CRITIQUE OF FLUID THEORY OF MAGNETOSPHERIC PHENOMENA

W. J. HEIKKILA
The University of Texas at Dallas, Dallas, Tex. U.S.A.

(Received 28 November, 1972; in revised form 8 February, 1973)

Abstract. It is pointed out that the fluid theory has been successful in magnetospheric problems (such as the shape of the magnetopause) which involve basic considerations such as the conservation of particles, of momentum, and of energy, but that it is inadequate for other problems (such as the energization of auroral particles). Difficulties arise from the fact that it is not always possible to specify 'a volume of plasma' because particles do not remain as neighbours. Misuse of the fluid theory has led to a number of fallacies, such as the idea that the causal order of physical events in cosmic electrodynamics is the reverse of that in the familiar laboratory electrodynamics. This mistaken idea comes from a confusion of a mathematical sequence of calculations with the causal order. Also, the importance of the magnetic field as an active element is over-emphasized. Appreciation of the fact that kinetic theory is the more fundamental seems to be widely lacking. A plea is made for a common sense approach to magnetospheric and auroral problems wherein the fluid theory is used whenever it can, but where it is not expected to be adequate for all purposes.

1. Introduction

In recent years it has become fashionable to discuss magnetospheric and auroral problems in terms of hydromagnetic-thermodynamic (or MHD) concepts, often to the exclusion of the more fundamental kinetic theory. This bias is perhaps the natural result of some rather spectacular successes of the fluid theory, such as the accurate specification of the shape of the magnetopause and the description of large scale magnetospheric convection. These successes have apparently blinded many workers to the limitations of the fluid theory.

We must however admit that the fluid theory does have limitations. It has brought no significant advances in auroral and magnetospheric theory for several years. It has led to an impasse on the question of auroral particle energization processes, and of substorm development. Our purpose here is to discuss these limitations, and thus to remove the fluid theory from the pedestal to which it has been raised.

We begin in the next section with a very brief and qualitative outline of the different theoretical approaches, showing how they are related. In Part 2 some of the limitations of the fluid theory as applied to magnetospheric problems are elaborated and in Part 3 some of the consequent fallacies are exposed. The idea of frozen field convection is criticized in Part 4 as is the concept of field line merging or annihilation. Finally, in the last part, we make a plea for a more balanced approach to magnetospheric problems wherein we use fluid theory wherever we can, but instead of carrying it on ad absurdum, we turn to the more basic kinetic theory when the capabilities of the fluid theory are exceeded.

Astrophysics and Space Science **23** (1973) 261–268. *All Rights Reserved*

2. Alternative Theoretical Formulations

The ultimate description of a gas or plasma would be in terms of the motions of all the individual particles, subject to external forces as well as to forces due to the other particles. This amount of detail is of course both unattainable and unnecessary. It is however often useful to consider a representative test particle in the gas, with the interactions between particles being represented by average or statistical force fields. The Langevin treatment of collisions is one example, while Parker's (1957) Newtonian analysis of a tenuous plasma is another.

Some kind of averaging over particles is required to get macroscopic quantities, and the most satisfactory method makes use of phase space distribution functions, which specify the numbers of particles with given positions and momenta. The development of the distribution function in space and time is described by Boltzmann's equation. Once the distribution function is evaluated, then the gas properties of interest can be deduced by suitable integrations. This statistical kinetic theory generally provides an accurate physical picture, when it can be worked out.

Unfortunately, the kinetic theory is mathematically very difficult, and it can be fully worked out for only a few simple and idealised cases. It is usually necessary to seek only partial answers. One important method is to evaluate the velocity moments of the Boltzmann equation, which are related to various invariants such as the number of particles, the total momentum, and the total energy. The chief difficulty in this approach is that each moment equation involves the next higher moment, and some method for truncating the series of equations must be found. The usual way is to make an appeal to thermodynamics, and to use an equation of state.

The moment equations in fact reduce to the so-called hydrodynamic or fluid equations, which can also be derived from first principles on the basis of fluid concepts. Wherever these fluid equations can be used in physical problems they should of course be used, in view of their relative simplicity, and also in view of the familiarity of concepts such as fluid flow and pressure.

