TUTORIAL REVIEW

Kramers–Krönig relations in nonlinear optics

D. C. HUTCHINGS, M. SHEIK-BAHAE, D. J. HAGAN*,
E. W. VAN STRYLAND†
*Center for Research in Electro-Optics and Lasers (CREOL),
University of Central Florida, Orlando, Florida 32826, USA*

We review dispersion relations, which relate the real part of the optical susceptibility (refraction) to the imaginary part (absorption). We derive and discuss these relations as applied to nonlinear optical systems. It is shown that in the nonlinear case, for self-action effects the correct form for such dispersion relations is nondegenerate, i.e. it is necessary to use multiple frequency arguments. Nonlinear dispersion relations have been shown to be very useful as they usually only require integration over a limited frequency range (corresponding to frequencies at which the absorption changes), unlike the conventional linear Kramers–Krönig relation which requires integration over all absorbing frequencies. Furthermore, calculation of refractive index changes using dispersion relations is easier than a direct calculation of the susceptibility, as transition rates (which give absorption coefficients) are, in general, far easier to calculate than the expectation value of the optical polarization. Both resonant (generation of some excitation that is long lived compared with an optical period) and nonresonant 'instantaneous' optical nonlinearities are discussed, and it is shown that the nonlinear dispersion relation has a common form and can be understood in terms of the linear Kramers–Krönig relation applied to a new system consisting of the material plus some 'perturbation'. We present several examples of the form of this external perturbation, which can be viewed as the pump in a pump–probe experiment. We discuss the two-level saturated atom model and bandfilling in semiconductors among others for the resonant case. For the nonresonant case some recent work is included where the electronic nonlinear refractive coefficient, n_2, is determined from the nonlinear absorption processes of two-photon absorption, Raman transitions and the a.c. Stark effect. We also review how the dispersion relations can be extended to give alternative forms for frequency summation which, for example, allows the real and imaginary parts of $\chi^{(2)}$ to be related.

1. Introduction

In this review we examine the application of causality to obtain dispersion relations for non linear optical properties. Linear dispersion relations were first derived for X-rays in the

*Also with the Department of Physics.
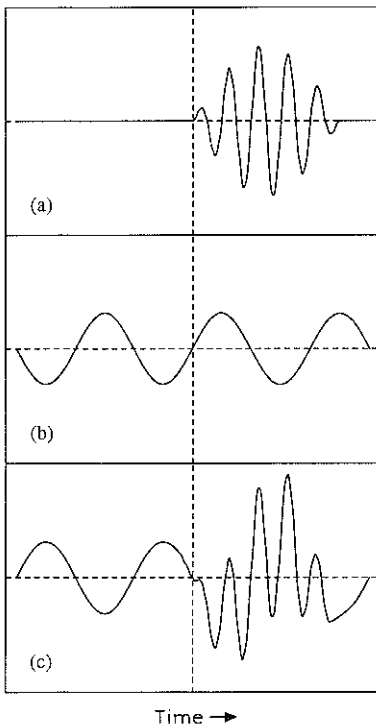†Also with the Departments of Physics and Electrical Engineering.

Figure 1 Illustration of the connection between causality and dispersion relations. An input (a), which is zero for times $t < 0$, is formed out of a superposition of many Fourier components, each of which will extend from $-\infty < t < \infty$. Now consider a perfect filter that absorbs just one of these Fourier components (b). Subtracting this from the input signal gives the output (c). However, the result violates causality, as now there is an output signal for $t < 0$ before any input. In order that this does not occur, the absorption of one Fourier component is accompanied by a phase shift in all of the remaining components in such a way that, upon summation, this also gives zero output for $t < 0$. These required phase shifts are prescribed by the dispersion relation for the system. (From [9].)

mid 1920s by Kramers [1] and Krönig [2] by considering scattering from a continuum of resonators. Krönig subsequently went on to prove the equivalence of causality and dispersion, showing that the dispersion relation is the necessary, as well as sufficient, condition for strict causality to be satisfied [3]. It was thus recognized that the same dispersion relations can be applied to optics in general. Causality refers to the statement that no output can occur before the input. Relativity extends this basic statement further since no signal can propagate faster than the speed of light in vacuum, c. In fact, similar dispersion relations exist for a number of different systems, for example electrical networks [4] and particle scattering [5–7]. The mathematical forms of these dispersion relations are known collectively as Hilbert transforms [8].

An interesting way of viewing the necessity of dispersion relations was given by Toll [9] as shown in Fig. 1. A wave train (a), consisting of a superposition of many frequencies, arrives at a medium that acts as a perfect filter. Thus, one frequency component (b) is completely absorbed and naively the output should be given by the difference between these (c). However, it can be seen that such an output would violate causality, with an output signal occurring at times before the incident wave train arrives. In order for causality to be satisfied, the absorption of one frequency component must be accompanied by a phase shift in all of the remaining components in such a fashion that, when the components are summed, zero output results for times before the arrival of the wave train.

The most frequent expression of the dispersion relation in optics relates the refractive index, n, to the absorption coefficient, α, over all frequencies, Ω

$$n(\omega) - 1 = \frac{c}{\pi} \mathscr{P} \int_0^\infty \frac{\alpha(\Omega)\,d\Omega}{\Omega^2 - \omega^2} \tag{1}$$

2

where $\mathscr{P}$ denotes the Cauchy principal value. It is this relation that is most commonly referred to when one speaks of Kramers–Krönig relations in optics and is the original form of the relation as given in [1, 2] and derived in Section 2.

Although dispersion relations for linear optics are well understood and documented, confusion has existed about applications to nonlinear optics. Clearly causality holds for nonlinear as well as linear systems. The question is, what form do the resulting dispersion relations take? For self-action nonlinear optical effects a nondegenerate form of dispersion relation is appropriate, where, for example, nonlinear refraction of two frequency arguments (the index change at ω due to the presence of a strong perturbing field at Ω) is related to an integral over ω' of the nonlinear absorption at ω' due to the presence of the same perturbing field at Ω. Thus, both the nonlinear absorption and the nonlinear refraction are equivalent to pump–probe spectra with a fixed pump frequency and variable probe frequency.

A first glance at the nonlinear optical susceptibility for a two-level saturated atom [10, 11] could lead one to the result that disperson relations are not valid for this nonlinear system. However, in the field of solid-state physics, dispersion relations have been used for resonant optical nonlinearities (where some real excitation is created within the material such as free carriers) (see, for example [12]). Furthermore, dispersion relations exist for harmonic generation where the real and imaginary parts of $\chi^{(2)}$ are related (see, for example [13]). It is the aim of this review to show how these different systems are related and to derive more-general dispersion relations. It is also indicated under precisely what conditions these dispersion relations can be used.

The result is that we are able to derive and utilize nonlinear optical Kramers–Krönig relations and apply them to various types of nonlinearities, resonant as well as nonresonant. We present below several extremely useful examples, some of which have only recently been recognized. For example, we show how the electronic Kerr effect and two-photon absorption are related by causality [14, 15].

This review is set out as follows. In Section 2 the linear Kramers–Krönig relations are derived from causality. These are extended to nonlinear optics in Section 3, and the nondegenerate dispersion relation is derived for nonlinear optics. In Section 4 it is shown how dispersion relations can be used in resonant nonlinear optics both for solid-state systems and for atomic systems. In Section 5 we give an example calculation for refractive index changes resulting from the Franz–Keldysh effect (d.c. field) in semiconductors. In Section 6 we consider dispersion relations for nonresonant (a.c. field) optical nonlinearities, and show as an example a calculation of the electronic Kerr effect in solids. We also extend the dispersion relations to derive a degenerate form that is applicable in some special cases such as harmonic generation.

This review uses Gaussian (CGS) units throughout. Appendices are also provided which introduce some of the nonlinear optics terminology, equivalent SI expressions, and some of the more detailed mathematics.

2. Linear Kramers–Krönig relations
In a dielectric medium the optical polarization, $P(t)$ can be obtained from the electric field, $E(t)$, by means of a response function

$$P(t) \;=\; \int_{-\infty}^{\infty} R(\tau)\,E(t\,-\,\tau)\,\mathrm{d}\tau \tag{2}$$

The response function, $R(\tau)$, is equivalent to a Green's function, as it gives the response

3

(polarization) resulting from a delta function input (electric field). This equation is more often stated in terms of its Fourier transform, where the convolution is transformed into a product

$$P(\omega) = \chi(\omega) E(\omega) \tag{3}$$

where $\chi(\omega)$ is the susceptibility defined in terms of the response function as

$$\chi(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{i\omega\tau} d\tau \tag{4}$$

Note that this is not the true Fourier transform as the factor of $(2\pi)^{-1}$ is omitted.

Causality states that the effect cannot precede the cause. In the above case this requires that $E(t - \tau)$ cannot contribute to $P(t)$ for $t < (t - \tau)$. Therefore, in order to satisfy causality, $R(\tau) = 0$ for $\tau < 0$. An easy way to see this is to consider the response to a delta function $E(T) = E_0 \delta(T)$, where the polarization would then follow $R(t)$. This has important consequences for the relation between the susceptibility, $\chi(\omega)$, and the response function, $R(\tau)$, since the integration need be performed only for positive times. Therefore, the lower limit in the integral in Equation 4 can in general be replaced by zero.

The usual method for deriving the Kramers–Krönig relation from this point is to consider a Cauchy integral in the complex frequency plane. We give the details of this calculation in Appendix A. However, in the Cauchy integral method, the physical principle from which dispersion relations result (namely causality) is not obvious. The principle of causality can be stated mathematically as

$$R(T) = R(T) \theta(T) \tag{5}$$

i.e. the response to an impulse at $t = 0$ must be zero for $t < 0$. Here $\theta(T)$ is the step function defined as $\theta(T) = 1$ for $T > 0$ and $\theta(T) = 0$ for $T < 0$. (It is also possible to use the 'sign' function at this point or any other function that requires $R = 0$ for $T < 0$.) On Fourier transforming this equation, the product in the time domain becomes a convolution in frequency space

$$\begin{aligned} \chi(\omega) &= \chi(\omega) * \left(\frac{\delta(\omega)}{2} + \frac{i}{2\pi\omega} \right) \\ &= \frac{\chi(\omega)}{2} + \frac{i}{2\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\chi(\Omega)}{\omega - \Omega} d\Omega \\ &= \frac{1}{i\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\chi(\Omega)}{\Omega - \omega} d\Omega \end{aligned} \tag{6}$$

which is the Kramers–Krönig relation for the linear optical susceptibility. Thus, the familiar Kramers–Krönig relation is simply a restatement of the causality condition (Equation 5) in the frequency domain.

This relation can also be obtained by starting from Equation 4, multiplying both sides by $(\omega - \Omega)^{-1}$ and integrating over all ω. Using the identity

$$\mathscr{P} \int_{-\infty}^{\infty} \frac{e^{i\omega\tau} d\omega}{\omega - \Omega} = i\pi e^{i\Omega\tau} \tag{7}$$

which is true only for $\tau > 0$ (which is required by causality) and again Equation 6 is obtained.

This form of relation is not confined to the field of optics only, but is a general property of Fourier transforms known as Hilbert transforms [8].

It is more usual to write the optical dispersion relations in terms of the more familiar quantities; the refractive index, $n(\omega)$, and absorption coefficient, $\alpha(\omega)$ [5]. In order to derive this, consider the propagation of a monochromatic plane wave through a thin slice of a dielectric medium of (complex) refractive index

$$\eta(\omega) = n(\omega) + i\alpha(\omega)c/2\omega \tag{8}$$

and thickness δz

$$E(z + \delta z, \omega) = \exp\left(\frac{i\eta(\omega)\omega}{c}\,\delta z\right) E(z, \omega) \tag{9}$$

This is essentially in the same form as Equation 3; that is, a product in frequency space that can be Fourier transformed into a convolution in time

$$E(z + \delta z, t) = \int_{-\infty}^{\infty} g(\delta z, \tau)\, E(z, t - \tau)\, d\tau \tag{10}$$

where the response function $g(\delta z, \tau)$ is defined

$$\exp\left(\frac{i\eta(\omega)\omega}{c}\,\delta z\right) = \int_{-\infty}^{\infty} g(\delta z, \tau)\, e^{i\omega\tau}\, d\tau \tag{11}$$

Now relativistic causality states that no signal can propagate faster than c, the speed of light in vacuum. This then requires that $g(\delta z, \tau) = 0$ for $\tau < \delta z/c$. Thus, the lower limit in Equation 11 can be replaced with $\delta z/c$. By substituting $T = \tau - \delta z/c$ we can write

$$\exp\left(\frac{i\omega}{c}[\eta(\omega) - 1]\delta z\right) = \int_0^{\infty} g\left(\delta z, T + \frac{\delta z}{c}\right) e^{i\omega T}\, dT \tag{12}$$

In a similar manner to the case for the susceptibility, $\chi(\omega)$, the response function, g (time domain), need only be integrated over positive times T in order to calculate its Fourier transform. In terms of complex variables (see Appendix A), this left-hand side of Equation 12 is regular and analytic in the positive imaginary frequency half-plane. This then requires that $\omega[\eta(\omega) - 1]$ can also be defined as an integral over positive times only (i.e. $\omega[\eta(\omega) - 1]$ is also regular and analytic in the upper frequency half-plane) and dispersion relations can be applied. This can be more clearly seen by taking the limit of a thin slice ($\delta z \to 0$) where the exponential can be expanded to a first-order Maclaurin series and we can write

$$\omega[\eta(\omega) - 1] = \int_0^{\infty} g^\dagger(T)\, e^{i\omega T}\, dT \tag{13}$$

where $g^\dagger(T)$ is the Fourier transform of the left-hand side of Equation 13, defined from Equation 12 as

$$g^\dagger(T) = -ic\left(\lim_{\delta z \to 0} \frac{g(\delta z, T + \delta z/c) - \delta(T)}{\delta z}\right) \tag{14}$$

The delta function is the Fourier transform of the zeroth-order term in the expansion of the exponential. The precise form of $g^\dagger(T)$ is unimportant, and the result we require is that $\omega[\eta(\omega) - 1]$ can be determined by a Fourier transform over positive times only.

