

X-602-76-175
PREPRINT

NASA TM X- 71193

THE PARTIALLY AVERAGED FIELD APPROACH TO COSMIC RAY DIFFUSION

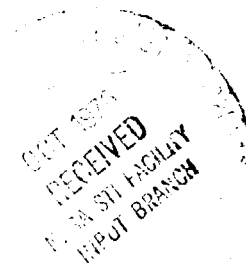
(NASA-TM-X-71193) THE PARTIALLY AVERAGED
FIELD APPROACH TO COSMIC RAY DIFFUSION
(NASA) 57 p HC \$4.50 CSCL 03B

N76-33118

Unclas
G3/93 04574

FRANK C. JONES
THOMAS J. BIRMINGHAM
THOMAS B. KAISER

AUGUST 1976



— GODDARD SPACE FLIGHT CENTER —

GREENBELT, MARYLAND

X-602-76-175
PREPRINT

THE PARTIALLY AVERAGED FIELD APPROACH
TO COSMIC RAY DIFFUSION

Frank C. Jones
Thomas J. Birmingham
Thomas B. Kaiser*

August 1976

*NAS-NRC Resident Research Associate. Present Address:
Lawrence Livermore Laboratory, Livermore, CA.

GODDARD SPACE FLIGHT CENTER

Greenbelt, Maryland 20771

CONTENTS

	<u>Page</u>
Abstract	iii
I. Introduction	1
II. Derivation of the Kinetic Equation	4
(a) Quasi-linear theory	6
(b) Partially averaged field theory	9
(c) Gaussian Statistics	13
III. The Magnetic Slab Model	13
(a) Results for the Perpendicular Slab Model	18
(b) Interpretation of the Differences between PAF and Quasi-linear $D_{\mu\mu}$'s	20
(c) Results for the Oblique Slab Model	26
IV. Spatial Transport and the Scaling of $D_{\mu\mu}^{\text{PA}}$ ($\mu \cong 0$)	27
V. Range of Validity of Partially Averaged Field Theory	32
APPENDIX A — Conditional Averaging for Multi-Dimensional Gaussian Processes	39
APPENDIX B — The Gyrotropy of $\langle f \rangle$	42

THE PARTIALLY AVERAGED FIELD APPROACH
TO COSMIC RAY DIFFUSION

Abstract

The kinetic equation for particles interacting with turbulent fluctuations is derived by a new non-linear technique which successfully corrects the difficulties associated with quasi-linear theory. In this new method the effects of the fluctuations are evaluated along particle orbits which themselves include the effects of a statistically averaged subset of the possible configurations of the turbulence. The new method is illustrated by calculating the pitch angle diffusion coefficient $D_{\mu\mu}$ for particles interacting with "slab model" magnetic turbulence, i.e., magnetic fluctuations linearly polarized transverse to a mean magnetic field $\langle \mathbf{B} \rangle$. Results are compared with those of quasi-linear theory and also with those of Monte Carlo calculations reported in a companion paper. The major effect of the non-linear treatment in this illustration is the determination of $D_{\mu\mu}$ in the vicinity of 90° pitch angles where quasi-linear theory breaks down. The spatial diffusion coefficient κ_{\parallel} parallel to $\langle \mathbf{B} \rangle$ is evaluated using $D_{\mu\mu}$ as calculated by our technique. It is argued that the partially averaged field method is not limited to small amplitude fluctuating fields and is hence not a perturbation theory.

THE PARTIALLY AVERAGED FIELD APPROACH TO COSMIC RAY DIFFUSION

I. Introduction

The systematic treatment of the transport of charged particles in disordered magnetic fields began about ten years ago.¹⁻³ These authors derived a Fokker-Planck equation for the time evolution of the particle distribution function by a method that is commonly referred to as the quasi-linear method.⁴

A more formal definition of the quasi-linear approach will be given in the next section but the basic idea may be quickly sketched. To compute the mean square effect of a random force field on a particle one must be able to follow the particle's orbit in that field for such a time as is required for the random force to become self-incoherent. This is an inherently non-linear problem and in general intractable; however, if the random force can be separated into an average part, for which orbits can be calculated, and a random part that is small (in some sense) one may use the orbit in the average field to compute the effects of the random part in the hope that the random part will become self-incoherent before it has a chance to significantly perturb the orbit.

There are two approaches one may take in applying these notions. One is to assume that the correct equation of evolution is a Fokker-Planck equation and proceed to calculate the appropriate coefficients. This approach has been developed primarily by Jokipii^{1,5} and by Hasselmann and Wibberenz.⁶ The difficulty in such an approach is the proper treatment of non-random forces.⁷ The second method is to begin with the Vlasov equation and derive a Fokker-Planck

equation to describe the evolution of the ensemble averaged distribution function. The ensemble is the collection of all possible random fields with the required statistical parameters. This approach was employed by Hall and Sturrock,³ Roelof,² Kulsrud and Pearce,⁸ and discussed in a general analysis by Kaufman.⁹ Most recent authors have employed this approach and we shall use it in this paper.

The first questions about the appropriateness of a Fokker-Planck equation for describing charged particle transport in random magnetic fields were raised by Klimas and Sandri.¹⁰ They pointed out that the presence of particle trajectories (in the average field) with pitch angles near or at 90° with respect to the average field invalidated the "adiabatic hypothesis." Since this hypothesis is necessary in order to derive the Fokker-Planck equation from a non-Markovian integro-differential equation resulting from previous steps in the derivation, these authors asserted that the correct transport equation was the more primitive (and more complicated) non-Markovian one.

The present authors then pointed out¹¹ that these same 90° pitch angle particles caused even worse problems. Such particles remain in a region of correlated field for an arbitrarily long time and thus their orbit perturbations could become important before the random force becomes self-incoherent, violating the assumptions of quasi-linear theory. Indeed we were able to show that for a particular, simple model the terms that are dropped in quasi-linear theory eventually dominate for such particles.

Although the difficulty with quasi-linear theory is confined to a region of phase space that is usually small, this region about 90° pitch angles can be quite important in considering the bulk transport of the charged particles. Earl¹² has shown that if the fluctuations in the magnetic field are depleted at high frequencies

(power spectrum steeper than k^{-2}) quasi-linear theory predicts no scattering through 90° pitch angles and hence a partially coherent propagation of the particles rather than a diffusive propagation.

That quasi-linear theory is inadequate to describe the behavior of particles with pitch angles near 90° is generally accepted and several authors have proposed modifications designed to obviate this problem. Some of the proposals have been purely phenomenological,¹³ some formal¹⁴ and some have been mixtures of the two.¹⁵ It is the purpose of the present paper to describe and present the results of a new method of calculating the diffusion coefficients of charged particles in a random magnetic field.

This method, which we call the partially averaged field (PAF) theory, has been briefly described in a previous report.¹⁶ Its fundamental idea may be summarized quite simply: while in quasi-linear theory the effects of the fluctuating field are computed along a particle orbit in the ensemble averaged field, in the present approach the orbit that is used is the particle's orbit in a partially ensemble averaged field. The partially averaged field is averaged only over a subset of the full ensemble, namely the subset of all realizations in which the value of the field at a certain fixed point has a fixed value. The results computed using this orbit are then finally averaged over the rest of the ensemble, i.e., all possible values of the field at the fixed point.

The resulting theory is non-linear in that effects of the fluctuating field are included in the particle's orbit in a fundamentally non-linear way and has the virtue of satisfying, at least marginally, the requirements for a useful, approximate theory.

These remarks will be made more precise in the next section where we will discuss the formalism employed in deriving the kinetic equations and show how the partially averaged field theory fits into this formalism. The conditions under which phase space diffusion takes place are developed. As an example of the application of our formalism we consider in Section III the "slab" model of the fluctuating field. Where existent, the pitch angle diffusion coefficient is explicitly evaluated and compared with the result of quasi-linear theory. The differences between the results of quasi-linear theory and partially averaged field theory are discussed in terms of the physical processes involved. In Section IV we review the relationship between pitch angle diffusion and spatial transport. For the pitch angle diffusion coefficients derived via the PAF theory in Section III, we derive in this section the corresponding spatial diffusion coefficients κ_{\parallel} parallel to the mean magnetic field. Since κ_{\parallel} depends importantly on the pitch angle diffusion coefficient in the vicinity of 90° pitch angles, we develop in this region the approximate scaling (from PAF theory) of the pitch angle diffusion coefficient with both the rms level of magnetic field fluctuations and with particle rigidity. Finally in Section V we argue that PAF theory is unrestricted in its validity by the rms value of the magnetic fluctuations and is - at least marginally - a genuine non-linear theory.

II. Derivation of the Kinetic Equation

Consider a collection of charged particles of rest mass m , charge q , and momentum \underline{p} [$\gamma = (1 + p^2/m^2c^2)^{1/2}$] which move in response to the electromagnetic force $\underline{F} = q(\underline{E} + \underline{v} \times \underline{B}/c)$. In our later cosmic ray considerations the electric force is negligible compared to the magnetic; for the time being, however,

we proceed with generality. \underline{F} has a component which fluctuates randomly in space and/or time. For our cosmic ray application of the theory \underline{F} is thought to be an "externally imposed" force, although the additional constraint of Maxwell's Equations can be added straightforwardly.

We describe the particles by their one particle distribution function $f(\underline{r}, \underline{p}, t)$ which, when the particles are interacting only with an externally imposed \underline{F} , obeys exactly the Vlasov Equation

$$\left(\frac{\partial}{\partial t} + \frac{\underline{p}}{m\gamma} \cdot \underline{\nabla} + \underline{F} \cdot \frac{\partial}{\partial \underline{p}} \right) f = 0 \quad (1)$$

We are interested in the coarse-grained behavior of f - its evolution on the time scale of several coherence times between the particles and the fluctuating force. We invoke the usual ergodic hypothesis and assert that this behavior is identically the behavior of $\langle f \rangle$, which is f averaged over an ensemble of \underline{F} 's. The ensemble of \underline{F} 's is specified by the values of the moments

$$\begin{aligned} & \langle \underline{F} \rangle \\ \langle \underline{F}\underline{F} \rangle & \equiv \langle (\langle \underline{F} \rangle + \delta \underline{F}) (\langle \underline{F} \rangle + \delta \underline{F}) \rangle = \langle \underline{F} \rangle \langle \underline{F} \rangle + \langle \delta \underline{F} \delta \underline{F} \rangle \\ \langle \underline{F}\underline{F}\underline{F} \rangle & = \langle \underline{F} \rangle \langle \underline{F} \rangle \langle \underline{F} \rangle + \langle \underline{F} \rangle \langle \delta \underline{F} \delta \underline{F} \rangle + \langle \delta \underline{F} \langle \underline{F} \rangle \delta \underline{F} \rangle + \langle \delta \underline{F} \delta \underline{F} \rangle \langle \underline{F} \rangle \\ & \quad + \langle \delta \underline{F} \delta \underline{F} \delta \underline{F} \rangle \end{aligned}$$

etc.

