Commun. Math. Phys. 241, 519–581 (2003)
Digital Object Identifier (DOI) 10.1007/s00220-003-0939-9

Communications in
**Mathematical
Physics**

Nonlinear Maxwell Equations in Inhomogeneous Media

Anatoli Babin, Alexander Figotin

Department of Mathematics, University of California at Irvine, Irvine, CA 92697-3875, USA

Received: 26 October 2002 / Accepted: 23 June 2003
Published online: 19 September 2003 –

Abstract: We study the basic properties of the Maxwell equations for nonlinear inhomogeneous media. Assuming the classical nonlinear optics representation for the nonlinear polarization as a power series, we show that the solution exists and is unique in an appropriate space if the excitation current is not too large. The solution to the nonlinear Maxwell equations is represented as a power series in terms of the solution of the corresponding linear Maxwell equations. This representation holds at least for the time period inversely proportional to the appropriate norm of the solution to the linear Maxwell equation. We derive recursive formulas for the terms of the power series for the solution including an explicit formula for the first significant term attributed to the nonlinearity.

1. Introduction

One of the motivations of this work is the growing interest in the theory of linear and nonlinear photonic crystals which are man-made periodic dielectric media, see [2, 7, 8, 12, 13, 18, 20, 24, 25, 30, 31, 36, 37, 41, 44, 47, 50]. In [5] we developed a framework for a consistent mathematical treatment of nonlinear interactions in periodic dielectric media. This paper provides rigorous proofs of the basic properties of nonlinear inhomogeneous Maxwell equations used in [5], including the existence of "well behaved" solutions for sufficiently long times. In addition, we consider here not only periodic but general inhomogeneous media.

We consider classical Maxwell equations ([27], Sect. 6.12)

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\partial_t \mathbf{B}(\mathbf{r}, t) - 4\pi \mathbf{J}_B(\mathbf{r}, t), \quad \nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0, \tag{1.1}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \partial_t \mathbf{D}(\mathbf{r}, t) + 4\pi \mathbf{J}_D(\mathbf{r}, t), \quad \nabla \cdot \mathbf{D}(\mathbf{r}, t) = 0, \tag{1.2}$$

where $\mathbf{H}$, $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{D}$ are respectively the magnetic and electric fields, magnetic and electric inductions, and $\mathbf{J}_D$ and $\mathbf{J}_B$ are excitation currents (current sources), $\mathbf{r} = (r_1, r_2, r_3)$.

It is assumed that the Maxwell equations (1.1) and (1.2) are written in dimensionless variables. We also assume that there are no free electric and magnetic charges, i.e.

$$\nabla \cdot \mathbf{J}_D(\mathbf{r}, t) = 0, \ \nabla \cdot \mathbf{J}_B(\mathbf{r}, t) = 0, \tag{1.3}$$

that is fully consistent with (1.1) and (1.2). Notice that Eqs. (1.1) and (1.2) require the fields **B** and **D** to be divergence free at all times. Following to our approach in [5] we use the excitation current

$$\mathbf{J} = \begin{pmatrix} 4\pi \mathbf{J}_D \\ 4\pi \mathbf{J}_B \end{pmatrix} \tag{1.4}$$

to produce non-zero solutions to the Maxwell equations (1.1) and (1.2), in particular wavepackets. We assume that $\mathbf{J}(t)$ vanishes for negative times, i.e.

$$\mathbf{J}_D(\mathbf{r}, t) = \mathbf{J}_B(\mathbf{r}, t) = 0, \ t \leq 0, \tag{1.5}$$

and we look for solutions satisfying the following rest condition:

$$\mathbf{D}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r}, t) = \mathbf{B}(\mathbf{r}, t) = 0 \quad \text{for} \ t \leq 0. \tag{1.6}$$

The dielectric properties of the medium are described by the constitutive (material) relations between the fields **E**, **D**, **H** and **B**, which can be nonlinear. For simplicity we consider the nonmagnetic media, i.e.

$$\mathbf{B}(\mathbf{r}, t) = \mu \mathbf{H}(\mathbf{r}, t), \ \mu = 1. \tag{1.7}$$

The constitutive relations between the fields **E** and **D** are of the standard form

$$\mathbf{D}(\mathbf{r}, t) = \boldsymbol{\varepsilon}(\mathbf{r}) \mathbf{E}(\mathbf{r}, t) + 4\pi \mathbf{P}_{\mathrm{NL}}(\mathbf{E}(\mathbf{r}, t)), \tag{1.8}$$

where

$$\boldsymbol{\varepsilon}(\mathbf{r}) = 1 + 4\pi \boldsymbol{\chi}^{(1)}(\mathbf{r}), \ \mathbf{r} = (r_1, r_2, r_3), \tag{1.9}$$

is the *electric permittivity tensor* (dielectric constant) describing the linear properties of the medium with $\boldsymbol{\chi}^{(1)}(\mathbf{r})$ being the *linear susceptibility tensor*, and $4\pi \mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ is the nonlinear component of the polarization total polarization **P**.

The electric permittivity tensor $\boldsymbol{\varepsilon}(\mathbf{r})$ is assumed to satisfy the following condition.

Condition 1.1. *The 3×3 matrix $\boldsymbol{\varepsilon}(\mathbf{r})$ with complex entries $\boldsymbol{\varepsilon}_{mn}(\mathbf{r})$ is a Hermitian matrix, i.e. $\boldsymbol{\varepsilon}_{mn}^*(\mathbf{r}) = \boldsymbol{\varepsilon}_{nm}(\mathbf{r})$, $\mathbf{r} \in \mathbb{R}^3$. It is bounded and positive definite, namely it satisfies for some constants $\varepsilon_+ \geq \varepsilon_- > 0$ the following inequalities:*

$$\varepsilon_- |\mathbf{e}|^2 \leq \sum_{m,n=1}^{3} \boldsymbol{\varepsilon}_{mn}(\mathbf{r}) e_m^* e_n \leq \varepsilon_+ |\mathbf{e}|^2, \ \mathbf{r} \in \mathbb{R}^3, \ \mathbf{e} = (e_1, e_2, e_3) \in \mathbb{C}^3. \tag{1.10}$$

The dependence of $\boldsymbol{\varepsilon}(\mathbf{r})$ on $\mathbf{r}$ is sufficiently smooth. Namely, there exists an integer $s > 3/2$ such that $\boldsymbol{\varepsilon}(\mathbf{r})$ and its inverse $\boldsymbol{\eta}(\mathbf{r}) = \boldsymbol{\varepsilon}^{-1}(\mathbf{r})$ have continuous, bounded over $\mathbb{R}^3$ derivatives of order up to s, that is as a function of $\mathbf{r}$ they have the following norms bounded:

$$\|\boldsymbol{\varepsilon}\|_{C^s(\mathbb{R}^3)}, \ \left\|\boldsymbol{\varepsilon}^{-1}\right\|_{C^s(\mathbb{R}^3)} < \infty. \tag{1.11}$$

We allow $\boldsymbol{\varepsilon}(\mathbf{r})$ to be Hermitian with complex entries, rather than simply real symmetric, since such permittivity tensors occur for a general dielectric (gyrotropic) media (see, for instance, [16] p.86 and [29], p. 49). We also allow for $\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ a general analytic dependence in $\mathbf{E}(\cdot)$,

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{E}) = \sum_{n=n_0}^{\infty} \mathcal{P}_n\left(\mathbf{E}^n\right),\ n_0 \geq 2,\ \mathcal{P}_n\left(\mathbf{E}^n\right) = \mathcal{P}_n(\mathbf{E}, \ldots, \mathbf{E}), \tag{1.12}$$

with $\mathcal{P}_n(\mathbf{E}_1, \ldots, \mathbf{E}_n)$ being a n-linear operator that acts on functions $\mathbf{E}_i(\mathbf{r}, t)$. The leading term of the degree $n_0 \geq 2$ in the nonlinear polarization $\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ in most of the applications is either quadratic, $n_0 = 2$, or cubic, $n_0 = 3$, [10, 11].

Following the classical nonlinear optics (see [11], Sect. 2) we assume the n-linear operators $\mathcal{P}_n(\mathbf{E})$ in (1.12) to be of the form

$$\begin{aligned} &\mathcal{P}_n(\mathbf{E})(\mathbf{r}, t) \\ &\quad = \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} P_n(\mathbf{r}; t - t_1, \ldots, t - t_n; \mathbf{E}(\mathbf{r}, t_1), \ldots, \mathbf{E}(\mathbf{r}, t_n)) \prod_{j=1}^{n} dt_j, \end{aligned} \tag{1.13}$$

$$P_n(\mathbf{r}; \tau_1, \ldots, \tau_n; \cdot) : \left(\mathbb{C}^3\right)^n \to \mathbb{C}^3,\ n \geq n_0.$$

The function $P_n(\mathbf{r}; \tau_1, \ldots, \tau_n; \mathbf{e}_1, \ldots, \mathbf{e}_n)$, which is a n-linear form (tensor) acting on $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{C}^3$, is called the *polarization response function* of the order n. For fixed $\mathbf{r}$ and τ_j the quantity $P_n(\mathbf{r}; \tau_1, \ldots, \tau_n; \mathbf{e}_1, \ldots, \mathbf{e}_n)$ is a n-linear function of $\mathbf{e}_j$ with values in $\mathbb{C}^3$. The Fourier transform of P_n in $(\tau_1, \ldots, \tau_n)$ is called the frequency dependent *susceptibility tensor* of the order n. We recall that the representation (1.13) takes explicitly into account *two fundamental properties of the medium: the time-invariance and the causality*, [11], Sect. 2. We refer to the series (1.12), (1.13) and the analytic function it defines as *causal*. Causality implies that $\mathcal{P}_n(\mathbf{E})(\cdot, t)$ depends only on $\mathbf{E}^{(j)}\left(\cdot, t_j\right)$ with $t_j \leq t$.

Note that (1.2) contains $\partial_t \mathbf{D}$ and by (1.8) the equation implicitly involves $\partial_t \mathbf{P}_{\mathrm{NL}}(\mathbf{E})$. According to (1.12) and (1.13) the time derivative $\partial_t \mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ equals the sum of terms of the form

$$\partial_t \mathcal{P}_n(\mathbf{E})(\mathbf{r}, t) = \mathcal{P}_{n,n-1}(\mathbf{E})(\mathbf{r}, t) + \mathcal{P}_{n,n}(\mathbf{E})(\mathbf{r}, t), \tag{1.14}$$

$$\begin{aligned} &\mathcal{P}_{n,n-1}(\mathbf{E})(\mathbf{r}, t) \\ &\quad = \sum_{l=1}^{n} \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} P_n(\mathbf{r}; t - t_1, ..., 0, \ldots, t - t_n; \mathbf{E}(\mathbf{r}, t_1), \ldots, \mathbf{E}(\mathbf{r}, t_n)) \prod_{j \neq l} dt_j, \\ &\mathcal{P}_{n,n}(\mathbf{E})(\mathbf{r}, t) \\ &\quad = \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} \dot{P}_n(\mathbf{r}; t - t_1, \ldots, t - t_n; \mathbf{E}(\mathbf{r}, t_1), \ldots, \mathbf{E}(\mathbf{r}, t_n)) \prod_{j} dt_j, \end{aligned}$$

where

$$\dot{P}_n(\mathbf{r}; t_1, \ldots, t_n; \cdot) = \sum_{l=1}^{n} \dot{P}_{nl}(\mathbf{r}; t_1, \ldots, t_n; \cdot) \tag{1.15}$$

and $\dot{P}_{nl}(\mathbf{r}; t_1, \ldots t_n; \cdot)$ is the derivative of the tensor $P_n(\mathbf{r}; t_1, \ldots t_n; \cdot)$ with respect to t_l. From (1.14) one can see that to provide the regularity of the multilinear operators we have to impose proper conditions on the time derivatives of the kernels $P_n(\mathbf{r}; t_1, \ldots t_n; \cdot)$ as well as their values at the boundary faces $t_l = 0$. The conditions on the polarization response functions (tensors) $P_n = P_n\left(\mathbf{r}; \overrightarrow{\tau}; \cdot\right)$ from (1.13) have to imply that the series (1.12) and a similar series for $\partial_t(\mathbf{P}_{\mathrm{NL}}(\mathbf{E}))$ converge. Here is the condition imposed on the polarization response functions.

Condition 1.2. *For every $n \geq n_0$ the tensor valued functions*

$$P_n = P_n\left(\mathbf{r}; \overrightarrow{\tau}\right), \ \mathbf{r} \in \mathbb{R}^3, \ \overrightarrow{\tau} = (\tau_1, \ldots, \tau_n) \in \mathbb{R}_+^n, \tag{1.16}$$

and their first time derivatives $\dot{P}_{nl}\left(\cdot; \overrightarrow{\tau}\right)$ with respect to τ_l, $l = 1, ..., n$, have the following properties:

(i) they belong for every fixed $\overrightarrow{\tau}$ to the space $C^s\left(\mathbb{R}^3\right)$, $s \geq 2$, consisting of s times, continuously differentiable, bounded functions of $\mathbf{r}$ *;*
(ii) tensors $P_n\left(\cdot; \overrightarrow{\tau}\right)$ and $\dot{P}_{nl}\left(\cdot; \overrightarrow{\tau}\right)$ as elements of the space $C^s\left(\mathbb{R}^3\right)$ continuously depend on $\overrightarrow{\tau} \in \mathbb{R}_+^n$ up to the boundary $\partial^{n-1}\mathbb{R}_+^n$;
(iii) $P_n\left(\mathbf{r}; \overrightarrow{\tau}\right)$ satisfy the causality condition

$$P_n\left(\mathbf{r}; \overrightarrow{\tau}\right) = 0, \ \overrightarrow{\tau} \in \mathbb{R}^n - \mathbb{R}_+^n, \mathbf{r} \in \mathbb{R}^3; \tag{1.17}$$

(iv) there exist constants $\beta_P > 0$, $C_P > 0$ such that P_n and $\dot{P}_n$ in (1.14), (1.15) satisfy

$$\int_{\mathbb{R}_+^n} \left(\|P_n\|_{C^s} + \left\|\dot{P}_n\right\|_{C^s}\right) d\overrightarrow{\tau} + \int_{\partial^{n-1}\mathbb{R}_+^n} \|P_n\|_{C^s} \, d\overrightarrow{\tau} < C_P \beta_P^{-n}. \tag{1.18}$$

Note that $\mathbb{R}_+^n$ is the set of vectors from $\mathbb{R}^n$ with nonnegative components $\tau_j \geq 0$, $j = 1, ..., n$. The $(n-1)$-dimensional boundary $\partial^{n-1}\mathbb{R}_+^n$ of this set is the union of n faces $f_i = \left\{\overrightarrow{\tau} \in \mathbb{R}_+^n : \tau_i = 0\right\}$.

A typical and rather common in optics example of the response function is

$$P_n\left(\mathbf{r}; \overrightarrow{\tau}; \overrightarrow{\mathbf{e}}\right) = \begin{cases} \exp\left\{-\sigma \sum_{j=1}^n \tau_j\right\} p_n\left(\mathbf{r}; \overrightarrow{\mathbf{e}}\right) & \text{if all } \tau_j \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{1.19}$$

where $p_n\left(\mathbf{r}; \overrightarrow{\mathbf{e}}\right)$ is a n -linear form of $\overrightarrow{\mathbf{e}} \in \left(\mathbb{C}^3\right)^n$, $\sigma > 0$ is a constant.

We study solutions $\{\mathbf{H}(t), \mathbf{E}(t), \mathbf{B}(t), \mathbf{D}(t)\}$ to the Maxwell equations on the time interval $-\infty < t \leq T$, $T > 0$. The solutions are continuous bounded functions of time t, taking on values in the Sobolev space $\mathbf{H}^s$ with an integer $s > 3/2$ and such that (1.6) holds. Using common notations we denote the corresponding Banach space of such functions by $\mathcal{C}_{0,\mathbf{H}^s}^T = C_0([-\infty, T]; \mathbf{H}^s)$. The full list of functional spaces and other related concepts is provided in the next section. Solutions of (1.1), (1.2), (1.7) and (1.8) are assumed to have time derivatives from $C_0\left([-\infty, T]; \mathbf{H}^{s-1}\right)$. Under natural assumptions, such as Condition 1.2, the series (1.12) converges in a ball in the Banach space $C_0([-\infty, T]; \mathbf{H}^s)$ and determines an analytic function $\mathbf{P}_{\mathrm{NL}}(\mathbf{E}(\cdot))$ of $\mathbf{E}(t)$. Since $\mathbf{P}_{\mathrm{NL}}$ includes integration with respect to time, its time derivative $\partial_t(\mathbf{P}_{\mathrm{NL}}(\mathbf{E}(\cdot)))$ also belongs to $C_0([-\infty, T]; \mathbf{H}^s)$. All differential operators and functions in (1.1), (1.2) and (1.8) are well-defined for such solutions (a detailed definition of a solution is given in Definition 2.2). In the following sections we discuss in detail the relevant concepts and properties of functions analytic in Banach spaces. We also analyze a special class of analytic functions arising in the classical nonlinear optics for which $\mathcal{P}^{(n)}$ is defined by (1.13).

In this paper we assume the space dimension $d = 3$. The space dimension $d = 1$ or $d = 2$ when the coefficients and the fields do not depend on r_2, r_3 or r_3 respectively. In these cases our results hold too, moreover, the condition $s > 3/2$ is replaced by $s > d/2$.

Using common notations (see the next section if needed) we can formulate one of our main results as follows.

Theorem 1.3. *Let $s > 3/2$ and Conditions 1.1 and 1.2 hold. Then the series (1.13) converges for $\|\mathbf{E}\|_{C_0([-\infty,T];\mathbf{H}^s)} < R_P$, where R_P depends on C_P, β_P in (1.18). Let $\mathbf{J} \in L_{1,0}([-\infty, T]; \mathbf{H}^s)$ and*

$$\|\mathbf{J}\|_{L_1([-\infty,T;]Y)} \leq \delta < \delta_0, \quad 1 + T < \frac{C}{\delta_0^{n_0-1}}, \tag{1.20}$$

where the constants C, δ_0 depend on n_0 and the constants C_P, β_P, $\|\boldsymbol{\varepsilon}(\cdot)\|_{C^s(\mathbb{R}^3)}$, ε_-, ε_+ from Conditions 1.1 and 1.2. Then there exists a unique solution

$$\mathbf{W}(\mathbf{r}, \mathbf{t}) = (\mathbf{D}(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t), \mathbf{E}(\mathbf{r}, t), \mathbf{H}(\mathbf{r}, t)) \in C_0\left([-\infty, T]; \left(\mathbf{H}^s\right)^2\right) \tag{1.21}$$

with $\|\mathbf{E}\|_{C_0([-\infty,T];\mathbf{H}^s)} < R_P$ to Eqs. (1.1), (1.2), (1.6), (1.7), (1.8). This solution $\mathbf{W}(\mathbf{r}, \mathbf{t})$ is an analytic function of $\mathbf{J}$ and it can be represented by a convergent power series

$$\mathbf{W} = \mathcal{W}(\mathbf{J}) = \sum_{n=n_0}^{\infty} \mathcal{W}_n(\mathbf{J}), \tag{1.22}$$

where $\mathcal{W}_n$ is a n-linear operator. The operators $\mathcal{W}_n$ can be explicitly expressed in terms of P_m by recursive relations (7.49) (see Theorem 7.8).

The proof of Theorem 1.3 is given in Sect. 8. More detailed statements are provided by Theorem 7.8 and Lemma 7.4. We would like to remark that the proof of the existence of solutions as well as the studies of their properties (see [5, 6]) are based on the reduction of the system (1.1), (1.2), (1.6), (1.7), (1.8) to the problem (7.36) for divergence-free variables **D, B.**

The primary focus of this paper is on the following subjects: (i) the existence and the uniqueness of the solution to the nonlinear Maxwell equations (1.1), (1.2), (1.6), (1.7), (1.8) for large time intervals; (ii) the representation of solutions in the form of convergent series involving causal operators. The proofs of the existence and the uniqueness provide a basis for a more detailed nonlinear interaction theory along the lines of [5, 6]. Our choice of the theory of analytic functions in infinite dimensional spaces as a technical tool is motivated primarily by the representation of the nonlinear polarization by the series (1.12), (1.13) which is standard in classical nonlinear optics. In addition to that, it turns out that the analytic approach based on representations of type (1.12), (1.13), (1.22) has additional advantages. In particular, it allows to give a rigorous meaning to some frequency-dependent nonlinearities, see [6] for details. It also allows to consider general nonlinearities as long as we can control their magnitude. In particular, we do not impose any specific structural conditions, such as the symmetry or skew-symmetry, sign conditions, etc., on the nonlinear tensors. Series expansions with the resulting analyticity naturally yield a rather constructive description of the solutions in the form common in the physical literature. Another important incentive for using the analytic functions approach is its usefulness in further analysis of the solutions, including their asymptotic approximations, when the excitation currents $\mathbf{J}(t)$ are nearly monochromatic wave packets, [5], with relative frequency bandwidth $\varrho = \frac{\Delta\omega}{\omega} \to 0$. It turns out, [5], that ϱ determines a naturally arising "slow time" $\tau = \varrho t$. It also follows from [5] that to analyze solutions of the Maxwell equation as $\varrho \to 0$, $t \to \infty$ with $\tau = \varrho t$ being fixed, one needs uniform with respect to ϱ approximations of the solutions as functions of the excitation currents $\mathbf{J}(t)$. The analytic function approach and series representations allow to get that kind of approximations.

To carry out the analytic function approach to the construction of the solutions we need to properly recast the original Maxwell equations (1.1), (1.2). This recasting is done in two steps.

The first step is to choose the divergence-free fields $\mathbf{D}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$ as the primary field variables. When changing the variables we keep in mind that the nonlinear polarization has the form (1.12), (1.13) implying that for any instant t the field $\mathbf{D}(\mathbf{r}, t)$ depends on all the values of the field $\mathbf{E}(\cdot, t')$ for prior times $t' < t$. This is one of the factors which has to be taken into account for the choice of functional spaces, namely, the spaces $C([-\infty, T]; \mathcal{H})$ with a suitable Hilbert space $\mathcal{H}$. An analysis shows that the choice of a suitable function space $\mathcal{H}$ of functions $\mathbf{V}(\mathbf{r})$ of the position variable $\mathbf{r}$ should be based on the following considerations. First, if $\mathbf{U}(\mathbf{r}, t)$ is a solution of the relevant linear Maxwell equations, $\mathbf{U}(\mathbf{r}, t)$ must remain in $\mathcal{H}$ at all times, and, more than that, the norm $\|\mathbf{U}(\cdot, t)\|_{\mathcal{H}}$ must remain bounded as time evolves. This property is important for the control of the magnitude of an $\mathcal{H}$ -valued solution for large time intervals that is crucial for existence on such intervals. The second condition on $\mathcal{H}$ is that the multilinear forms (1.13) must be continuous in $\mathcal{H}$. This requires that $\mathcal{H}$ must be closed with respect to the pointwise multiplication of functions. For instance, the space L_2 of square integrable functions is definitely not suitable. We show that the space $\mathcal{H} = \mathring{\mathbf{H}}^s_M$ with integer $s > 3/2$ introduced in Sect. 3 has both required properties.

The paper is structured as follows. In Sect. 2 we introduce function spaces, give a definition of a solution of the nonlinear Maxwell system and prove Theorem 2.4 on the uniqueness of such a solution. The equivalence of norms generated by the linear Maxwell operator and Sobolev norms and related issues are discussed in Sect. 3. In Sect. 4 we give necessary definitions and facts from the theory of analytic operators (functions) in Banach spaces. Then we provide the proof of the related Implicit Function Theorem with particular emphasis on explicit constructions of polynomial operators and explicit estimates on the radius of convergence of relevant power expansions. In Sect. 5 we consider the case of causal multilinear operators generalizing (1.12), (1.13). It is the most technical part of the paper. In Sect. 6 we consider Maxwell equations in a generalized operator setting. Section 7 is devoted to an integral form of the Maxwell equations involving only bounded operators. We call it a *regular integral form*. The reduction to this form essentially uses the fact that the nonlinear polarization is given by causal integral operators of the form (1.12), (1.13). Then we prove results for the original Maxwell equations (1.1), (1.2), (1.6), (1.7), (1.8), in particular Theorem 1.3 and more detailed statements such as Theorem 7.8.

2. Function Spaces for Solutions

In this section we define suitable function spaces for solutions to Maxwell equations and introduce notations.

2.1. Notations and function spaces. Below we provide a list of common and a few special notations needed for our analysis.

$\partial^{\alpha} = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\partial_3^{\alpha_3}$ where ∂_j is the partial derivative with respect to the space coordinate r_j and $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ is a multiindex with α_j being nonnegative integers and $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$.

$\mathbf{L}_2\left(\mathbb{R}^3\right) = L_2\left(\mathbb{R}^3, \mathbb{C}^3\right) = \mathbf{L}_2$ the Hilbert space of 3-dimensional vector fields $\mathbf{v}(\mathbf{r})$ with the scalar product

$$(\mathbf{u}, \mathbf{v}) = \int_{\mathbb{R}^3} \mathbf{u}^*(\mathbf{r})\mathbf{v}(\mathbf{r})\, d\mathbf{r} = \int_{\mathbb{R}^3} \overline{\mathbf{u}} \cdot \mathbf{v}(\mathbf{r})\, d\mathbf{r},\ \mathbf{u} \cdot \mathbf{v} = \sum_{j=1}^{3} u_j v_j, \tag{2.1}$$

where $\overline{\mathbf{u}}$ is a vector with components complex conjugate to the components of $\mathbf{u}$, and $\mathbf{u}^* = \overline{\mathbf{u}}^{\mathrm{T}}$ is the vector adjoint of $\mathbf{u}$ and for a vector (column) $\mathbf{u}$ the notation $\mathbf{u}^{\mathrm{T}}$ stands for a vector transposed to it. We will also use the notation $\boldsymbol{\varepsilon}^{\mathrm{T}}$ for a matrix transposed to the matrix $\boldsymbol{\varepsilon}$. If we have a term $\mathbf{uv}$, where $\mathbf{u}$ and $\mathbf{v}$ are matrices or vectors then vectors are treated as corresponding matrices and $\mathbf{uv}$ is understood as a standard matrix product. In our problems $\mathbf{v}(\mathbf{r})$ can be, for instance, the electric or magnetic field.

$\mathring{\mathbf{L}}_2\left(\mathbb{R}^3\right) = \mathring{\mathbf{L}}_2$ is the subspace of $\mathbf{L}_2\left(\mathbb{R}^3\right)$ consisting of the divergence free 3-dimensional fields, i.e. the subspace of $\mathbf{L}_2\left(\mathbb{R}^3\right)$ orthogonal to all the fields of the form grad $\varphi(\mathbf{r})$, where $\varphi(\mathbf{r}) \in C_0^{\infty}\left(\mathbb{R}^3\right)$.

Π_0 is the $\mathbf{L}_2$-orthogonal projection operator on $\mathring{\mathbf{L}}_2$.

$\mathbf{L}_2^2 = \mathbf{L}_2 \times \mathbf{L}_2$, $\mathring{\mathbf{L}}_2^2 = \mathring{\mathbf{L}}_2 \times \mathring{\mathbf{L}}_2$.

$\mathbf{H}_q^s\left(\mathbb{R}^3\right) = \mathbf{H}_q^s = \mathbf{H}^s$, $s = 0, 1, 2, \ldots$, is the Sobolev space of q-dimensional vector fields (or n-linear form (tensor) fields that are vectors of dimension q^n). For vector fields $\mathbf{V}(\mathbf{r}) = \left\{V_j(\mathbf{r}) : 1 \le j \le q\right\}$, $\mathbf{r} \in \mathbb{R}^3$ (index $q = 3, 6$ will be often suppressed) the Sobolev norm

$$\|\mathbf{V}\|_{\mathbf{H}^s}^2 = \sum_{|\alpha| \le s} \int_{\mathbb{R}^3} \left|\partial^{\alpha} V(\mathbf{r})\right|^2 \mathbf{dr}, \tag{2.2}$$

with $\partial^{\alpha} = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\partial_3^{\alpha_3}$, $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$ and

$$|\mathbf{v}|^2 = \overline{\mathbf{v}} \cdot \mathbf{v} = |v_1|^2 + \ldots + \left|v_q\right|^2 \tag{2.3}$$

being the standard Euclidean norm of a vector $\mathbf{v} \in \mathbb{C}^q$. For a n-linear form (tensor) field $\mathbf{V}(\mathbf{r}; \cdot)$ of tensors acting on vectors $\mathbf{e} \in \mathbb{C}^q$ the Sobolev norm $\|\mathbf{V}\|_{\mathbf{H}^s}^2$ is given by (2.2), where the norm of a $n-$ linear tensor $\mathbf{V}' = \partial^{\alpha} V(\mathbf{r})$ is given for any given α, $\mathbf{r}$ by

$$\left|\mathbf{V}'\right| = \sup_{|\mathbf{e}_1| = \ldots = |\mathbf{e}_n| = 1} \left|\mathbf{V}'(\mathbf{e}_1, \ldots, \mathbf{e}_n)\right| \tag{2.4}$$

with $\left|\mathbf{e}_j\right|$ being the standard Euclidean norm of a vector $\mathbf{e}_j \in \mathbb{C}^q$.

$\mathring{\mathbf{H}}^s = \mathbf{H}_3^s \cap \mathring{\mathbf{L}}_2$ and

$$\|\mathbf{u}\|_{\mathring{\mathbf{H}}^s} = \|\mathbf{u}\|_{\mathbf{H}^s}, \quad \mathbf{u} \in \mathring{\mathbf{H}}^s. \tag{2.5}$$

$\mathcal{C}_Y^T = C([-\infty, T]; Y)$, where Y is a Banach space, is the space of $Y-$ valued functions $y(t)$, $-\infty < t \le T$, with the norm defined by

$$\|y\|_{\mathcal{C}_Y^T} = \sup_{-\infty < t \le T} \|y(t)\|_Y . \tag{2.6}$$

In particular,

$\mathcal{C}^T_{\mathbf{H}^s} = C\left([-\infty, T]; \mathbf{H}^s\right)$, $T > 0$, is a Banach space of $\mathbf{H}^s$-valued continuous trajectories $\mathbf{U}(t)$, $-\infty < t \leq T$ in $\mathbf{H}^s$ with the norm

$$\|\mathbf{U}\|_{C([-\infty,T];\mathbf{H}^s)} = \sup_{-\infty<t\leq T} \|\mathbf{U}(t)\|_{\mathbf{H}^s}. \tag{2.7}$$

$\mathcal{C}^T_{0,\mathbf{H}^s} = C_0\left([-\infty, T]; \mathbf{H}^s\right)$, $T > 0$, is a Banach space of $\mathbf{H}^s$-valued continuous trajectories $\mathbf{U}(t)$, $-\infty < t \leq T$, such that $\mathbf{U}(t) = 0$, $-\infty < t \leq 0$ equipped with the norm (2.7).

$\mathcal{C}^T_{0,Y} = C_0\left([-\infty, T]; Y\right)$ is defined similarly for a Banach space Y.

$\mathcal{L}^T_Y = L_2\left([-\infty, T]; Y\right)$ is the space of Y-valued functions of $t \in [-\infty, T]$ that are square Lebesgue integrable; the norm in $\mathcal{L}^T_Y$ is defined by

$$||\mathbf{U}||^2_{\mathcal{L}^T_Y} = \int_{-\infty}^{T} ||\mathbf{U}(t)||^2_Y \, dt. \tag{2.8}$$

$L_{2,0}\left([-\infty, T]; Y\right)$ is the subspace of functions j from $L_2\left([-\infty, T]; Y\right)$ such that $j(t) = 0$, $-\infty < t \leq 0$.

$L_1\left([-\infty, T; Y]\right)$ is the space of Y-valued functions of $t \in [-\infty, T]$ with the norm

$$||j||_{L_1([-\infty,T;Y])} = \int_0^T \left\|j\left(t'\right)\right\|_Y \, dt'. \tag{2.9}$$

$L_{1,0}\left([-\infty, T]; Y\right)$ is the subspace of functions j from $L_1\left([-\infty, T]; Y\right)$ such that $j(t) = 0$, $-\infty < t \leq 0$.

$C^s\left(\mathbb{R}^3\right) = C^s$, $s = 1, 2, \ldots$, is the space of s times continuously differentiable vector fields or n-linear form (tensor) fields. The function norm in $C^s\left(\mathbb{R}^3\right)$ is defined by the formula

$$\|\mathbf{V}\|_{C^s} = \sup_{|\alpha|\leq s,\ \mathbf{r}\in\mathbb{R}^3} \left|\partial^\alpha V(\mathbf{r})\right|, \tag{2.10}$$

where for a vector $\partial^\alpha V(\mathbf{r})$ with given α, $\mathbf{r}$ the norm $|\partial^\alpha V(\mathbf{r})|$ is determined by (2.3) and for a n-linear form (tensor) field the norm $|\partial^\alpha V(\mathbf{r})|$ of a tensor $\partial^\alpha V(\mathbf{r})$ is determined using (2.4).

$\mathbf{H}_M$ is a Hilbert space consisting of the 6-dimensional fields from $\mathbf{L}^2_2$ but with the modified scalar product that includes a positive definite Hermitian matrix $\boldsymbol{\eta}(\mathbf{r})$ of the form

$$(\mathbf{U}, \mathbf{V})_{\mathbf{H}_M} = \int_{\mathbb{R}^3} \overline{\mathbf{U}}(\mathbf{r}) \cdot \Xi(\mathbf{r}) \mathbf{V}(\mathbf{r}) \, d\mathbf{r},\ \Xi(\mathbf{r}) = \begin{bmatrix} \boldsymbol{\eta}(\mathbf{r}) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \mathbf{U} = \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}. \tag{2.11}$$

Under the condition (3.2), which will be imposed on $\boldsymbol{\eta}(\mathbf{r})$, the norm $\|\cdot\|_{\mathbf{H}_M}$ is equivalent to the $\mathbf{L}_2$-norm $\|\cdot\|_{\mathbf{L}_2}$.

The space $\mathring{\mathbf{H}}_M$ consists of the 6-dimensional fields from $\mathring{\mathbf{L}}^2_2$ with the scalar product (2.11),

$$(\mathbf{U}, \mathbf{V})_{\mathbf{H}_M} = (\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M}. \tag{2.12}$$

$\mathring{\mathbf{H}}^s_M$ are the spaces generated by the linear Maxwell operator; they are considered in Sect. 3, see (3.7).