We can also obtain Equation 13 by considering the difference in the output field between the cases where the medium is present and absent (vacuum), again in the limit of a

vanishingly thin slice. Defining $E_{net} = (E_{medium} - E_{vacuum})$, we can write

$$E_{net}(z + \delta z, \omega) = (e^{i\eta(\omega)\omega\delta z/c} - e^{i\omega\delta z/c})E(z, \omega)$$

$$\approx \{i\omega[\eta(\omega) - 1]\delta z/c\}E(z, \omega) \tag{15}$$

and again causality can be applied and dispersion relations obtained.

As we are able to set $g^\dagger(T) = 0$ for $T < 0$ (as $g^\dagger(T)$ is required only for $T > 0$), the same analysis as was done on the response function, $R(t)$, can be performed, conditional on the refractive index and absorption coefficient being well behaved in the limit $\omega \to \pm\infty$, satisfying (see Appendix A)

$$\lim_{\omega \to \pm\infty} \{[\eta(\omega) - 1]/\omega\} \to 0 \tag{16}$$

which leads to the dispersion relation

$$\omega[\eta(\omega) - 1] = \frac{1}{i\pi}\int_{-\infty}^{\infty} \frac{\Omega[\eta(\Omega) - 1]}{\Omega - \omega}d\Omega \tag{17}$$

Note that there is no restriction on the comparitive magnitudes of the real and imaginary parts of $\eta(\omega)$.

It is usual to transform the negative frequency region into positive frequencies in the dispersion relations beween n and α. In this we use the reality conditions $n(-\omega) = n(\omega)$ and $\alpha(-\omega) = \alpha(\omega)$, which are a direct consequence of demanding that the observables $E(t)$ and $P(t)$ be real. The following relation results by equating the real parts of equation 17:

$$n(\omega) - 1 = \frac{c}{\pi}\mathscr{P}\int_0^\infty \frac{\alpha(\Omega)\,d\Omega}{\Omega^2 - \omega^2} \tag{18}$$

This relation was the original form of the dispersion relation as written down by Kramers [1] and Krönig [2]. A corresponding relation can be derived relating the absorption to the refractive index over all frequencies by equating the imaginary parts of Equation 17, but this is seldom used as absorption coefficients are, in general, easier to obtain than refractive indices, both experimentally and theoretically.

3. Nonlinear Kramers–Krönig formalism

The Kramers–Krönig relation can be used to calculate the change in refractive index from the change in absorption due to some external perturbation. The linear Kramers–Krönig relation can be applied both in the presence and in the absence of the perturbation, and the difference is taken between the two cases. Doing this, we can write down a modified form of Kramers–Krönig relation (which we also derive below specifically for an optical perturbation)

$$\Delta n(\omega; \zeta) = \frac{c}{\pi}\mathscr{P}\int_0^\infty \frac{\Delta\alpha(\omega'; \zeta)}{\omega'^2 - \omega^2}\,d\omega' \tag{19}$$

where ζ denotes the perturbation. An equivalent relation also exists whereby the change in absorption coefficient can be calculated from the change in the refractive index, but this is rarely used for the reasons described below. Note that it is essential that the perturbation be independent of the frequency of observation, ω', in the integral (i.e. the excitation must be held constant).

The main reason for calculating the refractive index change this way rather than directly
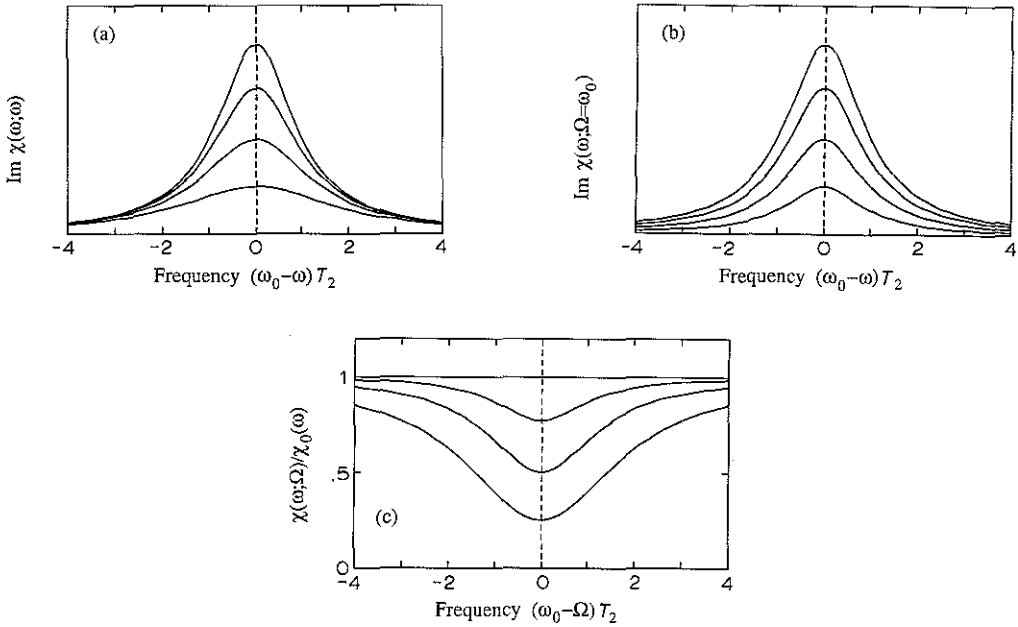
6

Figure 2 (a) Imaginary part of the degenerate optical susceptibility for a two-level atom versus detuning for various irradiance levels I/I_s = 0, 0.3, 1 and 3. As the irradiance increases, the oscillator strength of the transition is reduced in such a way that the width of the absorption line is broadened, which is often termed a 'power-broadened Lorentzian'. (b) Imaginary part of the nondegenerate optical susceptibility $\chi(\omega; \Omega)$ for a two-level atom versus probe detuning at the same irradiance levels as in (a) for the pump. Note that the Lorentzian lineshape and width is preserved at high pump irradiances and only the overall magnitude decreases, unlike the degenerate form shown in (a). This is also an indication that the Kramers–Krönig relation can be applied in the nondegenerate case. (c) Ratio of the nondegenerate optical susceptibility of a two-level atom (i.e. effect on a weak probe at ω from a strong pump at Ω) to the linear susceptibility as a function of pump detuning for the same irradiance levels as above. This quantity also gives the population difference between the ground and excited states (see text).

calculating the expectation value of the real part of the polarization (see, for example [10]) is that calculating the absorption via transition rates is, in general, far easier. Furthermore, this form of calculation of the refractive index for nonlinear optics is often more useful than for linear optics as absorption changes (which can be either calculated or measured) usually occur only over a limited frequency range and, thus, the integral in Equation 19 need be calculated only over this finite frequency range. In comparison, for the linear Kramers–Krönig calculation, absoption spectra tend to cover a very large frequency range and it is necessary to take account of this full range in order to obtain a quantitative fit for the dispersion, although a qualitative fit to the dispersion can often be obtained using a limited frequency range. Unfortunately, the converse is not true as refractive index changes are usually quite extensive in frequency, so a calculation of absorption changes from refractive index changes is seldom performed.

It is possible to use Equation 19 in nonlinear optics under resonant conditions, where the material is excited into some real state. This excitation can be treated as a perturbation itself. Essentially, the change in refractive index is calculated from the change in absorption,
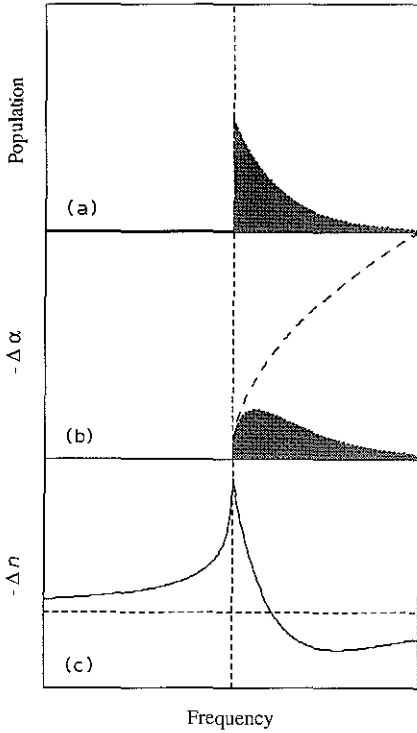
Figure 3 Diagrammatic representation of the blocking model for semiconductor optical nonlinearities. (a) The quasi-equilibrium electron (hole) population generated at the bottom (top) of the conduction (valence) band. When multiplied by the density of states, this gives the reduction in available states, and thus the reduction in absorption coefficient (b). Performing the Kramers–Krönig transform on this absorption change gives the reduction in refractive index (c).

treating this perturbation as constant, and only afterwards is the optical source of the perturbation considered.

This form of calculation will be applied to the two-level atom model. We also give three specific well-known examples for semiconductors: carrier-induced nonlinearities, where the absorption of light excites carriers in the material, which then cause changes in the optical
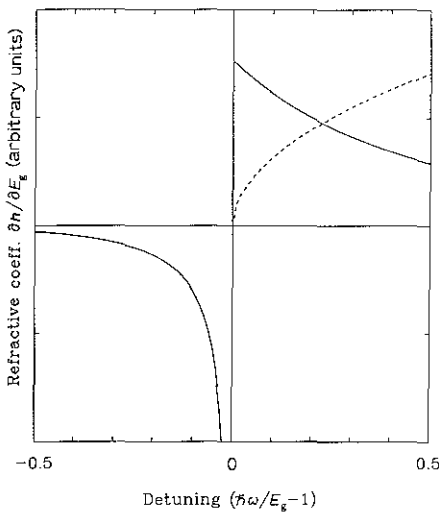

Figure 4 Rate of change of the refractive index with bandgap for an idealized square-root absorption edge as a function of detuning. This not only gives the contribution to the thermo-optic coefficient, $\partial n/\partial T$, but can be used for any mechanism that causes a change in the bandgap.
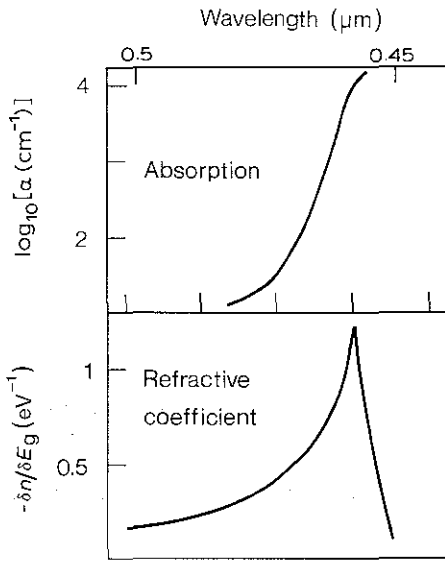
Figure 5 Calculation of the thermo-optic coefficient in a similar manner to Fig. 4, but now using an empirical form of the absorption edge consisting of an exponential tail with a square-root form at higher frequencies. This calculation was performed for thermally deposited ZnSe thin films. (From [38].)

properties; thermally induced nonlinearities, where the absorption of light raises the temperature, which then affects the optical properties; and field-induced nonlinearities, where the creation of carriers causes a change in the d.c. electric field in the material. Note that all three of these mechanisms will take some time to disappear after the excitation, either by carrier recombination or by diffusion, and they cannot be accurately described using a $\chi^{(3)}$ analysis.

The modified relation given by Equation 19 is an extension of the linear relation, and it has not yet been justified for nonresonant nonlinearities where the intermediate (perturbed) state is no longer well defined. Referring to Equation B1 of Appendix B, we now consider causality for the nonlinear contributions to the polarization. The notation used in this section for the nonlinear susceptibility, and its relation to other common descriptions in nonlinear optics, is discussed in Appendix B. Causality requires that no contribution can be made to the nth order polarization, $P^{(n)}(t)$, due to an electric field $E(t - \tau)$ for times $t < (t - \tau)$. This then requires that the response function, $R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)$, must be zero if any one of its arguments $(\tau_1, \tau_2, \ldots, \tau_n)$ is less than zero. Hence the nth order
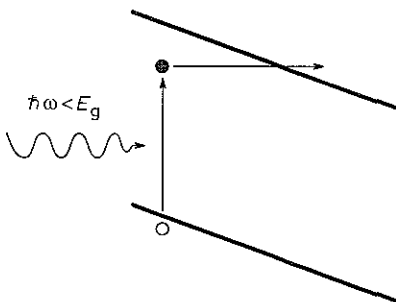

Figure 6 Schematic of the Franz–Keldysh effect where the presence of electric field and tunnelling can permit absorption at frequencies less than the band edge of semiconductors.
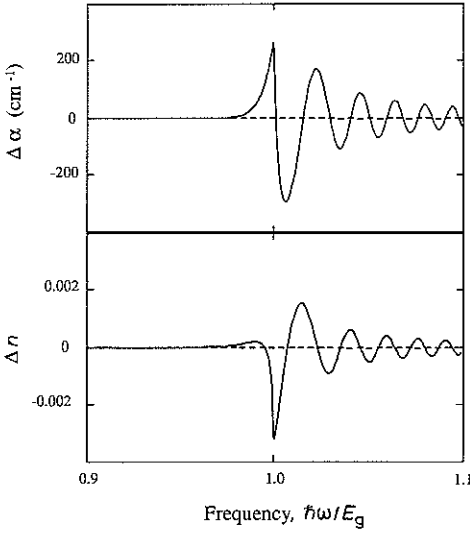
9

Figure 7 (a) Calculation of the change in the absorption coefficient of GaAs due to the Franz–Keldysh effect in an electric field $F = 30\,\text{kV}\,\text{cm}^{-1}$. (b) Kramers–Krönig transform of (a) to give the accompanying change in refractive index in a similar manner to [44].

susceptibility can be determined by integration over positive times only

$$\chi^{(n)}(\omega_1, \omega_2, \ldots, \omega_n) = \int_0^\infty d\tau_1 \int_0^\infty d\tau_2 \cdots \int_0^\infty d\tau_n \, R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) \, e^{i(\omega_1\tau_1 + \omega_2\tau_2 + \cdots + \omega_n\tau_n)} \tag{20}$$

It is now possible to use any of the methods that were used earlier for the linear susceptibility in order to derive a dispersion relation for the nonlinear susceptibility. For example, we can write

$$R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) = R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)\, \theta(\tau_i) \tag{21}$$

and then calculate the Fourier transform of this equation. Here i can apply to any one of the indices $1, 2, \ldots, n$. We could also use any number and combination of step functions; however, the simplest result is obtained by taking just one.