Equation (1) is readily split into an equation for $\langle f \rangle$

$$\left(\frac{\partial}{\partial t} + \frac{\underline{p}}{m\gamma} \cdot \underline{\nabla} + \langle \underline{F} \rangle \cdot \frac{\partial}{\partial \underline{p}} \right) \langle f \rangle + \langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}} \delta f \rangle = - \frac{\partial}{\partial \underline{p}} \cdot \langle \delta \underline{F} \delta f \rangle \quad (2)$$

($\partial/\partial \underline{p} \cdot \delta \underline{F} = 0$ for the electromagnetic force) and one for $\delta f = f - \langle f \rangle$

$$\left(\frac{\partial}{\partial t} + \frac{\underline{p}}{m\gamma} \cdot \underline{\nabla} + \langle \underline{F} \rangle \cdot \frac{\partial}{\partial \underline{p}} \right) \delta f + \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}} \delta f - \left\langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}} \delta f \right\rangle = - \delta \underline{F} \cdot \frac{\partial \langle f \rangle}{\partial \underline{p}} \quad (3)$$

Equations (2) and (3) together contain the full information content of (1).

An equation for $\langle f \rangle$ alone is obtained if Eq. (3) can be solved for δf in terms of its initial value $\delta f(\underline{r}, \underline{p}, t_0)$ and $\langle f \rangle$ and the result substituted on the right side of Eq. (2). An exact solution for δf is impossible, for Eq. (3) mixes derivative and ensemble average operations in a non-commutable fashion.

(a) Quasi-linear theory

The quasi-linear method approximates δf by the solution to the equation

$$\left(\frac{\partial}{\partial t} + \frac{\underline{p}}{m\gamma} \cdot \underline{\nabla} + \langle \underline{F} \rangle \cdot \frac{\partial}{\partial \underline{p}} \right) \delta f^{QL} = - \delta \underline{F} \cdot \frac{\partial \langle f \rangle}{\partial \underline{p}} \quad (4)$$

which neglects terms in Eq. (3) bi-linear in the fluctuations. The corresponding equation for $\langle f \rangle$ is

$$\left(\frac{\partial}{\partial t} + \frac{\underline{p}}{m\gamma} \cdot \underline{\nabla} + \langle \underline{F} \rangle \cdot \frac{\partial}{\partial \underline{p}} \right) \langle f \rangle = - \frac{\partial}{\partial \underline{p}} \cdot \left\langle \delta \underline{F} [U_0(t, t_0) \delta f(\underline{r}, \underline{p}, t_0)] \right\rangle + \frac{\partial}{\partial \underline{p}} \cdot \left\langle \delta \underline{F} \int_{t_0}^t d\tau U_0(t, \tau) \delta \underline{F}(\underline{r}, \underline{p}, \tau) \right\rangle \cdot \frac{\partial \langle f \rangle}{\partial \underline{p}}(\underline{r}, \underline{p}, t) \quad (5)$$

The quasi-linear propagator $U_0(t, t')$ propagates the $\underline{r}, \underline{p}$ dependence of functions to which it is applied backward in time from t to t' along trajectories in the force field $\langle \underline{F} \rangle$; these trajectories are the characteristic trajectories of Eq. (4).

We discard the initial value term from further consideration on the assumption that after a reasonable interval $t - t_0$ all correlation between δF and $U_0(t, t_0) \delta f(\underline{r}, \underline{p}, t_0)$ has died away so that ensemble averaging yields 0. If such correlations do in fact persist, we think that it is impossible to describe $\langle f \rangle$ by the simple Markovian description toward which we strive. We assume that a similar argument applies when U_0 is replaced by the more correct propagator that we introduce later.

It is the spirit of quasi-linear theory that scattering is small over a particle-fluctuation coherence time. The evolution of $\langle f \rangle$ in Eq. (5) is predominantly a convection of $\langle f \rangle(t_0)$ along the characteristic trajectories in $\langle F \rangle$. The quasi-linear collision term

$$C^{QL} = \frac{\partial}{\partial \underline{p}} \cdot \left\langle \delta F \int_{t_0}^t d\tau U_0(t, \tau) \delta F(\underline{r}, \underline{p}, \tau) \right\rangle \cdot \frac{\partial}{\partial \underline{p}} \langle f \rangle(\underline{r}, \underline{p}, \tau) \quad (6)$$

represents a small correction. It follows then that the adiabatic approximation $U_0(t, \tau) \langle f \rangle(\underline{r}, \underline{p}, \tau) \cong \langle f \rangle(\underline{r}, \underline{p}, t)$ can and should be made in every instance that the quasi-linear theory is valid.¹⁸ The quasi-linear collision term is thus properly written

$$C^{QL} = \frac{\partial}{\partial \underline{p}} \cdot \left\langle \delta F \int_{t_0}^t d\tau U_0(t, \tau) \left\{ \delta F(\underline{r}, \underline{p}, \tau) \cdot \frac{\partial}{\partial \underline{p}} \right\} \langle f \rangle(\underline{r}, \underline{p}, t) \right\rangle \quad (7)$$

The brackets $\langle \rangle$ indicate that U_0 now propagates δF in $\underline{r}/\underline{p}$ only.

We now examine the circumstances under which quasi-linear theory is valid, i.e. C^{QL} is "small." It can be shown¹⁸ that quasi-linear theory is valid for most regions of the particle phase space provided that δE and δB are small in the sense defined in Ref. 18). The reason is that the correlation $\langle \delta F(\underline{r}, \underline{p}, t) U_0(t, \tau) \delta F(\underline{r}, \underline{p}, \tau) \rangle$ diminishes rapidly with $t - \tau$. This decrease occurs in the

characteristic time t_c^{QL} and may result from (1) a short intrinsic correlation time, i.e., the correlation $\langle F(\underline{r}, \underline{p}, t) F(\underline{r}, \underline{p}, \tau) \rangle$ at fixed $\underline{r}, \underline{p}$ cuts off rapidly with increasing $t - \tau$, and/or (2) a short intrinsic correlation distance in \underline{r} through which the quasi-linear characteristic orbits move rapidly in time. The condition for the validity of quasi-linear theory may be represented concisely as $t_c^{QL} \ll t_{scat}$, where t_{scat} is the characteristic scattering time for particles in the presence of the fluctuating force. When $t_c^{QL} \ll t_{scat}$, scattering effects can be neglected in the computation of the scattering itself; quasi-linear theory is an iterative procedure.

The quasi-linear collision term Eq. (7) can be further simplified by noting that

$$U_0(t, \tau) \frac{\partial}{\partial \underline{p}} = \frac{\partial}{\partial \underline{p}(\tau)},$$

where $\underline{p}(\tau) = U_0(t, \tau)\underline{p}$ is the propagated momentum. Further

$$\frac{\partial}{\partial \underline{p}(\tau)} \Big|_{\underline{r}(\tau), \tau} = \frac{\partial \underline{p}}{\partial \underline{p}(\tau)} \cdot \frac{\partial}{\partial \underline{p}} + \frac{\partial \underline{r}}{\partial \underline{p}(\tau)} \cdot \frac{\partial}{\partial \underline{r}}$$

since a variation in $\underline{p}(\tau)$ at fixed $\underline{r}(\tau)$ results in a change in both \underline{r} and \underline{p} . The quasi-linear collision term is in general thus composed of two components

$$C^{QL} = \frac{\partial}{\partial \underline{p}} \cdot \int_{t_0}^t dt \langle F U_0(t, \tau) F(\underline{r}, \underline{p}, \tau) \rangle \cdot \frac{\partial \underline{p}}{\partial \underline{p}(\tau)} \cdot \frac{\partial (f)(\underline{r}, \underline{p}, t)}{\partial \underline{p}} \\ + \frac{\partial}{\partial \underline{p}} \cdot \int_{t_0}^t dt \langle F U_0(t, \tau) F(\underline{r}, \underline{p}, \tau) \rangle \cdot \frac{\partial \underline{r}}{\partial \underline{p}(\tau)} \cdot \frac{\partial (f)(\underline{r}, \underline{p}, t)}{\partial \underline{r}} \quad (8)$$

The first term represents pure momentum diffusion. For times $t-t_0 \gg \tau_c^{QL}$ this diffusion is described by the tensor

$$D_{\tilde{r}, \tilde{p}\tilde{p}}^{QL} = \int_0^x d\tau \left\langle \tilde{F}(\tilde{r}, \tilde{p}, t) U_0(t, t-\tau) \tilde{F}(\tilde{r}, \tilde{p}, t-\tau) \right\rangle \cdot \frac{\partial \tilde{p}}{\partial \tilde{p}(t-\tau)} \quad (9a)$$

There is in addition a mixed \tilde{p}, \tilde{r} diffusion, which for large $t-t_0$ is described by the tensor

$$D_{\tilde{r}, \tilde{p}\tilde{r}}^{QL} = \int_0^x d\tau \left\langle \tilde{F}(\tilde{r}, \tilde{p}, t) U_0(t, t-\tau) \tilde{F}(\tilde{r}, \tilde{p}, t-\tau) \right\rangle \cdot \frac{\tilde{r}}{\tilde{p}(t-\tau)} \quad (9b)$$

This mixed \tilde{p}, \tilde{r} diffusion disappears when $\langle f \rangle$ is spatially homogeneous and may be small for weak inhomogeneities. It is generally present, however, to the same degree as the pure momentum diffusion.

(b) Partially averaged field theory

In many models there are, however, pathological though important regions of phase space where τ_c^{QL} is quite long, so long in fact that $\tau_c^{QL} \geq \tau_{scat}$ and both the spirit and the substance of quasi-linear theory are violated.

We propose here an improved approximation, which includes the essential effect of the fluctuations themselves in the lowest order propagator. For reasons that will be soon evident, we call our method the partially averaged field approximation. In this formalism the coherence time τ_C^{PA} is naturally limited $\tau_C^{PA} \leq \tau_{scat}$ and the collision integral is validly calculated for (1) all regions of the particle phase space and (2) arbitrary amplitude of the fluctuations. Our method is not a weak-coupling approximation. It reduces to quasi-linear theory when $\tau_c^{QL} \ll \tau_{scat}$ holds.