2.2. Solutions and their uniqueness. To study the uniqueness problem we express $\mathbf{D}$ in terms of $\mathbf{E}$ using (1.8). Equation (1.1), (1.2), after taking into account (1.7), (1.8), turn into

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\partial_t \mathbf{B}(\mathbf{r}, t) - 4\pi \mathbf{J}_B(\mathbf{r}, t), \tag{2.13}$$

$$\nabla \times \mathbf{B}(\mathbf{r}, t) = \boldsymbol{\varepsilon}\partial_t \mathbf{E}(\mathbf{r}, t) + 4\pi\,\partial_t \mathbf{P}_{\mathrm{NL}}(\mathbf{E})(\mathbf{r}, t) + 4\pi \mathbf{J}_D(\mathbf{r}, t), \tag{2.14}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0, \nabla \cdot (\boldsymbol{\varepsilon}\mathbf{E} + 4\pi \mathbf{P}_{\mathrm{NL}}(\mathbf{E}))(\mathbf{r}, t) = 0. \tag{2.15}$$

It is assumed that $\mathbf{J}_B, \mathbf{J}_D \in L_{1,0}([-\infty, T]; \mathbf{H}^s)$ for an integer $s > 3/2$. We impose the following condition on the nonlinearity $\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$.

Condition 2.1. *We assume that the nonlinear operators $\mathbf{P}_{NL}(\mathbf{E})$ and $\partial_t \mathbf{P}_{NL}(\mathbf{E})$ are defined for $\|\mathbf{E}\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$ with $R_{\mathbf{P}} > 0$ and an integer $s > 3/2$. For every $T_1 \in [0, T]$ they satisfy the Lipschitz condition*

$$\int_0^{T_1} \|\partial_t \mathbf{P}_{NL}(\mathbf{E}_1)(t) - \partial_t \mathbf{P}_{NL}(\mathbf{E}_2)(t)\|^2_{\mathbf{H}^0}\, dt \leq K_L \int_0^{T_1} \|\mathbf{E}_1(t) - \mathbf{E}_2(t)\|^2_{\mathbf{H}^0}\, dt \tag{2.16}$$

for every $\mathbf{E}_1, \mathbf{E}_2 \in \mathcal{C}^T_{0,\mathbf{H}^s}$ such that $\|\mathbf{E}_1\|_{\mathcal{C}^T_{\mathbf{H}^s}}, \|\mathbf{E}_2\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$.

Now we are ready to define a solution to (2.13), (2.14), (2.15).

Definition 2.2. *A pair of functions $\mathbf{B}(\mathbf{r}, t)$, $\mathbf{E}(\mathbf{r}, t)$ is called a solution of (2.13), (2.14), (2.15) if for some $T > 0$ we have $\mathbf{B} \in \mathcal{C}^T_{0,\mathbf{H}^s} = C_0([-\infty, T]; \mathbf{H}^s)$, $\mathbf{E} \in \mathcal{C}^T_{0,\mathbf{H}^s} = C_0([-\infty, T]; \mathbf{H}^s)$, $\partial_t \mathbf{B} \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$, $\partial_t \mathbf{E} \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$ with $s \geq 2$ and $\|\mathbf{E}\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$. The corresponding quad $\mathbf{D}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$, $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{H}(\mathbf{r}, t)$ with $\mathbf{H}(\mathbf{r}, t)$ and $\mathbf{D}(\mathbf{r}, t)$ determined respectively by (1.7), (1.8) is called a solution to (1.1), (1.2), (1.6), (1.7), (1.8).*

Note that the curl $\nabla^{\times}$ and the divergency $\nabla\cdot$ are bounded operators from $\mathbf{H}^s$ to $\mathbf{H}^{s-1}$, and when they are applied to functions of $(\mathbf{r}, t)$ they become bounded operators from $\mathcal{C}^T_{\mathbf{H}^s} = C([-\infty, T]; \mathbf{H}^s)$ to $\mathcal{C}^T_{\mathbf{H}^{s-1}}$. Therefore for $\|\mathbf{E}\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$ the left-hand and right-hand sides of (2.13), (2.14), (2.15) are well-defined as elements of $\mathcal{C}^T_{\mathbf{H}^{s-1}}$.

The next lemma provides a sufficient condition for Condition 2.1 to hold.

Lemma 2.3. *Let Condition 1.2 hold. Then Condition 2.1 holds, and $R_{\mathbf{P}} = R_{\mathbf{P}}(\beta_P)$ is the same as in Lemma 7.4, $R_{\mathbf{P}}$ depends only on β_P from Condition 1.2.*

Proof. The statement follows from Lemma 7.4. □

The following theorem shows that Condition 2.1 (and consequently Condition 1.2) implies uniqueness of solutions.

Theorem 2.4. *Let Condition 2.1 hold together with (1.10) and all conditions from Condition 1.2 with only one exception, namely (1.11) holds for $s = 0$. Let $\mathbf{J}_D, \mathbf{J}_B \in \mathcal{C}^T_{0,\mathbf{H}^1}$, and suppose that $\mathbf{B}_1, \mathbf{E}_1 \in C_0\left([-\infty, T]; \mathbf{H}^2\right)$ and $\mathbf{B}_2, \mathbf{E}_2 \in C_0\left([-\infty, T]; \mathbf{H}^2\right)$ are two solutions to (2.13), (2.14). Then $\mathbf{B}_1 = \mathbf{B}_2$, $\mathbf{E}_1 = \mathbf{E}_2$.*

Proof. Note that by Definition 2.2 the solutions satisfy $||\mathbf{E}_1||_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$, $||\mathbf{E}_2||_{\mathcal{C}^T_{\mathbf{H}^s}} < R_{\mathbf{P}}$, $\partial_t \mathbf{B}_1 \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$, $\partial_t \mathbf{E}_1 \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$, $\partial_t \mathbf{B}_2 \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$, $\partial_t \mathbf{E}_2 \in \mathcal{C}^T_{0,\mathbf{H}^{s-1}}$. Therefore the difference $\mathbf{B}_3 = \mathbf{B}_1 - \mathbf{B}_2$, $\mathbf{E}_3 = \mathbf{E}_1 - \mathbf{E}_2$ of the solutions satisfies the system

$$\boldsymbol{\varepsilon}^{-1} (\nabla \times \mathbf{B}_3) = \partial_t \mathbf{E}_3 + 4\pi \boldsymbol{\varepsilon}^{-1} \partial_t (\mathbf{P}_{\mathrm{NL}} (\mathbf{E}_1) - \mathbf{P}_{\mathrm{NL}} (\mathbf{E}_2)), \tag{2.17}$$

$$\nabla \times \mathbf{E}_3 = -\partial_t \mathbf{B}_3, \tag{2.18}$$

$\mathbf{E}_3 \in C_0 \left([-\infty, T]; \mathbf{H}^2\right)$. Now we, first, dot-multiply Eqs. (2.17) and (2.18) by respectively $\boldsymbol{\varepsilon}^{\mathrm{T}} \overline{\mathbf{E}}_3$ and $-\overline{\mathbf{B}}_3$, and then add up them, integrate in $\mathbf{r}$ and t and take the real part. We have

$$\int \left[\overline{\mathbf{E}}_3 \cdot \nabla \times \mathbf{B}_3 - \overline{\mathbf{B}}_3 \cdot \nabla \times \mathbf{E}_3 + \mathbf{E}_3 \cdot \nabla \times \overline{\mathbf{B}}_3 - \mathbf{B}_3 \cdot \nabla \times \overline{\mathbf{E}}_3\right] d\mathbf{r} = \mathbf{0}, \tag{2.19}$$

hence, for $T_1 \leq T$,

$$\mathrm{Re} \int_{-\infty}^{T_1} \int \left(\overline{\mathbf{E}}_3 \cdot \boldsymbol{\varepsilon} \partial_t \mathbf{E}_3 + \mathbf{B}_3 \partial_t \mathbf{B}_3\right) d\mathbf{r} dt = \mathrm{Re} \int_0^{T_1} \int \overline{\mathbf{E}}_3 \cdot \mathbf{g}_2 d\mathbf{r} dt, \tag{2.20}$$

$$\mathbf{g}_2 = 4\pi \partial_t (\mathbf{P}_{\mathrm{NL}} (\mathbf{E}_1) - \mathbf{P}_{\mathrm{NL}} (\mathbf{E}_2)). \tag{2.21}$$

Let us introduce

$$N(\mathbf{E}, \mathbf{B}, T) = \int \left(\overline{\mathbf{E}} \cdot \boldsymbol{\varepsilon} \mathbf{E} + \overline{\mathbf{B}} \cdot \mathbf{B}\right) d\mathbf{r} \Big|_{t=T} = ||(\mathbf{E}, \mathbf{B})||^2_{\mathring{\mathbf{H}}_M}, \tag{2.22}$$

where the norm $\mathring{\mathbf{H}}_M$ is defined in (2.11). Since $\partial_t \mathbf{B}_3 \in \mathcal{C}^T_{0,\mathbf{H}^1}$, $\partial_t \mathbf{E}_3 \in \mathcal{C}^T_{0,\mathbf{H}^1}$, $\mathbf{B}_3 \in \mathcal{C}^T_{0,\mathbf{H}^2}$, $\mathbf{E}_3 \in \mathcal{C}^T_{0,\mathbf{H}^2}$ and $\boldsymbol{\varepsilon}$ is Hermitian we have

$$\mathrm{Re} \int_{-\infty}^{T_1} \int \left(\overline{\mathbf{E}}_3 \cdot \boldsymbol{\varepsilon} \partial_t \mathbf{E}_3 + \overline{\mathbf{B}}_3 \cdot \partial_t \mathbf{B}_3\right) d\mathbf{r} dt = \frac{1}{2} N(\mathbf{E}_3, \mathbf{B}_3, T_1). \tag{2.23}$$

Then we estimate the right-hand side of (2.20) using (1.10) as follows

$$\begin{aligned} \left| \int \overline{\mathbf{E}}_3 \cdot \mathbf{g}_2 \, d\mathbf{r} \right| &\leq \frac{1}{2} \left(||\mathbf{E}_3(t)||^2_{\mathbf{H}^0} + ||\mathbf{g}_2(t)||^2_{\mathbf{H}^0} \right) \\ &\leq \frac{1}{2} ||\mathbf{E}_3(t)||^2_{\mathbf{H}^0} + 2\pi^2 \int |\partial_t [\mathbf{P}_{\mathrm{NL}}(\mathbf{E}_1)(t) - \mathbf{P}_{\mathrm{NL}}(\mathbf{E}_2)(t)]|^2 \, d\mathbf{r}. \end{aligned} \tag{2.24}$$

Based on Condition 2.1 we get

$$\int_0^{T_1} \int |\partial_t [\mathbf{P}_{\mathrm{NL}}(\mathbf{E}_1)(t) - \mathbf{P}_{\mathrm{NL}}(\mathbf{E}_2)(t)]|^2 \, d\mathbf{r} dt \leq K_L \int_0^{T_1} ||\mathbf{E}_3(t)||^2_{\mathbf{H}^0} \, dt. \tag{2.25}$$

According to (1.10) we have

$$\varepsilon_- ||\mathbf{E}_3(t)||^2_{\mathbf{H}^0} \leq N^2(\mathbf{E}_3, \mathbf{B}_3, t). \tag{2.26}$$

Now combining (2.20) and (2.23), (2.24), (2.25) and (2.26) we obtain

$$N^2(\mathbf{E}, \mathbf{B}, T_1) \leq \left(1 + 4\pi^2 K_L\right) \varepsilon_-^{-1} \int_0^{T_1} N^2(\mathbf{E}, \mathbf{B}, t) \, dt. \tag{2.27}$$

Applying the Gronwall inequality to the estimate (2.27) we get for $T_1 \geq 0$,

$$N^2(\mathbf{E}_3, \mathbf{B}_3, T_1) \leq N^2(\mathbf{E}_3, \mathbf{B}_3, 0) \exp\left[\left(1 + 4\pi^2 K_L\right) \varepsilon_-^{-1} T_1\right]. \tag{2.28}$$

Since $N^2(\mathbf{E}_3, \mathbf{B}_3, 0) = 0$, the inequality (2.28) implies $N^2(\mathbf{E}_3, \mathbf{B}_3, T_1) = 0$ and, consequently $\mathbf{E}_3 = \mathbf{B}_3 = \mathbf{0}$. □

Remark 2.5. Note that the proof of the uniqueness of solutions does not use the divergence-free condition (2.15), which is very essential for the proof of their existence.

The proof of the existence uses a reduction of (1.1), (1.2), (1.6) and (1.8) to an integral equation for the divergence-free fields **B** and **D** considered in Sect. 7.

3. Linear Maxwell Operator

In this section we consider some important properties of the linear Maxwell equations. In the linear case (1.8) takes the form

$$\mathbf{E}(\mathbf{r}, t) = \boldsymbol{\eta}(\mathbf{r}) \mathbf{D}(\mathbf{r}, t), \quad \boldsymbol{\eta}(\mathbf{r}) = [\boldsymbol{\varepsilon}(\mathbf{r})]^{-1}, \tag{3.1}$$

where the Hermitian matrix $\boldsymbol{\eta}(\mathbf{r})$, called the *impermeability* tensor, satisfies the inequality

$$\varepsilon_+^{-1} |\mathbf{e}|^2 \leq \sum_{m,n=1}^{3} \boldsymbol{\varepsilon}_{mn}(\mathbf{r}) e_m^* e_n \leq \varepsilon_-^{-1} |\mathbf{e}|^2, \ \mathbf{r} \in \mathbb{R}^3, \mathbf{e} \in \mathbb{C}^3, \tag{3.2}$$

as it follows from (1.10). Based on (1.7), (3.1) we rewrite Eqs. (1.1),(1.2) in the form

$$\partial_t \mathbf{U}(t) = -i\mathbf{M}\mathbf{U}(t) - \mathbf{J}(t); \ \mathbf{U}(t) = 0 \text{ for } t \leq 0, \tag{3.3}$$

where

$$\mathbf{U} = \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}, \ \mathbf{M} = i \begin{bmatrix} \mathbf{0} & \nabla^\times \\ -\nabla^\times \boldsymbol{\eta} & \mathbf{0} \end{bmatrix}, \ \nabla^\times \mathbf{B} = \nabla \times \mathbf{B}, \mathbf{J} = 4\pi \begin{bmatrix} \mathbf{J}_D \\ \mathbf{J}_B \end{bmatrix}, \tag{3.4}$$

and $\boldsymbol{\eta}$ denotes the operator of multiplication by $\boldsymbol{\eta}(\mathbf{r})$. We write the linear Maxwell operator **M** in the form

$$\mathbf{M} = i\nabla^{\times\times}\Xi, \ \nabla^{\times\times} = \begin{bmatrix} \mathbf{0} & \nabla^\times \\ -\nabla^\times & \mathbf{0} \end{bmatrix},$$
$$[\Xi\mathbf{V}](\mathbf{r}) = \Xi(\mathbf{r})\mathbf{V}(\mathbf{r}), \ \Xi(\mathbf{r}) = \begin{bmatrix} \boldsymbol{\eta}(\mathbf{r}) & \mathbf{0} \\ 0 & \mathbf{1} \end{bmatrix}, \tag{3.5}$$

where $\nabla^\times$ is the curl operator. In view of (3.2) we have

$$\alpha_- I_6 \leq \Xi(\mathbf{r}) \leq \alpha_+ I_6, \ \mathbf{r} \in \mathbb{R}^3, \text{ with } I_6 \text{ being } 6 \times 6 \text{ identity matrix} \tag{3.6}$$

and $\alpha_+ = \max\left(1, \varepsilon_-^{-1}\right)$, $\alpha_- = \max\left(1, \varepsilon_+^{-1}\right)$. We introduce now the scale of Hilbert spaces $\mathring{\mathbf{H}}_M^s$, $s = 0, 1, \ldots$, consisting of divergence free $(3+3)$-dimensional vector-fields $\mathbf{V}(\mathbf{r})$ with the scalar product

$$(\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M^s} = (\mathbf{M}^s\mathbf{U}, \mathbf{M}^s\mathbf{V})_{\mathring{\mathbf{H}}_M} + (\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M}, \ s = 0, 1, \ldots, \tag{3.7}$$

where $(\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M}$ is defined by (2.11). Evidently, $(\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M^0} = 2(\mathbf{U}, \mathbf{V})_{\mathring{\mathbf{H}}_M}$. In the following subsections we study properties of the spaces $\mathring{\mathbf{H}}_M^s$.

3.1. Spaces of divergence-free fields. We consider the standard Hilbert space $\mathbf{L}_2 = L_2\left(\mathbb{R}^3, \mathbb{C}^3\right)$ of Lebesgue square-integrable 3-dimensional complex-valued vector fields in $\mathbb{R}^3$, and we consider a subspace $\mathring{\mathbf{L}}_2$ of $\mathbf{L}_2$, which is the closure of all smooth vector fields from $\mathbf{L}_2$ with zero divergence. We denote by Π_0 the $\mathbf{L}_2$-orthogonal projection operator onto $\mathring{\mathbf{L}}_2$. The space $\mathring{\mathbf{L}}_2$ can be equivalently defined as a space of all the fields orthogonal to every field of the form $\operatorname{grad}\varphi(\mathbf{r})$, where φ runs the space $C_0^\infty\left(\mathbb{R}^3\right)$ (the set of infinitely differentiable scalar functions with finite support). The space $\mathring{\mathbf{L}}_2$ can be explicitly described in terms of the Fourier transform $\mathcal{F}$ which is given by the following formula

$$\tilde{\mathbf{A}}(\mathbf{k}) = \mathcal{F}(\mathbf{A})(\mathbf{k}) = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} e^{-i\mathbf{k}\cdot\mathbf{r}}\mathbf{A}(\mathbf{r})\,d\mathbf{r},$$

$$\mathbf{A}(\mathbf{r}) = \mathcal{F}^{-1}\left(\tilde{\mathbf{A}}\right) = \int_{\mathbb{R}^3} e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{\mathbf{A}}(\mathbf{k})\,d\mathbf{r}. \tag{3.8}$$

Note that

$$\widetilde{\partial^\alpha \mathbf{A}}(\mathbf{k}) = i^{|\alpha|}\mathbf{k}^\alpha\widetilde{\mathbf{A}}(\mathbf{k}),\ \mathbf{k}^\alpha = k_1^{\alpha_1}k_2^{\alpha_2}k_3^{\alpha_3}. \tag{3.9}$$

By Plancherel's theorem

$$\frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3}|\mathbf{A}(\mathbf{r})|^2\,d\mathbf{r} = \int_{\mathbb{R}^3}\left|\tilde{\mathbf{A}}(\mathbf{k})\right|^2 d\mathbf{k}. \tag{3.10}$$

Hence the Sobolev norm (2.2) can be written in terms of the Fourier transform as

$$\|\mathbf{A}\|_{\mathbf{H}^s}^2 = (2\pi)^{3/2}\int_{\mathbb{R}^3}\sum_{|\alpha|\le s}\left|\mathbf{k}^\alpha\right|^2\left|\tilde{\mathbf{A}}(\mathbf{k})\right|^2 d\mathbf{k}. \tag{3.11}$$

An equivalent norm is given in terms of the Fourier transform by

$$\|\mathbf{A}\|_{s,F}^2 = (2\pi)^{3/2}\int\left(|\mathbf{k}|^{2s}+1\right)|\mathcal{F}(\mathbf{A})(\mathbf{k})|^2\,d\mathbf{k}. \tag{3.12}$$

Obviously,

$$c_H^{-1}\|\mathbf{A}\|_{\mathbf{H}^s}^2 \le \|\mathbf{A}\|_{s,F}^2 \le c_H\|\mathbf{A}\|_{\mathbf{H}^s}^2. \tag{3.13}$$

According to (3.9) the image $\mathcal{F}\mathring{\mathbf{L}}_2$ of the Fourier transform $\mathcal{F}$ is:

$$\mathcal{F}\mathring{\mathbf{L}}_2 = \left\{\tilde{\mathbf{A}}(\mathbf{k}) \in \mathbf{L}_2 : \tilde{\mathbf{A}}(\mathbf{k})\cdot\mathbf{k} = 0 \text{ for almost all } \mathbf{k} \in \mathbb{R}^3\right\}. \tag{3.14}$$

By (3.12) the space $\mathring{\mathbf{H}}^s$ consists of functions

$$\left\{\mathbf{U} \in \mathbf{L}_2 : \left(|\mathbf{k}|^2+1\right)^{s/2}\mathcal{F}(\mathbf{U})(\mathbf{k}) \in \mathbf{L}_2,\ \mathbf{k}\cdot\mathcal{F}(\mathbf{U})(\mathbf{k}) = 0\right\}. \tag{3.15}$$

The projection Π_0 in terms of the Fourier transform is written explicitly for every $\mathbf{k}$ as the orthogonal projection in $\mathbb{C}^3$ of a vector $\tilde{\mathbf{A}}(\mathbf{k})$ onto the plane $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k}) = 0$, i.e.

$$\Pi_0 = \mathcal{F}^{-1}\tilde{\Pi}_0\mathcal{F},\quad \tilde{\Pi}_0\tilde{\mathbf{A}}(\mathbf{k}) = \tilde{\mathbf{A}}(\mathbf{k}) - \frac{\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k})}{(\mathbf{k},\mathbf{k})}\mathbf{k}, \tag{3.16}$$

and the curl operator takes the form

$$\nabla^{\times} = \mathcal{F}^{-1}\tilde{\nabla}^{\times}\mathcal{F}, \qquad \tilde{\nabla}^{\times}\tilde{\mathbf{A}}(\mathbf{k}) = i\mathbf{k}\times\tilde{\mathbf{A}}(\mathbf{k}). \tag{3.17}$$

Evidently,

$$\Pi_0^2 = \Pi_0,\ \nabla^{\times}\Pi_0 = \Pi_0\nabla^{\times} = \Pi_0\nabla^{\times}\Pi_0 = \nabla^{\times}. \tag{3.18}$$

Since $\left|\tilde{\Pi}_0\tilde{\mathbf{A}}(\mathbf{k})\right| \leq \left|\tilde{\mathbf{A}}(\mathbf{k})\right|$ for every $\mathbf{k}$, the operator Π_0 has norm 1 in both norms (3.12) and (2.2) for every s,

$$\|\Pi_0\|_{\mathbf{H}^s,\mathbf{H}^s} = 1. \tag{3.19}$$

For any operator K acting in $\mathbf{L}_2$ we introduce an operator $\mathring{K} = \Pi_0 K \Pi_0$ that acts in $\mathring{\mathbf{L}}_2 \subset \mathbf{L}_2$, in particular,

$$\mathring{\boldsymbol{\eta}}\mathbf{v} = \Pi_0\boldsymbol{\eta}\Pi_0\mathbf{v} = \Pi_{\mathbf{0}}\boldsymbol{\eta}\mathbf{v}, \quad \mathbf{v}\in\mathring{\mathbf{L}}_2. \tag{3.20}$$

Notice that (3.16)–(3.20) imply

$$\Pi_0 K \Pi_0 = \mathring{K}\Pi_0,\ \mathring{\nabla}^{\times}\Pi_0 = \nabla^{\times}\Pi_0 = \nabla^{\times},\ \mathring{\Delta}\Pi_0 = \Delta\Pi_0. \tag{3.21}$$

Notice that if the medium is homogeneous and isotropic, i.e. $\Xi = I_6$, then the Maxwell operator has constant coefficients and Maxwell equations can be solved explicitly in terms of the Fourier transform determined by (3.8). In this case ellipticity of the curl operator $\nabla^{\times}$ on divergence free fields can be shown to be elementary using (3.17). According to the well-known property of the cross-product we have

$$\mathbf{k}\times\left(\mathbf{k}\times\tilde{\mathbf{A}}(\mathbf{k})\right) = -|\mathbf{k}|^2\tilde{\mathbf{A}}(\mathbf{k}) + \left(\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k})\right)\mathbf{k} \tag{3.22}$$

and, hence,

$$\begin{aligned}\left|\tilde{\nabla}^{\times}\tilde{\mathbf{A}}(\mathbf{k})\right|^2 &= \left(\mathbf{k}\times\tilde{\mathbf{A}}(\mathbf{k})\right)\cdot\left(\mathbf{k}\times\overline{\tilde{\mathbf{A}}}(\mathbf{k})\right)\\ &= -\left(\mathbf{k}\times\left(\mathbf{k}\times\tilde{\mathbf{A}}(\mathbf{k})\right)\right)\cdot\overline{\tilde{\mathbf{A}}}(\mathbf{k}) = \left|\tilde{\mathbf{A}}(\mathbf{k})\right|^2\cdot|\mathbf{k}|^2 - \left|\left(\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k})\right)\right|^2.\end{aligned}$$

Since $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k}) = 0$ for $\mathbf{A}\in\mathring{\mathbf{L}}_2$, the Sobolev norm in (3.12) coincides on $\mathring{\mathbf{L}}_2\times\mathring{\mathbf{L}}_2$ with the norm defined in terms of the curl operator $\nabla^{\times\times}$:

$$\begin{aligned}||\mathbf{U}||_{s,F}^2 &= \left|\left|\left(\nabla^{\times\times}\right)^s\mathbf{U}\right|\right|_{\mathring{\mathbf{L}}_2}^2 + ||\mathbf{U}||_{\mathring{\mathbf{L}}_2}^2 \\ &= (2\pi)^{3/2}\int\left[\left|\left(\tilde{\nabla}^{\times\times}\right)^s\mathcal{F}(\mathbf{U})(\mathbf{k})\right|^2 + |\mathcal{F}(\mathbf{U})(\mathbf{k})|^2\right]d\mathbf{k},\end{aligned} \tag{3.23}$$

and by (3.13) this norm is equivalent to the norm in (3.12), i.e. there exists a finite positive constant c_H such that

$$c_H^{-1}||\mathbf{U}||_{\mathring{\mathbf{H}}^s}^2 \leq \left|\left|\left(\nabla^{\times\times}\right)^s\mathbf{U}\right|\right|_{\mathring{\mathbf{L}}_2}^2 + ||\mathbf{U}||_{\mathring{\mathbf{L}}_2}^2 \leq c_H||\mathbf{U}||_{\mathring{\mathbf{H}}^s}^2. \tag{3.24}$$

An analogous property for the case of variable coefficients is given in the next subsection (see (3.41), (3.42)).

3.2. Maxwell operator with variable coefficients. When coefficients of the matrix η are variable we consider the linear Maxwell operator $\mathring{\mathbf{M}}$:

$$\mathring{\mathbf{M}} \text{ is the restriction of } \mathbf{M} = \nabla^{\times\times}\Xi \text{ to } \mathring{\mathbf{L}}_2 \times \mathring{\mathbf{L}}_2, \mathbf{M} = \nabla^{\times\times}\mathring{\Xi} \tag{3.25}$$

with $\nabla^{\times\times}$ and $\Xi = \Xi(\mathbf{r})$ defined in (3.5). The operator $\mathring{\mathbf{M}}$ is self-adjoint in $\mathring{\mathbf{H}}_M$, namely the following well-known lemma holds (for the proof see [9, 17]).

Lemma 3.1. *Assume that: (i) $\boldsymbol{\eta}(\mathbf{r})$ is a 3×3 Hermitian matrix that satisfies (3.2); (ii) $\boldsymbol{\eta}(\mathbf{r})$ has bounded measurable coefficients (in particular, it is sufficient that $\boldsymbol{\eta}(\cdot) \in C^0(\mathbb{R}^3)$). Then the operator $\mathring{\mathbf{M}}$ is self-adjoint in the space $\mathring{\mathbf{H}}_M$ with the scalar product defined by (2.11), the domain of $\mathring{\mathbf{M}}$ is $\mathring{\mathbf{H}}^1 \times \mathring{\mathbf{H}}^1$, $\mathring{\mathbf{H}}^1$ being defined by (3.15).*

The following lemma plays an important role in our analysis of the nonlinear Maxwell equations, in particular it is used to estimate the norms of $\mathring{\boldsymbol{\eta}}$ and $\mathring{\boldsymbol{\eta}}^{-1}$ that are included in the nonlinearity according to (7.35).

Lemma 3.2. *Let $s \geq 0$ be an integer, and $\boldsymbol{\eta}(\mathbf{r}) \in C^s$ be a 3×3 Hermitian matrix satisfying (3.2). Then there exist positive constants $c_\pm = c_\pm(s)$ such that*

$$c_- \|\mathbf{v}\|_{\mathring{\mathbf{H}}^s} \leq \left\|\mathring{\boldsymbol{\eta}}\mathbf{v}\right\|_{\mathring{\mathbf{H}}^s} \leq c_+ \|\mathbf{v}\|_{\mathring{\mathbf{H}}^s}, \quad \mathbf{v} \in \mathring{\mathbf{H}}^s. \tag{3.26}$$

The operator $\mathring{\boldsymbol{\eta}}$ in $\mathring{\mathbf{H}}^s$ has a bounded inverse $\mathring{\boldsymbol{\eta}}^{-1}$, $\left\|\mathring{\boldsymbol{\eta}}^{-1}\right\| \leq c_-^{-1}$. In addition to that, there exist positive constants $c'_\pm$ such that

$$c'_- \|\mathbf{v}\|_{\mathring{\mathbf{H}}^s} \leq \left\|\left(\mathring{\nabla}^\times\right)^s \mathring{\boldsymbol{\eta}}\mathbf{v}\right\|^2_{\mathring{\mathbf{L}}_2} + \left\|\mathring{\boldsymbol{\eta}}\mathbf{v}\right\|^2_{\mathring{\mathbf{L}}_2} \leq c'_+ \|\mathbf{v}\|_{\mathring{\mathbf{H}}^s}. \tag{3.27}$$

Proof. Let us show first that the statement of Lemma 3.2 holds for $s = 0$, namely

$$\varepsilon_+^{-1} \|\mathbf{v}\|_{\mathring{\mathbf{L}}_2} \leq \left\|\mathring{\boldsymbol{\eta}}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2} \leq \varepsilon_-^{-1} \|\mathbf{v}\|_{\mathring{\mathbf{L}}_2}, \quad \mathbf{v} \in \mathring{\mathbf{L}}_2. \tag{3.28}$$

It follows from (3.2) and (3.20) that $\mathring{\boldsymbol{\eta}}$ satisfies for $\mathbf{v} \in \mathring{\mathbf{L}}_2$ the inequality

$$\begin{aligned}(\mathbf{v}, \mathbf{v})_{\mathring{\mathbf{L}}_2} &= (\Pi_0\mathbf{v}, \Pi_0\mathbf{v})_{\mathbf{L}_2} \\ &\leq \varepsilon_+ (\Pi_0\mathbf{v}, \boldsymbol{\eta}\Pi_0\mathbf{v})_{\mathbf{L}_2} = \varepsilon_+ (\mathbf{v}, \Pi_0\boldsymbol{\eta}\Pi_0\mathbf{v})_{\mathbf{L}_2} = \varepsilon_+ \left(\mathbf{v}, \mathring{\boldsymbol{\eta}}\mathbf{v}\right)_{\mathring{\mathbf{L}}_2}.\end{aligned}$$

Note that $\mathring{\boldsymbol{\eta}}$ is a bounded self-adjoint positive operator in $\mathring{\mathbf{L}}_2$ and there exists the square root $\sqrt{\mathring{\boldsymbol{\eta}}}$ which is a bounded positive self-adjoint operator too. Taking $\mathbf{v} = \sqrt{\mathring{\boldsymbol{\eta}}}\mathbf{u}$ we get for any $\mathbf{u} \in \mathring{\mathbf{L}}_2$ the inequality

$$\varepsilon_+ \left(\sqrt{\mathring{\boldsymbol{\eta}}}\mathbf{u}, \mathring{\boldsymbol{\eta}}\sqrt{\mathring{\boldsymbol{\eta}}}\mathbf{u}\right)_{\mathring{\mathbf{L}}_2} \geq \left(\sqrt{\mathring{\boldsymbol{\eta}}}\mathbf{u}, \sqrt{\mathring{\boldsymbol{\eta}}}\mathbf{u}\right)_{\mathring{\mathbf{L}}_2} = \left(\mathbf{u}, \mathring{\boldsymbol{\eta}}\mathbf{u}\right)_{\mathring{\mathbf{L}}_2} \geq \varepsilon_+^{-1} (\mathbf{u}, \mathbf{u})_{\mathring{\mathbf{L}}_2}, \tag{3.29}$$

therefore for any $\mathbf{v} \in \mathring{\mathbf{L}}_2$,

$$\left\|\mathring{\boldsymbol{\eta}}\mathbf{v}\right\|^2_{\mathring{\mathbf{L}}_2} = \left(\mathring{\boldsymbol{\eta}}\mathbf{v}, \mathring{\boldsymbol{\eta}}\mathbf{v}\right)_{\mathring{\mathbf{L}}_2} \geq \varepsilon_+^{-2} (\mathbf{v}, \mathbf{v})_{\mathring{\mathbf{L}}_2}. \tag{3.30}$$

We also derive from (3.2) that $(\boldsymbol{\eta}\mathbf{v}, \boldsymbol{\eta}\mathbf{v})_{\mathring{\mathbf{L}}_2} \leq \varepsilon_-^{-2} (\mathbf{v}, \mathbf{v})_{\mathring{\mathbf{L}}_2}$ and obtain (3.28).