Alternatively, in terms of complex variables (Appendix A) the nonlinear susceptibility is regular and analytic in the upper half-plane of any one of the frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$.

TABLE I Frequency dependence of the nondegenerate absorption $F_2(\hbar\omega_1/E_g, \hbar\omega_2/E_g)$ as defined in Equation 34

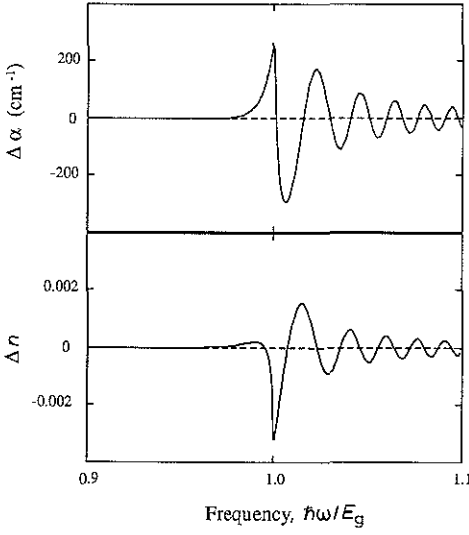| Contribution	$F_2(x_1, x_2)$
Two-photon absorption	$\dfrac{(x_1 + x_2 - 1)^{3/2}}{2^7 x_1 x_2^2}\left(\dfrac{1}{x_1} + \dfrac{1}{x_2}\right)^2$
Raman	$\dfrac{(x_1 - x_2 - 1)^{3/2}}{2^7 x_1 x_2^2}\left(\dfrac{1}{x_1} - \dfrac{1}{x_2}\right)^2$
Linear Stark	$-\dfrac{(x_1 - 1)^{3/2}}{2^6 x_1 x_2^2}\dfrac{1}{x_2^2}$
Quadratic Stark	$-\dfrac{1}{2^{10} x_1 x_2^2 (x_1 - 1)^{1/2}}\left(\dfrac{1}{x_1 - x_2} + \dfrac{1}{x_1 + x_2}\right)$

Figure 7 (a) Calculation of the change in the absorption coefficient of GaAs due to the Franz–Keldysh effect in an electric field $F = 30 \, \text{kV cm}^{-1}$. (b) Kramers–Krönig transform of (a) to give the accompanying change in refractive index in a similar manner to [44].

susceptibility can be determined by integration over positive times only

$$\chi^{(n)}(\omega_1, \omega_2, \ldots, \omega_n) = \int_0^\infty d\tau_1 \int_0^\infty d\tau_2 \cdots \int_0^\infty d\tau_n \, R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) \, e^{i(\omega_1 \tau_1 + \omega_2 \tau_2 + \cdots + \omega_n \tau_n)} \tag{20}$$

It is now possible to use any of the methods that were used earlier for the linear susceptibility in order to derive a dispersion relation for the nonlinear susceptibility. For example, we can write

$$R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) = R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) \, \theta(\tau_i) \tag{21}$$

and then calculate the Fourier transform of this equation. Here i can apply to any one of the indices $1, 2, \ldots, n$. We could also use any number and combination of step functions; however, the simplest result is obtained by taking just one.

Alternatively, in terms of complex variables (Appendix A) the nonlinear susceptibility is regular and analytic in the upper half-plane of any one of the frequencies $(\omega_1, \omega_2, \ldots, \omega_n)$.

TABLE I Frequency dependence of the nondegenerate absorption $F_2(\hbar\omega_1/E_g, \hbar\omega_2/E_g)$ as defined in Equation 34

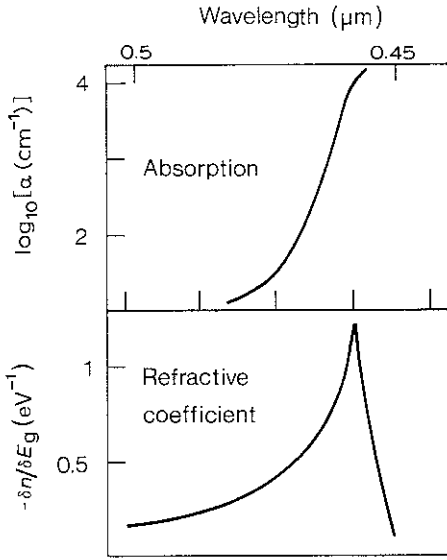| Contribution	$F_2(x_1, x_2)$
Two-photon absorption	$\dfrac{(x_1 + x_2 - 1)^{3/2}}{2^7 x_1 x_2^2}\left(\dfrac{1}{x_1} + \dfrac{1}{x_2}\right)^2$
Raman	$\dfrac{(x_1 - x_2 - 1)^{3/2}}{2^7 x_1 x_2^2}\left(\dfrac{1}{x_1} - \dfrac{1}{x_2}\right)^2$
Linear Stark	$-\dfrac{(x_1 - 1)^{3/2}}{2^6 x_1 x_2^2}\dfrac{1}{x_2^2}$
Quadratic Stark	$-\dfrac{1}{2^{10} x_1 x_2^2 (x_1 - 1)^{1/2}}\left(\dfrac{1}{x_1 - x_2} + \dfrac{1}{x_1 + x_2}\right)$

10

Figure 5 Calculation of the thermo-optic coefficient in a similar manner to Fig. 4, but now using an empirical form of the absorption edge consisting of an exponential tail with a square-root form at higher frequencies. This calculation was performed for thermally deposited ZnSe thin films. (From [38].)

properties; thermally induced nonlinearities, where the absorption of light raises the temperature, which then affects the optical properties; and field-induced nonlinearities, where the creation of carriers causes a change in the d.c. electric field in the material. Note that all three of these mechanisms will take some time to disappear after the excitation, either by carrier recombination or by diffusion, and they cannot be accurately described using a $\chi^{(3)}$ analysis.

The modified relation given by Equation 19 is an extension of the linear relation, and it has not yet been justified for nonresonant nonlinearities where the intermediate (perturbed) state is no longer well defined. Referring to Equation B1 of Appendix B, we now consider causality for the nonlinear contributions to the polarization. The notation used in this section for the nonlinear susceptibility, and its relation to other common descriptions in nonlinear optics, is discussed in Appendix B. Causality requires that no contribution can be made to the nth order polarization, $P^{(n)}(t)$, due to an electric field $E(t - \tau)$ for times $t < (t - \tau)$. This then requires that the response function, $R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)$, must be zero if any one of its arguments $(\tau_1, \tau_2, \ldots, \tau_n)$ is less than zero. Hence the nth order
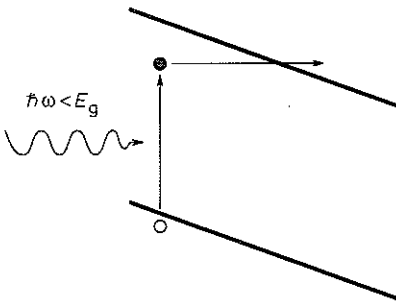

Figure 6 Schematic of the Franz–Keldysh effect where the presence of electric field and tunnelling can permit absorption at frequencies less than the band edge of semiconductors.
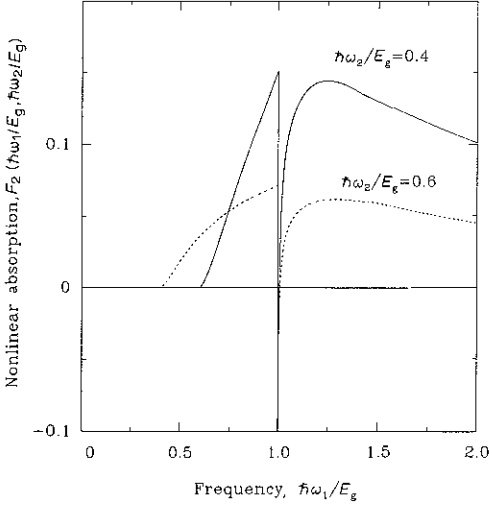
9

Figure 8 Theoretical frequency dependence of the nondegenerate nonlinear absorption for two different 'pump' frequencies $\hbar\omega_2 = E_g = 0.4$ and 0.6. Below the fundamental absorption edge $\hbar\omega_1 < E_g$ only two-photon absorption contributes to the nonlinear absorption.

We can thus obtain the generalized nonlinear Kramers–Krönig relation for a nondegenerate nonlinear susceptibility

$$\chi^{(n)}(\omega_1, \omega_2, \ldots, \omega_n) = \frac{1}{i\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\chi^{(n)}(\omega_1, \omega_2, \ldots, \Omega, \ldots, \omega_n)}{\Omega - \omega_i} \, d\Omega \qquad (22)$$

Note here that the integral is over only one frequency argument, Ω, and all other frequencies are held constant. Thus, we cannot obtain any relationship between the degenerate Kerr coefficient, $\gamma(\omega)$, and the degenerate two-photon absorption coefficient, $\beta(\omega)$, as defined in Appendix B. If multiple step functions had been used at an earlier stage in Equation 21, a multidimensional integral would have resulted. The derivation of the nonlinear Kramers–Krönig relationship in Equation 22 can also be found in [16–18].

Now consider the case where we have two monochromatic electromagnetic waves incident on a nonlinear material

$$E(t) = \tfrac{1}{2}(E_a \, e^{-i\omega_a t} + E_a^* \, e^{i\omega_a t}) + \tfrac{1}{2}(E_b \, e^{-i\omega_b t} + E_b^* \, e^{i\omega_b t}) \qquad (23)$$

On calculating the third-order nonlinear polarization (see Appendix B) and considering

TABLE II Dispersion of the nonlinear refraction $G_2(\hbar\omega/E_g)$ for frequencies below the band edge as defined in Equation 38. $\Theta(x)$ is the Heaviside or step function

Contribution	$G_2(x)$
Two-photon absorption	$[1/(2x)^6][-\tfrac{3}{8}x^2(1-x)^{-1/2} + 3x(1-x)^{1/2} - 2(1-x)^{3/2} + 2\Theta(1-2x)(1-2x)^{3/2}]$
Raman	$[1/(2x)^6][-\tfrac{3}{8}x^2(1+x)^{-1/2} - 3x(1+x)^{1/2} - 2(1+x)^{3/2} + 2(1+2x)^{3/2}]$
Linear Stark	$[1/(2x)^6][2 - (1-x)^{3/2} - (1+x)^{3/2}]$
Quadratic Stark	$[1/(2^{10}x^5)][(1-x)^{-1/2} - (1+x)^{-1/2} - \tfrac{1}{2}x(1-x)^{3/2} - \tfrac{1}{2}x(1+x)^{-3/2}]$
Divergent term	$\dfrac{1}{(2x)^6}\left(-2 - \dfrac{35x^2}{8} + \dfrac{x}{8}(3x-1)(1-x)^{-1/2} - 3x(1-x)^{1/2} + (1-x)^{3/2}\right.$ $\left. + \dfrac{x}{8}(3x+1)(1+x)^{-1/2} + 3x(1+x)^{1/2} + (1+x)^{3/2}\right)$

