

ON THE STATISTICAL MECHANICS OF MATTER IN AN ELECTROMAGNETIC FIELD. I

DERIVATION OF THE MAXWELL EQUATIONS FROM ELECTRON THEORY

by P. MAZUR and B. R. A. NIJBOER

Instituut voor theoretische natuurkunde, Utrecht, Nederland

Synopsis

Maxwell's macroscopic field equations are derived from the fundamental microscopic equations of electron theory in a new way. Instead of the usual space-time averaging procedure a statistical ensemble averaging method is applied, which is perhaps more satisfactory both from a physical and from a mathematical point of view. The treatment is valid for multi-component systems, in which every component may move in an arbitrary way, so that *e.g.* diffusion phenomena are included. Furthermore it would seem that the method given here is also suitable for the discussion of other problems connected with the behaviour of matter in an electromagnetic field.

1. *Introduction.* As is well-known Lorentz¹⁾ was the first to give a derivation of Maxwell's equations in material bodies from the fundamental equations of his electron theory by averaging the microscopic field quantities over physically infinitesimal space and time regions. This procedure has with only slight modifications been taken over by various authors^{2) 3) 4)}. In his recent monograph on the theory of electrons Rosenfeld⁵⁾ has given a very careful exposition of this derivation, which is based again upon the same kind of averaging procedure. However, even in this refined treatment the derivation is by no means straightforward from a mathematical point of view and requires rather too much verbal explanation. As a consequence inconsistencies in the frequently implicit assumptions easily escape notice and we shall see that even Rosenfeld's work is not completely blameless in this respect. It should also be noted that the actual averaging process is only performed over space regions and the time averaging is not essential for obtaining the

correct form of Maxwell's equations. Further, quantities like polarization and magnetization are defined here as space averages whereas in an actual evaluation of these quantities from a molecular theory they are defined as ensemble averages (cf. e.g. the theories of the dielectric constant by Yvon⁶⁾, Kirkwood⁷⁾, Brown⁸⁾, De Boer⁹⁾ and others.

In the present work it will be shown that a straightforward derivation of Maxwell's equations from electron theory can be given on the basis of ensemble averaging. We will treat a very general case here, in particular it will be shown that there is no need to restrict ourselves, as is done usually, to a medium at rest or moving as a whole with uniform velocity with respect to the observer. The formalism used is analogous to the one applied by Irving and Kirkwood¹⁰⁾ in their derivation of hydrodynamics from statistical mechanics. It will be seen in subsequent sections that the method applied here allows not only a neat derivation of Maxwell's equations, but moreover seems to offer a rather wide scope for the treatment of other problems concerned with the behaviour of material bodies in an electromagnetic field.

2. *Distribution functions.* Let us consider a system of N atoms which we number by the index k ($k = 1, 2, \dots, N$). The k -th atom consists of a point-nucleus with charge e_k , and of Z_k point-electrons with charges e_{ki} ($i = 1, 2, \dots, Z_k$). Position and momentum of the nucleus of the k -th atom are denoted by the vectors \mathbf{R}_k and \mathbf{p}_k respectively. The electrons of the k -th atom have positions relative to the nucleus denoted by \mathbf{r}_{ki} and canonically conjugate momenta \mathbf{p}_{ki} *) †).

The state of the whole system may be represented by a point in a phase space of $\sum_{k=1}^N 6(Z_k + 1)$ dimensions corresponding to all coordinates and momenta. The probability of finding our system in

*) The formalism to be developed here is obviously not restricted to a system consisting of identical and (or) neutral atoms. Mixtures of atoms and (or) ions are therefore included in this treatment. For simplicity's sake, polyatomic molecules have not been considered specifically, but the formalism could evidently be generalized so as to include the latter. In the above notation electrons with relative coordinates \mathbf{r}_{ki} are supposed to be bound to a nucleus. If free electrons do occur they should be included by treating them formally on the same footing as the nuclei.

†) It will be noted that our considerations are based upon a classical model of the atom. In connection with a possible extension of this work to a quantum statistical treatment we will come back to this point in section 7 (cf. also appendix I).

a certain state is obtained by means of a probability distribution function in this phase space which we denote by

$$f = f(\mathbf{R}_k^N, \mathbf{r}_{ki}^{NZ_k}, \mathbf{p}_k^N, \mathbf{p}_{ki}^{NZ_k}; t). \quad (1)$$

Here $\mathbf{r}_{ki}^{NZ_k}$, for instance, means that f depends on all coordinate vectors \mathbf{r}_{ki} where k runs from 1 to N and i from 1 to Z_k . Integration of any quantity over the whole phase space will be denoted by

$$\langle \dots \rangle = \int \dots \prod_k d\mathbf{R}_k \prod_{k,i} d\mathbf{r}_{ki} \prod_k d\mathbf{p}_k \prod_{k,i} d\mathbf{p}_{ki}. \quad (2)$$

The distribution function f is normalized in such a way that

$$\langle f \rangle = 1. \quad (3)$$

According to the principles of statistical mechanics the average of any dynamical quantity a

$$a = a(\mathbf{R}_k^N, \mathbf{r}_{ki}^{NZ_k}, \mathbf{p}_k^N, \mathbf{p}_{ki}^{NZ_k}; \mathbf{R}, t), \quad (4)$$

which may depend not only on the state of the system but also explicitly on the coordinates in ordinary three dimensional space and on the time, is then found from

$$\alpha_{Av}(\mathbf{R}, t) = \langle af \rangle. \quad (5)$$

If averages are to be taken of vectorial or tensorial quantities eq. (5) should be applied to each of their components.

