IOP Publishing European Journal of Physics

Eur. J. Phys. **31** (2010) 573–577 doi:10.1088/0143-0807/31/3/014

What did Kramers and Kronig do and how did they do it?

Craig F Bohren

Department of Meteorology, Pennsylvania State University, University Park, PA 16802 USA

E-mail: bohren@meteo.psu.edu

Received 2 January 2010, in final form 2 March 2010
Published 30 March 2010
Online at stacks.iop.org/EJP/31/573

Abstract

Over time the account of how the Kramers–Kronig (dispersion) relations between the real and imaginary parts of response functions were derived in 1926 and 1927 has been transmogrified into anecdotes about what might have been done but was not. Although Kramers obtained both members of a pair of relations, Kronig obtained only one. Both authors appealed to specific models of an atomic gas rather than to the general arguments about linearity, causality and analyticity in modern model-independent derivations. Kramers merely speculated on whether the specific results he obtained might have a more general validity. Neither author showed that a signal cannot travel faster than $c$in any medium for which the dispersion relations are satisfied. Indeed, they did not mention, even obliquely, signal speeds and causality. Despite their magical aura, Kramers–Kronig relations are translations into somewhat cryptic frequency language of statements clearer in time language.

Although advanced undergraduate students or beginning graduate students are likely to be taught about Kramers–Kronig relations in courses in electromagnetic theory, these relations usually are presented as mathematical theorems that, like Athena from the head of Zeus, emerged fully formed from the heads of their eponyms. These relations, however, were physically motivated, not the result of investigations of the properties of analytic functions. The purpose of this paper is to set the historical record straight and to make the Kramers–Kronig relations more palatable by showing that they are statements translated from a language that more clearly reveals their physical origins.

Kramers–Kronig relations (or dispersion relations) have an aura of magic because they are not physically transparent although they do rest on physical foundations: linearity, causality and the inability of physical systems to respond to excitation at indefinitely high frequencies.

Suppose that a physical quantity X at time t depends on another quantity Y at all other times t' by way of the linear functional relationship

$$X(t) = \int_{-\infty}^{\infty} R(t-t')Y(t')\,\mathrm{d}t', \tag{1}$$

where R is a response function; $R(t) = X(t)$ if $Y(t) = \delta(t)$. From the convolution theorem, the Fourier transforms from time t to frequency ω of these three functions are related by

$$\mathcal{X}(\omega) = \mathcal{R}(\omega)\mathcal{Y}(\omega). \tag{2}$$

For X and Y to be real, $\mathcal{X}(\omega) = \mathcal{X}^*(-\omega)$, $\mathcal{Y}(\omega) = \mathcal{Y}^*(-\omega)$ and hence $\mathcal{R}(\omega) = \mathcal{R}^*(-\omega)$, sometimes called a crossing condition. Strict causality—the past can determine the present but the future cannot—is expressed as $R(t - t') = 0$ for $t' \geqslant t$. Strict causality is to be distinguished from Einstein (or relativistic) causality, according to which no signal can propagate faster than c. Often the qualifiers modifying causality are omitted, context usually sufficient to convey which one is meant, although they are confused sometimes. The one does not imply the other. Einstein causality has withstood many attempts to topple it [1].

Because of causality, subject to proper behaviour of $\mathcal{R}(\omega)$ as $\omega \to \infty$ and possibly at some finite frequencies, the real and imaginary parts (or amplitude and phase) of $\mathcal{R}$ are connected by integral relations. For example, the real and imaginary parts of the electric susceptibility $\chi = \chi' + \mathrm{i}\chi''$ [$\lim_{\omega\to\infty}\chi(\omega) = 0$] of a linear, isotropic, non-chiral, optically homogeneous medium are related by

$$\chi'(\omega) = \frac{2}{\pi} P \int_0^\infty \frac{\Omega\chi''(\Omega)}{\Omega^2 - \omega^2}\, \mathrm{d}\Omega, \tag{3}$$

$$\chi''(\omega) = \frac{-2\omega}{\pi} P \int_0^\infty \frac{\chi'(\Omega)}{\Omega^2 - \omega^2}\, \mathrm{d}\Omega, \tag{4}$$

where P denotes the Cauchy principal value. Hu [2] derives equations (3) and (4) ‘in two lines’, and Sharnoff [3] discusses in detail the conditions under which they are valid. For an alternative derivation in many more than two lines, see King [4]. From equation (3), follows the sum rule

$$\chi'(0) = \frac{2}{\pi} \int_0^\infty \frac{\chi''(\Omega)}{\Omega}\, \mathrm{d}\Omega. \tag{5}$$

If $\mathcal{R}(\omega)$ does not behave nicely, it can be tamed by adding a function to it or multiplying it by a function, possibly changing the crossing condition with the result that its real and imaginary parts satisfy relations similar but not identical to equations (3) and (4).

