Ann. Geophys., 35, 683–690, 2017
www.ann-geophys.net/35/683/2017/
doi:10.5194/angeo-35-683-2017

Annales Geophysicae
Open Access
EGU

Causal kinetic equation of non-equilibrium plasmas

Rudolf A. Treumann[1,2] **and Wolfgang Baumjohann**[3]

[1]Department of Geophysics and Environmental Sciences, Ludwig-Maximilians-Universität München, Munich, Germany
[2]International Space Science Institute Bern, Bern, Switzerland
[3]Space Research Institute, Austrian Academy of Sciences, Graz, Austria

Correspondence to: Rudolf A. Treumann (art@geophysik.uni-muenchen.de)

Received: 20 April 2017 – Accepted: 2 May 2017 – Published: 24 May 2017

Abstract. Statistical plasma theory far from thermal equilibrium is subject to Liouville's equation, which is at the base of the BBGKY hierarchical approach to plasma kinetic theory, from which, in the absence of collisions, Vlasov's equation follows. It is also at the base of Klimontovich's approach which includes single-particle effects like spontaneous emission. All these theories have been applied to plasmas with admirable success even though they suffer from a fundamental omission in their use of the electrodynamic equations in the description of the highly dynamic interactions in many-particle conglomerations. In the following we extend this theory to taking into account that the interaction between particles separated from each other at a distance requires the transport of information. Action needs to be transported and thus, in the spirit of the direct-interaction theory as developed by Wheeler and Feynman (1945), requires time. This is done by reference to the retarded potentials. We derive the fundamental causal Liouville equation for the phase space density of a system composed of a very large number of charged particles. Applying the approach of Klimontovich (1967), we obtain the retarded time evolution equation of the one-particle distribution function in plasmas, which replaces Klimontovich's equation in cases when the direct-interaction effects have to be taken into account. This becomes important in all systems where the distance between two points $|\Delta \boldsymbol{q}| \sim ct$ is comparable to the product of observation time and light velocity, a situation which is typical in cosmic physics and astrophysics.

Keywords. Space plasma physics (General or miscellaneous)

1 Introduction

The starting point of (classical) kinetic theory is Liouville's equation. Written in terms of the N_a-particle Hamiltonian $\mathcal{H}_{N_a}(\boldsymbol{q},\boldsymbol{p},t)$ and defining the 6-D phase space density $\mathcal{N}_a(\boldsymbol{q},\boldsymbol{p},t)$ of species a, both functions of space $\boldsymbol{q}$ and momentum $\boldsymbol{p}$, it becomes

$$\dot{\mathcal{N}}_a \equiv \partial_t \mathcal{N}_a + \left[\mathcal{H}_{N_a}, \mathcal{N}_a\right] = 0, \tag{1}$$

where it is assumed that the particle number N_a of species a is conserved (along all dynamical phase space orbits). Otherwise the right-hand side would contain the difference of number sources and losses $\mathcal{S}_a - \mathcal{L}_a$. This equation, under the assumptions made, is completely general, applying to any system consisting of $N_a \gg 1$ particles in interaction with an external as well as with their mutual fields, of which they function as sources. These fields are contained in the Hamiltonian and act via the Poisson bracket [...].

In view of an application to plasmas, the relevant field is the electromagnetic field $\boldsymbol{E},\boldsymbol{B}$, with the particles carrying electric charges $e_a = \mp e \equiv ae$ (with $a = -, +$) being the sources of the field. For simplicity, in the following, we restrict ourselves to electrons and ions (protons) of mass m_a, and gravity can be neglected on all scales small enough for the electromagnetic fields to dominate. We also assume global quasi-neutrality and the absence of any external fields. Then $\boldsymbol{E},\boldsymbol{B} = (\boldsymbol{E},\boldsymbol{B})^m$ is the set of *microscopic* electromagnetic fields produced solely by the microscopic charge and current densities of the interacting particle components a, which serve as their sources:

AnGeo Communicates

Published by Copernicus Publications on behalf of the European Geosciences Union.

$$
\begin{aligned}
\rho_a^m(\boldsymbol{q},t) &= e_a \sum_{i=1}^{N_a} \delta\big(\boldsymbol{q}-\boldsymbol{q}_{ai}(t)\big) = e_a \int d^3p\ \mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t),\\
\boldsymbol{j}_a^m(\boldsymbol{q},t) &= \frac{e_a}{m_a}\sum_{i=1}^{N_a}\ \boldsymbol{p}_{ai}(t)\ \delta\big(\boldsymbol{q}-\boldsymbol{q}_{ai}(t)\big)\\
&= \frac{e_a}{m_a}\int d^3p\ \boldsymbol{p}\ \mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t),
\end{aligned}
\tag{2}
$$

where the exact 6-D phase space density is defined through

$$
\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t) = \sum_{i=1}^{N_a}\delta\big(\boldsymbol{p}-\boldsymbol{p}_{ai}(t)\big)\ \delta\big(\boldsymbol{q}-\boldsymbol{q}_{ai}(t)\big)
\tag{3}
$$

and $\boldsymbol{q}_{ai}(t)$, $\boldsymbol{p}_{ai}(t)$ are the spatial and momentum phase space trajectories which the particle ai performs in the phase space under the action of the complete microscopic electromagnetic field $(\boldsymbol{E},\boldsymbol{B})^m$, which it feels at its location $\boldsymbol{q}=\boldsymbol{q}_{ai}(t)$ at time t. Liouville's equation for the exact phase space density can then be written in the form

$$
\begin{aligned}
\frac{\partial\mathcal{N}_a}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}}\ \mathcal{N}_a + e_a\bigg\{\Big[\boldsymbol{E}^m(\boldsymbol{q},t)\\
+\frac{\boldsymbol{p}}{m_a}\wedge\boldsymbol{B}^m(\boldsymbol{q},t)\Big]\cdot\frac{\partial\mathcal{N}_a}{\partial\boldsymbol{p}}\bigg\} = 0.
\end{aligned}
\tag{4}
$$