Let us consider now $\partial^{\alpha}\left(\mathring{\boldsymbol{\eta}}\mathbf{v}\right)$ for a multiindex α such that $|\alpha| \leq s$. According to (3.9), (3.16) operators ∂^{α} commute with Π_0 and we have

$$\partial^{\alpha}\left(\mathring{\boldsymbol{\eta}}\mathbf{v}\right) - \mathring{\boldsymbol{\eta}}\partial^{\alpha}\mathbf{v} = \Pi_0\left[\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right) - \boldsymbol{\eta}\partial^{\alpha}\mathbf{v}\right], \text{ for } \mathbf{v} \in \mathring{\mathbf{H}}^s. \tag{3.31}$$

The relations (3.19) and (3.31) evidently imply the inequality

$$\left\|\partial^{\alpha}\left(\mathring{\boldsymbol{\eta}}\mathbf{v}\right) - \mathring{\boldsymbol{\eta}}\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2} \leq \left\|\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right) - \boldsymbol{\eta}\partial^{\alpha}\mathbf{v}\right\|_{\mathbf{L}_2}. \tag{3.32}$$

It follows from the Leibnitz formula applied to $\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right)$ that the difference $\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right) - \boldsymbol{\eta}\partial^{\alpha}\mathbf{v}$ will involve only the partial derivatives of $\mathbf{v}$ and $\boldsymbol{\eta}$ of respective order not exceeding $s-1$ and s. Combining this observation with (7.8) and the interpolation inequalities (3.64) we get the estimate

$$\left\|\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right) - \boldsymbol{\eta}\partial^{\alpha}\mathbf{v}\right\|_{\mathbf{L}_2} \leq \|\boldsymbol{\eta}\|_{C^s}\left(\varepsilon\|\mathbf{v}\|_{\mathbf{H}^s} + C_{s,\varepsilon}\|\mathbf{v}\|_{\mathbf{L}_2}\right) \tag{3.33}$$

which holds for any $0 < \varepsilon < 1$ with a constant $C_{s,\varepsilon}$ depending only on indicated parameters. For $\mathbf{v} \in \mathring{\mathbf{H}}^s$ evidently $\|\mathbf{v}\|_{\mathbf{H}^s} = \|\mathbf{v}\|_{\mathring{\mathbf{H}}^s}$ and, hence, (3.33) implies

$$\left\|\partial^{\alpha}\left(\boldsymbol{\eta}\mathbf{v}\right) - \boldsymbol{\eta}\partial^{\alpha}\mathbf{v}\right\|_{\mathbf{L}_2} \leq \|\boldsymbol{\eta}\|_{C^s}\left(\varepsilon\|\mathbf{v}\|_{\mathring{\mathbf{H}}^s} + C_{s,\varepsilon}\|\mathbf{v}\|_{\mathring{\mathbf{L}}_2}\right), \ 0 < \varepsilon < 1, \ \mathbf{v} \in \mathring{\mathbf{H}}^s. \tag{3.34}$$

Considering now $\mathring{\boldsymbol{\eta}}\partial^{\alpha}\mathbf{v}$ we notice that (3.28) implies

$$\varepsilon_+^{-1}\left\|\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2} \leq \left\|\mathring{\boldsymbol{\eta}}\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2} \leq \varepsilon_-^{-1}\left\|\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2}, \tag{3.35}$$

and, consequently,

$$\varepsilon_+^{-2}\sum_{|\alpha|\leq s}\left\|\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2}^2 \leq \sum_{|\alpha|\leq s}\left\|\mathring{\boldsymbol{\eta}}\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2}^2 \leq \varepsilon_-^{-2}\sum_{|\alpha|\leq s}\left\|\partial^{\alpha}\mathbf{v}\right\|_{\mathring{\mathbf{L}}_2}^2. \tag{3.36}$$

Combining (3.32), (3.33) and (3.36) we obtain

$$c_-\|\mathbf{v}\|_{\mathring{\mathbf{H}}^s} \leq \|\mathbf{v}\|_{\mathring{\mathbf{L}}_2} + \sum_{|\alpha|=s}\left\|\partial^{\alpha}\left(\mathring{\boldsymbol{\eta}}\mathbf{v}\right)\right\|_{\mathring{\mathbf{L}}_2} \leq c_+\|\mathbf{v}\|_{\mathring{\mathbf{H}}^s} \tag{3.37}$$

with constants $c_{\pm}$ depending only on s, $\varepsilon_{\pm}^{-1}$ and $\|\boldsymbol{\eta}\|_{C^s}$. The last inequalities readily imply the inequalities (3.26). By (3.30) the null-space of $\mathring{\boldsymbol{\eta}}$ is trivial and $\mathring{\boldsymbol{\eta}}^{-1}$ is bounded on the image of $\mathring{\boldsymbol{\eta}}$; since $\mathring{\boldsymbol{\eta}}$ is self-adjoint in $\mathring{\mathbf{L}}_2$ (3.30) implies that the image of $\mathring{\boldsymbol{\eta}}$ coincides with $\mathring{\mathbf{L}}_2$ and $\mathring{\boldsymbol{\eta}}^{-1}$ is a bounded operator defined on $\mathring{\mathbf{L}}_2$. Boundedness of $\mathring{\boldsymbol{\eta}}^{-1}$ in $\mathbf{H}^s$ follows from (3.26). To deduce (3.27) from (3.26) we apply (3.24). □

Let $\Pi_0^{(2)}$ be the orthogonal projector from $\mathbf{L}_2^2$ onto $\mathring{\mathbf{L}}_2^2$, i.e.

$$\Pi_0^{(2)}\begin{bmatrix}\mathbf{D}\\ \mathbf{B}\end{bmatrix} = \begin{bmatrix}\Pi_0\mathbf{D}\\ \Pi_0\mathbf{B}\end{bmatrix}. \tag{3.38}$$

Lemma 3.3. *Let $s \geq 0$ be an integer, $\Xi(\mathbf{r}) \in C^s$ be a 6×6 Hermitian matrix satisfying the condition*

$$\xi_- I_6 \leq \Xi(\mathbf{r}) \leq \xi_+ I_6 \quad \text{where } \xi_+ \geq \xi_- > 0, \mathbf{r} \in \mathbb{R}^3. \tag{3.39}$$

Then $\mathring{\Xi} = \Pi_0^{(2)} \Xi \mathring{\Pi}_0^{(2)}$ satisfies on $\mathring{\mathbf{L}}_2^2$ the following inequality analogous to (3.26):

$$c_- \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \left\| \mathring{\Xi} \mathbf{V} \right\|_{\mathring{H}^s \times \mathring{H}^s} \leq c_+ \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s}, \ \mathbf{V} \in \mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s, \tag{3.40}$$

and

$$c'_- \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \left\| \left(\mathring{\nabla}^{\times\times}\right)^s \mathring{\Xi} \mathbf{V} \right\|^2_{\mathring{\mathbf{L}}_2^2} + \|\mathbf{V}\|^2_{\mathring{\mathbf{L}}_2^2} \leq c'_+ \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s}. \tag{3.41}$$

The operator $\mathring{\Xi}$ in $\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s$ has a bounded inverse $\mathring{\Xi}^{-1}$, $\left\| \mathring{\Xi}^{-1} \right\| \leq c_-^{-1}$.

The proof of Lemma 3.3 is analogous to the proof of Lemma 3.2.

The following statement on the equivalence of the Hilbert spaces $\mathring{\mathbf{H}}^s_M$ and $\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s$ generalizes (3.24).

Lemma 3.4. *Suppose that $\boldsymbol{\eta}(\mathbf{r})$ satisfies all the conditions of Lemma 3.2 and the operator $\mathring{\mathbf{M}}$ is defined by (3.25). Then for any integer $s \geq 1$ there exist positive constants $c_\pm$ such that*

$$c_- \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s_M} \leq c_+ \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s}, \ \mathbf{V} \in \mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s. \tag{3.42}$$

The proof of Lemma 3.4 together with some auxiliary statements are subjects of the next subsection.

3.3. Abstract Sobolev spaces and the spaces equivalence. Notice that for integer values of $s \geq 0$ the spaces $\mathring{\mathbf{H}}^s$ and $\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s$ are generated respectively by the linear self-adjoint operators $\left[\mathring{\nabla}^{\times}\right]^s$ in $\mathring{\mathbf{L}}_2$ and $\left[i\mathring{\nabla}^{\times\times}\right]^s$ in $\mathring{\mathbf{L}}_2^2$. Indeed, it follows elementarily from the relations (3.8)–(3.18) that

$$\left[i\mathring{\nabla}^{\times\times}\right]^* i\mathring{\nabla}^{\times\times} = \left[i\mathring{\nabla}^{\times\times}\right]^2 = \begin{bmatrix} -\Delta\Pi_0 & \mathbf{0} \\ \mathbf{0} & -\Delta\Pi_0 \end{bmatrix}, \tag{3.43}$$

where Δ is the Laplace operator. From (3.24) we obtain that

$$c_H^{-1} \|\mathbf{V}\|^2_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \left\| \left(\mathring{\nabla}^{\times\times}\right)^s \mathbf{V} \right\|^2_{\mathring{\mathbf{L}}_2^2} + \|\mathbf{V}\|^2_{\mathring{\mathbf{L}}_2^2} \leq c_H \|\mathbf{V}\|^2_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s}. \tag{3.44}$$

To relate the Hilbert spaces $\mathring{\mathbf{H}}^s_M$ and $\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s$ when $\boldsymbol{\eta}(\mathbf{r})$ is not constant we use the concept of abstract Sobolev spaces generated by powers of linear operators, [49], Chapter 19.26. Namely, for a self-adjoint operator B in a Hilbert space $\mathcal{H}$ we consider its power B^s, $s \geq 1$ and equip its domain $\mathcal{D}(B^s) = \mathcal{H}_{B^s}$ with the graph norm and scalar product

$$\|u\|_{\mathcal{H}_{B^s}} = \sqrt{\|B^s u\|^2 + \|u\|^2}, \ (u, v)_{\mathcal{H}_{B^s}} = (Bu, Bv) + (u, v). \tag{3.45}$$

We recall that for a linear operator B its graph $G(B)$ is defined as the set of all pairs $\{u, Bu\}$ when u runs the domain $\mathcal{D}(B)$ of the operator B.

In view of (3.24) the norm equivalencies (3.24), (3.44) can be recast as

$$\mathring{\mathbf{H}}^s = \left[\mathring{\mathbf{L}}_2\right]_{\left[\mathring{\nabla}^{\times}\right]^s}, \ \mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s = \left[\mathring{\mathbf{L}}_2^2\right]_{\left[i\mathring{\nabla}^{\times\times}\right]^s}. \tag{3.46}$$

It is also evident, in view of (3.7) and (2.11), that

$$\mathring{\mathbf{H}}_M^s = \left[\mathring{\mathbf{L}}_2^2\right]_{\mathring{\mathbf{M}}^s}. \tag{3.47}$$

Hence, to show the equality of $\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s$ to $\mathring{\mathbf{H}}_M^s$, which is the statement of Lemma 3.4, it is sufficient to verify that

$$\left[\mathring{\mathbf{L}}_2^2\right]_{\mathring{\mathbf{M}}^s} = \left[\mathring{\mathbf{L}}_2^2\right]_{\mathring{\mathbf{M}}^s}. \tag{3.48}$$

The following abstract results are developed to establish (3.48).

Let us recall the basic concepts related to closed operators. For any linear operator B we consider, its domain $\mathcal{D}(B)$ is assumed to be dense in the Hilbert space $\mathcal{H}$ with the norm $\|\|_{\mathcal{H}} = \|\|$. A linear operator B is called *closable*, [28], Sect. III. 5.3, if and only if

$$u_n \in \mathcal{D}(B), \ u_n \to 0 \text{ and } Bu_n \to v \text{ imply } v = 0. \tag{3.49}$$

The *closure* $\bar{B}$ of a closable operator B has the graph $G(\bar{B})$ which is defined as the closure $\overline{G(B)}$ of the graph $G(B)$. For a closed operator B a set $\mathcal{D}$ is called its *core* if the closure of the restriction B on $\mathcal{D}$ is the operator B itself.

To deal with powers, products and sums of unbounded operators we introduce the following definitions.

Definition 3.5. *Let $B_1, B_2, \ldots, B_n$ be linear densely defined operators acting in $\mathcal{H}$. We define the product $B = B_1 B_2 \cdots B_n$ as a linear operator B acting naturally as*

$$Bu = B_1 (B_2 (\ldots (B_n u))) \textit{for } u \in \mathcal{D}(B), \tag{3.50}$$

where its domain $\mathcal{D}(B)$ is defined as the set of u such that

$$u \in \mathcal{D}(B_n), B_n u \in \mathcal{D}(B_{n-1}), \ldots, B_2 (\ldots (B_n u)) u \in \mathcal{D}(B_1). \tag{3.51}$$

If the domain $\mathcal{D}(B)$ is dense in $\mathcal{H}$ we call the product B densely defined.

Definition 3.6. *Let $B_1, B_2, \ldots, B_n$ be densely defined linear operators. Then the sum $B = B_1 + \cdots + B_n$ acts naturally as*

$$Bu = B_1 u + \cdots + B_n u, \tag{3.52}$$

where the domain $\mathcal{D}(B) = \mathcal{D}(B_1) \cap \cdots \cap \mathcal{D}(B_n)$. If the domain $\mathcal{D}(B)$ is dense in $\mathcal{H}$ we call the sum B densely defined.

Clearly the most important parts of the above definitions are the domains of the product and the sum, since the operators of interest are unbounded. One can easily verify that the above definitions are consistent and correct in the sense that the domains of the product and the sum are independent of how we group operators when forming the product and the sum. In particular, $\mathcal{D}(B_1(B_2B_3)) = \mathcal{D}((B_1B_2)B_3)$ with a similar equality holding for the sum.

To establish the identity of Hilbert spaces $\mathcal{H}_B$ generated by different operators B we introduce the following definition.

Definition 3.7. *We call two linear closed operators B_1 and B_2 equivalent and write $B_1 \sim B_2$ if $\mathcal{D}(B_1) = \mathcal{D}(B_2)$ and there exist positive constants γ_-, γ_+ $0 < \gamma_- \le \gamma_+$, such that for any $u \in \mathcal{D}(B_2)$,*

$$\gamma_- \left(\|B_2u\|^2 + \|u\|^2 \right) \le \|B_1u\|^2 + \|u\|^2 \le \gamma_+ \left(\|B_2u\|^2 + \|u\|^2 \right). \tag{3.53}$$

If B_1 and B_2 are linear operators defined on a dense domain $\mathcal{D}$, at least one of B_1, B_2 is closable, and the inequalities (3.53) hold for any $u \in \mathcal{D}$, then we write $B_1 \sim B_2$ on $\mathcal{D}$.

The following statement is useful for the verification of the equivalency of a two linear operators.

Lemma 3.8. *Let B_1 and B_2 be densely defined linear operators, and, in addition, B_2 be closable. Suppose that a set $\mathcal{D} \subseteq \mathcal{D}(B_1) \cap \mathcal{D}(B_2)$ is a core of $\bar{B}_2$. Suppose also that there exist positive numbers $\alpha_\pm$ and $\beta_\pm$ such that for any $u \in \mathcal{D}$,*

$$\alpha_- \|B_2u\|^2 - \beta_- \|u\|^2 \le \|B_1u\|^2 \le \alpha_+ \|B_2u\|^2 + \beta_+ \|u\|^2. \tag{3.54}$$

Then the following statements hold:

(i) the inequalities (3.53) and (3.54) are equivalent, and (3.54) implies (3.53) with $\gamma_+ = \max\{\alpha_+, 1+\beta_+\}$ and $\gamma_- = \min\{\alpha_-, \beta - \beta_-\}/\beta$, where $\beta = \beta_+ + \beta_- + 1$;
(ii) B_1 is closable, $B_1 \sim B_2$ on $\mathcal{D}$, and $\bar{B}_1 \sim \bar{B}_2$ including, in particular, $\mathcal{D}(\bar{B}_1) = \mathcal{D}(\bar{B}_2)$;
(iii) if we replace in the inequalities (3.53) and (3.54) B_1 and B_2 respectively with $\bar{B}_1$ and $\bar{B}_2$ these inequalities will hold for any $u \in \mathcal{D}(\bar{B}_1)$;
(iv) if $B_1 \sim B_2$ on $\mathcal{D}$ then both operators are closable and $\bar{B}_1 \sim \bar{B}_2$;
(v) the relation $B_1 \sim B_2$ between closed operators is an equivalency relation.

Proof. We begin with the statement (i). Indeed, (3.53) evidently implies (3.54). To show the opposite implication we set $\beta = \beta_+ + \beta_- + 1$ and, using the right-hand side of the inequality (3.54), obtain

$$\|B_1u\|^2 + \|u\|^2 \le \|B_2u\|^2 + (1+\beta_+)\|u\|^2 \le \gamma_+ \left(\|B_2u\|^2 + \|u\|^2 \right), \tag{3.55}$$

where $\gamma_+ = \max\{\alpha_+, 1+\beta_+\}$. Then using the left-hand side of the inequality (3.54) we get

$$\beta \left(\|B_1u\|^2 + \|u\|^2 \right) \ge \|B_1u\|^2 + \beta\|u\|^2 \ge \alpha_- \|B_2u\|^2 + (\beta - \beta_-)\|u\|^2 \tag{3.56}$$

implying, in turn,

$$\|B_1u\|^2 + \beta\|u\|^2 \ge \gamma_- \left(\|B_2u\|^2 + \|u\|^2 \right), \tag{3.57}$$

where $\gamma_- = \min\{\alpha_-, \beta - \beta_-\}/\beta$. This completes the proof of (i) and we may assume from now on that (3.53) holds.

Let us consider the graphs $G(B_1|_{\mathcal{D}})$ and $G(B_2|_{\mathcal{D}})$ of the corresponding restrictions of the operators B_1 and B_2 to the set $\mathcal{D}$. The inequalities (3.53) imply that for any sequence $u_n \in \mathcal{D}$ we have

$$\|u_n - u\|_{\mathcal{H}_{B_1}} \to 0 \text{ if and only if } \|u_n - u\|_{\mathcal{H}_{B_2}} \to 0 \tag{3.58}$$

and, in addition to that,

$$\gamma_- \left(\|v_2\|^2 + \|u\|^2\right) \leq \|v_1\|^2 + \|u\|^2 \leq \gamma_+ \left(\|v_2\|^2 + \|u\|^2\right), \tag{3.59}$$
$$\text{where } v_1 = \lim_{n\to\infty} B_1 u_n,\ v_2 = \lim_{n\to\infty} B_2 u_n.$$

Notice now that since B_2 is closed, then if $u = 0$ then, $v_2 = 0$ and, in view of (3.59), we may conclude that $v_1 = 0$, implying that B_1 is closable and $G\left(\bar{B}_1\right) = \overline{G(B_1|_{\mathcal{D}})}$. Since, according to the lemma conditions, $\mathcal{D}$ is a core of B_2 we also have $G(B_2) = \overline{G(B_2|_{\mathcal{D}})}$. Observe now that (3.58) implies: $\mathcal{D}\left(\bar{B}_1\right) = \mathcal{D}\left(\bar{B}_2\right)$; the inequalities (3.54) hold for any $u \in \mathcal{D}\left(\bar{B}_2\right)$. Hence, in accordance with Definition 3.7, we have $B_1 \sim B_2$ on $\mathcal{D}$, $\bar{B}_1 \sim \bar{B}_2$ and (ii) and (iii) are proven.

The proof of (iv) is based on the same arguments as the proofs of the statements (i)–(iii). The statement (v) follows from (i)–(iv) completing the lemma's proof. □

Definition 3.9. *Suppose that B and C are closed and densely defined operators. We say that C is subordinated to B, and write $C \prec B$, if $\mathcal{D}(B) \subseteq \mathcal{D}(C)$ and for every positive $\varepsilon < 1$ there exist a positive β_ε such that for any $u \in \mathcal{D}(B)$,*

$$\|Cu\|^2 \leq \varepsilon \|Bu\|^2 + \beta_\varepsilon \|u\|^2 . \tag{3.60}$$

If C and B are linear operators defined on a dense domain $\mathcal{D}$ and the inequalities (3.60) hold for every positive $\varepsilon < 1$ and every $u \in \mathcal{D}$, then we write $C \prec B$ on $\mathcal{D}$.

To verify the subordination of two operators we will be using the following statement.

Lemma 3.10. *Suppose that B and C are closable operators defined on a dense set $\mathcal{D}$ and that the inequalities (3.60) hold for every positive $\varepsilon < 1$ and every $u \in \mathcal{D}$. Then $\bar{C} \prec \bar{B}$.*

Proof. The proof follows immediately from Definition 3.9 and closability of B and C on $\mathcal{D}$. □

Now we prove a few technical statements.

Lemma 3.11. *Suppose that operators B and C are closable on a dense set $\mathcal{D}$, and that $C \prec B$ on $\mathcal{D}$. Then $B + C$ is also closable on $\mathcal{D}$ and $B + C \sim B$ on $\mathcal{D}$.*

Proof. The condition $C \prec B$ on $\mathcal{D}$ implies for any $u \in \mathcal{D}$,

$$\|Cu\|^2 \leq \varepsilon \|Bu\|^2 + \beta_\varepsilon \|u\|^2 . \tag{3.61}$$

Observe now that for any two vectors $v, w \in \mathcal{H}$ the following elementary inequalities hold:

$$\frac{3}{4}\|v\|^2 - 4\|w\|^2 \leq \|v\|^2 - 2\|v\|\|w\| \leq \|v + w\|^2 \leq 2\|v\|^2 + 2\|w\|^2 . \tag{3.62}$$

Combing (3.61) with (3.62) we get

$$\left(\frac{3}{4}-4\varepsilon\right)\|Bu\|^2-4\beta_\varepsilon\|u\|^2\leq\|Bu+Cu\|^2\leq(2+\varepsilon)\|Bu\|^2+\beta_\varepsilon\|u\|^2, \quad (3.63)$$

which together with Lemma 3.8 (i), (ii) imply that $B+C$ is closable on $\mathcal{D}$ and $B+C\sim B$ completing the proof. □

Recall now the following interpolation inequalities relating the L_2-norms of the derivatives of different orders (see [15], Chapter IV, Sect. 7, Corollary 4 or [19], Sect. 7, Theorem 7.27): for every $\varepsilon>0$ there exists $C_s(\varepsilon,d)$ such that for $0\leq s'<s$,

$$||u||_{H^{s}}\leq\varepsilon||u||_{H^s}+C_s(\varepsilon,d)||u||_{H^0},\ u\in H^s\left(\mathbb{R}^d\right). \quad (3.64)$$

In our case of the fields and functions over the entire space $\mathbb{R}^d$ the inequality (3.64) can be readily verified using (3.12), (3.13) together with the following elementary inequality,

$$|\mathbf{k}|^{2s'}\leq\varepsilon^2|\mathbf{k}|^{2s}+C_s^2(\varepsilon),\ \ 0\leq s'<s, \quad (3.65)$$

which holds for an appropriately chosen constant $C_s(\varepsilon)$. Using in a similar way the inequality (3.65) together with the standard spectral decomposition in place of the Fourier transform, one shows the validity of the following natural generalization of (3.64).

Lemma 3.12. *Let B be a self-adjoint operator and $1\leq s'<s$. Then $\mathcal{D}\left(B^{s'}\right)\supseteq\mathcal{D}(B^s)$ and for every positive $\varepsilon<1$ there exist a positive $\beta_{s,\varepsilon}$ such that*

$$\|u\|^2_{\mathcal{H}_{B^{s'}}}\leq\varepsilon\|u\|^2_{\mathcal{H}_{B^s}}+\beta_{s,\varepsilon}\|u\|^2_{\mathcal{H}},\ u\in\mathcal{D}\left(B^s\right)\text{ and }B^{s'}\prec B^s. \quad (3.66)$$

In addition to that, the restriction $B^{s'}\big|_{\mathcal{D}(B^s)}$ is closable on $\mathcal{D}(B^s)$, its closure is exactly $B^{s'}$, and

$$B^{s'}\prec B^{s''}\text{ on }\mathcal{D}\left(B^s\right),\ 1\leq s'\leq s''<s. \quad (3.67)$$

Theorem 3.13. *Let $s\geq1$ be an integer, B be a self-adjoint operator and $\mathcal{D}=\mathcal{D}(B^s)$. Suppose that A is a bounded operator such that $A\mathcal{D}\subset\mathcal{D}$, and that $B^{s'}A\sim B^{s'}$ on $\mathcal{D}$ for every $1\leq s'\leq s$. Suppose also that $1\leq m\leq s$ and $s_1,s_2,\ldots,s_m\geq1$ are integers such that $s_1+\cdots+s_m=s'\leq s$. Then*

$$\mathfrak{B}=\left(B^{s_1}A\right)\left(B^{s_2}A\right)\cdots\left(B^{s_m}A\right)\sim B^{s'}\text{ on }\mathcal{D}, \quad (3.68)$$

including the fact that $\mathcal{D}(\mathfrak{B})\supset\mathcal{D}$. In particular,

$$(BA)^{s'}\sim B^{s'}\text{ on }\mathcal{D}\text{ for every }1\leq s'\leq s. \quad (3.69)$$

Proof. We will refer to the numbers m and s' in the representation (3.68) for the operator $\mathfrak{B}$ respectively as its A-rank m and its power s'.

Let us look first at the domains of our operators. Since B is self-adjoint, we have

$$B^{s'}\mathcal{D}=B^{s'}\mathcal{D}\left(B^s\right)=\mathcal{D}\left(B^{s-s'}\right)\text{ for any }1\leq s'\leq s. \quad (3.70)$$

Based on the given conditions $B^{s'}A \sim B^{s'}$ on $\mathcal{D}$, $A\mathcal{D} \subset \mathcal{D}$ and with the help of Lemma 3.8 (ii) we may conclude that

$$B^{s'}A\mathcal{D} \subset B^{s'}\mathcal{D} = \mathcal{D}\left(B^{s-s'}\right) \text{ for every } 1 \leq s' \leq s. \tag{3.71}$$

The relation (3.71) readily implies that for every operator $\mathfrak{B}$ of the form (3.68) it is well defined on $\mathcal{D}$, i.e.

$$\mathcal{D}(\mathfrak{B}) = \mathcal{D}\left(B^{s-s'}\right) \supset \mathcal{D}\left(B^{s}\right) = \mathcal{D}. \tag{3.72}$$

We prove the main statement by the induction with respect to the A-rank m. Observe first, that the conditions of the theorem evidently imply the validity of (3.68) for $m = 1$. Suppose now that (3.68) holds for $1 \leq m < s' \leq s$, i.e.

$$\mathfrak{B}_1 = \left(B^{s_1}A\right)\left(B^{s_2}A\right)\cdots\left(B^{s_m}A\right) \sim B^{s'} \text{ on } \mathcal{D}, \text{ where } s' = s_1 + \ldots + s_m \leq s, \tag{3.73}$$

and let us show that it is true then for $m + 1$, i.e.

$$\mathfrak{B} = \left(B^{s_1}A\right)\mathfrak{B}_1 = \left(B^{s_1}A\right)\left(B^{s_2}A\right)\cdots\left(B^{s_{m+1}}A\right) \sim B^{s''} \text{ on } \mathcal{D}, \tag{3.74}$$
$$\text{where } s'' = s' + s_{m+1} = s_1 + \ldots + s_{m+1} \leq s.$$

Using Lemma 3.8 (i), (3.72) and the validity of (3.68) for $m = 1$ we obtain for any $u \in \mathcal{D}$,

$$\begin{aligned}\alpha_- \left\|\mathfrak{B}_1'u\right\| + \beta_- \left\|\mathfrak{B}_1u\right\| &\leq \left\|\mathfrak{B}u\right\| = \left\|B^{s_1}A\mathfrak{B}_1u\right\| \\ &\leq \alpha_+ \left\|\mathfrak{B}_1'u\right\| + \beta_+ \left\|\mathfrak{B}_1u\right\|, \quad \text{where } \mathfrak{B}_1' = B^{s_1}\mathfrak{B}_1,\end{aligned} \tag{3.75}$$

and the constants $\alpha_\pm$ and $\beta_\pm$ are, respectively, positive and real, depending only on B and s. Notice now that $\mathfrak{B}_1'$ has the same A-rank m as the operator $\mathfrak{B}_1$, and the power $s'' = s' + s_{m+1}$. Hence, in view of the induction hypothesis, the relation (3.73) applies for the both $\mathfrak{B}_1$ and $\mathfrak{B}_1'$. Using this fact, and once more Lemma 3.8 (i) and (3.75), we get for any $u \in \mathcal{D}$,

$$\begin{aligned}\beta_-' \|u\| + \alpha_-' \left\|B^{s'}u\right\| + \alpha_-'' \left\|B^{s''}u\right\| &\leq \|\mathfrak{B}u\| \\ &\leq \alpha_+'' \left\|B^{s''}u\right\| + \alpha_+' \left\|B^{s'}u\right\| + \beta_+' \|u\|,\end{aligned} \tag{3.76}$$

where the constants $\alpha_\pm''$ are positive and $\alpha_\pm'$, $\beta_\pm'$ are real. From (3.76) and Lemma 3.12 we get for any $u \in \mathcal{D}$,

$$\beta_-''' \|u\| + \alpha_-''' \left\|B^{s''}u\right\| \leq \|\mathfrak{B}u\| \leq \alpha_+''' \left\|B^{s''}u\right\| + \beta_+''' \|u\| \tag{3.77}$$

for some positive $\alpha_\pm'''$ and real β_-'''. Based on (3.77), Lemma 3.8 (i), (ii) we get the desired relation (3.74) that completes the proof of the theorem. □

Now we are ready to prove Lemma 3.4.

Proof (Proof of Lemma 3.4). The statement of the lemma follows from (3.46), (3.47), Lemma 3.3 and Theorem 3.13 where we set $\mathcal{H} = \mathring{\mathbf{L}}_2^2$, $B = i\mathring{\nabla}^{\times\times}$ and $A = \mathring{\Xi}$. □

Remark 3.14. Another way to prove Lemma 3.4 is by constructing a parametrix of the Stokes-type operator $\begin{bmatrix} \nabla^{\times}\eta & \nabla \\ \nabla & 0 \end{bmatrix}$ using the ellipticity of the operator and methods of the theory of pseudodifferential operators, for the exposition of the theory see, for example, [40, 34]. The proof of Lemma 3.4 we gave above is more elementary.

4. Analytic Operators and Series Expansions

4.1. Multilinear forms and polynomial operators. The nonlinear analysis of Maxwell equations requires the use of appropriate Banach spaces of time dependent fields, as well as multilinear and analytic functions in those spaces. It also requires an appropriate version of the implicit function theorem. For the reader's convenience we collect in this section the known concepts and statements on the above-mentioned subjects needed for our analysis.

Definition 4.1. *Suppose that $x_1, x_2, \ldots, x_n$ are vectors in a Banach space X. Let a function $F(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, be defined for all values of the variables $\boldsymbol{x} \in X^n$ and take values in Z. This function is called a n-linear form if it is linear in each variable separately. It is said to be bounded if the norm of F defined by*

$$\|F\|_{X,Z} = \sup_{\|x_1\|_X = \ldots = \|x_n\|_X = 1} \|F(x_1, x_2, \ldots, x_n)\|_Z \tag{4.1}$$

is finite. When the choice of the spaces X and Z is clear from the context, we simply write $\|F\|$.

Definition 4.2. *A function $P(x)$ from X to Z defined for all $x \in X$ is called a polynomial in x of degree n if for all $a, h \in X$ and all complex α,*

$$P(a + \alpha h) = \sum_{\nu=0}^{n} P_\nu(a, h)\, \alpha^\nu, \tag{4.2}$$

where $P_\nu(a, h) \in Z$ are independent of α. The degree is exactly n if $P_n(a, h)$ is not identically zero. $P(x)$ is a homogeneous polynomial of degree n if it is homogeneous of degree n,

$$f(\alpha x) = \alpha^n f(x) \tag{4.3}$$

and is a polynomial. A homogeneous polynomial is called bounded if its norm

$$\|f\|_* = \sup_{\|x\|_X = 1} \left\{\|f(x)\|_Z\right\} \tag{4.4}$$

is finite. For a given n-linear form $F_n(\boldsymbol{x}) = F_n(x_1, x_2, \ldots, x_n)$ we denote by $F_n(x^n)$ a homogeneous of degree n polynomial from X to Z,

$$F_n\left(x^n\right) = F_n(x, x, \ldots, x). \tag{4.5}$$

Usually we denote the multilinear operator $f_n(\boldsymbol{x})$ and the homogeneous polynomial $f_n(x^n)$ obtained by the restriction to the diagonal $x_1 = x_2 = \ldots = x_n$ by the same letter f_n. Obviously,

$$\|f_n\|_* \leq \|f_n\|. \tag{4.6}$$

Definition 4.3. *Let $f_m(x^m)$, $m = 2, 3, \ldots$ be a sequence of bounded m homogeneous polynomials from X to Z that satisfy the estimate*

$$\|f_m\|_* \leq C_{*f} R_{*f}^{-m}, m = 2, 3, \ldots. \tag{4.7}$$

We say that a function f that is defined for $\|x\|_X < R_{*f}$ by the series

$$f(x) = \sum_{n=2}^{\infty} f_n(x^n) \tag{4.8}$$

is in the analyticity class $A_*\left(C_{*f}, R_{*f}, X, Z\right)$ and write $f \in A_*\left(C_{*f}, R_{*f}, X, Z\right)$. We say that g is analytic in X if $g = L + f$, where L is a bounded linear operator in X and $f \in A_*\left(C_f, R_f, X, X\right)$ for some $C_f, R_f > 0$.

If $f \in A_*\left(C_{*f}, R_{*f}, X, Z\right)$ and $\|x\|_X < R_{*f}$ the series

$$\sum_{n=n_0}^{\infty} \|f_n(x^n)\|_Z \tag{4.9}$$

obviously converges, and, consequently, the series (4.8) converges in the Banach space Z. In addition to that we have the inequality

$$\|f(x)\|_Z \le C_{*f} \sum_{n=n_0}^{\infty} \|x\|_X^n R_{*f}^{-n} \le C_{*f} \frac{\|x\|_X^{n_0} R_{*f}^{-n_0}}{1 - \|x\|_X R_{*f}^{-1}}. \tag{4.10}$$

Definition 4.4. *If $f_m(\boldsymbol{x})$, $m = 2, 3, \ldots$, is a sequence of bounded m-linear operators from X^m to Z and*

$$\|f_m\| \le C_f R_f^{-m}, m = 2, 3...., \tag{4.11}$$

we say that a function f defined by the series (4.8) for $\|x\|_X < R_f$ belongs to the analyticity class $A\left(C_f, R_f, X, Z\right)$ and write $f \in A\left(C_f, R_f, X, Z\right)$.

When it does not lead to confusion, we write $A_*\left(C_{*f}, R_{*f}\right)$, $A\left(C_f, R_f\right)$ instead of $A_*\left(C_{*f}, R_{*f}, X, Z\right)$, $A\left(C_f, R_f, X, Z\right)$.

Note that obviously $A\left(C_f, R_f, X, Z\right) \subset A_*\left(C_f, R_f, X, Z\right)$. We often need to find a multilinear operator generating a given polynomial operator. Since different multilinear operators $G_m(x_1, \ldots, x_n)$ may result in the same polynomial operator $G_m(x, \ldots, x) = G_m(x^m)$, we assume the multilinear operator to be symmetric. It is called the polar form $\tilde{G}_m(x_1, x_2, \ldots, x_n)$ of $G_m(x^m)$. The existence of the polar form and an estimate of its norm is given by the following proposition (see [26], Sect. 26.2, [14], Sect. 1.1, 1.3 for details).

Proposition 4.5. *For any homogeneous polynomial $P_n(x)$ of degree n there is a unique symmetric n-linear form $\tilde{P}_n(x_1, x_2, \ldots, x_n)$, called the* **polar form** *of $P_n(x)$, such that $P_n(x) = \tilde{P}_n(x, x, \ldots, x)$. It is defined by the following* **polarization formula***:*

$$\tilde{P}_n(x_1, x_2, \ldots, x_n) = \frac{1}{2^n n!} \sum_{\xi_j = \pm 1} P_n\left(\sum_{j=1}^{n} \xi_j x_j\right). \tag{4.12}$$

In addition to that, the following estimate holds:

$$\|P_n\|_* \le \left\|\tilde{P}_n\right\| \le \frac{n^n}{n!} \|P_n\|_*. \tag{4.13}$$

Notice that using an expansion for the logarithm of the Gamma-function, [43], Sect. 12.33, we get for an integer $n \geq 1$,

$$\frac{n^n}{n!} \leq \frac{n^n}{(n+1)^{n+1/2}\, e^{-n-1}\sqrt{2\pi}} \leq \frac{n^n}{\sqrt{n+1}n^n e^{-n-1}\sqrt{2\pi}} = \frac{e}{\sqrt{2\pi}\,(n+1)} e^n \quad (4.14)$$

which, in turn, implies the inequality

$$\left\| \tilde{P}_n \right\| \leq e^n \left\| P_n \right\|_* . \quad (4.15)$$

Corollary 4.6. *If $f \in A_* \left(C_{*f}, R_{*f}, X, X\right)$ then the polynomial operator series (4.8) can be written in the form*

$$f(x) = \sum_{m=n_0}^{\infty} \tilde{f}_m \left(x^m\right), \quad (4.16)$$

where $\tilde{f}_m$ are m-linear symmetric operators and $f \in A\left(C_{*f}, R_{*f}/e, X, X\right)$.