11
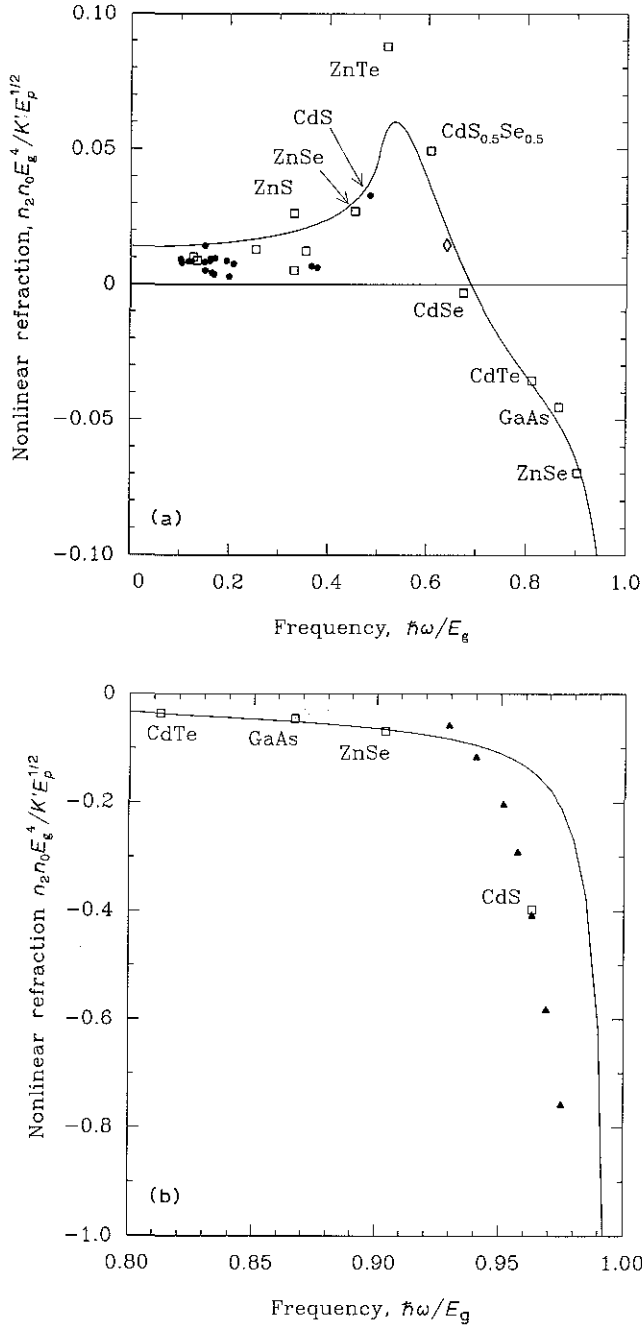
Figure 9 (a) Dispersion of the nonlinear refractive index n_2. Data for n_2 are scaled as $n_2 n_0 E_g^4 / K' E_p^{1/2}$. The circles are measurements in [58], the diamond is from [59] and the squares our own Z-scan measurements. We have labelled the semiconductor data. The full line is the $G_2 (\hbar\omega/E_g)$ function derived here for a two-band model of a semiconductor. (b) Extension of (a) to frequencies close to the band edge. The triangles are n_2 measurements of AlGaAs from [60].

Figure 10 A log–log plot showing the expected E_g^{-4} dependence of n_2. The data points are identical to those in Fig. 9a and b, but now scaled by the dispersion function $G_2(\hbar\omega/E_g)$. The full line is the function E_g^{-4} that translates into a straight line of slope -4 on a log–log plot.

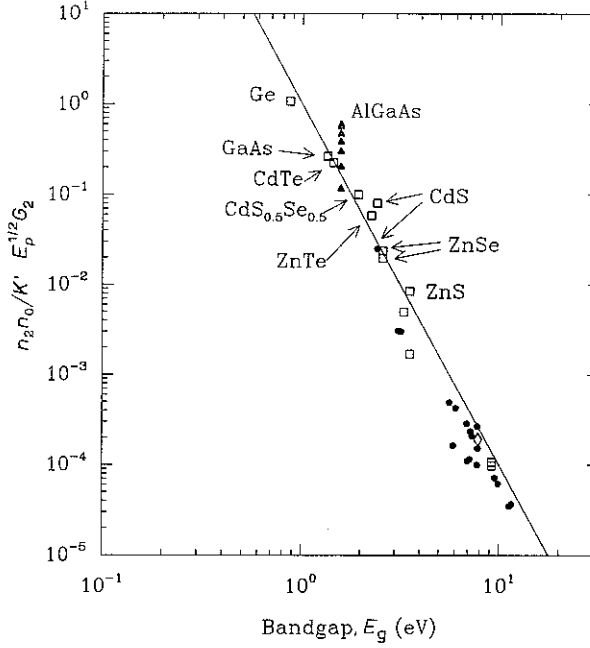only the mixed terms that give rise to contributions at frequencies $\pm\omega_a$ (i.e. ignoring all third harmonic and frequency mixing terms, such as $2\omega_b - \omega_a$)

$$P_a^{(3)}(\omega) = \tfrac{3}{4}|E_b|^2[\chi_{aabb}^{(3)}(\omega_a, \omega_b, -\omega_b)\delta(\omega - \omega_a)E_a + \chi_{aabb}^{(3)*}(\omega_a, \omega_b, -\omega_b)\delta(\omega + \omega_a)E_a^*] \tag{24}$$

These then lead to a change in the refractive index and absorption coefficient at frequency ω_a due to the presence of light at frequency, ω_b

$$\Delta n(\omega_a; \omega_b) = (24\pi^2/n_a n_b c)I_b \operatorname{Re}\chi_{aabb}^{(3)}(\omega_a, \omega_b, -\omega_b)$$

$$\Delta\alpha(\omega_a; \omega_b) = (48\pi^2\omega_a/n_a n_b c^2)I_b \operatorname{Im}\chi_{aabb}^{(3)}(\omega_a, \omega_b, -\omega_b) \tag{25}$$

where I_b is the irradiance of the light of frequency ω_b, and n_a and n_b are the linear refractive indices at frequencies ω_a and ω_b, respectively, and we have assumed that the linear absorption is sufficiently small that $\alpha(\omega_a)c/\omega_a \ll n_a$. This is not very restrictive, allowing validity even for $\alpha \approx 10^3\,\text{cm}^{-1}$ in the visible, and in most cases is a reasonable approximation. For example, in a semiconductor, changes in the refractive index beneath the bandgap are attributable to absorption changes at frequencies close to the band edge and not the high absorption transitions at higher frequencies. This will be more apparent in later sections. Note that these results in the nondegenerate case are a factor of 2 greater than the corresponding degenerate quantities given in Equation B14 of Appendix B.

Unfortunately, there is no universal convention for the definition of $\chi^{(3)}$, so some of the numerical prefactors may differ from those used in other papers. We discuss other common definitions in Appendix B.

Upon inserting these relations between the nondegenerate $\chi^{(3)}$ and the change in refraction and absorption into the nondegenerate Kramers–Krönig relation of Equation 22 and transforming to only positive frequencies in manner similar to Equation 18, one can obtain

$$\Delta n(\omega_a; \omega_b) = \frac{c}{\pi} \mathscr{P} \int_0^\infty \frac{\Delta \alpha(\omega; \omega_b)\, d\omega}{\omega^2 - \omega_a^2} \tag{26}$$

We note the similarities between Equations 26 and 19. This is not surprising, as we could consider the external perturbation in Equation 19 to be an electromagnetic field itself. Note that this perturbation should be constant on both sides of the relation and must not vary as the integral over ω is computed. Thus, a linear Kramers–Krönig calculation is being performed on a new system consisting of the material plus an electromagnetic field of fixed frequency and irradiance. This is a significant step, as we have extended the use of the modified linear relation as given in Equation 19 for resonant optical nonlinearities to the nonresonant case where there is no real excited intermediate state. However, by viewing the nonlinear optical process in terms of virtual excited states, this step is intuitively obvious.

These nondegenerate forms of the nonlinear refraction and absorption are equivalent to pump–probe spectra, in that $\Delta n(\omega_a; \omega_b)$ and $\Delta \alpha(\omega_a; \omega_b)$ describe the change in refractive index and absorption coefficient, respectively, for a weak optical probe of frequency ω_a when a strong pump of fixed frequency ω_b is applied.

4. Kramers–Krönig relations for resonant optical nonlinearities: examples
4.1. Two-level atom

A first glance at the result for the nonlinear optical susceptibility for a saturated two-level atom could lead to the assumption that nonlinear Kramers–Krönig relations cannot be used in such a system [10, 11]. The standard density matrix calculation [10, 19] results in a nonlinear optical susceptibility of the form

$$\chi(\omega) = \frac{\mu^2 \Delta N_0 (\omega_0 - \omega) + i/T_2}{\varepsilon_0 \hbar (\omega - \omega_0)^2 + (1 + I/I_s)/T_2^2} \tag{27}$$

where μ is the dipole moment for the transition of frequency ω_0, which has a dephasing time T_2 and a saturation irradiance I_s. ΔN_0 is the ground-state population density. For low excitation ($I \ll I_s$) this takes the form of a Lorentzian lineshape of width T_2^{-1}, but at higher excitation the lineshape is broadened as shown in Fig. 2a ('power-broadened Lorentzian'). Note here that in Appendix B we expand $\chi(\omega)$ as a power series in the electric field, but the present form corresponds to the nonlinear $\chi(\omega)$ before expansion, with the total polarization $P(\omega)$ (linear plus nonlinear) determined through Equation 3. By using a binomial expansion this form can be used to generate, for example, the $\chi^{(3)}$ term.

As Yariv [10] pointed out, this expression does not satisfy the Kramers–Krönig relations except under the trivial condition of zero excitation, $I = 0$. This can also be verified mathematically by considering the poles of $\chi(\omega)$, which in this case occur at

$$\omega = \omega_0 \pm (i/T_2)(1 + I/I_s)^{1/2} \tag{28}$$

Since a pole now exists in the upper half-plane, χ is no longer regular and analytic, and the contour integral around the upper half-plane is no longer zero. This method provides an easy check to see whether dispersion relations are valid (see Appendix A). This result is a consequence of the noncausal nature of the expression for $\chi(\omega)$.

14

It has been suggested that the contribution of this pole can be included in the contour integral, and one could derive Kramers–Krönig type relations where an additional constant is added to the usual integral [11]. However, since this constant is derived from the residue at the pole of $\chi(\omega)$, it is necessary to know fully the form of $\chi(\omega)$, i.e. both the real and imaginary parts. Since the usefulness of Kramers–Krönig relations is that, given only the imaginary part, the real part of $\chi(\omega)$ can be computed (or vice versa), such a relationship is of little practical use.

In order to apply the modified Kramers–Krönig relations as described in Equation 19 to the two-level atom problem, it is necessary to re-examine the source of the perturbation that causes the optical properties to change. In this instance the perturbation is a change in the excited-state population by optical excitation. Note, however, that if the frequency of the light varies, the excited state population varies also, due to the variation of the linear absorption coefficient with frequency. Thus, it should not be too surprising that the Kramers–Krönig relations do not apply to this form of nonlinear susceptibility, since the perturbation is changing over the frequency integral.

The perturbation (excited-state population) can be forced to be constant, however, by replacing the single pump field by a pump field at a fixed frequency plus a probe field (considered weak) whose frequency can vary. Thus, we require a pump–probe spectrum, where the two frequencies can be different. Returning to the two-level atom calculation [10], if the change in population is calculated solely from a pump field at frequency Ω and then the polarization is calculated at a different probe frequency ω, Equation 27 for the susceptibility now becomes

$$
\begin{aligned}
\chi(\omega; \Omega) &= \left(\frac{\mu^2 \Delta N_0}{\varepsilon_0 \hbar}\right)\left(\frac{(\omega_0 - \omega) + i/T_2}{(\omega_0 - \omega)^2 + 1/T_2^2}\right)\left(\frac{(\omega_0 - \Omega)^2 + 1/T_2^2}{(\omega_0 - \Omega)^2 + (1 + I/I_s)/T_2^2}\right) \\
&= \chi_0(\omega)\,\frac{(\omega_0 - \Omega)^2 + 1/T_2^2}{(\omega_0 - \Omega)^2 + (1 + I/I_s)/T_2^2}
\end{aligned} \tag{29}
$$

where I now refers to the pump irradiance. Here population pulsations have been ignored as discussed in the next paragraph. $\chi_0(\omega)$ refers to the susceptibility in the linear (low power, $I \to 0$) limit. We note that this form of nonlinear susceptibility does satisfy the Kramers–Krönig relations where the integral is over the probe frequency ω only. Figure 2b shows the imaginary part of the nondegenerate susceptibility at a fixed pump frequency ($\Omega = \omega_0$) for several pump irradiances. It is of interest to note that, in the nondegenerate form, the Lorentzian lineshape and width are preserved, which is another indication that dispersion relations can be applied. In fact, this susceptibility can be separated into a product of the linear susceptibility, multiplied by some fraction that depends only on the pump field (Ω), as shown in Equation 29. This fraction is precisely the fraction of atoms that occupy the ground state. Figure 2c demonstrates how this fraction varies with pump irradiance. This two-level description carries over directly to the solid state description of bandfilling, where nonlinear Kramers–Krönig relations have been utilized with great success.

In this example calculation we have neglected any effects of population pulsations (alternate time-ordering) caused by the beating of the two optical frequencies (ω and Ω) in the medium. Strictly, this approximation is valid only when the response of the medium is much slower than the beat frequency. Inclusion of population pulsations will lead to an enhancement in the nondegenerate susceptibility, typically by a factor of 2 for the third-order term $\chi^{(3)}$ in the expansion of the susceptibility. The main point of this example,

however, is not invalidated, in that dispersions relations can be applied providing it is in a nondegenerate form.

Although a general nonlinear $\chi(\omega; \Omega)$ cannot be determined when beating effects are included, Boyd and Mukamel [20] included such effects in their calculation for the nondegenerate third-order term $\chi^{(3)}(\omega; \Omega)$. It is interesting to note that their expression for $\chi^{(3)}$ satisfies the nondegenerate dispersion relations using ω but not Ω, which is precisely the result predicted here, indicating that the validity of the nondegenerate dispersion relations is independent of the nonlinear mechanism and the approximations used in its determination.

4.2. Carrier-induced optical nonlinearities in semiconductors

There are many mechanisms whereby the excitation of carriers leads to a change in optical properties. Broadly, these can be split into two categories. First, since electrons and holes are fermions, Pauli exclusion applies and no two identical particles can occupy the same state. Secondly, electrons and holes carry charge, and since upon excitation they are made mobile, they can influence electric fields in the material.