This is Klimontovich's equation for the exact microscopic phase space density $\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)$ in 6-D phase space (Klimontovich, 1967). It is a tautology because it does not say anything other than that particle number is conserved along all the dynamical orbits of the particles in phase space under the action of their mutual electromagnetic fields. The microscopic fields it contains are given by Maxwell's equations in differential form:

$$
\begin{aligned}
&\nabla_q\wedge\boldsymbol{B}^m = \mu_0\boldsymbol{j}^m + \mu_0\epsilon_0\partial_t\boldsymbol{E}^m, \quad \nabla_q\cdot\boldsymbol{B}^m = 0,\\
&\nabla_q\wedge\boldsymbol{E}^m = -\partial_t\boldsymbol{B}^m, \quad \nabla_q\cdot\boldsymbol{E}^m = \frac{1}{\epsilon_0}\sum_a\rho_a^m.
\end{aligned}
\tag{5}
$$

The solution of this set of equations is not possible as it requires knowledge of all microscopic particle orbits. One can, however introduce some coarse graining procedure and define integrated distribution functions which ultimately reduce the system to a set of equations known as Klimontovich–Vlasov equations for a one-particle phase space distribution in the presence of the average electromagnetic fields. This procedure is very efficient, and we will follow it below in a modified version.

2 Effect of retardation

The problem of the above equations is that they do not account for the fact that the electromagnetic signal of the presence and motion of the particles is transferred from the signal-emitting particles to the signal-receiving particles under consideration, i.e. the absorbers and reactors. Their sources are the charge and current densities $\rho_a^m(\boldsymbol{q},t)$, $\boldsymbol{j}^m(\boldsymbol{q},t)$, which are assumed to be known at any instant t in all space points $\boldsymbol{q}$. Obtaining this knowledge is impossible as it requires instantaneous measurements at time t of all positions $\boldsymbol{q}$ and momenta $\boldsymbol{p}$ of the particles present in real space. Instead, the information must be synchronized among all locations. This is taken care of in the Liénard–Wiechert potentials, which explicitly account for the transport of information from point $\boldsymbol{q}'$ to point $\boldsymbol{q}$. In this case in the Lorentz gauge

$$
\begin{aligned}
\boldsymbol{E}^m &= -\nabla_{\boldsymbol{q}}\phi^m(\boldsymbol{q},t) - \partial_t\boldsymbol{A}^m(\boldsymbol{q},t),\\
\boldsymbol{B}^m &= \nabla_{\boldsymbol{q}}\wedge\boldsymbol{A}^m(\boldsymbol{q},t),
\end{aligned}
\tag{6}
$$

the correct scalar and vector potentials are to be expressed by the *retarded* charge and current densities

$$
\begin{aligned}
\phi^m(\boldsymbol{q},t) &= \frac{1}{4\pi\epsilon_0}\sum_a\int d^3q'\ \frac{\rho_a^m(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|},\\
\boldsymbol{A}^m(\boldsymbol{q},t) &= \frac{\mu_0}{4\pi}\sum_a\int d^3q'\ \frac{\boldsymbol{j}_a^m(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|}
\end{aligned}
\tag{7}
$$

taken at the retarded time

$$
t' = t - |\boldsymbol{q}-\boldsymbol{q}'|/c, \quad \boldsymbol{q}\neq\boldsymbol{q}'
\tag{8}
$$

of arrival of *all the signals* emitted at t' from all the particles at spatial distance $|\boldsymbol{q}-\boldsymbol{q}'|$ from the location of particle a_i at $\boldsymbol{q}$ and at time t. This also implies that in the expressions for the charge and current densities $\mathcal{N}_a\to\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')$ is a function of the retarded time t'.

Since all particles serve both as field sources and actors, excluding their self-interaction, the use of the instantaneous fields ignores the time-consuming signal transport and thus cannot be correct. It is an approximation only that holds for comparably small volumes such that, in the expression for the retarded time, the spatial difference can be neglected. Thus, the restriction on the distance between particles is that

$$
|\Delta q| \ll c\Delta t.
\tag{9}
$$

Clearly, this condition will readily be violated in large volumes of cosmic and astrophysical size, where one must refer to the above precise potentials and the fields resulting from them in reference to the Lorentz gauge.