Our formalism proceeds as follows. We consider the ensemble averaging process to occur in two steps. A first partial average, denoted by $\langle \rangle_{\mathcal{E}}$, is performed over all realizations with the same fixed values of $\delta \underline{E}(\underline{r}, t)$ and $\delta \underline{B}(\underline{r}, t)$ at the point \underline{r} and time t at which Eq. (2) is valid. A second final average, indicated by $\langle \rangle_{\mathcal{F}}$, is then carried out over the distributions of $\delta \underline{E}(\underline{r}, t)$ and $\delta \underline{B}(\underline{r}, t)$. The collision term in Eq. (2) is identically written according to this prescription

$$C = - \frac{\partial}{\partial \underline{p}} \cdot \langle \delta \underline{F}(\underline{r}, \underline{p}, t) \langle \delta f \rangle_{\mathcal{E}} \rangle_{\mathcal{F}} \quad (10)$$

If $\delta f(\underline{r}, \underline{p}, t)$ depended only on $\delta \underline{F}(\underline{r}, \underline{p}, t)$ it would be the same for every realization included in the partial ensemble and partial averaging would be trivial. However, δf depends on $\delta \underline{F}$ globally. We therefore write Eq. (3) for the arbitrary point $\underline{r}', \underline{p}', t'$ and partially average to obtain

$$\begin{aligned} \left[\frac{\partial}{\partial t'} + \frac{\underline{p}'}{m\gamma'} \cdot \underline{\nabla}' + \langle \underline{F} \rangle \cdot \frac{\partial}{\partial \underline{p}'} \right] \langle \delta f \rangle_{\mathcal{E}} + \left\langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}'} \delta f \right\rangle_{\mathcal{E}} - \left\langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}'} \delta f \right\rangle \\ = - \langle \delta \underline{F} \rangle_{\mathcal{E}} \cdot \frac{\partial \langle \delta f \rangle}{\partial \underline{p}'} \quad (11) \end{aligned}$$

Note that partially averaged quantities depend functionally on $\underline{r}', \underline{p}',$ and t' which are the variables of Eq. (11) and parametrically on \underline{r} and t which are singled out by the partial averaging procedure. Thus far Eqs. (10) and (11) are exact, for all we have done is to break the full ensemble average procedure into two sequential steps: symbolically $\langle \rangle \Rightarrow \langle \langle \rangle_{\mathcal{E}} \rangle_{\mathcal{F}}$

The essence of our method is to approximate in Eq. (11) as follows

$$\left\langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}'} \delta f \right\rangle_{\mathcal{E}} - \left\langle \delta \underline{F} \cdot \frac{\partial}{\partial \underline{p}'} \delta f \right\rangle = \langle \delta \underline{F} \rangle_{\mathcal{E}} \cdot \frac{\partial \langle \delta f \rangle_{\mathcal{E}}}{\partial \underline{p}'} \quad (12)$$

This approximation maintains the properties of the exact terms in two essential limits. First, when \underline{r}', t' is much different from \underline{r}, t partial averaging and total averaging are equivalent and $\langle \delta \underline{F} \cdot \partial / \partial \underline{p}' \delta f \rangle_p - \langle \delta \underline{F} \cdot \partial / \partial \underline{p}' \delta f \rangle = 0$; however, $\langle \delta \underline{F} \rangle_p = 0$ also under such circumstances. Second, at $\underline{r}' = \underline{r}, t' = t$ we expect $\langle \delta \underline{F} \cdot \partial / \partial \underline{p}' \delta f \rangle_p (= \delta \underline{F} \cdot \partial / \partial \underline{p}' \langle \delta f \rangle_p)$ to dominate $\langle \delta \underline{F} \cdot \partial / \partial \underline{p}' \delta f \rangle$, particularly for those realizations that contribute importantly in Eq. (10); the right side of Eq. (12) is, however, identically $\delta \underline{F} \cdot \partial / \partial \underline{p}' \langle \delta f \rangle_p$ at $\underline{r}' = \underline{r}, t' = t$.

Our approximate equation for $\langle \delta f \rangle_p(\underline{r}', \underline{p}', t'; \underline{r}, t)$ is thus

$$\left[\frac{\partial}{\partial t'} + \frac{\underline{p}'}{m\gamma'} \cdot \underline{\nabla}' + (\langle \underline{F} \rangle + \langle \delta \underline{F} \rangle_p) \cdot \frac{\partial}{\partial \underline{p}'} \right] \langle \delta f \rangle_p = - \langle \delta \underline{F} \rangle_p \cdot \frac{\partial \langle f \rangle}{\partial \underline{p}'} \quad (13)$$

Neglecting the initial value term, whose contribution to C is again assumed to damp rapidly with time, we express the solution to Eq. (13) as

$$\langle \delta f \rangle_p(\underline{r}', \underline{p}', t'; \underline{r}, t) = - \int_{t_0}^{t'} d\tau U_p(t', \tau; t) \langle \delta \underline{F} \rangle_p(\underline{r}', \underline{p}', \tau; \underline{r}, t) \cdot \frac{\partial \langle f \rangle}{\partial \underline{p}'}(\underline{r}', \underline{p}', \tau) \quad (14)$$

The operator $U_p(t', \tau; t)$ is the propagator in the partially averaged field. Operating on an arbitrary function $g(\underline{r}', \underline{p}', \tau)$ it propagates the $\underline{r}', \underline{p}'$ dependence backward in time from t' to τ along orbits in the partially averaged force $\langle \underline{F} \rangle + \langle \delta \underline{F} \rangle_p$. U_p also depends parametrically on t because the partial average is carried out at constant $\delta \underline{E}(\underline{r}, t)$ and $\delta \underline{B}(\underline{r}, t)$.

For Eq. (10) we need $\langle \delta f \rangle_p(\underline{r}' = \underline{r}, \underline{p}' = \underline{p}, t' = t; \underline{r}, t)$. Substituting from (13), our new collision term becomes

$$C^{PA} = \frac{\partial}{\partial \underline{p}} \cdot \int_{t_0}^t d\tau \left\langle \delta \underline{F} U_p(t, \tau; t) \langle \delta \underline{F} \rangle_p(\underline{r}, \underline{p}, \tau; \underline{r}, t) \right. \\ \left. \frac{\partial \langle f \rangle(\underline{r}, \underline{p}, \tau)}{\partial \underline{p}} \right\rangle_{\mathfrak{B}} \quad (15)$$

We expect decay of $\langle \delta \underline{F} U_p(t, \tau; t) \langle \delta \underline{F} \rangle_p(\underline{r}, \underline{p}, \tau; \underline{r}, t) \rangle$ in the time $t - \tau = t_C^{PA}$ ($t_C^{PA} \leq t_C^{QL}$) sufficiently short that all characteristic orbits in the partially averaged force remain at least marginally valid for the entire time interval over which the integrand in (15) is non-zero. We hold this expectation irrespective of the size of $\delta \underline{F}$ and shall argue why in Section V.

For times $t - t_0 \gg \tau_C^{PA}$ we again write

$$C^{PA} = \frac{\partial}{\partial \underline{p}} \cdot \int_0^t d\tau \left\langle \delta \underline{F} U_p(t, t - \tau; t) \langle \delta \underline{F} \rangle_p(\underline{r}, \underline{p}, t - \tau; \underline{r}, t) \right. \\ \left. \frac{\partial \langle f \rangle(\underline{r}, \underline{p}, t - \tau)}{\partial \underline{p}} \right\rangle_{\mathfrak{B}} \quad (16)$$

The adiabatic approximation on $\langle f \rangle$ cannot always be made in C^{PA} . The physical reason is that when $\delta \underline{F}$ is substantial, scattering as described by Eq. (16) represents a significant correction to the convective motion in $\langle \underline{F} \rangle$. Our success in deriving an equation unrestricted by the size of $\delta \underline{F}$ is thus somewhat mitigated by the fact that this equation for $\langle f \rangle$ is an integro-differential equation. One must not conclude, however, that an adiabatic approximation is never possible: each model to which the $\langle f \rangle$ equation is applied must be examined individually to see whether an adiabatic ansatz can be made consistently with the resulting physics. Work reported in later sections is illustrative of this point.

(c) Gaussian statistics

It is impossible to proceed further with the partially averaged field method in full generality. We always consider the fluctuating electric and magnetic fields, $\delta \underline{\underline{E}}(\underline{r}, t)$ and $\delta \underline{\underline{B}}(\underline{r}, t)$, to be Gaussian processes. It can be shown (Appendix A) for a Gaussian process that

$$\langle \delta \underline{\underline{E}} \rangle_{\rho}(\underline{r}', t'; \underline{r}, t) = \delta \underline{\underline{E}}(\underline{r}, t) \cdot \underline{\underline{C}}_E(\underline{r}', t', \underline{r}, t) \quad (17a)$$

$$\langle \delta \underline{\underline{B}} \rangle_{\rho}(\underline{r}', t'; \underline{r}, t) = \delta \underline{\underline{B}}(\underline{r}, t) \cdot \underline{\underline{C}}_B(\underline{r}', t', \underline{r}, t) \quad (17b)$$

where $\underline{\underline{C}}_E$ and $\underline{\underline{C}}_B$ are the normalized correlation tensors for the fluctuating electric and magnetic fields respectively.

The assumption of a Gaussian process is a common one, partly because the central limit theorem¹⁹ states that a tendency towards Gaussian statistics is a common occurrence in random systems, but mostly because of its mathematical tractability which most other types of statistics do not share. Mathematically, Eqs. (17) follow uniquely – and $\delta \underline{\underline{E}}$ and $\delta \underline{\underline{B}}$ are Gaussian processes – provided that one demands that $\langle \delta \underline{\underline{E}} \rangle_{\rho}(\underline{r}', t'; \underline{r}, t)$ and $\langle \delta \underline{\underline{B}} \rangle_{\rho}(\underline{r}', t'; \underline{r}, t)$ be linearly dependent respectively on $\delta \underline{\underline{E}}(\underline{r}, t)$ and $\delta \underline{\underline{B}}(\underline{r}, t)$.

III. The Magnetic Slab Model

We have worked out the details of the partially averaged field formalism for the so-called magnetic slab model, and we shall discuss details in this section. The magnetic slab model was chosen because its simplicity facilitates calculation and direct comparison with Monte Carlo results (which are reported in the accompanying paper²⁰) and because it also contains some of the basic physics of the scattering of cosmic rays from irregularities in the interplanetary medium.

In the magnetic slab model there are no electric fields. The ensemble average magnetic field $\langle \underline{B} \rangle$ is spatially homogeneous and oriented along the z-axis of a Cartesian coordinate system; magnetic fluctuations are linearly polarized, constant in time, and vary in magnitude only with distance along $\langle \underline{B} \rangle$. Thus $\delta \underline{B} = \delta B(z) \hat{e}_x$. In this simplest form, there is no spatial variation in planar slabs oriented perpendicular to the z-axis and hence the coinage "slab model." At the end of this section we discuss a more generalized slab model in which $\delta \underline{B}$ is in the \hat{e}_x direction but planes of constant $|\delta \underline{B}|$ are \perp to the y-z plane and at a fixed angle (typically 30°) with respect to the z-axis.

The magnetic slab model is intended as a description of the interaction of cosmic rays with slow magnetic structures (tangential discontinuities or Alfvén waves) in the solar wind. Let us assume that the radial divergence of the solar wind flow can be neglected and that we can place ourselves therefore in a unique solar wind frame of reference. The amplitude of the electric fields associated with these slow structures is at most $|\underline{E}| = v_A |\underline{B}|/c$ and hence far smaller than the magnitude of the magnetic fields (v_A is the Alfvén speed and is typically 50-100 km/sec). The magnetic force on relativistic cosmic rays is therefore far larger than the electric force, and for that reason we neglect the latter.

Our assumption that $\langle \underline{B} \rangle$ is completely uniform breaks down, of course, on the 1 AU spatial scale of the classical Archimedes spiral. However, the pertinent scale for a description of scattering is the autocorrelation distance of magnetic fluctuations, which has been measured to be a few hundredths of an AU;²¹ to the extent that $\langle \underline{B} \rangle$ is uniform over this length our model is a good one. There is also evidence²² that the magnetic fluctuations are preferentially polarized in the direction normal to the ecliptic plane (the normal direction would thus correspond

to our x-direction); any component of $\delta \underline{B}$ in the ecliptic plane and normal to $\langle \underline{B} \rangle$ at its point of generation gradually develops a compressional component²³ as it propagates outward in the solar wind and tends to damp away.²⁴

Since electric fields are negligible, cosmic rays suffer no energy changes in their interactions with the fluctuations and $|\underline{p}|$ is simply a parameter in the equation for $\langle f \rangle$. Scattering in angle does occur and can be described in terms of the variables $\mu = \underline{p} \cdot \hat{e}_z / |\underline{p}|$ and $\phi = \tan^{-1} p_x / p_y$. μ is the cosine of the cosmic ray's pitch angle relative to $\langle \underline{B} \rangle$ and ϕ is the gyrophase angle measured from $\delta \underline{B}$ (the x-axis).