The concept of carriers inhibiting band edge absorption due to Pauli exclusion predates even the laser. Electrons can be placed in the conduction band by doping, and this leads to a blue shift of the band edge [21, 22], which is commonly known as the Burstein–Moss shift.

In the late 1970s very large self-defocusing refractive nonlinearities ($\Delta n < 0$) were observed in InSb [23] and GaAs [24] at wavelengths just beneath the band edge. These refractive nonlinearities were subsequently used to obtain bistability in a Fabry–Perot configuration [25, 26]. The origin of these nonlinearities is that a quasi-equilibrium population of excited carriers is maintained by the optical excitation, which in turn inhibits further transitions. This process has been termed the dynamic Burstein–Moss shift. Accompanying the change in absorption due to these carriers is a change in refractive index.

One model, which accurately accounts for the nonlinearities observed in InSb, is the bandfilling model [12, 27]. In semiconductors, carrier–carrier scattering and electron–phonon interaction tend to thermalize carrier distributions within a band on a timescale of typically < 1 ps. However, recombination of electrons and holes typically occurs on timescales of the order of ns to μs. This means that on the timescales of interest there is a quasi-equilibrium population of electrons at the bottom of the conduction band (and holes at the top of the valence band) which have a thermal distribution. It is important to note that these carriers will have no record of the optical pump field about 1 ps.

The change in absorption coefficient can be obtained using Fermi's Golden Rule, taking population factors into account

$$\Delta\alpha(\omega) = -\frac{4\pi^2 e^2}{n_0 \omega c m_0^2} \sum_k |\hat{a} \cdot p_{vc}(k)|^2 \delta(E_{vc}(k) - \hbar\omega)[f_c(k) + f_v(k)] \tag{30}$$

where $f_c(k)$ and $f_v(k)$ are the population factors, n_0 the background refractive index, m_0 the free electron mass, $\hat{a} \cdot p_{vc}(k)$ the extension of the momentum matrix element along the optical polarization vector (which can be determined for a direct-gap semiconductor using Kane's theory [28]) and $E_{vc}(k)$ the energy separation of the optically coupled states. The sum should be over all optically coupled states, but typically only the lowest conduction band (and highest valence bands) will have a significant population. Provided the quasi-equilibrium populations is not too large, Boltzmann statistics will be adequate to describe the population factors.

This calculation is shown diagrammatically in Fig. 3, in which the reduction in absorption is calculated from the product of the population distribution and the density of states. In other words, the distribution of electrons at the bottom of the conduction band removes the same distribution of states from the potential interband transitions. The modified form of the Kramers–Krönig relation can now be applied to produce the change in refractive index associated with this change in absorption as shown in Fig. 3.

Further examples of this type of calculation have been performed on a theoretical change in absorption due to an excited electron–hole plasma by Bányai and Koch for their plasma model, which includes Coulombic effects for bulk semiconductors [29] and for semiconductor doped glasses [30], and the many-body calculations of Löwenau *et al.* [31].

This form of calculation need not be restricted to theoretical differential spectra. Provided the complete differential absorption spectrum is known (i.e. measured over all frequencies for $\Delta\alpha \neq 0$), the associated refractive index change can be calculated using Equation 19. The converse is also true, but this is unusual as it is far easier to measure absorption spectra (which also tend to be more restricted in frequency) than the dispersion of a material. Examples of this calculation can be found in [32] (bulk GaAs), [33] (saturation of GaAs quantum wells), [34] (quantum-confined Stark effect in GaAs quantum wells) and [35] (semiconductor clusters).

4.3. Thermally induced optical nonlinearities in semiconductors

Changes in the optical properties of a material occur not only for changes in carrier population as described previously, changes of temperature also have an effect. Often this can be a problem for applications involving optical switching; however, thermal effect alone can be used, for example in optical bistability [36, 37].

As an example calculation, consider light incident on a semiconductor. The energy from the illumination usually ends up as heat (for example, light generates carriers which recombine and give up their energy to the lattice). Close to the band edge the optical properties of the semiconductor change, principally through the thermal shift of the bandgap

$$\Delta n = (\partial n/\partial T)\Delta T$$

$$\frac{\partial n}{\partial T} = \frac{\partial n}{\partial E_g}\frac{\partial E_g}{\partial T} + \frac{\partial n}{\partial T}\bigg|_b \tag{31}$$

where we have also allowed for some background contribution to the thermo-optic coefficient, $\partial n/\partial T$, such as thermal expansion or lattice contributions.

The coefficient $\partial n/\partial E_g$ can be obtained from a Kramers–Krönig transform of the band edge absorption [38]

$$\frac{\partial n}{\partial E_g}(\omega) = \lim_{\delta E_g \to 0}\frac{c}{\pi}\mathscr{P}\int_0^\infty \frac{\alpha(\omega'; E_g + \delta E_g) - \alpha(\omega'; E_g)}{\omega'^2 - \omega^2}\,d\omega' \tag{32}$$

Using a square-root absorption edge for a direct-gap material, the dispersion of $\partial n/\partial T$ around the band edge can be determined and is shown plotted in Fig. 4. Note the resonance as the frequency approaches the band edge; in fact, for a sharp square-root absorption this dispersion function is discontinuous at $\hbar\omega = E_g$. Note that since $\partial E_g/\partial T$ is usually negative [39], this means that in most cases $\partial n/\partial T$ is positive beneath the band edge of a semiconductor.

Of course, this form of nonlinear dispersion can be applied to any process that causes a shift in the band edge, e.g. bandgap renormalization. It is also possible to perform the same

17

calculation using a more realistic form for the absorpton coefficient; for example, in [38] the thermo-optic coefficient, $\partial n/\partial T$, is calculated using a semi-empirical form for the band edge absorption which has an exponential Urbach-like tail. The result of this calculation is shown in Fig. 5.

5. Changes in optical properties induced by a d.c. electric field

The optical properties of a semiconductor can also be altered through the applications of an electric field. As was first noted by Franz [40] and Keldysh [41], transitions just beneath the band edge can occur in the presence of an electric field due to tunnelling (Fig. 6). In the weak field limit for a direct gap semiconductor, Tharmalingam [42] gives for the absorption coefficient

$$\alpha(\omega, F) = (C\theta_F^{1/2}/\omega)[|Ai'(\eta)|^2 - \eta|Ai(\eta)|^2] \tag{33}$$

where $\theta_F^3 = e^2 F^2/2m_{vc}\hbar$, $\eta = (E_g - \hbar\omega)/\hbar\theta_F$, F is the d.c. electric field, C a material constant that involves the matrix elements and $Ai(x)$ is the Airy function [43]. Fig. 7a shows the calculated change in absorption coefficient for frequencies around the bandgap calculated for GaAs.

As before, we can use this change in absorption to calculate the change in refractive index via the modified Kramers–Krönig relations [44]. These integrals can be performed numerically (see Appendix D) and the associated refractive index changes associated with the absorption changes in Fig. 7a are shown in Fig. 7b.

The Franz–Keldysh effect has practical uses (for example, as a modulator by using an interferometric configuration) and we discuss it briefly for two reasons. First, there are devices and materials in which the intermediate state, upon application of a strong light field, is a change in the static electric field usually due to the movement of charge [45]. Good examples of this are the photorefractive effect [46] and self-electro-optic devices (SEEDs) [47, 48]. Although not all of these applications use bulk semiconductors, the principle of using the modified Kramers-Krönig relation to calculate the change in refractive index from absorption changes remains the same. Secondly, since light is in itelf an oscillating electric field, the Franz–Keldysh effect can be used to determine the change in refraction and absorption resulting from the interaction of a light field in the low frequency limit [49]. Note that what is obtained from this calculation is the effect on a weak probe at an arbitrary frequency ω due to a strong pump at frequency Ω in the limit $\Omega \to 0$.

6. Kramers–Krönig relations for nonresonant optical nonlinearities (a.c. electric field)

6.1. Nonlinear absorption and refraction in solids

In what follows we consider the electronic Kerr effect in more detail, and in particular its calculation from nonlinear absorption processes. We have already shown how the refractive index change can be calculated using the modified Kramers–Krönig relation from the effect of an electric field on the absorption (Franz–Keldysh effect). This calculation refers to the limit $\omega_b \to 0$ in the above dispersion relations. By also taking the limit $\omega_a \to 0$ this gives the low frequency limit of the electronic Kerr effect. In what follows, this form of calculation is extended to other frequencies.

In a recent paper we used Equation 26 in a theoretical calculation to determine the dispersion and scaling of n_2 in solids [14, 15]. It is necessary to account for all forms of nonlinear absorption in this calculation, and hence we include two-photon and Raman

transitions and the a.c. stark effect. It is important to emphasize again that in this calculation it is the nondegenerate form of the nonlinear absorption that is required. Existing expressions for nonlinear absorption, e.g. multiphoton absorption [41, 50, 51], are derived only for the degenerate form.

The nondegenerate absorption was derived using tunnelling theory based on an $A \cdot p$ perturbation for a two-band model of a semiconductor. The details of the calculation are not relevant to this paper, but the result can be expressed in the scaling form as [15]

$$\Delta\alpha(\omega_1; \omega_2) = 2K \frac{E_p^{1/2}}{n_1 n_2 E_g^3} F_2\left(\frac{\hbar\omega_1}{E_g}; \frac{\hbar\omega_2}{E_g}\right) I_2 \tag{34}$$

Here E_g is the energy gap and E_p is related to the momentum matrix element and is nearly constant at $E_p \approx 21$ eV over the vast majority of semiconductors. The refractive indices n_1 and n_2 refer to the frequencies ω_1 and ω_2, respectively, and I_2 is the irradiance at frequency ω_2. The constant K is material-independent and was determined from the fit to degenerate two-photon absorption measurements [52] which gives $K = 3100 \, \text{cm} \, \text{GW}^{-1} \, \text{eV}^{5/2}$. It should be noted, however, that theoretical calculations of this quantity are not too different from this value. The dimensionless spectral function F_2 can be determined for each of the nonlinear optical processes and is the sum of the terms shown in Table I.

The spectrum of the nonlinear absorption $F_2(\hbar\omega_1/E_g; \hbar\omega_2/E_g)$ is shown in Fig. 8 for two different 'pump' frequencies ω_2. Note that for $\hbar\omega_2 < E_g$ the Raman and Stark terms can contribute only to the nonlinear absorption for $\hbar\omega_1 > E_g$. The negative (i.e. decreasing absorption) divergence at the bandgap results from the quadratic Stark shift causing a blue shift of the band edge.

By using Equation 26 and substituting for the degenerate Δn, the following form for the degenerate Kerr coefficient is obtained [15]:

$$\gamma = K \frac{\hbar c \, E_p^{1/2}}{2n_0^2 E_g^4} G_2\left(\frac{\hbar\omega}{E_g}\right) \tag{35}$$

where the dispersion function G_2 is given by

$$G_2(x) = \frac{2}{\pi} \int_0^\infty \frac{F_2(x'; x) \, dx'}{x'^2 - x^2} \tag{36}$$

We have neglected any dispersion in the linear refractive index n_0 in the integral. The magnitude of the dispersion is typically only 10% of the background refractive index around the band edge of semiconductors, so we do not anticipate any significant error. the factor of $1/2$ is introduced when the degenerate coefficient is calculated from the nondegenerate as beating terms (cross-modulation or grating terms), double the magnitude of the nonlinearity in the nondegenerate case when the material can respond to the beat frequency. This is often referred to as weak wave retardation [53]. This factor has also been included in the determination of the nonlinear absorption (Equation 34). The origin of this factor can be seen by comparing the degenerate results in Equation B14 of Appendix B to the nondegenerate results given in Equation 25. Note that this factor of $1/2$ was omitted in the nonlinear refractive index calculation in [15]. The individual contributions to G_2 are set out in Table II.

On examining the low frequency limit it is found that these terms diverge as $\omega \to 0$. In order to investigate such non-physical 'infrared' divergence we go one step back and examine the nondegenerate case. It is found that $\Delta n(\omega; \Omega)$ is not divergent in ω, and

therefore any zero-frequency divergence must be in the pump frequency Ω as it appears in $\Delta\alpha$. To identify this divergence further, $\Delta n(\omega; \Omega)$ can be expanded as a Laurent series around $\Omega = 0$. On summing the contributions to G_2, the divergent term is determined to be

$$G_2^{\mathrm{div}} = \frac{g(\omega)}{\Omega^2} \tag{37}$$

where $g(\omega)$ has no divergence at $\omega = 0$. Now by setting $\Omega = \omega$ one arrives at the degenerate divergence function as is shown in Table II. This diverging term is not unexpected as $\boldsymbol{A} \cdot \boldsymbol{p}$ perturbation theory has been used in the transition rate calculation and it is well known that a divergence of this order can be introduced [54], whereas the comparable $\boldsymbol{E} \cdot \boldsymbol{r}$ perturbation theory avoids such a divergence. The latter perturbation technique, however, is not suitable for solids with extended wavefunctions and simple scaling rules cannot be easily derived. In a similar manner to Moss *et al.* [55] we treat such a divergence as non-physical and subtract it from the result for the nonlinear refraction.

