengths of the two components are related in exactly the same way as in more conventional nonlinear materials. In Section 6, the comparison between different nonlinearities is performed using common figures of merit and the niches where ENZ materials have advantages are identified. Finally, the conclusions are drawn in Section 7.

2. Fast and Slow Nonlinearities: General Discussion

Third-order nonlinearities are the lowest order nonlinear phenomena intrinsic to all media (with or without inversion symmetry). Any third-order optical process is characterized by the nonlinear susceptibility $\chi^{(3)}(\omega_4 = \omega_1 \pm \omega_2 \pm \omega_3)$ which reveals that in essence it is a four-wave interaction. However, since two or more of the waves can be degenerate, this susceptibility can describe self- and cross-phase modulation, third harmonic genera-

tion, and quite a few other nonlinear phenomena. Furthermore, being a complex number, the susceptibility can describe optically induced change in both refraction and absorption.

Third-order processes are usually subdivided into two broad categories—*fast* or instantaneous nonlinearities, and *slow* or delayed nonlinearities. The *fast* or *instantaneous* nonlinearity is associated with *virtual* processes, i.e., not mediated by real transitions. Familiar examples of such nonlinearities are the optical Kerr effect[35] as well as harmonic generation in dielectrics[36] or semiconductors excited well below the bandgap.[35] These processes are usually associated with low optical losses and a limited magnitude of the nonlinearity. The second is the variety of *slow* nonlinearities where real excitations take place and persist over a certain lifetime, which may be the recombination time or thermal diffusion time. A typical example is the nonlinear index for semiconductors excited near the bandgap.[37] The magnitude of these nonlinearities is relatively high and they exhibit resonances. For these processes the response is slow, on the scale of recombination time (picosecond to nanosecond), and losses are large as the nonlinear effects are based on the absorption and excitation of real carriers in the band. The speed can be enhanced by reducing the recombination time, e.g., including defects,[38] although this comes, at the expense of a decreased magnitude. Yet, the ability to engineer the response of slow nonlinearity, e.g., by annealing,[38,39] makes them attractive for optical control applications.
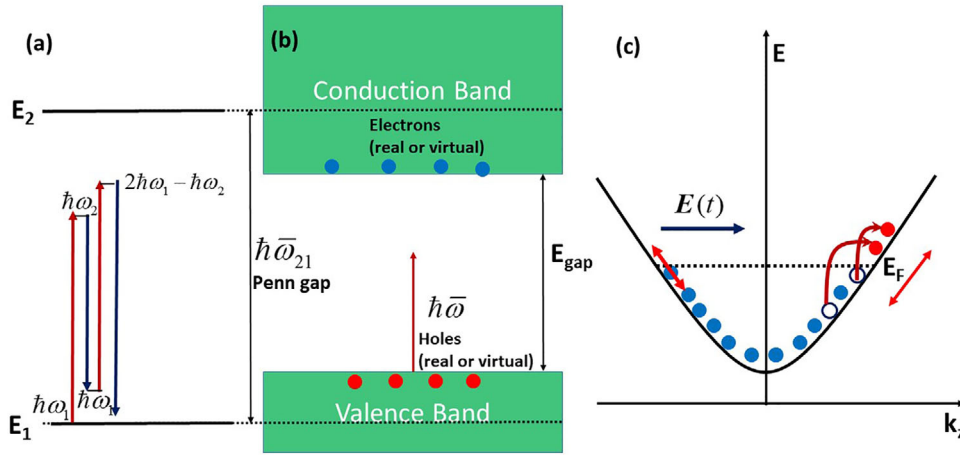
While the fast and slow nonlinearities are usually treated differently, they are in fact closely related, and as we show in this section, practically every nonlinear process has both a fast and a slow component.

The nonlinearity that is perhaps the most familiar originates from the transitions between discrete atomic or molecular energy levels, or between the energy bands in dielectrics or semiconductors. As an example of this process, we consider a degenerate four wave mixing (FWM) in the two-level system (**Figure 1**a), the process in which two waves with frequencies ω_1 and ω_2 mix and build up a time-dependent polarization oscillating with frequency components $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$. This choice allows us to explore the temporal response of the nonlinear susceptibility $\chi^{(3)}(\omega_3 = \omega_1 - \omega_2 + \omega_1)$ by varying the beat frequency $\Delta\omega = \omega_2 - \omega_1$.

The nonlinear medium can be modeled with a two-level system characterized by the usual set of density matrix equations[40,41]

$$\frac{d\Delta\rho}{dt} = -(\Delta\rho - 1)/T_1 + i\frac{\mu_{21} \cdot \mathbf{E}}{\hbar}(\rho_{21} - \rho_{12})$$

$$\frac{d\rho_{21}}{dt} = -\gamma\rho_{21}(\mathbf{r}) - i\omega_{21}\rho_{21}(\mathbf{r}) + i\frac{\mu_{21} \cdot \mathbf{E}}{\hbar}\Delta\rho$$

$$\rho_{21} = \rho_{12}^* \tag{1}$$

where $\hbar\omega_{21}$ is the transition energy, $\Delta\rho = \rho_{11} - \rho_{22}$ is the population difference between the lower and upper levels. The off-diagonal terms ρ_{21} and ρ_{12} are sometimes called coherences and describe the net atomic polarization; $\mu_{21} = ez_{21}$ is the transition dipole matrix element; T_1 is the interlevel relaxation (recombination) time; $\gamma = 1/T_2$ is the scattering rate, where T_2 is the dephasing time. We can write the electric field $E(t) = E_1 e^{-i\omega_1 t} +$

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**LASER
& PHOTONICS
REVIEWS**

www.lpr-journal.org

Figure 1. a) Two-level system and a diagram of a degenerate four wave mixing process in it. b) Two-band solid-state nonlinear material that is approximated by a two-level system c) intraband nonlinearity in a nonparabolic conduction band. On the left, electrons near Fermi level move in the nonparabolic energy band in the presence of optical field $E(t)$—origin of fast nonlinearity. On the right, the electrons get promoted from below the Fermi level to above Fermi level where they are considered "hot" and have higher effective mass—the origin of slow nonlinearity.

$E_2 e^{-i\omega_2 t}$ + c.c., and in the rotating wave approximation we can expand the matrix elements adopting a perturbative approach

$$
\rho_{21} = \sigma_1 e^{-i\omega_1 t} + \sigma_2 e^{-i\omega_2 t} + \sigma_{112} e^{-i(2\omega_1-\omega_2)t} + \sigma_{221} e^{-i(2\omega_2-\omega_1)t} + \cdots
$$
$$
\Delta\rho = 1 - \delta_{12} e^{-i(\omega_1-\omega_2)t} - \delta_{12}^* e^{+i(\omega_1-\omega_2)t} + \cdots
$$
. (2)

The first equation illustrates that the net polarization is a sum of both linear (two first terms) and nonlinear polarizations. Here, we have kept the linear polarization terms $\sigma_{1(2)}$ oscillating at ω_1, ω_2 and FWM terms $\sigma_{112(221)}$ oscillating at $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$, which are the subject of this section. However, one should keep in mind that other third-order terms corresponding to third harmonic, sum frequencies, and others are also present. Similarly, the net population difference between states 1 and 2, $\Delta\rho$, can also be expanded into terms which oscillate at different frequencies. Here, we have kept only the second-order terms δ_{12} describing population variation at the beat frequency $\Delta\omega = \omega_1 - \omega_2$ because it is these population pulsations that engender FWM.