We shall also need a kind of atomic distribution function

$$g_k = g_k(\mathbf{R}_k, \mathbf{r}_{ki}; t) = \int f \prod_{j \neq k} d\mathbf{R}_j \prod_{\substack{j,i \\ j \neq k}} d\mathbf{r}_{ji} \prod_j d\mathbf{p}_j \prod_{j,i} d\mathbf{p}_{ji}, \quad (6)$$

where the integration on the right-hand side is to be performed over all variables with the exception of the coordinates \mathbf{R}_k and \mathbf{r}_{ki} of nucleus and electrons of the k -th atom.

One can now also average a dynamical quantity a (cf. eq. (4)) over this restricted number of variables. We then obtain a new kind of average, denoted by \bar{a} , which still depends on \mathbf{R}_k and \mathbf{r}_{ki} (and maybe also on \mathbf{R} and t) and is given by

$$g_k \bar{a}(\mathbf{R}_k, \mathbf{r}_{ki}; \mathbf{R}, t) = \int af \prod_{j \neq k} d\mathbf{R}_j \prod_{\substack{j,i \\ j \neq k}} d\mathbf{r}_{ji} \prod_j d\mathbf{p}_j \prod_{j,i} d\mathbf{p}_{ji}. \quad (7)$$

We then have

$$\alpha_{Av} = \langle af \rangle = \int \bar{a} g_k d\mathbf{R}_k \prod_i d\mathbf{r}_{ki}, \quad (8)$$

as can be seen immediately from (5) and (7).

§ 3. *Time derivative of a dynamical quantity.* From conservation of probability in phase space it follows that

$$\frac{\partial f}{\partial t} = - \sum_k \nabla_{\mathbf{R}_k} \cdot f \dot{\mathbf{R}}_k - \sum_{k,i} \nabla_{\mathbf{r}_{ki}} \cdot f \dot{\mathbf{r}}_{ki} - \sum_k \nabla_{\mathbf{p}_k} \cdot f \dot{\mathbf{p}}_k - \sum_{k,i} \nabla_{\mathbf{p}_{ki}} \cdot f \dot{\mathbf{p}}_{ki}, \quad (9)$$

where, for instance, the differential operator $\nabla_{\mathbf{R}_k}$ is the ordinary nabla operator and $\dot{\mathbf{R}}_k$ the velocity in the \mathbf{R}_k -subspace.

We are now able to evaluate the time derivative of the average of a dynamical quantity a as follows

$$\frac{\partial \alpha_{Av}}{\partial t} = \frac{\partial}{\partial t} \langle a f \rangle = \left\langle \frac{\partial a}{\partial t} f \right\rangle + \langle a \frac{\partial f}{\partial t} \rangle, \quad (10)$$

where (5) has been used.

Substituting (9) into the second term of the last member of (10) and applying Gauss' theorem we obtain, since f falls off rapidly as coordinates and momenta tend to infinity

$$\frac{\partial \alpha_{Av}}{\partial t} = \left\langle \frac{da}{dt} f \right\rangle, \quad (11)$$

where d/dt is the total time derivative in phase space, i.e.

$$\frac{da}{dt} = \frac{\partial a}{\partial t} + \sum_k \dot{\mathbf{R}}_k \cdot \nabla_{\mathbf{R}_k} a + \sum_{k,i} \dot{\mathbf{r}}_{ki} \cdot \nabla_{\mathbf{r}_{ki}} a + \sum_k \dot{\mathbf{p}}_k \cdot \nabla_{\mathbf{p}_k} a + \sum_{k,i} \dot{\mathbf{p}}_{ki} \cdot \nabla_{\mathbf{p}_{ki}} a. \quad (12)$$

The simple result (11) has also been obtained by Irving and Kirkwood¹⁰⁾ by making use of Liouville's theorem. This, however, as the above derivation shows, is not necessary at all, the law of conservation of probability being already sufficient.

§ 4. *Definitions of electric and magnetic moments.* As has already been done in a similar case by Irving and Kirkwood¹⁰⁾, we can express the probability per unit volume, that the nucleus of the k -th atom (whatever its momentum and whatever the position of all other particles, including its own electrons) be at \mathbf{R} at time t , in the following way

$$\langle \delta(\mathbf{R}_k - \mathbf{R}) f \rangle, \quad (13)$$

where $\delta(\mathbf{R}_k - \mathbf{R})$ represents Dirac's δ -function.

Now an atom may be said to lie within a certain volume whenever its nucleus has its position inside that volume. Therefore we may

define the average number of atoms per unit volume at position \mathbf{R} at time t , which we will call the number density n , as

$$n(\mathbf{R}, t) = \sum_k \langle \delta(\mathbf{R}_k - \mathbf{R}) \rangle. \quad (14)$$

The charge of the k -th atom is given by

$$\sum_i (e_k + e_{ki}).$$

The average true charge density ϱ at \mathbf{R} is obtained on multiplying this expression by (13) and summing over k

$$\begin{aligned} \varrho(\mathbf{R}, t) &= \langle \sum_{k,i} (e_k + e_{ki}) \delta(\mathbf{R}_k - \mathbf{R}) \rangle \\ &= \sum_k \int \sum_i (e_k + e_{ki}) g_k \Pi \, d\mathbf{r}_{kj}, \end{aligned} \quad (15)$$

where (6) has been applied.