The general scaled form for n_2 is given by

$$n_2(\mathrm{e.s.u.}) = K'(E_p^{1/2}/n_0 E_g^4) G_2(\hbar\omega/E_g) \tag{38}$$

where, using a fit to n_2 in semiconductors the constant $K' = 1.50 \times 10^{-8}$ when E_g and E_p are defined in eV. Using the value of K obtained by fitting two-photon absorption measurements in semiconductors, a value for K' is obtained that is within a factor of 2 of the above. A graphical comparison of the dispersion function $G_2(\hbar\omega/E_g)$ with measured values of n_2 is shown in Fig. 9a. The values for semiconductors (squares) were obtained from Z-scan measurements at 1.06 and 0.53 μm [56, 57]. We also show 'nearly degenerate three-wave mixing' n_2 measurements of large gap optical materials [58] (closed circles) and a measurement of n_2 in silica at 249 nm [59] (diamond).

Figure 9b shows the extension to Fig. 9a for frequencies close to the band edge, where the bound electronic refractive nonlinearity shows a resonance due to the quadratic a.c. Stark effect. This graph also includes recent measurements of n_2 in AlGaAs by LaGasse *et al.* [60] using femtosecond time division interferometry (closed triangles).

Hidden in Fig. 9 is the E_g^{-4} scaling of n_2 that gives a variation of n_2 from 2.5×10^{-14} e.s.u. for a material such as MgF_2 at 1.06 μm to -2.6×10^{-9} e.s.u. for AlGaAs at 810 nm [60] and 2.7×10^{-10} e.s.u. for Ge at 10.6 μm, which we measured with a picosecond CO_2 laser. This five orders of magnitude variation of n_2 is displayed better by plotting n_2 scaled by n_0 and G_2 as a function of E_g on a log–log plot as shown in Fig. 10. In spite of this very large variation in magnitude of n_2 (and change in sign), this extremely simple model gives good agreement with the data for materials including both semiconductors and insulators. Additionally, note that although the measured values of n_2 for ZnSe at 1.06 and 0.53 μm have different signs, both measurements are consistent with the scaling law

6.2. Alternative formalism for frequency summation

It has been demonstrated that nonlinear Kramers–Krönig relations can be applied, provided that they take a nondegenerate form. In this section we show how this can be extended to an alternative form in the case of frequency summation in the same manner as in [13, 61, 62]. Starting from the Fourier transform of the nonlinear susceptibility as defined

in Equation 20, a change in variables is performed

$$\Omega_i = \sum_j M_{ij}\omega_j$$

$$T_i = \sum_j (\tilde{M}^{-1})_{ij}\tau_j \tag{39}$$

which ensures that $\Sigma_i\omega_i\tau_i = \Sigma_j\Omega_jT_j$. This now gives

$$\hat{\chi}^{(n)}(\Omega_1, \Omega_2, \ldots, \Omega_n) = \int dT_1 \int dT_2 \cdots \int dT_n \hat{R}^{(n)}(T_1, T_2, \ldots, T_n)\, e^{i(\Omega_1 T_1 + \Omega_2 T_2 + \cdots + \Omega_n T_n)} \tag{40}$$

where the integration region is the equivalent in T-space to the positive quadrant in τ-space. Now consider a special case of the above transform where, for some k, $(\tilde{M}^{-1})_{kj} \geqslant 0$ for all possible j. Since the response function is integrated over only positive τ_j (as causality demands R be zero for any $\tau_j < 0$), then the transformed variable T_k will be integrated over only positive values. That is, $\hat{R}^{(n)}(T_1, T_2, \ldots, T_k, \ldots, T_n)$ can be set to zero for $T_k < 0$. Hence, in a similar manner to the linear Kramers–Krönig derivation, by transforming this property to the frequency domain, we obtain the dispersion relation for $\hat{\chi}^{(n)}(\Omega_1, \Omega_2, \ldots, \Omega_n)$

$$\hat{\chi}^{(n)}(\Omega_1, \Omega_2, \ldots, \Omega_k, \ldots, \Omega_n) = \frac{1}{i\pi}\mathscr{P}\int_{-\infty}^{\infty}\frac{\hat{\chi}^{(n)}(\Omega_1, \Omega_2, \ldots, \Omega, \ldots, \Omega_n)}{\Omega - \Omega_k}\,d\Omega \tag{41}$$

Note that this is formally identical to the nondegenerate nonlinear Kramers–Krönig relation dervied above. On transforming back to real frequencies, we obtain $\omega_i = \Sigma_j(M^{-1})_{ij}\Omega_j$. Since we have previously required $(\tilde{M}^{-1})_{kj} \geqslant 0$, the coefficient of Ω_k, $(M^{-1})_{ik}$ is always positive (or zero). This results in a generalized form for the nonlinear Kramers–Krönig transform

$$\chi^{(n)}(\omega_1 + p_1\omega, \omega_2 + p_2\omega, \ldots, \omega_n + p_n\omega)$$

$$= \frac{1}{i\pi}\mathscr{P}\int_{-\infty}^{\infty}\frac{\chi^{(n)}(\omega_1 + p_1\Omega, \omega_2 + p_2\Omega, \ldots, \omega_n + p_n\Omega)}{\Omega - \omega}\,d\Omega \tag{42}$$

where $p_i \geqslant 0$ for all possible i, provided that at least one $p_i \neq 0$. For the case where only one of the coefficients p_i is nonzero and set equal to unity, we regain the nondegenerate form of the relationship.

We also give here a new, simpler method for obtaining the same relationship in Equation 42 using the Fourier transform as in Equation 7. Starting with the definition of $\chi^{(n)}$ from the response function $R^{(n)}$ in Equation 20, multiplying both sides by $(\Omega - \omega)^{-1}$ and integrating over Ω

$$\mathscr{P}\int_{-\infty}^{\infty}\frac{\chi^{(n)}(\omega_1 + p_1\Omega, \omega_2 + P_2\Omega, \ldots, \omega_n + p_n\Omega)}{\Omega - \omega}\,d\Omega$$

$$= \int_0^{\infty}d\tau_1\int_0^{\infty}d\tau_2\cdots\int_0^{\infty}d\tau_n R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)\, e^{i(\omega_1\tau_1 + \omega_2\tau_2 + \cdots \omega_n\tau_n)}$$

$$\times\ \mathscr{P}\int_{-\infty}^{\infty}\frac{e^{i\Omega(p_1\tau_1 + p_2\tau_2 + \cdots + p_n\tau_n)}}{\Omega - \omega}\,d\Omega$$

$$= i\pi\int_0^{\infty}d\tau_1\int_0^{\infty}d\tau_2\cdots\int_0^{\infty}d\tau_n R^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)e^{i(\omega_1\tau_1 + \omega_2\tau_2 + \cdots + \omega_n\tau_n)}e^{i\omega(p_1\tau_1 + p_2\tau_2 + \cdots + p_n\tau_n)}$$

$$= i\pi\chi^{(n)}(\omega_1 + p_1\omega, \omega_2 + p_2\omega, \ldots, \omega_n + p_n\omega) \tag{43}$$

which is conditional on $(p_1\tau_1 + p_2\tau_2 + \cdots + p_n\tau_n) > 0$. Since causality already requires $\tau_1, \tau_2, \ldots, \tau_n > 0$, this requires that all of the coefficients $p_1, p_2, \ldots, p_n \geqslant 0$.

It can be seen that no combination can give rise to a relation for $\chi^{(3)}(\omega, \omega, -\omega)$ as the lower limit in the Fourier transform would be $-\infty$ rather than zero, and the implications resulting from causality cannot be used. In general there is no valid relation in the case of a nonlinear $\chi^{(n)}$ where the integral is over a single frequency ω that contains a mixture of positive and negative terms in $\chi^{(n)}$. Thus, we can extend degenerate nonlinear Kramers–Krönig relations only to the case of frequency summation or a partially degenerate case such as using ω_1 as the integral variable in $\chi^{(3)}(\omega_1, \omega_1, -\omega_2)$, which gives a polarization at the new frequency $2\omega_1 - \omega_2$. Dispersion relations between self-action processes such as nonlinear refraction and absorption are valid only in the nondegenerate form.

As an example we can write for $\chi^{(2)}$, which governs second harmonic generation (i.e. a polarization is produced at a frequency 2ω for an input at ω),

$$\chi^{(2)}(\omega, \omega) = \frac{1}{i\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\chi^{(2)}(\Omega, \Omega)}{\Omega - \omega} \, d\Omega \tag{44}$$

Thus, we can relate the real and imaginary parts of $\chi^{(2)}$. However, since a new frequency is generated, we are only usually concerned with the magnitude of $\chi^{(2)}$ and not its phase (since there is no initial field at frequency 2ω for it to interfere with). This tends to restrict the use of Equation 44, as experimentally it is the quantity $|\chi^{(2)}|$ that is measured. However, Equation 44 has been used in theoretical calculations where it is far easier to determine $\mathrm{Im}\,\chi^{(2)}$ alone [63]. Similar calculations have also been performed for third-harmonic generation [55].

7. Conclusions

The Kramers–Krönig dispersion relation and its derivation from the principle of causality for linear optics has been well understood for a number of decades. Although nonlinear Kramers–Krönig relations have found considerable use in resonant semiconductor non-linearities, there is some confusion regarding the application of causality to other optical nonlinearities such as nonresonant 'instantaneous' nonlinearities. It is the aim of this review to demonstrate that such relations exist and there is a common form that is applicable to all nonlinear optical mechanisms.

The key relation for the calculation of refractive index changes from absorption coefficient changes is given by Equation 19

$$\Delta n(\omega; \zeta) = \frac{c}{\pi} \mathscr{P} \int_0^{\infty} \frac{\Delta\alpha(\omega'; \zeta)}{\omega'^2 - \omega^2} \, d\omega' \tag{45}$$

where ζ denotes the perturbation that is the source of the change. This relation has been utilized in the calculation of resonant optical nonlinearities. We outline specific examples of this, namely the bandfilling model for resonant semiconductor nonlinearities, thermo-optic nonlinearities in semiconductors and d.c. electric field-induced nonlinearities. We also extend the standard calculation of the two-level atom and have written it in such a form that the Kramers–Krönig relation is valid. Of course, Equation 45 is applicable not only to nonlinear optics where the intense light causes a perturbation that then affects the optical characteristics of the material, but to any perturbation, whether it is optically generated or not. Thus, for example, one could calculate the change in refraction for use in an optical

phase modulator from changes in the absorption spectra of the desired material through a variation in a d.c. electric field.

An important step discussed here is that the perturbation can be another optical field and, thus, extend this relation to nonresonant (fast) optical nonlinearities. This can be obtained from Equation 45, but we also include a derivation based on the causality of the system. However, since the perturbation has to remain constant over the integration, it is necessary to have a nondegenerate form for the change in absorption, i.e. a pump–probe spectrum, with the pump at a fixed frequency. Unfortunately, conventional experiments and theories tend to determine the degenerate (single frequency argument) form only. It is likely that this contributes to the confusion over nonlinear Kramers–Krönig relations, since these conventional results cannot be utilized, leading to the incorrect assumption that nonlinear Kramers–Krönig relations are invalid.

A recent calculation using this dispersion relation is in the calculation of the scaling and dispersion of the electronic nonlinear refractive coefficient n_2 in solids. We reproduce an outline of this calculation in this review. First, all of the relevant nondegenerate nonlinear absorption contributions are determined. This includes terms from two-photon absorption, Raman transitions and the a.c. Stark effect. Then the nonlinear Kramers–Krönig transform is computed and an analytic expression is obtained for the degenerate n_2. On comparison with experimentally measured values, excellent agreement is obtained, considering the uncertainty in some of the materials parameters and the fact that a simple two-band model was used to calculate the nonlinear absorptive contributions. Note, although we have calculated the degenerate form of the electronic n_2 in order to provide a comparison with experimental data, the nonlinear dispersion relation is much more general. No new information is necessary for the nondegenerate form of n_2 to be calculated.

As a final note, although in general nonlinear dispersion relations must take the nondegenerate form, it can be shown through a change of variables that in the particular case of frequency summation a degenerate form can be used. This has limited applications although as this is not a self-action effect, so one is usually concerned with the magnitude of χ only and not the real and imaginary components.

Acknowledgements

References

1. H. A. KRAMERS, *Atti Congr. Int. Fis. Como* **2** (1927) 545.
2. R. DE L. KRÖNIG, *J. Opt. Soc. Am. Rev. Scient. Instrum.* **12** (1926) 547.
3. *Idem, Ned. Tijdschr. Natuurk* **9** (1942) 402.
4. H. BODE, 'Network Analysis and Feedback Amplifier Design' (Van Nostrand, New York, 1945).
5. H. M. NUSSENZVEIG, 'Causality and Dispersion Relations' (Academic Press, New York, 1972).
6. W. SCHÜTZER and J. TIOMNO, *Phys. Rev.* **83** (1951) 249.
7. E. P. WIGNER, *Am. J. Phys.* **23** (1955) 371.
8. E. C. TITCHMARSH, 'Introduction to the Theory of Fourier Integrals' (Oxford University Press, Oxford, 1948).
9. J. S. TOLL, *Phys. Rev.* **104** (1956) 1760.
10. A. YARIV, 'Quantum Electronics', 2nd Edn (Wiley, New York, 1975).
11. K. A. SHORE and D. A. S. CHAN, *Electron. Lett.* **26** (1990) 1206.
12. D. A. B. MILLER, C. T. SEATON, M. E. PRISE and S. D. SMITH, *Phys. Rev. Lett.* **47** (1981) 197.
13. J. F. L. RIDENER and J. R. H. GOOD, *Phys. Rev.* **B10** (1974) 4980.