We consider here the situation where $\langle f \rangle$ is spatially homogeneous. Such a situation arises when $\langle f \rangle$ is homogeneous at $t = t_0$ and the spectrum of the fluctuations is itself spatially homogeneous. This assumption of spatial homogeneity is artificial and unrealistic: as a result of it the bulk spatial transport of cosmic rays along and across field lines is precluded. It is made solely to achieve a simple model in which to explore and justify the partially averaged field method. The spatially inhomogeneous situation will be discussed in Section IV.

With the aforementioned approximations the equation for $\langle f \rangle$ is

$$\frac{\partial \langle f \rangle}{\partial t} - \langle \omega \rangle \frac{\partial \langle f \rangle}{\partial \phi} = D(\mu, t) \left\langle \int_0^\tau d\tau U_{\perp}(\underline{r}, \underline{p}, t - \tau) \delta \omega \right\rangle_{D(\mu, t) \langle f \rangle(\underline{r}, \underline{p}, t - \tau)} \quad (18)$$

In Eq. (18), $\langle \omega \rangle = q \langle B \rangle / \gamma mc$ is the gyrofrequency in the ensemble average magnetic field $\langle B \rangle \hat{e}_z$, $\delta \omega$ is the gyrofrequency $q |\delta \underline{B}| / \gamma mc$ that would prevail if $|\delta \underline{B}| \hat{e}_x$ were the only magnetic field, and D is the differential operator

$$D = (1 - \mu)^{1/2} \sin \phi \frac{\partial}{\partial \mu} - \frac{\mu \cos \phi}{(1 - \mu^2)^{1/2}} \frac{\partial}{\partial \phi} \equiv \frac{\partial}{\partial \mu} (1 - \mu^2)^{1/2} \sin \phi - \frac{\partial}{\partial \phi} \frac{\mu \cos \phi}{(1 - \mu^2)^{1/2}} \quad (19)$$

We next approximate further by neglecting the gyrophase dependence of $\langle f \rangle$. This is accomplished formally by first expanding $\langle f \rangle$ in the Fourier series

$$\langle f \rangle = \sum_{n=-\infty}^{+\infty} \langle f \rangle_n \exp i n \phi,$$

next substituting into Eq. (18) and projecting out the $n = 0$ component

$$\begin{aligned} \frac{\partial \langle f \rangle_0}{\partial t} &= \frac{1}{2\pi} \int_0^{2\pi} d\phi \frac{\partial}{\partial \mu} (1 - \mu^2)^{1/2} \sin \phi \left\langle \int_0^\infty d\tau U_\omega(t, t - \tau; t) \right. \\ &\quad \left. \left[(1 - \mu^2)^{1/2} \sin \phi \frac{\partial}{\partial \mu} - \frac{\mu \cos \phi}{(1 - \mu^2)^{1/2}} \frac{\partial}{\partial \phi} \right] \right. \\ &\quad \left. \sum_{n=-\infty}^{+\infty} \langle f \rangle_n \exp i n \phi \right\rangle_0 \quad (20) \end{aligned}$$

and finally neglecting the $n \neq 0$ terms on the right hand side of Eq. (21). Fourier amplitudes $\langle f \rangle_n$, $n > 0$, are small compared to $\langle f \rangle_0$ in the approximate ratio $\eta^2 = \langle \delta B^2 \rangle / \langle B \rangle^2$ (cf. Appendix B) provided that they are initially small; the physical reason being that the convective term $(v \cdot \nabla) \langle f \rangle / \partial \phi$ keeps particles well stirred in ϕ . Thus

$$\frac{\partial \langle f \rangle_0}{\partial t} = \frac{1}{2\pi} \int_0^{2\pi} d\phi \frac{\partial}{\partial \mu} (1 - \mu^2)^{1/2} \sin \phi \left\langle \delta \omega \int_0^x d\tau U_p(t, t - \tau; t) \right. \\ \left. \delta \omega (1 - \mu^2)^{1/2} \sin \phi \frac{\partial \langle f \rangle_0}{\partial \mu}(\mu, \tau) \right\rangle \quad (21)$$

Next note

$$\dot{\mu} = \frac{\dot{p}_z}{p} = - \frac{\delta \omega p_y}{p} = - \delta \omega (1 - \mu^2)^{1/2} \sin \phi \quad (22)$$

so that

$$\frac{\partial \langle f \rangle_0}{\partial t} = \frac{\partial}{\partial \mu} \frac{1}{2\pi} \int_0^{2\pi} d\phi \left\langle \dot{\mu} \int_0^x d\tau U_p \frac{\partial \langle f \rangle_0(\mu, \tau)}{\partial \mu} \right\rangle \quad (23)$$

To the right side of Eq. (23) must be added any effects such as sinks and sources of particles which are not included in the original Vlasov description Eq. (1).

In the steady state $\langle f \rangle_0$ no longer depends explicitly on time and (23) can therefore be written in the more transparent form

$$\frac{\partial}{\partial \mu} \left[\frac{1}{2\pi} \int_0^{2\pi} d\phi \left\langle \dot{\mu} \langle \Delta \langle f \rangle_0 \rangle \right\rangle \right] = 0 \quad (24)$$

Here

$$\langle \Delta \langle f \rangle_0 \rangle_\mu = \int_0^x d\tau U_p \dot{\mu} \frac{\partial \langle f \rangle_0(\mu)}{\partial \mu} = \int_0^x d\tau \frac{d \langle f \rangle_0}{d\tau}$$

is the change in $\langle f \rangle_0$ apparent to an observer moving from the point μ at $t = 0$ along an orbit in the partially averaged field $\langle \delta \mathbf{B} \rangle_\mu$.

To the extent that $\langle f \rangle_0(\mu)$ is a linear function of μ - and we have found from our Monte Carlo calculations that such is approximately the case for a wide variety of conditions -

$$\langle \Delta \langle f \rangle_0 \rangle_\tau \approx \langle \Delta \mu \rangle_\tau \frac{\partial \langle f \rangle_0}{\partial \mu}$$

and Eq. (24) is of diffusion form

$$\frac{\partial}{\partial \mu} \left(D_{\mu\mu}^{\text{PA}} \frac{\partial \langle f \rangle_0}{\partial \mu} \right) = \frac{\partial}{\partial \mu} \left[\frac{1}{2\pi} \int_0^{2\pi} d\phi \langle \dot{\mu} \langle \Delta \mu \rangle_\tau \rangle ; \frac{\partial \langle f \rangle_0}{\partial \mu} \right] = 0 \quad (25)$$

The diffusion coefficient $D_{\mu\mu}^{\text{PA}}$ has exactly the same form as the quasi-linear diffusion coefficient $D_{\mu\mu}^{\text{QL}}$ but differs in that $\Delta\mu$ is the change in μ evaluated following an orbit in the partially averaged field rather than following a helical orbit in $\langle \underline{B} \rangle$ as quasi-linear theory prescribes. For a $\mu = 0$ particle in the slab model $\Delta\mu$ approaches no finite limit as $\tau \rightarrow \infty$ in quasi-linear theory²⁵ because such a particle suffers exactly canceling deflections over its circular orbit at constant z in $\langle \underline{B} \rangle$; no such pathological behavior occurs in the partially averaged field method because even $\mu = 0$ particles move freely in z and hence enter regions of uncorrelated $\delta \underline{B}$.

(a) Results for the Perpendicular Slab Model

In order to evaluate $\langle \Delta \mu \rangle_\tau$ and hence the diffusion coefficient

$$D_{\mu\mu}^{\text{PA}} = \frac{1}{2\pi} \int_0^{2\pi} d\phi \langle \dot{\mu} \langle \Delta \mu \rangle_\tau \rangle ; \quad (26)$$

Newton's Equations of motion

$$\frac{d\mathbf{p}'}{d\tau} = \mathbf{p}' \times \left[\langle \mathbf{v} \rangle \hat{\mathbf{e}}_z + \delta v(z) C(z-z') \hat{\mathbf{e}}_x \right] \quad (27a)$$

$$\frac{d\mathbf{r}'}{d\tau} = \frac{\mathbf{p}'}{m} \quad (27b)$$

are integrated numerically backward in time from the boundary conditions $\underline{p}'(\tau = t) = \underline{p}$, $\underline{r}'(\tau = t) = \underline{r}$. The correlation function C is chosen to be the exponential

$$C(z - z') = \frac{1}{z_c} \exp - \frac{|z - z'|}{z_c} \quad (28)$$

since we compare $D_{\mu\mu}^{PA}$ with the results of a Monte Carlo simulation,²⁰ for which it is simplest to generate numerically magnetic fields with an exponential correlation function. The correlation distance z_c is the distance scale and the correlation time $t_c = z_c \gamma m/p$ the unit of time.

We first pick a $\delta\phi(z)$ and the boundary values $\mu(\tau = t)$ and $\dot{\mu}(\tau = t)$. $\dot{\mu}$ for inclusion in (26) follows immediately from Eq. (22). $\langle \Delta\mu \rangle$, on the other hand is the change in μ which results from the integration of Eqs. (27). This integration is carried out until the particle has moved typically a distance $z' - z \approx 16z_c$ in the direction of varying $\delta\phi(z')$.

The entire procedure is next averaged over a uniform distribution of $\dot{\mu}(t)$. The density with which the $\dot{\mu}(t)$ must be chosen depends on $\mu(t)$. When the latter is small ($\mu(t) \lesssim \dot{\mu}(z)/\langle \dot{\mu} \rangle$) particle mirroring is an important effect, as we shall see shortly. For small μ , the $\dot{\mu}$'s are taken at 1° intervals in order to achieve convergence for the $\dot{\mu}$ -averaging process. For larger μ ($\mu \gtrsim \dot{\mu}(z)/\langle \dot{\mu} \rangle$) mirroring does not occur and it is sufficient to take $\dot{\mu}(t)$'s which are 2° apart.

The last step in the evaluation of $D_{\mu\mu}^{PA}$ is the final averaging procedure, in which we perform the steps of the preceding two paragraphs for several different values of $\delta\phi(z)$, typically 20, each weighted by the distribution

$$P(\delta\phi) = \frac{1}{\sqrt{2\pi} \langle \delta\phi^2 \rangle} \exp - \frac{\delta\phi^2}{2 \langle \delta\phi^2 \rangle}$$

Figures 1 to 3 depict $D_{\mu\mu}^{\text{PA}}$ thus computed as a function of μ . Each figure corresponds to particular values of $\gamma = \langle \delta v^2 \rangle^{1/2} / \langle v \rangle$ and $\tau = \gamma m \langle v \rangle z_c / p$. The discrete points at which $D_{\mu\mu}^{\text{PA}}$ is calculated are indicated. The solid curve is a spline fit to these points. Shown also on each graph is the curve of $D_{\mu\mu}^{\text{QL}}$ vs μ . This quasi-linear diffusion coefficient is computed from the formula

$$D_{\mu\mu}^{\text{QL}} = \frac{1 - \mu^2}{2} |\mu| \frac{\gamma^2 \epsilon}{\epsilon^2 + \mu^2} \quad (29)$$

which results when the exponential correlation function (28) is assumed.