However, the fluid theory is only an approximation to the more complete kinetic theory. The central point to be made is that "the (kinetic) description is the more fundamental and only by means of it [or by direct experimental verification] can the more approximate fluid theory be justified in detail" (Hines, 1963). Approximations are of course generally required, and justified, in physical theories, but "the use of an approximation for one purpose by no means implies that the neglected process – or term, or quantity – is negligible for all purposes. [The approximations] must always provide a source of suspicion when the relevant parameters are unknown" (Hines, 1964). Regrettably, the limitations of the fluid theory are too often forgotten.

3. Limitations of the Fluid Theory

There is a fundamental difference between a gas or fluid and a plasma (Bohm and Gross, 1949); the properties of an ideal gas are largely determined by close collisions, whereas those of a plasma result from the long range Coulomb collisions. In a plasma each particle moves in an average randomly changing field determined by many neighbouring particles. This difference is so important that it prompted Stix (1962) to say that "perhaps the most remarkable thing about the two-fluid description of a plasma is its validity, for such a description ignores the very nature of a plasma."

We may well ask at this point how the fluid theory can be so successful if it is on such a shaky basis. However, the answer to that question is fairly obvious; the fluid equations express the conservation of particles, of momentum, and of energy, and these conservation relations are of fundamental importance. They are sufficient for many purposes, such as for the specification of the shape of the magnetopause.

The first difficulty that arises is in the choice of the equation of state (the alternative being the inclusion of the next higher order moment equation describing the heat flow). Even in the adiabatic case it is not always obvious what should be the value of γ, the ratio of specific heats. When interactions between particles (and between species of particles) are involved γ may not be any simple function at all; in complex Fourier integral representation it may not even be real (Pavkovich, 1964). Here the difficulty arises from the fact that groups of particles in a plasma do not generally move together, and it is therefore not possible to specify a 'volume of plasma' that is to be followed in its motion.

This is in fact the essential difficulty in the application of fluid theory to the magnetosphere: it is not always possible to specify a 'volume of plasma' whose history is to be described. Each particle finds that in the course of its motion it is continually changing its neighbours.

For the same reason there is serious difficulty in the choice of a reference frame in which to describe the electrodynamics. Here we turn to the 'frozen field' concept, wherein a magnetic field line is identified by the plasma located on it. In such a description the magnetic field lines are said to convect with the electric drift velocity $\mathbf{V}_E = \mathbf{E} \times \mathbf{B}/B^2$. This hydromagnetic description is a beautiful example of an approximation that is made for a specific purpose, as pointed out by Parker (1967). In it Maxwell's and Newton's equations are first examined, and all the 'nonessential field quantities' are eliminated. 'Only the field of principal interest, the magnetic field, remains. The dynamics of the geomagnetic field is reduced to a discussion of the balance of forces.'

Parker's words point directly at the reason why hydromagnetic theory cannot possibly be adequate for the analysis of auroral particle energization: it is not the magnetic field that is the field of principal interest. There is no possibility of particle energization by a static magnetic field, since the Lorentz force on a particle is orthogonal to the velocity vector. On the other hand, a significant induced electric field requires a rather rapid change in magnetic field; such rapid change is observed only during the expansive phase of a substorm. However, there is plenty of evidence for the existence of a large scale magnetospheric electric field at all times, which *can* provide the necessary energization. To understand this electric field it is necessary to keep track of charge separation and electric currents (Alfvén and Fälthammar, 1971). All of these quantities are explicitly left out of the hydromagnetic formulation. It is thus apparent that the auroral particle problem is entirely different in character from the magnetopause problem, and one wonders why people are so bent on using the same theory for the two.

It is often stated that the hydromagnetic theory is equivalent to the kinetic or particle drift analysis, and two papers which are purported to show this equivalence are those of Parker (1957) and Hines (1964). For the limited purposes which those authors pursued the two approaches were indeed found to be equivalent, but their theoretical development included assumptions which made them inapplicable to the plasma sheet in the magnetotail. Specifically, Parker neglected particle pressure in comparison with the magnetic pressure, as in the step from his Equation (44) to Equation (45), or alternatively, he assumed isotropy of the particle pressure. Both Parker and Hines made frequent use of the condition curl $\mathbf{B}=0$, which also is inappropriate to the magnetotail where the importance of current systems is beyond question. Thus, although there is some equivalence between the two approaches, that equivalence holds only for limited applications which do not include the problem of auroral particle energization in the plasma sheet.