For single-particle–particle interactions, this problem has been discussed in depth in seminal papers by Wheeler and Feynman (1945, 1949). They showed that in a closed system where no information is lost to the outside eliminating any self-interaction of a particle with its proper electromagnetic field implies that the fields are properly described via *retarded* potentials only as done above. These potentials account for the emission of a signal by one particle and the absorption of the signal after some travel time by the target particle, causing this particle to interact. The emitted signals belong to advanced potentials which, when correctly included, subtract out, thereby restoring the required real-world causality. It is incorrect to assume that the information arrives microscopically instantaneously at the target, causing this to act. The electromagnetic fields following from the Lorentz gauge in the microscopic domain are

$$\begin{aligned} \boldsymbol{E}^m(\boldsymbol{q},t) =& \frac{1}{4\pi\epsilon_0}\int d^3q'\Bigg[\Bigg(\frac{\rho^m(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} \\ &+\frac{\partial_t\rho^m(\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg)(\boldsymbol{q}-\boldsymbol{q}')-\frac{\partial_t \boldsymbol{j}^m(\boldsymbol{q}',t')}{c^2|\boldsymbol{q}-\boldsymbol{q}'|}\Bigg], \\ \boldsymbol{B}^m(\boldsymbol{q},t) =& \frac{\mu_0}{4\pi}\int d^3q'\Bigg[\frac{\boldsymbol{j}^m(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3}+\frac{\partial_t \boldsymbol{j}^m(\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg] \\ &\wedge(\boldsymbol{q}-\boldsymbol{q}'), \end{aligned} \tag{10}$$

which were first given, independently, by Panofsky and Phillips (1962) and Jefimenko (1966). One should note that in these expressions the charge ρ^m and current densities $\boldsymbol{j}^m$ are summed over all particle species a.

This explicit representation of the microscopic fields accounts properly for the time delay between the signal emitted from the total compound of primed particles to arrive at the location $\boldsymbol{q}$ of the particle under consideration. Since the microscopic charge and current densities are functionals of the phase space density, these expressions contain the latter, though in a more involved manner than when using the differential forms of the electrodynamic equations, which do not show where and whether the retardation of the signal is taken into account. It is clear from these expressions that particles which are far away from the target do not affect it. The main effect will always come from close neighbours.

3 Retarded charge and current densities

Taking the divergence of the microscopic electric field and the curl of the microscopic magnetic field, one readily reads the correct microscopic charge and current densities when comparing the expressions with the microscopic Maxwell equations:

$$\begin{aligned} \rho^m(\boldsymbol{q},t) =& \frac{1}{4\pi}\sum_a \nabla_q\cdot\int d^3q'\Bigg[\Bigg(\frac{\rho^m_a(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} \\ &+\frac{\partial_t\rho^m_a(\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg)(\boldsymbol{q}-\boldsymbol{q}')-\frac{\partial_t \boldsymbol{j}^m_a(\boldsymbol{q}',t')}{c^2|\boldsymbol{q}-\boldsymbol{q}'|}\Bigg], \end{aligned} \tag{11}$$

$$\begin{aligned} \boldsymbol{j}^m(\boldsymbol{q},t) =& \frac{1}{4\pi}\sum_a\Bigg\{\nabla_q\wedge\int d^3q'\Bigg[\frac{\boldsymbol{j}^m_a(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3}+\frac{\partial_t \boldsymbol{j}^m_a(\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg] \\ &\wedge(\boldsymbol{q}-\boldsymbol{q}')-\partial_t\int d^3q'\Bigg[\Bigg(\frac{\rho^m_a(\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} \\ &+\frac{\partial_t\rho^m_a(\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg)(\boldsymbol{q}-\boldsymbol{q}')-\frac{\partial_t \boldsymbol{j}^m_a(\boldsymbol{q}',t')}{c^2|\boldsymbol{q}-\boldsymbol{q}'|}\Bigg]\Bigg\}. \end{aligned} \tag{12}$$

These are the correct forms of the charge and current densities summed over species a which have to be used in Maxwell's equations in order to account for the retarded transfer of information between the particles in the plasma. These expressions are implicit for both the charge and current densities. In order to relate them to the exact microscopic phase space distribution $\mathcal{N}_a$ as defined in Eq. (3), one refers to the representations (Eq. 2) of the charge and current densities. This shows that the functional dependence of the phase space density is itself implicit. It depends on itself, taken at all the positions $\boldsymbol{q}'$ and retarded times t'.

The proper way of dealing with this problem is to focus on the microscopic picture for as long as possible. There, all the charged particles can be imagined as moving in a vacuum as long as the medium is sufficiently dilute. By progressing to a coarse-grained picture one may afterwards advance to considering a more continuous medium in which ultimately the propagation properties of the signals will become modified by the collective properties of the matter.