Let us consider now analytic functions of two vector variables $(x, z) \in X \times X$. We define the norm on $X \times X$ as follows: $\|(x, z)\|_{X\times X} = \|x\|_X + \|z\|_X$. Let $F_m\left(\overrightarrow{(x, z)}\right)$, $m = 1, 2...$, be a sequence of bounded m-linear operators from $(X \times X)^m$ to X, and assume that $F(x, z)$ defined by (4.8) belongs to $A(C_F, R_F, X \times X, X)$. The formula (4.8) for an analytic operator of two vector variables takes the form

$$F(x, z) = \sum_{m=2}^{\infty} F_m \left[(x, z)^m\right], \ x \in X, \ z \in X. \quad (4.17)$$

The series converges if $\|(x, z)\|_{X\times X} < R_F$. When $\|(x, z)\|_{X\times X} \leq r < R_F$ the convergence is uniform. Using the multi-linearity of F_m and splitting $(x_i, z_i) = (x_i, 0)+(0, z_i)$, the m- linear operator on $(X \times X)^m$ can be rewritten in the form

$$\begin{aligned} F_m\left(\overrightarrow{(x, z)}\right) &= F_m\left((x_1, z_1), \ldots, (x_n, z_n)\right) \\ &= \sum_{\delta_1,\ldots\delta_m} F\left((\delta_1 x_1, (1-\delta_1) z_1), \ldots, (\delta_m x_m, (1-\delta_m) z_m)\right), \ \delta_i \in \{0, 1\}. \end{aligned} \quad (4.18)$$

This sum contains 2^m terms. Collecting the terms of the homogeneity s in x (and $m - s$ in z) we write

$$F_m\left(\overrightarrow{(x, z)}\right) = \sum_{s=0}^{m} F_{ms}(\boldsymbol{x}; \boldsymbol{z}), \quad (4.19)$$

where F_{ms} is s-linear in x and $m - s$-linear in z, and, in particular,

$$F_{ms}(t\boldsymbol{x}; r\boldsymbol{z}) = t^s r^{m-s} F_{ms}(\boldsymbol{x}; \boldsymbol{z}). \quad (4.20)$$

Note that

$$F_m\left[(x, z)^m\right] = \sum_{s=0}^{m} F_{ms}\left(x^s; z^{m-s}\right). \quad (4.21)$$

We renumerate the variables for every term in (4.18) so that $F_{ms}(\boldsymbol{x};\boldsymbol{z})$ depends only on $x_1, ..., x_s$ and $z_1, ..., z_{m-s}$. Clearly, it does not change (4.21) and the norm of every term in (4.18). Since the operator F_{ms} involves $\binom{s}{m}$ terms from (4.18) (here $\binom{s}{m} = \frac{m!}{s!(m-s)!}$ is the binomial coefficient), its norm can be estimated as follows:

$$\|F_{ms}(\boldsymbol{x};\boldsymbol{z})\|_X \le \binom{s}{m} \|F_m\| \prod_{i=1}^{s} \|x_i\|_X \prod_{i=1}^{m-s} \|z_i\|_X . \tag{4.22}$$

Hence, we can recast (4.17) as

$$F(x,z) = \sum_{m=1}^{\infty} \sum_{s=0}^{m} F_{ms}\left(x^s; z^{m-s}\right). \tag{4.23}$$

Lemma 4.7. *Let $F \in A(C_F, R_F, X \times X, X)$. Then the series (4.23) converges when $\|x\|_X + \|z\|_X < R_F$ and the sums (4.23) and (4.17) coincide.*

Proof. By (4.22) and (4.11) we get

$$\begin{aligned}\sum_{s=0}^{m} \left\|F_{ms}\left(x^s; z^{m-s}\right)\right\|_X &\le C_F \sum_{s=0}^{m} \frac{m!}{s!\,(m-s)!} R_F^{-m} \|x\|_X^s \|z\|_X^{m-s} \\ &= C_F R_F^{-m} \left(\|x\|_X + \|z\|_X\right)^m\end{aligned}$$

and, hence, the series converges. Since (4.21) holds for every m, the sums (4.23) and (4.17) coincide. □

4.2. Implicit Function Theorem. Let us consider the equation

$$z = Lx + F(x,z), \tag{4.24}$$

where L is a bounded linear operator and $F(x,z)$ is a nonlinear operator such that $F(0,0) = 0$. We single out the linear part since the norm which estimates the linear term is somewhat different. Sometimes though it is convenient to include Lx into F replacing it with a single term $F_1 x$. We seek the solution $z(x)$ to Eq. (4.24) for small $\|x\|$. The following implicit function theorem holds.

Theorem 4.8. *Let $F \in A(C_F, R_F, X \times X, X)$. Then there exists a solution $z = Lx + G(x)$ of (4.24) with $G \in A_*(C_{*G}, R_{*G}, X, X)$, $G \in A(C_{*G}, R_{*G}/e, X, X)$, where*

$$R_{*G} = \frac{R_F + 2C_F - 2\sqrt{R_F C_F + C_F^2}}{1+\gamma_L}, \quad \gamma_L = \|L\|, \tag{4.25}$$

$$\begin{aligned}C_{*G} &= \frac{R_F}{2(R_F + C_F)} \left(R_F + (1+\gamma_L) R_{*G}\right) - R_{*G} \\ &\le \frac{1}{2}\left(R_F + \gamma_L R_{*G} - R_{*G}\right).\end{aligned} \tag{4.26}$$

In particular,

$$G(x) = \sum_{m=2}^{\infty} G_m\left(x^m\right) \text{ for } \|x\|_X < R_{*G}, \tag{4.27}$$

and its norm satisfies

$$\|G(x)\|_X \leq C_{*G} \frac{\|x\|_X^2 R_{*G}^{-2}}{1 - \|x\|_X R_{*G}^{-1}}. \tag{4.28}$$

The polynomial operators $G_m(x^m)$ satisfy the following recursive relations: $G_1 = L$,

$$G_m\left(x^m\right) = \sum_{j=1}^{m} \sum_{s=0}^{j} \sum_{i_1+\cdots+i_{j-s}=m-s} F_{js}\left(x^s; G_{i_1}\left(x^{i_1}\right), \ldots, G_{i_{j-s}}\left(x^{i_{j-s}}\right)\right), m \geq 2, \tag{4.29}$$

where $F_{11} = L$, $F_{10} = 0$, F_{js} are the same as in (4.23). The operator $G(x)$ is unique in the classes $A_*(C, R, X, X)$, $C > 0$, $R > 0$.

Proof. It is convenient to denote the linear operator L in (4.24) by F_1, i.e. $L = F_1$. Note that according to (4.23) $F_1 = F_{11}$ and $F_{10} = 0$ since Lx does not depend on z. Observe also that the recursive relations (4.29) are obtained by formally collecting terms of the homogeneity m in x from Eq. (4.24), where z and F are given respectively by (4.27) and (4.23). In other words they are equivalent to the formal equality

$$\sum_{m=1}^{\infty} G_m\left(x^m\right) = \sum_{j=1}^{\infty} \sum_{s=0}^{j} F_{js}\left(x^s; \left[\sum_{i=0}^{\infty} G_i\left(x^i\right)\right]^{j-s}\right). \tag{4.30}$$

Let us study now the issue of convergence of the series (4.30). First we notice that $G_1 = L$ since (4.30) $F_{11} = L$ and $F_{10} = 0$. To estimate $\|G_m\|_*$ defined by (4.4) let us estimate $\|G_m(x^m)\|_X$ for $\|x\|_X = 1$. Evidently

$$\|G_1(x)\|_X = \|Lx\|_X \leq \gamma_L, \quad \gamma_L = \|L\|. \tag{4.31}$$

For $m > 1$ using (4.22) we get

$$\begin{aligned}\left\|G_m\left(x^m\right)\right\|_X &\leq \sum_{j=2}^{m} \sum_{s=0}^{j} \sum_{i_1+\cdots+i_{j-s}=m-s} \left\|F_{js}\left(x^s; G_{i_1}\left(x^{i_1}\right), \ldots, G_{i_{j-s}}\left(x^{i_{j-s}}\right)\right)\right\|_X \\ &\leq \|x\|_X^m \sum_{j=2}^{m} \sum_{s=0}^{j} \sum_{i_1+\cdots+i_{j-s}=m-s} C_F R_F^{-j} \binom{s}{j} \prod_{l=1}^{j-s} \left\|G_{i_l}\right\|_*.\end{aligned}$$

Hence, we have the following recursive estimate:

$$\|G_m\|_* \leq \sum_{j=2}^{m} \sum_{s=0}^{j} \sum_{i_1+\cdots+i_{j-s}=m-s} C_F R_F^{-j} \binom{s}{j} \prod_{l=1}^{j-s} \left\|G_{i_l}\right\|_*, \; m = 2, 3, \ldots. \tag{4.32}$$

Let us introduce a sequence of majorants g_m by

$$g_1 = \gamma, \quad g_m = \sum_{j=2}^{m} \sum_{s=0}^{j} \sum_{i_1+\cdots+i_{j-s}=m-s} C_F R_F^{-j} \binom{s}{j} \prod_{l=1}^{j-s} g_{i_l}, \; m \geq 2. \tag{4.33}$$

Obviously,

$$\|G_m\|_* \leq g_m, \ m = 1, 2, \ldots . \tag{4.34}$$

Then we introduce an auxiliary function

$$Z(r) = \sum_{m=1}^{\infty} g_m r^m. \tag{4.35}$$

Note that (4.33) can be obtained by equating m^{th} powers of r from the equation

$$\sum_{m=1}^{\infty} g_m r^m = \gamma r + \sum_{j=2}^{\infty}\sum_{s=0}^{j} C_F R_F^{-j} \binom{s}{j} r^s \left[\sum_{i=1}^{\infty} g_i r^i\right]^{j-s}. \tag{4.36}$$

The right-hand side of (4.36) equals

$$\gamma r + \sum_{j=2}^{\infty} C_F R_F^{-j} \left[r + \sum_{i=1}^{\infty} g_i r^i\right]^j = \gamma r + C_F \left[\frac{((r + Z(r))/R_F)^2}{1 - (r + Z(r))/R_F}\right]. \tag{4.37}$$

Hence, Eq. (4.36) is equivalent to the equation for $Z(r)$ given by (4.35)

$$Z(r) = \gamma r + C_F \left[\frac{((r + Z(r))/R_F)^2}{1 - (r + Z(r))/R_F}\right]. \tag{4.38}$$

The numbers g_m then are the Maclaurin coefficients of the solution of this algebraic equation. The estimates of g_m are provided below in the following Lemma 4.9 where we set $C = C_F$, $R = R_F$, $C_2 = C_{*G}$, $r_0 = R_{*G}$. These estimates and (4.34) imply (4.25) and (4.26). Hence $G \in A_*(C_{*G}, R_{*G}, X, X)$. Using Corollary 4.6 we obtain also that $G \in A(C_{*G}, R_{*G}/e, X, X)$.

The sums in m in the left-hand and right-hand sides of (4.29), (4.30) converge, yielding $G(x) = Lx + F(x, G(x))$ and, hence, $G(x)$ is a solution of (4.24). From (4.27) and (4.10) we obtain that for $\|x\|_X < R_{*G}$ (4.28) holds. The uniqueness of $G(x)$ follows from the fact that if $G \in A_*(C, R, X, X)$ with $C, R > 0$ is a solution of the equation $G(x) = Lx + F(x, G(x))$ then it must satisfy the recursive relations (4.29). □

Lemma 4.9. *The analytic solution $Z(r)$ of the equation*

$$Z(r) = \gamma r + C\left[\frac{(r + Z(r))^2}{R(R - r - Z(r))}\right], \ Z(0) = 0, \tag{4.39}$$

with constants $C, R > 0$ and $\gamma \geq 0$, expands into the Maclaurin series $\sum_n g_n r^n$ with $g_n \geq 0$ and the radius of convergence

$$r_0 = \frac{R + 2C - 2\sqrt{CR + C^2}}{\gamma + 1}. \tag{4.40}$$

The coefficients g_n satisfy the inequalities

$$|g_n| \leq C_2 r_0^{-n}, \ C_2 = \frac{R(R + r_0(1+\gamma))}{2(R + C)} - r_0, n = 1, 2, \ldots . \tag{4.41}$$

Proof. Eq. (4.39) is reducible to a quadratic equation

$$R\,(R-r-Z)\,(Z+r-(1+\gamma)\,r)=C\,(r+Z)^2 \tag{4.42}$$

that is equivalent to

$$\left(z-\gamma' r\right)R\,(R-z)=Cz^2,\quad \gamma'=1+\gamma,\quad z=r+Z. \tag{4.43}$$

The solution $Z\,(r)$ that satisfies (4.39) and $Z\,(0)=0$ corresponds to

$$z\,(r)=\frac{R}{2\,(R+C)}\left(R+\gamma' r-\sqrt{R^2-2\gamma' rR+(\gamma')^2r^2-4\gamma' rC}\right). \tag{4.44}$$

By the recursive relations (4.33) all the Maclaurin coefficients g_n of $Z\,(r)$ are non-negative, $g_n\geq 0$. The same is true for $z\,(r)=r+Z\,(r)$. Notice that the functions $Z\,(r)$ and $z\,(r)$ have the same branching points. The branching points of $z\,(r)$ are given by the discriminant equation

$$R^2-2\gamma' rR+(\gamma')^2r^2-4\gamma' rC=0. \tag{4.45}$$

The branching point r_0 with the minimal modulus is

$$r_0=\frac{R+2C-2\sqrt{CR+C^2}}{\gamma'}=\frac{R^2}{\gamma'\left(R+2C+2\sqrt{CR+C^2}\right)} \tag{4.46}$$

that yields (4.40). In as much as the function $Z\,(r)$ is analytic for $|r|<r_0$ and is bounded for $|r|\leq r$, using the Cauchy formula we obtain

$$|g_m|\leq\max_{|r|=r_0}|Z\,(r)|\,r_0^{-m}. \tag{4.47}$$

Since all $g_n\geq 0$ the maximum of $|Z\,(r)|$ over $|r|=r'<r_0$ is attained at a real positive $r=r'$. Obviously, $z\,(r)$ given by (4.44) for $|r|\leq r_0$ is continuous and we get

$$\max_{|r|=r_0}|Z\,(r)|=Z\,(r_0)=z\,(r_0)-r_0=\frac{R^2}{2\,(R+C)}+\frac{r_0R\gamma'}{2\,(R+C)}-r_0. \tag{4.48}$$

□

According to Theorem 4.8 a solution of (4.24) of the form $z=G\,(x)$, $G\in A_*\,(C,R,X,X)$ is unique, but more general solutions may be not unique. Though, the next lemma shows that solutions z are unique if $\|z\|_X+\|x\|_X$ is small enough.

Lemma 4.10. *Let $F\in A\,(C_F,R_F,X\times X,X)$ and z_1,z_2 be two solutions to Eq. (4.24) with $\|z_1\|_X+\|x\|_X\leq r$, $\|z_2\|_X+\|x\|_X\leq r$, $r<R_F$. If*

$$\frac{C_F}{R_F}\left[\frac{1}{(1-r/R_F)^2}-1\right]<1, \tag{4.49}$$

then $z_2=z_1$.

Proof. Evidently,

$$\|z_1 - z_2\| = \|F(x, z_1) - F(x, z_2)\|. \tag{4.50}$$

Notice then that the following identity holds for an n^{th} order multilinear form F_n:

$$\begin{aligned} F_n\left((x, z_1)^n\right) - F_n\left((x, z_2)^n\right) &= F_n((x, z_1), ..., (x, z_1)) - F_n((x, z_2), .., (x, z_2)) \\ &= F_n((x, z_1), (x, z_1), ..., (x, z_1)) - F_n((x, z_2), (x, z_1), ..., (x, z_1)) + ... \\ &\quad + F_n((x, z_2), ..., (x, z_2), (x, z_1)) - F_n((x, z_2), ..., (x, z_2)) \\ &= F_n((x, z_1 - z_2), (x, z_1), ..., (x, z_1)) + ... \\ &\quad + F_n((x, z_2), ..., (x, z_2), (x, z_1 - z_2)), \end{aligned} \tag{4.51}$$

implying

$$\left\| F_n\left((x, z_1)^n\right) - F_n\left((x, z_2)^n\right)\right\|_X \le n \|F_n\| r^{n-1} \|z_1 - z_2\|_X. \tag{4.52}$$

Summing up with respect to n the terms in the previous inequality we get

$$\begin{aligned} &\|F(x, z_1) - F(x, z_2)\|_X \\ &\le \sum_{n=2}^{\infty} n r^{n-1} C_F R_F^{-n} = \frac{C_F}{R_F}\left[\frac{1}{(1 - r/R_F)^2} - 1\right]\|z_1 - z_2\|_X. \end{aligned} \tag{4.53}$$

Hence $\|z_1 - z_2\|_X \le c\|z_1 - z_2\|_X$ with $c < 1$ implying $\|z_1 - z_2\|_X = 0$. □

Now let us consider the case when the analytic $F(x, z)$ is of order $n_0 > 2$ at $z = 0$, namely

$$F_2(x, z) = ... = F_{n_0 - 1}(x, z) = 0, \ n_0 > 2. \tag{4.54}$$

It is convenient to rescale the variables

$$z = \alpha z', \ x = \alpha x', \ 0 < \alpha \le 1, \tag{4.55}$$

and to consider the following corollary of Theorem 4.8.

Corollary 4.11. *Assume that the conditions of Theorem 4.8 are fulfilled, and, in addition to that, (4.54) holds. Then for all* $\alpha \in [0, 1]$, *the operator* G *belongs to* $A_*\left(C'_{*G'}, \alpha R_{*G'}, X \times X, X\right)$, *where*

$$R_{*G'} = \frac{R_F + 2\alpha^{n_0-1}C_F - 2\sqrt{R_F\alpha^{n_0-1}C_F + \alpha^{2n_0-2}C_F^2}}{1 + \gamma_L}, \tag{4.56}$$

$$C'_{*G'} = \frac{1}{2}\left(R_F + \gamma_L R_{*G'} - R_{*G'}\right). \tag{4.57}$$

Proof. First, we rewrite Eq. (4.24) in the form

$$z' = Lx' + \alpha^{-1}F\left(\alpha x', \alpha z'\right), \tag{4.58}$$

and introduce

$$F'\left(z'\right) = \alpha^{-1}F\left(\alpha z'\right). \tag{4.59}$$

Since

$$\alpha^{-1} F_n\left(\left(\alpha z'\right)^n\right) = \alpha^{n-1} F_n\left(\left(z'\right)^n\right), n \geq n_0, \tag{4.60}$$

we conclude that

$$\text{if } F \in A\left(C_F, R_F, X \times X, X\right) \quad \text{then} \quad F' \in A\left(\alpha^{n_0-1} C_F, R_F, X \times X, X\right). \tag{4.61}$$

Note that after rescaling (4.55) the solution $z = G(x)$ of (4.24) takes the form $z' = \alpha^{-1} G\left(\alpha x'\right) = G'\left(x'\right)$. Since (4.58) has the form of (4.24), formula (4.25) of Theorem 4.8 gives an estimate (4.56) of the radius $R_{*G'}$ of convergence of the power expansion of $G'\left(x'\right)$. Further, $G(x) = \alpha G'\left(\alpha^{-1} x\right)$ and we obtain that $G \in A_*\left(\alpha C_{*G'}, \alpha R_{*G'}, X, X\right)$, where $C_{*G'}$ is defined by (4.26) with C_F replaced by $\alpha^{n_0-1} C_F$ and, consequently, $C_{*G'} \leq C'_{*G'}$ where $C'_{*G'}$ is defined in (4.57). □

Let us consider a slightly more general case than Eq. (4.24), namely, the equation

$$b_1 z = Lx + F(x, z), \tag{4.62}$$

where $b_1$is a linear operator and $F(x, z)$ is as in (4.24). When b_1 has a bounded inverse b_1^{-1} (this is the standard condition of the implicit function theorem), we reduce (4.62) to (4.24). Namely, we rewrite (4.62) in the form

$$z = b_1^{-1} Lx + b_1^{-1} F(x, z). \tag{4.63}$$

This equation is of form (4.24) with a modified right-hand side, the modified multilinear operators F_m^b are

$$F_m^b = b_1^{-1} F_m \tag{4.64}$$

and the constants R_F, C_F and γ_L in (4.25), (4.26) are replaced respectively by

$$R_{bF} = R_F, C_{bF} = \left\| b_1^{-1} \right\| C_F, \gamma_b = \left\| b_1^{-1} \right\| \gamma_L. \tag{4.65}$$

Consider now the composition $G = F\left(S_1 z + S(z)\right)$ of two analytic operators.

It is well-known that the composition is analytic (see [26]). In the following theorem we give an explicit estimate of the radius of convergence of G.

Theorem 4.12. *Let $F \in A\left(C_F, R_F, X, X\right)$, $S \in A_*\left(C_S, R_S, X, X\right)$, S_1 be a linear bounded operator in X and $\|S_1\| \leq C_S R_S^{-1}$. Let*

$$G_m\left(x^m\right) = \sum_{j=1}^{m} \sum_{i_1+\cdots+i_j=m} F_j\left(S_{i_1}\left(x^{i_1}\right), \ldots, S_{i_{j-s}}\left(x^{i_j}\right)\right), \; G(x) = \sum_{m=1}^{\infty} G_m\left(x^m\right). \tag{4.66}$$

Then $G \in A_*\left(C_G, R_G, X, X\right)$, $G \in A\left(C_G, R_G/e, X, X\right)$, where

$$R_G = \frac{R_F R_S}{R_F + C_S}, C_G = \frac{C_S C_F}{R_F + C_S}. \tag{4.67}$$

The operator $G(x)$ coincides with $F\left(S_1 x + S(x)\right)$ for $\|x\|_X < R_G$.

Proof. The formula (4.66) is obtained by collecting terms of the homogeneity m in the identity $G(x) = F(S(x) + S_1 x)$. Let us estimate $G_m(x^m)$ with $\|x\|_X = 1$:

$$\|G_m(x^m)\|_X \leq \sum_{j=2}^{m} \sum_{i_1+\cdots+i_j=m} \left\| F_j\left(S_{i_1}\left(x^{i_1}\right), \ldots, S_{i_{j-s}}\left(x^{i_j}\right)\right)\right\|_X$$
$$\leq C_F \sum_{j=0}^{m} R_F^{-j} C_S^j \sum_{i_1+\cdots+i_j=m} R_S^{-(i_1+\cdots+i_j)}. \tag{4.68}$$

For further estimation we introduce majorants

$$g_m = C_F \sum_{j=0}^{m} R_F^{-j} C_S^j \sum_{i_1+\cdots+i_j=m} R_S^{-(i_1+\cdots+i_j)}. \tag{4.69}$$

One can see that the right-hand side of the equality (4.69) coincides with the coefficient at λ^m of the formal series obtained after substitution of the number series

$$z(\lambda) = C_S \sum_{i=1}^{\infty} R_S^{-i} \lambda^i \tag{4.70}$$

into the series

$$f(z) = C_F \sum_{j=0}^{m} R_F^{-j} z^j. \tag{4.71}$$

Both series determine respectively the analytic functions

$$z(\lambda) = C_S R_S^{-1} \frac{\lambda}{1 - \lambda R_S^{-1}}, \quad f(z) = C_F \left[\frac{1}{1 - z R_F^{-1}}\right]. \tag{4.72}$$

Therefore s_m coincide with the Maclaurin coefficients of $f(z(\lambda))$ which is a rational function of λ :

$$f(z(\lambda)) = C_F R_F \left[\frac{R_S - \lambda}{R_F R_S - (R_F + C_S)\lambda}\right]$$
$$= \frac{C_F R_F}{R_F + C_S} + \frac{C_F C_S}{R_F + C_S} \frac{1}{1 - \lambda (R_F + C_S)/(R_F R_S)}.$$

The series expansion of $f(z(\lambda))$ yields $g_m = C_G R_G^{-m}$, $m \geq 1$, where C_G and R_G are defined by (4.67). Hence

$$\|G_m\|_* \leq g_m \leq C_G R_G^{-m}, \tag{4.73}$$

and $G \in A_*(C_G, R_G, X, X)$. Applying Corollary 4.6 we obtain that G belongs to $A(C_G, R_G/e, X, X)$. □

Remark 4.13. The condition $\|S_1\| \leq C_S R_S^{-1}$ of the previous theorem can be always satisfied by decreasing R_S since $A_*(C_S, R_S, X, X) \subset A_*\left(C_S, R_S', X, X\right)$ when $R_S' \leq R_S$. One can treat a general case $\|S_1\| = \gamma_S$ directly as in the proof of Theorem 4.12 using a more general majorant $z(\lambda)$ and still can get explicit formulas for R_G and C_G, but they are more involved. For example, the expression for the radius of convergence becomes

$$R_G = \frac{2 R_S R_F}{R_F + \gamma_S R_S + \sqrt{(R_F + \gamma_S R_S)^2 + 4 R_F (C_S - \gamma_S R_S)}}. \tag{4.74}$$

4.3. Further properties of analytic operators. The operators (functions) in Banach spaces that we construct in this paper are in the form of convergent series of polynomials. We consider and study power series primarily at zero, that correspond to the state of complete rest for the medium. Such and more general operators are the subject of the theory of analytic operators (functions). Many properties of scalar analytic functions of a complex variable can be extended to such vector-to-vector functions. In particular, they are continuous, complex differentiable. Among other properties, complex analyticity and boundedness imply analyticity, the Cauchy formula is valid, Taylor series converge in a neighborhood of a point of analyticity, etc. The reader can find details on the properties of abstract analytic functions in [26], Sects. 3.10–3.19 and Sects. 26.1–26.7. All these results are applicable to the analytic operators we construct in this paper.

5. Abstract Causal Power Series

This section provides a systematic analysis of an abstract version of power series similar to (1.12), (1.13) representing the nonlinear polarization. We will refer to power series similar to (1.12), (1.13) and the functions they define as respectively *causal* series and *causal* polynomial operators. In the literature (see, for instance, [21], Chapter IV, and [23], Chapters I and II) the equations involving similar operators are called *retarded* or *Volterra.*

Let us introduce the following notations that are used in the definition of causal operators:

$$\mathbb{R}_+^n = \left\{ \overrightarrow{\tau} \in \mathbb{R}^n : \tau_1, \dots, \tau_n \geq 0 \right\}, \tag{5.1}$$

$$\overrightarrow{\tau} = (\tau_1, \dots, \tau_n), \ \ \overrightarrow{1} = (1, \dots, 1), \ \ \overrightarrow{\mathbf{E}} = (\mathbf{E}_1, \dots, \mathbf{E}_n). \tag{5.2}$$

Let Y be a Banach space. We consider trajectories $x = x(t)$, $-\infty < t \leq T$, which are continuous Y-valued functions of t. Let us recall basic definitions of continuity and strong continuity.

A Y-valued function $x(t)$ defined on an open interval $I \subseteq \mathbb{R}$ is called continuous on I if for any $t_0 \in I$ we have $\lim_{t \to t_0} \|x(t) - x(t_0)\| = 0$. Similar definitions of the continuity and the differentiability are assumed for Y-valued functions $x(t_1, \dots, t_n)$ of n real variables.

If $f(\theta)$ is a bounded n-linear form in a Banach space Y that depends on a parameter θ from a domain Θ in $\mathbb{R}^m$, it is called strongly continuous at a point θ_0 if $f(\theta)(\boldsymbol{x}) \to f(\theta_0)(\boldsymbol{x})$ as $\theta \to \theta_0$ for any $\boldsymbol{x} \in Y^n$. The differentiability and the partial derivatives in the strong sense are defined in a similar way. Note that when we say that an operator $f(\theta)$ depends on θ continuously we understand continuous dependence in the operator norm topology (not in the strong sense).

In most applications considered in this paper we have vector functions $x(t)$, $-\infty < t \leq T$ that satisfy the following rest condition:

$$x(t) = 0, \ t \leq 0, \tag{5.3}$$

and, in this case, it is sufficient to consider the restriction $x(t)$ for $t \in [0, T]$. It is still convenient though to keep (5.3) for simplicity in writing integrals such as (1.13) and their abstract counterparts.

For a given Banach space Y and a positive time T we introduce the Banach space $\mathcal{C}_Y^T = C([-\infty, T]; Y)$ of bounded continuous Y-valued functions $x(t)$, $-\infty < t \leq T$, with the norm defined by (2.6) and the Banach subspace $\mathcal{C}_{0,Y}^{T} = C_0\left([-\infty, T]; Y\right) \subset \mathcal{C}_Y^T$ of functions $x(t)$ satisfying the rest condition (5.3). In our problems the Banach space Y is usually either the Sobolev space H^s or the Hilbert space $\mathcal{H}_{\mathfrak{m}}^s$ defined in terms of the Maxwell operator $\mathfrak{m}$. An electromagnetic field at any fixed time is an element of $\mathcal{H}_{\mathfrak{m}}^s$. A trajectory $x(t)$, $-\infty < t \leq T$, then describes the field evolution up to the time T.

Let us introduce strictly causal n-linear operators $\mathfrak{p}_n$ that act on $\boldsymbol{x} \in \left(\mathcal{C}_Y^T\right)^n$, and take values in $\mathcal{C}_Y^T$. They are abstract versions of the nonlinear polarization operators (1.13):

$$\mathfrak{p}_n(\boldsymbol{x})(t) = \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} p_n\left[t - t_1, \ldots, t - t_n; x_1(t_1), \ldots, x_n(t_n)\right] dt_1 \cdots dt_n, \tag{5.4}$$

where $p_n\left[\tau_1, \ldots, \tau_n; z\right]$ are Y-valued bounded n-linear forms of

$$z = (z_1, \ldots, z_n) \in Y^n \tag{5.5}$$

that continuously depend on $(\tau_1, \ldots, \tau_n) = \boldsymbol{\tau} \in \mathbb{R}_+^n$. We refer to p_n as the *density forms*, *density operators*, or just the *densities*. Evidently, a polynomial form $\mathfrak{p}_n$ corresponding to (5.4) is given by

$$\begin{aligned} \mathfrak{p}_n(x)(t) &= \mathfrak{p}_n\left(x^n\right)(t) \\ &= \int_{\mathbb{R}_+^n} p_n\left[\tau_1, \ldots, \tau_n; x(t - \tau_1), \ldots, x(t - \tau_n)\right] d\tau_1 \cdots d\tau_n. \end{aligned} \tag{5.6}$$

Note that the integral operator (1.14) involves integration over faces on the boundary $\partial\mathbb{R}_+^n$ of the generalized n-dimensional octant $\mathbb{R}_+^n$. Compositions of such operators could involve the integration over more general faces and the corresponding presentation of the operators takes a more general form. To describe this more general form we consider all the faces of all dimensions on the boundary $\partial\mathbb{R}_+^n$. Namely, let

$$\operatorname{sign}(\tau) = \begin{cases} 1 & \text{if} \quad \tau > 0 \\ 0 & \text{if} \quad \tau = 0 \end{cases}, \tag{5.7}$$

and let us introduce for $\nu = 0, 1, \ldots, n$ the following ν-dimensional manifold:

$$\partial^{\nu}\mathbb{R}_+^n = \left\{ \tau_1 \geq 0, \ldots, \tau_n \geq 0 : \sum_{j=1}^{n} \operatorname{sign}\left(\tau_j\right) = \nu \right\} \tag{5.8}$$

with the Lebesgue measure $d^{\nu}\tau$ on it. The manifold $\partial^{\nu}\mathbb{R}_+^n$ is a union of *rectilinear faces* f (generalized octants) of dimension ν

$$f(j_1, \ldots, j_\nu) = \left\{\boldsymbol{\tau} \in \mathbb{R}_+^n : \tau_{j_1} > 0, \ldots, \tau_{j_\nu} > 0, \quad \tau_j = 0, j \neq j_1, \ldots, j_\nu\right\}, \tag{5.9}$$

$$f(j_1, \ldots, j_\nu) \subset \mathbb{R}^n(j_1, \ldots, j_\nu) = \left\{\boldsymbol{\tau} \in \mathbb{R}^n : \tau_j = 0, j \neq j_1, \ldots, j_\nu\right\}. \tag{5.10}$$

Notice that

$$\partial^n\mathbb{R}_+^n = \mathbb{R}_+^n - \partial\mathbb{R}_+^n. \tag{5.11}$$

Using the notations (5.1), (5.2), (5.8) we introduce for $0 \leq \nu \leq n$ the following, more general multilinear forms

$$\mathfrak{p}_{n,\nu}(\boldsymbol{x})(t) = \int_{\partial^\nu \mathbb{R}_+^n} p_{n,\nu}\left[\overrightarrow{\tau}; x_1(t-\tau_1), \ldots, x_n(t-\tau_n)\right] d^\nu \tau$$
$$= \sum_{f \subset \partial^\nu \mathbb{R}_+^n} \int_f p_{n,\nu,f}\left[\overrightarrow{\tau}; x_1(t-\tau_1), \ldots, x_n(t-\tau_n)\right] d^\nu \tau, \quad (5.12)$$

where the densities $p_{n,\nu}\left[\overrightarrow{\tau}; \overrightarrow{z}\right]$ are n-linear forms in $\overrightarrow{z} \in Y^n$ continuously depending on $\tau_1, \ldots, \tau_n \in f \subset \mathbb{R}_+^n$, $p_{n,\nu,f}$ is a restriction of $p_{n,\nu}$ to a face $f \subset \partial^\nu \mathbb{R}_+^n$. These operators involve the integration over the faces $\partial^\nu \mathbb{R}_+^n$, an example of such a form occurs in (1.14). The corresponding polynomial form is given by

$$\mathfrak{p}_{n,\nu}(x)(t) = \mathfrak{p}_{n,\nu}(x, \ldots, x)(t) = \int_{\partial^\nu \mathbb{R}_+^n} p_{n,\nu}\left[\overrightarrow{\tau}; x\left(t\overrightarrow{1} - \overrightarrow{\tau}\right)\right] d^\nu \tau. \quad (5.13)$$

We introduce now forms $\mathfrak{p}_n$ involving the integration over faces of all dimensions

$$\mathfrak{p}_n(x_1, \ldots, x_n) = \sum_{\nu=0}^{n} \mathfrak{p}_{n,\nu}(x_1, \ldots, x_n), \quad (5.14)$$

and the corresponding polynomials

$$\mathfrak{p}_n(x)(t) = \mathfrak{p}_n(x, \ldots, x)(t) = \sum_{\nu=0}^{n} \int_{\partial^\nu \mathbb{R}_+^n} p_{n,\nu}\left[\overrightarrow{\tau}; x\left(t\overrightarrow{1} - \overrightarrow{\tau}\right)\right] d^\nu \tau. \quad (5.15)$$

The form (5.15) can be recast as

$$\mathfrak{p}_n(x)(t) = \sum_{\nu=0}^{n} \int_{t\overrightarrow{1} - \partial^\nu \mathbb{R}_+^n} p_{n,\nu}\left[t\overrightarrow{1} - \overrightarrow{\tau}; x\left(\overrightarrow{\tau}\right)\right] d^\nu \tau, \quad (5.16)$$

which is useful when we differentiate it with respect to time.