14. M. SHEIK-BAHAE, D. J. HAGAN and E. W. VAN STRYLAND, *Phys. Rev. Lett.* **65** (1990) 96.
15. M. SHEIK-BAHAE, D. C. HUTCHINGS, D. J. HAGAN and E. W. VAN STRYLAND, *IEEE J. Quantum Electron.* **QE-27** (1991) 1296.
16. S. M. KOGAN, *Sov. Phys. JETP* **16** (1963) 217.
17. P. J. PRICE, *Phys. Rev.* **130** (1963) 1792.
18. W. J. CASPERS, *ibid.* **A133** (1964) 1249.
19. P. MEYSTRE and M. SARGENT III, in 'Elements of Quantum Optics' (Springer–Verlag, Berlin, 1990).
20. R. W. BOYD and S. MUKAMEL, *Phys. Rev.* **B29** (1984) 1973.
21. T. S. MOSS, *Proc. Phys. Soc. Lond.* **B67** (1954) 775.
22. E. BURSTEIN, *Phys. Rev.* **93** (1954) 632.
23. D. A. B. MILLER, M. H. MOZOLOWSKI, A. MILLER and S. D. SMITH, *Opt. Commun.* **27** (1978) 133.
24. H. M. GIBBS, T. N. C. VENKATESAN, S. L. McCALL, A. PASSNER, A. C. GOSSARD and W. WIEGMANN, *Appl. Phys. Lett.* **34** (1979) 511.
25. H. M. GIBBS, S. L. McCALL, T. N. C. VENKATESAN, A. C. GOSSARD, A. PASSNER and W. WIEGMANN, *ibid.* **35** (1979) 451.
26. D. A. B. MILLER, S. D. SMITH and A. M. JOHNSTON, *ibid.* **35** (1979) 658.
27. B. S. WHERRETT and N. A. HIGGINS, *Proc. R. Soc.* **A379** (1982) 67.
28. E. O. KANE, *J. Phys. Chem. Solids* **1** (1957) 249.
29. L. BÁNYAI and S. W. KOCH, *Z. Phys. B: Condens. Matter* **63** (1986) 283.
30. N. FINLAYSON, W. C. BANYAI, C. T. SEATON, G. I. STEGEMAN, M. O'NEILL, T. J. CULLEN and C. N. IRONSIDE, *J. Opt. Soc. Am.* **B6** (1989) 675.
31. J. P. LÖWENAU, S. SCHMITT-RINK and H. HAUG, *Phys. Rev. Lett.* **49** (1982) 1511.
32. Y. H. LEE, A. CHAVEZ-PIRSON, S. W. KOCH, H. M. GIBBS, S. H. PARK, J. MORHANGE, A. JEFFREY, N. PEYGHAMBARIAN, L. BANYAI, A. C. GOSSARD and W. WIEGMANN, *ibid.* **57** (1986) 2446.
33. D. S. CHEMLA, D. A. B. MILLER, P. W. SMITH, A. C. GOSSARD, and W. WIEGMANN, *IEEE J. Quantum Electron.* **QE-20** (1984) 265.
34. J. S. WEINER, D. A. B. MILLER and D. S. CHEMLA, *Appl. Phys. Lett.* **50** (1987) 842.
35. Y. WANG, N. HERRON, W. MAHLER and A. SUNA, *J. Opt. Soc. Am.* **B6** (1989) 808.
36. F. V. KARPUSHKO and G. V. SINITSYN, *J. Appl. Spectrosc. USSR* **29** (1978) 1323.
37. D. C. HUTCHINGS, A. D. LLOYD, I. JANOSSY and B. S. WHERRETT, *Opt. Commun.* **61** (1987) 345.
38. B. S. WHERRETT, D. HUTCHINGS and D. RUSSELL, *J. Opt. Soc. Am.* **B3** (1986) 351.
39. LANDOLT-BÖRSTEIN, 'Numerical Data and Functional Relationships in Science and Technology', Vols 17a and 17b 'Group III' (Springer-Verlag, Berlin, 1982).
40. W. FRANZ, *Z. Naturforsch.* **13a** (1958) 484.
41. L. V. KELDYSH, *Sov. Phys. JETP* **34** (1958) 788.
42. K. THARMALINGAM, *Phys. Rev.* **130** (1963) 2204.
43. M. ABRAMOWITZ and I. A. STEGUN, 'Handbook of Mathematical Functions' (Dover, New York, 1964).
44. B. O. SERAPHIN and N. BOTTKA, *Phys. Rev.* **139** (1965) 560.
45. *Opt. Quantum Electron.* **22** (special issue on charge transport nonlinearities) (1990).
46. P. GUNTER and J. P. HUIGNARD (editors), 'Photorefractive Materials and their Applications', Vol. 61 (Springer-Verlag, Berlin, 1988).
47. B. S. RYVKIN, *Sov. Phys. Semicond.* **15** (1981) 796.
48. D. A. B. MILLER, D. S. CHEMLA, T. C. DAMEN, T. H. WOOD, C. A. BURRUS, JR, A. C. GOSSARD and W. WIEGMANN, *IEEE J. Quantum Electron.* **QE-21** (1985) 1462.
49. J. A. VAN VECHTEN and D. E. ASPNES, *Phys. Lett.* **30A** (1969) 346.
50. M. H. WEILER, *Solid St. Commun.* **39** (1981) 937.
51. B. S. WHERRETT, *J. Opt. Soc. Am.* **B1** (1984) 67.
52. E. W. VAN STRYLAND, H. VANHERZEELE, M. A. WOODALL, M. J. SOILEAU, A. L. SMIRL, S. GUHA and T. F. BOGGESS, *Opt. Engng* **24** (1985) 613.
53. E. W. VAN STRYLAND, A. L. SMIRL, T. F. BOGGESS, M. J. SOILEAU, B. S. WHERRETT and F. A. HOPF, in 'Chemistry and Physics of Picosecond Phenomena III', Vol. 23, edited by K. B. Eisenthal, R. M. Hochstrasser, W. Kaiser and A. Laubereau (Springer-Verlag, Berlin, 1982) p. 368.
54. J. M. WORLOCK, in 'Laser Handbook', edited by F. T. Arecchi and E. D. Schulz-DuBois (North-Holland, Amsterdam, 1972) p. 1323.
55. D. J. MOSS, E. GHAHRAMANI, J. E. SIPE and H. M. VAN DRIEL, *Phys. Rev.* **B41** (1990) 1542.

56. M. SHEIK-BAHAE, A. A. SAID, T. H. WEI, D. J. HAGAN and E. W. VAN STRYLAND, *IEEE J. Quantum Electron.* **26** (1990) 760.
57. M. SHEIK-BAHAE, A. A. SAID and E. W. VAN STRYLAND, *Opt. Lett.* **14** (1989) 955.
58. R. ADAIR, L. L. CHASE and S. A. PAYNE, *Phys. Rev.* **B39** (1989) 3337.
59. I. N. ROSS, W. T. TONER, C. J. HOOKER, J. R. M. BARR and I. COFFEY, *J. Modern Opt.* **37** (1990) 555.
60. M. J. LaGASSE, K. K. ANDERSON, C. A. WANG, H. A. HAUS and J. G. FUJIMOTO, *Appl. Phys. Lett.* **56** (1990) 417.
61. J. F. L. RIDENER and J. R. H. GOOD, *Phys. Rev.* **B11** (1975) 2768.
62. F. SMET and A. VAN GROENENDAEL, *ibid.* **A19** (1979) 334.
63. D. J. MOSS, J. E. SIPE and H. M. VAN DRIEL, *ibid.* **B36** (1987) 9708.
64. P. N. BUTCHER and D. COTTER, 'The Elements of Nonlinear Optics' (Cambridge University Press, Cambridge, 1990).
65. F. A. HOPF and G. I. STEGEMAN, 'Applied Classical Electrodynamics', Vol. 2: 'Nonlinear Optics' (Wiley, New York, 1986).
66. S. CHAPMAN, *Am. J. Phys.* **24** (1956) 162.

Appendices

A. Cauchy integral derivation of dispersion relations

The susceptibility $\chi(\omega)$ is derived from the response function by means of the Fourier transform in Equation 4. However, because of causality, the integral need over run only positive times. Generalizing this for complex ω, for the condition $\mathrm{Im}\,\omega > 0$, the convergence of Equation 4 can only be improved. This can easily be seen by writing $\omega = \omega' + i\omega''$, which leads to a factor $e^{-\omega''T}$ in the integrand. Thus, $\chi(\omega)$ has a regular analytic continuation in the positive imaginary plane of ω.

Consider the Cauchy integral round the contour Γ in the frequency plane as shown in Fig. A1.

$$\int_\Gamma \frac{\chi(\omega)\,d\omega}{\omega - \Omega} = 0 \tag{A1}$$

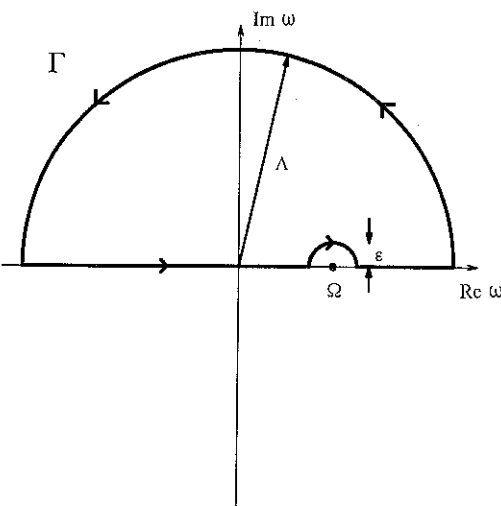which is zero since ω is regular and analytic with no poles within the contour. First, consider

Figure A1 The contour Γ in the complex frequency plane used in the derivation of dispersion relations of the large semicircle of radius Λ, the axis $\mathrm{Im}\,\omega = 0$ and the small semicircle of radius ε centred around the point Ω.

25

the contribution from the large semicircle $\omega = \Lambda e^{i\theta}$, where θ runs from 0 to π

$$\lim_{\Lambda \to \infty} \left| \int_0^\pi \frac{\chi(\Lambda e^{i\theta})\,\Lambda e^{i\theta}\,d\theta}{\Lambda e^{i\theta} - \Omega} \right| \leqslant \max_{0 \leqslant \theta \leqslant \pi} |\chi(\Lambda e^{i\theta})| \int_0^\pi \frac{\Lambda e^{i\theta}\,d\theta}{\Lambda e^{i\theta} - \Omega}$$

$$\leqslant \max_{0 \leqslant \theta \leqslant \pi} |\chi(\Lambda e^{i\theta})| \ln \left(\frac{\Lambda + \Omega}{\Lambda - \Omega} \right) \Big|_0^\infty \to 0 \tag{A2}$$

which vanishes provided that $\chi(\omega)$ is reasonably well-behaved at $\pm \infty$

$$\lim_{\omega \to \pm \infty} [\chi(\omega)/\omega] = 0 \tag{A3}$$

Secondly, consider the contribution from the small semicircle centred on $\omega = \Omega$. In the limit of a vanishingly small radius, this is just a half-residue of the integrand at Ω. Thus, the total of all the contributions is

$$0 - i\pi\chi(\Omega) + \lim_{\varepsilon \to 0} \left(\int_{-\infty}^{\omega-\varepsilon} \frac{\chi(\omega)\,d\omega}{\omega - \Omega} + \int_{\omega+\varepsilon}^{\infty} \frac{\chi(\omega)\,d\omega}{\omega - \Omega} \right) = 0 \tag{A4}$$

This can be rewritten as

$$\chi(\Omega) = \frac{1}{i\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\chi(\omega)\,d\omega}{\omega - \Omega} \tag{A5}$$

where $\mathscr{P}$ denotes the Cauchy principal value. As a result of the factor i in this relation, the real part of $\chi(\omega)$ can be expressed solely in terms of its imaginary part, and vice versa. Thus, given knowledge of either the real or imaginary part of $\chi(\omega)$ over all frequencies, it is possible to determine the other completely.

This derivation of the optical susceptibility also provides an easy check to see whether dispersion relations exist for an arbitrary $\chi(\omega)$. Since it is necessary for $\chi(\omega)$ to be regular and analytic in the upper half-plane, it should have no poles in that plane. Therefore, if such poles exist there can be no dispersion relations. This result has been used in the discussion on the two-level atom.

