Infinities in molecular quantum electrodynamics and generalized functions

R. Guy Woolley *

School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, United Kingdom

(Received 12 April 2024; accepted 18 June 2024; published 3 July 2024)

The Power-Zienau-Woolley Hamiltonian for the quantum electrodynamics of atoms and molecules is written in terms of purely transverse electromagnetic field variables and so-called polarization fields for the charged particles. It is well known that the attempt at finding solutions to the coupled equations that arise from the Hamiltonian is marred by the occurrence of infinite "self-energies" for both particles and the field. Because of the occurrence of the Dirac δ function in the nonzero Poisson-brackets/commutation relation for the fields, and in the definition of the polarization fields, these variables, classical and quantum, must be identified as distributions, in the mathematical sense. The Schwartz "impossibility theorem" shows that there is no general multiplication rule for distributions, so one has to find a framework that gives meaning to the Hamiltonian The energy of the electric polarization field is analyzed in the Colombeau algebra and shown to be finite; in particular Coulomb's law $(1/r, r > 0)$ with a *finite* self-energy $(r = 0)$ is obtained. How these ideas could be extended to the free-field Hamiltonian is discussed. A *finite* zero-point energy for the electromagnetic field is to be expected. Relevant mathematical results are summarized in an Appendix.

DOI: 10.1103/PhysRevA.110.012204

I. INTRODUCTION

For a closed system of $N \geqslant 1$ spinless charges in a radiation field $(\mathbf{E}^{\perp}, \mathbf{B})$, the general nonrelativistic Hamiltonian for electrodynamics may be written [1,2]

$$\begin{aligned}\mathsf{H}_{\mathbf{P}} = &\sum_{n=1}^{N} \frac{|\mathbf{p}_n|^2}{2m_n} + \frac{1}{2}\epsilon_0 \int (|\mathbf{E}^{\perp}|^2 + c^2|\mathbf{B}|^2)\, \mathrm{d}^3\mathbf{x} \\ &- \int \mathbf{P}\cdot\mathbf{E}^{\perp}\, \mathrm{d}^3\mathbf{x} - \int \mathbf{M}\cdot\mathbf{B}\, \mathrm{d}^3\mathbf{x} + \iint \boldsymbol{\mathcal{X}} : \mathbf{B}\mathbf{B}\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{x}' \\ &+ \frac{1}{2\epsilon_0}\int \mathbf{P}\cdot\mathbf{P}\, \mathrm{d}^3\mathbf{x}. \end{aligned} \tag{1}$$

The Hamiltonian scheme is completed by giving the equal-time Poisson-brackets/commutators of the dynamical variables, which for QED are

$$[\mathsf{E}(\mathbf{x})^r, \mathsf{E}(\mathbf{x}')^s] = [\mathsf{B}(\mathbf{x})^r, \mathsf{B}(\mathbf{x}')^s] = 0, \tag{2}$$

$$\left[\mathsf{q}_m^r, \mathsf{p}_n^s\right] = \mathrm{i}\hbar\delta_{mn}\delta_{rs}, \tag{3}$$

$$[\mathsf{E}(\mathbf{x})^r, \mathsf{B}(\mathbf{x}')^s] = \mathrm{i}\hbar\epsilon_0^{-1}\epsilon_{rst}\nabla^t_{\mathbf{x}'}\delta^3(\mathbf{x}-\mathbf{x}'). \tag{4}$$

In (1) the first term is the total kinetic energy for N free charges, and the second term is the usual Hamiltonian for free radiation. The next three terms couple the charges to the radiation, while the last term has no dependence on the field nor on the particles' motion; it is of a purely static nature. $\mathbf{M}$ is a magnetization density linear in the charge e that involves the particles' position and momentum variables, and $\boldsymbol{\mathcal{X}}$ is a generalized diamagnetic susceptibility tensor that is proportional to e^2. Their particular forms depend on the choice made for the electric polarization field $\mathbf{P}$ which is also linear in the charge e. The polarization field, like the vector potential $\mathbf{a}$, is essentially an arbitrary working variable inasmuch that one cannot specify an experimental setup that corresponds to some definite choice of $\mathbf{P}$ (or $\mathbf{a}$).

Routine calculation yields the equations of motion as the Maxwell equations for the fields associated with the polarization fields $(\mathbf{P}, \mathbf{M})$, and the Lorentz force law for the particle motion in the fields $(\mathbf{E}, \mathbf{B})$ [2]. Of course, these must be solved in a self-consistent manner for the closed system, and one learns from the conventional calculations that both classical and quantum formulations lead to infinite quantities, which physically is a nonsense. We explore here some ideas about the origin of the infinities which can be traced to invalid assumptions in the calculations. In order to make this account reasonably self-contained, some relevant mathematical results are collected together in the Appendix. Proofs can be found in the cited mathematical literature.

Some results for specific choices for the polarization field $\mathbf{P}$ were discussed in [1,2], and, in particular, the last term in the Hamiltonian

$$\mathcal{E}_{\mathbf{P}} = \frac{1}{2\epsilon_0}\int \mathbf{P}\cdot\mathbf{P}\, \mathrm{d}^3\mathbf{x} \tag{5}$$

was examined. For an overall neutral multicharge system (atom, molecule), $\mathbf{P}$ may be written as a sum of line integrals with endpoints corresponding to the particle position variables, so it suffices to examine a typical term (see Sec. II)

$$\mathbf{P} = e\int_{\mathbf{X}_1}^{\mathbf{X}_2} \delta(.;\mathbf{z})\, \mathrm{d}\mathbf{z}. \tag{6}$$

Although it was noted that the components of the vector $\mathbf{P}$ in the point particle model were distributions, the subsequent calculations were performed as though they were continuous

*Contact author: faps1995@cantab.net

 ©2024 American Physical Society

functions; $\mathcal{E}_{\mathbf{P}}$ was found to have δ function and explicitly divergent contributions,

$$\mathcal{E}_{\mathbf{P}} \sim \frac{e^2}{4\pi\epsilon_0}\delta^2(0)\int \mathrm{d}l + \cdots, \tag{7}$$

where l is the arc length along the integration path, and $\delta^2(0)$ is the singular spatial δ function in two dimensions evaluated at the origin with dimension L^{-2} [1].

Distributions are a particular type of "generalized function" produced by a construction which allows one to work with irregular functions in familiar ways; for example addition and differentiation are defined. They are linear functionals on the Schwartz space S and belong to the vector space S', the dual of S [3].

We examine first (Sec. II) some of the implications of the distributional nature of the polarization field; we concentrate on $\mathbf{P}$ though similar concerns also apply to the magnetic variables ($\mathbf{M}$, $\mathcal{X}$) derived from it. Conventionally the properties of $\mathbf{P}$ are determined by a vector quantity, $\mathbf{g}$, which is a Green's function or fundamental solution of the divergence equation [1]. However, once one recognizes that the Green's function is really a distribution $\mathbf{g}$ with which $\mathbf{g}$ is associated it becomes clear why the calculation of $\mathcal{E}_{\mathbf{P}}$ in the usual manner is problematic.

This is because the Schwartz "impossibility theorem" [4,5] shows that in general distributions cannot be multiplied unambiguously, unlike the continuous functions we are used to; thus there is the obvious question: how should $\mathcal{E}_{\mathbf{P}}$, (5), be understood? The Schwartz theorem can be evaded if the factors u, v in the distributional product $(u * v)$ satisfy certain regularity conditions. For example, Hörmander gave a criterion in terms of the "wave front sets" of the two distributions. Roughly speaking when the Fourier transform of a factor, u, around any point does not decay exponentially in the direction of a particular wave vector (an element of its wave-front set), the Fourier transform of the other factor, v, *must* decay exponentially in the opposite direction of that wave vector [6]. This key insight is used to regularize the singular Feynman diagrams in the usual perturbation theory approach to Lorentz invariant QED and to carry through the renormalization program [7–9]. However, the evaluation of $\mathcal{E}_{\mathbf{P}}$ is not amenable to this approach since (7) results from an integral over a product of δ functions if ordinary pointwise multiplication is used [see (A40)].

There are ways of extending irregular functions to a larger class of "generalized functions" than the distributions for which a product *is* defined. By changing the natural assumptions Schwartz required for his interdict against multiplication of distributions one can make a modified construction that leads to an algebra of generalized functions. This is the approach to generalized functions of the *Colombeau algebra* which is sketched in the Appendix. The formation of a product of distributions in the Colombeau algebra is unrestricted in that it is not subject to the regularity condition mentioned above. It does, however, involve radically new mathematics since the familiar notion of pointwise product is given up. The construction does not simply regularize a singular quantity by imposing some arbitrary large momentum cutoff to suppress ultraviolet infinities. The whole, usual calculus on continuous functions is moved into an entirely new mathematical framework while keeping many familiar features including, importantly, the notion of *product*. This is enough to show that $\mathcal{E}_{\mathbf{P}}$ is finite, but not to give a definite numerical value. The details are given in several monographs [10–12], and applications, developments, and further literature can be found in [5,13,14].

In Sec. III we look at the polarization field from the perspective of the Colombeau algebra having shown (Sec. II) that the earlier attempt at regularization [1,2] does not satisfy the algebra's conditions. In Sec. IV there are some thoughts about other terms in (1) in the framework of generalized functions.