Substituting the expansion (2) into (1) and equating the terms oscillating at the same frequency on the l.h.s. and r.h.s. results first in the steady state solution for the linear atomic polarization

$$
\sigma_{1(2)} = \frac{\mu_{21} E_{1(2)} \hbar^{-1}}{(\omega_{21} - \omega_1) - i\gamma} \tag{3}
$$

which is the traditional Lorentzian. Second, the expression for the oscillations of population difference that occur at the beat frequency is found as

$$
\delta_{12} = \frac{\omega_1 - \omega_2 + 2i\gamma}{\omega_1 - \omega_2 + i/T_1} \frac{\mu_{21}^2 E_1 E_2 / \hbar^2}{(\omega_{21} - \omega_1 - i\gamma)(\omega_{21} - \omega_2 + i\gamma)} \tag{4}
$$

Once this population beating is mixed with a third electric field at frequency ω_1 or ω_2 the third-order coherence at the intermod-

ulation frequency $2\omega_1 - \omega_2$ or $2\omega_2 - \omega_1$ is produced

$$
\sigma_{112} = \frac{\omega_1 - \omega_2 + 2i\gamma}{\omega_1 - \omega_2 + i/T_1}
$$
$$
\times \frac{\mu_{21}^3 E_1^2 E_2 / \hbar^3}{(\omega_{21} - \omega_1 - i\gamma)(\omega_{21} - \omega_2 + i\gamma)(\omega_{21} - 2\omega_1 + \omega_2 - i\gamma)} \tag{5}
$$

Now, the material polarization is $P = N\mu_{21}\rho_{21}$ + c.c., where N is the density of the two-level entities. Therefore, the first-order (linear) polarization can be found as
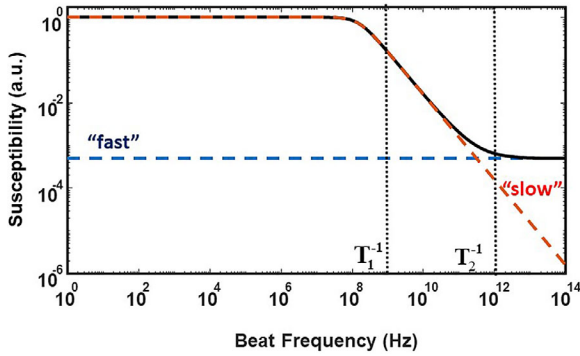
$$
P^{(1)}(\omega_1) = N\mu_{21}\sigma_1 e^{-\omega_1 t} + \text{c.c.} \equiv \varepsilon_0 \chi^{(1)}(\omega_1) E_1^1 e^{-\omega_1 t} + \text{c.c.} \tag{6}
$$

and the third-order nonlinear polarization responsible for FWM can be found as

$$
P^{(3)}(2\omega_1 - \omega_2) = N\mu_{21}\sigma_{112} e^{-i(2\omega_1-\omega_2)t} + \text{c.c.}
$$
$$
\equiv \varepsilon_0 \chi^{(3)}(\omega_3 = \omega_1 - \omega_2 + \omega_1) E_1^2 E_2^* e^{-i(2\omega_1-\omega_2)t} + \text{c.c.} \tag{7}
$$

where we have introduced the first- and third-order susceptibilities, $\chi^{(1)}$ and $\chi^{(3)}$, respectively. If we assume that the frequencies are reasonably close, i.e., $\Delta\omega \ll \omega_1, \omega_2$, then one can approximate all the optical frequencies in the denominator of (5) by the mean frequency $\bar{\omega} = (\omega_1 + \omega_2)/2$, and obtain the expression for third-order susceptibility for the FWM process

$$
\chi^{(3)}(\omega_3 = \omega_1 - \omega_2 + \omega_1) \sim \frac{\Delta\omega + 2j\gamma}{\Delta\omega + i/T_1} \frac{N\mu_{21}^4 / \varepsilon_0 \hbar^3}{\left[(\omega_{21} - \bar{\omega})^2 + \gamma^2 \right]^{3/2}} e^{j\Phi}
$$
$$
= \left(1 + \frac{2T_1/T_2 - 1}{1 - i\Delta\omega T_1} \right) \frac{N\mu_{21}^4 / \varepsilon_0 \hbar^3}{\left[(\omega_{21} - \bar{\omega})^2 + \gamma^2 \right]^{3/2}} e^{j\Phi} = \chi^{(3)}_{\text{fast}} + \chi^{(3)}_{\text{slow}}
$$
(8)

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

Figure 2. Frequency dependence of the FWM susceptibility (solid line) and its fast and slow components (dashed lines) for the material system with $T_1 = 1$ns and $T_2 = 1$ps.

where Φ is the phase that is not important for the present discussion.

The nonlinear polarization contains two terms that have common dependence on material parameters but differ in their dependence of beat frequency as shown in **Figure 2**. There is a fast (almost instantaneous) nonlinear response $\chi_{\text{fast}}^{(3)}$ that is frequency independent, and a slow response $\chi_{\text{slow}}^{(3)}$ whose frequency response has 3 dB bandwidth $1/2\pi T_1$. The slow term is related to absorption and the so-called "real" excitations, where electrons get excited (instantaneously) to the upper state and then decay after the time T_1. The fast term is often related to "virtual" excitations, where carriers spend very short time in the upper state, determined by the uncertainty principle.

At low beat frequencies, $\Delta\omega < 1/T_1$, the slow term is larger than the fast one by $2T_1/T_2 - 1$. For most nonlinear media the ratio $2T_1/T_2 \gg 1$ and the slow nonlinearity is much larger than the fast one, with the possible exception of low-pressure atomic gases. However, once the beat frequency $\Delta\omega$ exceeds $2/T_2$, the fast nonlinearity becomes dominant.

Notice that even when the two frequencies are the same and one talks about, for example, self-phase modulation, there are still two components to $\chi^{(3)}$—the fast and slow one. It can also be seen that for other nonlinear processes, for example, third-harmonic generation, $\Delta\omega$ gets replaced by 2ω in (8) and the "slow" response becomes extremely small, leaving only the instant nonlinearity.

We may also consider the relative magnitudes of the individual contributions. For the *fast* effect, the third-order susceptibility can be expressed via the first-order susceptibility $\chi^{(1)}(\omega) = \varepsilon(\omega) - 1$ (6) as

$$\chi_{\text{fast}}^{(3)} \sim \chi^{(1)} \frac{e^2 z_{21}^2}{\left[\left(\omega_{21} - \bar{\omega} \right)^2 + \gamma^2 \right] \hbar^2} \tag{9}$$

where z_{21} is on the scale of the interatomic distance, i.e., 0.1 nm and for the case of condensed matter and $\hbar(\omega_{21} - \omega)$ is on the scale of an eV, while $\chi^{(1)}$ is on the scale of 10 eV which means the scale $\chi_{\text{fast}}^{(3)} \approx 10^{19} \text{m}^2\text{V}^{-2}$, which is easy to interpret as the inverse square of the internal field in a typical polarizable bond.

Although Equation (8) has been derived in the case of a two-level system, the same model can be applied to describe the non-

linear effects of standard dielectric materials. For these materials the upper and lower states are distributed over the valence and conduction bands that extend over ranges comparable to or even wider than the optical bandgap (width of transmission region). In this case, for transitions that are sufficiently detuned from the absorption edge, $\hbar\omega < E_{\text{gap}}$, one can approximate the band by a two-level system with some average optical transition energy, often referred to as the "Penn gap" $\hbar\bar{\omega}_{21}$[42] which is significantly larger than E_{gap} (by a few eV), see Figure 1b. When $\bar{\omega}_{21}$ is substituted for ω_{21} in (9), and, since $\hbar\omega \leq E_{\text{gap}}$, the denominator in (9) is always on the scale of a few eV^2 and is nearly frequency independent. This approximation has been used with a great degree of accuracy to describe both the linear[43,44] and nonlinear[45,46] properties of many crystals and can be used in conjunction with Equations (8) and (9) to estimate the fast nonlinearity for dense materials.

For the "slow" nonlinearity, the situation is different because the absorption underpinning the slow effect does not follow the Lorentz behavior seen in the two-level system. Rather, the absorption changes with frequency, for example, it decreases exponentially below the bandgap as described by the Urbach rule.[47,48] Nevertheless, Equation (8) can still be used by allowing the effective scattering rate $\gamma = T_2^{-1}$ in (8) to be a function of optical frequency $\gamma(\omega)$ (as was done in refs. [49, 50] where it was shown that the effective scattering rate decreases almost exponentially with detuning from the bandgap $E_{\text{gap}} - \hbar\omega$), so at a particular frequency ω the relation between the fast and slow nonlinearities $\chi_{\text{slow}}^{(3)}(\omega)/\chi_{\text{fast}}^{(3)}(\omega) \approx 2T_1/T_2(\omega)$ is conserved.