Although the complex refractive index $n + \mathrm{i}k$ is not a fundamental material response function, it is an analytic function of such a function, namely $\sqrt{1+\chi}$ (if the permeability is that of free space), provided that $1 + \chi \neq 0$ in the upper half of the complex frequency plane, and hence its real and imaginary parts satisfy equations of the same form as equations (3) and (4):

$$n(\omega) - 1 = \frac{2}{\pi} P \int_0^\infty \frac{\Omega k(\Omega)}{\Omega^2 - \omega^2}\, \mathrm{d}\Omega = \frac{c}{\pi} P \int_0^\infty \frac{\alpha(\Omega)}{\Omega^2 - \omega^2}\, \mathrm{d}\Omega, \tag{6}$$

$$k(\omega) = \frac{\alpha(\omega)c}{2\omega} = -\frac{2\omega}{\pi} P \int_0^\infty \frac{n(\Omega)}{\Omega^2 - \omega^2}\, \mathrm{d}\Omega, \tag{7}$$

where α is the absorption coefficient. The upper limits of integration should not be interpreted literally. At sufficiently high frequencies, the corresponding wavelength is much less than molecular diameters and hence the concept of an unrestricted refractive index of an optically homogeneous medium becomes shaky, which was recognized by Kronig [5]. But we still can define$n$as a phase-shift parameter in the *exact* forward direction, and k determines the spatial decrease in the amplitude of a plane wave in this direction.

The integral of $n(\omega) - 1$ over *all* frequencies vanishes [6], and hence n must be less than 1 at *some* frequencies (or 1 at *all* frequencies).

The usefulness of Kramers–Kronig relations and sum rules in analyzing optical data is discussed by Smith [7], for example. Among other duties, they constrain optical properties and provide means by which measuring one frequency-dependent quantity can yield another, for example, the phase of a reflection coefficient at a given frequency is determined by the integral of its amplitude over a sufficiently large neighbourhood of that frequency.

The physical content of Kramers–Kronig relations is manifest in equation (1), a non-local relation in time between an output X and an input Y. It makes physical sense that an output is not determined by the instantaneous value of an input. For example, an oscillator driven by a time-varying force cannot follow it in lockstep. Because of inertia and damping, it takes time for the oscillator to respond to a change in the force, during which time it has, so to speak, moved on. Kramers–Kronig relations are translations into somewhat cryptic frequency language of statements clearer in time language. This translation is needed because measurements are made and theories formulated much more often in the frequency domain than in the time domain. Although spacetime is the stage on which electromagnetic fields act, time often is backstage.

The susceptibility $\chi(\omega)$ is the Fourier transform of the response function relating the electric polarization to the electric field. For these fields to be synchronous would require this response function to be a delta function, which in turn would require χ to be real and independent of frequency. But then time-harmonic fields would be in phase for *all* frequencies, which is not possible (except, trivially, in free space).

Kramers and Kronig sometimes are credited posthumously with what they might have done but did not, at least not in their original papers, cited more often than read. Both authors appealed to *specific* atomic models of matter rather than to the *general* arguments about linearity, causality and analyticity in modern model-independent derivations.

Presenting Kramers–Kronig relations as if they were derived as theorems in the theory of functions of a complex variable obscures their origins. There are no complex variables in Kronig's 1926 paper [5]. His interest was in x-rays ($n \approx 1$) and his atomic absorption coefficient is an integral over a narrow line, because of which the polarizability he used to determine n is real. He went directly to n instead of indirectly by way of χ, and obtained an expression similar to equation (6) if α is interpreted as the number of atoms per unit volume of a *gas* times the sum of absorption coefficients associated with all transitions. Nussenzveig [8] correctly points out that equation (6) was the 'first-known dispersion relation'. Note the singular 'relation': Kronig [5] gave only *one* of the eponymous relations. Kramers and Kronig knew of each other's work but Kronig published first, and hence subsequently has been credited with two relations for the price of one.

In a frequently cited paper on the foundations of causality and dispersion relations, Toll [9] asserts that 'Kramers' used the notion of the complex refractive index defined by analytic continuation in the complex frequency plane to show that a signal cannot travel faster than c in any medium for which the dispersion relation is satisfied'. The assertion that signals cannot be propagated faster than c in a medium for which equations (6) and (7) are valid does not hold up under scrutiny if for no other reason than lack of universal agreement about the definition of signal speed or how to measure it. Although the signal speed—regardless of how defined—in a material for which these equations are satisfied may be less than c, this does not prove that they are *necessary* for Einstein causality. Analytic continuation is a technique to extend the domain of definition of an analytic function. Although the domain of definition of $f(\omega)$ can be extended from the real ω line to the complex $\tilde{\omega}$ plane by setting $\omega = \tilde{\omega} = \omega_r + \mathrm{i}\omega_i$, and the result *may* be an analytic function $f(\tilde{\omega})$, this is not analytic continuation.