II. THE POLARIZATION FIELD, P, AND ITS ENERGY

The electric polarization field, $\mathbf{P}$, of N charged particles $\{e_i,\ i = 1, \ldots, N\}$ is any solution of the divergence equation

$$\nabla \cdot \mathbf{P} = -\rho, \tag{8}$$

where ρ is the charge density. For *point* charges at positions $\{\mathbf{X}_i\ i = 1, \ldots, N\}$ it is customary to take ρ to be given by

$$\rho = \sum_i^N e_i \delta^3(.;\mathbf{X}_i), \tag{9}$$

where δ^3 is the three-dimensional Dirac δ function and the $\{\mathbf{X}_i\}$ are parameters. Since (8) is linear we may write

$$\mathbf{P} = \sum_i^N \mathbf{P}_i \tag{10}$$

so that

$$\frac{1}{e_i}\nabla \cdot \mathbf{P}_i = -\delta^3(.;\mathbf{X}_i). \tag{11}$$

This equation is, to within a constant, the defining equation for the Green's function or fundamental solution for the divergence equation,

$$\nabla \cdot \mathbf{g}(.;\mathbf{x}') = -\delta^3(.;\mathbf{x}'), \tag{12}$$

so if we can find a Green's function $\mathbf{g}$ we have

$$\mathbf{P}_i = e_i \mathbf{g}(.;\mathbf{X}_i). \tag{13}$$

Equation (11) is also, to within a constant, the same as Gauss's Law for the electric field intensity, $\mathbf{E}$, in the Maxwell equations, so what is said about $\mathbf{g}$ here also applies to $\mathbf{E}$.

The literature identifies the vector-valued function

$$\mathbf{g}(\mathbf{x};\mathbf{x}')^{\parallel} = \nabla_{\mathbf{x}}\left(\frac{1}{4\pi\,|\mathbf{x}-\mathbf{x}'|}\right) \tag{14}$$

valid for $\mathbf{x} \neq \mathbf{x}'$ as a Green's function for (12); it is not defined when $\mathbf{x}$ and $\mathbf{x}'$ coincide. The solution set of (12) is much more general than purely (14). A transverse vector field defined by $\mathbf{g}(\mathbf{x};\mathbf{x}')^{\perp} = \mathrm{Curl}_{\mathbf{x}}\mathbf{f}(\mathbf{x},\mathbf{x}')$ where $\mathbf{f}$ is any differentiable vector field in the variable $\mathbf{x}$ can be added to $\mathbf{g}^{\parallel}$ along with any solution $\mathbf{g}_0$ of the homogeneous equation associated with (12).

The line integral form

$$\mathbf{g}(\mathbf{x};\mathbf{x}')^{C} = \nabla_{\mathbf{x}}\left(\frac{1}{4\pi\,|\mathbf{x}-\boldsymbol{O}|}\right) + \int_{C_{[O]}}^{\mathbf{x}'} \delta^3(\mathbf{x}-\mathbf{z})\,\mathrm{d}\mathbf{z} \tag{15}$$

for paths $C_{[\boldsymbol{O}]}$ from some origin $\boldsymbol{O}$ to the field point $\mathbf{x}'$ is particularly important in electrodynamics. If the Dirac δ function is multiplied by the unit dyadic and then decomposed into longitudinal and transverse components

$$\delta_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{y}) = \delta^{\parallel}_{\alpha\beta}(\mathbf{x}-\mathbf{y}) + \delta^{\perp}_{\alpha\beta}(\mathbf{x}-\mathbf{y}), \tag{16}$$

(15) becomes, in component form ($\alpha, \beta = 1, 2, 3$),

$$g(\mathbf{x};\mathbf{x}')^C_\alpha = \nabla_{\mathbf{x},\alpha}\left(\frac{1}{4\pi\,|\mathbf{x}-\boldsymbol{O}|}\right) + \int_{C_{[\boldsymbol{O}]}}^{\mathbf{x}'} \delta^{\parallel}_{\alpha\beta}(\mathbf{x}-\mathbf{z})\,\mathrm{d}z_\beta + \int_{C_{[\boldsymbol{O}]}}^{\mathbf{x}'} \delta^{\perp}_{\alpha\beta}(\mathbf{x}-\mathbf{z})\,\mathrm{d}z_\beta. \tag{17}$$

The first two terms combine to give precisely $\mathbf{g}(\mathbf{x};\mathbf{x}')^{\parallel}$, (14), and the third term is purely transverse by construction.

If one chooses $\mathbf{g} = \mathbf{g}^{\parallel}$ the polarization field (10) is that appropriate to electrostatics. When radiation is involved (moving charges) the polarization field may have a transverse component, and this can be accommodated with the use of the line integral form (15) for $\mathbf{g}$. The Coulomb gauge version of QED corresponds to choosing the purely *longitudinal* polarization field. Since the origin $\boldsymbol{O}$ and the choice of path $C_{[\boldsymbol{O}]}$ in (15) are arbitrary, this freedom is an expression of the gauge symmetry of electrodynamics.

A useful simplification for the line integral Green's function follows from the recognition that the arbitrary origin $\boldsymbol{O}$ should not appear in the final result. For an overall neutral system of charges this can be achieved by a reordering of the terms in the charge density so that the limits in every line integral are associated with coordinates of charges, and terms involving $\boldsymbol{O}$ no longer appear [1]. Thus for the neutral two-particle system, the function

$$\mathbf{g}(\mathbf{x};\mathbf{x}';\mathbf{x}'') = \int_{\mathbf{x}'}^{\mathbf{x}''} \delta^3(\mathbf{x}-\mathbf{z})\,\mathrm{d}\mathbf{z} \tag{18}$$

derived directly from the Green's function, and the charge density

$$\rho(\mathbf{x}'-\mathbf{X}) = e\delta^3(\mathbf{x}'-\mathbf{X}_1) - e\delta^3(\mathbf{x}'-\mathbf{X}_2) \tag{19}$$

yields the polarization field in the well-known form (6). For the straight line path from $\mathbf{X}_1$ to $\mathbf{X}_2$ one has [15]

$$\mathbf{P}(\mathbf{x}) = \frac{e\hat{\mathbf{r}}\delta^2(1-\cos(\omega))}{|\mathbf{X}_1-\mathbf{x}|^2}, \quad |\mathbf{X}_1-\mathbf{x}| \leqslant |\mathbf{r}| \tag{20}$$

and otherwise 0 where $\mathbf{r} = \mathbf{X}_2 - \mathbf{X}_1$, and ω is the angle between the vectors $\mathbf{r}$ and $\mathbf{X}_1 - \mathbf{x}$. $\mathbf{P}$ has dimensions of Q/L^2 since $\delta^2(\theta)$ is dimensionless.

Let C_1 and C_2 be two distinct paths from the charge at $\mathbf{X}_1$ to the charge at $\mathbf{X}_2$ with C_2 the straight line path between the two charges so that $C_1 - C_2$ is a closed loop. Then formally

$$\mathbf{P}(\mathbf{x};C_1) = \mathbf{P}(\mathbf{x};C_2) + e\left(\int_{\Sigma_{12}} \nabla_{\mathbf{z}} \wedge \delta^3(\mathbf{z}-\mathbf{x})\,\mathrm{d}\mathbf{S}\right), \tag{21}$$

where Σ_{12} is a surface bounded by the closed path $\mathbf{z}$ formed from C_1 and $\mathbf{C}_2$. The arbitrariness in $\mathbf{P}$ is carried by the surface integral.

Using (10) and (13) the energy $\mathcal{E}_{\mathbf{P}}$ can be put in the form

$$\mathcal{E}_{\mathbf{P}} = \frac{1}{2\epsilon_0}\sum_{i,j}^{N} e_i e_j \int \mathbf{g}(\mathbf{x};\mathbf{X}_i)\cdot\mathbf{g}(\mathbf{x};\mathbf{X}_j)\,\mathrm{d}^3\mathbf{x}, \tag{22}$$

so it suffices to examine the integral

$$J(\mathbf{g})_{\mathbf{x}',\mathbf{x}''} = \int \mathbf{g}(\mathbf{x};\mathbf{x}')\cdot\mathbf{g}(\mathbf{x};\mathbf{x}'')\,\mathrm{d}^3\mathbf{x}. \tag{23}$$

This was done in [1,2] and results in not only the Coulomb interaction between pairs of charges and the usual infinite "self-energies" (with $\mathbf{g} \equiv \mathbf{g}^{\parallel}$) but also a variety of other divergent terms (with the line-integral form).

We now review the formalism above in the light of the Appendix, focusing on the implications of the use of the Dirac δ function in (9) and (12). The δ is *not* a function in the classical sense and has to be understood as a tempered distribution acting on functions belonging to the Schwartz space, $\mathcal{S}$, according to the rule (A32). Since (12) is an equality the Green's function is also a (vector-valued) distribution $\mathfrak{g}$ in the variable $\mathbf{x}$ providing a distributional solution to the differential equation

$$\boldsymbol{\nabla}\cdot\mathfrak{u} = -\delta_{\mathbf{x}'}, \tag{24}$$

the vector $\mathbf{x}'$ being regarded as a parameter. In terms of $\mathfrak{g}$, (24) is explicitly, for all $s \in \mathcal{S}$

$$\langle \boldsymbol{\nabla}\cdot\mathfrak{g}(\cdot;\mathbf{x}');s\rangle = -\langle\delta_{\mathbf{x}';}s\rangle \equiv -s(\mathbf{x}'). \tag{25}$$