3. Carrier Motion in the Nonparabolic Band

Now, that we have established that the interband nonlinearity has two components—a slow term, associated with excitation of real carriers, and a fast term, associated with virtual carrier excitation—the question is whether the same approach can be applied to the intraband nonlinearity, based on the motion of free carriers inside the band, which is responsible for the extraordinary properties of ENZ materials.

To answer this question, the motion of electrons in the isotropic band with dispersion $E(k)$, where k is a wavevector, is considered when the harmonic field $E_z(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} +$ c.c. is applied along the z-direction (Figure 1c). The equation of motion for a carrier with the wavevector k_0 in this band is
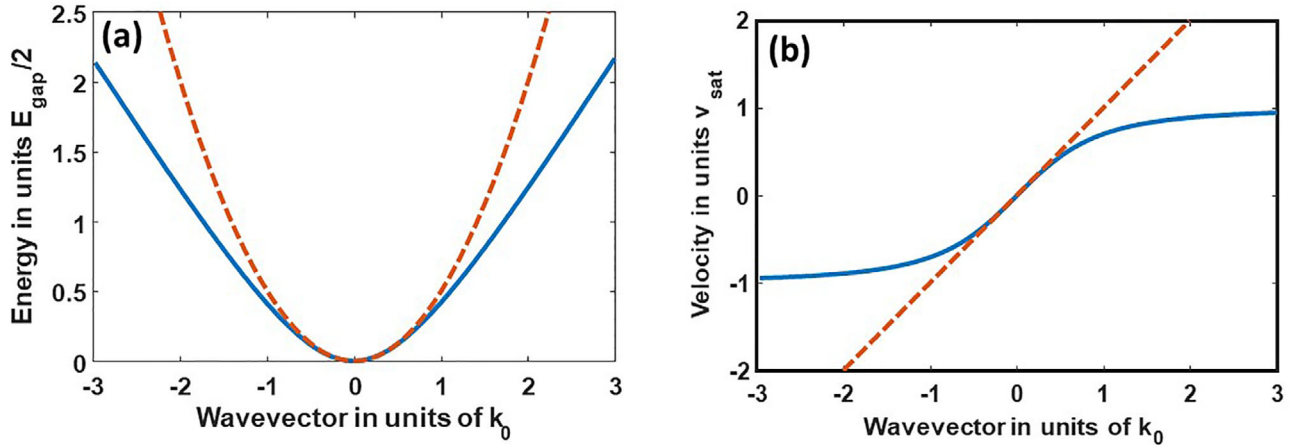
$$\frac{dk_z}{dt} = -\frac{e}{\hbar} E_z - \gamma(k_z - k_{z0}) \tag{10}$$

where γ is the momentum scattering rate. The solution for a harmonic input field is $k_z(t) = k_{z0} + \delta k(t)$, where

$$\delta k(t) = \frac{e}{\hbar(i\omega_1 - \gamma)} E_1 e^{-i\omega_1 t} + \frac{e}{\hbar(i\omega_2 - \gamma)} E_2 e^{-i\omega_2 t} + c.c. \tag{11}$$

The electron velocity in the band $E(k)$ and in the direction of the applied electric field is $v_z = \hbar^{-1}\partial E/\partial k_z$, and can be expanded into the power series around k_0 as

$$v_z(t) = v_z(k_{z0}) + \frac{\partial v_z}{\partial k_z}\delta k(t) + \frac{1}{2}\frac{\partial^2 v_z}{\partial k_z^2}\delta k^2(t) + \frac{1}{6}\frac{\partial^3 v_z}{\partial k_z^3}\delta k^3(t) + \dots \tag{12}$$

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

Figure 3. Dispersion of a) energy and b) velocity in the nonparabolic band (solid lines) and its parabolic approximation (dashed lines).

Since due to time reversal symmetry $v(k) = -v(-k)$, summation over all the filled states in the band will produce a net zero velocity for the time independent term, we may drop the first term in (12). The first remaining term that is linear in $\delta k(t)$ can then be represented by introducing the (inverse) transport effective mass as

$$m_t^{-1} = \hbar^{-1} \frac{\partial v_z}{\partial k_z} \tag{13}$$

so that $\delta v_z = \hbar m_t^{-1} \delta k(t)$. The expression for the transport effective mass is derived as[23]

$$m_t^{-1} = \hbar^{-1} \left[\frac{2}{3} \frac{v(k)}{k} + \frac{1}{3} \frac{dv(k)}{dk} \right] = \frac{1}{3\hbar k^2} \frac{d}{dk} \left(k^2 v(k) \right) \tag{14}$$

This definition is somewhat different from the conventional definition of the effective mass $m_c(k) = \hbar^{-1} dv/dk$ as well as from the alternative definition of the optical effective mass $m_{opt}(k) = \hbar^{-1} v/k$. However, for the case of an ideal parabolic band with a linear velocity–momentum relation, all three definitions result in the same value. For the nonparabolic band, though there is an important difference. Using the definition in (14), when the velocity saturates, i.e., $dv/dk = 0$, the effective mass does not go to infinity. Even when $dv/dk < 0$ $m_t(k)$ stays finite and positive. As we shall see further on, for a typical ENZ material, the difference is not drastic and does not change the conclusion.

Now let us introduce the dispersion of a nonparabolic band derived from a simple two-band $\mathbf{k} \cdot \mathbf{P}$ model. According to Kane[51]

$$E(k) = \frac{E_{gap}}{2} \sqrt{1 + \frac{2\hbar^2 k^2}{m^* E_{gap}}} = \frac{E_{gap}}{2} \sqrt{1 + (k/k_0)^2} \tag{15}$$

where m^* is the effective mass at $k = 0$, E_{gap} is the bandgap energy and we have introduced the "nonparabolicity" wavevector $k_0 = \sqrt{m^* E_{gap}/2\hbar^2}$ at which the dispersion changes from "parabola-like" to "linear-like" as shown in **Figure 3**a, where the energy is measured relative to the bottom of conduction band, wavevectors

are normalized to k_0, and the energy is normalized to $E_{gap}/2$. The velocity, plotted in Figure 3b, is

$$v_z(k) = v_{sat} \frac{(k/k_0)}{\sqrt{1 + (k/k_0)^2}} \tag{16}$$

where the saturation velocity is

$$v_{sat} = \frac{1}{2} \frac{E_{gap}}{\hbar k_0} = \frac{\hbar k_0}{m^*} \tag{17}$$

The effective optical mass can be calculated according to (14)

$$m_t^{-1} = \frac{v_{sat}}{\hbar k_0} \frac{2(k/k_0)^2 + 3}{3\left[1 + (k/k_0)^2\right]^{3/2}} = m^{*-1} \frac{2(k/k_0)^2 + 3}{3\left[1 + (k/k_0)^2\right]^{3/2}} \tag{18}$$

and is shown in **Figure 4**a compared to the other two definitions of the effective mass mentioned above ($m_c(k) = \hbar^{-1} dv/dk$ and $m_{opt}(k) = \hbar^{-1} v/k$). We also compute the first and second derivative of the inverse effective mass as it is required below. The first derivative

$$\alpha = \frac{dm_t^{-1}}{dk} = \hbar^{-1} \frac{d^2 v_z}{dk^2} = \frac{m^{*-1}}{k_0} \frac{k}{k_0} \frac{2(k/k_0)^2 + 5}{3\left[1 + (k/k_0)^2\right]^{5/2}} \tag{19}$$

is plotted in Figure 4b (in units of m^{*-1}/k_0), and the second derivative

$$\beta = \frac{d^2 m_t^{-1}}{dk^2} = \hbar^{-1} \frac{d^3 v_z}{dk^3} = \frac{m^{*-1}}{k_0^2} \frac{4(k/k_0)^4 + 14(k/k_0)^2 - 5}{3\left[1 + (k/k_0)^2\right]^{7/2}} \tag{20}$$

is plotted in Figure 4c, normalized to m^{*-1}/k_0^2. Note that in their normalized form both coefficients are of the order of unity, and this fact will play an important role further on when we compare the slow nonlinearity that depends on α with the fast one that depends on β.