With these results it is convenient to express the microscopic electromagnetic fields through the microscopic phase space densities of the particle species:

$$\begin{aligned} \boldsymbol{E}^m(\boldsymbol{q},t) =& \sum_a\frac{e_a}{4\pi\epsilon_0}\int d^3p\, d^3q'\Bigg[\Bigg(\frac{\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} \\ &+\frac{\partial_t\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg)(\boldsymbol{q}-\boldsymbol{q}')-\frac{\boldsymbol{p}\ \partial_t\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{c^2|\boldsymbol{q}-\boldsymbol{q}'|}\Bigg], \end{aligned} \tag{13}$$

$$\begin{aligned} \boldsymbol{B}^m(\boldsymbol{q},t) =& \sum_a\frac{e_a\mu_0}{4\pi}\int d^3p\, d^3q'\Bigg[\frac{\boldsymbol{p}\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3}+ \\ &+\frac{\boldsymbol{p}\ \partial_t\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg]\wedge(\boldsymbol{q}-\boldsymbol{q}'). \end{aligned} \tag{14}$$

These are the expressions of the electromagnetic field which have to be used in the microscopic Liouville Eq. (4) for the microscopic N-particle phase space density. Not only do they couple the different particle species, thus leading to a coupling between their phase space distributions, they also make each microscopic distribution $\mathcal{N}_a$ a functional of the distributions taken at all different phase space locations which are causally accessible via their retarded times of signal propagation t'. Clearly, this is a substantial complication, which is introduced into kinetic theory by the requirement of causality.

It is quite inconvenient to deal with all microscopic phase space densities. We would rather have separate equations for them. This can be achieved when observing that Eq. (4) is an equation for $\mathcal{N}_a$. Thus, putting $a\to b$ in the last expressions, which means that we sum over all particle species b including also $b=a$ (with self-interaction excluded by the definition of the retarded time), we have

$$
\begin{aligned}
&\frac{\partial \mathcal{N}_a}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}}\mathcal{N}_a + \sum_b \frac{e_a e_b}{4\pi}\int d^3p'\, d^3q' \Bigg\{ \\
&\quad \frac{1}{\epsilon_0}\Bigg[\Bigg(\frac{\mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} + \frac{\partial_t \mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg)(\boldsymbol{q}-\boldsymbol{q}') \\
&\quad - \frac{\boldsymbol{p}'\ \partial_t \mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')}{c^2|\boldsymbol{q}-\boldsymbol{q}'|}\Bigg] - \frac{\mu_0}{m_a}\boldsymbol{p}\wedge\Bigg[(\boldsymbol{q}-\boldsymbol{q}')\wedge \\
&\quad \Bigg(\frac{\boldsymbol{p}'\mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|^3} + \frac{\boldsymbol{p}'\ \partial_t \mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')}{c|\boldsymbol{q}-\boldsymbol{q}'|^2}\Bigg) \\
&\quad \Bigg]\Bigg\}\cdot\frac{\partial}{\partial \boldsymbol{p}}\mathcal{N}_a\big(\boldsymbol{p},\boldsymbol{q},t\big) = 0. \qquad (15)
\end{aligned}
$$

Here $\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)$, while $\mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t')$ depends on the dummy coordinates of all particles b of integration and on the retarded time $t' = t - |\boldsymbol{q}-\boldsymbol{q}'|/c,\ \boldsymbol{q}\neq\boldsymbol{q}'$. Thus, in the q' integration all particles of sort a are also included, with the exception of the particle located at $\boldsymbol{q}=\boldsymbol{q}'$ at time t.

The above Eq. (15) is the *causal* Liouville equation acting on the microscopic N_a-particle phase space density $\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)$ in the presence of a large number of charged particles interacting via their self-consistently generated electromagnetic fields. It extends Klimontovich's equation to the correct inclusion of the retardation effect of transmission of information between the particles via electromagnetic fields.

The inclusion of information transport between the interacting particles substantially complicates the basic kinetic equation. It causes a delay in response, and thus refers to a natural measuring process in which the particles are not only generators of the electromagnetic field but also measure its effect over a causal distances accessible to them. The delay must thus necessarily cause decorrelation of the response.

There is another complication with this picture which comes into play when considering large compounds of particles rather than single particles. Single charged particles are assumed to move in the vacuum; the signal propagation between them takes place at light speed c. Immersed in the comparably dense environment of all the other charged particles, any light or radiation experiences radiation transport, which is dominated by scattering, reflection, transmission and absorption, processes that occur due to the active response of the environment to the presence of radiation and depend on the capabilities of the medium to let electromagnetic signals pass. In these processes various proper electromagnetic modes excited in the medium become involved. These are solutions of the dispersion properties of the matter. Hence, correctly accounting for the signal transport becomes rather involved. For this reason the theory even in this complex version applies to sufficiently dilute media to allow the assumption of signal propagation in a vacuum.

In the following we will proceed along the same lines as Klimontovich (1967) but will in the end refer to the above field equations. This means that in defining the average distributions, we will consider Liouville's equation without explicit reference to the fields.

4 Average distribution functions

Dealing with the causal N-particle kinetic Eq. (15) is impractical. One wants to reduce it to an equation for a one-particle distribution function in 6-D phase space for indistinguishable particles of sort a. This is done by integrating out in Eq. (3) all particle coordinates $i > 1$. Defining phase space coordinates $\boldsymbol{x} = (\boldsymbol{p},\boldsymbol{q}),\ \boldsymbol{x}_{ai}(t) = \big(\boldsymbol{p}_{ai}(t),\boldsymbol{q}_{ai}(t)\big)$, the N-particle density becomes

$$
\mathcal{N}_a(\boldsymbol{x},t) = \sum_{i=1}^{N_a}\delta\big(\boldsymbol{x}-\boldsymbol{x}_{ai}(t)\big). \qquad (16)
$$