Toll's citation is to a 1927 paper presented by Kramers [10] at an International Physics Congress. Although Kramers's derivation is much closer to later derivations, neither he nor Kronig mentioned, even obliquely, causality or the speed of signals. Kramers argued that (ξ, η), the real and imaginary parts of the atomic polarizability, are Hilbert transform pairs (not called such). He ends by asserting that 'it would be interesting to know… [if] the real and imaginary parts of $\varepsilon - 1$ $[\chi]$… are still connected in the same way as ξ and η' for a medium denser than a gas. In 1927, equations (3) and (4) were still a gleam in Kramers's eye.

Kronig and Kramers obtained relations of general validity by way of specific models of a polarizable atomic gas because their point of departure was Sellmeier's equation (i.e. the frequency response of an undamped harmonic oscillator), which contains the quotient $1/(\Omega^2 - \omega^2)$ appearing in dispersion relations.

In 1942, Kronig [11] published a paper in a Dutch-language journal in which he derived equations (3) and (4) without invoking a specific model of matter. He recognized that $\chi(\omega)$ is the Fourier transform of a polarization response function and invoked causality by assuming that this function vanishes for negative time, because of which the inverse cosine transform of $\chi'(\omega)$ is equal to the inverse sine transform of $\chi''(\omega)$. He did not evaluate contour integrals and hence his integrals are not denoted explicitly as Cauchy principal values. The only slightly tricky step (which he omitted) is showing that

$$\lim_{t\to\infty} \int_0^\infty f(\Omega) \frac{\cos(\Omega \pm \omega)t}{\Omega \pm \omega}\, \mathrm{d}\Omega = 0, \tag{8}$$

which follows from changing the variable of integration to $\mu/t \mp \omega$. His derivation is as clean and compact as any to be found. But he makes neither explicit mention of causality nor implicit mention of signal propagation speed.

Kronig himself had a hand in distorting the history of the Kramers–Kronig relations. In a 1936 paper with Gorter [12] the cited source for equation (3) is Kronig's 1926 paper in which only equation (6) is derived, and equation (4) is said to have been 'first derived by Kramers', whereas equations (3) and (4) were still speculations in his 1927 paper.

Acknowledgments

I am grateful to Gail Brown Bayler for considerable help with the translation of Kramers's paper, to Jan Tobochnik for encouraging me to provide 'physical insight' into Kramers–Kronig relations, to Akhlesh Lakhtakia for comments and suggestions, and to Hans Verlinde for a rough and ready translation of Kronig's 1942 paper.

References

[1] Stenner M D, Gauthier D J and Nelfeld M A 2003 The speed of information in a 'fast-light' medium *Nature* **425** 695–8
[2] Hu B Y-K 1989 Kramers–Kronig in two lines *Am. J. Phys.* **57** 821
[3] Sharnoff M 1964 Validity conditions for the Kramers–Kronig relations *Am. J. Phys.* **32** 40–4
[4] King F W 2006 Alternative approach to the derivation of dispersion relations for optical constants *J. Phys. A: Math. Gen.* **39** 10427–35
[5] Kronig R de L 1926 On the theory of dispersion of x-rays *J. Opt. Soc. Am. Rev. Sci. Instrum.* **12** 547–57
[6] Altarelli M, Dexter D L, Nussenzveig H M and Smith D Y 1972 Superconvergence and sum rules for the optical constants *Phys. Rev.* B **6** 4502–9
[7] Smith D Y 1998 Dispersion theory, sum rules, and their application to the analysis of optical data *Handbook of Optical Constants of Solids I* ed E D Palik (San Diego: Academic) pp 35–68
[8] Nussenzveig H M 1972 *Causality and Dispersion Relations* (New York: Academic) p 45
[9] Toll J S 1956 Causality and the dispersion relation: logical foundations *Phys. Rev.* **104** 1760–70

[10] Kramers H A 1927 La diffusion de la lumiére par les atomes *Atti del Congresso Internazionale dei Fisici, Como-Pavia-Roma,* Vol. 2 (Bologna: Nicola Zanichelli) pp 545–57 (English translation in Ter Haar D, 1998 Master of Modern Physics: The Scientific Contributions of H. A. Kramers (Princeton University Press) Appendix D)

[11] Kronig R de L 1942 Algemeene theorie der diëlectrische en magnetische verliezen (General theory of dielectric and magnetic losses) *Ned. Tijdschr. Natuurkd.* **9** 402–9

[12] Gorter C J and Kronig R de L 1942 On the theory of absorption and dispersion in paramagnetic and dielectric media *Physica* **3** 1009–20