The vector quantity in (14) is locally integrable, and so can be associated with a tempered distribution $\mathfrak{g}(\cdot;\mathbf{x}')^{\parallel}$. Similarly, $\mathrm{Curl}_{\mathbf{x}}\mathbf{f}$, for suitable $\mathbf{f}$, will also be locally integrable and hence is associated with a tempered distribution $\mathfrak{g}(\cdot;\mathbf{x}')^{\perp}$, and one can try

$$\mathfrak{g}(\cdot;\mathbf{x}') = \mathfrak{g}(\cdot;\mathbf{x}')^{\parallel} + \mathfrak{g}(\cdot;\mathbf{x}')^{\perp}, \tag{26}$$

where

$$\mathfrak{g}(\cdot;\mathbf{x}')^{\parallel} = \mathbf{T}_{\mathbf{g}(\cdot;\mathbf{x}')^{\parallel}} \tag{27}$$

and similarly for the transverse component. With these associations, (25) can be rewritten as

$$\langle\boldsymbol{\nabla}\cdot\mathbf{T}_{\mathbf{g}(\cdot;\mathbf{x}')};s\rangle = -\int_{\mathbb{R}^3}\mathbf{g}(\mathbf{x};\mathbf{x}')\cdot\boldsymbol{\nabla}s(x)\,\mathrm{d}x = \int_{\mathbb{R}^3}\boldsymbol{\nabla}_{\mathbf{x}}\cdot\mathbf{g}(\mathbf{x};\mathbf{x}')s(x)\,\mathrm{d}x \tag{28}$$

after integration by parts and use of the differentiation rule for distributions, (A28),

$$\langle\boldsymbol{\nabla}\cdot\mathfrak{g}(\cdot;\mathbf{x}');s\rangle = -\langle\mathfrak{g}(\cdot;\mathbf{x}');\boldsymbol{\nabla}s\rangle. \tag{29}$$

In the same way, Eq. (15) becomes

$$\mathfrak{g}(\cdot;\mathbf{x}')^C = \mathfrak{g}(\cdot;\boldsymbol{O})^{\parallel} + \int_{C_{[\boldsymbol{O}]}}^{\mathbf{x}'}\delta_{\mathbf{z}}\,\mathrm{d}\mathbf{z}. \tag{30}$$

This is a distributional solution of (25) because

$$\langle\boldsymbol{\nabla}\cdot\mathfrak{g}(\cdot;\boldsymbol{O})^{\parallel};s\rangle = -s(\boldsymbol{O}), \tag{31}$$

while

$$\begin{aligned}\langle \nabla_{\mathbf{x}} \cdot \int_{C[O]}^{\mathbf{x}'} \delta_{\mathbf{z}}\, \mathrm{d}\mathbf{z}; s\rangle &= -\left\langle \int_{C[O]}^{\mathbf{x}'} \delta_{\mathbf{z}}\, \mathrm{d}\mathbf{z}; \nabla s \right\rangle \\ &= -\int_{C[O]}^{\mathbf{x}'} \mathrm{d}\mathbf{z} \cdot \langle \delta_{\mathbf{z}}; \nabla s\rangle \\ &= -\int_{C[O]}^{\mathbf{x}'} \mathrm{d}\mathbf{z} \cdot \nabla_{\mathbf{z}} s(\mathbf{z}) \\ &= -s(\mathbf{x}') + s(\boldsymbol{O}). \end{aligned} \tag{32}$$

The conclusion then is that the electric polarization field, $\mathbf{P}$, must also be thought of as a distributional solution to (8) defined by

$$\mathfrak{P} = \sum_i^N \mathfrak{P}_i \equiv \sum_i^N e_i \mathrm{g}(.; \mathbf{X}_i). \tag{33}$$

In view of Schwartz's theorem [4] there is no distributional transcription for (5), so what does $\mathcal{E}_{\mathbf{P}}$ mean?

In [1,2] the Fourier transform of $\mathbf{P}$ was "regulated" by a function $\hat{\chi}_a(k)$ obtained by Fourier transformation of some function that "broadens out" the δ functions in the charge density for point particles. Taking the Fourier transform of $\mathbf{g}(\mathbf{x}; \mathbf{x}')$, (14), as

$$\hat{\mathbf{g}}(\mathbf{k}; \mathbf{x}')^{\parallel} = -\mathrm{i}\frac{\mathbf{k}}{k^2} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}'} \tag{34}$$

the idea was to regularize (34) by putting

$$\hat{\mathbf{g}}(\mathbf{k}, \mathbf{x}')^{\parallel} \to \hat{\mathbf{g}}_a(\mathbf{k}, \mathbf{x}')^{\parallel} = -\mathrm{i}\mathbf{k}\frac{\hat{\chi}_a(k)}{k^2} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}'}. \tag{35}$$

$\hat{\chi}_a(k)$ was chosen as a real function and required to have the properties

$$\begin{aligned} &|\hat{\chi}_a(k)| \leqslant 1, \quad \hat{\chi}_a(0) = \hat{\chi}_0(k) = 1, \\ &\int_0^\infty \hat{\chi}_a(k)^2\, \mathrm{d}k = \frac{\pi}{2a}, \quad k = |\mathbf{k}|, \end{aligned} \tag{36}$$

which are such that the unmodified equations can be recovered in the limit $a \to 0$. The regulator in [1,2] was

$$\hat{\chi}_\lambda(k) = \mathrm{e}^{-\lambda k}, \quad \lambda = \frac{a}{\pi}, \quad k = |\mathbf{k}|. \tag{37}$$

From (36) we see that $\hat{\chi}_a$ is square integrable so that

$$\chi_a(|\mathbf{x}|) = \frac{1}{(2\pi)^3} \int \hat{\chi}_a(|\mathbf{k}|) \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}^3\mathbf{k} \tag{38}$$

is defined.

Combining (37) and (38) the integral is elementary with the result

$$\chi_\lambda(x) = \frac{\lambda}{\pi^2} \frac{1}{(x^2 + \lambda^2)^2}, \quad x = |\mathbf{x}|. \tag{39}$$

Although this satisfies [cf. (A36)]

$$\int_{\mathbb{R}^3} \chi_\lambda(x)\, \mathrm{d}x = 1 \tag{40}$$

we see that $\chi_\lambda(x)$ is nonzero on all of $\mathbb{R}^3$ and does not decay sufficiently fast at ∞ to be an element of the Schwartz space, $\mathcal{S}$. We conclude the calculations of $\mathcal{E}_{\mathbf{P}}$ in [1,2] do not resolve the ambiguity in the multiplication of the distribution $\mathfrak{P}$ required for the evaluation of $\mathcal{E}_{\mathbf{P}}$.

III. THE POLARIZATION FIELD REVISITED

The Colombeau framework offers a way to regularize the Green's function, $\mathbf{g}$, discussed in Sec. II and eliminate the infinities. The aim is to find a replacement for the scalar product

$$J_{\mathbf{x}'\mathbf{x}''} = \int \mathbf{g}(\mathbf{x}; \mathbf{x}') \cdot \mathbf{g}(\mathbf{x}; \mathbf{x}'')\, \mathrm{d}^3\mathbf{x}. \tag{41}$$

As before the substitutions $\mathbf{x}' \to \mathbf{X}_i$, $\mathbf{x}'' \to \mathbf{X}_j$ in J when multiplied by $e_i e_j/2\epsilon_0$ gives $\mathcal{E}_{\mathbf{P}}$ for two charges at $\mathbf{X}_i$, $\mathbf{X}_j$.

Instead of working with

$$\mathbf{g}(\mathbf{x}; \mathbf{x}') = \frac{1}{(2\pi)^3} \int \mathbf{g}(\mathbf{k}; \mathbf{x}') \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}^3\mathbf{k} \tag{42}$$

as in Sec. II, we view it as a tempered distribution in the variable $\mathbf{x}$ and transform it into a generalized function $\mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}')$. This amounts to choosing a Schwartz function according to the requirements discussed in the Appendix. We denote the mollifier as s, and the main conditions are that its Fourier transform $\hat{s}$ should be smooth and compact, that is, $\hat{s} \in \mathcal{D}$, with $\hat{s}(k) = 1$ in a finite neighborhood of $k = 0$.

A representative of $\mathbf{g}$ can be constructed by convolution with a suitable mollifier

$$\begin{aligned} \mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}') &= \mathbf{g}(\mathbf{x}; \mathbf{x}') * s_\lambda \\ &= \frac{1}{\lambda^3} \int s\left(\frac{\mathbf{z} - \mathbf{x}}{\lambda}\right) \mathbf{g}(\mathbf{z}; \mathbf{x}')\, \mathrm{d}^3\mathbf{z} \\ &= \int s(\mathbf{u}) \mathbf{g}(\mathbf{x} + \lambda\mathbf{u}; \mathbf{x}')\, \mathrm{d}^3\mathbf{u} \end{aligned} \tag{43}$$

after an obvious change of integration variable. Combining (42) and (43) and performing the integration over $\mathbf{u}$ (a Fourier transform) yields

$$\mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}') = \frac{1}{(2\pi)^3} \int \mathbf{g}(\mathbf{k}; \mathbf{x}') \hat{s}(\lambda\mathbf{k}) \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}^3\mathbf{k}. \tag{44}$$

This is an embedding of the vector $\mathbf{g}$ in the Colombeau set $\mathcal{G}$. It is possible to show that all moderate embeddings (43) are equivalent modulo infinitesimal quantities that depend on λ (see Proposition 2.1.2 in [14]), so that there is no single choice of s here.

In terms of the generalized vector function R the integrand in (41) is replaced by the density

$$\mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}') \cdot \mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}''), \tag{45}$$

and $J_{\mathbf{x}'\mathbf{x}''}$ is the integral of this density over the whole of $\mathbb{R}^3$. At this point we follow [14] and perform this integration with the inclusion of a "damper" in the integrand, setting

$$J(s)^\lambda_{\mathbf{x}'\mathbf{x}''} = \int \mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}') \cdot \mathrm{R}(\mathbf{g}; s_\lambda; \mathbf{x}; \mathbf{x}'') \hat{\psi}(\lambda\mathbf{x})\, \mathrm{d}^3\mathbf{x}, \tag{46}$$

where $\hat{\psi}(\mathbf{x})$ is a smooth real Schwartz function on $\mathbb{R}^3$ with compact support ($\hat{\psi} \in \mathcal{D}$), and identical to 1 in a 0-neighborhood. The physical interpretation of the inclusion of the damper is that it has the same effect as restricting the integration over the spatial coordinates to a large but finite

volume [14]. Such a restriction is understood tacitly for (1) since otherwise the integrals would not be finite.

