Broadband electromagnetic scattering by particles

Michael I. Mishchenko

NASA Goddard Institute for Space Studies, 2880 Broadway, New York, New York 10025, USA
(mmishchenko@giss.nasa.gov)

Received July 17, 2008; accepted September 18, 2008;
posted September 24, 2008 (Doc. ID 98905); published November 4, 2008

The so-called sum rule for the extinction cross section has been the subject of several publications. However, it has not been obtained directly from the macroscopic Maxwell equations but rather follows from heuristic causality considerations. It is argued that these causality considerations are, in fact, questionable and do not follow from the fundamental concept of electromagnetic scattering by a particle. Therefore, the resulting sum rule should be considered an unproven hypothesis rather than an outcome of a rigorous derivation from first principles.

OCIS codes: 030.5620, 260.2110, 260.5430, 280.1310, 290.4210, 290.5850.

1. INTRODUCTION

Purcell [1] appears to be the first to introduce the so-called sum rule for the extinction cross section by expressing the integral of the extinction cross section over all wavelengths in terms of a simple product of the particle volume and a coefficient depending on the particle shape and static dielectric function. This result was rederived in [2] and generalized in [3,4] (see also [5–12]).

Unfortunately, the final results of [1–4] have not been obtained directly from the macroscopic Maxwell equations but rather are based on heuristic causality considerations. In the following I argue that the causality arguments used in [1–4] are, in fact, quite questionable and are inconsistent with the fundamental concept of electromagnetic scattering by particles. The main objective of this paper is not to disprove the extinction sum rule but rather to demonstrate that the existing derivations of this mathematical theorem are inadequate.

2. ANALYSIS

The derivations presented in [1–4] essentially use as a template the Kramers–Kronig relations. The latter are a direct consequence of postulating appropriate constitutive relations, which state that the electric displacement $\mathbf{D}$, the magnetic induction $\mathbf{B}$, and the macroscopic charge density $\mathbf{J}$ at a point $\mathbf{r}$ and at a moment t depend on the values of the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{H}$, at this same point but over the entire time interval $(-\infty, t]$. Note that $\mathbf{D}$, $\mathbf{B}$, $\mathbf{J}$, $\mathbf{E}$, and $\mathbf{H}$ are actual coexisting physical fields and depend explicitly on time and spatial coordinates.

It is very important to recognize that the constitutive relations are indispensable axioms of Maxwell's macroscopic electromagnetics intended to ensure that the number of unknown fields does not exceed the number of independent equations from which the fields are determined [13,14]. It can then be shown that the set of equations of macroscopic electromagnetics has a unique solution provided that initial and boundary conditions are appropriately formulated. From the perspective of electromagnetic scattering, this solution is fully defined by determining $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ everywhere in space and at all moments in time. One can then use this solution to calculate any observable quantity.

The constitutive relations are often characterized as representing physical causality by expressing an outcome (the local response of the medium) in terms of an income (the local electric and magnetic fields). After this causal relation has been established (in fact, postulated), the temporal evolution of the entire system (field+matter) fully and uniquely follows from the macroscopic Maxwell equations supplemented by appropriate initial and boundary conditions.

Of course, it can also be said that the resulting solution of the Maxwell equations is the causal outcome of the initial conditions. The mathematical expression of this overarching causality can be very complex. However, in the framework of the standard frequency-domain electromagnetics this latter causality becomes tautological because the fields do not change in time except for the time harmonic factor $\exp(-i\omega t)$, where ω is the angular frequency and $i=(-1)^{1/2}$.

Purcell [1] derived the sum rule for the extinction cross section directly from the Kramers–Kronig relations by applying them to the so-called effective dielectric constant of a discrete random medium populated by sparsely and randomly distributed particles. The medium is assumed to be illuminated by a plane electromagnetic wave, and this wave is assumed to be exponentially attenuated inside the random medium due to the fact that the imaginary part of the effective dielectric constant is nonzero owing to the presence of the randomly positioned particles. The finite size of the particles is neglected and the medium is assumed to be continuous.

It should be recognized, however, that the incident plane electromagnetic wave is not attenuated exponentially inside the discrete random medium but rather remains unchanged, as it should be with any incident field (see below). What is attenuated exponentially according to the effective dielectric constant with a nonzero imaginary part is the time-independent so-called coherent field $\mathbf{E}_c(\mathbf{r})$ [15–17]. The latter is obtained by

- writing the total electric field inside the medium as $\mathbf{E}_0(\mathbf{r},t)\exp(-i\omega t)$, where the electric field amplitude $\mathbf{E}_0(\mathbf{r},t)$ is a "slowly varying" function of time provided that significant changes in particle positions occur over time intervals much longer than the period of time-harmonic oscillations $2\pi/\omega$;
- artificially neglecting the time-harmonic factor $\exp(-i\omega t)$;
- expressing the random amplitude $\mathbf{E}_0(\mathbf{r},t)$ as a sum of the time-independent coherent (average) field $\mathbf{E}_c(\mathbf{r})$ and a fluctuating field $\mathbf{E}_f(\mathbf{r},t)$ caused by random changes in particle positions; and
- calculating $\mathbf{E}_c(\mathbf{r})$ as the average of $\mathbf{E}_0(\mathbf{r},t)$ over a time interval long enough to establish full ergodicty of the medium.