It is instructive to look at the simplest case of a general quadratic causal polynomial

$$\mathfrak{p}_2(x)(t) = p_{2,0}(x(t), x(t)) + \int_0^t p_{2,1,f_1}[t-\tau_1, t; x(\tau_1), x(t)]\, d\tau_1$$
$$+ \int_0^t p_{2,1,f_2}[t, t-\tau_2; x(t), x(\tau_2)]\, d\tau_2$$
$$+ \int_0^t \int_0^t p_{2,2}[t-\tau_1, t-\tau_2; x(\tau_1), x(\tau_2)]\, d\tau_1 d\tau_2. \quad (5.17)$$

In this case $\partial^\nu \mathbb{R}_+^n$ is a quadrant when $\nu = 2$, and the faces of the boundary consist of the union of two rays f_1, f_2 with $\nu = 1$ and one point (the origin) with $\nu = 0$. Notice that the first three terms in (5.17) depend explicitly on $x(t)$ at the instant t whereas the fourth term does not, it is strictly causal. The significance of that becomes clear when we differentiate $\mathfrak{p}_2$ with respect to time. Indeed, the time derivative of the first two terms involves the time derivative of $x(t)$ whereas the double integral term does not. Therefore, the double integral term (strictly causal) provides a priori time smoothness of $\mathfrak{p}_{2,2}(x)(t)$. Notice that the classical optics representation for the polarization (1.12), (1.13) involves only such integrals! For this reason we single out as a special class the causal series of the form (5.4) which involve only volume (the highest possible dimension) integrations.

Definition 5.1. *We call* **causal** *the forms and polynomials defined by (5.4), (5.6), (5.13), (5.14) and (5.15) with the densities* $p_{n,\nu}\left[\overrightarrow{\tau};\overrightarrow{z}\right]$, $\nu = 0, 1, ...n$, *that (i) are bounded n-linear forms in* $\overrightarrow{z} \in Y^n$, *(ii) continuously depend on* $\overrightarrow{\tau}$ *from every rectilinear face of* $\partial^\nu \mathbb{R}^n_+$. *The series, forms and polynomials that include only volume integrals as in (5.4) will be called* **strictly causal**. *We extend the forms* $p_{n,\nu}\left[\overrightarrow{\tau};\cdot\right]$ *for* $\overrightarrow{\tau} \in \mathbb{R}^n$ *by assigning them zero values outside* $\partial^\nu \mathbb{R}^n_+$, *i.e. we set*

$$p_{n,\nu}\left[\tau_1, \ldots, \tau_n; z_1, \ldots, z_n\right] = 0 \text{ if } (\tau_1, \ldots, \tau_n) \notin \partial^\nu \mathbb{R}^n_+. \tag{5.18}$$

We introduce the norm of a density $p_{n,\nu}$

$$\left\| p_{n,\nu} \right\| = \left\| p_{n,\nu} \right\|_{Y,Y} = \int_{\partial^\nu \mathbb{R}^n_+} \left\| p_{n,\nu}\left[\overrightarrow{\tau};\cdot\right] \right\|_{Y,Y} d^\nu \tau, \tag{5.19}$$

where for every fixed $\overrightarrow{\tau}$ *the norm* $\left\| p_{n,\nu}\left[\overrightarrow{\tau};\cdot\right] \right\|_{Y,Y}$ *is given by (4.1); we assume that* $\left\| p_{n,\nu}\left[\overrightarrow{\tau};\cdot\right] \right\|_{Y,Y}$ *is bounded uniformly in* $\overrightarrow{\tau}$ *and we assume that for causal forms the norm (5.19) is finite:*

$$\left\| p_{n,\nu} \right\| = \left\| p_{n,\nu} \right\|_{Y,Y} = \int_{\partial^\nu \mathbb{R}^n_+} \left\| p_{n,\nu}\left[\overrightarrow{\tau};\cdot\right] \right\| d^\nu \overrightarrow{\tau} < \infty. \tag{5.20}$$

Notice that the continuity together with (5.20) are sufficient for the Bochner integrability of forms with respect to $\overrightarrow{\tau}$ (see [26], Sects. 3.1–3.93, or [45], Chapter IV).

Note that when (5.18) is fulfilled, the integration in (5.12) over a face f of $\partial^\nu \mathbb{R}^n_+$ can be replaced by the integration over a subspace $\mathbb{R}^\nu_f$ that contains f and integrals in (5.12) after a renumeration of the variables τ take the form

$$\begin{aligned} &\mathfrak{p}_{n,\nu,f}\left(x_1, \ldots, x_n\right)(t) \\ &\quad = \int_{\mathbb{R}^\nu_f} p_{n,\nu,f}\Big[\tau_1, ..., \tau_\nu, 0, ..., 0; x_1\left(t-\tau_1\right), \ldots, x_\nu\left(t-\tau_\nu\right), x_{\nu+1}(t), \\ &\qquad\qquad \ldots, x_n(t)\Big] d\tau^\nu \end{aligned} \tag{5.21}$$

$$\quad = \int_{\mathbb{R}^\nu_f} p_{n,\nu,f}\left[t\overrightarrow{1}_f - \overrightarrow{\tau}; I_f \boldsymbol{x}\left(\overrightarrow{\tau}\right) + \left(I - I_f\right)\boldsymbol{x}(t)\right] d\tau^\nu, \tag{5.22}$$

where I_f denotes the projection in $\mathbb{R}^n$ onto the subspace that spans f.

Lemma 5.2. *Let (5.20) hold. Then the operator* $p_{n,\nu}$ *is bounded from* $\left(\mathcal{C}^T_Y\right)^n$ *into* $\mathcal{C}^T_Y$ *for every* $T \geq 0$ *and the norm of this operator defined by (4.1) admits the estimate*

$$\left\| \mathfrak{p}_{n,\nu} \right\|_{\mathcal{C}^T_Y, \mathcal{C}^T_Y} \leq \left\| p_{n,\nu} \right\|_{Y,Y}. \tag{5.23}$$

The polynomial operator $\mathfrak{p}_{n,\nu}\left(x^n\right)$ *leaves* $\mathcal{C}^T_{0,Y}$ *invariant, i.e.* $\mathfrak{p}_{n,\nu}\left(\mathcal{C}^T_{0,Y}\right) \subset \mathcal{C}^T_{0,Y}$.

Proof. Note that $\partial^\nu \mathbb{R}^n_+$ consists of $\binom{\nu}{n}$ different faces f for which $n-\nu$ variables τ_j equal zero, and, hence, the integral (5.12) splits into the sum of $\binom{\nu}{n}$ integrals $\mathfrak{p}_{n,\nu,f}$ over the faces. We estimate first the integral of $\mathfrak{p}_{n,\nu}\left(x_1, \ldots, x_n\right)(t)$ over one of the faces f. Without loss of generality we assume that for this face $\tau_{\nu+1} = ... = \tau_n = 0$ and, hence,

$$\mathfrak{p}_{n,\nu,f}\left(x_1,\ldots,x_n\right)(t)=\int_0^\infty\ldots\int_0^\infty p_{n,\nu,f}\Big[\tau_1,\ldots,\tau_\nu,0,\ldots,0;x_1\left(t-\tau_1\right),\\ \ldots,x_n\left(t-\tau_\nu\right)\Big]d\tau_1 d\tau_\nu, \quad (5.24)$$

where $p_{n,\nu,f}$ is the density restricted to the face f. According to (2.6),

$$\left\|x_j\left(t-\tau\right)\right\|_Y\leq\left\|x_j\right\|_{\mathcal{C}_Y^T}\quad\text{for } t\leq T,\tau\geq 0, \quad (5.25)$$

and, hence, we have for $t\leq T$,

$$\begin{aligned}&\left\|\mathfrak{p}_{n,\nu,f}\left(x_1,\ldots,x_n\right)(t)\right\|_Y\\ &\leq\int_0^\infty\ldots\int_0^\infty\left\|p_{n,\nu,f}\left[\tau_1,\ldots,\tau_\nu,0,\ldots,0;x_1\left(t-\tau_1\right),\ldots,x_n\left(t-\tau_\nu\right)\right]\right\|_Y d\tau_1\ldots d\tau_\nu\\ &\leq\int_0^\infty\ldots\int_0^\infty\left\|p_{n,\nu,f}\left(\tau_1,\ldots,\tau_\nu,0,\ldots,0\right)\right\|d\tau_1\ldots d\tau_\nu\prod_{j=1}^n\left\|x_j\right\|_{\mathcal{C}_Y^T}\\ &=\left\|p_{n,\nu,f}\right\|\prod_{j=1}^n\left\|x_j\right\|_{\mathcal{C}_Y^T}. \end{aligned} \quad (5.26)$$

Since

$$\left\|p_{n,\nu}\right\|=\sum_{f\subset\partial^\nu\mathbb{R}_+^n}\left\|p_{n,\nu,f}\right\|, \quad (5.27)$$

then summing up the terms (5.26) over all $f\subset\partial^\nu\mathbb{R}_+^n$ we obtain

$$\left\|\mathfrak{p}_{n,\nu}\left(x_1,\ldots,x_n\right)(t)\right\|_Y\leq\left\|p_{n,\nu}\right\|\prod_{j=1}^n\left\|x_j\right\|_{\mathcal{C}_Y^T}, \quad (5.28)$$

implying (5.23). Since in the integral (5.12) $t-\tau_j\leq t$, then if $x_j(t)=0$ for $t\leq 0$ then $\mathfrak{p}_{n,\nu}\left(x_1,\ldots,x_n\right)(t)=0$ for $t\leq 0$ and the subspace $\mathcal{C}_{0,Y}^T$ is invariant under the action of $\mathfrak{p}_{n,\nu}$. □

Now we define a causal power series which is an abstract version of the nonlinear polarization series (1.12). Namely,

$$\mathfrak{p}(x)=\sum_{n\geq n_0}\mathfrak{p}_n(x),\ \mathfrak{p}_n(x)=\mathfrak{p}_n\left(x^n\right)=\mathfrak{p}_n(x,\ldots,x), \quad (5.29)$$

where $\mathfrak{p}_n$ are given by (5.14).

We call a series (5.29) *strictly causal* if it involves only polynomials of the form (5.4). Strictly causal series and polynomials form subsets of respectively sets of the causal series and causal polynomials, as defined by (5.13), (5.29), and, evidently, are singled out by the condition

$$p_{n,\nu}=0\text{ if }0\leq\nu\leq n-1\text{ for all }n. \quad (5.30)$$

In other words, for strictly causal series and polynomials the only nonzero densities are $p_{n,n}=p_n$. From Lemma 5.2 we readily obtain the following statement.

Lemma 5.3. *Let the densities $p_{n,\nu}$ satisfy*

$$\sum_{\nu=0}^{n} \left\| p_{n,\nu} \right\| \leq C_p \beta_p^{-n}, \; n = 0, \ldots . \tag{5.31}$$

Then $\mathfrak{p}$ defined by (5.29) is an analytic operator, $\mathfrak{p} \in A\left(C_p, \beta_p, \mathcal{C}_Y^T, \mathcal{C}_Y^T\right)$ for every $T > 0$. The subspace $\mathcal{C}_{0,Y}^T$ is invariant and $\mathfrak{p} \in A\left(C_p, \beta_p, \mathcal{C}_{0,Y}^T, \mathcal{C}_{0,Y}^T\right)$.

In the next section we consider operators involving time derivatives. The following statements deal with such operators. First we give an abstract version of (1.14).

Lemma 5.4. *Let for an integer $n \geq 1$ the form $\mathfrak{p}_n$, $n \geq 1$, be strictly causal (see (5.4) and Definition 5.1) with a density $p_n\left(\overrightarrow{\tau}; \cdot\right) = p_{n,n}\left(\overrightarrow{\tau}; \cdot\right)$ being continuous and continuously differentiable in $\overrightarrow{\tau} \in \mathbb{R}_+^n$ up to the boundary $\partial\mathbb{R}_+^n$ and such that*

$$\left\| \dot{p}_{n,n} \right\| < \infty, \tag{5.32}$$

where

$$\dot{p}_{n,n}\left(\overrightarrow{\tau}; \cdot\right) = \sum_{j=1}^{n} \partial_{\tau_j} p_{n,n}\left(\overrightarrow{\tau}; \cdot\right). \tag{5.33}$$

Then the composition $\partial_t \circ \mathfrak{p}_n = \partial_t \mathfrak{p}_n$ of the form $\mathfrak{p}_n$ and the time differentiation operator ∂_t is a causal form with the density $\dot{p}_{n,n}\left(\overrightarrow{\tau}; \cdot\right)$ given by (5.33) and

$$\begin{aligned} \dot{p}_{n,n-1}\left(\overrightarrow{\tau}; \cdot\right) &= p_n\left(\overrightarrow{\tau}; \cdot\right) \text{ for } \overrightarrow{\tau} \in \partial^{n-1}\mathbb{R}_+^n, \\ p_{n,\nu}\left(\overrightarrow{\tau}; \cdot\right) &= 0 \text{ for } 0 < \nu \leq n-2. \end{aligned} \tag{5.34}$$

Proof. The statements of the lemma follow straightforwardly from the representation (5.4) for the strictly causal forms $\mathfrak{p}_n$ and the conditions of the lemma. □

Condition 5.5. *Let $\mathfrak{q}$ be a strictly causal (see Definition 5.1) operator. The densities $q_n\left(\overrightarrow{\tau}; \cdot\right)$, $\overrightarrow{\tau} = (\tau_1, \ldots, \tau_n)$, $n \geq n_0$, are assumed to be continuously differentiable in $(\tau_1, \ldots, \tau_n)$ on $\mathbb{R}_+^n$ up to the boundary $\partial\mathbb{R}_+^n$. We assume that there exist $\beta > 0$, $C_q > 0$ such that*

$$\int_{\mathbb{R}_+^n} \left(\|q_n\| + \left\| \sum_{j=1}^{n} \partial_{\tau_j} q_n \right\| \right) d\overrightarrow{\tau} + \int_{\partial^{n-1}\mathbb{R}_+^n} \|q_n\| \, d\overrightarrow{\tau} < C_q \beta^{-n}, \tag{5.35}$$

with $\|q_n\| = \left\| q_n\left(\overrightarrow{\tau}; \cdot\right) \right\|_{Y,Y}$ for $n = n_0, n_0 + 1, \ldots$.

Lemma 5.6. *Let Condition 5.5 hold. Then $\mathfrak{q} \in A\left(C_\mathfrak{q}, R_\mathfrak{q}, C_{0,Y}^T, C_{0,Y}^T\right)$, $\partial_t \circ \mathfrak{q} \in A\left(C_\mathfrak{q}, R_\mathfrak{q}, C_{0,Y}^T, C_{0,Y}^T\right)$ with $R_\mathfrak{q} = \beta$, $C_\mathfrak{q} = C_q$.*

Proof. The statement directly follows from Lemmas 5.2, 5.3 and 5.4. □

The following lemma shows that a composition of causal operators is again a causal operator.

Lemma 5.7. *Let forms* $\mathfrak{p}_{n,\mu}$ *and* $\mathfrak{g}_{m_1,\nu_1},\dots,\mathfrak{g}_{m_n,\nu_n}$ *be causal. Then the composition of the forms*

$$\mathfrak{s}_{N,\nu}(\cdot)=\mathfrak{p}_{n,\mu}\left(\mathfrak{g}_{m_1,\nu_1}(\cdot),\dots,\mathfrak{g}_{m_n,\nu_n}(\cdot)\right),\ N=m_1+\dots+m_n, \tag{5.36}$$

is an N*-linear causal form as in (5.14) with a density* $s_{N,\nu}$*; if* $\mathfrak{p}_{n,\nu_0}$ *and* $\mathfrak{g}_{m_1,\nu_1},\dots,\mathfrak{g}_{m_n,\nu_n}$ *are strictly causal then* $\mathfrak{s}_{N,\nu}$ *is strictly causal too. In addition to that, if we introduce integers*

$$M_0=0,\ M_j=m_1+\dots+m_j,\ j=1,\dots,n, \tag{5.37}$$

then for $0\leq\nu\leq n$*,*

$$\begin{aligned} s_{N,\nu}\left(\overrightarrow{\tau};\overrightarrow{z}\right)&=\sum_{\nu_1+\dots+\nu_n=\nu}\int_{\partial^\mu\mathbb{R}^n_+}p_{n,\mu}\left(\overrightarrow{\tau}';\check{g}_{\nu_1},\dots,\check{g}_{\nu_n}\right)d\overrightarrow{\tau}',\\ \check{g}_{\nu_j}&=g_{m_j,\nu_j}\left(\tau_{M_{j-1}+1}-\tau'_j,\dots,\tau_{M_j}-\tau'_j;z_{M_{j-1}+1},\dots,z_{M_j}\right)\end{aligned} \tag{5.38}$$

with the convention (5.18) applied to all forms under the integral. The norm of the density $s_{N,\nu}$ *satisfies*

$$\left\|s_{N,\nu}\right\|\leq\|\mathfrak{p}_n\|\left(\sum_{\nu_1+\dots+\nu_n=\nu}\prod_{j=1}^n\left\|g_{m_j,\nu_j}\right\|\right). \tag{5.39}$$

Proof. Since the integral over $\partial^\mu\mathbb{R}^n_+$ equals the sum of the integrals over the faces $f\subset\partial^\mu\mathbb{R}^n_+$, and the formula (5.38) is linear with respect to the densities $p_{n,\mu}$ and $\check{g}_{\nu_j}$, the integral (5.38) expands into a sum of integrals over faces and it is sufficient to consider the case when the density is non-zero only on one face. Namely, $p_n=p_{n,\mu}=p_{n,\mu,f}$ and $g_{m_j,\nu_j}=g_{m_j,\nu_j,f_j}$ are supported respectively on faces f and f_j. To verify the representation (5.38) for $s_{N,\nu}$ we plug the expressions for $\mathfrak{g}_{m_j,\nu_j}$ with the densityg_{m_j,ν_j,f_j} into the integral representation (5.12), (5.21) of $\mathfrak{p}_{n,\mu}$ with the integral over the face f of $\partial^\mu\mathbb{R}^n_+$. We get then for such $\mathfrak{p}_n\left(\mathfrak{g}_{m_1},\dots,\mathfrak{g}_{m_n}\right)(t)$ the following "long" expression

$$\int_{\mathbb{R}^\mu_f}p_{n,\mu,f}\begin{pmatrix}\overrightarrow{\tau};\int_{\mathbb{R}^{\nu_1}_{f_1}}g_{m_1,\nu_1,f_1}\left[\boldsymbol{\tau}_1,\boldsymbol{x}_1\left((t-\tau_1)\overrightarrow{1}-\overrightarrow{\tau}_1\right)\right]d^{\nu_1}\boldsymbol{\tau}_1,\dots,\\ \int_{\mathbb{R}^{\nu_n}_{f_n}}g_{m_n,\nu_n,f_n}\left[\boldsymbol{\tau}_n,\boldsymbol{x}_n\left((t-\tau_n)\overrightarrow{1}-\overrightarrow{\tau}_n\right)\right]d^{\nu_n}\boldsymbol{\tau}_n\end{pmatrix}d^\mu\boldsymbol{\tau}. \tag{5.40}$$

Note that

$$\boldsymbol{x}_j\left((t-\tau_j)\overrightarrow{1}-\overrightarrow{\tau}_j\right)=\boldsymbol{x}_j\left(\left(t\overrightarrow{1}-\overrightarrow{\tau}_j\right)-\tau_j\overrightarrow{1}\right). \tag{5.41}$$

After the integration with respect to τ_j the result depends only on $p_{n,\mu,f}$, g_{m_j,ν_j,f_j}, $t\overrightarrow{1}-\overrightarrow{\tau}_j$ and on $\boldsymbol{x}_j(\cdot)$. To check that the result can be written in the form of a causal integral we use (5.21). For simplicity, we take

$$f=\left\{\tau_{\mu+1}=\dots=\tau_n=0,\tau_1>0,\dots,\tau_\mu>0\right\}, \tag{5.42}$$

and, hence, $I_f x_1 = x_1, ..., I_f x_n = 0$. Then we recast (5.40) changing the order of integration

$$\int_{\mathbb{R}_{f_1}^{\nu_1}} d^{\nu_1}\boldsymbol{\tau}_1 ... \int_{\mathbb{R}_{fn}^{\nu_n}} d^{\nu_n}\boldsymbol{\tau}_n \int_{\mathbb{R}_f^{\mu}} d^{\mu}\boldsymbol{\tau}$$

$$p_{n,\mu,f}\left(\begin{array}{c} t\overrightarrow{1}_f - \overrightarrow{\tau}; g_{m_1,\nu_1,f_1}\left[\tau_1\overrightarrow{1}_{f_1} - \overrightarrow{\tau}_1; I_{f_1}\boldsymbol{x}_1\left(\overrightarrow{\tau}_1\right) + \left(I - I_{f_1}\right)\boldsymbol{x}_1(t)\right], \\ ..., g_{m_n,\nu_n,f_n} \end{array}\right). \tag{5.43}$$

Note that in the process of the integration with respect to $d^{\mu}\boldsymbol{\tau}$ the functions $\boldsymbol{x}_j\left(\overrightarrow{\tau}_j\right)$ are constant, and the result is a multi-linear form with respect to $z_j = I_{f_j}\boldsymbol{x}_j\left(\overrightarrow{\tau}_j\right) + \left(I - I_{f_j}\right)\boldsymbol{x}_j(t)$. It has the following form:

$$\int_{\mathbb{R}_{f_1}^{\nu_1}} d^{\nu_1}\boldsymbol{\tau}_1 ... \int_{\mathbb{R}_{fn}^{\nu_n}} d^{\nu_n}\boldsymbol{\tau}_n$$

$$\int_{\mathbb{R}_f^{\mu}} p_{n,\mu,f}\left(\begin{array}{c} t\overrightarrow{1}_f - \overrightarrow{\tau}; g_{m_1,\nu_1,f_1}\left(\tau_1\overrightarrow{1}_{f_1} - \overrightarrow{\tau}_1; z_1\right), \\ ..., g_{m_n,\nu_n,f_n}\left(\tau_n\overrightarrow{1}_{f_n} - \overrightarrow{\tau}_n, z_n\right) \end{array}\right) d^{\mu}\boldsymbol{\tau}.$$

Since

$$\tau_j\overrightarrow{1}_{f_j} - \overrightarrow{\tau}_j = \boldsymbol{\eta}_j - \left(t - \tau_j\right)\overrightarrow{1}_{f_j}, \quad \boldsymbol{\eta}_j = t\overrightarrow{1}_{f_j} - \overrightarrow{\tau}_j, \tag{5.44}$$

the integral with respect to $\boldsymbol{\tau}$ equals the convolution

$$\begin{aligned} &s_{n,\mu,f,m_1,\nu_1,f_1,...,m_n,\nu_n,f_n}\left(\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_n; z_1, ..., z_n\right) \\ &= \int_{\mathbb{R}_f^{\mu}} p_{n,\mu,f}\Big(\overrightarrow{\tau}; g_{m_1,\nu_1,f_1}\left(\boldsymbol{\eta}_1 - \tau_1\overrightarrow{1}_{f_1}; z_1\right), ..., g_{m_\mu,\nu_\mu,f_\mu}\left(\boldsymbol{\eta}_\mu - \tau_\mu\overrightarrow{1}_{f_\mu}, z_\mu\right), \\ &\qquad ..., g_{m_n,\nu_n,f_n}\left(\boldsymbol{\eta}_n, z_n\right)\Big) d^{\mu}\boldsymbol{\tau}. \end{aligned} \tag{5.45}$$

Thus (5.40) equals

$$\int_{\mathbb{R}_{f_1}^{\nu_1}} d^{\nu_1}\boldsymbol{\tau}_1 ... \int_{\mathbb{R}_{fn}^{\nu_n}} d^{\nu_n}\boldsymbol{\tau}_n$$

$$s_{n,\mu,f,m1,\nu_1,f_1,...}\left(\begin{array}{c} t\overrightarrow{1}_{f_1} - \overrightarrow{\tau}_1, ..., t\overrightarrow{1}_{f_n} - \overrightarrow{\tau}_n; I_{f_1}\boldsymbol{x}_1\left(\overrightarrow{\tau}_1\right) + \left(I - I_{f_1}\right)\boldsymbol{x}_1(t), \\ ..., I_{f_n}\boldsymbol{x}_n\left(\overrightarrow{\tau}_n\right) + \left(I - I_{f_n}\right)\boldsymbol{x}_n(t) \end{array}\right).$$

This integral has the form of the right-hand side (5.22) of (5.21) with $\overrightarrow{\tau}$ replaced by $\overrightarrow{\boldsymbol{\tau}} = \left(\overrightarrow{\tau}_1, ..., \overrightarrow{\tau}_n\right)$ and $\boldsymbol{x}$ replaced by $\overrightarrow{\boldsymbol{x}} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$. Therefore (5.40) coincides with a causal integral, in particular $s_{n,\mu,,...} = 0$ when one of $t\overrightarrow{1}_{f_j} - \overrightarrow{\tau}_j \notin f_j$ since $t\overrightarrow{1}_{f_j} - \overrightarrow{\tau}_j - \tau_j\overrightarrow{1}_{f_j} \notin f_j$ as well when $\tau_j \geq 0$ and $g_{m_j,\nu_j,f_j} = 0$ in this case. Formula (5.38) is obtained by the summation of (5.40) with respect to $f, f_1, ..., f_n$. The boundedness and the continuity properties can also be verified straightforwardly based on (5.45). A direct estimation of the norm of (5.45) yields the inequality

$$\begin{aligned} &\sup_{\boldsymbol{\eta}_1,...,\boldsymbol{\eta}_n} \left\|s_{n,\mu,f,m1,\nu_1,f_1,...}\left(\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_n; z_1, ..., z_n\right)\right\| \\ &\leq \left\|p_{n,\mu,f}\right\| \prod_{j=1}^{n} \sup_{\boldsymbol{\eta}_j} \left\|g_{m_j,\nu_j,f_j}\left(z_j\right)\right\|. \end{aligned} \tag{5.46}$$

For the density norm of $s_{n,\dots}$ defined by (5.19) we get

$$\int \left\| s_{n,\dots}\left(\overrightarrow{\boldsymbol{\eta}}, \overrightarrow{z}\right)\right\| d\overrightarrow{\boldsymbol{\tau}}$$
$$\leq \int\int \left\| p_{n,\mu,f}\left(\overrightarrow{\tau};\cdot\right)\right\| \prod_{j=1}^{n} \left\| g_{m_j,\nu_j,f_j}\left(\boldsymbol{\eta}_j - \tau_j \overrightarrow{1}_{f_j}, z_j\right)\right\| d\overrightarrow{\boldsymbol{\eta}}\, d^{\mu}\boldsymbol{\tau}$$
$$= \int\int \left\| p_{n,\mu,f}\left(t\overrightarrow{1}_f - \overrightarrow{\tau};\cdot\right)\right\| \prod_{j=1}^{n} \left\| g_{m_j,\nu_j,f_j}\left(\tau_j \overrightarrow{1}_{f_j} - \overrightarrow{\tau}_j, z_j\right)\right\| d\overrightarrow{\boldsymbol{\tau}}\, d^{\mu}\boldsymbol{\tau}$$
$$\leq \int \left\| p_{n,\mu,f}\left(t\overrightarrow{1}_f - \overrightarrow{\tau};\cdot\right)\right\| d^{\mu}\boldsymbol{\tau} \prod_{j=1}^{n} \left\| g_{m_j,\nu_j,f_j}\right\| \prod_{i=1}^{m_j} \left\| z_{ji}\right\|.$$

Therefore,

$$\left\| s_{n,\mu,f,m1,\nu_1,f_1,\dots}\right\| \leq \left\| p_{n,\mu,f}\right\| \prod_{j=1}^{n} \left\| g_{m_j,\nu_j,f_j}\right\|. \tag{5.47}$$

To get (5.39) from (5.47) by the summation with respect to $f, f_1, \dots, f_n$ we use (5.27). The continuity of $s_{n,\mu,f,m1,\nu_1,f_1,\dots}$ is proven in the following Lemma 5.8. □

Lemma 5.8. *Let $s\left(\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_n; z_1, \dots, z_n\right)$ be given by (5.45). Then $s\left(\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_n; \cdot\right)$ is continuous up to the boundary with respect to every $\boldsymbol{\eta}_j \in f_j$, $j = 1, \dots, \mu$.*

Proof. Let us rewrite (5.45) in the form

$$s\left(\overrightarrow{\boldsymbol{\eta}}; z_1, \dots, z_n\right) = \int p\left(\overrightarrow{\tau}; g_1\left(\boldsymbol{\eta}_1 - \tau_1 \overrightarrow{1}_{f_1}; z_1\right), \right.$$
$$\left. \dots, g_{\mu}\left(\boldsymbol{\eta}_{\mu} - \tau_{\mu} \overrightarrow{1}_{f_{\mu}}, z_{\mu}\right), \dots, g_n\left(\boldsymbol{\eta}_n, z_n\right)\right) d^{\mu}\boldsymbol{\tau}, \tag{5.48}$$

where

$$g_1 = g_{m_1,\nu_1,f_1}, \dots, g_n = g_{m_n,\nu_n,f_n}, p = p_{n,\mu,f}, \overrightarrow{\boldsymbol{\eta}} = \left(\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_n\right). \tag{5.49}$$

Let us pick any $z_1, \dots, z_n$ such that $\left\| z_{ji}\right\|_Y = 1$. Below we will skip $z_1, \dots, z_n$ in the notation. Consider a sequence $\overrightarrow{\boldsymbol{\eta}}_l \to \overrightarrow{\boldsymbol{\eta}}_0$, $l \to \infty$. Let us also pick a small $\epsilon > 0$ and show that for large N,

$$\left\| s\left(\overrightarrow{\boldsymbol{\eta}}_l\right) - s\left(\overrightarrow{\boldsymbol{\eta}}_0\right)\right\|_Y < \epsilon, \; l \geq N. \tag{5.50}$$

Notice that since

$$\left\| g_{m_j,\nu_j,f_j}\left(\boldsymbol{\eta}_j\right)\right\| \leq C_n, \; j = 1, \dots, n \tag{5.51}$$

and (5.20) holds, we can always find large enough ρ to get for all $\overrightarrow{\boldsymbol{\eta}}$ the following inequality:

$$\int_{\{|\boldsymbol{\tau}|\geq\rho\}} \left\| p\left(\overrightarrow{\tau}; g_1\left(\boldsymbol{\eta}_1 - \tau_1 \overrightarrow{1}_{f_1};\right), \dots, g_{\mu}\left(\boldsymbol{\eta}_{\mu} - \tau_{\mu} \overrightarrow{1}_{f_{\mu}},\right), \dots, g_n\left(\boldsymbol{\eta}_n,\right)\right)\right\|_Y d^{\mu}\boldsymbol{\tau}$$
$$< \epsilon/6.$$

Let the number T_{0j} be defined to satisfy the following relations $\boldsymbol{\eta}_{j0} - \tau_j \overrightarrow{1}_{f_j} \in f_j$ for $0 \leq \tau_j < T_{0j}, \boldsymbol{\eta}_{j0} - \tau_j \overrightarrow{1}_{f_j} \notin f_j$ for $\tau_j > T_{0j}$. Such a number T_{0j} exists and is unique since f_j is convex. For $\delta > 0$ we set

$$\Omega_\delta = \left\{ \boldsymbol{\tau} \in f : |\boldsymbol{\tau}| \leq \rho, \ \left|\tau_j - T_{0j}\right| < \delta, \text{ for one of } j = 1, ..., \mu \right\} \tag{5.52}$$

and we choose δ to be so small that

$$\int_{\Omega_\delta} \left\| p\left(\overrightarrow{\tau}; g_1\left(\boldsymbol{\eta}_1 - \tau_1 \overrightarrow{1}_{f_1}\right), ..., g_\mu\left(\boldsymbol{\eta}_\mu - \tau_\mu \overrightarrow{1}_{f_\mu}\right), ..., g_n\left(\boldsymbol{\eta}_n\right)\right)\right\|_Y d^\mu \boldsymbol{\tau} < \epsilon/6.$$

We choose N_0 to be so large that $\left|\boldsymbol{\eta}_{jl} - \boldsymbol{\eta}_{j0}\right| < \delta/2$ when $l \geq N_0$ for all j. Notice that $g_j\left(\boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j}\right)$ vanishes when $\tau_j > T_{0j} + \delta$ according to (5.18) since $\left|\boldsymbol{\eta}_{jl} - \boldsymbol{\eta}_{j0}\right| < \delta$ and $\boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j} \notin f_j$ in this case. Therefore the integrand in (5.48) with $\overrightarrow{\boldsymbol{\eta}} = \overrightarrow{\boldsymbol{\eta}}_l$ is non-zero only when $0 \leq \tau_j \leq T_{0j} + \delta/2, \ j = 1, ..., \mu$; we denote this domain by D. Consider now

$$\int_{D \cap \{|\boldsymbol{\tau}| \leq \rho\} \setminus \Omega_\delta} \|\bar{p}\|_Y \, d^\mu \boldsymbol{\tau}, \tag{5.53}$$

$$\begin{aligned}\bar{p} = p&\left(\overrightarrow{\tau}; g_1\left(\boldsymbol{\eta}_{1l} - \tau_1 \overrightarrow{1}_{f_1}\right), ..., g_\mu\left(\boldsymbol{\eta}_{\mu l} - \tau_\mu \overrightarrow{1}_{f_\mu}\right), ..., g_n\left(\boldsymbol{\eta}_{nl}\right)\right) \\ &- p\left(\overrightarrow{\tau}; g_1\left(\boldsymbol{\eta}_{10} - \tau_1 \overrightarrow{1}_{f_1}\right), \ldots, g_\mu\left(\boldsymbol{\eta}_{\mu 0} - \tau_\mu \overrightarrow{1}_{f_\mu}\right), ..., g_n\left(\boldsymbol{\eta}_{n0}\right)\right)\end{aligned}$$

with $l \geq N_0$. We would like to show that the domain of this integral is such that $\boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j} \in f_j, \ j = 1, ..., \mu$. Indeed, if $\tau_j \leq T_{0j} + \delta/2, \ \boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j} \notin f_j$ then $\boldsymbol{\eta}_{jl} - \tau_j' \overrightarrow{1}_{f_j} = \xi \in \partial f_j, 0 \leq \tau_j' \leq \tau_j$, and we have all $\xi_i \geq 0, \xi_{i_0} = 0$ for some $i_0 \leq \nu_j$. Since $\left|\boldsymbol{\eta}_{jl} - \boldsymbol{\eta}_{j0}\right| < \delta/2$, the i_0^{th} coordinate η_{j0i_0} of $\boldsymbol{\eta}_{j0} - \tau_j' \overrightarrow{1}_{f_j}$ satisfies $\left|\eta_{j0i_0}\right| < \delta/2$, and, hence, $\boldsymbol{\eta}_{j0} - \tau_j' \overrightarrow{1}_{f_j} - \beta \overrightarrow{1}_{f_j} \notin f_j$ for $\beta \geq \delta/2$, therefore $\tau_j + \delta/2 > T_{0j}$. Therefore $\left|T_{0j} - \tau_j\right| \leq \delta/2$ and $\boldsymbol{\tau} \in \Omega_\delta$. This contradicts the requirement $\boldsymbol{\tau} \in D \cap \{|\boldsymbol{\tau}| \leq \rho\} \setminus \Omega_\delta$.