The electric current due to the motion of the k -th atom as a whole is analogously given by

$$\sum_i (e_k + e_{ki}) \dot{\mathbf{R}}_k.$$

Therefore the mean electric current at \mathbf{R} and time t is

$$\frac{\langle \sum_{k,i} (e_k + e_{ki}) \dot{\mathbf{R}}_k \delta(\mathbf{R}_k - \mathbf{R}) \rangle}{\langle \sum_k \delta(\mathbf{R}_k - \mathbf{R}) \rangle}. \quad (16)$$

From this expression the average true current density \mathbf{i} at \mathbf{R} and t is obtained by multiplication with the number density (14), so that we find, on application of (7),

$$\begin{aligned} \mathbf{i}(\mathbf{R}, t) &= \langle \sum_{k,i} (e_k + e_{ki}) \dot{\mathbf{R}}_k \delta(\mathbf{R}_k - \mathbf{R}) \rangle = \\ &= \sum_k \int \sum_i (e_k + e_{ki}) \overline{\dot{\mathbf{R}}_k} g_k(\mathbf{R}, \mathbf{r}_{kj}; t) \Pi \, d\mathbf{r}_{kj}, \end{aligned} \quad (17)$$

the average $\overline{\dot{\mathbf{R}}_k}$ as well as the atomic distribution function g_k being functions of \mathbf{R} , \mathbf{r}_{kj} and t . From (15) and (17) one may with the help of (11) easily verify the law of conservation of charge

$$\frac{\partial \varrho}{\partial t} = -\nabla_{\mathbf{R}} \cdot \mathbf{i}$$

In a similar way starting from the usual definitions of atomic

electric dipole and quadrupole and magnetic dipole moments μ_k^{el} , q_k^{el} and μ_k^m

$$\mu_k^{el} = \sum_i e_{ki} \mathbf{r}_{ki}, \quad (18)$$

$$q_k^{el} = \sum_i e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki}^*), \quad (19)$$

$$\mu_k^m = \frac{1}{2}c^{-1} \sum_i e_{ki} \mathbf{r}_{ki} \wedge \dot{\mathbf{r}}_{ki}, \quad (20)$$

(where the moments of the k -th atom are defined with respect to the position \mathbf{R}_k of its nucleus as origin and where c is the velocity of light) we obtain for the densities of these moments

$$\begin{aligned} \mathbf{P}(\mathbf{R}, t) &= \langle \sum_{k,i} e_{ki} \mathbf{r}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle = \\ &= \sum_k \int \sum_i e_{ki} \mathbf{r}_{ki} g_k(\mathbf{R}, \mathbf{r}_{kj}; t) \Pi \, d\mathbf{r}_{kj}, \end{aligned} \quad (21)$$

$$\begin{aligned} \mathbf{Q}(\mathbf{R}, t) &= \langle \frac{1}{2} \sum_{k,i} e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle = \\ &= \sum_k \int \frac{1}{2} \sum_i e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki} g_k(\mathbf{R}, \mathbf{r}_{kj}; t) \Pi \, d\mathbf{r}_{kj}, \end{aligned} \quad (22)$$

$$\begin{aligned} \mathbf{M}(\mathbf{R}, t) &= \langle \frac{1}{2}c^{-1} \sum_{k,i} e_{ki} \mathbf{r}_{ki} \wedge \dot{\mathbf{r}}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle = \\ &= \sum_k \int \frac{1}{2}c^{-1} \sum_i e_{ki} \mathbf{r}_{ki} \wedge \dot{\mathbf{r}}_{ki} g_k(\mathbf{R}, \mathbf{r}_{kj}; t) \Pi \, d\mathbf{r}_{kj}. \end{aligned} \quad (23)$$

It is seen that an atom gives a contribution to the moment densities at the point \mathbf{R} only when its nucleus is at \mathbf{R} .

We shall also need the following space and time derivatives of the above mentioned moment densities (compare appendix II on tensor notation):

The divergence of the polarization \mathbf{P} is given by (cf. (21))

$$\nabla_{\mathbf{R}} \cdot \mathbf{P} = \sum_k \int \sum_i e_{ki} \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} g_k \Pi \, d\mathbf{r}_{kj}. \quad (24)$$

The time derivative of \mathbf{P} can be evaluated on application of (11) and (12) to the second member of (21)

$$\begin{aligned} \frac{\partial \mathbf{P}}{\partial t} &= \langle \int \frac{d}{dt} \sum_{k,i} e_{ki} \mathbf{r}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle \\ &= \langle \int \sum_{k,i} e_{ki} \mathbf{r}_{ki} \{ \dot{\mathbf{R}}_k \cdot \nabla_{\mathbf{R}_k} \delta(\mathbf{R}_k - \mathbf{R}) \} \rangle + \langle \int \sum_{k,i} e_{ki} \dot{\mathbf{r}}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle. \end{aligned} \quad (25)$$

*) cf. appendix II.

Integrating the first term of this expression by parts we obtain

$$\begin{aligned} \frac{\partial \mathbf{P}}{\partial t} &= - \langle \sum_{k,i} e_{ki} \mathbf{r}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \nabla_{\mathbf{R}_k} \cdot \dot{\mathbf{R}}_k f \rangle + \langle f \sum_{k,i} e_{ki} \dot{\mathbf{r}}_{ki} \delta(\mathbf{R}_k - \mathbf{R}) \rangle = \\ &= - \sum_k \int \sum_i e_{ki} \mathbf{r}_{ki} \nabla_{\mathbf{R}} \cdot \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj} + \sum_k \int \sum_i e_{ki} \overline{\dot{\mathbf{r}}}_{ki} g_k \Pi \, d\mathbf{r}_{kj}, \quad (26) \end{aligned}$$

using also (7). Eq. (26) may be considered as a balance equation for the polarization density.