From this we can see two possible avenues from which optical nonlinearities can arise, similar to the ideal two-level system
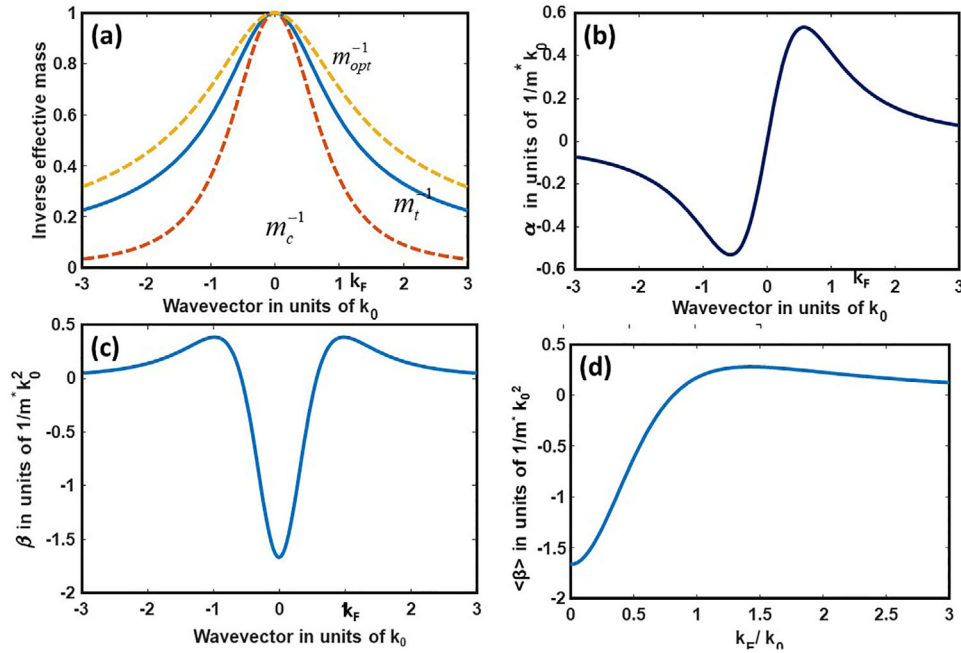
Figure 4. a) Inverse effective masses m_t^{-1} (solid line), m_c^{-1} and m_{opt}^{-1} (dashed lines) normalized to m^{*-1}. b) Normalized first and c) second derivatives of the inverse effective mass versus k_0. d) Mean value of the second derivative versus Fermi wavevector.

described in Section 2. The first avenue is through an applied electric field which effectively polarizes the conduction band electrons in the nonparabolic band (see (12) and Figure 3b). Note that the expansion of carrier velocity, i.e., conductivity current in (12) is conceptually no different from the expansion of polarization (and displacement current) in (6) and (7), and involves no characteristic time. The second avenue is through the absorption of energy which redistributes carriers within the nonparabolic band (see (18) and Figure 4a). As we will show, the first constitutes the *fast* component of the nonlinearity while the second constitutes the *slow* effect as its relaxation is governed by the hot carrier lifetime T_1.

4. Fast Nonlinearity in ENZ Material

To uncover the fast component of the nonlinearity, we return to (12) and express the electron velocity as a function of the first and second derivative of the effective mass

$$v_z(k,t) = \hbar \left[m_t^{-1}\delta k(t) + \frac{1}{2}\alpha \delta k^2(t) + \frac{1}{6}\beta \delta k^3(t) \right] \tag{21}$$

The current density can then be computed as

$$J(t) = -e \sum_k v_z(k,t) = -e \sum_k^{k_F} \hbar \left[m_t^{-1}\delta k(t) + \frac{1}{6}\beta \delta k^3(t) \right]$$

$$\approx -Ne\hbar \left[\langle m_t^{-1} \rangle \, \delta k(t) + \frac{1}{6} \langle \beta \rangle \, \delta k^3(t) \right] \tag{22}$$

where k_F is Fermi wavevector. We note that the summation over k cancels for the second-order term as $\alpha(k)$ is an odd function,
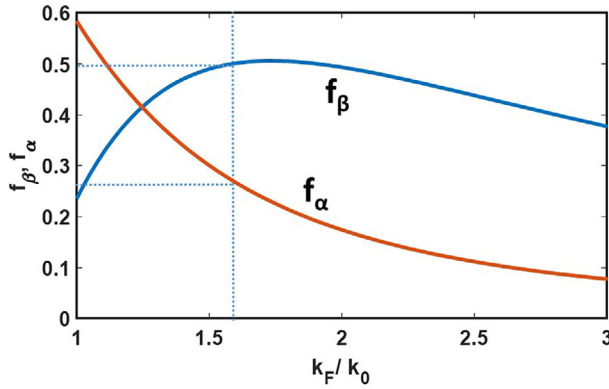
hence the current has only odd order terms—first-order (or linear) $J^{(1)}$ and the third-order one $J^{(3)}$. Here, N is the carrier density and the averaging is done over the distribution of carriers in the band where

$$\langle m_t^{-1} \rangle = \frac{3}{k_F^2} \int_0^{k_F} m_t^{-1}(k) k^2 \mathrm{d}k = \frac{m^{*-1}}{\left[1 + (k_F/k_0)^2 \right]^{1/2}}$$

$$\langle \beta \rangle = \frac{3}{k_F^2} \int_0^{k_F} \beta k^2 \mathrm{d}k$$

$$= \frac{m^{*-1}}{k_0^2} \left\{ \frac{4\sinh^{-1}(k_F/k_0)}{(k_F/k_0)^3} - \frac{4(k_F/k_0)^4 + 11(k_F/k_0)^2 + 4}{(k_F/k_0)^2 \left[1 + (k_F/k_0)^2 \right]^{1/2}} \right\} \tag{23}$$

The latter is plotted in Figure 4d, and once again, when properly normalized it is of the order of unity. Note that approximation (23) is valid for degenerate doping with a Fermi energy approaching 1 eV, i.e., much larger than the thermal energy of electrons and typical ENZ material operating in visible or near-infrared falls into this category.

Substituting (11) into (22), we obtain the linear (first order) current response described by

$$J^{(1)}(\omega_{1,2}) = \frac{iNe^2 \langle m_t^{-1} \rangle}{\omega_{1,2} + i\gamma} E_{1,2} e^{-i\omega_{1,2}t} + \text{c.c.} \tag{24}$$

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

Figure 5. The coefficients f_β for fast nonlinearity and f_α for slow nonlinearity as functions of the Fermi level position.

and the corresponding free carrier part of the susceptibility reads

$$\chi_{fc}^{(1)}(\omega) = -\frac{Ne^2 \langle m_t^{-1} \rangle}{\varepsilon_0 \omega(\omega + i\gamma)} \tag{25}$$

while the total dielectric constant is $\varepsilon(\omega) = \varepsilon_\infty + \chi_{fc}^{(1)}(\omega)$.