4. Common Fallacies

There are many fallacies which have arisen from the promiscuous application of fluid theory. A common factor behind these seems to be the mistaken identification of a *mathematical sequence* of calculations with the *causal order* of physical phenomena. This error is made for example by Dungey (1958) in the first chapter of his book. Dungey throws out physical experience as formulated in our very familiar approach to electrodynamics as 'inappropriate to cosmic electrodynamics'.

Perhaps an example will help to put this criticism on a firm footing. By consideration of the balance of forces between the solar wind and the geomagnetic field it has been possible to describe the magnetopause in quantitative terms, including the magnitude of the discontinuity in the total magnetic field; having obtained this discontinuity it is then possible to calculate the strength of the surface current that must flow, by the use of $\nabla\times\mathbf{H}=\mathbf{J}$. It is then stated that 'the magnetic field causes the current' (Dungey, 1958, p. 15). We contend that this statement on causal order is a complete non-sequitur.

For a better understanding of the physics of the magnetopause it is necessary to turn to a microscopic analysis, based on a consideration of particle trajectories (see the review by Willis, 1971). Such an analysis indicates that the surface current is carried by the solar wind particles as they are deflected by the magnetic field in the surface layer. On this basis we can readily understand that the current is *driven* by the solar wind, and that this current is the direct cause of the discontinuity in the total magnetic field. Its magnitude and location at equilibrium must of course reflect 'the balance of forces'. In this picture there is no reason to throw away our previous concepts based on our experience in the laboratory. This way of looking at the problem is also in accord with the mathematical formulation in which a vector field is said to be the result of sources which determine its divergence and curl. Using Dungey's notation we can now indicate the causal sequence as

$$\mathbf{u} \to \mathbf{j} \to \mathbf{H}$$

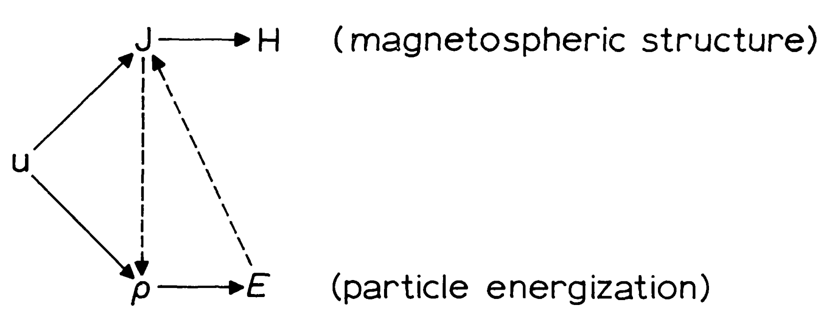
which is the usual relationship, and is the very reverse of that proposed by Dungey.

On the basis of computational expediency Dungey thus proceeds to reverse causal relationships that have come to be accepted on the basis of considerable experience. On the surface the various equations such as Maxwell's equations would appear to be reversible but that is a paradox which is true of any microscopic analysis. On a macroscopic scale relationships are not reversible in general, as can best be understood by entropy considerations. The explanation of this difference between the microscopic and the macroscopic is not easy, but we can avoid that difficult task by pointing out that the problem at hand involves *dissipation* of energy, and therefore all doubts as to causal order can be resolved quite easily.