Written out in full (46) is

$$J(s)^{\lambda}_{\mathbf{x}'\mathbf{x}''} = \frac{1}{(2\pi)^6}\iiint \mathbf{g}(\mathbf{k}_1;\mathbf{x}')\cdot\mathbf{g}(\mathbf{k}_2;\mathbf{x}'')\hat{s}(\lambda\mathbf{k}_1)\hat{s}(\lambda\mathbf{k}_2) \times \mathrm{e}^{-\mathrm{i}(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{x}}\hat{\psi}(\lambda\mathbf{x})\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{k}_1\,\mathrm{d}^3\mathbf{k}_2. \tag{47}$$

Because of the presence of the mollifiers/damper the integrations can be done in any order, so we choose to integrate over $\mathbf{x}$ first. By inverse Fourier transformation we have

$$\frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{x}\mathrm{e}^{-\mathrm{i}(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{x}}\hat{\psi}(\lambda\mathbf{x}) = \frac{1}{\lambda^3}\psi\left(\frac{\mathbf{k}_1+\mathbf{k}_2}{\lambda}\right). \tag{48}$$

If we were to put $\lambda = 0$ the function ψ would reduce to a Dirac δ function and one of the wave vector integrations would be trivial, ($\mathbf{k}_1 = -\mathbf{k}_2$), For $\lambda > 0$ this simplifying result is not immediate, but with the aid of a proposition in [14] it can be recovered [its basis lies in (A47)].

Taking account of (48) we now write (47) in the form

$$J(s)^{\lambda}_{\mathbf{x}'\mathbf{x}''} = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}_1\hat{s}(\lambda\mathbf{k}_1)\int \mathrm{d}^3\mathbf{k}_2\hat{s}(\lambda\mathbf{k}_2) \times \mathrm{M}(\mathbf{k}_1,\mathbf{k}_2)\frac{1}{\lambda^3}\psi\left(\frac{\mathbf{k}_1+\mathbf{k}_2}{\lambda}\right) \tag{49}$$

with

$$\mathrm{M}(\mathbf{k}_1,\mathbf{k}_2) = \mathbf{g}(\mathbf{k}_1;\mathbf{x}')\cdot\mathbf{g}(\mathbf{k}_2;\mathbf{x}''). \tag{50}$$

Then by Proposition 1.4.2 in [14] the integration over $\mathbf{k}_2$ for fixed $\mathbf{k}_1$ is

$$\int \mathrm{d}^3\mathbf{k}_2\,\hat{s}(\lambda\mathbf{k}_2)\mathrm{M}(\mathbf{k}_1,\mathbf{k}_2)\frac{1}{\lambda^3}\psi\left(\frac{\mathbf{k}_1+\mathbf{k}_2}{\lambda}\right) = \hat{s}(-\lambda\mathbf{k}_1)\mathrm{M}(\mathbf{k}_1,-\mathbf{k}_1) + O(\lambda^{q+1}) \tag{51}$$

so that $J(s)^{\lambda}_{\mathbf{x}'\mathbf{x}''}$ reduces to

$$J(s)^{\lambda}_{\mathbf{x}'\mathbf{x}''} = \frac{1}{(2\pi)^3}\int \mathbf{g}(\mathbf{k};\mathbf{x}')\cdot\mathbf{g}(-\mathbf{k};\mathbf{x}'')\hat{s}(\lambda\mathbf{k})\hat{s}(-\lambda\mathbf{k})\,\mathrm{d}^3\mathbf{k} \tag{52}$$

with an infinitesimal remainder. As noted in the Appendix, a product of Schwartz functions is another Schwartz function, so we can simplify (52) to

$$J(s)^{\lambda}_{\mathbf{x}'\mathbf{x}''} = \frac{1}{(2\pi)^3}\int \mathbf{g}(\mathbf{k};\mathbf{x}')\cdot\mathbf{g}(-\mathbf{k};\mathbf{x}'')\hat{s}(\lambda\mathbf{k})\,\mathrm{d}^3\mathbf{k} \tag{53}$$

for some test function $\hat{s}(\lambda\mathbf{k}) \in \mathcal{D}$. This result is of the same form as the "regulated" version of these calculations in [1,2]; the crucial difference is that the classical function $\hat{\chi}_\lambda(k)$, defined by point values, is replaced by a family of mollifiers $\{\hat{s}(\lambda\mathbf{k})\}$.

We will assume that $s(x)$ is real and $x = |\mathbf{x}|$, so that its Fourier transform is also radial. The angular integrations are then elementary and one is left with

$$J(s)^{\lambda}_{r} = \frac{1}{2\pi^2}\int_0^\infty h(kr)\hat{s}(k\lambda)\,\mathrm{d}k, \quad r = |x''-x'|. \tag{54}$$

Here the function $h(kr)$ depends on the particular choice of Green's function, $\mathbf{g}(\mathbf{x};\mathbf{x}')$, that was made originally, for example,

$$\begin{aligned}(1):&\quad \mathbf{g} = \mathbf{g}^{\parallel} \to h(kr) = \frac{\sin(kr)}{kr},\\ (2):&\quad \mathbf{g} = \mathbf{g}^{C} \to h(kr) = 2[kr\mathrm{Si}(kr) + \cos(kr) - 1].\end{aligned} \tag{55}$$

In case (1), the purely longitudinal Green's function, we have

$$h(0) = 1, \quad h(x) < 1 \quad \text{for} \quad x > 0 \tag{56}$$

so $J(s)^{\parallel}_{r}$ for $r = 0$ and $\lambda > 0$ has a finite real bound since $\hat{s}$ is enough to ensure convergence of the integral. Hence the calculation of $\mathcal{E}_{\mathbf{P}}$, which simply involves interpreting r as $|\mathbf{X}_2 - \mathbf{X}_1|$ yields a *finite real generalized number* for the self-energy of the charges. On the other hand, for $r \neq 0$ one can take λ as an infinitesimal, i.e., in the neighborhood of 0, and so replace $\hat{s}$ by 1, and then the energy $\mathcal{E}_{\mathbf{P}}$ will appear as the familar Coulombic form

$$\mathcal{E}_{\mathbf{P}} = \frac{e_1 e_2}{4\pi\epsilon_0 r}, \quad r = |\mathbf{X}_2 - \mathbf{X}_1|, \quad \mathbf{X}_2 \neq \mathbf{X}_1 \tag{57}$$

since the remaining integral yields simply $\pi/2r$.

Case (2) is not so straightforward as explicit formulas depend on the particular form of $\hat{s}$. However, it is evident that at $r = 0$, $h = 0$ so the integral vanishes; for $r > 0$ the result is finite since the mollifier ensures that the integral is convergent. Hence when $\mathcal{E}_{\mathbf{P}}$ is calculated one will get an $1/r$ dependence modulated by some r-dependent coefficient. This is due to the introduction of a transverse component for the polarization field; of course one has to remember that $\mathbf{P}^{\perp}$ also enters in the interaction terms in (1). In the conventional perturbation theory one can show a formal cancellation of part of the transverse contributions and one is left with (57), and perhaps that will be true in a more rigorous treatment of the fields in (1) along the lines sketched.

IV. THE HAMILTONIAN

Maxwell's equations for the field of a point charge at rest at the origin of the coordinates reduce to simply Gauss's Law:

$$\epsilon_0\nabla\cdot\mathbf{E} = \rho, \quad \rho(\mathbf{x}) = e\delta^3(\mathbf{x}). \tag{58}$$

The "solution" that vanishes at ∞ is

$$\mathbf{E}(\mathbf{x}) = \frac{e\hat{\mathbf{x}}}{4\pi\epsilon_0 x^2}, \tag{59}$$

which leads to a divergent energy integral

$$\mathcal{E} = \frac{1}{2}\epsilon_0\int \mathbf{E}(\mathbf{x})\cdot\mathbf{E}(\mathbf{x})\,\mathrm{d}^3\mathbf{x} = \infty. \tag{60}$$

As noted earlier, this application of Gauss's law involves the same mathematics as the equation defining the polarization field and is subject to the same critique. The electric field should be viewed as a tempered distribution and the discussion in Sec. III applies verbatim. In the Colombeau algebra the solution of (58) leads to the usual Coulombic expression for $|\mathbf{x}| \neq 0$, and a *finite* generalized number at the origin. As for the particle's dynamics, Hamilton's equations for the charge should be a pair of first-order differential equations for its position and momentum variables; they are, however, pathological since they include a term proportional to the particle's

acceleration, $\ddot{\mathbf{x}}$, which leads to a runaway solution for the orbit [2].