The third-order nonlinear response is

$$J^{(3)} = -i\frac{Ne^4 \langle \beta \rangle}{6\hbar^2} \left(\frac{E_1 e^{-i\omega_1 t}}{\omega_1 + i\gamma} + \frac{E_2 e^{-i\omega_2 t}}{\omega_2 + i\gamma} + \text{c.c.} \right)^3 \tag{26}$$

Selecting the FWM term we get

$$J_{FWM}^{(3)} \approx -i\frac{Ne^4 \langle \beta \rangle}{2\hbar^2(\bar\omega^2 + \gamma^2)^{3/2}} \left(E_1^2 E_2 e^{-i(2\omega_1 - \omega_2)t} + E_2^2 E_1 e^{-i(2\omega_2 - \omega_1)t} \right) e^{i\Phi}$$
$$+ \text{c.c.} \tag{27}$$

and the *fast* FWM susceptibility

$$\chi_{fast}^{(3)}(\omega_3 = \omega_1 - \omega_2 + \omega_1) = \frac{Ne^4 \langle \beta \rangle}{2\varepsilon_0 \hbar^2(\bar\omega^2 + \gamma^2)^{3/2}}$$
$$\approx \chi_{fc}^{(1)}(\omega)\frac{e^2 f_\beta}{2\hbar^2 k_0^2 (\bar\omega^2 + \gamma^2)} \tag{28}$$

where

$$f_\beta(k_F) = k_0^2 \langle \beta \rangle \langle m_t \rangle \tag{29}$$

is plotted in **Figure 5**, and is once again on the order of unity. Obviously, Equation (28) describes the fast nonlinearity as its magnitude does not depend on the beat frequency $\Delta\omega = \omega_1 - \omega_2$. Additionally, it can be seen from (26) that the third harmonic susceptibility will have a similar magnitude, which would not be the case for the slow component.

To evaluate the reasonability of the approach, we may consider ITO with a carrier density $N = 1.1 \times 10^{21} \text{cm}^{-3}$, i.e., $k_F = (3\pi^2 N)^{1/3} = 3.2 \text{ nm}^{-1}$. The value of k_0 can be obtained by fitting

as roughly 2 nm^{-1} which makes $k_F/k_0 \approx 1.6$ and $f_\beta \approx 0.5$. Furthermore, since we are concerned with the response near the ENZ condition, $\chi_{fc}^{(1)}(\omega) \sim -\varepsilon_\infty \approx -4$ and so for 1 eV excitation we obtain

$$\chi_{fast}^{(3)}(\omega_3 = \omega_1 - \omega_2 + \omega_1) \approx \varepsilon_\infty \frac{e^2}{2\hbar^2 \omega^2 k_0^2} \langle f_\beta \rangle \approx 2.5 \times 10^{-19} \text{ m}^2\text{V}^{-2} \tag{30}$$

This instant nonlinearity due to free electrons has roughly the same order of magnitude as the fast nonresonant nonlinearity due to bound electrons as estimated in (9). Comparing (30) with (9) one can notice that ε_∞ for ENZ material is comparable to $\chi^{(1)}$ of traditional material, therefore it appears that z_{21} and k_0 are related as $z_{12}k_0 \approx 1$. We shall return with an explanation of this remarkable fact in Section 6.

5. Slow (Thermal) Nonlinearity in ENZ Material

Next, we move to an estimate of the *slow* nonlinearity. To do so, we must calculate the amount of absorbed power per unit volume, found from (24) as

$$\delta P(t) = J(t)E(t) = \frac{2Ne^2\gamma \langle m_t^{-1} \rangle}{\bar\omega^2 + \gamma^2} \left[E_1^2 + E_2^2 + 2E_1 E_2 \cos[(\omega_1 - \omega_2)t] \right] \tag{31}$$

where the oscillations at frequencies of the order 2ω are neglected. The absorbed power is transferred to the hot electrons excited from below to above Fermi level as shown in Figure 1c, left panel. The excess energy density of hot carriers can be found from

$$\frac{dU_{hot}(t)}{dt} = \delta P(t) - \frac{U_{hot}}{\tau_{el}} \tag{32}$$

where τ_{el} is the energy relaxation rate between hot carriers and lattice. We are interested in the energy density oscillations at beat frequency $\Delta\omega = \omega_1 - \omega_2$

$$U_{hot}(t) = \frac{\tau_{el}}{1 - i\Delta\omega\tau_{el}} \frac{2Ne^2\gamma \langle m_t^{-1} \rangle}{\bar\omega^2 + \gamma^2} E_1 E_2 e^{-i(\omega_1 - \omega_2)t} + \text{c.c.} \tag{33}$$

This energy is not shared equally by all the conduction electrons as the hot electrons include both the "primary," generated when photon is absorbed, and the "secondary" carriers, generated via fast (100 fs or less) electron–electron collisions. Thus, the number of excited carriers is not fixed. Usually it is assumed that the electrons thermalize with a certain electron temperature T_e which needs to be evaluated. However, even though electron–electron collissions are fast, it may take longer than 100 fs to establish thermal equilibrium. Therefore, it is preferable to evaluate the nonlinearity without making any assumption of thermal equilibrium.

Let us say the fraction of the "hot" carriers promoted from below the Fermi levels is f_{hot} and their density is $f_{hot}N$. Then the

energy of the average hot carrier oscillates as $\delta E(t) = U(t)/f_{\text{hot}}N$, and the average time-dependent change of the wavevector is

$$\delta k(t) = \delta E(t)/\hbar v_{\text{F}} \tag{34}$$

which causes the change of the inverse effective mass of the hot carriers

$$\delta m_t^{-1}(t) = \frac{\mathrm{d}m_t^{-1}}{\mathrm{d}k}\delta k(t) = \alpha_{\text{F}}\frac{U_{\text{hot}}(t)}{f_{\text{hot}}N\hbar v_{\text{F}}} \tag{35}$$

where $\alpha_{\text{F}} = |\alpha(k_{\text{F}})|$, and describes first derivative of inverse transport effective mass af Fermi level. The change in susceptibility, and hence dielectric constant, can then be estimated as

$$\delta\varepsilon(t) = \delta\chi^{(1)}(t) \approx f_{\text{hot}}\frac{\delta m_t^{-1}}{\langle m_t^{-1}\rangle}\chi^{(1)} = \alpha_{\text{F}}\frac{U_{\text{hot}}(t)}{\langle m_t^{-1}\rangle N\hbar v_{\text{F}}}\chi^{(1)} \tag{36}$$

where the transport mass is averaged over the electron distribution. Note that, as expected, the change in the susceptibility does not depend on f_{hot}. Hence, as long as the electron temperature is not excessively high, $T_e \ll E_{\text{F}}/k_{\text{B}} \approx 15\,000$ K, it is not required to compute the exact distribution of the carriers that are excited to obtain a decent estimate of nonlinearity, and the calculation of the electron temperature is, therefore, unnecessary. Substituting (33) we obtain the beat frequency oscillations of the dielectric constant

$$\delta\varepsilon(\omega_1 - \omega_2) \approx \alpha_{\text{F}}\frac{1}{N\hbar v_{\text{F}}}\frac{\gamma\tau_{\text{el}}}{1 - i\Delta\omega\tau_{\text{el}}}\frac{2Ne^2}{(\bar{\omega}^2 + \gamma^2)}\chi^{(1)}E_1E_2e^{-i(\omega_1-\omega_2)t}$$
$$+ \text{c.c.} \tag{37}$$

When the electric field $E_1e^{-i\omega_1 t}$ scatters from these oscillations, the nonlinear polarization arises, with a *slow* FWM susceptibility given by

$$\chi^{(3)}_{\text{slow}}(\omega_3 = \omega_1 - \omega_2 + \omega_1) = \frac{\gamma\tau_{\text{el}}}{1 - i\Delta\omega\tau_{\text{el}}}\frac{2\alpha_{\text{F}}}{\hbar v_{\text{F}}}\frac{e^2}{\bar{\omega}^2 + \gamma^2}\chi^{(1)} \tag{38}$$

From (16) and (19), we get

$$\frac{\alpha_{\text{F}}}{v_{\text{F}}} = \frac{m^{*-1}}{v_{\text{sat}}k_0}f_\alpha(k_{\text{F}}) = \frac{1}{\hbar k_0^2}f_\alpha(k_{\text{F}}) \tag{39}$$

where

$$f_\alpha(k_{\text{F}}) = \frac{2(k_{\text{F}}/k_0)^2 + 5}{3\left[1 + (k_{\text{F}}/k_0)^2\right]^2} \tag{40}$$

and is plotted in Figure 5, and once again is on the order of unity. Substituting (39) into (38), we obtain the final expression for the slow nonlinearity

$$\chi^{(3)}_{\text{slow}}(\omega_3 = \omega_1 - \omega_2 + \omega_1) = \frac{2\tau_{\text{el}}/\tau_{\text{s}}}{1 - i\Delta\omega\tau_{\text{el}}}\chi^{(1)}\frac{e^2 f_\alpha}{\hbar^2 k_0^2(\bar{\omega}^2 + \gamma^2)} \tag{41}$$

where $\tau_{\text{s}} = \gamma^{-1}$ is the momentum scattering time. Momentum scattering typically occurs due to scattering on impurities and phonons and can be estimated from the mobility measurements.