Following Klimontovich (1967), let us define the one-particle distribution of sort a of indistinguishable particles by

$$
\begin{aligned}
f_a(\boldsymbol{x}_{a1},t) = V_a \int f_N d^6\boldsymbol{x}_{a2}\ldots d^6\boldsymbol{x}^6_{aN_a}\times \\
\times\prod_{b\neq a} d^6\boldsymbol{x}_{b1}\ldots d^6\boldsymbol{x}^6_{bN_b}. \qquad (17)
\end{aligned}
$$

The N-particle probability distribution f_N depends on all the particle coordinates in phase space which have been integrated out in the last expression, including $\boldsymbol{x}_{a1}$, and V is the spatial volume of particle $a1$, i.e. the volume all indistinguishable particles occupy. With its help the averaged phase space density yields directly

$$
\frac{N_a}{V_a} f_a(\boldsymbol{p},\boldsymbol{q},t) = \big\langle \mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)\big\rangle. \qquad (18)
$$

Here, the right-hand side is the ensemble-averaged one-particle phase space density $\langle\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)\rangle$, which is a functional only of the indistinguishable dynamics of the particles indexed by $i = 1$. Accordingly, averaging the product of two phase space densities $\mathcal{N}_a(\boldsymbol{x},t)$ and $\mathcal{N}_b(\boldsymbol{x}',t)$ yields

$$
\begin{aligned}
\big\langle\mathcal{N}_a(\boldsymbol{x},t)\mathcal{N}_b(\boldsymbol{x}',t)\big\rangle = n_a\delta_{ab}\delta(\boldsymbol{x}-\boldsymbol{x}')f_a(\boldsymbol{x},t) \\
+ n_a n_b f_{ab}(\boldsymbol{x},\boldsymbol{x}',t), \qquad (19)
\end{aligned}
$$

where the partial densities are defined as $n_a = N_a/V_a,\ n_b = N_b/V_b$ and $f_{ab}(\boldsymbol{x},\boldsymbol{x}',t)$ is the two-particle distribution function (Klimontovich, 1967). In the same way, higher-order average products of phase space densities can be reduced to sums of distribution functions.

This procedure must be applied to the causal N-particle kinetic Eq. (15). This is a formidable task if using the N-particle kinetic equation in its explicit form. As stated earlier, it is more convenient to remain with the implicit versions of the Lorentz gauge (Eq. 6) and the retarded potentials in which we replace the charge end current densities by the general expressions given in Eq. (2). From Eq. (7), this yields

$$\phi^m(\boldsymbol{q},t) = \sum_a \frac{e_a}{4\pi\epsilon_0}\int d^3p\, d^3q' \frac{\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|},$$
$$\boldsymbol{A}^m(\boldsymbol{q},t) = \sum_a \frac{e_a\mu_0}{4\pi}\int d^3p\, d^3q' \frac{\boldsymbol{p}\,\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q}',t')}{|\boldsymbol{q}-\boldsymbol{q}'|}, \qquad (20)$$

with the time taken as the retarded time t' thus depending on the spatial coordinate $\boldsymbol{q}'$, which is to be integrated out.

5 Causal one-particle kinetic equation

These expressions are to be used in the Lorentz gauge (Eq. 6) when expressing the electromagnetic fields in the N-particle kinetic Eq. (4). Formally, this is the same as if we used Eq. (15) directly in deriving the corresponding causal equation for the one-particle distribution function of indistinguishable particles of sort a. It is only the electromagnetic fields in Eq. (4) which depend on the retarded time. Therefore, one can formally calculate the average to obtain

$$\frac{\partial\langle\mathcal{N}_a\rangle}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}}\langle\mathcal{N}_a\rangle + e_a\left\langle\left[\boldsymbol{E}^m(\boldsymbol{q},t) + \frac{\boldsymbol{p}}{m_a}\wedge\boldsymbol{B}^m(\boldsymbol{q},t)\right]\cdot\frac{\partial\mathcal{N}_a}{\partial\boldsymbol{p}}\right\rangle = 0. \qquad (21)$$

The last term in this equation contains particles of kind a and b as well as the retarded time coordinate. Nevertheless, by carefully ordering the different contributions and variables of integration one can bring it into a more convenient form. For this we indicate all integration variables by primes $'$ and rename the retarded time variable by a superscript R. Then $t' \to t^R = t - |\boldsymbol{q}-\boldsymbol{q}'|/c$. After expressing the last term in angular brackets for the average phase space density, this yields :

$$\frac{\partial\langle\mathcal{N}_a\rangle}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}}\langle\mathcal{N}_a\rangle - \sum_b \frac{e_a e_b}{4\pi\epsilon_0}\int d^3p' d^3q' \left\{ \left(\nabla_{\boldsymbol{q}} + \frac{\boldsymbol{p}'}{m_b c^2}\frac{\partial}{\partial t}\right) - \frac{1}{m_b c^2}\nabla_{\boldsymbol{q}}\wedge\boldsymbol{p}' \right\} \cdot\frac{\partial}{\partial\boldsymbol{p}}\left(\frac{\langle\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)\mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t^R)\rangle}{|\boldsymbol{q}-\boldsymbol{q}'|}\right) = 0. \qquad (22)$$