The situation in QED is rather different. In the usual treatment, the transverse electromagnetic field variables $\mathbf{A}$ (the Coulomb gauge vector potential), $\mathbf{E}^{\perp}$, $\mathbf{B}$ are represented as Fourier series derived from the standing waves in a "box" of finite volume Ω [2]. On passing to the continuum limit these quantities satisfy the bracket relation (4). As operator-valued quantities the transverse electric field operator, for example, is given the Fourier expansion

$$\mathbf{E}(\mathbf{x})^{\perp} = \int \left(\hat{\mathbf{E}}(\mathbf{k})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} + \mathrm{H.c.}\right)\mathrm{d}^3\mathbf{k}, \tag{61}$$

where

$$\hat{\mathbf{E}}^{\perp}(\mathbf{k}) = \mathrm{i}\sqrt{\frac{\hbar kc}{(2\pi)^3 2\epsilon_0}}\mathbf{c}(\mathbf{k}). \tag{62}$$

The vector $\mathbf{c}(\mathbf{k})$ satisfies $\mathbf{k}\cdot\mathbf{c} = 0$ and so can be expressed in terms of components with respect to the usual "polarization" unit vectors that span the plane orthogonal to the propagation direction $\mathbf{k}$. The components are the familiar annihilation, $\mathsf{c}(\mathbf{k})_{\sigma}$, and creation, $\mathsf{c}(\mathbf{k})^{+}_{\sigma}$, operators for a photon with momentum $\mathbf{k}$ and polarization σ, ($\sigma = 1, 2$), with commutator

$$[\mathsf{c}(\mathbf{k})_{\sigma}, \mathsf{c}(\mathbf{k}')^{+}_{\sigma'}] = \delta(\mathbf{k}-\mathbf{k}')\delta_{\sigma\sigma'}. \tag{63}$$

There is a similar expansion for the magnetic field operator $\mathbf{B}$ ($\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$)

$$\hat{\mathbf{B}}(\mathbf{k}) = \mathrm{i}\sqrt{\frac{\hbar k}{(2\pi)^3 2\epsilon_0 c}}\hat{\mathbf{k}}\wedge\mathbf{c}(\mathbf{k}). \tag{64}$$

In QED if either $\mathbf{E}^{\perp}$ or $\mathbf{B}$ is applied to the field's vacuum state $|0\rangle$, the resulting state contains the factor $\sqrt{k}\exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{x})$ for each mode and, with an infinite number of modes, it is not square integrable, so the state does not belong to Fock space; unsurprisingly

$$\langle 0|\mathbf{E}^{\perp}\cdot\mathbf{E}^{\perp}|0\rangle = \infty. \tag{65}$$

In terms of the photon operators the free-field Hamiltonian is [16]

$$\mathsf{H}_0 = \int \hbar ck\left(\mathsf{c}(\mathbf{k})^{+}\mathsf{c}(\mathbf{k}) + \frac{1}{2}\delta^3(0)\right)\mathrm{d}^3\mathbf{k} \tag{66}$$

so that even in the vacuum state the field energy is infinite,

$$\mathsf{H}_0|0\rangle = \frac{1}{2}\hbar c\delta^3(0)\int k\,\mathrm{d}^3\mathbf{k}. \tag{67}$$

The δ function arises from the integration over all space ($\mathbb{R}^3$), and the momentum integral diverges for large k. Although the practical response is simply to wave away the offending infinite contribution, this does not really dispose of the underlying reasons for its occurrence, which manifest themselves again when interactions are introduced. Unlike a classical field, the fluctuations in the quantized field do not die out as length scales are reduced [7] so one cannot assume the field variable to be continuous. In mathematical terms the field is neither absolutely integrable nor square integrable; consequently the conventional Fourier representations (61) and (64) are not defined.

The field variables are, however, locally integrable on any compact subspace of $\mathbb{R}^3$ and so may be viewed as *tempered distributions* in the space variable $\mathbf{x}$. Indeed, the usual response to the difficulties of infinities in electrodynamics is to give up the idea that the fields are continuous vector-valued functions/operators, and reinterpret them as distributions [17]. As we have seen this means "smearing" the field variables with a function belonging to the Schwartz space, $\mathcal{S}$; in the notation of the Appendix this is explicitly

$$\mathbf{E} \to \mathbf{T}_{\mathbf{E}}, \quad s \to \langle \mathbf{T}_{\mathbf{E}}; s\rangle = \int_{\mathbb{R}^3}\mathbf{E}(x)s(x)\,\mathrm{d}x \tag{68}$$

as in (A21), and similarly for the magnetic field $\mathbf{B}$. That such a step is necessary is also evident from the appearance of a "Dirac δ function" in the commutator (4). The Dirac δ is a distribution, and since Eq. (4) is an equality the l.h.s. of the commutator must also involve distributions. Likewise with (63). But this does not solve the problem of giving meaning to the nonlinear terms in the Hamiltonian (1). In order to solve the problem of multiplication they can be embedded in the Colombeau algebra in exactly the same way as $\mathbf{g}$ in (43) and (44), that is, by convolution with a suitable mollifier $s_{\lambda}(\mathbf{x})$ followed by integration over the variable $\mathbf{u}$. The result is that the integrand in (61) is modified simply by inclusion of the factor $\hat{s}(\lambda\mathbf{k})$.

The work in [14] is concerned with translating the formal calculations of a model quantum field theory—the original Heisenberg-Pauli (HP) quantum theory of a scalar boson field—into the Colombeau algebra to start a mathematically rigorous justification for what is done conventionally. The boson field operator and its conjugate are reinterpreted as distributions in the space variable $\mathbf{x}$ and then transformed into elements of the Colombeau algebra. The "free-field" part of the HP Hamiltonian is closely related to the free-field part of the QED Hamiltonian (1) since both involve only quadratic combinations of the field operators; they can be written as Hamiltonian *densities* which when integrated over all space ($\mathbb{R}^3$) gives the Hamiltonian as the energy (i.e., the energy operator when quantized).

V. DISCUSSION

The nonrelativistic Hamiltonian (1) is of interest in its own right; for example, the question as to whether it has a ground state, which cannot be answered by perturbation theories, is important for understanding the stability of matter [17]. Besides there is no Lorentz invariant account of atoms, molecules, condensed matter, etc., interacting with the electromagnetic field, so (1) cannot be viewed simply as some limit of the Lorentz invariant theory of electrons and photons.

The infinities found in nonrelativistic quantum electrodynamics based on the Hamiltonian (1) are due to the neglect of the true mathematical nature of the field operators (electromagnetic and matter polarization) that it is formulated in terms of. Since the familiar Coulomb gauge form is simply a special case of (1) this remark is quite general. Even without considering interactions there is the infinite zero-point energy of the free electromagnetic field. At nonrelativistic energies the electrons and nuclei appear to have no structure, and it is natural to describe them as "pointlike." Having said that there

is something paradoxical about associating mass to entities that have no extension in space. This tension manifests itself in the appearance of the Dirac "δ function" in the formalism, an object that needs to be handled with great care. Calculation treating the operators as ordinary continuous functions in the point-particle model leads to an infinite "electromagnetic mass," and to the problems for the energy $\mathcal{E}_{\mathsf{P}}$ discussed in Sec. II [2].

In practice one uses the observed values for the charge and mass parameters of the particles (a necessary step to be sure) and simply drops the troublesome infinite terms as they arise in perturbation theory. Nonperturbative analysis of nonrelativistic QED is almost invariably based on the Coulomb gauge Hamiltonian, and divergent momentum integrals are simply cut off to maintain nonrelativistic energies and ensure that the Hamiltonian is self-adjoint [2,17].

The Colombeau algebra discussed here offers a means to address these foundational problems at the expense of an unfamiliar mathematical framework. The zero-point energy of the free HP Hamiltonian was shown to be finite in [14], and one can reasonably expect the same result for the free quantized electromagnetic field since it is also quadratic in the field operators; the self-interactions of charged particles could also be expected to be finite. This is only a qualitative result, however; these calculations do not lead to a precise (ordinary) numerical value for the self-energies because there is no unique Schwartz function to be used in integrals like (43). The nonrelativistic perturbation theory can be approached via a diagrammatic technique and perhaps divergent terms could be dealt with using the Hörmander criterion; that has yet to be done.

One further unresolved question that involves the distributional nature of the polarization field should be mentioned. The Hamiltonian (1) is related to the familiar Coulomb gauge Hamiltonian by the Power-Zienau-Woolley transformation with the operator

$$\Lambda_{\mathrm{pzw}} = \exp\left(-\frac{\mathrm{i}}{\hbar}\int \mathbf{P}(\mathbf{x})\cdot\mathbf{A}(\mathbf{x})\,\mathrm{d}^3\mathbf{x}\right), \tag{69}$$

where $\mathbf{A}$ is the Coulomb gauge vector potential. Power and Zienau noted that the transformation could be thought of as a redefinition of the modes of the field that incorporates the atomic system as a whole as a source of the field [18]. One can view the PZW transformation as a coherent state boson translation which, for any polarization field, creates a corresponding Fock space from the original Fock space of the Coulomb gauge theory. The resulting coherent state operators involve a mixture of the original particle and field variables and make sense only for the interacting system. The integration in (69) has always been interpreted using the usual product for continuous functions. For point charges and using (6) one finds that the transformed space and the original Fock space have orthogonal vacuum states,

$$\langle 0|\Lambda_{\mathrm{pzw}}|0\rangle = 0. \tag{70}$$

This imples that unitarity is lost in this limiting case [1,2]. The underlying reason must be that the polarization field is a distribution, and the vector potential does not belong to the Schwartz space. How this should be resolved is an open question.

ACKNOWLEDGMENT

I am extremely grateful to Prof. Thierry Jecko for all the advice on mathematical questions that he has freely offered.