Since all the arguments $\boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j}$ and $\boldsymbol{\eta}_{j0} - \tau_j \overrightarrow{1}_{f_j}$ in (5.53) are shown to be in f_j, the functions g_j are continuous on the closed bounded set $D \cap \{|\boldsymbol{\tau}| \leq \rho\} \setminus \Omega_\delta$. Hence

$$\left\| g_j\left(\boldsymbol{\eta}_{jl} - \tau_j \overrightarrow{1}_{f_j}\right) - \left(\boldsymbol{\eta}_{j0} - \tau_j \overrightarrow{1}_{f_j}\right)\right\|_Y \leq \epsilon_l, \ j = 1, ..., n, \tag{5.54}$$

where $\epsilon_l \to 0$ as $l \to \infty$. Since $\left\| p\left(\overrightarrow{\tau}\right)\right\|$ is bounded, the integral (5.53) is not greater than $\epsilon/6$ when $l \geq N_1$ for large enough N_1. Splitting $s\left(\overrightarrow{\boldsymbol{\eta}}_l\right) - s\left(\overrightarrow{\boldsymbol{\eta}}_0\right)$ into the sum of integrals over the following three domains $D \cap \{|\boldsymbol{\tau}| \geq \rho\}$, $D \cap \Omega_\delta$ and $D \cap \{|\boldsymbol{\tau}| \leq \rho\} \setminus \Omega_\delta$ and using the above estimates we obtain (5.50). □

6. Abstract Nonlinear Maxwell Equations

A substantial part of the nonlinear analysis of the Maxwell equations can be carried out in an abstract and simpler form as it is shown below. Conditions imposed on quantifies of interest are motivated by the original nonlinear Maxwell equations.

Let $\mathfrak{m}$ be a linear self-adjoint operator in $\mathcal{H}$, and let us consider the equation

$$\partial_t u = -i\mathfrak{m}u - j(t). \tag{6.1}$$

We assume that for negative times everything is at rest, i.e.

$$j(t) = 0,\ u(t) = 0,\ t \leq 0. \tag{6.2}$$

The properties of solutions (6.1) are described in terms of the nested Hilbert spaces

$$\mathcal{H}_{\mathfrak{m}}^s = \left\{u : \mathfrak{m}^s u \in \mathcal{H}\right\}, \quad \mathcal{H}_{\mathfrak{m}}^s = \mathcal{H}_{\mathfrak{m}^s} \subset \mathcal{H},\ s = 1, 2, ..., \tag{6.3}$$

with the norms defined by (3.45) with $B = \mathfrak{m}$. In terms of the spectral projections $P(\lambda)$ associated with the operator $\mathfrak{m}$ we can write

$$\mathfrak{m}^s u = \int_{-\infty}^{\infty} \lambda^s dP(\lambda) u, \quad \|u\|_{\mathcal{H}_{\mathfrak{m}}^s}^2 = \int_{-\infty}^{\infty} \left(\lambda^{2s} + 1\right) d\left(P(\lambda) u, u\right)_{\mathcal{H}}. \tag{6.4}$$

Observe that, in view of the self-adjointness of $\mathfrak{m}$ in $\mathcal{H}$, the operator of linear evolution $e^{i\mathfrak{m}t}$ is unitary in $\mathcal{H}$. Besides, every $\mathcal{H}_{\mathfrak{m}}^s$ is invariant under the action of $e^{i\mathfrak{m}t}$, and $e^{i\mathfrak{m}t}$ is also unitary on $\mathcal{H}_{\mathfrak{m}}^s$ since

$$\left(e^{i\mathfrak{m}t}u, e^{i\mathfrak{m}t}v\right)_{\mathcal{H}_{\mathfrak{m}}^s} = \left(\mathfrak{m}^s e^{i\mathfrak{m}t}u, \mathfrak{m}^s e^{i\mathfrak{m}t}v\right) = \left(\mathfrak{m}^s u, \mathfrak{m}^s v\right) = (u, v)_{\mathcal{H}_{\mathfrak{m}}^s}. \tag{6.5}$$

The solution to Eq. (6.1) under the conditions (6.2) takes the form

$$u_0(t) = -\int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')} j\left(t'\right) dt'. \tag{6.6}$$

The validity of this representation is given by the following lemma.

Lemma 6.1. *Let $\mathcal{H}$ be a separable Hilbert space and $\mathfrak{m}$ be a self-adjoint operator in it. Let $j(\cdot) \in C_{0,\mathcal{H}^s}^T$, $s \geq 1$. Then $e^{i\mathfrak{m}t} j(t)$ is a continuous $\mathcal{H}_{\mathfrak{m}}^s$-valued function of t. Let*

$$u(t) = -\int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')} j\left(t'\right) dt'. \tag{6.7}$$

Then $u(\cdot) \in C_{0,\mathcal{H}_{\mathfrak{m}}^s}^T$, $\partial_t u(\cdot) \in C_{0,\mathcal{H}_{\mathfrak{m}}^{s-1}}^T$,

$$\begin{aligned}\|u(t)\|_{\mathcal{H}_{\mathfrak{m}}^s} &\leq \int_0^t \left\|j\left(t'\right)\right\|_{\mathcal{H}_{\mathfrak{m}}^s} dt',\ \|\partial_t u(t)\|_{\mathcal{H}_{\mathfrak{m}}^{s-1}} \\ &\leq \int_0^t \left\|j\left(t'\right)\right\|_{\mathcal{H}_{\mathfrak{m}}^s} dt' + \|u(t)\|_{\mathcal{H}_{\mathfrak{m}}^s},\end{aligned} \tag{6.8}$$

and $u(\cdot)$ *satisfies Eq. (6.1) in* $C^T_{\mathcal{H}^{s-1}_{\mathfrak{m}}}$. *The solution operator* $(\partial_t + i\mathfrak{m})^{-1} : j(\cdot) \mapsto u(\cdot)$ *given by (6.7) extends by continuity to a bounded operator from* $L_1\left([-\infty, T]; \mathcal{H}^s_{\mathfrak{m}}\right)$ *to* $C^T_{\mathcal{H}^s_{\mathfrak{m}}}$ *with the norm*

$$\left\| (\partial_t + i\mathfrak{m})^{-1} j \right\|_{\mathcal{C}^T_{\mathcal{H}^s_{\mathfrak{m}}}} \leq \|j\|_{L_1([-\infty,T];\mathcal{H}^s_{\mathfrak{m}})} \text{ for any } T \geq 0, \tag{6.9}$$

and $(\partial_t + i\mathfrak{m})^{-1} L_{1,0}\left([-\infty, T]; \mathcal{H}^s_{\mathfrak{m}}\right) \subset C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$. *The operator* $\partial_t (\partial_t + i\mathfrak{m})^{-1}$ *extends to a bounded operator from* $L_1\left([-\infty, T]; \mathcal{H}^s_{\mathfrak{m}}\right)$ *to* $C^T_{\mathcal{H}^{s-1}_{\mathfrak{m}}}$.

Proof. The linear operator $e^{i\mathfrak{m}t}$ continuously depends on t in the strong operator topology and is uniformly bounded (it is unitary in $\mathcal{H}^s_{\mathfrak{m}}$ for every t), $j(t)$ continuously depends on t. Therefore $e^{i\mathfrak{m}t} j(t)$ is a continuous function of t. The function

$$\int_{-\infty}^{t} e^{i\mathfrak{m}t'} j\left(t'\right) dt' \tag{6.10}$$

is a continuously differentiable $\mathcal{H}^s_{\mathfrak{m}}$-valued function of t and

$$\partial_t \int_{-\infty}^{t} e^{i\mathfrak{m}t'} j\left(t'\right) dt' = e^{i\mathfrak{m}t} j(t). \tag{6.11}$$

Let us introduce

$$u_0(t) = -\int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')} j\left(t'\right) dt' = -e^{-i\mathfrak{m}t} \int_{-\infty}^{t} e^{i\mathfrak{m}t'} j\left(t'\right) dt', \tag{6.12}$$

which is a continuous function of t in $\mathcal{H}^s_{\mathfrak{m}}$. The operator $e^{-i\mathfrak{m}t}$ considered as an operator from $\mathcal{H}^s_{\mathfrak{m}}$ to $\mathcal{H}^{s-1}_{\mathfrak{m}}$ is strongly differentiable with respect to t and

$$\partial_t u_0(t) = i\mathfrak{m} e^{-i\mathfrak{m}t} \int_{-\infty}^{t} e^{i\mathfrak{m}t'} j\left(t'\right) dt' - j(t). \tag{6.13}$$

Consequently, (6.1) holds with both parts being in $\mathcal{H}^{s-1}_{\mathfrak{m}}$. Obviously, $u_0(t) = 0$, $t < 0$. The inequalities (6.8) follow straightforwardly from (6.6), (6.13). Note that (6.8) implies

$$\|u_0\|_{\mathcal{C}^T_{\mathcal{H}^s_{\mathfrak{m}}}} \leq \int_{-\infty}^{T} \|j(t)\|_{\mathcal{H}^s_{\mathfrak{m}}} dt \text{ for any } T \geq 0, \tag{6.14}$$

which, in turn, yields the boundedness of the operator $(\partial_t + i\mathfrak{m})^{-1}$ together with the inequality (6.9). Using this inequality we extend $(\partial_t + i\mathfrak{m})^{-1}$ to functions $j(t)$ having the following norm bounded:

$$\|j\|_{L_1([-\infty,T];\mathcal{H}^s_{\mathfrak{m}})} = \int_{-\infty}^{T} \left\| j\left(t'\right) \right\|_{\mathcal{H}^s_{\mathfrak{m}}\mathcal{H}^s} dt' < \infty. \tag{6.15}$$

The inequalities (6.8) imply the boundedness of the operator $\partial_t (\partial_t + i\mathfrak{m})^{-1}$ from $L_1([-\infty, T]; \mathcal{H}^s)$ to $\mathcal{C}^T_{\mathcal{H}^{s-1}_{\mathfrak{m}}}$. □

We use the formula (6.7) to define the action of the operator $(\partial_t + i\mathfrak{m})^{-1}$ on $L_1\left(\left[-\infty, T; \mathcal{H}^s_{\mathfrak{m}}\right]\right)$ yielding a solution to (6.1). In particular, we use (6.12) for $j \in L_{1,0}\left(\left[-\infty, T; \mathcal{H}^s_{\mathfrak{m}}\right]\right)$.

Let us look at a modification of the linear evolution equation (6.1) by inserting there a nonlinearity $\mathfrak{q}$ to get an abstract version of the Maxwell equations (7.36). An abstract version $\mathfrak{q}$ of the operator $\mathfrak{q}$ given by (7.37) acts on trajectories $u(t)$, $-\infty < t < T$, in $\mathcal{H}^s_{\mathfrak{m}}$ rather than just states in $\mathcal{H}^s_{\mathfrak{m}}$. We assume $\mathfrak{q}$ to be a strictly causal analytic function of $u \in \mathcal{C}^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$ represented by a power series as in (5.29) (see Definition 5.1).

Our abstract version of the nonlinear Maxwell equations is

$$\partial_t u = -i\mathfrak{m}\{u + \mathfrak{q}(u)\}(t) - j(t), \quad j(t) = u(t) = 0, \ t \leq 0. \tag{6.16}$$

Now we recast Eq. (6.16) to "eliminate" the action of the unbounded operator $\mathfrak{m}$ onto the nonlinearity. We introduce 4, 5

$$w = u + \mathfrak{q}(u) \tag{6.17}$$

and recast (6.16) as

$$\partial_t w = -i\mathfrak{m}w(t) + \partial_t \mathfrak{q}(u) - j(t) \tag{6.18}$$

(see Lemma 6.2 for a justification). By Lemma 6.1 Eq. (6.18) is equivalent to the following equation:

$$w(t) = u_0(t) + \int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')}\partial_t \mathfrak{q}(u)\left(t'\right) dt'. \tag{6.19}$$

Expressing w in terms of u we get

$$u(t) = u_0(t) - \mathfrak{q}(u) + \int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')}\partial_t \mathfrak{q}(u)\left(t'\right) dt'. \tag{6.20}$$

To write (6.20) in the form of (4.24) we introduce the operator

$$\mathfrak{R}(u)(t) = -\mathfrak{q}(u) + \int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')}\partial_t \mathfrak{q}(u)\left(t'\right) dt', \ t \leq T \tag{6.21}$$

which allows us to rewrite (6.20) as

$$u = u_0 + \mathfrak{R}(u). \tag{6.22}$$

The next lemma shows Eqs. (6.16), (6.20) under natural conditions are equivalent.

Lemma 6.2. *Let $j \in L_{1,0}\left(\left[-\infty, T; \mathcal{H}^s_{\mathfrak{m}}\right]\right)$, operators $u \to \partial_t \mathfrak{q}(u)$ and $u \to \mathfrak{q}(u)$ act from the neighborhood $\Omega_{r_1} = \left\{u : \|u\|_{C^T_{\mathcal{H}^s_{\mathfrak{m}}}} \leq r_1\right\}$ of zero in $C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$ into $C^T_{\mathcal{H}^s_{\mathfrak{m}}}$. Let $u \in \Omega_{r_1} \subset C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$ be a solution to (6.20). Then, first, $\partial_t u \in C^T_{0,\mathcal{H}^{s-1}_{\mathfrak{m}}}$ and, second, $u(t)$ is a solution to (6.16) for $t \leq T$. Conversely, if u belongs to $\Omega_{r_1} \subset C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$, $\partial_t u \in C^T_{0,\mathcal{H}^{s-1}_{\mathfrak{m}}}$ and $u(t)$ is a solution of (6.16) then u is a solution of (6.20).*

Proof. If u is a solution of (6.20) then by Lemma 6.1 $\partial_t u \in C^T_{0,\mathcal{H}^{s-1}_{\mathfrak{m}}}$. For $w = u + \mathfrak{q}(u)$ obviously $w \in \Omega_{r_1+r_2}$ and $\partial_t w \in \mathcal{H}^{s-1}_{\mathfrak{m}}$. By the assumptions $\mathfrak{q}(u), \partial_t \mathfrak{q}(u) \in C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$ and the equality (6.19) holds. By Lemma 6.1 we get (6.18) and since $\partial_t w = \partial_t u + \partial_t \mathfrak{q}(u)$ we get (6.16). Conversely, if $u \in \Omega_{r_1}$ we get (6.18), then by Lemma 6.1 we get the equalities (6.19) and (6.20). □

Now we prove our main statements on the abstract Maxwell equations (6.16).

Theorem 6.3. *Let $\mathcal{H}$ be a separable Hilbert space and $\mathfrak{m}$ be a self-adjoint operator in it. Let $T > 0$, $\mathfrak{q}$ be an analytic function in $C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$, and constants $C_{\mathfrak{q}}$ and $R_{\mathfrak{q}}$ be such that $\mathfrak{q} \in A\left(C_{\mathfrak{q}}, R_{\mathfrak{q}}, C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}, C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}\right)$ and $\partial_t \mathfrak{q} \in A\left(C_{\mathfrak{q}}, R_{\mathfrak{q}}, C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}, C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}\right)$. Let $n_0 \geq 2$, $\mathfrak{q}_n = 0$ for $n \leq n_0 - 1$, $\mathfrak{q}_{n_0} \neq 0$. Let $j \in L_{1,0}\left([-\infty, T]; \mathcal{H}^s_{\mathfrak{m}}\right)$ and*

$$\|j\|_{L_1([-\infty,T];\mathcal{H}^s_{\mathfrak{m}})} = \int_{-\infty}^{T} \|j(t)\|_{\mathcal{H}^s_{\mathfrak{m}}}\, dt \leq \delta_0, \tag{6.23}$$

where δ_0 is small enough for the following condition to hold:

$$1 + T < \frac{1}{\delta_0} \frac{R_{\mathfrak{q}}\left(R_{\mathfrak{q}} - 4\delta_0\right)}{8C_{\mathfrak{q}}}. \tag{6.24}$$

Then there exists a solution $u \in C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$ of the the abstract Maxwell equation (6.20). The solution $u = \mathcal{U}(u_0)$ is a uniquely determined analytic function of u_0 with $\mathcal{U}$ being an analytic operator in $C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}$. In particular, u expands into the convergent series

$$u(t) = \mathcal{U}(u_0)(t) = u_0(t) + \sum_{n \geq n_0} \mathcal{U}_n(u_0)(t), \tag{6.25}$$

$$\|\mathcal{U}_n(u_0)\|_{C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}} \leq C_{*G} R_{*G}^{-n} \|u_0\|^n_{C^T_{0,\mathcal{H}^s_{\mathfrak{m}}}}, \quad n \geq n_0, \tag{6.26}$$

with C_{*G} and R_{*G} as in Theorem 4.8, where $R_F = R_{\mathfrak{q}}$, $C_F = (1+T)\, C_{\mathfrak{q}}$, $\gamma_L = 1$. The operators $\mathcal{U}_n$ can be found from the following recursive relations:

$$\mathcal{U}_n = \sum_{m \geq n_0,\ n_1+\cdots+n_m = n} \mathfrak{R}_m\left(\mathcal{U}_{n_1}, \ldots, \mathcal{U}_{n_m}\right), \ \mathcal{U}_1 \text{ is the identity operator,} \tag{6.27}$$

with $\mathfrak{R}_n$ being given by

$$\mathfrak{R}_n(u)(t) = -\mathfrak{q}_n(u)(t) + \int_{-\infty}^{t} e^{-i\mathfrak{m}(t-t')} [\partial_t \mathfrak{q}_n(u)]\left(t'\right) dt', \quad n \geq n_0 \geq 2. \tag{6.28}$$

The first significant term $\mathcal{U}_{n_0}$ in (6.25), **the first nonlinear response**, *has the following representation:*

$$\begin{aligned} \mathcal{U}_{n_0}(u_0)(t) &= \mathfrak{R}_{n_0}(u_0)(t) \\ &= -\mathfrak{q}_{n_0}(u_0) + \int_0^t e^{-i\mathfrak{m}(t-t')} \left[\partial_{t'} \mathfrak{q}_{n_0}(u_0)\right]\left(t'\right) dt' \\ &= -i \int_0^t e^{-i\mathfrak{m}(t-t')} \mathfrak{m}\mathfrak{q}_{n_0}(u_0)\left(t'\right) dt'. \end{aligned} \tag{6.29}$$

Any solution of (6.20) satisfying $\|u\|_{C^{T_\delta}_{0,\mathcal{H}^s_{\mathrm{m}}}} \leq \delta$ *with a sufficiently small* δ *is unique. If* $n_0 > 2$ *and* $\delta_0 \leq R_{\mathrm{q}}/8$ *the condition (6.24) can be replaced by*

$$1 + T < \frac{R_{\mathrm{q}} \left(8\delta_0/R_{\mathrm{q}}\right)^{1-n_0}}{2C_{\mathrm{q}}}. \tag{6.30}$$

Proof. Equation (6.20) when rewritten as (6.22) is of the form of Eq. (4.24) with $z = u$, $X = C^T_{0,\mathcal{H}^s_{\mathrm{m}}}$, $x = u_0$, L being the identity operator. The n-linear operators $F_n = \mathfrak{R}_n$ in the series decomposition of $F = \mathfrak{R}$ are defined by (6.28). To apply Theorem 4.8 we need to estimate the norms of the n-linear operators F_n. Since $\partial_t \mathrm{q}_n(u) \in C^T_{0,\mathcal{H}^s_{\mathrm{m}}}$, $[\partial_t \mathrm{q}_n(u)](t') = 0$ for $t \leq 0$ and $e^{-i\mathrm{m}(t-t')}$ are unitary in $\mathcal{H}^s_{\mathrm{m}}$, then for $t \geq 0$, $t \leq T$, in view of (5.35) we have

$$\begin{aligned}\left\| \int_{-\infty}^{t} e^{-i\mathrm{m}(t-t')} \partial_t \mathrm{q}_n(u)(t')\,dt' \right\|_{\mathcal{H}^s_{\mathrm{m}}} &\leq \int_0^t \left\| \partial_t \mathrm{q}_n(u)(t') \right\|_{\mathcal{H}^s_{\mathrm{m}}} dt' \\ &\leq t \sup_{t' \leq t} \left\| \partial_t \mathrm{q}_n(u)(t') \right\|_{\mathcal{H}^s_{\mathrm{m}}} \leq T \left\| \partial_t \mathrm{q}_n(u) \right\|_{C^T_{0,\mathcal{H}^s_{\mathrm{m}}}}.\end{aligned}$$

Since $\mathrm{q}, \partial_t \mathrm{q} \in A\left(C_{\mathrm{q}}, R_{\mathrm{q}}, C^T_{0,\mathcal{H}^s_{\mathrm{m}}}, C^T_{0,\mathcal{H}^s_{\mathrm{m}}}\right)$ we obtain from (6.28) that

$$\|\mathfrak{R}_n(u)\|_X \leq (1+T)\, C_{\mathrm{q}} R_{\mathrm{q}}^{-n} \|u\|_X^n, \; n = 2, 3, \ldots \quad . \tag{6.31}$$

Observe that $\|L\| = 1$. The inequality (6.31) implies that $\mathfrak{R}$ belongs to the class $A\left((1+T)\, C_{\mathrm{q}}, R_{\mathrm{q}}, X, X\right)$. We would like to apply now Theorem 4.8 and Corollary 4.11 with $R_F = R_{\mathrm{q}}$, $C_F = (1+T)\, C_{\mathrm{q}}$, $\gamma_L = 1$. By Theorem 4.8 $u = G(u_0)$, $G \in A_*(C_{*G}, R_{*G}, X, X)$, and, hence, setting $G = \mathcal{U}$ we obtain the relations (6.25), (6.26). Note that according to (6.23) and (6.14) $||u_0||_Y \leq \delta_0$. By Corollary 4.11 $G(u_0)$ is defined for

$$\|u_0\|_Y \leq \delta_0 < \alpha R_{*G'}, \tag{6.32}$$

where $R_{*G'}$ is given by (4.56), that is

$$R_{*G'} = \frac{R_F^2}{2\left(R_F + 2C_{F'} + 2\sqrt{R_F C_{F'} + C_{F'}^2}\right)}, \; C_{F'} = \alpha^{n_0-1} C_F. \tag{6.33}$$

Note that $2\sqrt{C_{F'}}\sqrt{R_F + C_{F'}} \leq R_F + C_{F'} + C_{F'}$, and, hence, the condition $\delta_0 < \alpha R_{*G}$ is satisfied for $\delta_0/\alpha < R_F^2/[4(R_F + 2C_{F'})]$. Consequently, a sufficient condition for the solvability of (6.22) takes the form $2C_{F'} < \frac{\alpha R_{\mathrm{q}}^2}{4\delta_0} - R_{\mathrm{q}}$ that is

$$2\alpha^{n_0-1}(1+T)\, C_{\mathrm{q}} < \frac{\alpha R_{\mathrm{q}}^2}{4\delta_0} - R_{\mathrm{q}}. \tag{6.34}$$

If $n_0 = 2$ we set $\alpha = 1$ and obtain the condition (6.24). If $n_0 > 2$ and $\delta_0 \leq R_{\mathrm{q}}/8$ we set $\alpha = 8\delta_0/R_{\mathrm{q}}$ and obtain the condition (6.30). If conditions (6.24) or (6.30) are satisfied Theorem 4.8 implies the existence of the solution $u = G(u_0) = \mathcal{U}(u_0)$ of (6.20) written in the form (6.22). We obtain then (6.25) from (4.27), and the inequality (6.26) follows from the definition of the class $A_*(C_G, R_G, X, X)$, see Definition 4.3. Formula (6.27) follows from (4.29). Note that in (4.29) $F_{js} = F_j$ and $s = 0$ since F does not depend on u_0, and (4.29) takes the form of (6.27). The uniqueness of a small solution follows from Lemma 4.10. □

The next theorem adds some more details on properties of the solution for the case when $\mathfrak{q}$ and $\partial_t \mathfrak{q}$ are causal.

Theorem 6.4. *Assume that Condition 5.5 is satisfied. Then statements of Theorem 6.3 hold. In addition to that, we have* $\mathcal{U} \in A\left(C_{*G}, R_{*G}/e, \mathcal{C}^T_{0,\mathcal{H}^s_{\mathrm{m}}}, \mathcal{C}^T_{0,\mathcal{H}^s_{\mathrm{m}}}\right)$, *where* C_{*G}, R_{*G} *are given by (4.25), (4.26) with*

$$R_F = R_{\mathfrak{q}} = \beta, C_F = (1+T)\, C_{\mathfrak{q}} = (1+T)\, C_q, \tag{6.35}$$

and the constants $C_{\mathfrak{q}}$ *and* β *are from Condition 5.5.*

Proof. By Lemma 5.6 $\mathfrak{q}, \partial_t \mathfrak{q} \in A\left(C_{\mathfrak{q}}, R_{\mathfrak{q}}, C^T_{0,\mathcal{H}^s_{\mathrm{m}}}, C^T_{0,\mathcal{H}^s_{\mathrm{m}}}\right)$ and Theorem 6.3 can be applied. □

7. Analysis of the Original Maxwell Equations

In this section we provide the proofs of Theorem 1.3 and a more detailed Theorem 7.8 assuming that Conditions 1.1 and 1.2 are satisfied. Our analysis of the Maxwell equations (7.29) and its regularized form (7.42) is based on their reduction to the abstract Maxwell equation (6.16) and consequent use of Theorems 6.3 and 6.4. First we show that the constituency relation (1.8) is given by a causal analytic operator considered in the previous section.

7.1. Analyticity of the constituency relation. In this subsection we study operators (1.12), (1.13) that are involved in the nonlinear constituency relation (1.8). To use results of Sects. 4 and 5 we set

$$Y = \mathbf{H}^s, X = \mathcal{C}^T_{0,\mathbf{H}^s} = C_0\left([-\infty, T]; \mathbf{H}^s\right). \tag{7.1}$$

The operators corresponding to (1.13) fit into the abstract framework of the previous section and possess an additional property. The multi-linear forms of the type (1.13), (1.14) define operators

$$\mathfrak{P}_n\left(\mathbf{E}_1, \dots, \mathbf{E}_n\right)(t) = \sum_{\nu=0}^{n} \int_{\partial^\nu \mathbb{R}^n_+} P_{n,\nu}\left[\mathbf{r}, \overrightarrow{\tau}; \mathbf{E}_1(t-\tau_1), \dots, \mathbf{E}_n(t-\tau_n)\right] d^\nu \tau, \tag{7.2}$$

where the densities $P_{n,\nu}\left[\mathbf{r}, \overrightarrow{\tau}; \overrightarrow{\mathbf{e}}\right]$ are n-linear forms in $\overrightarrow{\mathbf{e}} \in \mathbb{C}^n$ that depend on variables: $\overrightarrow{\tau} = (\tau_1, \dots, \tau_n) \in \partial^\nu \mathbb{R}^n_+$ and act on $\overrightarrow{\mathbf{e}} = (\mathbf{e}_1, \dots, \mathbf{e}_n) \in \left(\mathbb{C}^3\right)^n$. Causality implies that $\mathfrak{P}_n\left(\overrightarrow{\mathbf{E}}(\cdot)\right)(\cdot, t)$ depends only on $\mathbf{E}^{(j)}\left(\cdot, t_j\right)$ with $t_j \leq t$. The operators $\mathfrak{P}_n$ defined by (7.2) have an important property: they are *spatially local,* namely

$$\text{the value of } \mathfrak{P}_n\left(\overrightarrow{\mathbf{E}}(\cdot)\right)(\mathbf{r}, t) \text{ depends only on } \mathbf{E}^{(j)}\left(\mathbf{r}, t_j\right) \text{ with the same } \mathbf{r}. \tag{7.3}$$

In this subsection we consequently establish the analyticity and the causality for the operators $\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$, $\partial_t \mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ from (1.8). Further, we prove that (1.8) determines $\mathbf{E}$ as an analytical function of $\mathbf{D}$, namely that $\mathbf{E} = \mathcal{S}(\mathbf{D})$, where $\mathcal{S}(\mathbf{D})$, $\partial_t \mathcal{S}_{\mathrm{NL}}(\mathbf{D})$ are analytic causal operators. According to the following proposition the Sobolev space $H^s\left(\mathbb{R}^d\right)$ of scalar functions with $s > d/2$ is a generalized Banach algebra (see [49], Sect. 21.21 or [42], Sect. 2.8.3)

Proposition 7.1. *For any integer $s > d/2$ there exists a constant γ_0 depending only on s and d such that*

$$\|uv\|_{H^s} \leq \gamma_0 \|u\|_{H^s} \|v\|_{H^s}, \tag{7.4}$$

and, consequently,

$$\|u_1 ... u_n\|_{H^s} \leq \gamma_0^{n-1} \|u_1\|_{H^s} ... \|u_n\|_{H^s}, \; n \geq 2. \tag{7.5}$$

To verify the continuity of multilinear polarization forms such as in (1.13) we use Proposition 7.1 and obtain the following lemma.

Lemma 7.2. *Let $P_n\left(\mathbf{r}; \overrightarrow{\mathbf{e}}\right)$, $\overrightarrow{\mathbf{e}} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \in \mathbb{C}^{mn}$, be n-linear operators (tensors) from $\mathbb{C}^{mn}$ to $\mathbb{C}^m$ with coefficients that depend on $\mathbf{r} \in \mathbb{R}^d$. If $\mathbf{E}_j = \mathbf{E}_j(\mathbf{r})$, $j = 1, ..., n$, are functions from $\mathbf{H}^s$, $s > d/2$, $P_n\left(\cdot; \overrightarrow{\mathbf{E}}(\cdot)\right)$ belongs to $\mathbf{H}^s = \mathbf{H}^s_m$. The mapping $P_n : \overrightarrow{\mathbf{E}}(\mathbf{r}) \mapsto P_n\left(\mathbf{r}; \overrightarrow{\mathbf{E}}(\mathbf{r})\right)$ determines a bounded n-linear operator P_n from $\left(\mathbf{H}^s_m\right)^n$ to $\mathbf{H}^s_m$ and there exist positive constants C and γ depending only on s, m and d such that*

$$\|P_n\|_{\mathbf{H}^s, \mathbf{H}^s} \leq C_s \gamma^{-n} \|P_n\|_{C^s}, \tag{7.6}$$

where the norm $\|P_n\|_{C^s}$ of the tensor $P_n(\mathbf{r})$ is defined by (2.10). In addition, for any i, $1 \leq i \leq n$,

$$\left\| P_n\left(\overrightarrow{\mathbf{E}}\right)\right\|_{\mathbf{H}^0} \leq C'_s \gamma^{-n} \|P_n\|_{C^0} \|\mathbf{E}_i\|_{\mathbf{H}^0} \prod_{j \neq i} \left\|\mathbf{E}_j\right\|_{\mathbf{H}^s}. \tag{7.7}$$

Proof. Since the tensors can be written in coordinates, it is sufficient to consider a scalar case. Notice first that there exists a constant γ_1 depending only on $s \geq 0$ such that

$$\|uv\|_{H^s} \leq \gamma_1 \|u\|_{C^s} \|v\|_{H^s}. \tag{7.8}$$

Combining (7.8) with (7.5) we get (7.6). To get (7.7) we note that by the Sobolev embedding theorem in $\mathbb{R}^d$,

$$\|u\|_{C^0} \leq C' \|u\|_{H^s}, \; s > d/2 \tag{7.9}$$

and

$$\|uv\|_{H^0} \leq \|u\|_{C^0} \|v\|_{H^0} \leq C' \|u\|_{H^s} \|v\|_{H^0}. \tag{7.10}$$

When u is a product of functions we apply (7.5) and obtain (7.7). □

As a direct corollary we obtain the following lemma.

Lemma 7.3. *Let densities* $P_n(\mathbf{r};\boldsymbol{\tau}) = P_n\left(\mathbf{r};\boldsymbol{\tau};\overrightarrow{\mathbf{e}}\right)$ *that depend on the parameters* $\boldsymbol{\tau} = (\tau_1,\ldots,\tau_\nu) \in \mathbb{R}_+^n$ *and act on* $\overrightarrow{\mathbf{e}} = (\mathbf{e}_1,\ldots,\mathbf{e}_n) \in \mathbb{C}^{3n}$ *satisfy Condition 1.2. If* $\mathbf{E}_j = \mathbf{E}_j(\mathbf{r})$ *are functions from* $\mathbf{H}^s$, $s > 3/2$, *then for any* $\boldsymbol{\tau} = \tau_1,\ldots,\tau_n$ $P_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}(\cdot)\right)$ *belongs to* $\mathbf{H}^s$ *and there exist positive constants* C, C' *and* γ *depending only on* s *such that*

$$\left\|P_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)\right\|_{\mathbf{H}^s} + \left\|\dot{P}_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)\right\|_{\mathbf{H}^s} \tag{7.11}$$

$$\leq \frac{C_s}{\gamma^n}\left(\|P_n(\boldsymbol{\tau})\|_{C^s} + \left\|\dot{P}_n(\boldsymbol{\tau})\right\|_{C^s}\right)\prod_{j=1}^n \|\mathbf{E}_j\|_{\mathbf{H}^s}. \tag{7.12}$$

In addition, for any $1 \leq i \leq n$,

$$\left\|P_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)\right\|_{\mathbf{H}^0} + \left\|\dot{P}_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)\right\|_{\mathbf{H}^0} \tag{7.13}$$

$$\leq \frac{C_s}{\gamma^n}\|\mathbf{E}_i\|_{\mathbf{H}^0}\left(\|P_n(\boldsymbol{\tau})\|_{C^0} + \left\|\dot{P}_n(\boldsymbol{\tau})\right\|_{C^0}\right)\prod_{j\neq i}\|\mathbf{E}_j\|_{\mathbf{H}^s}. \tag{7.14}$$

Lemma 7.4. *Let Condition 1.2 hold and* $\mathcal{P}_n(\mathbf{E})$ *be defined by (1.13). Then Condition 5.5 holds for the densities* $q_n\left(\overrightarrow{\tau};\cdot\right) = P_n(\cdot;\boldsymbol{\tau};\cdot)$ *with* $Y = \mathbf{H}^s$, $\beta = \gamma\beta_P$, $C_q = C_s C_P$. *The constants* C_s *and* γ *are the same as in (7.11),* C_P, β_P *are the same as in (1.18). For any* $T > 0$ *the series (1.12) determines an analytic operator* $\mathbf{P}_{NL} \in A(C_{\mathbf{P}}, R_{\mathbf{P}}, X, X)$ *where* $C_{\mathbf{P}} = C_s C_P$, $R_{\mathbf{P}} = \gamma\beta_P$, $X = \mathcal{C}_{0,\mathbf{H}^s}^T = C_0([-\infty,T];\mathbf{H}^s)$, $s > 3/2$. *The operator* $\partial_t\mathbf{P}_{NL}$ *is also an analytic operator in* $\mathcal{C}_{0,\mathbf{H}^s}^T$, $\partial_t\mathbf{P}_{NL} \in A(C_{\mathbf{P}}, R_{\mathbf{P}}, X, X)$. *The operators* $\mathbf{P}_{NL}$ *and* $\partial_t\mathbf{P}_{NL}$ *are respectively strictly causal and causal; they satisfy Condition 2.1.*

Proof. By (1.13) $\mathbf{P}_{\mathrm{NL}}$ is strictly causal. By Condition 1.2 and Lemma 7.3 $P_n = q_n$ satisfy the inequality (5.35), therefore Condition 5.5 holds. By Lemma 5.6 $\mathbf{P}_{\mathrm{NL}}, \partial_t\mathbf{P}_{\mathrm{NL}} \in A(C_s C_P, \gamma\beta_P, X, X)$. The fact that $\partial_t\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$ is causal follows from Lemma 5.4.