In a similar way we obtain for the derivatives of \mathbf{Q}

$$\nabla_{\mathbf{R}} \cdot \mathbf{Q} = \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} (\mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} g_k) \Pi \, d\mathbf{r}_{kj}, \quad (27)$$

$$\nabla_{\mathbf{R}} \cdot (\nabla_{\mathbf{R}} \cdot \mathbf{Q}) = \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki} : \nabla_{\mathbf{R}} \nabla_{\mathbf{R}} g_k \Pi \, d\mathbf{r}_{kj} \quad (*), \quad (28)$$

$$\begin{aligned} \frac{\partial \mathbf{Q}}{\partial t} &= - \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki} \nabla_{\mathbf{R}} \cdot \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj} + \\ &+ \sum_k \int \sum_i \frac{1}{2} e_{ki} (\overline{\dot{\mathbf{r}}}_{ki} \mathbf{r}_{ki} + \mathbf{r}_{ki} \overline{\dot{\mathbf{r}}}_{ki}) g_k \Pi \, d\mathbf{r}_{kj}, \quad (29) \end{aligned}$$

$$\nabla_{\mathbf{R}} \cdot \frac{\partial \mathbf{Q}}{\partial t} = - \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \{ \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} (\nabla_{\mathbf{R}} \cdot \overline{\mathbf{R}}_k g_k) \} \Pi \, d\mathbf{r}_{kj} \quad (30)$$

$$+ \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \nabla_{\mathbf{R}} \cdot \overline{\dot{\mathbf{r}}}_{ki} g_k \Pi \, d\mathbf{r}_{kj} + \sum_k \int \sum_i \frac{1}{2} e_{ki} (\mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}}) \overline{\dot{\mathbf{r}}}_{ki} g_k \Pi \, d\mathbf{r}_{kj}.$$

Finally the curl of the magnetization \mathbf{M} (cf. (23)) is found to be

$$\begin{aligned} \nabla_{\mathbf{R}} \wedge \mathbf{M} &= \sum_k \int \sum_i \frac{1}{2} c^{-1} e_{ki} \mathbf{r}_{ki} \nabla_{\mathbf{R}} \cdot \overline{\dot{\mathbf{r}}}_{ki} g_k \Pi \, d\mathbf{r}_{kj} - \\ &- \sum_k \int \sum_i \frac{1}{2} c^{-1} e_{ki} (\mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}}) \overline{\dot{\mathbf{r}}}_{ki} g_k \Pi \, d\mathbf{r}_{kj}. \quad (31) \end{aligned}$$

These expressions will appear as a consequence of the averaging process of the Lorentz equations to be performed in § 6.

§ 5. The microscopic field quantities and the equations of Lorentz.

The microscopic electric and magnetic field strengths will as is customary be indicated by small symbols \mathbf{e} and \mathbf{b} respectively

$$\mathbf{e} = \mathbf{e}(\mathbf{R}_k^N, \mathbf{r}_{ki}^{NZk}, \mathbf{p}_k^N, \mathbf{p}_{ki}^{NZk}; \mathbf{R}, t) \quad (32)$$

$$\mathbf{b} = \mathbf{b}(\mathbf{R}_k^N, \mathbf{r}_{ki}^{NZk}, \mathbf{p}_k^N, \mathbf{p}_{ki}^{NZk}; \mathbf{R}, t), \quad (33)$$

i.e., these field strengths depend on the state of the system and also explicitly on space and time. Thus \mathbf{e} and \mathbf{b} are dynamical quantities of the type (4) to which the various averaging procedures discussed

*) cf. appendix II.

in § 2, and further in particular formula (11) for the time derivative of an average, can be applied.

The Lorentz field equations can now be written in the following form (in rationalized gaussian units)

$$\nabla_{\mathbf{R}} \cdot \mathbf{b} = 0, \quad (34)$$

$$\frac{d\mathbf{b}}{dt} = -c \nabla_{\mathbf{R}} \wedge \mathbf{e}, \quad (35)$$

$$\nabla_{\mathbf{R}} \cdot \mathbf{e} = \sum_k e_k \delta(\mathbf{R}_k - \mathbf{R}) + \sum_{k,i} e_{ki} \delta(\mathbf{R}_k + \mathbf{r}_{ki} - \mathbf{R}), \quad (36)$$

$$\frac{d\mathbf{e}}{dt} = c \nabla_{\mathbf{R}} \wedge \mathbf{b} - \sum_k e_k \dot{\mathbf{R}}_k \delta(\mathbf{R}_k - \mathbf{R}) - \sum_{k,i} e_{ki} (\dot{\mathbf{R}}_k + \dot{\mathbf{r}}_{ki}) \delta(\mathbf{R}_k + \mathbf{r}_{ki} - \mathbf{R}). \quad (37)$$

With respect to the time derivatives occurring in (35) and (37) one should remember that the field quantities \mathbf{e} and \mathbf{b} do not only depend explicitly on time (because of the presence of time-dependent external fields), but that they also vary with time owing to the fact that the state of the system varies. Therefore $d\mathbf{b}/dt$ and $d\mathbf{e}/dt$ should be evaluated as total time derivatives in phase space (cf. (12)). Furthermore the microscopic charge and current densities in equations (36) and (37), which usually are treated as continuous functions of space and time, have been represented here consistently, with the help of Dirac's δ -function, as due to point charges and the motions thereof.

§ 6. *Derivation of the Maxwell equations.* In this section we will apply the actual ensemble averaging process defined by (5) to the microscopic equations (34)–(37) in succession.

For (34) and (35) this method yields simply, when (5) and (11) are used, two of the Maxwell equations

$$\nabla_{\mathbf{R}} \cdot \mathbf{B} = 0, \quad (38)$$

and

$$\partial \mathbf{B} / \partial t = -c \nabla_{\mathbf{R}} \wedge \mathbf{E}, \quad (39)$$

where

$$\mathbf{B} = \langle \mathbf{b} \rangle, \quad (40)$$

and

$$\mathbf{E} = \langle \mathbf{e} \rangle \quad (41),$$

are the macroscopic magnetic and electric field strengths.