APPENDIX: FUNCTIONS, FUNCTIONALS AND GENERALIZED FUNCTIONS

Informally, a classical *function* f is a relation that takes an input (the "argument"), commonly denoted by x, and outputs a value $f(x)$, expressed symbolically as

$$y = f(x) \text{ or } x \mapsto f(x) \tag{A1}$$

so that the pair (x, y) [or $(x, f(x))$] belongs to the set of pairs defining f. The *domain* of a function f is the set of all values for which the function is defined; the *range* is the set of values $f(x)$, that is, the set of output values. If x_i is in the domain of f so that $f(x_i)$ is defined and equal to $\lim_{x\to x_i} f(x)$, f is *continuous* at $x = x_i$. Such functions belong to a set $\mathcal{C}$ (in the usual terminology). Functions that have (partial) derivatives of all orders are said to be *smooth*; they belong to a subset of $\mathcal{C}$ denoted $\mathcal{C}^\infty$, and one has the obvious inclusion $\mathcal{C}^\infty \subset \mathcal{C}$. Smooth functions are particularly important in the discussion here. Sometimes there are points x_i outside the domain of f such that the values $f(x)$ of f tend to ∞, as x tends to x_i. This situation is often realized in physically interesting cases; the electrostatic potentials of a point charge ($\propto 1/|\mathbf{x}|$) and a point dipole ($\propto 1/|\mathbf{x}|^2$) which are infinite at the origin, $x_i = 0$, are important examples in electromagnetic theory.

A function belongs to the space of absolutely integrable functions, $L^1(\mathbb{R}^n)$, if

$$\int_{\mathbb{R}^n} |f(x)|\,\mathrm{d}x \;<\; \infty \tag{A2}$$

and to the space of square integrable functions, $L^2(\mathbb{R}^n)$, if

$$\int_{\mathbb{R}^n} |f(x)|^2\,\mathrm{d}x < \infty. \tag{A3}$$

The classical theory of the Fourier transform shows that for functions that satisfy (A2) the transform is defined properly by

$$(\mathrm{F}f)(x) \mapsto \hat{f}(k) = \int_{\mathbb{R}^n} f(x)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x. \tag{A4}$$

If $\hat{f}$ is absolutely integrable, the inverse transform F^{-1} is defined to be

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}(k)\mathrm{e}^{-\mathrm{i}xk}\,\mathrm{d}k. \tag{A5}$$

If f is also square integrable the inverse transform always exists, a fact that is very important in quantum mechanics where the Fourier transform expresses the unitary transformation connecting position and momentum representations. Heisenberg's Uncertainty Principle is a special case of a general rule that the more localized one member of a Fourier transform pair $(f, \hat{f})$ is, the more spread out its transform partner is. For functions that satisfy (A3), but not (A2), the Fourier transform and its inverse are still defined formally by (A4) and (A5) though only via an indirect approach (Parseval-Plancherel theorem) since in this case (A4) is not a convergent

integral. The *convolution* of two functions s, χ is defined as

$$(s * \chi)(x) = \int_{\mathbb{R}^n} s(y-x)\chi(y)\,\mathrm{d}y; \tag{A6}$$

the convolution theorem states that the Fourier transform of the l.h.s. satisfies

$$\mathrm{F}(s * \chi) = (\mathrm{F}s)(\mathrm{F}\chi). \tag{A7}$$

In much broader mathematical settings the input in (A1) may itself be a function f, and we write

$$F \mapsto F[f] \text{ or } f \mapsto \langle F, f\rangle; \tag{A8}$$

we then speak of F as a *functional*. The definite integral, F, of a continuous function f over some region Ω where f is defined

$$\langle F, f\rangle = \int_\Omega f(\omega)\,\mathrm{d}\omega \tag{A9}$$

is such a functional relation; in this case it is a *linear* one.

The reason why these "abstract" details are of interest is that there are familiar physical quantities that do not satisfy either of the conditions (A2) or (A3) such as the Coulomb potential $1/|\mathbf{x}|$ with $n = 3$ mentioned earlier. One has only the more limited statement that it is *locally integrable*,

$$\int_\Omega \frac{1}{|\mathbf{x}|}\,\mathrm{d}x < \infty, \tag{A10}$$

where the integral is taken over any compact subset, Ω, in $\mathbb{R}^3$. An elementary calculation shows the difficulty; if we define the Coulomb potential as

$$V(\mathbf{x}) = \frac{1}{|\mathbf{x}|}, \tag{A11}$$

its Fourier transform should be

$$\hat{V}(\mathbf{k}) = \int \frac{1}{|\mathbf{x}|}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}^3\mathbf{x} \tag{A12}$$

in accordance with (A4). Using polar coordinates

$$\mathrm{d}^3\mathbf{x} = x^2\,\mathrm{d}x\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi, \tag{A13}$$

the angular integration is immediate and we are left with

$$\begin{aligned}\hat{V}(k) &= \lim_{R\to\infty} 4\pi\int_0^R \frac{\sin(kx)}{kx}x\,\mathrm{d}x \\ &= \frac{4\pi}{|\mathbf{k}|^2}[1-\cos(|\mathbf{k}|R)]\Big|_{R\to\infty};\end{aligned} \tag{A14}$$

this limit does not exist.

A widely used "remedy" is to express the Coulomb potential, V, as a limit of the Yukawa potential which does have a classical Fourier transform,

$$\int \frac{\mathrm{e}^{-m|\mathbf{x}|}}{|\mathbf{x}|}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}^3\mathbf{x} = \frac{4\pi}{|\mathbf{k}|^2+m^2}; \tag{A15}$$

the $m \to 0$ limit of this integral gives just $4\pi/|\mathbf{k}|^2$, which is taken to be the transform of the Coulomb potential. Notice, however, a sleight of hand here; obviously the Coulomb potential is related to the Yukawa potential by

$$\lim_{m\to 0}\frac{\mathrm{e}^{-m|\mathbf{x}|}}{|\mathbf{x}|} = \frac{1}{|\mathbf{x}|} = V(\mathbf{x}). \tag{A16}$$

However,

$$\begin{aligned}&\lim_{m\to 0}\int \frac{\mathrm{e}^{-m|\mathbf{x}|}}{|\mathbf{x}|}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}^3\mathbf{x} = \frac{4\pi}{|\mathbf{k}|^2}, \\ &\int \lim_{m\to 0}\frac{\mathrm{e}^{-m|\mathbf{x}|}}{|\mathbf{x}|}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}^3\mathbf{x} \to ?,\end{aligned} \tag{A17}$$

since the second integral obtained by reversing the order of limit and integration has no meaning as seen in (A14)! What then to do?

The *Schwartz space*, $S(\mathbb{R}^n)$, is the vector space of smooth, complex-valued functions on $\mathbb{R}^n$ that, together with all their derivatives, decrease at infinity faster than any polynomial; the (n-dimensional) function $f(x) = g(x)e^{-x^2}$ for any polynomial $g(x)$ is an example. More precisely this means that such a function f has the property that for any partial derivative ∂^α and any integer m, there is a constant $C_{\alpha,m} > 0$ such that for all x

$$(1+|x|)^m|(\partial^\alpha f)(x)| \leqslant C_{\alpha,m}. \tag{A18}$$

We denote a general element of S by $s = s(x)$; they are essential elements of the mathematical theory of "generalized functions." From their definition we see that Schwartz functions satisfy (A2); hence their Fourier transforms $\{\hat{s}\}$ are well defined, and they also are elements of $\mathcal{S}$. The same is true for the sum, product, and convolution of two Schwartz functions. The subset of elements, $\{x \mapsto \phi(x)\}$, of $\mathcal{S}$ that are not only smooth but also have *compact support* are usually referred to as "test functions." This subset is also a vector space that is usually denoted $\mathcal{D}$.

A *distribution* is a particular type of linear functional and can be thought of as a development of the classical idea of a function summarized above. Instead of f "acting on" point values to give an outcome, a distribution "acts on" elements of a set of functions [3]. A *regular* distribution is a continuous, linear functional on the set of smooth functions with compact support, that is, the elements $\{\phi\}$ of the set $\mathcal{D}$; it is associated with a locally integrable function f according to the pairing formula

$$\phi \mapsto \langle \mathbf{T}_f; \phi\rangle = \int_{\mathbb{R}^n} f(x)\phi(x)\,\mathrm{d}x. \tag{A19}$$

In general the result of such an integration is a complex number. The collection of continuous, linear functionals on $\mathcal{D}$ is a vector space denoted $\mathcal{D}'$. A simpler notation which maintains the distinction between a distribution and the function it is associated with is to write the distribution in a different typeface so that, for example, using fraktur

$$\mathfrak{f} \equiv \mathbf{T}_f. \tag{A20}$$

A *tempered* distribution may be defined in the same way as (A19) except that the integrand involves the elements $\{s\}$ of the Schwartz space $\mathcal{S}$,

$$s \mapsto \langle \mathbf{T}_f; s\rangle = \int_{\mathbb{R}^n} f(x)s(x)\,\mathrm{d}x. \tag{A21}$$

Such functionals belong to a vector space denoted $\mathcal{S}'$. The tempered distributions are important because they provide the basis for the modern account of the Fourier transform for functions that fall outside the classical definition described

earlier. The Fourier transform of a tempered distribution $\mathbf{T}$ is another tempered distribution, $\hat{\mathbf{T}}$, acting on s as

$$\langle \hat{\mathbf{T}}, s\rangle = \langle \mathbf{T}, \hat{s}\rangle, \tag{A22}$$

where $\hat{s}$ is the usual classical Fourier transform of s.