To check that Condition 2.1 is fulfilled for $\partial_t\mathbf{P}_{\mathrm{NL}}$ we use the fact that the multilinear operators $\partial_t\mathcal{P}_n\left(\overrightarrow{\mathbf{E}}\right)$ are represented in the form (5.13), (5.33) by the explicit formulas (1.14) involving the densities $P_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)$, $\dot{P}_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\right)$ to which we can apply the inequalities (7.13). We pick a test function $\psi(t,\mathbf{r})$ and, then, taking the strictly causal part $\mathcal{P}_{n,n}$ of (1.14) proceed similarly to (5.26):

$$\begin{aligned}
&\int_0^{T_1}\int \psi(t,\mathbf{r})\,\mathcal{P}_{n,n}\left(\overrightarrow{\mathbf{E}}\right)(\mathbf{r},t)\,d\mathbf{r}dt \\
&= \int_0^{T_1}\int \psi(t,\mathbf{r})\int_{-\infty}^t\cdots\int_{-\infty}^t \dot{P}_n\left(\mathbf{r};t\overrightarrow{1}-\boldsymbol{t};\overrightarrow{\mathbf{E}}(\mathbf{r},\boldsymbol{t})\right)\prod_j dt_j d\mathbf{r}dt \\
&\leq \int_0^{T_1}\int_0^\infty\cdots\int_0^\infty \left\|\dot{P}_n\left(\cdot;\boldsymbol{\tau};\overrightarrow{\mathbf{E}}\left(\cdot,t\overrightarrow{1}-\boldsymbol{\tau}\right)\right)\right\|_{\mathbf{H}^0}\|\psi(\cdot,t)\|_{\mathbf{H}^0}\prod_j d\tau_j dt.
\end{aligned}$$

Using the inequality (7.13) for a given i we find that the right-hand side is not greater than $C_s \gamma^{-n}$ times

$$\prod_{j \neq i} \left\| \mathbf{E}_j \right\|_{\mathcal{C}^T_{\mathbf{H}^s}} \int_0^{T_1} \int_0^{\infty} \cdots \int_0^{\infty} \left\| \dot{P}_n (\boldsymbol{\tau}) \right\|_{C^0} \| \mathbf{E}_i (t - \tau_i) \|_{\mathbf{H}^0} \| \psi (t) \|_{\mathbf{H}^0} \prod_j d\tau_j dt$$

$$\leq \prod_{j \neq i} \left\| \mathbf{E}_j \right\|_{\mathcal{C}^T_{\mathbf{H}^s}} ||\psi||_{L_2([0,T_1],\mathbf{H}^0)} \| \mathbf{E}_i \|_{L_2([0,T_1],\mathbf{H}^0)} \int_0^{\infty} \cdots \int_0^{\infty} \left\| \dot{P}_n (\boldsymbol{\tau}) \right\|_{C^0} \prod_j d\tau_j .$$

Using the inequality (1.18) we get

$$\int_0^{T_1} \int \psi (t, \mathbf{r}) \mathcal{P}_{n,n} \left(\overrightarrow{\mathbf{E}} \right) (\mathbf{r}, t) \, d\mathbf{r} dt$$
$$\leq C_s \gamma^{-n} \prod_{j \neq i} \left\| \mathbf{E}_j \right\|_{\mathcal{C}^T_{\mathbf{H}^s}} \| \psi \|_{L_2([0,T_1],\mathbf{H}^0)} \| \mathbf{E}_i \|_{L_2([0,T_1],\mathbf{H}^0)} C_P \beta_P^{-n} .$$

Setting $\psi (t, \mathbf{r}) = \mathcal{P}_{n,n} \left(\overrightarrow{\mathbf{E}} \right) (\mathbf{r}, t)$ we obtain

$$\left\| \mathcal{P}_{n,n} \left(\overrightarrow{\mathbf{E}} \right) \right\|_{L_2([0,T_1],\mathbf{H}^0)} \leq C_P C_s (\gamma \beta_P)^{-n} \prod_{j \neq i} \left\| \mathbf{E}_j \right\|_{\mathcal{C}^T_{\mathbf{H}^s}} \| \mathbf{E}_i \|_{L_2([0,T_1],\mathbf{H}^0)} . \tag{7.15}$$

Observe that a similar estimate holds for the term $\mathcal{P}_{n,n-1} \left(\overrightarrow{\mathbf{E}} \right)$ in (1.14). Hence

$$\left\| \partial_t \mathcal{P}_n \left(\overrightarrow{\mathbf{E}} \right) \right\|_{L_2([0,T_1],\mathbf{H}^0)} \leq 2 C_P C_s R_{\mathbf{P}}^{-n} \| \mathbf{E}_i \|_{L_2([0,T_1],\mathbf{H}^0)} \prod_{j \neq i} \left\| \mathbf{E}_j \right\|_{\mathcal{C}^T_{\mathbf{H}^s}} , \tag{7.16}$$

where $R_{\mathbf{P}} = \gamma \beta_P$. Using this inequality and the evaluations similar to (4.51)–(4.53) we obtain the following estimate:

$$\left\| \partial_t \mathcal{P}_n \left(\mathbf{E}_1^n \right) - \partial_t \mathcal{P}_n \left(\mathbf{E}_2^n \right) \right\|_{L_2([0,T_1],\mathbf{H}^0)}$$
$$\leq 2 C_P C_s R_{\mathbf{P}}^{-2n} \max \left(\| \mathbf{E}_1 \|_{\mathcal{C}^T_{\mathbf{H}^s}} , \| \mathbf{E}_2 \|_{\mathcal{C}^T_{\mathbf{H}^s}} \right)^{n-1} \| \mathbf{E}_1 - \mathbf{E}_2 \|_{L_2([0,T_1],\mathbf{H}^0)} ,$$

which, after the summation in n, allows to conclude that Condition 2.1 holds. □

Lemma 7.5. *Let $\mathcal{P}_n (\mathbf{E})$ be defined by (1.13) and Condition 1.2 hold with some $s \geq 2$. Let*

$$R_S = \frac{R_{\mathbf{P}} + 2 C_{\pi \mathbf{P}} - 2 \sqrt{R_{\mathbf{P}} C_{\pi \mathbf{P}} + C_{\pi \mathbf{P}}^2}}{1 + \gamma_\eta}, \quad \gamma_\eta = \| \boldsymbol{\eta} \|_{\mathbf{H}^s, \mathbf{H}^s} , \tag{7.17}$$

$$C_S = \frac{R_{\mathbf{P}}}{2 (R_{\mathbf{P}} + C_{\pi \mathbf{P}})} \left(R_{\mathbf{P}} + \left(1 + \gamma_\eta \right) R_S \right) - R_S, \tag{7.18}$$

where $C_{\pi \mathbf{P}} = 4 \pi \gamma_\eta C_{\mathbf{P}}$, and the constants $R_{\mathbf{P}}$, $C_{\mathbf{P}}$ are the same as in Lemma 7.4. Then for every $T > 0$ there exists a unique analytic operator $\mathcal{S}_{NL}$ in the space $X = \mathcal{C}^T_{0,\mathbf{H}^s}$ such that, first, $\mathcal{S}_{NL} \in A_* (C_S, R_S, X, X)$, $\mathcal{S}_{NL} \in A (C_S, R_S/e, X, X)$, and, second,

$$\mathbf{E} = \mathcal{S} (\mathbf{D}) = \boldsymbol{\eta} \mathbf{D} + \mathcal{S}_{NL} (\mathbf{D}) \tag{7.19}$$

solves Eq. (1.8) for $\|\mathbf{D}\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_S$. *The operator* $\mathcal{S}_{NL}(\mathbf{D})$ *is a strictly causal analytic function of* $\mathbf{D}$ *represented by the convergent power series*

$$\mathcal{S}_{NL}(\mathbf{D}) = \sum_{n \geq n_0} \mathcal{S}_n(\mathbf{D}), \ \|\mathbf{D}\|_{\mathcal{C}^T_{\mathbf{H}^s}} < R_S. \tag{7.20}$$

The operators $\mathcal{S}_n$ *satisfy the following recursive formulas:*

$$\mathcal{S}_1 = \boldsymbol{\eta}, \mathcal{S}_n(\cdot) = -4\pi\boldsymbol{\eta} \sum_{m \geq n_0, \ n_1+\cdots+n_m=n} \mathcal{P}_m\left(\mathcal{S}_{n_1}(\cdot), \ldots, \mathcal{S}_{n_m}(\cdot)\right), \ n \geq 2, \tag{7.21}$$

in particular

$$\mathcal{S}_n = 0, \ 2 \leq n \leq n_0 - 1, \ \mathcal{S}_{n_0}(\mathbf{D}) = -4\pi\boldsymbol{\eta}\mathcal{P}_{n_0}\left((\boldsymbol{\eta}\mathbf{D})^{n_0}\right). \tag{7.22}$$

The polynomials $\mathcal{S}_n(\mathbf{D}(\cdot))$ *are spatially local as in (1.13), (7.3).*

Proof. Let us rewrite the equality (1.8) to make it fit the form (4.24):

$$\mathbf{E}(\mathbf{r}, t) = \boldsymbol{\eta}(\mathbf{r})\mathbf{D}(\mathbf{r}, t) - 4\pi\boldsymbol{\eta}(\mathbf{r})\mathbf{P}_{\mathrm{NL}}(\mathbf{E})(\mathbf{r}, t), \tag{7.23}$$

where $\boldsymbol{\eta}(\mathbf{r})$ is defined by (3.2). Then we apply Theorem 4.8 with $L = \boldsymbol{\eta}$, $x = \mathbf{D}$, $z = \mathbf{E}$, $F(z) = -4\pi\boldsymbol{\eta}\mathbf{P}_{\mathrm{NL}}(\mathbf{E})$, $X = \mathcal{C}^T_{0,\mathbf{H}^s} = C_0([-\infty, T]; \mathbf{H}^s)$. Note that in (4.29) $F_{js} = 0$ when $s \neq 0$, $F_{j0} = F_j$ and since F does not depend on **D,** and (4.29) takes the form of (7.21). By Theorem 4.8 we obtain that (i) the series (7.20) converges; (ii) $\mathcal{S}_{\mathrm{NL}} \in A_*(C_S, R_S, X, X)$ and $\mathbf{E} = \mathcal{S}(\mathbf{D}) = \boldsymbol{\eta}\mathbf{D} + \mathcal{S}_{\mathrm{NL}}(\mathbf{D})$ is a solution of (7.23). By Corollary 4.6 we have $\mathcal{S}_{\mathrm{NL}} \in A_*(C_S, R_S/e, X, X)$. By Lemma 5.7 operators $\mathcal{S}_n$ are strictly causal. Since a composition of spatially local operators is spatially local, $\mathcal{S}_n$ defined by (7.21) are spatially local. □

Notice that the statements of Lemma 7.5 imply that the function $\mathcal{S}_{\mathrm{NL}}(\mathbf{D})$ has the radius of convergence which does not depend on T.

Lemma 7.6. *Assume that Condition 1.2 is satisfied. Then the operator* $\partial_t \circ \mathcal{S}_{NL} = \partial_t\mathcal{S}_{NL}$ *is an analytic operator such that* $\partial_t\mathcal{S}_{NL} \in A_*(C_E, R_E, X, X)$, $\partial_t\mathcal{S}_{NL} \in A(C_E, R_E/e, X, X)$, *where* $X = \mathcal{C}^T_{0,\mathbf{H}^s}$,

$$R_E = \frac{R_{\mathbf{P}}R_S}{R_{\mathbf{P}} + C_S}, C_E = 4\pi\gamma_\eta\frac{C_S C_{\mathbf{P}}}{R_{\mathbf{P}} + C_S}, \tag{7.24}$$

and the constants $C_{\mathbf{P}}, R_{\mathbf{P}}, C_E, R_E, \gamma_\eta$ *are as in Lemmas 7.4 and 7.5. The operator* $\partial_t\mathcal{S}_{NL}(\mathbf{D})$ *is represented by the power series*

$$\partial_t\mathbf{S}_{NL}(\cdot) = \sum_{n \geq n_0} \partial_t\mathcal{S}_n(\cdot), \tag{7.25}$$

where $\partial_t\mathcal{S}_n$ *satisfy the formulas based on* $\mathbf{P}_{NL}$ *and* $\mathcal{S}(\mathbf{D})$ *defined by (7.20)–(7.22)*

$$\partial_t\mathcal{S}_n(\cdot) = -\sum_{m \geq 1}\sum_{n_1+\cdots+n_m=n} 4\pi\boldsymbol{\eta}\partial_t\mathcal{P}_m\left(\mathcal{S}_{n_1}(\cdot), \ldots, \mathcal{S}_{n_m}(\cdot)\right), n \geq n_0, \tag{7.26}$$

where $\partial_t\mathcal{P}_m$ *are given in (1.14).*

Proof. According to Lemma 7.5 we can rewrite (7.23) in the form

$$\mathbf{E} = \boldsymbol{\eta}\mathbf{D} - 4\pi\boldsymbol{\eta}\mathbf{P}_{\mathrm{NL}}\left(\mathcal{S}\left(\mathbf{D}\right)\right), \tag{7.27}$$

and comparing with $\mathbf{E} = \boldsymbol{\eta}\mathbf{D} + \mathcal{S}_{\mathrm{NL}}\left(\mathbf{D}\right)$ we conclude that

$$\mathcal{S}_{\mathrm{NL}}\left(\mathbf{D}\right) = -4\pi\boldsymbol{\eta}\mathbf{P}_{\mathrm{NL}}\left(\mathcal{S}\left(\mathbf{D}\right)\right), \quad \partial_t\mathcal{S}_{\mathrm{NL}}\left(\mathbf{D}\right) = -4\pi\boldsymbol{\eta}\partial_t\mathbf{P}_{\mathrm{NL}}\left(\mathcal{S}\left(\mathbf{D}\right)\right). \tag{7.28}$$

Using Lemma 7.4 we conclude that $-4\pi\boldsymbol{\eta}\mathbf{P}_{\mathrm{NL}} \in A\left(4\pi\gamma_\eta C_{\mathbf{P}}, R_{\mathbf{P}}/e, X, X\right)$. According to Theorem 4.12 the composition $-4\pi\boldsymbol{\eta}\partial_t\mathbf{P}_{\mathrm{NL}}\left(\mathcal{S}\right)$ belongs to the classes $A_*\left(C_S, R_S, X, X\right)$ and $A\left(C_S, R_S/e, X, X\right)$. □

7.2. Nonlinear Maxwell equations in divergence-free variables and existence of a solution. To construct and study solutions to the Maxwell equations we need to recast the equations to an equivalent form which, firstly, involves only divergence free fields, and, secondly, provides means to control the spatial regularity of the fields as they evolve in time. For a spatially inhomogeneous medium when the material constants, in particular $\boldsymbol{\varepsilon}$, depend on the position vector $\mathbf{r}$, there is an advantage in selecting the electric inductance $\mathbf{D}$ rather than electric field $\mathbf{E}$ to be the primary field variable, because of the simplicity of the condition $\nabla \cdot \mathbf{D} = 0$ compared with $\nabla \cdot \left(\boldsymbol{\varepsilon}\left(\mathbf{r}\right)\mathbf{E}\right) = 0$. This advantage is even greater in the nonlinear case since Eq. (2.15) becomes nonlinear. For this reason we recast the constitutive relations (1.8) to express $\mathbf{E}\left(\cdot\right)$ as a function of $\mathbf{D}\left(\cdot\right)$.

Substituting the expression $\mathbf{E} = \mathcal{S}\left(\mathbf{D}\right)$ given by (7.19) into the Maxwell equations (1.1), (1.2) and (1.5) we get the following operator form of the Maxwell equations:

$$\partial_t\mathbf{U}\left(t\right) = -i\mathbf{M}\mathbf{U}\left(t\right) + iQ\left(\mathbf{U}\right)\left(t\right) - \mathbf{J}\left(t\right); \;\; \mathbf{U}\left(t\right) = 0 \text{ for } t \leq 0, \tag{7.29}$$

where $\mathbf{U}$, $\mathbf{M}$, $\mathbf{J}$ are given in (3.4),

$$Q\left(\mathbf{U}\right) = i\begin{bmatrix}\mathbf{0}\\ \nabla^{\times}\mathbf{S}_{\mathrm{NL}}\left(\mathbf{D}\right)\end{bmatrix}, \tag{7.30}$$

and

$$\nabla \cdot \mathbf{D} = \nabla \cdot \mathbf{B} = \nabla \cdot \mathbf{J}_D = \nabla \cdot \mathbf{J}_B = 0, \tag{7.31}$$

$$\mathbf{J}\left(t\right) = 0 \text{ for } t \leq 0. \tag{7.32}$$

We look for a solution $\mathbf{U}\left(t\right), -\infty < t \leq T$ that belongs to $C_0\left(\left[-\infty, T\right]; \mathbf{H}^s\right)$ with $s \geq 2$ and $\partial_t\mathbf{U} \in C_0\left(\left[-\infty, T\right]; \mathbf{H}^{s-1}\right)$. Evidently, (7.32) is consistent with the requirement $\mathbf{U}\left(t\right) = 0$ for $t \leq 0$ in (7.29). Note that for any t the function $\mathbf{S}_{\mathrm{NL}}\left(\mathbf{D}\right)\left(t\right)$ depends on the values of the field $\mathbf{U}\left(t'\right)$ at times $t' \leq t$ as in (1.13). Since the expression (7.30) for the nonlinearity $Q\left(\mathbf{U}\right)$ involves the curl operator $\nabla^{\times}$, it acts from $C_0\left(\left[-\infty, T\right]; \mathbf{H}^s\right)$ to $C_0\left(\left[-\infty, T\right]; \mathbf{H}^{s-1}\right)$. To reduce (7.29) to a regular integral form (7.42) with a bounded nonlinearity that acts in the space $C_0\left(\left[-\infty, T\right]; \mathbf{H}^s\right)$ we need to apply some transformations. We can do that by recasting the Maxwell system (7.29) in the form of their abstract version (6.16) considered in Sect. 6 by setting

$$\mathfrak{m} = \mathring{\mathbf{M}}, \; \mathcal{H}_{\mathfrak{m}}^s = \mathring{\mathbf{H}}_M^s, \; \mathcal{C}_{0,\mathcal{H}_{\mathfrak{m}}^s}^T = C\left(\left[-\infty, , T\right]; \mathcal{H}_{\mathfrak{m}}^s\right) = C\left(\left[-\infty, T\right]; \mathring{\mathbf{H}}_M^s\right), \tag{7.33}$$

where $\mathring{\mathbf{M}}$ is defined by (3.25). Below we verify the conditions imposed on abstract operators in Sects. 5, 6.

Recall that the Banach space $C_0\left([-\infty, T]; \mathring{\mathbf{H}}_M^s\right)$ consists of $\mathring{\mathbf{H}}_M^s$ -continuous trajectories $\mathbf{U} = \mathbf{U}(t)$, $-\infty < t \leq T$ in $\mathring{\mathbf{H}}_M^s$ satisfying the rest condition $\mathbf{U}(t) = 0$ for $t < 0$ with the norm

$$\|\mathbf{U}(\cdot)\|_{C\left([0,T];\mathring{\mathbf{H}}_M^s\right)} = \sup_{-\infty<t\leq T} \|\mathbf{U}(t)\|_{\mathring{\mathbf{H}}_M^s} = \sup_{0\leq t\leq T} \|\mathbf{U}(t)\|_{\mathring{\mathbf{H}}_M^s}. \tag{7.34}$$

According to (3.21) we have in (7.30),

$$\nabla^{\times}\mathbf{S}_{\mathrm{NL}} = \nabla^{\times}\Pi_0\mathbf{S}_{\mathrm{NL}} = \nabla^{\times}\Pi_0\mathring{\boldsymbol{\eta}}\mathring{\boldsymbol{\eta}}^{-1}\Pi_0\mathbf{S}_{\mathrm{NL}} = \nabla^{\times}\mathring{\boldsymbol{\eta}}\mathring{\boldsymbol{\eta}}^{-1}\Pi_0\mathbf{S}_{\mathrm{NL}}, \tag{7.35}$$

where $\boldsymbol{\eta}$ is the operator of multiplication by a matrix $\boldsymbol{\eta}(\mathbf{r})$,$\mathring{\boldsymbol{\eta}}$ is the restriction of $\Pi_0\boldsymbol{\eta}$ to $\mathring{\mathbf{L}}_2$, and $\mathring{\boldsymbol{\eta}}^{-1}$ is the inverse of $\mathring{\boldsymbol{\eta}}$ on $\mathring{\mathbf{L}}_2$, the inverse exists according to Lemma 3.2. Using this identity we rewrite the Maxwell equations (7.29)– (7.31) in the form

$$\partial_t\mathbf{U}(t) = -i\mathring{\mathbf{M}}[\mathbf{U}(t) + \mathfrak{q}(\mathbf{U})(t)] - \mathbf{J}(t), \tag{7.36}$$

where $\mathring{\mathbf{M}}$ is defined by (3.25)

$$\mathring{\mathbf{M}} = i\begin{bmatrix} \mathbf{0} & \mathring{\nabla}^{\times} \\ -\nabla^{\times}\mathring{\boldsymbol{\eta}} & \mathbf{0}, \end{bmatrix}, \ \mathfrak{q}(\mathbf{U}) = \sum_{n=n_0}^{\infty} \mathfrak{q}_n(\mathbf{U}), \ \mathfrak{q}_n(\mathbf{U}) = \mathring{\Xi}^{-1}\begin{bmatrix} \Pi_0\mathcal{S}_n(\mathbf{D}) \\ \mathbf{0} \end{bmatrix}, \tag{7.37}$$

where $\mathring{\Xi}^{-1}$ is the inverse on $\mathring{\mathbf{L}}_2^2$ of $\mathring{\Xi} = \Pi_0^{(2)}\Xi\mathring{\Pi}_0^{(2)}$ from (3.5), $\Pi_0^{(2)}$ is defined in (3.38); the inverse exists according to Lemma 3.3. We assume that the impressed currents in (1.4) satisfy the following condition:

$$\mathbf{J}_D, \mathbf{J}_B \in L_{1,0}\left([-\infty, T]; \mathring{\mathbf{H}}^s\right), \tag{7.38}$$

and we look for a solution

$$\mathbf{U} = \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}, \ \mathbf{D}, \mathbf{B} \in C_0\left([-\infty, T]; \mathring{\mathbf{H}}^s\right). \tag{7.39}$$

Recall that the conditions $\mathbf{J}_D, \mathbf{J}_B \in L_{1,0}\left([-\infty, T]; \mathring{\mathbf{H}}^s\right)$ and, similarly, $\mathbf{D}, \mathbf{B} \in C_0\left([-\infty, T]; \mathring{\mathbf{H}}^s\right)$ include the divergence-free conditions (7.31) as well as the rest conditions (1.6) and (7.32).

By Lemma 3.1 the operator $\mathring{\mathbf{M}}$ in (3.25) is self-adjoint, therefore Lemma 6.1 is applicable and transformations (6.16)–(6.20) are applicable, with notations (7.33). Clearly, (7.36) has the form of (6.16). Let us look now at the nonlinear equation (7.36) and, in particular, on its term $\mathring{\mathbf{M}}\mathfrak{q}$ which involves the differentiation with respect to space variables of the nonlinear function of the fields. Following (6.16) –(6.20) we can recast Eqs. (7.36) and trade off the space derivatives in $\mathring{\mathbf{M}}\mathfrak{q}$ for the time derivative $\partial_t\mathfrak{q}(\mathbf{U})$ which, as we show, results in an analytic operator with respect to $\mathbf{U}$ for $\mathfrak{q}$ of the form (7.37). Note that in the framework of Sect. 6 Eq. (6.12) becomes

$$\mathbf{U}_0(t) = -\int_{-\infty}^{t} e^{-i(t-t')\mathring{\mathbf{M}}}\mathbf{J}(t')\,dt', \tag{7.40}$$

obviously $\mathbf{U}_0(t) = 0$ for $t < 0$ since $\mathbf{J}(t) = 0$ for $t < 0$. Notice that (6.14) implies

$$\|\mathbf{U}_0\|_{C_0\left([-\infty,T];\mathring{\mathbf{H}}^s_M\right)} \leq \int_{-\infty}^{T} \|\mathbf{J}(t)\|_{\mathring{\mathbf{H}}^s_M} \, dt \tag{7.41}$$

for any $T \geq 0$. Using (6.20) we rewrite (7.36) in the *regular integral form*

$$\mathbf{U}(t) = \mathbf{U}_0(t) - \mathfrak{q}(\mathbf{U}(t)) + \int_{-\infty}^{t} e^{-i(t-t')\mathring{\mathbf{M}}} \partial_{t'} [\mathfrak{q}(\mathbf{U})]\left(t'\right) dt'. \tag{7.42}$$

By Lemma 6.2 this equation is equivalent to (7.36). Lemma 6.2 is applicable since the time derivative $\partial_t \circ \mathfrak{q}(\mathbf{U})$ is a bounded, analytic operator according to the following lemma.

Lemma 7.7. *The function* $\mathfrak{q}$ *defined by (7.37) is an analytic function in* $\mathcal{C}^T_{\mathring{\mathbf{H}}^s_M}$ *for every* $T > 0$. *We also have that* $\mathfrak{q} \in A_*\left(C_Q, R_Q, \mathcal{C}^T_{\mathbf{H}^s}, \mathcal{C}^T_{\mathbf{H}^s}\right)$ *and* $\mathfrak{q}$ *belongs to* $A\left(C_Q, R_Q/e, \mathcal{C}^T_{\mathring{\mathbf{H}}^s_M}, \mathcal{C}^T_{\mathring{\mathbf{H}}^s_M}\right)$, *where*

$$R_Q = R_S c_-, C_Q = C_S c_+ \left\|\mathring{\Xi}^{-1}\right\|_{\mathring{\mathbf{H}}^s_M, \mathring{\mathbf{H}}^s_M} \tag{7.43}$$

with R_S, C_S *being as in Lemma 7.5,* $\left\|\mathring{\Xi}^{-1}\right\|_{\mathring{\mathbf{H}}^s_M, \mathring{\mathbf{H}}^s_M}$ *being the norm of the operator* $\mathring{\Xi}^{-1}$ *in* $\mathring{\mathbf{H}}^s_M$ *and* c_+, c_- *being as in Lemma 3.4. The operators* $\partial_t \mathfrak{q} \in A_*\left(C'_Q, R'_Q, \mathcal{C}^T_{\mathbf{H}^s}, \mathcal{C}^T_{\mathbf{H}^s}\right)$, $\partial_t \mathfrak{q} \in A\left(C'_Q, R'_Q/e, \mathcal{C}^T_{\mathring{\mathbf{H}}^s_M}, \mathcal{C}^T_{\mathring{\mathbf{H}}^s_M}\right)$,

$$R'_Q = R_E c_-, C'_Q = C_E c_+ \left\|\mathring{\Xi}^{-1}\right\|_{\mathring{\mathbf{H}}^s_M, \mathring{\mathbf{H}}^s_M}, \tag{7.44}$$

where R_E, C_E *are the same as in Lemma 7.6.*

Proof. Operator $\mathfrak{q}$ is obtained from $\mathcal{S}_{\text{NL}}$ by applying the linear operator $-\mathring{\Xi}^{-1}$ to $\begin{pmatrix} \Pi_0 \mathcal{S}_{\text{NL}} \\ 0 \end{pmatrix}$. By Lemma 3.4 the norms in $\mathring{\mathbf{H}}^s_M$, $\mathring{\mathbf{H}}^s$ and $\mathbf{H}^s$ are equivalent on $\mathring{\mathbf{H}}^s$, and, hence, for $\mathbf{V} \in \mathring{\mathbf{H}}^s$,

$$\begin{aligned} \|\Pi_0 \mathcal{S}_n \mathbf{V}\|_{\mathring{\mathbf{H}}^s_M} &\leq c_+ \|\Pi_0 \mathcal{S}_n \mathbf{V}\|_{\mathring{\mathbf{H}}^s} \leq c_+ \|\mathcal{S}_n \mathbf{V}\|_{\mathbf{H}^s} \leq c_+ \|\mathcal{S}_n\|_{\mathbf{H}^s, \mathbf{H}^s} \|\mathbf{V}\|^n_{\mathbf{H}^s} \\ &= c_+ \|\mathcal{S}_n\|_{\mathbf{H}^s, \mathbf{H}^s} \|\mathbf{V}\|^n_{\mathring{\mathbf{H}}^s} \leq c_+ c_-^{-n} \|\mathcal{S}_n\|_{\mathbf{H}^s, \mathbf{H}^s} \|\mathbf{V}\|^n_{\mathring{\mathbf{H}}^s_M} \leq c_+ c_-^{-n} C_S R_S^{-n} \|\mathbf{V}\|^n_{\mathring{\mathbf{H}}^s_M} \end{aligned}$$

with similar inequalities holding for $\partial_t \mathcal{S}_n$, $n \geq n_0$. The norm of $\mathring{\Xi}^{-1}$ in $\mathring{\mathbf{H}}^s$ is bounded according to Lemma 3.3. Since norms in $\mathring{\mathbf{H}}^s_M$, and $\mathring{\mathbf{H}}^s$ are equivalent, the norm of $\mathring{\Xi}^{-1}$ in $\mathring{\mathbf{H}}^s_M$ is bounded too. This implies the statements of the lemma for $\mathfrak{q}$. Taking into account that the operator ∂_t commutes with $\mathring{\Xi}^{-1}$ and Π_0 and using Lemma 7.6 we obtain the statements for $\partial_t \mathfrak{q}$. □

Here are our main statements on the Maxwell equations (7.42).

Theorem 7.8. *Assume that Conditions 1.1 and 1.2 are satisfied with* $s \geq 2$*. Let* R_Q, C_Q, R'_Q, C'_Q *be the same as in Lemma 7.7,*

$$R_{\mathfrak{q}} = \min\left(R_Q, R'_Q\right), \quad C_{\mathfrak{q}} = C_Q + C'_Q. \tag{7.45}$$

Let $T > 0$*,* $\mathbf{J} \in L_{1,0}\left([-\infty, T;];\mathring{\mathbf{H}}^s_M\right)$*,*

$$\|\mathbf{J}\|_{L_1\left([-\infty,T;];\mathring{\mathbf{H}}^s_M\right)} = \int_{-\infty}^{T} \|\mathbf{J}(t)\|_{\mathring{\mathbf{H}}^s_M}\, dt < \delta_0, \tag{7.46}$$

where $\delta_0 \leq R_{\mathfrak{q}}/8$ *is small enough for the following condition to hold:*

$$1 + T < \frac{R_{\mathfrak{q}}\left(8\delta_0 R_{\mathfrak{q}}/R_{\mathfrak{q}}\right)^{1-n_0}}{2C_{\mathfrak{q}}}. \tag{7.47}$$

Let $\mathbf{U}_0 \in \mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M} = C_0\left([-\infty, T], \mathring{\mathbf{H}}^s_M\right)$ *be given in terms of* $\mathbf{J}$ *by (7.40). Then there exists a uniquely determined analytic operator* $\mathcal{U}(\mathbf{U}_0)$ *in the space* $C^T_{0,\mathring{\mathbf{H}}^s_M}$ *such that* $\mathbf{U} = \mathcal{U}(\mathbf{U}_0)$ *gives a solution to Eq. (6.21) and (7.36). In addition to that,* $\|\mathbf{U}\|_{\mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}} < R_Q$ *and* $\mathfrak{q}(\mathbf{U})$ *is well-defined. The solution* $\mathbf{U}$ *expands into series*

$$\begin{gathered}\mathbf{U}(t) = \mathcal{U}(\mathbf{U}_0)(t) = \mathbf{U}_0(t) + \sum_{n\geq n_0} \mathcal{U}_n(\mathbf{U}_0)(t), \\ \|\mathcal{U}_n(\mathbf{U}_0)\|_{\mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}} \leq C_G R_G^{-n} \|\mathbf{U}_0\|^n_{\mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}}, \quad n \geq n_0,\end{gathered} \tag{7.48}$$

where C_G, R_G *are as in Theorem 4.8 and* $R_F = R_{\mathfrak{q}}$*,* $C_F = (1+T)\, C_{\mathfrak{q}}$*,* $\gamma_L = 1$*. The operators* $\mathcal{U}_n$*,* $n = 1, 2, \ldots$ *satisfy the following recursive formulas:*

$$\mathcal{U}_1 \text{ is the identity, } \mathcal{U}_n(\cdot) = \sum_{m\geq n_0,\ n_1+\cdots+n_m=n} \mathfrak{R}_m\left(\mathcal{U}_{n_1}(\cdot), \ldots, \mathcal{U}_{n_m}(\cdot)\right), \ n \geq n_0, \tag{7.49}$$

where $\mathfrak{R}_m$ *are the relevant terms of the analytic function* $\mathfrak{R}$ *defined by (6.21), (6.28) with* $\mathfrak{q}_n$ *defined by (7.37). The first significant term* $\mathcal{U}_{n_0}$*,* the first nonlinear response*, in (7.48) is represented by*

$$\mathcal{U}_{n_0}(\mathbf{U}_0)(t) = \mathbf{U}_0(t) - \mathfrak{q}_{n_0}(\mathbf{U}_0)(t) + \int_0^t e^{-i\mathring{\mathbf{M}}(t-t')}\partial_{t'}\mathfrak{q}_{n_0}(\mathbf{U}_0)\left(t'\right) dt' \tag{7.50}$$

$$= \mathbf{U}_0(t) - i\int_0^t e^{-i\mathring{\mathbf{M}}(t-t')}\mathring{\mathbf{M}}\mathfrak{q}_{n_0}(\mathbf{U}_0)\left(t'\right) dt'. \tag{7.51}$$

Proof. Lemma 7.7 implies that the Maxwell equation (7.36) is a particular case of the abstract Maxwell equation (6.16). The conditions of Theorem 6.3 where $R_{\mathfrak{q}}$ and $C_{\mathfrak{q}}$ are given in (7.45) are satisfied. By Theorem 6.3 there exists an operator $\mathcal{U} \in A\left(C_G, R_G, \mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}, \mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}\right)$, and, hence, (7.48) holds. This theorem implies that $\mathbf{U} = \mathcal{U}(\mathbf{U}_0)$ is a solution of (7.42). According to Lemma 6.2, $\mathbf{U}$ is a solution of (7.36). Formulas (7.48)–(7.51) follow from (6.25)–(6.27) in Theorem 6.3. The formula (7.51) is obtained from (7.50) using integration by parts. □

Now we give the proof of Theorem 1.3. Note that since $\mathbf{U}_0 = \left(\partial_t + i\mathring{\mathbf{M}}\right)^{-1} \mathbf{J}$, where $\left(\partial_t + i\mathring{\mathbf{M}}\right)^{-1}$ is a linear bounded operator in $\mathcal{C}^T_{0,\mathring{\mathbf{H}}^s_M}$ (see Lemma 6.1) analyticity and power expansions in terms of $\mathbf{U}_0$ imply also the analytic dependence with corresponding power expansion with respect to $\mathbf{J}$.