For (36) and (37) the derivation is somewhat more complicated.

Taking the ensemble average of both members of (36) we obtain, using the definitions (41) and (6),

$$\begin{aligned} \nabla_{\mathbf{R}} \cdot \mathbf{E} &= \langle \sum_k e_k \delta(\mathbf{R}_k - \mathbf{R}) \rangle + \langle \sum_{k,i} e_{ki} \delta(\mathbf{R}_k + \mathbf{r}_{ki} - \mathbf{R}) \rangle = \\ &= \sum_k \int \{ e_k g_k(\mathbf{R}, \mathbf{r}_{kj}; t) + \sum_i e_{ki} g_k(\mathbf{R} - \mathbf{r}_{ki}, \mathbf{r}_{kj}; t) \} \Pi \, d\mathbf{r}_{kj}. \end{aligned} \quad (42)$$

In this expression the first arguments of the function g_k arise as a consequence of the integration with respect to \mathbf{R}_k , taking into account the δ -functions occurring in the integrand.

Expanding the function $g_k(\mathbf{R} - \mathbf{r}_{ki}, \mathbf{r}_{kj}; t)$ in a Taylor-series with respect to its first argument around the point \mathbf{R} , and limiting ourselves to the first three terms of this expansion we find, on substitution into (42)

$$\begin{aligned} \nabla_{\mathbf{R}} \cdot \mathbf{E} &= \sum_k \int (e_k + \sum_i e_{ki}) g_k \Pi \, d\mathbf{r}_{kj} - \sum_k \int \sum_i e_{ki} \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} g_k \Pi \, d\mathbf{r}_{kj} + \\ &\quad + \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \mathbf{r}_{ki} : \nabla_{\mathbf{R}} \nabla_{\mathbf{R}} g_k \Pi \, d\mathbf{r}_{kj}. \end{aligned} \quad (43)$$

This immediately reduces, as can be seen from the previously derived formulae (15), (24) and (28), to

$$\nabla_{\mathbf{R}} \cdot \mathbf{E} = \rho - \nabla_{\mathbf{R}} \cdot (\mathbf{P} - \nabla_{\mathbf{R}} \cdot \mathbf{Q}). \quad (44)$$

Introducing the electric displacement vector

$$\mathbf{D} = \mathbf{E} + \mathbf{P} - \nabla_{\mathbf{R}} \cdot \mathbf{Q}, \quad (45)$$

we may write (44) in the Maxwellian form

$$\nabla_{\mathbf{R}} \cdot \mathbf{D} = \rho \quad (46)$$

As has been remarked already by Rosenfeld⁵⁾ the definition (45) of \mathbf{D} differs from the usual one by the inclusion of the quadrupole term which is ordinarily neglected. As Rosenfeld pointed out, however, it is consistent to retain the electric quadrupole term when in the fourth Maxwell equation, to be considered below, the magnetization, which is of the same order, is taken into account. Of course the use of a Taylor-expansion makes only sense when g_k does not change appreciably over distances comparable to the atomic dimensions, i.e. to the mean values of \mathbf{r}_{ki} . In other words all macroscopic properties should be slowly varying functions of space.

Let us now proceed with the last microscopic equation (37). Averaging again both members, we obtain with (7), (40) and (41)

$$\begin{aligned} \partial \mathbf{E} / \partial t &= c \nabla_{\mathbf{R}} \wedge \mathbf{B} - \\ &- \sum \int \{ e_k \overline{\mathbf{R}}_k g_k(\mathbf{R}, \mathbf{r}_{kj}; t) + \sum_i e_{ki} (\overline{\mathbf{R}}_k + \overline{\mathbf{r}}_{ki}) g_k(\mathbf{R} - \mathbf{r}_{ki}, \mathbf{r}_{kj}; t) \} \Pi \, d\mathbf{r}_{kj}. \end{aligned} \quad (47)$$

One should note that the quantities $\overline{\mathbf{R}}_k$ and $\overline{\mathbf{r}}_{ki}$ in (47) are functions of the same arguments as the functions g_k by which they are multiplied. Therefore $\overline{\mathbf{R}}_k$ in the first term has not the same value as $\overline{\mathbf{R}}_k$ in the second one.

Expanding $(\overline{\mathbf{R}}_k + \overline{\mathbf{r}}_{ki}) g_k$ in a Taylor-series again, and neglecting, as we did before (cf. (43)), all terms of degree higher than the second in \mathbf{r}_{ki} and $\dot{\mathbf{r}}_{ki}$, we get

$$\begin{aligned} \partial \mathbf{E} / \partial t = & c \nabla_{\mathbf{R}} \wedge \mathbf{B} - \sum_k \int (e_k + \sum_i e_{ki}) \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj} - \sum_k \int \sum_i e_{ki} \overline{\mathbf{r}}_{ki} g_k \Pi \, d\mathbf{r}_{kj} + \\ & + \sum_k \int \sum_i e_{ki} (\mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}}) \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj} + \sum_k \int \sum_i e_{ki} (\mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}}) \overline{\mathbf{r}}_{ki} g_k \Pi \, d\mathbf{r}_{kj} - \\ & - \sum_k \int \sum_i \frac{1}{2} e_{ki} (\mathbf{r}_{ki} \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} \nabla_{\mathbf{R}}) \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj}. \quad (48) \end{aligned}$$