For ITO, AZO, and many other transparent oxides, it is on the scale of 10 fs or less. The electron-lattice relaxation time τ_{el} is also determined by the electron–phonon scattering, but is typically much longer than τ_s and is on the scale of 100s of femtoseconds. The reason for it is threefold. First of all, momentum scattering occurs independently of whether the scattering event involves the absorption or emission of a phonon, but the energy transfer from hot carriers to lattice is determined by the net emission of phonons, i.e., the difference of phonon emission and absorption rates. Second, the phonon energy is usually less than thermal energy of hot carriers $k_{\text{B}}T_e$, so it takes quite a few phonon emission events per carrier to cool them down. Finally, the impurity and surface scattering contributing to momentum relaxation are elastic processes and do not provide a channel for energy transfer to the lattice.

From Figure 5 one can note that for the example of ITO with $k_{\text{F}}/k_0 = 1.6$, $f_\alpha \approx f_\beta/2$, and for other ENZ materials there is not much difference. Therefore, one can write the relation between the slow (41) and fast (28) nonlinearities as

$$\chi^{(3)}_{\text{slow}} \approx \frac{2\tau_{\text{el}}/\tau_s}{1 - i\Delta\omega\tau_{\text{el}}}\chi^{(3)}_{\text{fast}} \tag{42}$$

which is quite similar to the relation between the slow (real) and fast (virtual) interband nonlinearities in the more conventional materials discussed in Section 2, see (8). It follows that $\chi^{(3)}_{\text{slow}}$ is about two order of magnitude higher than $\chi^{(3)}_{\text{fast}}$, up to $10^{-17}\,\text{m}^2\text{V}^{-2}$. Clearly, the momentum scattering time τ_s plays the role of the dephasing time T_2 and the electron-lattice energy relaxation time τ_{el} plays the role of the recombination time T_1. But the magnitudes of these times are quite different in ENZ materials. Typical values of the recombination times in dielectrics and semiconductors are on the nanosecond scale, while τ_{el} in TCOs is less than a picosecond, hence even the "slow" nonlinearity in TCO can be considered "ultrafast" as the term "ultrafast" is defined today (sub-picosecond, or THz). It is quite conceivable that as ultrafast science progresses, the definition of "ultrafast" may shift into femtosecond domain. But, of course, in TCOs the nonlinearity enhancement is weaker than in the nonlinearity based on saturation of absorption—the usual gain-bandwidth compromise.

6. Discussion

Through this discussion it is clear that there is a strong connection between the nonlinearities in traditional materials and ENZ materials. Let us now consider how the nonlinearity can be tailored for a given application.

The fast nonlinearity does not radically change whether it be from one material to another or if one considers traditional nonlinearities due to bound carriers, or nonlinearities due to free carriers in ENZ materials. We can understand this by considering the terms that play into the strength of the fast nonlinearity—for odd-order traditional nonlinearities only the matrix element of the dipole transition between valence and conduction bands $\mu_{\text{cv}} = ez_{\text{cv}}$, the density, and the detuning from resonance play a role. Among them, the dipole z_{cv} is proportional to the bond length and the density is inversely proportional to its cube, but

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

in all nonlinear materials the bond length remains roughly constant varying between 1.6 and 2.1 Å.[52] Similarly, the effective detuning $\bar{\omega}_{21} - \omega$ is commensurate with bandgap, which is again roughly the same for the materials with the same transparency bandwidth. Together, this explains why the magnitudes of nonlinearities in dielectrics are all similar and only show some increase with an increase in wavelength as one can use materials with narrower bandgap.

For the nonlinearities in ENZ materials, the magnitude depends upon the density of free carriers and the band nonparabolicity. Although the free carrier density is not fixed, if one is to approach the ENZ condition in the telecommunications spectrum (as is the case for most TCOs) the density must be high enough to cancel the positive dielectric constant which forces the majority of materials to maintain a comparable carrier density $\approx 10^{21}$ cm^{-3}. What is important though is that linear susceptibility of the ENZ material due to free carriers is comparable to the linear susceptibility of traditional nonlinear materials. Additionally, the nonlinearity hinges on the nonparabolicity of the band, as expressed by the value of k_0^{-1} in (15). However, the strength of the interband nonlinear and the band nonparabolicity are connected by the same oscillator sum rule[53]

$$\frac{2}{m_0} \sum_{k \neq n} \frac{P_{kn}^2}{E_k - E_n} + \frac{m_0}{m_n^*} = 1 \tag{43}$$

where the first term contains the strength of all the interband transitions originating in the band n and m_n^* is the effective mass of that band. For the two band model, using definition of k_0 this expression can be simplified to

$$\frac{m_0 E_{\text{gap}}}{2\hbar^2 k_0^2} = 1 + \frac{2P_{\text{cv}}^2}{m_0 E_{\text{gap}}} \tag{44}$$

where P_{cv} is the matrix element of the momentum for the valence-to-conduction band transition given by $P_{\text{cv}} = m_0 z_{\text{cv}} E_{\text{gap}}/\hbar$ according to Kane's $\mathbf{k} \cdot \mathbf{P}$[51] theory. Combining these two expressions we can see that $k_0^{-1} \sim z_{\text{cv}}$. As a result, fast nonlinearities in traditional and ENZ materials are both determined by the strength of the interband transition, and thus it follows that their magnitudes are quite similar.

The main difference between fast intraband (free carriers) and interband (bound carriers) nonlinearities is that the loss in the intraband nonlinearities is unavoidable simply because for any photon frequency there always exist plenty of filled initial and empty final states. In the interband nonlinearities, as long as one operates far enough from the absorption edge the loss is minimized. That explains why for the femtosecond nonlinear processes, such as supercontinuum generation and optical frequency combs, or for the harmonic generation, the low loss transparent materials, such as SiO_2, SiN, diamond, and $LiNbO_3$ remain dominant and it would be a trall order for lossy materials like ENZ will replace them.

As for slow nonlinearities, they rely on absorption which provides a broad range of tunability in the nonlinear response for both the traditional and ENZ nonlinearities, with ENZ materials, as explained next, having distinct advantages. To identify the optimal material properties, we approximate the propagation (ab-

sorption) length as $L_a \sim c\chi^{(1)}/\gamma$ which can be used to introduce a figure of merit (FOM)—the product of maximum nonlinearity, propagation length, and the signal bandwidth B—to evaluate and optimize materials. For ENZ materials the FOM becomes

$$\text{FOM} = \chi_{\text{slow}}^{(3)} L_a B \approx \frac{n_g e^2 c f_\alpha}{2\pi\hbar^2 k_0^2 \bar{\omega}^2}, B = 1/2\pi\tau_{\text{el}} \tag{45}$$

and for the traditional nonlinearity

$$\text{FOM} = \chi_{\text{slow}}^{(3)} L_a B \approx \frac{n_g e^2 z_{21}^2 c\bar{\omega}}{2\pi\hbar^2 (\omega_{21} - \bar{\omega})^3}, B = 1/2\pi T_1 \tag{46}$$

where an additional enhancement by the group index n_g has been included due to enhanced interaction time between light and matter.[32,54] For ENZ materials, this can be a factor of a few in thin film form, while for most traditional nonlinearities, a large group index can be attained only by utilizing specially fabricated structures such as photonic crystals.

As one can see, the FOM is rather well defined and constant for a given wavelength, therefore a prudent way to optimization leads to making the bandwidth just sufficient for a given task. In this sense, the ENZ response time, 100 fs < τ_{el} < 1 ps, is almost ideal to allow the nonlinear effects to accumulate over the time τ_{el} while remaining fast enough for many applications in telecommunications such as all-optical switching, adiabatic frequency shifting, and FWM at a speed of a few THz.