B. Relationships between various nonlinear optical descriptions

In nonlinear optics we examine regions where there is no longer a simple linear relation between the polarization and electric field as given in Equation 3. It is conventional to write the polarization in terms of an expansion in the electric field. Thus, we can write the various contributions to the polarization $P_i(t) = P_i^{(1)}(t) + P_i^{(2)}(t) + P_i^{(3)}(t) + \cdots$ in the time domain as [64]

$$P_i^{(1)}(t) = \int_{-\infty}^{\infty} d\tau_1\, R_{ij}^{(1)}(\tau_1)\, E_j(t - \tau_1)$$

$$P_i^{(2)}(t) = \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2\, R_{ijk}^{(2)}(\tau_1, \tau_2)\, E_j(t - \tau_1)\, E_k(t - \tau_2)$$

$$P_i^{(3)}(t) = \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \int_{-\infty}^{\infty} d\tau_3\, R_{ijkl}^{(3)}(\tau_1, \tau_2, \tau_3)\, E_j(t - \tau_1)\, E_k(t - \tau_2)\, E_l(t - \tau_3)$$

$$\vdots$$

$$P_i^{(n)}(t) = \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \cdots \int_{-\infty}^{\infty} d\tau_n\, R_{ijk\ldots n}^{(n)}(\tau_1, \tau_2, \ldots, \tau_n)\, E_j(t - \tau_1)$$
$$\times E_k(t - \tau_2) \ldots E_n(t - \tau_n) \tag{B1}$$

Here the summation over the various directions j, k, l is implied for the various tensor elements of χ. Upon Fourier transformation we obtain the equivalent relations in the frequency domain

$$P_i^{(1)}(\omega) = \int_{-\infty}^{\infty} d\omega_1 \, \chi_{ij}^{(1)}(\omega_1) \, E_j(\omega_1) \, \delta(\omega - \omega_1)$$

$$P_i^{(2)}(\omega) = \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \, \chi_{ijk}^{(2)}(\omega_1, \omega_2) \, E_j(\omega_1) \, E_k(\omega_2) \, \delta(\omega - \omega_1 - \omega_2)$$

$$P_i^{(3)}(\omega) = \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \int_{-\infty}^{\infty} d\omega_3 \, \chi_{ijkl}^{(3)}(\omega_1, \omega_2, \omega_3) \, E_j(\omega_1) \, E_k(\omega_2)$$
$$\times \, E_l(\omega_3) \, \delta(\omega - \omega_1 - \omega_2 - \omega_3)$$

$$\vdots$$

$$P_i^{(n)}(\omega) = \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \cdots \int_{-\infty}^{\infty} d\omega_n \, \chi_{ijk \ldots n}^{(n)}(\omega_1, \omega_2, \ldots, \omega_n)$$
$$\times \, E_j(\omega_1) \, E_k(\omega_2) \ldots E_n(\omega_n) \, \delta(\omega - \omega_1 - \omega_2 - \cdots - \omega_n) \tag{B2}$$

where δ is the Dirac delta function. The nth order susceptibility is generally obtained from the Fourier transform of the nth order response function

$$\chi_{ijk \ldots n}^{(n)}(\omega_1, \omega_2, \ldots, \omega_n) = \sum_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \cdots$$
$$\times \int_{-\infty}^{\infty} d\tau_n \, R_{ijk \ldots n}^{(n)}(\tau_1, \tau_2, \ldots, \tau_n) \, e^{i(\omega_1 \tau_1 + \omega_2 \tau_2 + \cdots + \omega_n \tau_n)} \tag{B3}$$

It can be seen that on evaluating the frequency integral for the linear polarization $P^{(1)}(\omega)$, Equation 3 is again obtained.

In this paper we are mainly concerned with monochromatic electromagnetic fields defined as

$$E(t) = \tfrac{1}{2}(E_a e^{-i\omega_a t} + E_a^* e^{i\omega_a t}) \tag{B4}$$

whose Fourier transform is

$$E(\omega) = \tfrac{1}{2}[E_a \delta(\omega - \omega_a) + E_a^* \delta(\omega + \omega_a)] \tag{B5}$$

Throughout much of this paper we are concerned with third-order nonlinearities. The reason for this is that this is the lowest order of nonlinearity that produces a nonlinear polarization at the same frequency as the applied electromagnetic field. This results in self-action effects such as nonlinear absorption and refraction.

Upon inserting the monochromatic field defined in Equation B5 into Equation B2 for the third-order nonlinear polarization and performing the various frequency integrals, one obtains for the component $P^{(3)}(\omega)$ parallel to E

$$P^{(3)}(\omega) = \tfrac{1}{8}[3\chi^{(3)}(\omega_a, \omega_a, -\omega_a) E_a^2 E_a^* \delta(\omega - \omega_a) + 3\chi^{(3)*}(\omega_a, \omega_a, -\omega_a) E_a E_a^{*2} \delta(\omega + \omega_a)$$
$$+ \chi^{(3)}(\omega_a, \omega_a, \omega_a) E_a^3 \delta(\omega - 3\omega_a) + \chi^{(3)*}(\omega_a, \omega_a, \omega_a) E_a^{*3} \delta(\omega + 3\omega_a)] \tag{B6}$$

Here we have used permutation symmetry [64] to simplify the above result, where for example

$$\chi_{ijkl}^{(3)}(\omega_1, \omega_2, \omega_3) = \chi_{ijlk}^{(3)}(\omega_1, \omega_3, \omega_2)$$
$$= \chi_{ilkj}^{(3)}(\omega_3, \omega_2, \omega_1) \tag{B7}$$

27

This symmetry arises from the fact the E-fields and indices in Equation B1 can be interchanged without affecting the result. In addition, we have also applied the reality condition [64] that comes from the requirement that the response function $R(t)$ must be real since $E(t)$ and $P(t)$ are both real

$$\chi^{(3)}*(\omega_1, \omega_2, \omega_3) = \chi^{(3)}(-\omega_1, -\omega_2, -\omega_3) \tag{B8}$$

It can be seen from Equation B6 that there are components produced at the same frequency as the fundamental ($\pm \omega_a$) as well as third-harmonic terms ($\pm 3\omega_a$).

In nonlinear optics, effects are often expressed in terms of a nonlinear refractive index or absorption coefficient rather than in terms of $\chi^{(3)}$, especially for self-action third-order effects, which we examine in the following. In Gaussian units the refractive index and absorption coefficient can be obtained from the definitions for the electric displacement $D(\omega)$

$$D(\omega) = [n(\omega) + i\alpha(\omega)c/2\omega]^2 E(\omega) = E(\omega) + 4\pi P(\omega) \tag{B9}$$

from which we can obtain the following

$$\Delta n = (3\pi/2n_0)|E_a|^2 \operatorname{Re} \chi^{(3)}(\omega_a, \omega_a, -\omega_a)$$
$$\Delta \alpha = (3\pi\omega_a/n_0 c)|E_a|^2 \operatorname{Im} \chi^{(3)}(\omega_a, \omega_a, -\omega_a) \tag{B10}$$

assuming the background absorption is small, $\alpha_b c/\omega \ll n_0$, which is usually necessary anyway for the observation of nonlinear optical effects. Here n_0 refers to the background (linear) refractive index. The nonlinear refractive effect is commonly referred to as the optical Kerr effect. One of the possible contributions to the nonlinear absorptive effect is two-photon absorption.

An alternative description for the optical Kerr effect defines

$$\Delta n = n_2 \langle E^2(t) \rangle = \tfrac{1}{2} n_2 |E_a|^2 \tag{B11}$$

where $\langle E^2(t) \rangle$ denotes the time average for the square of the electric field. This then gives the relation between n_2 and $\chi^{(3)}$ in the present notation as

$$n_2(\omega) = (3\pi/n_0)\operatorname{Re} \chi^{(3)}(\omega, \omega, -\omega) \tag{B12}$$

This relation between n_2 and $\chi^{(3)}$ varies considerably in the literature as the definition of n_2 (Equation B11) is sometimes defined to be a factor of 2 greater, the definition of monochromatic electric field (Equation B4) is sometimes a factor of 2 greater and the summation over the various permutations of the frequency arguments in calculating the nonlinear polarization sometimes is not performed (Equation B2), which results in a factor of 3 difference. For instance, one often sees $\chi_{\text{eff}}^{(3)}$ defined for monochromatic light as (for example, see [65])

$$P(\omega) = \chi_{\text{eff}}^{(3)}(\omega)\langle E^2(t) \rangle E(\omega) \tag{B13}$$

Comparing this with our definition gives $\chi_{\text{eff}}^{(3)}(\omega) = \tfrac{3}{2}\chi^{(3)}(\omega, \omega, -\omega)$. In addition, definitions vary when SI units are used instead of Gaussian units, and n_2 is sometimes used with irradiances, which would be equivalent to γ as defined below. However, our form is completely consistent with our basic definitions (particularly the use of the response function) and is extendable to any degree of nonlinearity.

The nonlinear optical effects can also be expressed in terms of irradiances, i.e. $\Delta\alpha = \beta I$ and $\Delta n = \gamma I$. By using the relation between irradiance and electric field, $I = (n_0 c/8\pi)|E_a|^2$,

the relation between these quantities and $\chi^{(3)}$ can be written as

$$\beta(\omega) = (24\pi^2 \omega/n_0^2 c^2)\text{Im}\,\chi^{(3)}(\omega, \omega, -\omega)$$

$$\gamma(\omega) = (12\pi^2/n_0^2 c)\text{Re}\,\chi^{(3)}(\omega, \omega, -\omega) \tag{B14}$$

Note that for the equivalent nondegenerate quantities derived in Equation 25, a factor of 2 appears due to weak wave retardation [53].

C. Conversion from Gaussian to SI units

This review uses Gaussian (CGS) units and in this appendix we indicate the differences that occur when SI (MKS) units are used instead. A good review of the differences in expressions under the systems can be found in [66] for linear quantities. However, in the field of nonlinear optics there are no consistent definitions over all publications, even within one set of units. In this appendix the notation of [64] is used, which is the direct SI equivalent of the notation used throughout here.

In SI units the permittivity of free space, ε_0, appears in the relationship between the polarization and electric field, and hence, for example, for the first-order (linear) polarization we have

$$P_i^{(1)}(t) = \varepsilon_0 \int_{-\infty}^{\infty} d\tau_1\, R_{ij}^{(1)}(\tau_1)\, E_j(t - \tau_1) \tag{C1}$$

and in frequency space

$$P_i^{(1)}(\omega) = \varepsilon_0\, \chi_{ij}^{(1)}(\omega)\, E_j(\omega) \tag{C2}$$

Furthermore, the relation between electric displacement, polarization and field is now written as

$$D(\omega) = \varepsilon_0[n(\omega) + i\alpha(\omega)c/2\omega]^2 E(\omega) = \varepsilon_0 E(\omega) + P(\omega) \tag{C3}$$

Under SI units the irradiance is now given by $I = (\varepsilon_0 n_0 c)/2)|E_a|^2$ when the definition $E(t) = (E_a e^{i\omega_a t} + E_a^* e^{-i\omega_a t})/2$ is used.

Using the same analysis as previously, which leads to a summation over the three indices for the third-order terms, for a $\chi^{(3)}$ nonlinearity it can be shown that the changes in refractive index and absorption coefficient are now given by

$$\Delta n = \frac{3}{8n_0}\,\text{Re}\,\chi^{(3)}(\omega_a, \omega_a, -\omega_a)|E_a|^2 = \frac{3}{4}\frac{I_a}{\varepsilon_0 c n_0^2}\,\text{Re}\,\chi^{(3)}(\omega_a, \omega_a, -\omega_a)$$

$$\Delta\alpha = \frac{3}{4n_0}\,\text{Im}\,\chi^{(3)}(\omega_a, \omega_a, -\omega_a)|E_a|^2 = \frac{3}{2}\frac{\omega_a I_a}{\varepsilon_0 c^2 n_0^2}\,\text{Im}\,\chi^{(3)}(\omega_a, \omega_a, -\omega_a) \tag{C4}$$

Thus, the definitions for the Kerr and two-photon absorption coefficients will be modified accordingly. It is important to note, however, that the Kramers–Krönig relation between n and α is unaltered in SI units.

Both in the literature and this review, two principal forms have been used for quoting values for nonlinear refraction: n_2 and γ. n_2 is most often quoted in e.s.u. and γ is usually quoted in SI or some derived unit (e.g. $\text{cm}^2\,\text{GW}^{-1}$). By comparing Equations B12 and B14 and converting units accordingly, we can write the conversion between these two forms as

$$n_2(\text{e.s.u.}) = (n_0 c/40\pi)\gamma \quad (\text{SI}) \tag{C5}$$

where all of the quantities on the right-hand side are quoted in SI.

D. Numerical evaluation of principal parts integrals

The Kramers–Krönig relation is an example of a principal parts integral. Most numerical integration routines have difficulty in handling poles in the integrand. Here we present a method that transforms the principal parts integral into integrals that have no poles.

Consider the principal parts integral for $d > 0$

$$I = \mathscr{P} \int_0^\infty \frac{f(x)\,dx}{x - d} \overset{\text{def}}{=} \lim_{\delta \to 0} \left(\int_0^{d-\delta} \frac{f(x)\,dx}{x - d} + \int_{d+\delta}^\infty \frac{f(x)\,dx}{x - d} \right) \tag{D1}$$

This integral can be split into the regions

$$I = \lim_{\delta \to 0} \left(\int_0^{d-\delta} \frac{f(x)\,dx}{x - d} + \int_{d+\delta}^{2d} \frac{f(x)\,dx}{x - d} \right) + \int_{2d}^\infty \frac{f(x)\,dx}{x - d} \tag{D2}$$

In the second region above, perform the transformation of the integration variable $x \to 2d - x$. This then transforms the limits of the second integral to be the same as the first

$$I = \lim_{\delta \to 0} \int_0^{d-\delta} \frac{f(x) - f(2d - x)}{x - d}\,dx + \int_{2d}^\infty \frac{f(x)\,dx}{x - d} \tag{D3}$$

Now consider the integrand of the first integral in the limit

$$\lim_{x \to d} \frac{f(x) - f(2d - x)}{x - d} \overset{\text{def}}{=} 2 \left. \frac{df}{dx} \right|_{x=d} \tag{D4}$$

In other words, provided the derivatives of $f(x)$ exists at the point $x = d$, we have effectively removed this pole from the integral. This requires that our original integrand should have only a first-order pole for this expression to be useful. Hence, Equation D3 provides the form for the numerical evaluation of principal parts integrals consisting of an integral over a finite range and one over a semi-infinite range. If the integration routine used evaluates the end-point (e.g. Simpson's rule), it will be necessary to substitute the limit in Equation D4 for this one point. However, this will not be necessary for routines where the end-point is not evaluated (e.g. Gaussian quadrature). For the semi-infinite integral it is possible to use a transform of coordinates to turn this into an integral over a finite range, but many numerical packages exist in which this is performed automatically.