Introducing into this result the previously obtained formulae (17), (25), (30) and (31), we can write alternatively (after some straightforward transformations)

$$\begin{aligned} \partial \mathbf{E} / \partial t = & c \nabla_{\mathbf{R}} \wedge \mathbf{B} - \mathbf{i} - \partial(\mathbf{P} - \nabla_{\mathbf{R}} \cdot \mathbf{Q}) / \partial t - c \nabla_{\mathbf{R}} \wedge \mathbf{M} - \\ & - \sum_k \int \nabla_{\mathbf{R}} \wedge (\sum_i e_{ki} \mathbf{r}_{ki} \wedge \overline{\mathbf{R}}_{ki} g_k) \Pi \, d\mathbf{r}_{kj} - \\ & - \sum_k \int \sum_i \frac{1}{2} e_{ki} (\mathbf{r}_{ki} \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} \nabla_{\mathbf{R}}) \overline{\mathbf{R}}_k g_k \Pi \, d\mathbf{r}_{kj} + \\ & + \sum_k \int \sum_i \frac{1}{2} e_{ki} \mathbf{r}_{ki} \{ \mathbf{r}_{ki} \cdot \nabla_{\mathbf{R}} (\nabla_{\mathbf{R}} \cdot \overline{\mathbf{R}}_k g_k) \} \Pi \, d\mathbf{r}_{kj}. \quad (49) \end{aligned}$$

The sum of the last two terms in (49) may be rearranged with the help of the vector formula (see appendix II)

$$\mathbf{a} \{ \mathbf{a} \cdot \nabla (\nabla \cdot \mathbf{b}) \} - (\mathbf{a} \mathbf{a} \cdot \nabla \nabla) \mathbf{b} = \nabla \wedge \{ \nabla \cdot \mathbf{a} (\mathbf{a} \wedge \mathbf{b}) \} \quad (50)$$

valid for any two vectors \mathbf{a} and \mathbf{b} , when \mathbf{a} is a constant with respect to the differentiation involved in ∇ .

We then obtain the final result

$$\begin{aligned} \partial \mathbf{E} / \partial t = & c \nabla_{\mathbf{R}} \wedge \mathbf{B} - \mathbf{i} - \partial(\mathbf{P} - \nabla_{\mathbf{R}} \cdot \mathbf{Q}) / \partial t - c \nabla_{\mathbf{R}} \wedge \mathbf{M} - \\ & - \nabla_{\mathbf{R}} \wedge [\sum_k \int \sum_i e_{ki} \{ (\mathbf{r}_{ki} \wedge \overline{\mathbf{R}}_k g_k) - \frac{1}{2} \nabla_{\mathbf{R}} \cdot \mathbf{r}_{ki} (\mathbf{r}_{ki} \wedge \overline{\mathbf{R}}_k g_k) \} \Pi \, d\mathbf{r}_{kj}]. \quad (51) \end{aligned}$$

Taking into account the definition (45) of the electric displacement \mathbf{D} , and introducing the magnetic displacement \mathbf{H} (which is commonly called the magnetic field strength) by

$$\mathbf{H} = \mathbf{B} - \mathbf{M} - c^{-1} \sum_k \int \sum_i e_{ki} \{ (\mathbf{r}_{ki} \wedge \overline{\mathbf{R}}_k g_k) - \frac{1}{2} \nabla_{\mathbf{R}} \cdot \mathbf{r}_{ki} (\mathbf{r}_{ki} \wedge \overline{\mathbf{R}}_k g_k) \} \Pi \, d\mathbf{r}_{kj}, \quad (52)$$

the last equation (51) reduces to the common Maxwellian form

$$\partial \mathbf{D} / \partial t = c \nabla_{\mathbf{R}} \wedge \mathbf{H} - \mathbf{i}. \quad (53)$$

It is very important to remark that again our definition of \mathbf{H} (52) differs from the customary one. Here, however, this is not due, as previously with respect to \mathbf{D} (cf. (43)), to the inclusion of higher order multipole moments, but to the fact that all possible motions of the particles are taken into consideration in a consistent way. In the usual space-time averaging treatment only two special cases are considered i.e. systems at rest and systems moving with uniform velocity. It should be stressed, however, that in the first case, though the electric current is supposed to be due to moving particles (ions or electrons), the particles contributing to the electric or magnetic polarization are supposed to be at rest. This of course is inconsistent as the ions contributing to the true current density may well have dipole and higher moments. A similar objection can, *mutatis mutandis*, be made with regard to the case of uniformly moving bodies. It will be remembered (cf. note p. 972) that conduction electrons are to be treated formally as nuclei. As a consequence, for certain values of the index k , e_k and \mathbf{R}_k simply refer to the charge and the position of a free electron. For these same values of k , therefore, no bound electrons with charges e_{ki} and positions \mathbf{r}_{ki} exist. Thus the free electrons may well contribute to the mean electric charge and current densities (15) and (17), but not to the higher moment densities (21), (22), (23).

In a few particular cases our general definition (52) reduces to a more familiar form.

a. If in a certain medium the condition

$$\overline{\mathbf{R}_k} \equiv 0 \quad (54)$$

holds for all k not referring to conduction electrons, it will be seen immediately from (52), that then

$$\mathbf{H} = \mathbf{B} - \mathbf{M}. \quad (55)$$

We thus have specified clearly the condition under which the relation (55) between \mathbf{H} and \mathbf{B} , together with the definition (23) of \mathbf{M} , holds. This relation is usually (but not quite correctly) stated as being valid for any system at rest, including systems with ionic conduction. Of course, on changing the definition of \mathbf{M} in an appropriate manner, one could always justify the relation (55).

b. Let us now suppose that, for all k not referring to conduction electrons,

$$\overline{\mathbf{R}}_k = \mathbf{v}, \quad (56)$$

where the common mean velocity \mathbf{v} is constant with respect to time and space. It can then be shown very simply (cf. (21), (27) and appendix II) that (52) reduces to

$$\mathbf{H} = \mathbf{B} - \mathbf{M} - c^{-1} (\mathbf{P} - \nabla_{\mathbf{R}} \cdot \mathbf{Q}) \wedge \mathbf{v}. \quad (57)$$

This relation is usually derived^{1) 2)} with the definitions (21), (22) and (23) for the moment densities, for "systems moving with uniform velocity". Again our condition (56) gives a more exact qualification of such systems than the usual one. Here again it is true, that by changing the definition of \mathbf{M} , (57) could be made to hold in a more general case.