Proof (Proof of Theorem 1.3). The existence of $\mathbf{B}, \mathbf{D} \in C_0\left([-\infty, T]; \mathbf{H}^s\right)$ that solve (7.42) follows from Theorem 7.8, and, hence, $(\mathbf{B}, \mathbf{D}) = \mathcal{U}(\mathbf{U}_0)$ and $\partial_t \mathbf{B}$, $\partial_t \mathbf{D}$ belong to $C_0\left([-\infty, T]; \mathbf{H}^{s-1}\right)$. We define $\mathbf{E} = \mathcal{S}(\mathbf{D})$ as a solution of (1.8) with already found $\mathbf{D}$, according to Lemma 7.5 $\mathbf{E} \in C_0\left([-\infty, T]; \mathbf{H}^s\right)$, $\partial_t \mathbf{E} \in C_0\left([-\infty, T]; \mathbf{H}^{s-1}\right)$. In turn, the function $\mathbf{H}$ is defined by (1.7) in terms of $\mathbf{B}$. A pair of functions $\mathbf{B}, \mathbf{E}$ is a solution of (2.13), (2.14), (2.15) and $(\mathbf{D}, \mathbf{B}, \mathbf{E}, \mathbf{H}) \in C_0\left([-\infty, T]; \mathbf{H}^s\right)$ is a solution of (1.1), (1.2), (1.7), (1.8), (1.6) from the class considered in Definition 2.2. Since $\mathbf{B}, \mathbf{E}$ are unique by Theorem 2.4 and $\mathbf{D}, \mathbf{H}$ are uniquely determined by $\mathbf{B}, \mathbf{E}$ from (1.7), (1.8) the solution is unique. □

8. Extension to More General Cases

8.1. General dielectric media. When analyzing nonlinear dielectric media we assumed for the sake of simplicity that the medium is not magnetic with the magnetic permeability $\mu = 1$. In fact, all the results still hold if the medium is a general bianisotropic (magnetoelectric), inhomogeneous and nonlinear medium with the material relations more general than in (1.8), namely of the form (see [29], Sect. 1.1)

$$\mathbf{V}(\mathbf{r}, t) = \Xi(\mathbf{r}) \mathbf{U}(\mathbf{r}, t) + \mathbf{K}_{\mathrm{NL}}(\mathbf{V})(\mathbf{r}, t), \ \mathbf{U} = \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}, \tag{8.1}$$

where $\Xi = \Xi(\mathbf{r})$, $\mathbf{r} \in \mathbb{R}^3$ is a Hermitian 6×6 matrix (not necessarily of the form (2.11)) and $\mathbf{K}_{\mathrm{NL}}$ is the nonlinear component of the material relations. The expressions for $\mathbf{K}_{\mathrm{NL}}$ are of the form similar to (1.12), (1.13):

$$\mathbf{K}_{\mathrm{NL}}(\mathbf{U}) = \sum_{n=n_0}^{\infty} \mathcal{K}_n(\mathbf{U}), \quad n_0 \geq 2, \quad \mathcal{K}_n(\mathbf{U}) = \mathcal{K}_n(\mathbf{U}, \ldots, \mathbf{U}), \tag{8.2}$$

$$\mathcal{K}_n(\mathbf{U}) = \int_{-\infty}^{t} \cdots \int_{-\infty}^{t} K_n\left(\mathbf{r}; t - t_1, \ldots, t - t_n; \mathbf{U}(\mathbf{r}, t_1), \ldots, \mathbf{U}(\mathbf{r}, t_n)\right) \prod_{j=1}^{n} dt_j, \tag{8.3}$$

$$K_n(\mathbf{r}; \tau_1, \ldots, \tau_n; \cdot) : \left(\mathbb{C}^6\right)^n \to \mathbb{C}^6, \ n \geq n_0.$$

Note that like in Lemma 7.5 it is easy to show that (8.1) is equivalent to the relation

$$\mathbf{V}(\mathbf{r}, t) = \Xi(\mathbf{r}) \mathbf{U}(\mathbf{r}, t) + \mathcal{Q}_{\mathrm{NL}}(\mathbf{U})(\mathbf{r}, t). \tag{8.4}$$

The Maxwell equation can be written in the same operator form as (7.36), namely

$$\partial_t \mathbf{U}(t) = -i\mathring{\mathbf{M}}\left[\mathbf{U}(t) + \mathfrak{q}(\mathbf{U})(t)\right] - \mathbf{J}(t), \tag{8.5}$$

where

$$\mathbf{U} \in C_0\left([-\infty, T]; \mathbf{H}^s \times \mathbf{H}^s\right), \tag{8.6}$$

$\mathring{\mathbf{M}}$ is defined by

$$\mathring{\mathbf{M}} = i\mathring{\nabla}^{\times\times}\Xi, \ \mathring{\nabla}^{\times\times} = \begin{bmatrix} \mathbf{0} & \mathring{\nabla}^{\times} \\ -\mathring{\nabla}^{\times} & \mathbf{0} \end{bmatrix}, \ [\Xi\mathbf{V}](\mathbf{r}) = \Xi(\mathbf{r})\mathbf{V}(\mathbf{r}), \tag{8.7}$$

and the nonlinearity is given by

$$\mathfrak{q}(\mathbf{U}) = \sum_{n=n_0}^{\infty} \mathfrak{q}_n(\mathbf{U}), \quad \mathfrak{q}_n(\mathbf{U}) = i\mathring{\Xi}^{-1}\Pi_0^{(2)}\mathcal{Q}_n(\mathbf{U}). \tag{8.8}$$

We assume that $\Xi(\mathbf{r})$ satisfies the following condition.

Condition 8.1. *Let $\Xi(\mathbf{r})$ be a positive definite Hermitian 6×6 matrix, which is a measurable function of $\mathbf{r}$ and satisfies (3.6) We also assume that there exists an integer $s \geq 2$ such that*

$$\|\Xi\|_{C^s(\mathbb{R}^3)} < \infty, \quad \left\|\Xi^{-1}\right\|_{C^s(\mathbb{R}^3)} < \infty. \tag{8.9}$$

An examination of the arguments shows that the statements of Theorem 7.8 still hold for the general Maxwell equations (8.5)–(8.8) provided that the linear generalized polarization $\Xi(\mathbf{r})$ satisfies Condition 8.1 and the nonlinear generalized polarization $\mathbf{K}_{\mathrm{NL}}(\mathbf{U})$ satisfies Condition 1.2 where P_n are replaced with K_n.

8.2. Coefficients from Sobolev classes. The smoothness requirements on dependence on $\mathbf{r}$ of the medium coefficients $\boldsymbol{\varepsilon}(\mathbf{r})$ and $P_n(\mathbf{r}, \cdot)$ were imposed in Conditions 1.1 and 1.2. The conditions are formulated in terms of the spaces $C^s\left(\mathbb{R}^3\right)$ of s times continuously differentiable functions, namely they require that $\boldsymbol{\varepsilon} \in C^s\left(\mathbb{R}^3\right)$ and $P_n \in C^s\left(\mathbb{R}^3\right)$. These conditions can be relaxed allowing the coefficients to be in the local Sobolev spaces $\mathbf{WB}_2^s\left(\mathbb{R}^d\right)$ of bounded functions defined as follows. The space $\mathbf{WB}_2^s\left(\mathbb{R}^d\right)$ consists of functions that are locally in $\mathbf{W}_2^s\left(\mathbb{R}^d\right)$ with the local $\mathbf{W}_2^s$-norms being uniformly bounded, namely

$$\|\mathbf{V}\|^2_{\mathbf{WB}_2^s(\mathbb{R}^d)} = \sup_{\mathbf{y}\in\mathbb{R}^d} \sum_{0\leq l_1+\ldots+l_d\leq s} \int_{|r|\leq 1} \left|\partial_1^{l_1}\ldots\partial_d^{l_d}\mathbf{V}(\mathbf{r}+\mathbf{y})\right|^2 \mathbf{dr}, \ s = 1, 2, \ldots. \tag{8.10}$$

The following statements are proven in [6].

Lemma 8.2. *Let $s > d/2$, $f \in \mathbf{W}_2^s\left(\mathbb{R}^d\right)$, $g \in \mathbf{WB}_2^s\left(\mathbb{R}^d\right)$. Then*

$$\|fg\|_{\mathbf{W}_2^s(\mathbb{R}^d)} \leq C_1 \|f\|_{\mathbf{W}_2^s(\mathbb{R}^d)} \|g\|_{\mathbf{WB}_2^s(\mathbb{R}^d)}, \tag{8.11}$$

where C_1 depends only on s and d.

Lemma 8.3. *Let f be an h-linear tensor in $\mathbb{R}^n$ with coefficients that depend on the variable $\mathbf{r} \in \mathbb{R}^d$. Assume that the coefficients belong to $\mathbf{WB}_2^s\left(\mathbb{R}^d\right)$ with $s > d/2$. Then for $g \in \left(\mathbf{W}_2^s\left(\mathbb{R}^d\right)\right)^h$,*

$$\|f(g_1, ..., g_h)\|_{\mathbf{W}_2^s(\mathbb{R}^d)} \leq C_2 C_3^{h-1} \|f\|_{\mathbf{WB}_2^s(\mathbb{R}^d)} \|g_1\|_{\mathbf{W}_2^s(\mathbb{R}^d)} \cdots \|g_h\|_{\mathbf{W}_2^s(\mathbb{R}^d)}, \tag{8.12}$$

where C_2, C_3 depend only on s, n and d.

The following lemma shows that the spaces $\mathring{\mathbf{H}}^s$ and $\mathring{\mathbf{H}}^s_M$ are equivalent when $\Xi \in \mathbf{WB}_2^s$.

Lemma 8.4. *Let $s \geq 2$, $\|\Xi\|_{\mathbf{WB}_2^s} < \infty$. Then (3.42) is still true.*

Proof. According to (3.7) the inequality (3.42) has the form

$$c_- \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \left(\mathring{\mathbf{M}}^s \mathbf{U}, \Xi \mathring{\mathbf{M}}^s \mathbf{V}\right)_{\mathbf{L}_2} + (\mathbf{U}, \Xi \mathbf{V})_{\mathbf{L}_2} \leq c_+ \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s}. \tag{8.13}$$

By (3.6) inequality (8.13) is equivalent to

$$c'_- \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \leq \left(\mathring{\mathbf{M}}^s \mathbf{U}, \mathring{\mathbf{M}}^s \mathbf{V}\right)_{\mathbf{L}_2} + (\mathbf{U}, \mathbf{V})_{\mathbf{L}_2} \leq c'_+ \|\mathbf{V}\|_{\mathring{\mathbf{H}}^s \times \mathring{\mathbf{H}}^s} \tag{8.14}$$

with $c'_-, c'_+ > 0$. Since infinitely smooth functions that belong to $\cap_l \mathring{\mathbf{H}}^l$ are dense in $\mathring{\mathbf{H}}^s$, it is sufficient to prove (8.14) for such functions $\mathbf{V}$. By (3.21) $\mathring{\mathbf{M}}^s \mathbf{U} = M^s \mathbf{U}$, where M is the same as in Lemma 8.6, therefore the Ξ -dependent part of (8.14) coincides with $(M^s \mathbf{U}, M^s \mathbf{V})_{\mathbf{L}_2}$. By Lemma 8.6 it depends continuously on Ξ in $\mathbf{WB}_2^s\left(\mathbb{R}^3\right)$. By Lemma 3.4 (8.14) holds for $\Xi \in C^s\left(\mathbb{R}^3\right)$; since $C^s\left(\mathbb{R}^3\right)$ is dense in $\mathbf{WB}_2^s\left(\mathbb{R}^3\right)$, passing to the limit in (8.14) we obtain (8.14) for Ξ in $\mathbf{WB}_2^s\left(\mathbb{R}^3\right)$. □

In our results concerning the Maxwell equations the condition $\boldsymbol{\varepsilon}, P_n \in C^s\left(\mathbb{R}^3\right)$ that requires continuity of $s-$ th order spatial derivatives of the coefficients can be relaxed to a less restrictive condition $\boldsymbol{\varepsilon}, P_n \in \mathbf{WB}_2^s\left(\mathbb{R}^3\right)$ that requires local square integrability of the derivatives. The exact statements are given in Theorem 8.5; their proof is based on Lemmas 8.2, 8.3, 8.6 and 8.4. The proof of Lemma 8.6 is rather technical and is given after the proof of Theorem 8.5.

Theorem 8.5. *Let $s \geq 2$. Assume that Conditions 1.1 and 1.2 hold with the following quantities being replaced: $\|\boldsymbol{\varepsilon}\|_{C^s}$ by $\|\boldsymbol{\varepsilon}\|_{\mathbf{WB}^s}$, $\|\boldsymbol{\eta}\|_{C^s}$ by $\|\boldsymbol{\eta}\|_{\mathbf{WB}^s}$, $\|P_n\|_{C^s}$ by $\|P_n\|_{\mathbf{WB}^s}$, $\left\|\dot{P}_n\right\|_{C^s}$ by $\left\|\dot{P}_n\right\|_{\mathbf{WB}_2^s}$. Then the statements of Theorems 1.3 and 7.8 are true (with modified constants R_Q, C_Q, R'_Q, C'_Q).*

Proof. The proofs of Theorems 1.3 and 7.8 are based on the properties of the linear Maxwell operator $\mathring{\mathbf{M}}$ and the nonlinearity $\mathbf{P}_{\mathrm{NL}}$ described by Lemmas 3.4 and 7.3, respectively. By Lemmas 8.2 and 8.3 the inequality (7.6) can be replaced by

$$\|P_n\|_{\mathbf{H}^s, \mathbf{H}^s} \leq C_s \gamma_1^{-n} \|P_n\|_{\mathbf{WB}_2^s}. \tag{8.15}$$

Hence, $\left\|P_{n,\nu}(\cdot; \boldsymbol{\tau}; \cdot)\right\|_{C^s}$ in Lemma 7.3 can be replaced by $\left\|P_{n,\nu}(\cdot; \boldsymbol{\tau}; \cdot)\right\|_{\mathbf{WB}_2^s}$ with a modified γ. According to Lemma 8.4 the inequality (3.42) of Lemma 3.4 holds too. Therefore the statements of Lemmas 3.4 and 7.3 can be applied in this case and Theorems 1.3 and 7.8 hold too. □

Lemma 8.6. *Let* $M = \nabla^{\times\times} \circ \Theta$, *where* $\nabla^{\times\times}$ *is defined by (3.5),* $\Theta = \Theta(\mathbf{r})$ *is a* 6×6 *matrix with smooth* $\mathbf{r}$*-dependent coefficients from* $C^s(\mathbb{R}^3)$, *let* $\mathbf{V} \in \cap_s \mathbf{H}^s$. *Then* $(M^{\mathbf{s}}\mathbf{V}, M^{\mathbf{s}}\mathbf{V})_{\mathbf{L}_2}$ *continuously depends on the matrix* Ξ *in the metric of* $\mathbf{WB}_2^s(\mathbb{R}^3)$ *for* $s \geq 2$.

Proof. Let us introduce

$$h(\Theta_1, ..., \Theta_{2s}) = \left(\prod_{i=1}^{s} \left(\nabla^{\times\times}\Theta_i\right) \mathbf{V}, \prod_{i=s+1}^{2s} \left(\nabla^{\times\times}\Theta_i\right) \mathbf{V} \right)_{\mathbf{L}_2}, \tag{8.16}$$

where Θ_i, $i = 1, ..., 2s$ are matrices with entries from $C^s(\mathbb{R}^3)$. Obviously, $h_s(\boldsymbol{\Theta})$ is a $2s$-linear form of Θ_i, $i = 1, ..., 2s$. The continuity of this form is equivalent to its boundedness. Note that using Leibnitz formula we can obtain the following representation:

$$\prod_{i=1}^{s} \left(\nabla^{\times\times}\Theta_i\right) \mathbf{V} = \sum_{|\beta| \leq s} A_\beta(\Theta_1, ..., \Theta_s)\, \partial^\beta \mathbf{V}, \tag{8.17}$$

where $A_\beta(\Theta_1, ..., \Theta_s)$ is a matrix of the form

$$A_\beta(\Theta_1, ..., \Theta_s) = \sum K_1\Theta_1 ... K_s\Theta_s \tag{8.18}$$

and K_j are matrix differential operators with constant coefficients of order $n_j \geq 0$ satisfying

$$n_1 + ... + n_s = s - |\beta|, \tag{8.19}$$

and the number of terms in the sum depends only on s.

To prove boundedness of (8.16) we substitute (8.17) into (8.16) and get

$$\begin{aligned} h(\Theta_1, ..., \Theta_{2s}) &\leq C \sum_{|\beta| \leq s} h_\beta(\Theta_1, ..., \Theta_{2s}), \\ h_\beta(\Theta_1, ..., \Theta_{2s}) &= \int \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\partial^\beta \mathbf{V}\right|^2 d\mathbf{r}. \end{aligned} \tag{8.20}$$

Let $\phi_0(\mathbf{r})$ be an infinitely smooth, nonnegative function such that

$$\phi_0(\mathbf{r}) \geq 0,\ \phi_0(\mathbf{r}) = 1, \text{as } |\mathbf{r}| \leq 3, \quad \phi(\mathbf{r}) = 0, \text{as } |\mathbf{r}| \geq 6. \tag{8.21}$$

Notice that the supports of all functions $\phi_0(\mathbf{r} - \mathbf{l})$, when $\mathbf{l}$ runs the set $\mathbb{Z}^3$ of 3 -dimensional integer valued vectors cover the entire $\mathbb{R}^3$. We also use the function

$$\phi_1(\mathbf{r}) = \phi_0(\mathbf{r}/(6)). \tag{8.22}$$

Clearly, $\phi_1(\mathbf{r}) \geq 1$ when $\phi_0(\mathbf{r}) \neq 0$, therefore there exists a constant $C(s)$,

$$\sup_{|\alpha| \leq s} \left|\partial^\alpha \phi_0(\mathbf{r})\right| \leq C(s)\, \phi_1(\mathbf{r}), \mathbf{r} \in \mathbb{R}^3. \tag{8.23}$$

In addition, let

$$\Phi(\mathbf{r}) = \sum_{\mathbf{l}\in\mathbb{Z}^d} \phi_0^2(\mathbf{r}-\mathbf{l}), \quad \Phi_1(\mathbf{r}) = \sum_{\mathbf{l}\in\mathbb{Z}^d} \phi_1(\mathbf{r}-\mathbf{l}). \tag{8.24}$$

$\Phi(\mathbf{r}), \Phi_1(\mathbf{r})$ are infinitely smooth periodic functions, $\Phi(\mathbf{r}), \Phi_1(\mathbf{r}) \geq 1$. Note that

$$h_\beta(\Theta_1, ..., \Theta_{2s}) = \sum_{\mathbf{l}\in\mathbb{Z}^d} \int \frac{\phi_0^2(\cdot-\mathbf{l})}{\Phi} \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\partial^\beta \mathbf{V}\right|^2 d\mathbf{r}$$
$$\leq \left\|\frac{1}{\Phi}\right\|_{C^0} \sum_{\mathbf{l}\in\mathbb{Z}^d} \int \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r}.$$

We consider one term in the above sum

$$\int \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r}. \tag{8.25}$$

When $s - |\beta| < 3/2$ we use Holder inequality with

$$1/p + 1/p' = 1, \quad 3/(2p) = 3/2 - (s - |\beta|) \tag{8.26}$$

$$\int \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r}$$
$$\leq \left(\int_{|\mathbf{r}-\mathbf{l}|\leq 6} \left|A_\beta(\Theta_1, ..., \Theta_s)\right|^{p'} \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right|^{p'} d\mathbf{r}\right)^{1/p'} \Psi_{\beta 1}, \tag{8.27}$$

where

$$\Psi_{\beta 1}(\phi_0(\cdot-\mathbf{l})\,\mathbf{V}) = \left(\int \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^{2p} d\mathbf{r}\right)^{1/p}. \tag{8.28}$$

When $s - |\beta| \geq 3/2$, we take $p' = 1$ and

$$\int \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r}$$
$$\leq \int_{|\mathbf{r}-\mathbf{l}|\leq 6} \left|A_\beta(\Theta_1, ..., \Theta_s)\right| \left|A_\beta(\Theta_{s+1}, ..., \Theta_{2s})\right| d\mathbf{r}'\,\Psi_{\beta 2} \tag{8.29}$$

with

$$\Psi_{\beta 2}(\phi_0(\cdot-\mathbf{l})\,\mathbf{V}) = \sup_{|\mathbf{r}-\mathbf{l}|\leq 6} \left|\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right|^2. \tag{8.30}$$

In both cases using the Sobolev embedding theorem and (8.23) we obtain for $\Psi_{\beta i}$, $i = 1, 2$, the estimate

$$\Psi_{\beta i} \leq C \left\|\phi_0(\cdot-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right\|^2_{H^{s-|\beta|}(B_l)}$$
$$= C \sum_{|\alpha|\leq s-|\beta|} \int \left|\partial^\alpha\left(\phi_0(\mathbf{r}-\mathbf{l})\,\partial^\beta \mathbf{V}(\mathbf{r})\right)\right|^2 d\mathbf{r}$$

$$\leq C_1 \sum_{|\alpha|\leq s} \int \phi_1 (\mathbf{r}-\mathbf{l}) \left|\partial^\alpha \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r}.$$

When $s-|\beta|<3/2$ using again Holder inequality with $1/q_1+...+1/q_{2s}=1$ we get

$$\left(\int_{|\mathbf{r}-\mathbf{l}|\leq 6} |K_1\Theta_1...K_s\Theta_s|^{p'} |K_{s+1}\Theta_{s+1}...K_{2s}\Theta_{2s}|^{p'} d\mathbf{r}\right)^{1/p'}$$
$$\leq \prod_{i=1}^{2s} \left(\int_{|\mathbf{r}-\mathbf{l}|\leq 6} |K_i\Theta_i|^{q_i p'} d\mathbf{r}\right)^{1/(p'q_i)} = \prod_{i=1}^{2s} \|K_i\Theta_i\|_{L_{q_i p'}(B_l)}. \tag{8.31}$$

We take

$$\frac{1}{q_i} = \frac{n_i}{\bar{n}}, \bar{n} = n_1 + ...n_{2s}, \tag{8.32}$$

where n_i are from (8.19); $\frac{1}{q_i}=0$ corresponds to the L_∞ norm of $K_i\Theta_i$ (it coincides with the C^0 norm since $K_i\Theta_i(\mathbf{r})$ are continuous). By the Sobolev embedding theorem (see [42, 35]) in the domain

$$B_l = \left\{\mathbf{r}\in\mathbb{R}^3 : |\mathbf{r}-\mathbf{l}|\leq 6\right\}, \tag{8.33}$$

$$\|V\|_{L_p(B_l)} \leq C(p)\|V\|_{H^l(B_l)}, \quad -\frac{3}{p} \leq l - \frac{3}{2}, 1\leq p<\infty, \tag{8.34}$$

$$\|V\|_{C^0(B_l)} \leq C\|V\|_{H^l(B_l)}, \quad 0 < l - \frac{3}{2}. \tag{8.35}$$

By (8.19) $\bar{n}=2s-2|\beta|$ and by (8.26) $1/p = 1-2(s-|\beta|)/3$, therefore

$$\frac{3}{q_i p'} = \frac{2(s-|\beta|)n_i}{\bar{n}} = \frac{2(s-|\beta|)n_i}{2s-2|\beta|} = n_i, \tag{8.36}$$

and since $s>3/2$

$$\frac{3}{q_i p'} = n_i < n_i + \frac{3}{2} - s. \tag{8.37}$$

Applying (8.34) with $l = s-n_i$ we get from (8.31),

$$\left(\int_{|\mathbf{r}-\mathbf{l}|\leq 6} |K_1\Theta_1...K_s\Theta_s|^{p'} |K_{s+1}\Theta_{s+1}...K_{2s}\Theta_{2s}|^{p'} d\mathbf{r}\right)^{1/p'}$$
$$\leq C_1' \prod_{i=1}^{2s} \|K_i\Theta_i\|_{H^{s-n_i}} \leq C_1'' \prod_{i=1}^{2s} \|\Theta_i\|_{H^s}.$$

By (8.19) we get

$$\left(\int_{|\mathbf{r}-\mathbf{l}|\leq 6} \left|A_\beta(\Theta_1,...,\Theta_s)\right|^{p'} \left|A_\beta(\Theta_{s+1},...,\Theta_{2s})\right|^{p'} d\mathbf{r}\right)^{1/p'}$$
$$\leq C_2 \prod_{i=1}^{s} \|\Theta_i\|_{H^s(B_l)} \leq C_3 \prod_{i=1}^{s} \|\Theta_i\|_{\mathbf{WB}_2^s(\mathbb{R}^3)}.$$

Therefore

$$\begin{aligned} h_\beta(\Theta_1, ..., \Theta_{2s}) &\leq C_4 \prod_{i=1}^{s} \|\Theta_i\|_{\mathbf{WB}_2^s(\mathbb{R}^3)} \sum_{\mathbf{l}\in\mathbb{Z}^d} \sum_{|\alpha|\leq s} \int \phi_1(\mathbf{r}-\mathbf{l}) \left|\partial^\alpha \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r} \\ &= C_4 \prod_{i=1}^{s} \|\Theta_i\|_{\mathbf{WB}_2^s(\mathbb{R}^3)} \sum_{|\alpha|\leq s} \int \Phi_1(\mathbf{r}) \left|\partial^\alpha \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r} \\ &\leq C_5 \prod_{i=1}^{s} \|\Theta_i\|_{\mathbf{WB}_2^s(\mathbb{R}^3)} \sum_{|\alpha|\leq s} \int \left|\partial^\alpha \mathbf{V}(\mathbf{r})\right|^2 d\mathbf{r} \\ &= \|\mathbf{V}\|^2_{\mathbf{H}^s(\mathbb{R}^3)} C_5 \prod_{i=1}^{s} \|\Theta_i\|_{\mathbf{WB}_2^s(\mathbb{R}^3)} . \end{aligned}$$

After summation in β we obtain boundedness of $h(\Theta_1, ..., \Theta_{2s})$ which implies the continuity of the $2s$ - linear form $h(\Theta_1, ..., \Theta_{2s})$. □

Acknowledgement and Disclaimer. Effort of A. Babin and A. Figotin is sponsored by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number F49620-01-1-0567. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government.

References

1. Adams, D., Hedberg, L.: *Function Spaces and Potential Theory*. Berlin-Heidelberg-New York: Springer, 1996
2. Aközbek, N., John, S.: Optical solitary waves in two- and three-dimensional nonlinear photonic band-gap structures. Phys. Rev. E **57**(2), 2287–2319 (1998)
3. Axmann, W., Kuchment, P.: An efficient finite element method for computing spectra of photonic and acoustic band-gap materials. J. Comp. Phys. **150**, 468–481 (1999)
4. Atiyah, M.F.: Resolution of singularities and division of distributions. Commun. Pure Appl. Math. **23**(2), 145–150 (1970)
5. Babin, A., Figotin, A.: Nonlinear Photonic Crystals I. Quadratic nonlinearity. Waves in Random Media **11**, R31–R102 (2001)
6. Babin, A., Figotin, A.: *Multilinear Spectral Decomposition for Nonlinear Maxwell Equations*. Am. Math. Soc. Transl. (2), **206**, Providence, RI: Am. Math. Soc., 2002
7. Bulgakov, A., Bulgakov, S., Vazquez, L.: Second-harmonic resonant excitation in optical periodic structures with nonlinear anisotropic layers. Phys. Rev. E. **58**(6), 7887–7898 (1998)
8. Berger, V.: Nonlinear Photonic Crystals. Phys. Rev. Lett. **81**, 4136–4139 (1999)
9. Birman, M, Sh. Solomyak, M.Z.: L^2-theory of the maxwell operator in arbitrary domains. Russ. Math. Surv. **42**(6), 75–96 (1987)
10. Boyd, R.: *Nonlinear Optics*. London-New York: Academic Press, 1992
11. Butcher, P., Cotter, D.: *The Elements of Nonlinear Optics*. Cambridge: Cambridge University Press, 1990
12. Bennink, R., Yoon, Y., Boyd, R., Sipe, J.: Accessing the Optical Nonlinearity of Metals with Metal-Dielectric Photonic Bandgap Structures. Optics Letters **24**(20), 1416–1418 (1999)
13. Centini, M. et al.: Dispersive properties of finite, One-dimensional photonic bandgap structures: Applications to nonlinear quadratic interactions. Phys. Rev. E. **60**(4), 4891–4898 (1999)
14. Dineen, S.: *Complex Analysis on Infinite Dimensional Spaces*. Berlin-Heidelberg-New York: Springer, 1999

15. Dautray, R., Lions, J.-L.: *Mathematical Analysis and Numerical Methods for Science and Technology*. Vol. **2**, Functional and Variational Methods, Berlin-Heidelberg-New York: Springer-Verlag, 1988
16. Felsen, L., Marcuvits, N.: *Radiation and Scattering of Waves*. Oxford: Oxford University Press, IEEE Press, 1994
17. Figotin, A., Klein, A.: Localization of classical waves II. electromagnetic waves. Commun. Math. Phys. **184**, 411–441 (1997)
18. Fogel, I. et al.: Spontaneous emission and nonlinear effects in photonic bandgap materials. Pure Appl. Opt. **7**, 393–407 (1998)
19. Gilbarg, D., Trudinger, N.: *Elliptic Partial Differential Equations of Second Order*. Berlin-Heidelberg-New York: Springer-Verlag, 1983
20. Georgieva, A., Kricherbauer, T., Venakides, S.: Wave propagation and resonance in one-dimesional nonlinear discrete periodic medium. SIAM J. Appl. Math. **60**(1), 272–294 (1999)
21. Hale, J.: *Asymptotic Behavior of Dissipative Systems*. Providence, RI: AMS, 1988
22. Hale, J.: *Functional Differential Equations*. New York: Springer-Verlag, 1971
23. Hale, J.K., Verduyn Lunel, S.M.: *Introduction To Functional Differential Equations*. New York: Springer-Verlag, 1993
24. Hattori, T., Tsurumachi, N., Nakatsuka, H.: Analysis of optical nonlinearity by defect states in one-dimensional photonic crystals. J. Opt. Soc. Am. B. **14**(2), 348–355 (1997)
25. Haus, J., Viswanathan, Scalora, M., Kalocsai, A., Cole, J., Theimer, J.: Enhanced second-harmonic generation in media with a weak periodicity. Phys. Rev. A **57**(3), 2120–2128 (1998)
26. Hille, E., Phillips, R.S.: *Functional Analysis and Semigroups*. Providence, RI: AMS, 1991
27. Jackson, J.: *Classical Electrodynamics*. New York: Wiley, 1975
28. Kato, T.: *Perturbation Theory for Linear Operators*. Berlin-Heidelberg-New York: Springer, 1995
29. Kong, J.A.: *Electromagnetic Wave Theory*. New York: Wiley, 1990
30. Martorell, J., Vilaceca, R., Corbalan, R.: Second harmonic generation in a photonic crystal. Appl. Phys. Lett. **70**(6), 702–704 (1997)
31. Nelson, R., Boyd, R.: Enhanced third-order nonlinear optical response of photonic bandgap materials. J. Modern Optics **46**(7), 1061–1069 (1999)
32. Reed, M., Simon, B.: *Functional Analysis*. Vol. **1**, New York: Academic Press, 1972
33. Reed, M., Simon, B.: *Analysis of Operators*. Vol. **4**, New York: Academic Press, 1978
34. Shubin, M.A.: *Pseudo-differential operators and spectral theory*. Berlin: Springer Verlag, 2001
35. Stein, E.M.: *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. **30**, Princeton, NJ: Princeton University Press, 1970
36. Scalora, M., Bloemer, M.J., Manka, A.S. et al.: Pulsed second-harmonic generation in nonlinear, one-dimensional, periodic structures. Phys. Rev. A **56**, 3166–3174 (1997)
37. Tarasishin, A., Zheltikov, A., Magnitskii, S.: Matched second-harmonic generation of ultrasoft laser pulses in photonic crystals. JETP Lett. **70**(12), 819–825 (1999)
38. Taylor, M.E.: *Partial Differential Equations, **I**. Basic Theory*. Berlin-Heidelberg-New York: Springer, 1996
39. Taylor, M.E., *Partial Differential Equations, **III**. Nonlinear Equations*. Berlin-Heidelberg-New York: Springer, 1996
40. Taylor, M.E.: *Pseudodifferential operators and nonlinear PDE*. Progress in Mathematics V. **100**, Boston: Birkhauser, 1991
41. Tran, P.: Optical limiting and switching of short pulses by use of a nonlinear photonic bandgap structures with a defect. J. Opt. Soc. B **14**(10), 2589–2595 (1997)
42. Triebel, H.: *Theory of Function Spaces*. Basel-Boston: Birkhauser, 1983
43. Whittaker, E., Watson G.: *A Course of Modern Analysis*. Cambridge: Cambridge Univ. Press, 1996
44. Winn, J., Fan, S., Joannopoulos, J.: Interband transitions in photonic crystals. Phys. Rev. B **59**(3), 1551–1554 (1999)
45. Wloka, J.T.: *Partical Differential Equations*. Cambridge: Cambridge University Press, 1992
46. Wloka, J.T., Rowley, B., Lawruk, B.: *Boundary Problems for Elliptic Systems*. Cambridge: Cambridge University Press, 1995
47. Wang, Z. et al.: Nonlinear transmission resonance in a two-dimensional periodic structure with photonic band gap. Phys. Rev. B. **56**(15), 9185–9188 (1997)
48. Zeidler, E.: *Nonlinear Functional Analysis and its Applications*. Vol. **I**: Fixed-Point Theorems, Berlin-Heidelberg-New York: Springer, 1986
49. Zeidler, E.: *Nonlinear Functional Analysis and its Applications*. Vol. **IIA**: Linear Monotone Operators, Berlin-Heidelberg-New York: Springer, 1990
50. Zhu, Y., Ming, N.: Dielectric superlattices for nonlinear optical effects. Optical and Quantum Electronics **31**, 1093–1128 (1999)

Communicated by P. Constantin