For the interband transitions, the recombination time T_1 is about three orders of magnitude longer for direct bandgap materials (and even longer for indirect bandgap), and that is way too long for all optical processing, but may be sufficiently fast for spatial light modulators and or all optical routers. Of course, the speed of the interband nonlinearity can be enhanced if the recombination time is quenched in structures full of defects, for instance, low-temperature growth GaAs,[55] but that introduces significant background absorption, and the only application of these materials is as saturable absorbers, which of course may be extremely useful for some applications, like passive mode-locking,[56] and for photoconductive switching for the generation of THz waves.

Aside from having just about the ideal response time, an important advantage of the ENZ "slow" nonlinearity is that it is fairly broadband (in the sense of optical bandwidth B_{opt}) since the absorption follows the smooth Drude dispersion $\sim \omega^{-2}$. That is dramatically different from the traditional slow nonlinearity whose magnitude may change by orders of magnitude over a few tens of nanometers. Of course, since the absorption length also changes by the equal amount, the FOM is broadband, but that means that the length of the device should be changed when wavelength changes, which makes it impractical. Therefore, as mentioned above slow interband nonlinearities are used almost exclusively as saturable absorbers,[56] an important, but a relatively narrow application niche.

Before concluding, one can also make an interesting comparison with intersubband nonlinearities in semiconductor quantum wells, investigated at length in 1990s[57,58] but given a new life more recently in combination with metasurfaces.[59,60] On one hand, these are intraband nonlinearities (like ENZ), but on the

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

other hand the carriers in them are confined and the transitions are between discrete levels. However, the intersubband relaxation time T_1 and intrasubband scattering time T_2 in square QWs are roughly of the same order of magnitude (100s of fs), while the lineshape is perfectly Lorentzian. As a result, in simple structures there is no huge difference between the "slow" and "fast" nonlinearities in terms of magnitude, although by means of band engineering one can increase the intersubband relaxation time and enhance the slow nonlinearity,[61,62] obviously at the price of making it slower, but still sufficiently fast for THz applications. In that respect, intersubband nonlinearities and nonlinearities in ENZ are very similar (although intersubband nonlinearities are currently limited to mid-IR by the finite values of band off-sets in semiconductors). The main difference lies in the fact that intersubband transitions are relatively narrow and thus one can provide additional resonant enhancement, but obviously at the expense of optical bandwidth, while ENZ nonlinearities are inherently broadband. Therefore, rather than competing, ENZ and intersubband nonlinearities complement each other depending on what is more important: sheer magnitude of the effect or optical bandwidth, and what spectral range is being considered.

One cannot also omit the fact that the hot carrier nonlinearity observed in ENZ materials can and has been observed in metals.[9] However, nonparabolicity in noble metals, most often used in combination with plasmonic enhancement, is much less than in ENZ materials (as well as in other semiconductors). For example, the conduction band in silver has effective mass almost equal to that of free electron in vacuum, meaning that the band is almost perfectly parabolic. In addition, in metals there is no nonlinearity enhancement associated with slow propagation which is one of the main advantages of ENZ materials.

7. Conclusions

In this study, we have placed the nonlinearity in ENZ materials based on conductive oxides into the conventional framework of nonlinear susceptibilities. While the mechanism of nonlinearity in ENZ materials (band nonparabolicity and carrier heating) is seemingly different than the mechanism in more conventional nonlinear materials, ENZ nonlinearities can still be expressed in terms of nonlinear susceptibilities. Just as in conventional materials, there are two components to the effect—a "fast," or "instant," and a relatively "slow" one. Despite apparently different mechanisms, the magnitudes of "fast" nonlinearity at a given wavelength for the ENZ and conventional nonlinearity are roughly of the same magnitude. A deeper look reveals that in the end, intraband ENZ and conventional interband mechanisms rely on the same dipole matrix element of the interband transition to achieve their fast nonlinearity, hence their similar magnitudes (for a given wavelength) should not come as a surprise.

Another important conclusion is that, for both the nonlinearity in ENZ and conventional nonlinearities, the relative strength of slow and fast components is determined by the ratio of two characteristic times, T_1/T_2. T_1 is the energy relaxation time that determines the speed of the "slow" nonlinearity while T_2 is momentum relaxation (or dephasing) time, which also determines the absorption strength. As a result, the figure-of-merit, defined as the product of nonlinear susceptibility, signal bandwidth, and absorption length, is roughly the same for any nonlinear mechanism for a given wavelength. However, ENZ does have three important advantages.

First of all, the energy relaxation time in ENZ being a few hundred femtoseconds happens to be just right for many important applications, which means that the relatively "slow" ENZ nonlinearity is still ultrafast on the absolute scale. In other words, for THz scale processes the nonlinearity in ENZ is as high as it can be (for a given wavelength). Second, the "slow" nonlinearity in ENZ materials is very broadband due to the nonresonant character of absorption. And finally, both the fast and slow nonlinearities in ENZ are enhanced by low group velocity achieved without any required nanofabrication. As a result, many experimental results achieved in the most recent past have made use of these unique benefits to demonstration exceptional nonlinear interactions in ENZ thin films including adiabatic frequency conversion,[32] negative refraction,[29,30] and phase conjugation, all-optical switching,[26,28] bicolor switching,[63] and more.[64,65]

Subsequently demystified, ENZ materials may not be a magic solution for all tasks nonlinear optics is expected to perform, but they do have an important niche to occupy—the niche in which one is looking for a temporal response on the scale of a few hundreds of femtoseconds and compact size. Combined with the fact that many ENZ materials achieve these properties in the near-infrared spectral range using well-established CMOS-compatible materials bodes well for the continued exploration for THz switching and routing in telecommunications as well as compact wave mixing and harmonic generation in both integrated and free-space platforms. As a result, in contrast to the numerous promising platforms explored in the past which have come and gone, we expect ENZ materials to be a mainstay in the nonlinear community for the foreseeable future and continue driving advancements in key application spaces of nonlinear optics.

Conflict of Interest

The authors declare no competing financial interest.

[1] N. Bloembergen, *Nonlinear Optics*, 4th ed., World Scientific, Singapore **1996**, p. 172.

[2] Y. R. Shen, *The Principles of Nonlinear Optics*, Wiley-Interscience, Hoboken, NJ **2003**, p. 563.

2000291 (10 of 11)

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

LASER
& PHOTONICS
REVIEWS

www.lpr-journal.org

[3] D. E. Spence, P. N. Kean, W. Sibbett, *Opt. Lett.* **1991**, *16*, 42.

[4] T. Fortier, E. Baumann, *Commun. Phys.* **2019**, *2*, 153.

[5] S. E. Harris, *Proc. IEEE* **1969**, *57*, 2096.

[6] S. Ghimire, D. A. Reis, *Nat. Phys.* **2019**, *15*, 10.

[7] H. S. Nalwa, *Adv. Mater.* **1993**, *5*, 341.

[8] S. Schmitt-Rink, D. S. Chemla, D. A. B. Miller, *Adv. Phys.* **1989**, *38*, 89.

[9] M. Kauranen, A. V. Zayats, *Nat. Photonics* **2012**, *6*, 737.

[10] J. Wang, Y. Chen, W. J. Blau, *J. Mater. Chem.* **2009**, *19*, 7425.

[11] J. B. Khurgin, *Appl. Phys. Lett.* **2014**, *104*, 161116.

[12] Z. Zeng, X. Sun, D. Zhang, W. Zheng, X. Fan, M. He, T. Xu, L. Sun, X. Wang, A. Pan, *Adv. Funct. Mater.* **2019**, *29*, 1806874.

[13] R. S. Weis, T. K. Gaylord, *Appl. Phys. A* **1985**, *37*, 191.

[14] A. Dubietis, G. Jonušauskas, A. Piskarskas, *Opt. Commun.* **1992**, *88*, 437.

[15] R. C. Eckardt, H. Masuda, Y. X. Fan, R. L. Byer, *IEEE J. Quantum Electron.* **1990**, *26*, 922.