Equation (29) is the limiting $\tau \rightarrow \infty$ value of the expression

$$\left\langle \dot{\mu}(t) \Delta_{\mu}^{\text{QL}}(t - \tau) \right\rangle$$

wherever that limit exists. As we have mentioned earlier and demonstrated elsewhere²⁵ this limit does not exist as $\mu \rightarrow 0$ and indeed a quasi-linear diffusion coefficient is not properly defined at such small μ 's.

We would point out three aspects of these figures: (1) the peaking and definitely non-zero value of $D_{\mu\mu}^{\text{PA}}$ as $\mu \rightarrow 0$, (2) the fact that $D_{\mu\mu}^{\text{PA}}$ is approximately equal to but slightly less than $D_{\mu\mu}^{\text{QL}}$ for $\mu > 0$, and (3) the essential equality of $D_{\mu\mu}^{\text{PA}}$ and $D_{\mu\mu}^{\text{QL}}$ as $\mu \rightarrow 1$. These aspects are also characteristic of the diffusion coefficient $D_{\mu\mu}$ obtained from Monte Carlo computer experiments.²⁰

(b) Interpretation of the Differences between PAF and Quasi-Linear $D_{\mu\mu}$'s.

We interpret the bump which appears in the vicinity of $\mu = 0$ as due to particle mirroring: particles with pitch angles near 90° are dominantly affected by the maximum in $\langle B \rangle_{\parallel}$ at $z = 0$ and their $\dot{\mu}$'s are determined by magnetic moment conservation to be

$$\langle \Delta \mu \rangle_{\tilde{\eta}} = \mp \left[1 - \frac{1 - \mu'^2}{(1 + \tilde{\eta}^2)^{1/2}} \right]^{1/2} - \mu = \langle \mu \rangle_{\tilde{\eta}} - \mu \quad (30)$$

(Without loss of generality we assume here that $z = 0$ is the point at which the diffusion coefficient is to be evaluated.) Here $\tilde{\eta} = \delta B(0) / \langle B \rangle$ and μ' is the cosine of the particle's pitch angle with respect to the total (partially averaged) field $\langle \tilde{B} \rangle_{\tilde{\eta}}$:

$$\mu' = \frac{\mu}{(1 + \tilde{\eta}^2)^{1/2}} + \frac{(1 - \mu^2)^{1/2} \tilde{\eta}}{(1 + \tilde{\eta}^2)^{1/2}} \cos \psi \quad (31)$$

Using Eq. (30) for $\Delta \mu$ and Eq. (22) for $\dot{\mu}$, we determine the mirroring contribution to $D_{\mu\mu}^{PA}$ thus to be (cf. Eq. (26))

$$D_{\mu\mu}^{PA} = \frac{(1 - \mu^2)^{1/2}}{2\pi} \left\langle \dot{\mu} \int_0^{2\pi} d\psi \sin \psi \left\{ \mp \left[1 - \frac{1 - \mu'^2}{(1 + \tilde{\eta}^2)^{1/2}} \right]^{1/2} \right\} \right\rangle \quad (32)$$

The \mp sign is included within the integral sign because the algebraic sign of $\langle \mu \rangle_{\tilde{\eta}}$ is dependent on ψ ; if it were not, (32) would clearly integrate to zero because μ' depends only on $\cos \psi$, and there would be no contribution of mirroring to the diffusion coefficient. Furthermore the sign cannot be just a function of $\cos \psi$ (as is the sign of μ') or the same reasoning gives a zero result once again. We shall see that it is just the act of mirroring that injects the needed asymmetry about $\psi = 0$, i.e., a dependence on $\sin \psi$, to produce the peak of $\mu = 0$.

To determine the sign of $\langle \mu \rangle_{\tilde{\eta}}$ we proceed in the following heuristic fashion. Consider the $\mu > 0$ (but $|\mu| \cong 0$) particle at $z = 0$ at $t = 0$ in the partially averaged field. Depending on its ψ and μ , this particle's guiding center is at positive or negative z .

$$z_{GC} \approx \frac{1}{\epsilon} \frac{(1-\mu^2)^{1/2}}{\left(1 + \frac{\phi B^2(0)}{\langle B \rangle^2}\right)} \frac{\phi B(0)}{\langle B \rangle} \sin \theta z_C = \frac{1}{\epsilon} \frac{(1-\mu^2)^{1/2}}{(1+\tilde{\eta}^2)} \tilde{\eta} \sin \theta z_C \quad (33)$$

and its μ' (cf. Eq. (31)) is also positive or negative. If z_{GC} and μ' are both > 0 , its initial motion carries the particle to the right (to larger z) and we assert that $\langle \mu \rangle_{z_0} > 0$. Similarly, if z_{GC} and μ' are both < 0 , $\langle \mu \rangle_{z_0} < 0$. On the other hand, if z_{GC} and μ' have opposite algebraic signs, the particle's initial motion carries its guiding center toward $z = 0$ and stronger magnetic field so that mirroring is a possibility. If mirroring does occur – and we shall presently develop criteria for mirroring – $\langle \mu \rangle_{z_0}$ has the same sign as z_{GC} . If mirroring does not occur, $\langle \mu \rangle_{z_0}$ has the same sign as μ' . In Figure 4, which depicts the transverse p_x, p_y plane, mirroring may occur for $\tilde{\eta} > 0$ in those sectors denoted by an "M?". (Because of the symmetry of the distribution of $\phi B(0)$, we may restrict our consideration to $\tilde{\eta} > 0$.) The sign of $\langle \mu \rangle_{z_0}$ in the complementary sectors is also indicated. The angle A is given by

$$A = \sin^{-1} \left[\frac{\mu'}{(1-\mu^2)^{1/2}} \frac{1}{\tilde{\eta}} \right]$$

When does mirroring in fact occur? We first applied the criterion that predicts mirroring when

$$\frac{1-\mu'^2}{\langle B \rangle \cdot (z_{GC})} > \frac{1}{\langle B \rangle \cdot (z=0)} \quad (34)$$

On using the approximation

$$\langle B \rangle \cdot (z_{GC}) \approx \langle B \rangle \cdot (z=0) \left(1 - \frac{z_{GC}}{z_C} \frac{\tilde{\eta}^2}{1+\tilde{\eta}^2} \right)$$

plus the expressions (31) and (33) for μ' and z_{GC} , the criterion for mirroring is equivalently written

$$\left| \mu + (1 - \mu^2)^{1/2} \tilde{\gamma} \cos \theta \right|^2 < \frac{\tilde{\gamma}^3}{\epsilon} \frac{(1 - \mu^2)^{1/2}}{(1 + \tilde{\gamma}^2)} |\sin \theta| \quad (35)$$

The θ -integration in Eq. (32) has been performed numerically. The ambiguity in the sign of $\langle \dots \rangle$ in the regions of Figure 4 marked by "M?" is resolved by testing for the validity of (35) and assigning $\langle \dots \rangle$ the sign of z_{GC} if (35) is indeed satisfied.

Following the procedure just prescribed, we have calculated the mirroring contribution to $D_{\mu\mu}^{PA}$ and found it in the vicinity of $\mu = 0$ to be typically a factor of 2 larger than the $D_{\mu\mu}^{PA}$ determined by exact integration of orbits in the partially averaged magnetic field: mirroring has been introduced too strongly in our heuristic procedure. Much better agreement is obtained when we modify the mirroring requirement (34) as follows:

$$\frac{1 - \mu'^2}{\langle B \rangle(z_{GC})} > \frac{1}{\langle B \rangle(z=0) \left[1 - \frac{(1 - \mu^2)^{1/2} \tilde{\gamma}^3}{(1 + \tilde{\gamma}^2)^{3/2}} \frac{2}{\tilde{\gamma}} \right]} \quad (36)$$

The additional factor in the denominator of the right side accounts for the fact that even when the guiding center is located at $z = 0$ the particle itself is almost always at a weaker field strength, owing to the $\exp - (|z|/z_C)$ fall-off in B_x .

We attribute the fact that $D_{\mu\mu}^{PA} < D_{\mu\mu}^{OL}$ in the regions $\mu > 0$ of Figs. 1-3 to another effect, which we shall try to explain heuristically in this section. In quasi-linear theory, a particle moves with the assumed constant momentum p_{\parallel} in the z -direction. As time increases from $t = 0$, the particle interacts (in the

ensemble average) with a portion of the power spectrum of the fluctuations of ever-decreasing width: the effect on $D_{\mu\mu}^{QL}$ of all Fourier components except that at the resonant wave number $k = \epsilon/|\mu| z_c$ ultimately averages to zero.

The constancy of μ and p_z are, however, only approximate. More accurately, the particle attempts to maintain the constancy of μ' , Eq. (31), which we write in a slightly modified form as

$$\frac{\gamma m}{p} \frac{dz}{d\tau} + (1 - \mu^2)^{1/2} \tilde{\eta}_l \cos \phi = \text{const.} \quad (37)$$

Equation (37) can be integrated approximately to give

$$z = \left[\frac{p}{\gamma m} \mu + (1 - \mu^2)^{1/2} \tilde{\eta}_l \cos \phi \right] \tau - \frac{p}{\gamma m} \frac{(1 - \mu^2)^{1/2} \tilde{\eta}_l}{\langle \omega \rangle} [\sin(\phi + \langle \omega \rangle \tau) - \sin \phi] \quad (38)$$

Note that this z-motion includes an oscillation at the gyrofrequency $\langle \omega \rangle$. As a result, the particle never attains perfect resonance with a single Fourier mode and its interaction with the fluctuations is on the average diminished.

To see this semi-quantitatively, we note that by approximately integrating Eq. (22) $\Delta\mu$ is given as

$$\Delta\mu \simeq - (1 - \mu^2)^{1/2} \int_0^t d\tau \int_{-\infty}^{+\infty} dk \delta\omega_k \sin(\phi + \langle \omega \rangle \tau) \exp\left[\frac{i k p}{\gamma m} \left\{ \mu \tau - \frac{(1 - \mu^2)^{1/2} \tilde{\eta}_l}{\langle \omega \rangle} [\sin(\phi + \langle \omega \rangle \tau) - \sin \phi] \right\} \right] \quad (39)$$

Equation (39) differs from its quasi-linear analogue only in the fact that the oscillatory component of the z-motion is included, $\tilde{\eta}$ in (38) has been approximated by η and the η contribution to the rectilinear motion has been discarded.