We thus see that the coherent field is an artificial mathematical construction rather than an actual time-dependent physical field [17]. In particular, it is not a time-harmonic plane electromagnetic wave. The only reason to even consider this purely mathematical quantity is that it turns out to be useful in the derivation of formulas for certain optical observables in the context of the radiative transfer theory [17].

Therefore, we must conclude that the effective dielectric constant is not a physical dielectric constant appearing in expressions for actual time-dependent physical fields, and there is no reason whatsoever to state that it satisfies the Kramers–Kronig relations. This argument completely negates the derivation of the sum rule in [1].

The derivations in [2–4] are essentially based on an alleged causal relation between the incident and scattered fields formulated as follows: the scattered field cannot precede in time the incident field that excited it. This causality is obviously assumed to exist in the time domain [5], which makes the first unnumbered formula in [4] [or the unnumbered formula between Eqs. (2.1) and (2.2) in [3]] the very basis of the derivations of the extinction sum rule in [2–4]. This formula (hereinafter referred to as SGK) is intended to relate the electric field of the outgoing spherical scattered wave and the electric field of the incident plane wave. Both waves are not, in general, time harmonic.

Let us analyze the alleged causal relation between the incident and scattered fields from the perspective of what electromagnetic scattering actually is in the framework of macroscopic electromagnetics. It should be recognized that the solution of the Maxwell equations is always formulated in terms of the total electric and magnetic fields. Any measurement of electromagnetic scattering is always reduced to the measurement of certain optical observables first in the absence of the scattering object and then in the presence of the object. It is the fact that the two measurements yield, in general, different results that is usually referred to as the "phenomenon of electromagnetic scattering." Therefore, in order to facilitate the theoretical interpretation of such measurements, the Maxwell equations are solved twice. The first solution, $\{\mathbf{E}_1,\mathbf{H}_1\}$, corresponds to the situation with no scattering object, e.g., to free space. The second solution, $\{\mathbf{E}_2,\mathbf{H}_2\}$, corresponds to the situation with a scattering object present and is intentionally sought in the form $\{\mathbf{E}_2=\mathbf{E}_1+\mathbf{E}_3,\ \mathbf{H}_2=\mathbf{H}_1+\mathbf{H}_3\}$, where the fields $\mathbf{E}_3$ and $\mathbf{H}_3$ are required to satisfy certain radiation conditions at infinity. It is understood that the difference between the solutions $\{\mathbf{E}_1,\mathbf{H}_1\}$ and $\{\mathbf{E}_2,\mathbf{H}_2\}$ is caused by the differences in the corresponding initial and boundary conditions.

It is now clear that the fields $\mathbf{E}_3$ and $\mathbf{H}_3$ are obtained by means of a purely mathematical subtraction of electric and magnetic fields corresponding to two quite different physical situations: $\mathbf{E}_3=\mathbf{E}_2-\mathbf{E}_1$ and $\mathbf{H}_3=\mathbf{H}_2-\mathbf{H}_1$. Calling $\{\mathbf{E}_1,\mathbf{H}_1\}$ the "incident field" and $\{\mathbf{E}_3,\mathbf{H}_3\}$ the "scattered field" is a matter of convention and does not mean that the scattered field actually exists and is caused by the incident field. Indeed, there can be no temporal causal relation between two separate solutions of the Maxwell equations $\{\mathbf{E}_1,\mathbf{H}_1\}$ and $\{\mathbf{E}_2,\mathbf{H}_2\}$. Therefore, there can be no temporal causal relation between $\{\mathbf{E}_1,\mathbf{H}_1\}$ and $\{\mathbf{E}_3,\mathbf{H}_3\}$. We thus have to conclude that SGK cannot be a direct and unequivocal consequence of verbally formulated physical causality.

To emphasize the alleged causality of the "scattering process," it is often said that the incident field is transformed into the scattered field upon its interaction with the object (e.g., [18]). However, the very representation of the total field in the presence of the object as the superposition of the incident and scattered fields implies that the incident field remains unchanged rather than is transformed into the scattered field.

3. CONCLUSION

It should be recognized that the extinction cross section appearing in [2–4] is defined in the framework of macroscopic Maxwell's electromagnetics and that the latter is mathematically self-contained and self-sufficient. In other words, after the Maxwell equations and the appropriate constitutive relations have been formulated and supplemented by appropriate initial and boundary conditions, no further postulates are required in order to obtain a unique solution for the electric and magnetic fields at all times and everywhere in space and, thus, to calculate any observable quantity, including the extinction cross section. In particular, all physical causality contained in macroscopic electromagnetics is fully defined by the initial conditions and constitutive relations. There is no need to postulate any additional causality that may, in fact, imply results inconsistent with the direct solution of the Maxwell equations.