In this version of the phase space (ensemble) averaged equation for the time and one-particle phase space evolution of the (ensemble) averaged one-particle phase space density $\langle\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)\rangle$, the retarded time appears only in the averaged product. This equation is the master equation for constructing the kinetic equation for the particle distribution function. Defining the fluctuation of particle number density as $\delta\mathcal{N}_a(\boldsymbol{x},t) = \mathcal{N}_a(\boldsymbol{x},t) - \langle\mathcal{N}_a(\boldsymbol{x},t)\rangle$ and referring to the correlation function $g_{ab}(\boldsymbol{x},\boldsymbol{x}',t)$ defined through the average of the product of the fluctuations (Klimontovich, 1967)

$$\langle\delta\mathcal{N}_a(\boldsymbol{x},t)\delta\mathcal{N}_b(\boldsymbol{x}',t)\rangle = n_a n_b g_{ab}(\boldsymbol{x},\boldsymbol{x}',t) + n_a\delta_{ab}\delta(\boldsymbol{x}-\boldsymbol{x}')f_a, \qquad (23)$$

we finally arrive at the desired causal kinetic equation:

$$\frac{\partial f_a}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}} f_a - \sum_b \frac{n_b e_a e_b}{4\pi\epsilon_0}\int d^3p' d^3q' \left\{ \left(\nabla_{\boldsymbol{q}} + \frac{\boldsymbol{p}'}{m_b c^2}\frac{\partial}{\partial t}\right) - \frac{1}{m_b c^2}\nabla_{\boldsymbol{q}}\wedge\boldsymbol{p}' \right\} \cdot\frac{\partial}{\partial\boldsymbol{p}}\left(\frac{f_a(\boldsymbol{p},\boldsymbol{q},t)f_b(\boldsymbol{p}',\boldsymbol{q}',t^R)}{|\boldsymbol{q}-\boldsymbol{q}'|}\right) = \mathcal{C}_a(\boldsymbol{p},\boldsymbol{q},t). \qquad (24)$$

The interaction term on the right-hand side arises from the various interparticle collisions which are mediated by the electromagnetic field. From the above definition of the fluctuations and correlations, it is given by

$$\mathcal{C}_a(\boldsymbol{p},\boldsymbol{q},t) \equiv \frac{1}{n_a}\sum_b \frac{n_b e_a e_b}{4\pi\epsilon_0}\int d^3p' d^3q' \left\{ \left(\nabla_{\boldsymbol{q}} + \frac{\boldsymbol{p}'}{m_b c^2}\frac{\partial}{\partial t}\right) - \frac{1}{m_b c^2}\nabla_{\boldsymbol{q}}\wedge\boldsymbol{p}' \right\} \cdot\frac{\partial}{\partial\boldsymbol{p}}\left(\frac{\langle\delta\mathcal{N}_a(\boldsymbol{p},\boldsymbol{q},t)\ \delta\mathcal{N}_b(\boldsymbol{p}',\boldsymbol{q}',t^R)\rangle}{|\boldsymbol{q}-\boldsymbol{q}'|}\right). \qquad (25)$$

Formally, these expressions, as claimed in the previous sections, are rather similar to those which, for the non-retarded interactions, had already been obtained by Klimontovich (1967), with the main difference being that here they are written in terms of the full electromagnetic field and contain the spatial integration over all the remote particle space. Referring to the full electromagnetic fields is necessary because of their role in the information transport and maintenance of causality in absorber theory. The expressions above are, however, very different from Klimontovich's because they account for the necessary causal relation between the interacting particles, which is contained in their dependence on the retarded time t^R, by which the particles respond to the transport of information. As a result of this response, the spatial integral appearing in these expressions contains an integration over t^R and thus also the primed space coordinate $\boldsymbol{q}'$. This complicates the calculation substantially and in an analytical treatment possibly requires the introduction of further approximations. Nevertheless, the above final equation with the implicitly given collision term extends Klimontovich's theory to the explicit reference to causality.

Referring to Eq. (22) the collision term can also be expressed via the fluctuations of the phase space density $\delta\mathcal{N}_a = \mathcal{N}_a - \langle\mathcal{N}_a\rangle$ and the fluctuations of the electromagnetic fields

$(\delta\boldsymbol{E},\delta\boldsymbol{B}) = (\boldsymbol{E},\boldsymbol{B})^m - \langle(\boldsymbol{E},\boldsymbol{B})^m\rangle$. This yields

$$\mathcal{C}_a(\boldsymbol{x},t) = -\frac{e_a}{n_a}\int d^3q' d^3p' \bigg(\frac{\partial}{\partial \boldsymbol{p}}\langle\delta\boldsymbol{E}\delta\mathcal{N}_a\rangle - \frac{\boldsymbol{p}}{m_a}\cdot\frac{\partial}{\partial \boldsymbol{p}}\wedge\langle\delta\boldsymbol{B}\delta\mathcal{N}_a\rangle\bigg),$$

where the average refers to the integration over all particle space $i>1$, and all quantities still depend on the retarded time t^R which requires integration with respect to q'. It is, however, more convenient to make use of the representation via the correlation function, in which case, from Eqs. (23) and (25) we have

$$\mathcal{C}_a(\boldsymbol{x},t) \equiv \sum_b \frac{n_b e_a e_b}{4\pi\epsilon_0}\int d^3p' d^3q' \bigg\{\bigg(\nabla_{\boldsymbol{q}} + \frac{\boldsymbol{p}'}{m_b c^2}\frac{\partial}{\partial t}\bigg) - \frac{1}{m_b c^2}\nabla_{\boldsymbol{q}}\wedge\boldsymbol{p}'\bigg\}\cdot\frac{\partial}{\partial \boldsymbol{p}}\frac{g_{ab}(\boldsymbol{x},\boldsymbol{x}',t^R)}{|\boldsymbol{q}-\boldsymbol{q}'|}. \tag{26}$$

This is the general causal collision integral term including the interactions between particles indexed by $i=1$ and $i=2$.