It is now straightforward to calculate, using Eq. (26), the resonance contribution to $D_{\mu\mu}$. We use (22) for $\dot{\mu}$, (39) for $\Delta\mu$, assume no correlation between different Fourier modes (the random phase approximation), and employ the expansion

$$\exp(-i\alpha \sin \phi) = \sum_{m=-\infty}^{+\infty} \exp(-im\phi) J_m(\alpha)$$

where the J_m 's are the ordinary Bessel Functions. The result obtained for the exponential correlation function Eq. (28) is

$$D_{\mu\mu}^{\text{res}} = \frac{1-\mu^2}{2} |\mu| \eta^2 \epsilon \sum_{s=1}^{\infty} \frac{1}{\mu^2 + s^2 \epsilon^2} \left\{ J_{s+1} \left[\frac{s(1-\mu^2)^{1/2}}{|\mu|} \eta \right] - J_{s-1} \left[\frac{s(1-\mu^2)^{1/2}}{|\mu|} \eta \right] \right\}^2 \quad (40)$$

The quasi-linear result is recovered by letting $\eta \rightarrow 0$ inside of the summation, only the $s = 1$ term survives and Eq. (29) readily follows.

It is not obvious from Eq. (40) that the resonance interruption results in a diminished diffusion coefficient. We have, however, evaluated numerically as a function of μ , the ratio

$$(\mu^2 + \epsilon^2) \sum_{s=1}^{\infty} \frac{1}{\mu^2 + s^2 \epsilon^2} \left\{ J_{s+1} \left[\frac{s(1-\mu^2)^{1/2}}{|\mu|} \eta \right] - J_{s-1} \left[\frac{s(1-\mu^2)^{1/2}}{|\mu|} \eta \right] \right\}^2$$

for several values of η and ϵ . Its value is indeed ≤ 1 and its general behavior is quite similar to that of $D_{\mu\mu}^{PA}/D_{\mu\mu}^{QL}$ in Figs. 1-3 in the $\mu > \eta$ range. This finding is the basis of our interpretation that resonance interruption is the reason that $D_{\mu\mu}^{PA}$ falls beneath $D_{\mu\mu}^{QL}$ in this range.

(c) Results for the Oblique Slab Model

We have also evaluated $D_{\mu\mu}^{PA}$ using Eqs. (22), (26), and (27) for the situation where

$$\langle \delta \mathbf{B} \rangle_{\mu\mu} = \delta B(s) C(s - s') \hat{e}_x \quad (41)$$

The correlation C is again exponential (cf. Eq. (28)) but in this instance

$$s = z \cos \alpha + y \sin \alpha$$

so that planes of constant $|\langle \delta \mathbf{B} \rangle|$ are perpendicular to the y - z plane and inclined at angle α with respect to the z -axis. (The perpendicular slab model considered in IIIa corresponds to $\alpha = 0$). Since the final averaging is over an assumed spatially homogeneous distribution of $\delta B^2(s)$, the position s is again irrelevant and we may without loss of generality in (41) take $z = y = s = 0$. The processes of averaging numerically over t and $\delta B(s)$ are exactly as detailed in IIIa.

We display in Figure 5 results for $D_{\mu\mu}^{PA}$ and $D_{\mu\mu}^{QL}$ for the case $\epsilon = 1.0$, $\eta = 0.1$, and $\alpha = 30^\circ$. Figure 5 is thus the counterpart of Figure 2 for the oblique situation. The obvious difference between the oblique and perpendicular models is the strong suppression of diffusion at small μ 's which non-perpendicularity introduces. This is true for both $D_{\mu\mu}^{PA}$ and $D_{\mu\mu}^{QL}$. In addition, non-perpendicularity apparently leads to enhancements of $D_{\mu\mu}$ in the region of maximum around $\mu = 0.5$. The discrete values of μ for which $D_{\mu\mu}$'s were explicitly evaluated are indicated in Fig. 5. The

distribution and number of such points is dictated by the length of the numerics involved in calculating $D_{\mu\mu}^{\text{PA}}$. For comparison we have also calculated $D_{\mu\mu}^{\text{QL}}$ at the same μ values. It is probable that the relatively small number of points in the region $\mu \cong 0$ leads to the suppression of small amplitude oscillations which characterize the higher harmonic (of $\langle \epsilon \rangle$) interaction of particles and magnetic perturbations with spatial structure perpendicular to $\langle \mathbf{B} \rangle$. It is known that such oscillations characterize $D_{\mu\mu}^{\text{QL}}$ ²⁶ and it is probable that they would also appear if $D_{\mu\mu}^{\text{PA}}$ were calculated sufficiently densely in μ . Our attitude is that if these oscillations have the same amplitude in $D_{\mu\mu}^{\text{PA}}$ as they have in $D_{\mu\mu}^{\text{QL}}$ they are inconsequential by comparison with the magnitude of $D_{\mu\mu}^{\text{PA}}$ ($\mu \cong 0$).

As we shall see in Section IV, the spatial diffusion coefficient κ_{\parallel} parallel to $\langle \mathbf{B} \rangle$ varies inversely with $D_{\mu\mu}$. The important inference to be gained therefore from Fig. 5 is that the oblique slab situation is more conducive to parallel spatial transport than is the perpendicular slab. Extrapolating to the case of the interplanetary medium, we expect that quasi-planar magnetic structures propagating radially outward from the sun undergo transition from a perpendicular slab-like situation to the oblique slab case as the idealized $\langle \mathbf{B} \rangle$ develops spirally.²² Given constant η and ϵ , κ_{\parallel} should therefore ideally decrease with approaching distance to the sun.

IV. Spatial Transport and the Scaling of $D_{\mu\mu}^{\text{PA}}$ ($\mu \cong 0$)

Let us assume that pitch angle diffusion theory is true locally and that $D_{\mu\mu}$ is properly given by partially averaged field theory. We re-introduce a spatial gradient in the z-direction with scale much longer than z_c , and suppressing the suffix "0" rewrite Eq. (23) as

$$\frac{\partial \langle f \rangle}{\partial t} + \mu v \frac{\partial \langle f \rangle}{\partial z} = \frac{\partial}{\partial \mu} D_{\mu\mu}^{PA}(z, v, \mu) \frac{\partial \langle f \rangle}{\partial \mu} \quad (42)$$

Contact with the macroscopic variables – density n , flux Φ etc. – is made by taking moments of Eq. (42) with respect to μ . Particle conservation

$$\begin{aligned} \frac{\partial n}{\partial t} &= - \int_{-1}^{+1} d\mu \mu v \frac{\partial \langle f \rangle}{\partial z} = - v \frac{\partial}{\partial z} \int_{-1}^{+1} d\mu \mu \langle f \rangle \\ &= - v \frac{\partial}{\partial z} \Phi \end{aligned} \quad (43)$$

follows exactly by taking the zeroth moment and noting that there can be no diffusive flux $D_{\mu\mu}^{PA} \partial \langle f \rangle / \partial \mu$ at the boundaries $\mu = \pm 1$ in μ space.

There exist a variety of methods for expressing $\Phi \equiv \int_{-1}^{+1} d\mu \mu \langle f \rangle$ in terms of n and thus obtaining a closed equation for n . The standard transport theory method²⁷ is to expand $\langle f \rangle$ in Legendre Polynomials and truncate the series after $P_1(\mu)$. This yields a standard diffusion equation for n

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial z} \kappa_{\parallel}^L \frac{\partial n}{\partial z} \quad (44)$$

where

$$\kappa_{\parallel}^L = \frac{\frac{2}{9} v^2}{\int_{-1}^{+1} d\mu D_{\mu\mu}(\mu)} \quad (45)$$

Alternatively one could expand $\langle f \rangle$ in eigenfunctions of the diffusion operator $\partial/\partial\mu D_{\mu\mu} \partial/\partial\mu$ and obtain an alternate expression for κ_{\parallel}^E ²⁸:

$$\kappa_{\parallel}^E = \sum_{i = \text{odd}} \tau_i V_{0i}^2 \quad (46)$$

where the τ_i are the eigenvalues of the equation

$$\frac{\partial}{\partial \mu} D_{\mu\mu} \frac{\partial}{\partial \mu} R_i(\mu) = -\frac{1}{\tau_i} R_i(\mu) \quad (47)$$

and¹²

$$V_{ij} = v \int_{-1}^1 \mu R_i(\mu) R_j(\mu) d\mu \quad (48)$$

Due to the difficulty in obtaining solutions to (47), however, the usefulness of this approach is limited.

A third approach, due to Jokipii,¹ is to solve the approximate equation

$$\mu v \frac{\partial}{\partial z} \langle f \rangle = \frac{\partial}{\partial \mu} D_{\mu\mu} \frac{\partial}{\partial \mu} \langle f \rangle \quad (49)$$

by inverting the diffusion operator to obtain

$$\langle f \rangle = -\frac{v}{4} \int_0^\mu d\mu' (1 - \mu'^2) / D_{\mu\mu}(\mu') \frac{\partial n}{\partial z} \quad (50)$$

where we further approximate

$$\frac{\partial \langle f \rangle}{\partial z} \approx \frac{1}{2} \frac{\partial n}{\partial z}$$

Inverting (50) into (43) yields (44) once again but with

$$\kappa_{\parallel}^J = \frac{v^2}{8} \int_{-1}^1 \frac{d\mu (1 - \mu^2)^2}{D_{\mu\mu}(\mu)}. \quad (51)$$

In all of the above methods the time derivatives of all of the components of $\langle f \rangle$ with the exception of the isotropic one, n , are neglected when compared to the streaming and diffusive terms.

Earl¹² has compared the three approaches extensively and has concluded that the third one is the correct approach to employ. While we hesitate to call any approximation scheme "correct," we do concur that the third method has many arguments in its favor and is the result of a fully systematic approximation scheme.

However, it is known that if $D_{\mu\mu}(\mu) \propto (1 - \mu^2)$ representing isotropic scattering, all three methods lead to the identical value for κ_{\parallel} . Therefore if the form of $D_{\mu\mu}(\mu)$ is not too different from that of isotropic scattering one would not expect it to matter very much which approach was used and convenience should dictate, especially in the field of astrophysics where the quantitative input is usually of a somewhat approximate nature.

We have computed values of κ_{\parallel} by both the Legendre expansion method (κ_{\parallel}^L) and the third or inversion method (κ_{\parallel}^J) for the four sets of parameters for which we have complete curves of $D_{\mu\mu}^{PA}(\mu)$. The results are shown in Table 1.

As would be expected, the differences are most pronounced where $D_{\mu\mu}^{PA}(\mu = 0)$ is the smallest. The case $\epsilon = 1$, $\eta = .1$ for the perpendicular slab may also be compared with the results of Goldstein²⁹ who calculated $D_{\mu\mu}$ by an independent non-linear method¹⁵ and used Eqs. (45) and (51) to calculate κ_{\parallel}^L and κ_{\parallel}^J (Goldstein uses the symbol κ^D rather than our κ_{\parallel}^J .) Goldstein obtains $\kappa_{\parallel}^L = 115$ and $\kappa_{\parallel}^J = 165$. The near identity of κ_{\parallel}^L with our results reflects the insensitivity of κ_{\parallel}^L (cf. Eq. 45) to $D_{\mu\mu}(\mu = 0)$, while his larger κ_{\parallel}^J is the result of the fact that his non-linear method leads to a smaller $D_{\mu\mu}$ than ours in the region $\mu \approx 0$ (cf. Eq. 51).