With the help of the familiar formula (57), where it is seen that as a consequence of motion the polarization contributes to the difference of \mathbf{B} and \mathbf{H} (that is gives rise to an effective magnetization), our general formula (52) can be easily understood. Indeed, (52) simply expresses the fact that every electrically polarized particle (atom or ion) with a non-vanishing mean velocity gives its own contribution to the effective magnetization. We finally note that relation (45) between \mathbf{D} and \mathbf{E} was not affected by any motion of the particles. This is in agreement with earlier derivations of the Maxwell equations from electron theory. As is well-known from a consistent relativistic treatment, first given by D ä l l e n b a c h^{11) 12)}, also moving magnetic dipoles contribute to the effective electric polarization, in agreement with the macroscopic Minkowski theory^{2) 13)}.

§ 7. *Concluding remarks.* After having accomplished formally the derivation of the Maxwell equations and having defined the macroscopic field quantities in terms of ensemble averages, it is perhaps useful to add a few remarks which may help to clarify the relation between the earlier Lorentz treatment and the present one.

In the first place we may repeat here the observation made by I r v i n g and K i r k w o o d¹⁰⁾ in an analogous case, that the densities and average field quantities defined previously are point functions which already satisfy the macroscopic Maxwell equations. According to the principles of statistical mechanics these ensemble averages correspond, speaking somewhat loosely, to time averages.

In order to obtain the macroscopic observables an additional space averaging process would be necessary. This second averaging process, however, does not modify anymore the form of the equations. On the other hand, in the Lorentz type of treatment, though various authors indicate that space as well as time averages must be taken in order to obtain the macroscopic quantities, only the first one is actually performed, the second one again not modifying the form of the equations. As V a n V l e c k ³⁾ pointed out already, the smoothing out over infinitesimal space regions may be interpreted in a way as corresponding to a smoothing out in time. One could also say that space averaging corresponds to a peculiar kind of ensemble averaging, with a distribution function such that all configurations taken into account are related to each other by a rigid displacement of the system as a whole. From the foregoing it will be evident that the present method introducing an arbitrary distribution function in phase space, apart from allowing a mathematically more rigorous treatment, enables one to perform the averaging process in a more fundamental way.

It has been stated already in the introduction that the method applied in this paper is based on classical statistical mechanics and on a classical model of the atom. Actually, however, the motion of electrons bound in the atoms should be treated on the basis of quantum mechanics. The same thing applies to the motion of conduction electrons in a metal, and at low temperatures even to the motion of the atoms themselves. Therefore it would certainly be gratifying if the present formalism could be extended to the case where the particles are treated quantummechanically. At present we will not attempt an exhaustive discussion of this problem. It may be remarked already here, however, that — as is well-known — the role played by the probability distribution function f in classical statistical mechanics, is, in quantum statistics, taken over by the density matrix P first introduced by V o n N e u m a n n. It will be shown in appendix I that the theorem (11), which enables one to evaluate the time derivative of the average of a dynamical variable and which plays a fundamental role in the present formalism, remains valid also in quantum statistics. Hence it will be clear that the derivation of three of the Maxwell equations (namely (38), (39) and (46)) remains formally unchanged. The quantum statistical derivation of the fourth Maxwell equation (53), however, requires a closer investi-

gation; we intend to come back to this point in a subsequent publication.

As has been mentioned before the method developed in this work seems to be suitable for the discussion of other problems concerned with matter in an electromagnetic field. One might think for instance of transport phenomena in electromagnetic fields, including a treatment of ponderomotive forces, further of the theory of the dielectric constant and its extension to dispersion phenomena. We plan to deal with some of these problems in another communication.

APPENDIX I

A quantum statistical theorem. We shall prove here briefly the quantum statistical analogue of the classical theorem (11), which permits to evaluate the time derivative of the ensemble average of a dynamical quantity.

The density matrix for a system of N particles with coordinates \mathbf{R}_k can be defined by the ensemble average ¹⁴⁾

$$P(\mathbf{R}^N, \mathbf{R}'^N) = \{\psi^*(\mathbf{R}^N, t) \psi(\mathbf{R}'^N, t)\}_{Av},$$

where $\psi(\mathbf{R}^N, t)$ is the wave function depending on all particle coordinates $\mathbf{R}_1, \dots, \mathbf{R}_N$ and on the time.

The average of an operator $a(\mathbf{R}^N, \mathbf{p}^N; \mathbf{R}, t)$ is now expressed as the trace of the product of two matrices:

$$a_{Av} = \text{Tr} \{aP\}$$

Because of the well-known relation ¹⁴⁾

$$\frac{\partial P}{\partial t} = -\frac{i}{\hbar} (HP - PH),$$

where H is the Hamiltonian, we have

$$\begin{aligned} \frac{\partial a_{Av}}{\partial t} &= \text{Tr} \left\{ \frac{\partial a}{\partial t} P \right\} + \text{Tr} \left\{ a \frac{\partial P}{\partial t} \right\} = \text{Tr} \left\{ \frac{\partial a}{\partial t} P \right\} - \frac{i}{\hbar} \text{Tr} \{a(HP - PH)\} \\ &= \text{Tr} \left\{ \frac{\partial a}{\partial t} P \right\} + \frac{i}{\hbar} \text{Tr} \{P(Ha - aH)\} = \text{Tr} \left\{ \frac{da}{dt} P \right\}, \end{aligned}$$

which is formally identical with the classical result (11).