From all these expressions one can again obtain an equation for the fluctuation of phase space density $\delta\mathcal{N}_a$ as well as for the fluctuating fields expressed through the space charge density and current fluctuations.

With knowledge of the collision term on the right or some of its approximations, Eq. (24) provides the basis for a linearized kinetic theory to any order, including particle and time-retarded interaction effects. For this, one defines the fluctuations of the one-particle distribution function in the usual way as

$$\delta f_a = f_a - \bar{f}_a, \tag{27}$$

where $\bar{f}_a$ is the one-particle "equilibrium" distribution around which the fluctuations occur. The distribution is either an equilibrium solution of the stationary kinetic Eq. (24) or some of its large-scale solutions with scales exceeding those of the fluctuations such that the average of the fluctuation taken over these scales $\overline{\Delta f}_a = 0$ vanishes. Neglecting the collision term by putting $\mathcal{C}_a = 0$ and subtracting the fluctuation averaged kinetic equation, the causal collisionless kinetic equation for the fluctuations then becomes

$$\frac{\partial\delta f_a}{\partial t} + \frac{\boldsymbol{p}}{m_a}\cdot\nabla_{\boldsymbol{q}}\delta f_a - \sum_b \frac{n_b e_a e_b}{4\pi\epsilon_0}\int d^3p' d^3q' \bigg\{ \times\bigg(\nabla_{\boldsymbol{q}} + \frac{\boldsymbol{p}'}{m_b c^2}\frac{\partial}{\partial t}\bigg) - \frac{1}{m_b c^2}\nabla_{\boldsymbol{q}}\wedge\boldsymbol{p}'\bigg\}\cdot\frac{\partial}{\partial \boldsymbol{p}}\bigg(\times\frac{\delta f_a(\boldsymbol{x},t)\ \overline{f_b(\boldsymbol{x}',t^R)} + \overline{f_a(\boldsymbol{x},t)}\ \delta f_b(\boldsymbol{x}',t^R) - \overline{\delta f_a\delta f_b}}{|\boldsymbol{q}-\boldsymbol{q}'|}\bigg) = 0. \tag{28}$$

This equation contains the correlations of the fluctuations $\overline{\delta f_a \delta f_b}$, which are neglected in a linearized theory.

Clearly, the above equations resemble the well-known approach to plasma kinetic theory. It should, however, be pointed out that even when dropping the collision term on the right in a Klimontovich–Vlasov approach in linear theory, the retardation effect remains in the third term on the left-hand side in Eq. (24), which is the lowest-order electromagnetic field-charged particle interaction term.

6 Discussion

6.1 Remarks

The one-particle kinetic equation (Eq. 24) obtained here is fundamental to all electromagnetic plasma interactions. Since these are electromagnetic, the purely electrostatic approximation when applied must be justified separately. This is not easy because in a strictly electrostatic approach the field response is instantaneous, which contradicts electrodynamics and relativity, on which it is based. It can be upheld if the information transport occurs by electrostatic waves only but still requires some assumption about the brevity of time delay. This assumption is that the electrostatic fluctuations occur on a vastly longer timescale than the travel time of light from the remotest position of particles. Thus, one restricts oneself to sufficiently small plasma volumes in which the information transport may occur without some remarkable delay.

Under such conditions Klimontovich–Vlasov theory applies, and the complications introduced by reference to the retarded time can be neglected. On the other hand, in very large volumes like in cosmical and astrophysical applications transport of information is provided by radiation transport and becomes rather slow. Hence, remote volumes will not respond immediately and not even within light-propagation time, which can then be treated again in the simplified theory.

However, the current investigation is necessary as a clarification of two points: Firstly, that the interaction among different volumes in plasma in principle cannot be considered to occur instantaneously. Secondly, the inclusion of retarded times gives a clue to the direction of time – as briefly discussed below – which in many-particle systems has only one direction, forward. Events are delayed by information transport and thus decorrelate even though they become relativistically synchronized by accounting for the information transport. This should necessarily contribute to dissipation because information becomes diffused by passing across the plasma from one particle to another.