The framework expressed by (A22) can be applied to give a meaning to the Fourier transform of the Coulomb potential. In $\mathbb{R}^3$ the Coulomb potential, (A11) is locally integrable and bounded at infinity, so we may view it, and its Fourier transform, as tempered distributions, specifically

$$\langle \mathbf{T}_V, s\rangle = \int \frac{1}{|\mathbf{x}|} s(x)\, \mathrm{d}^3\mathbf{x}. \tag{A23}$$

From (A22) we have

$$\langle \widehat{\mathbf{T}_V}, s\rangle = \int_{\mathbb{R}^3} \frac{1}{|\mathbf{x}|}\hat{s}(x)\,\mathrm{d}x = \lim_{R\to\infty}\int_{|\mathbf{x}|<R} \frac{1}{|\mathbf{x}|}\hat{s}(x)\,\mathrm{d}x. \tag{A24}$$

Now for $R > 0$

$$\begin{aligned}\int_{|\mathbf{x}|<R} \frac{1}{|\mathbf{x}|}\hat{s}(x)\,\mathrm{d}x &= \int_{|\mathbf{x}|<R}\frac{1}{|\mathbf{x}|}\int \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\mathbf{k}} s(\mathbf{k})\,\mathrm{d}^3\mathbf{k}\,\mathrm{d}^3\mathbf{x}\\ &= \int \mathrm{d}^3\mathbf{k}\, s(\mathbf{k})\int_{|\mathbf{x}|<R}\frac{1}{|\mathbf{x}|}\mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\mathbf{k}}\,\mathrm{d}^3\mathbf{x}\\ &= \int \mathrm{d}^3\mathbf{k}\, s(\mathbf{k})\frac{4\pi}{|\mathbf{k}|^2}[1-\cos(|\mathbf{k}|R)]\end{aligned} \tag{A25}$$

from (A14). The difference from the earlier calculation is that the integral involving the cosine factor in (A25) vanishes for $R\to\infty$, as follows from the properties of s and an integration by parts. Hence

$$\int_{\mathbb{R}^3} \widehat{\mathbf{T}_V} s(x)\,\mathrm{d}x = \int_{\mathbb{R}^3}\frac{4\pi}{|\mathbf{k}|^2}s(k)\,\mathrm{d}k, \tag{A26}$$

where now the Fourier transform of the Coulomb potential is understood as the distribution associated with the function $k\mapsto 4\pi/|\mathbf{k}|^2$.

The convolution of a tempered distribution with a function $s\in\mathcal{S}$ is given by

$$s' * \langle \mathbf{T}, s\rangle = \langle \mathbf{T}, \tilde{s}' * s\,\rangle, \tag{A27}$$

where $\tilde{s}'(x) = s'(-x)$ is the reflection of s' about the origin. The *distributional derivative* is defined in similar fashion by passing the differentiation through to the Schwartz function; we set

$$(\mathbf{T}_f)' = -\langle \mathbf{T}_f; s'\rangle. \tag{A28}$$

If f is a differentiable function, the derivative of the distribution associated with it, $\mathbf{T}_f$, is defined to be the distribution associated with the usual derivative of the function f; that is, the r.h.s. of (A28) is simply $\langle \mathbf{T}_{f'}; s\rangle$. This follows from an integration by parts and recognition that the boundary term vanishes because s belongs to the Schwartz space.

It should be noted that there are numerous distributions that have no associated function and cannot be represented as in (A21); familiar examples are the Cauchy principal value and the Dirac δ function, which perhaps is the most well-known generalized function. The "δ function," $\delta(x)$, was introduced by Dirac for handling the continuous spectrum in quantum mechanics with the definition

$$\int_{\mathbb{R}}\delta(x)\,\mathrm{d}x = 1,\quad \delta(x) = 0,\ x\neq 0 \tag{A29}$$

but with no value defined at $x = 0$. Integral equations represent an example of a functional relationship where corresponding to a function f there is a function g defined by

$$g(x) = \int_{\mathbb{R}} K(x,y)f(y)\,\mathrm{d}y \tag{A30}$$

for some kernel K. The kernel of the identity transformation is formally the Dirac δ function since

$$f(x) = \int_{\mathbb{R}}\delta(z-x)f(z)\,\mathrm{d}z \tag{A31}$$

provided f is continuous at x, a relation given by Dirac [19]. Equation (A31) defines what is usually called the "sifting" property of the δ function. Equation (A28) is valid irrespective of whether the distribution is defined by the integral formula; thus distributional derivatives of the Dirac δ function are well defined.

That the conventional description of the Dirac δ function is problematic can be seen as follows; the definition, (A29), is consistent with the criterion (A2) for a function to be integrable. However, the associated tempered distribution $\mathbf{T}_\delta$ defined by (A21) is trivial, $\mathbf{T}_\delta = 0$, since an integral over an interval of zero length (the point 0) is zero whatever $s(x)$ is; instead the precise definition of the δ function relies on its action on elements s of S with (A31) replaced by

$$\langle \boldsymbol{\delta}_{x_0}, s\rangle = s(x_0). \tag{A32}$$

Just as a classical function f is defined by the complete collection of its values $f(x)$ at all points $\{x\}$ in its domain, so the "δ function" $\delta(x-x_0)$ is defined by the set of values $\{s(x_0)\}$ of all the functions $\{s\}$ in S; however, by analogy with (A21) one commonly writes

$$\int_{\mathbb{R}}\delta(x-x_0)s(x)\,\mathrm{d}x = s(x_0). \tag{A33}$$

This is a purely formal statement since there is no function δ satisfying (A29) according to the classical definition of a function, and the "integral" in (A33) cannot be interpreted in the usual way as a Riemann or Lebesgue integral. To indicate that the "δ function" is a distribution we write it as $\boldsymbol{\delta}$.

The Fourier transform of the delta distribution is simply a constant, the value of which depends on how the factor of 2π is shared between the transform and its inverse. The tempered distribution T_δ is the identity operation for convolution in the sense that

$$s' * \langle T_\delta, s\rangle = \langle T_\delta, s' * s\rangle = \int_{\mathbb{R}} s'(y)s(y)\,\mathrm{d}y \equiv \langle T_{s'}, s\rangle. \tag{A34}$$

Distributions can be given a concrete realization in the following way which is an alternative view to the "abstract" description earlier. For simplicity of exposition we restrict the discussion to distributions on $\mathbb{R}$. Let $s(x)\in\mathcal{S}$ be a normalized function in the Schwartz space

$$\int s(x)\,\mathrm{d}x = 1, \tag{A35}$$

and note that its translations, $s(z-x)$, also belong to the set of Schwartz functions, $\mathcal{S}$. By dilation of s with a parameter $0<\lambda<1$, we obtain the scaled function

$$s_\lambda = \frac{1}{\lambda} s\left(\frac{x}{\lambda}\right) \tag{A36}$$

with the same normalization. s_λ can be viewed as a representation of the Dirac δ which has been "broadened out" or *mollified* about $x=0$. It has area 1 [normalization (A35)], approximate height $1/2\lambda$, and approximate width 2λ. A function such as (A36) is called a *mollifier*.

Given a continuous function or distribution f we can construct "representatives" of it as a sequence of functions that are *smooth* in the variable x using the convolution integral with s_λ,

$$\mathcal{F}_\lambda(s,x) = \int_{\mathbb{R}} \frac{1}{\lambda} s\left(\frac{z-x}{\lambda}\right) f(z)\,\mathrm{d}z, \tag{A37}$$

which is an obvious echo of the sifting property of the Dirac δ (A31). With a change of integration variable we have

$$\mathcal{F}_\lambda(s,x) = \int_{\mathbb{R}} s(y) f(x+\lambda y)\,\mathrm{d}y, \tag{A38}$$

which is such that

$$f = \lim_{\lambda\to 0} \mathcal{F}_\lambda(s) \tag{A39}$$

irrespective of the particular Schwartz function s in the convolution. Such representatives are very convenient when derivatives of distributions are required, but as they stand there is a severe limitation, namely, representatives when multiplied together do not give a representative of a distribution.

As an example, consider the Dirac δ which is the $\lambda\to 0$ limit of s_λ, (A36). According to the above discussion its square should be available from the square of its representative, so taking $\xi\in\mathcal{S}$ we should have for a distribution

$$\langle \delta^2, \xi\rangle = \lim_{\lambda\to 0}\int_{\mathbb{R}} \xi(x)\frac{1}{\lambda} s\left(\frac{x}{\lambda}\right)^2 \mathrm{d}x = \lim_{\lambda\to 0}\frac{\xi(0)}{\lambda}\int_{\mathbb{R}} s(z)^2\,\mathrm{d}z, \tag{A40}$$

which is evidently ∞. Thus δ^2 does not belong to the space of distributions, $\mathcal{D}'$. This illustrates a quite general result: there is no multiplication on all of $\mathcal{D}'$ giving a result that is in $\mathcal{D}'$.

As another example, take (A21) again for two functions, f_1, f_2, on some set $\Omega\subset\mathbb{R}^n$. When considered as distributions they are the linear forms

$$\langle \mathbf{T}_{f_1}, s\rangle = \int_\Omega f_1(x)s(x)\,\mathrm{d}x, \quad \langle \mathbf{T}_{f_2}, s\rangle = \int_\Omega f_2(x)s(x)\,\mathrm{d}x. \tag{A41}$$

Their product is then

$$s \mapsto \int_\Omega f_1(x)s(x)\,\mathrm{d}x \int_\Omega f_2(x)s(x)\,\mathrm{d}x, \tag{A42}$$

while the classical product of f_1, f_2 interpreted as a distribution leads to

$$s \mapsto \int_\Omega f_1(x)f_2(x)s(x)\,\mathrm{d}x, \tag{A43}$$

which in general is not the same as (A42); thus the notion of "product" is ambiguous. The idea of a distribution as a generalization of the classical notion of a function has been very fruitful in analysis; however. some *nonlinear* problems require new mathematics that goes beyond the notion of a distribution as a linear form so that the problem that *multiplication* is not defined for distributions can be overcome [4].

An essential mathematical notion that is required for this development is that of "embedding" elements of a set into another set. Roughly speaking the relationship between two sets $\mathcal{X}$ and $\mathcal{Y}$ is an embedding if the map $f:\mathcal{X}\to\mathcal{Y}$ [$x\to f(x)$ in the notation of (A1)] has the properties:

(1) For every $x_1, x_2\in\mathcal{X}$ such that $x_1\neq x_2$, $f(x_1)\neq f(x_2)$, that is, different elements of $\mathcal{X}$ correspond to distinct elements of $\mathcal{Y}$.