Since κ_{\parallel} depends inversely on the magnitude of $D_{\mu\mu}$ a knowledge of how $D_{\mu\mu}$ scales with the parameters ϵ and η would be of use in estimating the way κ_{\parallel} varies with the power in the fluctuating fields and particle rigidity. Since

κ_{\parallel} depends on the shape of $D_{\mu\mu}$ as a function of μ as well as the overall magnitude a simple scaling will not tell the whole story. However, since the value of κ_{\parallel} in Eq. (51) depends rather sensitively on the value of $D_{\mu\mu}$ at $\mu = 0$ we have chosen this value as the scaling quantity of interest. We have therefore computed $D_{\mu\mu}^{PA}(\mu = 0)$ for the nine different cases resulting from the combination of the 3 ϵ -values .25, 1, and 4 with the 3 η -values .05, .1, and .3, and we have attempted from these 9 cases to fit $D_{\mu\mu}^{PA}(\mu = 0)$ to the form

$$D_{\mu\mu}^{PA}(\mu = 0) \sim \tau_{\eta}^m \epsilon^q$$

For each value of ϵ we determined that m which gave best fit to the theoretically determined values for the 3 η 's. The linearity of $\log [D_{\mu\mu}^{PA}(\mu = 0)]$ vs $\log \tau_{\eta}$ is excellent for each value of ϵ . The optimum m -values are 2.24, 2.30, and 2.34 for the cases $\epsilon = .25, 1$ and 4 respectively.

Similarly we determined q 's holding each of the 3 values of τ_{η} fixed. The linearity of $\log [D_{\mu\mu}^{PA}(\mu = 0)]$ vs $\log \epsilon$ at constant τ_{η} is again excellent. The optimum q -values are .58 ($\eta = .05$), .60 ($\eta = .1$) and .64 ($\eta = .3$). Combining fits we propose the empirical scaling

$$D_{\mu\mu}^{PA}(\mu = 0) \sim \tau_{\eta}^{2.3} \epsilon^{.6} \quad (52)$$

A similar effort to scale the values of $D_{\mu\mu}$ obtained from the Monte Carlo simulations described in the companion paper yields²⁰

$$D_{\mu\mu}^{MC}(\mu = 0) \sim \tau_{\eta}^{2.7} \epsilon^{.3} \quad (53)$$

It is difficult to say whether the difference between the indices in Eqs. (52) and (53) is significant. But it is interesting to note that if one adopts the heuristic,

mirroring picture described in Section III one comes up with a scaling

$$D_{\mu\mu}(\mu = 0) \sim \eta^{2.5} \epsilon^{.5}$$

intermediate between the two.

V. Range of Validity of Partially Averaged Field Theory

All theories, whether quasilinear or non-linear, compute the effects of the random force field on a collection of particles by summing its instantaneous effects along some particular particle orbit or collection of orbits until the random force on the particle becomes self-incoherent. In the quasilinear theory the orbit employed is the one the particle would have if the random part of the force were absent and we shall investigate how rapidly the actual orbit deviates in a significant way from this approximate orbit. In the following we shall continue to limit ourselves to the slab model of Section III. This limitation is for the sake of illustration only and does not limit the validity of our conclusions.

The quasilinear orbit is characterized by $z = \mu vt$, $\mu = \text{const.}$ but in fact, μ will be changed by the random field by an amount

$$\delta\mu \approx \delta\omega t, \quad (54)$$

where in this section $\delta\omega$ refers to the RMS value. If $\delta\omega \gtrsim \langle \omega \rangle$ then $\delta\mu$ will become of order unity in a time τ_D which we shall call the deflection time where

$$\tau_D = (\delta\omega)^{-1} \text{ for } \delta\omega \gtrsim \langle \omega \rangle \quad (55)$$

If $\delta\omega \ll \langle \omega \rangle$ then $\delta\mu$ will oscillate with a frequency $\langle \omega \rangle$ and an amplitude of order $\delta\omega / \langle \omega \rangle = \tau_f$. However, due to the random scattering of the pitch angle, the deflection in μ will grow in time as

$$\delta\mu \approx [D_{\mu\mu} t]^{1/2} \quad (56)$$

where $D_{\mu\mu}$ is the pitch angle diffusion coefficient calculated by the theory in question. In this case the deflection time is given by

$$\tau_D = (D_{\mu\mu})^{-1} \text{ for } \delta\mu \ll \langle \mu \rangle \quad (57)$$

Clearly the random force must become incoherent in a time short compared to τ_D but there is another time of interest. The deflection time is a measure of how rapidly the particle orbit becomes unreliable in velocity space but the orbit can also become unreliable in coordinate space (in this model the z coordinate is the only one that matters). If μ is uncertain by an amount $\delta\mu(t)$ the z coordinate will be uncertain by an amount

$$\delta z = v \int_0^t \delta\mu(t') dt' \quad (58)$$

$$= \frac{1}{2} v \delta\mu t^2 \text{ for } \delta\mu \gg \langle \mu \rangle$$

or

$$= vt \delta\mu / \langle \mu \rangle \text{ for } \delta\mu \ll \langle \mu \rangle$$

If this uncertainty in coordinate space grows to be of the order of the correlation length z_c which is the characteristic scale of variation of the random field, the unperturbed orbit becomes quite useless. The time in which this occurs is called the deviation time and is given by

$$\tau_d = \left(\frac{z_c}{v \delta\mu} \right)^{1/2} \text{ for } \delta\mu \geq \langle \mu \rangle$$

$$= \frac{z_c \langle \mu \rangle}{v \delta\mu} \text{ for } \delta\mu \ll \langle \mu \rangle. \quad (59)$$

We further require that this deviation time be long compared to the time required for the random force to become self-incoherent.

Just as we have defined two time scales over which the approximate particle orbit becomes unreliable we may define two coherence time scales, one for configuration space and one for velocity space. This is possible because the random force $\delta F = (q/c) \mathbf{v} \times \delta B(z)$ depends on both z and v . Clearly the shorter of these two times is the time over which δF becomes incoherent.

The correlation length z_c is defined as the length scale over which the random field δB becomes self-incoherent, it is therefore natural to define the correlation time τ_c as the time required for the particle to travel one correlation length along its unperturbed orbit. In quasi-linear theory this is given by

$$\tau_c = \frac{z_c}{\mu v}. \quad (60)$$

We may further define the coherence time τ_c^v as the time required for the velocity of the particle to become self-incoherent as it travels along its unperturbed orbit. In quasi-linear theory the orbit is fixed for all realizations of the ensemble so the particle's velocity has no stochastic component and hence never becomes self-incoherent. Therefore obviously in quasi-linear theory $\tau_c^v = \infty$ and τ_c is the determining coherence scale. This is not the case in PAF theory.

We have defined two time scales over which the orbit used in calculations becomes unreliable, one for configuration space and one for velocity space. In the same vein we have defined two time scales over which the random force becomes self-incoherent, again one for configuration space and one for velocity space. In Table 2 we list these four time scales for future reference. We require of any theory that

$$\text{shorter of } \tau_c, \tau_c' < \text{shorter of } \tau_D, \tau_d. \quad (61)$$

In Table 2 we have listed the symbol, name and meaning of each of the four time scales that we have defined here.

We turn now to a comparison of the ratios $\tau_c/\tau_D, \tau_c/\tau_d$ etc. as given by quasi-linear theory and PAF theory. In the interest of saving space we shall not write down all of the time scale ratios that can be formed although it is a simple matter to do so. We shall rather only explore the consequences of the constraint (61).

In quasi-linear theory the requirement $\tau_c < \tau_D$ leads to the constraint

$$\mu > \frac{Z_c \omega}{v} \quad \text{for } \omega \gtrsim \omega_c$$

or

$$\mu > \left(\frac{Z_c}{v} \right) D_{\text{th}}(\omega) \quad \text{for } \omega \ll \omega_c \quad (62)$$

For strong random fields this condition cannot be met for all (if any) values of μ but for weak random fields the condition can be met if the power spectrum goes as k^{-2} or steeper for high frequencies (i.e. $D_{\text{th}}(\omega)$ goes to zero with ω linearly or faster).

If, however, one turns to the requirement that $\tau_c < \tau_d$ one finds for $\omega \gtrsim \omega_c$

$$\mu > \left(\frac{Z_c \omega}{v} \right)^{1/2} \quad (63)$$

not much different than before but for $\omega \ll \omega_c$ the requirement is

$$\mu > \omega \quad (64)$$

a condition that cannot be met for all values of μ . We see therefore that quasilinear theory will always have trouble as $\mu \rightarrow 0$.

The various non-linear theories and PAF theory in particular were introduced to obviate this problem in the vicinity of $\mu = 0$. This is achieved in PAF theory by allowing the fluctuating part of the field to influence the orbit to lowest order. If $\omega \ll \langle \omega \rangle$ this influence may be approximated

$$z(t) \approx [\mu^2 + \omega^2 + \langle \omega \rangle^2]^{1/2} vt \quad (65)$$

and hence

$$\tau_c' = \frac{\tau_c}{[\mu^2 + \omega^2 + \langle \omega \rangle^2]^{1/2} v} \quad (66)$$

Since this is equivalent to replacing μ with an effective $\mu_{\text{eff}} = [\mu^2 + \omega^2 + \langle \omega \rangle^2]^{1/2}$ a glance at (62) and (64) indicates that the troubles at $\mu = 0$ are avoided.

However, for the case of strong fluctuating fields, $\omega \gtrsim \langle \omega \rangle$ this substitution of μ_{eff} for μ is not a valid approximation. In fact for certain configurations of the fluctuating fields (other than the slab model) the particles can be trapped in the vicinity of the origin for an indefinitely long time and τ_c' can become essentially infinite. In this situation we are saved by the fact that the coherence time τ_c' which was infinite in quasilinear theory is finite in PAF theory and can be the shorter of the correlation/coherence times.

In PAF theory the velocity vector is turned not at the fixed frequency $\langle \omega \rangle$ but at the frequency characteristic of the partially averaged field

$$\omega = \frac{q}{mc} [\langle B \rangle^2 + B^2]^{1/2}. \quad (67)$$

During final averaging this frequency will be spread over an amount $\Delta\omega$ where

$$\Delta\omega \approx \delta\omega \text{ for } \delta\omega \gtrsim \langle \omega \rangle$$

$$\Delta\omega \approx \delta\omega^2 / \langle \omega \rangle \text{ for } \delta\omega \ll \langle \omega \rangle \quad (68)$$

The velocity vector will thus become self-incoherent in a time τ_c' given by

$$\begin{aligned} \tau_c' &= (\Delta\omega)^{-1} = \delta\omega^{-1} \text{ for } \delta\omega \gtrsim \langle \omega \rangle \\ &= \langle \omega \rangle / \delta\omega^2 \text{ for } \delta\omega \ll \langle \omega \rangle. \end{aligned} \quad (69)$$

We now find for the various ratios

$$\begin{aligned} \tau_c' / \tau_D &= 1 \quad \delta\omega \gtrsim \langle \omega \rangle \\ &= \langle \omega \rangle / D_{\mu\mu} \delta\omega^2 \quad \delta\omega \ll \langle \omega \rangle \end{aligned}$$

$$\tau_c' / \tau_d = \left(\frac{v}{Z_c \delta\omega} \right)^{1/2} \quad \delta\omega \gtrsim \langle \omega \rangle$$

$$\frac{v}{Z_c \delta\omega} \quad \delta\omega \ll \langle \omega \rangle$$

Since as $v \rightarrow 0$, $D_{\mu\mu} \propto \frac{v^2}{\langle \omega \rangle} \left(\frac{v}{\langle \omega \rangle Z_c} \right)^2$

where $0 < \nu$, the only ratio that does not go to zero as $v \rightarrow 0$ is τ_c' / τ_D in the strong field case. So we see that even in the case of strong random fields and low rigidity particles there exists an effect that destroys the coherence of the random force at least on the same time scale that the orbit becomes unreliable.