APPENDIX II

On tensor notation. The system of vector and tensor notation used in the present paper is the same as that explained in the appendix of Rosenfeld's book ⁵⁾, which in its turn is essentially that of Milne and Chapman ¹⁵⁾.

Thus, the exterior product of an ordered pair of vectors **a**, **b** is a tensor

$$\mathbf{T} = \mathbf{ab} \quad T_{ik} = a_i b_k.$$

The scalar product of two vectors **a**, **b** is written as

$$\mathbf{a} \cdot \mathbf{b} = \sum_i a_i b_i,$$

while the vector product is denoted by

$$\mathbf{a} \wedge \mathbf{b} \rightarrow (\mathbf{a} \wedge \mathbf{b})_i = \sum_{k,l} \epsilon_{ikl} a_k b_l,$$

where ϵ_{ikl} represents the well-known isotropic tensor of the third order discussed among others by Jeffreys and Jeffreys ¹⁶⁾, for which the relation $\sum_l \epsilon_{ikl} \epsilon_{\mu\nu l} = \delta_{i\mu} \delta_{k\nu} - \delta_{i\nu} \delta_{k\mu}$ holds. The divergences of a vector and a tensor are written respectively as

$$\text{div } \mathbf{a} = \nabla \cdot \mathbf{a} = \sum_i \frac{\partial a_i}{\partial x_i}$$

and

$$\text{Div } \mathbf{T} = \nabla \cdot \mathbf{T} \rightarrow (\text{Div } \mathbf{T})_k = \sum_i \frac{\partial T_{ik}}{\partial x_i}.$$

Further

$$\mathbf{S} \cdot \mathbf{T} = \sum_{i,k} S_{ik} T_{ki},$$

so that an expression like that occurring in eq. (28) has the meaning

$$\mathbf{ab} \cdot \nabla \nabla g = \sum_{i,k} a_i b_k \frac{\partial^2 g}{\partial x_k \partial x_i}$$

Formula (50) in the text can now be proved in the following way: The *i*-th component of the right-hand member is

$$\begin{aligned} & \sum_{k,l,m,\mu,\nu} \epsilon_{ikl} \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_m} a_m \epsilon_{i\mu\nu} a_\mu b_\nu = \\ & \sum_{k,m,\mu,\nu} \left(\sum_l \epsilon_{ikl} \epsilon_{\mu\nu l} \right) \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_m} a_m a_\mu b_\nu = \\ & \sum_{k,m,\mu,\nu} (\delta_{i\mu} \delta_{k\nu} - \delta_{i\nu} \delta_{k\mu}) \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_m} a_m a_\mu b_\nu = \\ & \sum_{k,m} \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_m} (a_m a_i b_k - a_m a_k b_i) \end{aligned}$$

which for constant vector \mathbf{a} is the i -th component of the left-hand member of (50).

Finally the formula ($\mathbf{b} = \text{constant}$)

$$\nabla \cdot \mathbf{a} (\mathbf{a} \wedge \mathbf{b}g) = (\nabla \cdot \mathbf{a} \mathbf{a} g) \wedge \mathbf{b},$$

applied in the transition from (52) to (57), can be proved as follows:

The i -th component of the left-hand member is

$$\sum_{k,l,m} \frac{\partial}{\partial x_k} a_k \epsilon_{ilm} a_l b_m g = \sum_{l,m} \epsilon_{ilm} \sum_k \frac{\partial}{\partial x_k} a_k a_l g b_m$$

and this is, for constant \mathbf{b} , the i -th component of $(\nabla \cdot \mathbf{a} \mathbf{a} g) \wedge \mathbf{b}$.

Received 23-5-53.

REFERENCES

- 1) Lorentz, H. A., Proc. Amsterdam (1902) 254; Enc. Math. Wiss., V 2, Heft 1, S. 200.
- 2) Fokker, A. D., Relativiteitstheorie, Groningen 1929. Becker, R., Theorie der Elektrizität, Band II, Leipzig 1933.
- 3) Van Vleck, J. H., The theory of electric and magnetic susceptibilities, Oxford, 1931.
- 4) Sauter, F., Z. Phys., **126** (1949) 207.
- 5) Rosenfeld, L., Theory of electrons, Amsterdam, 1951.
- 6) Yvon, J., Actualités Scientifiques et Industrielles, nos. 542 et 543, Paris, 1937.
- 7) Kirkwood, J. G., J. chem. Phys. **4** (1936) 592.
- 8) Brown, W. F., J. chem. Phys. **18** (1950) 1193.
- 9) De Boer, J., Van der Maesen, F. and Ten Seldam, C. A., Physica **19** (1953) 265.
- 10) Irving, J. H. and Kirkwood, J. G., J. chem. Phys. **18** (1950) 817.
- 11) Dällenbach, W., Ann. Physik **58** (1919) 523.
- 12) Pauli, W., Relativitätstheorie, Leipzig 1921; Enc. Math. Wiss. V 2, Heft 4, S. 539.
- 13) Sommerfeld, A., Vorlesungen über theoretische Physik, Bd III, Wiesbaden, 1948.
- 14) Cf. e.g. De Boer, J., Reports on progress in physics, **12** (1949) 305.
- 15) Cf. Chapman, S. and Cowling, T., The mathematical theory of non-uniform gases, Cambridge 1939.
- 16) Jeffreys, H. and Jeffreys, B. S., Methods of mathematical physics, Cambridge 1946.