6.2 Direction of time

Reference to the retarded potentials and the effect of emission and absorption implies a distinction between advanced and retarded effects. This in itself unexpectedly brings up the problem of direction of time, this time not in electrodynamics like in absorber theory, but also and directly in the microscopic theory of phase space evolution. The delayed and integrated response of the charge and current densities at location $\boldsymbol{q}$ and time t to the variation in the corresponding densities at all locations $\boldsymbol{q}'$ and t' takes account of causality and thus of the direction of time. Ignoring the effect of time retardation, the original Liouville equation is clearly symmetric in time. It does not distinguish between processes proceeding forward and backward in time. This is one of the big problems in physics, which possibly only resolves on a macroscopic level. When making reference to signal retardation in absorber theory, this symmetry might be broken from the retarded time Eq. (8) as suggested. By replacing $t \to -t$, one has

$$t' = -\big(t + |\boldsymbol{q} - \boldsymbol{q}'|/c\big), \tag{29}$$

and thus, with constant velocity of light c, the negative retarded time $t' \to -t'$ becomes advanced. In order to restore retardation as required by the Wheeler–Feynman absorber theory, one needs to redefine the velocity of light as $c \to -c$. In a time-symmetric many-particle theory, the negative time direction would come into accord with absorber theory only under the requirement that time velocity c is negative there, i.e. one has to take the negative root $c = -1/\sqrt{\epsilon_0\mu_0}$. There is no obvious reason why this should be imposed, and it thus becomes a philosophical question. Should c be considered the inverse positive or negative root of the product of susceptibilities of the vacuum, or should c be interpreted as a positive speed, the speed of light, with reference to a distance travelled by time in either positive or negative time?

This question cannot be answered a priori. Absorber theory is restored in the second case in the causal many-particle theory. When considering the vacuum as a medium in which the dispersion of electromagnetic waves is described by a dispersion relation $\omega^2 = k^2/\epsilon_0\mu_0$, interpreting this as the relation between photon energy and momentum, one has $\hbar\omega = \pm\hbar k/\sqrt{\mu_0\epsilon_0}$. Since photon energies should be real and positive, a negative sign of the root implies negative wave numbers or negative photon momenta and thus also spatial inversion.

6.3 Conclusions

The present investigation extends Klimontovich's approach to kinetic plasma theory to the inclusion of signal retardation effects. It applies to systems of indistinguishable charged particles interacting via their self-consistent electromagnetic fields. One can trivially extend it to the presence of external fields like stationary or variable magnetic fields caused by external sources.

A number of points may be worth mentioning. First, the result looks simple as it seems that simple replacement of time with retarded time would have been sufficient to obtain it. This is true, but it is not proof for the result's correctness. For this reason we have chosen to follow the derivation step by step, which is the usual way of confirming a hypothesis. This required using the basic equations derived by Klimontovich (1967) in his fundamental approach; it also required reference to the famous absorber theory (Wheeler and Feynman, 1945, 1949) and the Liénard–Wiechert potentials on which it was based. The result is, however, substantially more complicated than Klimontovich's by the fact that the retarded time itself depends on the primed space coordinate $\boldsymbol{q}'$, which is an integration variable, and on space itself. This complicates any calculation. The main physical consequence is, however, that the retardation effect restricts applications to the domain which in observation time can be accessed by the propagation of light. For example, the observation of plasma waves in the Earth's foreshock at a frequency of $\omega/2\pi \sim 30\,\mathrm{kHz}$ implies that the observed source of emission must have been located in a region of distance $\Delta x < 10\,\mathrm{km}$. Though this is not a severe restriction for Langmuir waves, which are locally excited, application to astrophysical conditions is more interesting. The mechanism of the modulation of solar radio emissions at, say, 300 MHz with a frequency of 10 Hz is restricted to a region substantially smaller than 30 000 km, which in any theory of such a mechanism must be taken into account in the calculation. In galactic astrophysics, typical scales are several parsecs, referring to times of a few lightyears, which sets bounds on plasma mechanisms which could be evoked to participate in a causal relation.

The same procedure may also be applied to other classical fields since in all interactions the transport of information from the agent to the absorber takes time. This is the case in gases where sound waves or gravity waves can be excited and these transport the information from one fluid element to another place to affect the dynamics of other elements. In these cases it is not the photons but phonons that transport energy and information. Application to these systems lies outside the intention of the present work.

Data availability. No data sets were used in this article.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgement. This work was part of a Visiting Scientist Programme at the International Space Science Institute Bern in 2007. The interest of the ISSI Directorate is acknowledged, as, and in particular, is the friendly hospitality of the ISSI staff. Thanks are directed to the ISSI system administrator S. Saliba for technical support and to the librarians Andrea Fischer and Irmela Schweizer for access to the library and literature.

The topical editor, E. Roussos, thanks P. Yoon for help in evaluating this paper.

References

Jefimenko, O. D.: Electricity and Magnetism: An Introduction to the Theory of Electric and Magnetic Fields, Appleton-Century-Crofts, New York, 1966.

Klimontovich, Y. L.: The Statistical Theory of Non-equilibrium Processes in a Plasma, The MIT Press, Cambridge, MA, 1967.

Panofsky, W. K. H. and Phillips, M.: Classical Electricity and Magnetism, Addison-Wesley, New York, 1962.

Wheeler, J. A. and Feynman, R. P.: Interaction with the absorber as the mechanism of radiation, Rev. Mod. Phys., 17, 157–161, doi:10.1103/RevModPhys.17.157, 1945.

Wheeler, J. A. and Feynman, R. P.: Classical electrodynamics in terms of direct interparticle action, Rev. Mod. Phys., 21, 425–433, doi:10.1103/RevModPhys.21.425, 1949.