(2) If some property holds for $x_1, x_2, \ldots, x_n$ the same property holds for $f(x_1), f(x_2), \ldots, f(x_n)$.

A subgroup in a larger group, the integers in relation to the rational numbers, the rationals in relation to real numbers are all examples of embeddings.

The aim is to define an embedding that gives a set, $\mathcal{G}$, of generalized functions containing the distributions and ordinary functions such that the usual rules of differentiation apply, and multiplication is defined; such a set is called a *differential algebra*. The Schwartz impossibility theorem [4] is the demonstration that there is no differential algebra in which the ordinary product of continuous functions is equal to the corresponding product of generalized functions they are related to [5]. Colombeau's insight was to recognize that the Schwartz theorem did not apply to *smooth* functions ($\mathcal{C}^\infty$). Using the properties of the Schwartz space he demonstrated that the product of two smooth functions embedded in $\mathcal{G}$ coincides with their ordinary product (in $\mathcal{C}^\infty$).

The technical details of this novel mathematics can be found in the cited literature [5–14]. The first step in constructing an element of an algebra of generalized functions from f (function or distribution) is the definition of representatives as in (A37); such representatives can be freely multiplied. Colombeau's aim was to replace f by some $\mathcal{F}_\lambda$ (a generalized function) with an infinitesimal error for finite λ, and this requires further restrictions on the admissible functions $\{s\}$. Note, however, that there is not a single s to be considered; the construction is available to any element of the set of mollifiers. For ease of presentation the following is restricted to the one-variable case ($\mathbb{R}$); it can easily be extended to many variables ($\mathbb{R}^n$). If one puts $\lambda=0$ the conventional account in terms of functions and/or distributions is recovered.

(1) Let $\mathrm{F}s_\lambda$ be the Fourier transform of (A36); from (A4) and (A37) it is

$$\mathrm{F}s_\lambda := \hat{s}_\lambda(k) \equiv \hat{s}(\lambda k). \tag{A44}$$

A mollifier s is said to be "suitable" if the transform $\hat{s}(k)=1$ in a *finite neighborhood of* $k=0$ and not just at the *point* 0. Such a transform has height 1, approximate area $2/\lambda$ and approximate width $2/\lambda$. It can be viewed as a "cutoff" for large k values that vanishes smoothly as $k\to\infty$; for this reason the Fourier transforms $\{\hat{s}\}$ are referred to as *dampers*. As with the mollifiers a product of dampers belongs to the set of dampers. A Fourier transform of an element of $\mathcal{S}$ belongs to $\mathcal{S}$; in the following we will require the further condition that the transform $\hat{s}$ be a *test function*, that is, it has compact support ($\hat{s}\in\mathcal{D}\subset\mathcal{S}$) [14] [see below, (A48)].

(2) Since $\mathcal{F}_\lambda(s,x)$ is *smooth* it has a Taylor series expansion in powers of λ with remainder

$$\mathcal{F}_\lambda(s,x) = f(x) + \cdots \frac{\lambda^m}{m!} f^m(x) \int_{\mathbb{R}} y^m s(y)\,\mathrm{d}y + \mathrm{O}_x(\lambda^{m+1}), \tag{A45}$$

where as usual $f^m(x)$ is the mth derivative of $\mathcal{F}_\lambda$ evaluated at x. It follows that

$$\mathcal{F}_\lambda(s,x) - f(x) = \sum_{m=1} \frac{\lambda^m}{m!} f^m(x) \int_{\mathbb{R}} y^m s(y)\,\mathrm{d}y + \mathrm{O}_z(\lambda^{m+1}). \tag{A46}$$

(3) Colombeau showed that it is possible to construct a set of Schwartz functions $\{s\}$ such that their first m moments vanish [11],

$$\int_{\mathbb{R}} y^k s(y)\,\mathrm{d}y = 0, \quad 1 \leqslant k \leqslant m. \tag{A47}$$

This means that the sum term in (A46) can be made to vanish and $\mathcal{F}_\lambda(x) = f(x) + \mathrm{O}_x(\lambda^{m+1})$. Hence the remainder can be made as small as we please for m large enough for any λ.

(4) An equivalent formulation of the conditions (A47) can be expressed in terms of the Fourier transformation [5]; using the "hat" notation (A4), for the transform of any function $s(x)$ in the Schwartz space S, we have the relations

$$\hat{s}(0) = \int_{\mathbb{R}} s(x)\,\mathrm{d}x, \quad (-\mathrm{i})^n \mathrm{D}^n \hat{s}(0) = \int_{\mathbb{R}} x^n s(x)\,\mathrm{d}x, \tag{A48}$$

where D stands for any derivative operator. Taking $s(x)$ with $\hat{s}(0) = 1$, the conditions (A47) are satisfied for any m as large as we like. All the mollifiers can be assumed to belong to the set

$$\mathcal{A}_\infty = \{s(x) \in \mathcal{S}, \hat{s} \in \mathcal{D}, \text{with } \hat{s}(0) = 1\}, \tag{A49}$$

where "0" implies a finite neighborhood of zero, not just the point 0. This is the viewpoint adopted in [14].

(5) Among the functionals $\mathcal{F}_\lambda(s;x)$ we focus on two categories, $\mathcal{F}_\lambda(s;x) \in \mathcal{E}_M$, called *moderate*, and $\mathcal{F}_\lambda(s;x) \in \mathcal{N}$, called *negligible*. The precise distinction between these two categories can be found in the references to the Colombeau algebra cited previously. The important point is that the Colombeau algebra $\mathcal{G}$ is the quotient $\mathcal{E}_M/\mathcal{N}$. The moderate functionals have polynomial growth in $1/\lambda$ as $\lambda \to 0$, whereas the negligible ones decay faster than any power of λ as $\lambda \to 0$. Heuristically, the moderate functionals are the ones of interest, and the negligible ones effectively play the role of the generalized number "0" in the algebra. The important point is that the product of two moderate functionals is again moderate, whereas if the product contains at least one negligible functional the result is negligible. Two functionals are said to be "equivalent" if their difference, $\mathcal{F} - \mathcal{F}'$, is negligible.

A generalized function is associated with a distribution $\mathbf{T}_f$ if it has a representative $\mathcal{F}_\lambda$ belonging to the class of moderate functionals such that

$$\lim_{\lambda \to 0} \mathcal{F}_\lambda = \mathbf{T}_f \in \mathcal{D}'. \tag{A50}$$

Every distribution can be converted to a moderate family through the construction of a representative by convolution as above. However, not every moderate family is the regularization of a distribution. The square of the Dirac δ is not a distribution, (A40), but its representative (A36) squared is moderate. Whereas there is only one Dirac δ distribution in $\mathcal{D}'$ there is an infinity of Dirac δ like generalized functions in $\mathcal{G}$, and likewise for any other distribution. Colombeau emphasized that "equivalent" functionals in $\mathcal{G}$ are not necessarily equal, in the sense of the classical equality denoted by $=$, because they may differ by infinitesimal quantities; rather there is a *weak* equality for which he proposed the relational symbol $\approx$ [13].

The general idea then is that the functions $\{f\}$ are transformed to $\{\mathcal{F}_\lambda\}$ as in (A37) with mollifiers taken from (A49), and the appropriate $\mathcal{F}_\lambda$ is used in all calculations with λ finite until the end of the calculation. If the model is linear, the results will be the same as though one stayed within distribution theory, which is recovered in the limit $\lambda \to 0$. In linear or nonlinear models which result in divergences as $\lambda \to 0$, the parameter λ can be kept finite, and the results will be "generalized numbers" or "generalized functions," genuinely new mathematical objects.

[1] R. G. Woolley, Phys. Rev. Res. **2**, 013206 (2020).
[2] R. G. Woolley, *Foundations of Molecular Quantum Electrodynamics* (Cambridge University Press, Cambridge, 2022), Chap. 2 and 3.
[3] P. Roman, *Some Modern Mathematics for Physicists and Other Outsiders* (Pergamon Press, Oxford, 1965), Vol. 2.
[4] L. Schwartz, C. R. Acad. Sci. **239**, 847 (1954).
[5] A. Gsponer, arXiv:0807.0529.
[6] G. Hörmann and M. Kunzinger, J. Math. Anal. Appl. **261**, 254 (2001)
[7] G. P. Lepage, arXiv:hep-ph/0506330v1.
[8] S. Weinberg, *The Quantum Theory of Fields*, Foundations (Cambridge University Press, Cambridge, 1995), Vol. 1.
[9] B. Delamotte, Am. J. Phys. **72**, 170 (2004).
[10] J. F. Colombeau, *New Generalized Functions and Multiplication of Distributions*, North-Holland Mathematics Studies Vol. 84 (North-Holland, Amsterdam, 1984).
[11] J. F. Colombeau, *Elementary Introduction to New Generalized Functions* (North-Holland, Amsterdam, 1985).
[12] J. F. Colombeau, *Multiplication of Distributions* (Springer-Verlag, Berlin, 1992).
[13] J. F. Colombeau, arXiv:math-ph/0612077.
[14] J. F. Colombeau and A. Gsponer, arXiv:0807.0289v2.
[15] R. G. Woolley, Mol. Phys. **22**, 1013 (1971).
[16] S. Weinberg, *Lectures on Quantum Mechanics* (Cambridge University Press, Cambridge, 2013), p. 323.
[17] E. H. Lieb and R. Seiringer, *The Stability of Matter in Quantum Mechanics* (Cambridge University Press, Cambridge, 2010).
[18] E. A. Power and S. Zienau, Philos. Trans. R. Soc. London Ser. A **251**, 427 (1959).
[19] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1930).