Thus we can probably agree that energy flows from the solar wind into the magnetosphere, and not vice versa. In particular, the main source of energy is the bulk motion of the solar wind protons, and the ultimate recipient of this energy is the ensemble of particles in the magnetotail. Within the magnetosphere it must be the electric field that is the principal agent for the transfer of this energy, as pointed out above. An electrostatic field must in turn arise from the separation of charge, as given by Poisson's equation. Undoubtedly, such charge separation in the magnetosphere involves currents (Heikkila, 1973); such currents must in the steady state be entirely due to nonconservative forces (Panofsky and Phillips, 1962, p. 119), probably due to the Lorentz force on a flowing plasma, as in a MHD generator. We therefore infer the following sequence for this energization process:

$$\mathbf{u} \to \varrho \to \mathbf{E} \to \text{particle kinetic energy}.$$

We can combine these relationships into one diagram, indicating a couple of secondary relationships by dotted arrows, and adding some comments on the phenomena ensuing from the indicated processes:



This scheme does not violate any of the familiar concepts of electrodynamics.

5. Comments on Frozen Field Convection

The identification of a magnetic field line is rather arbitrary, since only the magnitude and direction at a point are specified by the fundamental relations. In the frozen field description the field lines and plasma are pictured as convecting together, as a fluid, with the electric drift velocity v_E. In this drifting frame of reference there is no electric field, in conformity with the characteristic of a plasma that electric fields are quickly shielded out by redistribution of charge.

The above concept is to a large extent satisfactory for a cold plasma with large external dimensions in a smoothly and slowly varying magnetic field; this situation applies for example to the problem of the convection of ionospheric plasma within the magnetosphere. The convection is of course driven by the magnetospheric electric field; in passing we note that there seems to be some tendency to think of the convection as producing the electric field, and caution is advised in use of the term convection electric field at least for the main magnetospheric electric field; the term is somewhat more appropriate to the corotation electric field, which does have its origin in the convection process. MHD theory has great merit for such applications in that many relationships between quantities are automatically satisfied.

When particles of higher energy (or alternatively a plasma with smaller external dimensions in a magnetic field with greater structure) are considered then the MHD model is no longer adequate. In the extreme this becomes quite obvious, as it becomes permissible to ignore the electric field entirely, or at least to treat it as a minor perturbation, when trapped energetic particles in the radiation belts are considered. For such particles in the magnetosphere, with energies exceeding some 100 keV, the major causes of drift are the gradient and the curvature of the magnetic field lines. It turns out that for auroral particles with keV energies the electric and magnetic drifts are of comparable importance. Furthermore, the relative importance of the three (i.e. the electric drift and the magnetic gradient and curvature drifts) depends strongly on the particle energy, mass, sign of charge, and pitch angle. The particles do not travel together, as they are assumed to do in the fluid theory, and it is essential that each particle be followed in its motion separately. Failure to do so dooms any attempt to explain in detail the processes of auroral particle energization and precipitation. Diagrams of convection patterns obtained solely from $\mathbf{E} \times \mathbf{B}$ are inappropriate.

There are of course other possible frames of reference, such as the one in which the total momentum is zero; this frame has the merit that it does take into account the actual motions of the plasma particles. However, this is still not a frame in which all particles move together; in particular, particles of interest (such as precipitating auroral particles with high energies, and small pitch angles) must move relative to such a frame. The description of the forces which govern their motion is therefore no simpler in this frame than it is in a stationary frame.

There is still another line of thought which has become fashionable, that magnetic field lines are merged or annihilated during their convection across x-type neutral lines. The magnetic field energy so annihilated is thought to be converted to particle kinetic energy. In spite of the prevalence of this idea for some ten years, and after many attempts to use it to describe auroral particle energization, there has nevertheless been no clarification as to the meaning of such a process. In fact there are conflicting ideas in that some define merging as existing when an electric field is applied across a plasma in which an x-type neutral line occurs (Vasyliunas, private communication), whereas others attempt to explain the electric field as originating from the merging process (Dessler, 1971). It is our contention that such attempts are futile because they are based on false notions. The only known energization process by means of a magnetic field is through the induced electric field, and this induced electric field is very small except possibly during the expansive phase. Even then it may be better to turn to a picture based on inductive current systems rather than to persist in ignoring the currents altogether.