[16] P. Canarelli, Z. Benko, R. Curl, F. K. Tittel, *J. Opt. Soc. Am. B* **1992**, *9*, 197.

[17] G. P. Agrawal, *Nonlinear Science at the Dawn of the 21st Century*, Springer, Berlin **2000**.

[18] D. J. Moss, R. Morandotti, A. L. Gaeta, M. Lipson, *Nat. Photonics* **2013**, *7*, 597.

[19] J. Leuthold, C. Koos, W. Freude, *Nat. Photonics* **2010**, *4*, 535.

[20] E. W. Van Stryland, H. Vanherzeele, M. A. Woodall, M. J. Soileau, A. L. Smirl, S. Guha, T. F. Boggess, *Opt. Eng.* **1985**, *24*, 244613.

[21] T. D. Krauss, F. W. Wise, *Appl. Phys. Lett.* **1994**, *65*, 1739.

[22] B. J. Eggleton, B. Luther-Davies, K. Richardson, *Nat. Photonics* **2011**, *5*, 141.

[23] R. Secondo, J. Khurgin, N. Kinsey, *Opt. Mater. Express* **2020**, *10*, 1545.

[24] A. Capretti, Y. Wang, N. Engheta, L. Dal Negro, *Opt. Lett.* **2015**, *40*, 1500.

[25] Y. Yang, J. Lu, A. Manjavacas, T. S. Luk, H. Liu, K. Kelley, J.-P. Maria, M. B. Sinclair, S. Ghimire, I. Brener, *Nat. Phys.* **2019**, *15*, 1022.

[26] L. Caspani, R. P. Kaipurath, M. Clerici, M. Ferrera, T. Roger, J. Kim, N. Kinsey, M. Pietrzyk, A. Di Falco, V. M. Shalaev, A. Boltasseva, D. Faccio, *Phys. Rev. Lett.* **2016**, *116*, 233901.

[27] M. Z. Alam, I. De Leon, R. W. Boyd, *Science* **2016**, *352*, 795.

[28] N. Kinsey, C. DeVault, J. Kim, M. Ferrera, V. M. Shalaev, A. Boltasseva, *Optica* **2015**, *2*, 616.

[29] S. Vezzoli, V. Bruno, C. DeVault, T. Roger, *Phys. Rev. Lett.* **2018**, *120*, 043902.

[30] V. Bruno, C. DeVault, S. Vezzoli, Z. Kudyshev, T. Huq, S. Mignuzzi, A. Jacassi, S. Saha, Y. D. Shah, S. A. Maier, D. R. S. Cumming, A. Boltasseva, M. Ferrera, M. Clerici, D. Faccio, R. Sapienza, V. M. Shalaev, *Phys. Rev. Lett.* **2020**, *124*, 043902.

[31] V. Bruno, S. Vezzoli, C. DeVault, E. Carnemolla, M. Ferrera, A. Boltasseva, V. M. Shalaev, D. Faccio, M. Clerici, *Appl. Sci.* **2020**, *10*.

[32] J. B. Khurgin, M. Clerici, V. Bruno, L. Caspani, C. DeVault, J. Kim, A. Shaltout, A. Boltasseva, V. M. Shalaev, M. Ferrera, D. Faccio, N. Kinsey, *Optica* **2020**, *7*, 226.

[33] Y. Zhou, M. Z. Alam, M. Karimi, J. Upham, O. Reshef, C. Liu, A. E. Willner, R. W. Boyd, *Nat. Commun.* **2020**, *11*, 2180.

[34] T. S. Luk, D. de Ceglia, S. Liu, G. A. Keeler, R. P. Prasankumar, M. A. Vincenti, M. Scalora, M. B. Sinclair, S. Campione, *Appl. Phys. Lett.* **2015**, *106*, 151103.

[35] M. Sheik-Bahae, D. C. Hutchings, D. J. Hagan, E. W. Van Stryland, *IEEE J. Quantum Electron.* **1991**, *27*, 1296.

[36] R. H. Stolen, A. Ashkin, *Appl. Phys. Lett.* **1973**, *22*, 294.

[37] D. A. B. Miller, C. T. Seaton, M. E. Prise, S. D. Smith, *Phys. Rev. Lett.* **1981**, *47*, 197.

[38] E. S. Harmon, M. R. Melloch, J. M. Woodall, *Appl. Phys. Lett.* **1993**, *63*, 2248.

[39] E. R. Brown, K. A. McIntosh, K. B. Nichols, C. L. Dennis, *Appl. Phys. Lett.* **1995**, *66*, 285.

[40] A. Yariv, *Quantum Electronics*, 2nd ed., Wiley, New York **1975**, 570.

[41] L. Allen, J. H. Eberly, *Optical Resonance and Two-Level Atoms*, Dover, New York **1987**, p. 233.

[42] D. R. Penn, *Phys. Rev.* **1962**, *128*, 2093.

[43] J. C. Phillips, *Rev. Mod. Phys.* **1970**, *42*, 317.

[44] J. A. Van Vechten, *Phys. Rev.* **1969**, *182*, 891.

[45] J. C. Phillips, J. A. Van Vechten, *Phys. Rev.* **1969**, *183*, 709.

[46] D. A. Kleinman, *Phys. Rev. B* **1970**, *2*, 3139.

[47] F. Urbach, *Phys. Rev.* **1953**, *92*, 1324.

[48] T. H. Keil, *Phys. Rev.* **1966**, *144*, 582.

[49] J. B. Khurgin, *Appl. Phys. Lett.* **2014**, *104*, 221115.

[50] J. B. Khurgin, *Phys. Rev. B* **2008**, *77*, 235206.

[51] E. O. Kane, *J. Phys. Chem. Solids* **1957**, *1*, 249.

[52] K. P. Huber, G. Herzberg, *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*, Springer US, Boston, MA **1979**, pp. 8–689.

[53] J. M. Luttinger, W. Kohn, *Phys. Rev.* **1955**, *97*, 869.

[54] N. Kinsey, J. Khurgin, *Opt. Mater. Express* **2019**, *9*, 2793.

[55] M. Haiml, U. Siegner, F. Morier-Genoud, U. Keller, *Appl. Phys. Lett.* **1999**, *74*, 3134.

[56] M. Leitner, P. Glas, T. Sandrock, M. Wrage, G. Apostolopoulos, A. Riedel, H. Kostial, J. Herfort, K.-J. Friedland, L. Däweritz, *Opt. Lett.* **1999**, *24*, 1567.

[57] J. Khurgin, *J. Opt. Soc. Am. B* **1989**, *6*, 1673.

[58] G. Almogy, A. Yariv, *J. Nonlinear Opt. Phys. Mater.* **1995**, *04*, 401.

[59] J. Lee, M. Tymchenko, C. Argyropoulos, P.-Y. Chen, F. Lu, F. Demmerle, G. Boehm, M.-C. Amann, A. Alù, M. A. Belkin, *Nature* **2014**, *511*, 65.

[60] N. Nookala, J. Lee, M. Tymchenko, J. S. Gomez-Diaz, F. Demmerle, G. Boehm, K. Lai, G. Shvets, M.-C. Amann, A. Alu, M. Belkin, *Optica* **2016**, *3*, 283.

[61] G. Sun, J. B. Khurgin, R. A. Soref, *Appl. Phys. Lett.* **2005**, *87*, 201108.

[62] W. S. Rabinovich, G. Beadie, D. S. Katzer, *IEEE J. Quantum Electron.* **1998**, *34*, 975.

[63] M. Clerici, N. Kinsey, C. DeVault, J. Kim, E. G. Carnemolla, L. Caspani, A. Shaltout, D. Faccio, V. Shalaev, A. Boltasseva, M. Ferrera, *Nat. Commun.* **2017**, *8*, 15829.

[64] N. Kinsey, C. DeVault, A. Boltasseva, V. M. Shalaev, *Nat. Rev. Mater.* **2019**, *4*, 742.

[65] O. Reshef, I. De Leon, M. Z. Alam, R. W. Boyd, *Nat. Rev. Mater.* **2019**, *4*, 535.

2000291 (11 of 11)