Thus the argument for SGK appears to be based on some intuitive picture of a preceding incoming wave, an interaction of this wave with the particle, and a subsequent outgoing scattered wave, thus allowing for the description of the process by two separate, causally related fields. However, this picture is misleading because it suggests that the incident and scattered fields are what we would observe in the physical world. There is, however, only one observable electromagnetic field in the physical world, which is the total field. Its dynamical behavior is fully described by the Maxwell equations and constitutive/boundary conditions that automatically ensure causality. The incident and scattered fields are only formally defined quantities with no direct counterparts in the physical world. It is, therefore, not self-evident that there must exist a causal physical cross correlation between these two constructed fields.

It is, of course, likely that there is a *mathematical* relation between $\{\mathbf{E}_1, \mathbf{H}_1\}$ and $\{\mathbf{E}_3, \mathbf{H}_3\}$, the general time-domain form of which, to the best of my knowledge, is yet to be derived. This relation would be a time-domain extension of the frequency-domain formulas (2.17) and (2.18) in [19] and would require proving mathematically that the Lippmann–Schwinger equation for the time-domain dyad transition operator has a unique solution. It is even conceivable that for some special cases of the incident field this relation may be reduced to SGK. However, the correctness and range of applicability of SGK must be established by deriving it from time-domain macroscopic electromagnetics rather than by postulating it on the basis of alleged causality. Until and unless this has been done, the sum rules presented in [1–4,6–12] should be considered unproven hypotheses rather than an outcome of a rigorous derivation from first principles.

ACKNOWLEDGMENTS

I thank Michael Kahnert and Daniel Mackowski for helpful comments on the initial version of this paper. This research was funded by the NASA Radiation Sciences Program managed by Hal Maring.

REFERENCES

1. E. M. Purcell, “On the absorption and emission of light by interstellar grains,” Astrophys. J. **158**, 433–440 (1969).
2. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, 1983).
3. C. Sohl, M. Gustafsson, and G. Kristensson, “Physical limitations on broadband scattering by heterogeneous obstacles,” J. Phys. A **40**, 11165–11182 (2007).
4. C. Sohl, M. Gustafsson, and G. Kristensson, “Physical limitations on metamaterials: restrictions on scattering and absorption over a frequency interval,” J. Phys. D **40**, 7146–7151 (2007).
5. N. G. van Kampen, “*S*-matrix and causality conditions. I. Maxwell field,” Phys. Rev. **89**, 1072–1079 (1953).
6. B. H. J. McKellar, M. A. Box, and C. F. Bohren, “Sum rules for optical scattering amplitudes,” J. Opt. Soc. Am. **72**, 535–538 (1982).
7. J. C. Ku and J. D. Felske, “Determination of refractive indices of Mie scatterers from Kramers–Kronig analysis of spectral extinction data,” J. Opt. Soc. Am. A **3**, 617–623 (1986).
8. M. A. Box, P. Attard, G. Bryant, and B. H. J. McKellar, “Sum rules for Mie scattering,” J. Opt. Soc. Am. A **4**, 795–799 (1987).
9. Z.-F. Xing and J. M. Greenberg, “An efficient method for the calculation of mean extinction. I. The analyticity of the complex extinction efficiency of homogeneous spheres,” Astrophys. J. **399**, 164–175 (1992).
10. A. J. Sievers, “Extinction sum rule and optical moment for an ellispsoidal particle of arbitrary size,” Opt. Commun. **109**, 71–74 (1994).
11. R. Lai and A. J. Sievers, “Mie computational test of the extinction cross-section sum rules and optical moments for large dielectric spheres and shells,” Opt. Commun. **116**, 72–76 (1995).
12. R. Lai and A. J. Sievers, “Sum rules and optical moments for a coarse scattering medium,” J. Opt. Soc. Am. A **13**, 1036–1045 (1996).
13. J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, 1941).
14. E. J. Rothwell and M. J. Cloud, *Electromagnetics* (CRC, 2001).
15. V. Twersky, “On propagation in random media of discrete scatterers,” Proc. Symp. Appl. Math. **16**, 84–116 (1964).
16. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (IEEE, 1997).
17. M. I. Mishchenko, L. D. Travis, and A. A. Lacis, *Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering* (Cambridge U. Press, 2006).
18. R. Pike and P. Sabatier, *Scattering: Scattering and Inverse Scattering in Pure and Applied Mathematics* (Academic, 2002), p. xi.
19. M. I. Mishchenko, L. D. Travis, and A. A. Lacis, *Scattering, Absorption, and Emission of light by Small Particles* (Cambridge U. Press, 2002) (available in the PDF format at http://www.giss.nasa.gov/~crmim/books.html).