The MHD theory places undue importance on the magnetic field. When we consider the magnetospheric problem in the Earth's frame it becomes obvious that the magnetic field is more properly to be considered as a passive element, in much the same way that a solid wall such as a rocket nozzle acts passively in ordinary hydrodynamics. The structure of the magnetic field is of course of great importance (as is the shape of the nozzle) as shown by the appearance of various geometric integrals in the theory of adiabatic motion of particles in the geomagnetic field (Roederer, 1970). But there is little or no work done by the magnetic field, again excepting certain phases of the substorm. The work is done first of all by the flowing solar wind plasma in its interaction with the magnetosphere, and subsequently by the magnetospheric electric field as particles cross equipotential surfaces.

Turning then to the particle drift approach, we need to consider a quasistatic nonconvecting magnetic field on which is superimposed a large scale electric field. This frame of reference is an entirely natural one, in that it is the one in which the particles are observed, and in which their energies are measured. It is also the one in which the current systems causing the structure in the magnetic field are fixed, since the magnetopause, ring, and tail currents are all quasi-steady in the Earth's frame. Even during the expansive phase of a substorm, when there are large changes in the magnitude and topology of the current systems, there is still no other frame of reference which is more appropriate to the over-all problem. In this frame an electric field does appear, and it can readily provide the energization required (Sharber and Heikkila, 1972; Heikkila, 1973).

6. Plea for Common Sense

That the hydromagnetic formulation might lead us into difficulty should not come as any surprise. As has been pointed out by Hines (1963) "in the hydromagnetic approach, the only component of $\mathbf{V}$ relevant to the discussion was the $\mathbf{E}\times\mathbf{B}/B^2$ term, whereas in the particle-drift approach this term contributed nothing and the remaining terms alone were relevant". Any theory in which the basic processes are so distorted must surely be suspect. Statements such as that of Axford (1967) that "this does not imply that microscopic processes are unimportant or uninteresting, but rather that they are secondary..." are indicative of the reversed logic which is so characteristic of this field.

Let us then use the fluid theory wherever we can (as in the description of ionospheric plasma convection), but let us not regard it as the fundamental description. Let us be prepared to find that it can be deficient (as in the case of auroral particle energization), and so let us use its results with suspicion until they can be substantiated by more basic theory or by observation. Above all, when the fluid theory fails let us neither force it by Procrustean measures nor castigate it, let us simply resign ourselves to the use of the more basic and difficult particle-drift or kinetic theory.

Acknowledgements

I wish to express my thanks to many persons for stimulating and sometimes heated discussions. I am particularly indebted to Colin Hines, Bill Olson, Juan Roederer and Vytenis Vasyliunas for helpful discussions and comments.

This work was done under NASA Contract NAS 5-11011.

References

Alfvén, H. and Fälthammar, C. G.: 1971, *Cosmic Electrodyn.* **2**, 78.
Axford, W. I.: 1967, *Space Sci. Rev.* **7**, 149.
Bohm, D. and Gross E. P.: 1949, *Phys. Rev.* **75**, 1851.
Dessler, A. J.: 1971, *J. Geophys. Res.* **76**, 3174.
Dungey, J. W.: 1958, *Cosmic Electrodynamics*, Cambridge at the University Press.
Heikkila, W. J.: 1973, 'Penetration of Particles into the Polar Cap and Auroral Regions', to be published.
Hines, C. O.: 1963, *Planetary Space Sci.* **10**, 239.
Hines, C. O.: 1964, *Space Sci. Rev.* **3**, 342.
Panofsky, Wolfgang K., and Phillips Melba: 1962, *Classical Electricity and Magnetism*, Addison-Wesley Publishing Company, Inc.
Parker, E. B.: 1957, *Phys. Rev.* **107**, 924.
Pavkovich, J. M.: 1964, 'Numerical Calculations Related to the RF Properties of the Plasma Sheet', Stanford University Report No. 1093.
Roederer, J. G.: 1970, *Adiabatic Motion of Geomagnetically Trapped Particles*, Springer-Verlag, New York.
Sharber, J. R. and Heikkila, W. J.: 1972, *J. Geophys. Res.* **77**, 3397,
Stix, T. H.: 1962, *The Theory of Plasma Waves*, McGraw-Hill, New York, p. 107.
Willis, D. M.: 1971, *Rev. Geophys. Space Phys.* **9**, 953.