Non-linear theories represent an improvement over quasi-linear theory for one primary reason: they employ a more realistic description of particle orbits

in determining the effects of the fluctuating fields on the particles themselves. We have seen that one of the effects of this description is the limiting of the correlation time to a more realistic value such that the resulting calculation remains of marginal validity for a much wider range of parameters. This is true for PAF theory and would presumably be true for any non-linear theory that correctly describes the loss of coherence over the ensemble of the particle's velocity vector.

APPENDIX A
 Conditional Averaging for Multi-Dimensional
 Gaussian Processes

Consider a six dimensional vector \underline{x} made up of two three dimensional vectors \underline{x}_a and \underline{x}_b ,

$$\underline{x} = (\underline{x}_a, \underline{x}_b) \quad (A1)$$

We wish to find the average value of \underline{x}_b under the condition that $\underline{x}_a = \underline{x}_{ay}$. If the joint probability density for \underline{x}_a and \underline{x}_b is given as a function of \underline{x} as $P(\underline{x})$ we have

$$\langle \underline{x}_b \rangle_{\underline{x}_a = \underline{x}_{ay}} = \int d^6 \underline{x} \delta(\underline{x}_a - \underline{x}_{ay}) \underline{x}_b P(\underline{x}) / P(\underline{x}_{ay}) \quad (A2)$$

Note that

$$\begin{aligned} \delta(\underline{x}_b) &= -i \frac{\delta}{\delta \underline{k}_b} e^{i \underline{k}_b \cdot \underline{x}_b} \Big|_{\underline{k}_b = 0} \\ &= -i \frac{\delta}{\delta \underline{k}_b} e^{i \underline{k} \cdot \underline{x}} \Big|_{\underline{k} = 0} \end{aligned}$$

where \underline{k} like \underline{x} is six dimensional. Further, writing the delta function as a Fourier integral

$$\begin{aligned} \delta(\underline{x}_a - \underline{x}_{ay}) &= \frac{1}{(2\pi)^3} \int d^3 \underline{k}'_a e^{i \underline{k}'_a \cdot (\underline{x}_a - \underline{x}_{ay})} \\ &= \frac{1}{(2\pi)^3} \int d^3 \underline{k}'_a e^{i \underline{k}' \cdot (\underline{x} - \underline{x}_{ay})} \Big|_{\underline{k}'_b = 0} \end{aligned}$$

where \underline{k}' is also six dimensional, we may insert these expressions into (A2) to obtain

$$\langle \underline{x}_b \rangle_{\underline{x}_a = \bar{x}_a} = \frac{-i}{(2\pi)^3} \frac{\int d^3 \underline{k}'_a e^{-i \underline{k}'_a \cdot \bar{x}} \frac{\partial}{\partial \underline{k}_b} \left[\int d^6 \underline{x} e^{i \underline{\kappa} \cdot \underline{x}} P(\underline{x}) \right] \Big|_{\underline{k} = \underline{k}' = 0}}{P(\bar{x}_a)} \quad (A3)$$

where $\underline{\kappa} = \underline{k} + \underline{k}'$.

The quantity in square brackets is the characteristic function of the distribution function $P(\underline{x})$

$$\text{Characteristic fn. } \tilde{P}(\underline{\kappa}) = \langle e^{i \underline{\kappa} \cdot \underline{x}} \rangle$$

If \underline{x} is a Gaussian process the distribution $P(\underline{x})$ is a multivariate Gaussian³¹

$$P(\underline{x}) = (2\pi)^{-3} |M|^{-1} \exp \left(-\frac{1}{2} \underline{x} \cdot \underline{M}^{-1} \cdot \underline{x} \right) \quad (A4)$$

Where \underline{M}^{-1} is the inverse of the covariance matrix \underline{M} , i.e., $M_{ij} = \langle x_i x_j \rangle$ an obviously symmetric matrix. In complete analogy with a one dimensional Gaussian distribution it is straightforward to show that

$$\tilde{P}(\underline{\kappa}) = \exp \left(-\frac{1}{2} \underline{\kappa} \cdot \underline{M} \cdot \underline{\kappa} \right). \quad (A5)$$

If we now break \underline{M} up into the three by three matrices $\underline{S} = \langle \underline{x}_a \underline{x}_a \rangle = \langle \underline{x}_b \underline{x}_b \rangle$ and $\underline{C} = \langle \underline{x}_a \underline{x}_b \rangle$, $\underline{C}^T = \langle \underline{x}_b \underline{x}_a \rangle$ where we have assumed that the marginal distributions for \underline{x}_a and \underline{x}_b are identical we have

$$\underline{M} = \begin{pmatrix} \underline{S} & \underline{C} \\ \underline{C}^T & \underline{S} \end{pmatrix}$$

and $\underline{\kappa} \cdot \underline{M} \cdot \underline{\kappa} = \underline{\kappa}_a \cdot \underline{S} \cdot \underline{\kappa}_a + \underline{\kappa}_b \cdot \underline{S} \cdot \underline{\kappa}_b + \underline{\kappa}_a \cdot \underline{C} \cdot \underline{\kappa}_b + \underline{\kappa}_b \cdot \underline{C}^T \cdot \underline{\kappa}_a$

Thus

$$\begin{aligned}
\left. \frac{\partial}{\partial \underline{k}_b} \tilde{P}(\underline{\kappa}) \right|_{\underline{k} = \underline{k}_b'} = 0 &= -\frac{1}{2} \left(\frac{\partial}{\partial \underline{k}_b} \underline{\kappa} \cdot \underline{M} \cdot \underline{\kappa} \right) \tilde{P}(\underline{\kappa}) \Big|_{\underline{k} = \underline{k}_b'} = 0 \\
&= -(\underline{\kappa}_b \cdot \underline{S} + \underline{\kappa}_a \cdot \underline{C}) \tilde{P}(\underline{\kappa}) \Big|_{\underline{k} = \underline{k}_b'} = 0 \\
&= -\underline{k}_a' \cdot \underline{C} \exp \left(-\frac{1}{2} \underline{k}_a' \cdot \underline{S} \cdot \underline{k}_a' \right).
\end{aligned} \tag{A6}$$

We have now

$$\begin{aligned}
\langle \underline{x}_b \rangle_{\underline{x}_a = \underline{\bar{x}}_a} &= \left(\frac{i}{(2\pi)^3} \frac{\int d^3 \underline{k}'_a e^{-i \underline{k}'_a \cdot \underline{\bar{x}}_a} e^{\frac{1}{2} \underline{k}'_a \cdot \underline{S} \cdot \underline{k}'_a} \underline{k}'_a \cdot \underline{C}}{P(\underline{\bar{x}}_a)} \right) \\
&= -\frac{\left(\frac{\partial}{\partial \underline{\bar{x}}_a} \frac{1}{(2\pi)^3} \int d^3 \underline{k}'_a e^{-i \underline{k}'_a \cdot \underline{\bar{x}}_a} \tilde{P}(\underline{k}'_a) \right)}{P(\underline{\bar{x}}_a)} \cdot \underline{C} \\
&= -\frac{\left[\frac{\partial}{\partial \underline{\bar{x}}_a} P(\underline{\bar{x}}_a) \right]}{P(\underline{\bar{x}}_a)} \cdot \underline{C} \\
&= (\underline{\bar{x}}_a \cdot \underline{S}^{-1}) \cdot \underline{C} \\
&= \underline{\bar{x}}_a \cdot (\underline{S}^{-1} \cdot \underline{C})
\end{aligned}$$

APPENDIX B
The Gyrotropy of $\langle f \rangle$

Consider the equation for $\langle f \rangle$

$$\frac{\partial \langle f \rangle}{\partial t} + \langle \omega \rangle \frac{\partial \langle f \rangle}{\partial \phi} = D(\phi) \langle f \rangle \quad (\text{B1})$$

where $D(\phi)$ represents the complete collision operator. If $D(\phi)$ were not a function of the gyro-phase ϕ , we could average this equation over ϕ and remove it from the problem but the ϕ dependence of D and $\langle f \rangle$ will average together to produce an unknown phase correlation.

Let us break $\langle f \rangle$ into a gyrotropic part $\langle f \rangle_0$ and a ϕ dependent part $\langle f \rangle_1$. If we indicate the average over gyrophase of a quantity by placing a bar over it, we further specify that $\overline{\langle f \rangle} = \langle f \rangle_0$ and $\overline{\langle f \rangle_1} = 0$.

Inserting in (B1) and averaging over ϕ gives

$$\frac{\partial \langle f \rangle_0}{\partial t} = \bar{D} \langle f \rangle_0 + \overline{D \langle f \rangle_1} \quad (\text{B2})$$

Subtracting (B2) from (B1) gives

$$\frac{\partial \langle f \rangle_1}{\partial t} + \langle \omega \rangle \frac{\partial \langle f \rangle_1}{\partial \phi} = (D - \bar{D}) \langle f \rangle_0 + D \langle f \rangle_1 - \overline{D \langle f \rangle_1} \quad (\text{B3})$$

We now shall estimate the terms in (B3) as follows

$$\frac{\partial \langle f \rangle_1}{\partial t} \approx \langle \omega \rangle \frac{\partial \langle f \rangle_1}{\partial \phi}; \quad D \langle f \rangle \approx \eta^2 \langle \omega \rangle \langle f \rangle$$

to obtain

$$\frac{\partial \langle f \rangle_1}{\partial t} + \langle \omega \rangle \frac{\partial \langle f \rangle_1}{\partial \phi} \approx \eta^2 \langle \omega \rangle \langle f \rangle_0 \quad (\text{B4})$$

where we now assume $\eta^2 \ll 1$ and we neglect $\eta^2 \langle \omega \rangle \langle f \rangle_1$ compared to $\langle \omega \rangle \langle f \rangle_1$.

Integrating (B4) with $\langle f \rangle_1(t=0) = 0$ gives

$$\begin{aligned} \langle f \rangle_1 &\approx \eta^2 \langle \omega \rangle \int_0^t \exp[-\langle \omega \rangle (t-\tau)] \langle f \rangle_0(\tau) d\tau \\ &\approx \eta^2 \langle f \rangle_0 + O(\eta^4) \end{aligned} \tag{B5}$$

for $t \gg \langle \omega \rangle^{-1}$.

We see from (B5) that $\langle f \rangle_1$ is $O(\eta^2)$ if $\langle f \rangle_0$ is $O(1)$; therefore the second term on the right hand side of (B2) is of $O(\eta^4)$ and may be neglected.

Table 1
 The Parallel Spatial Diffusion Coefficients as Calculated
 Using $D_{\mu\mu}^{PA}$ by the Legendre and Inversion Methods

Case	$\kappa_{\parallel}^L \langle v \rangle / v^2$	$\kappa_{\parallel}^J \langle v \rangle / v^2$
$\epsilon = 1, \eta = .05, \perp$ slab	456	622
$\epsilon = 1, \eta = .1, \perp$ slab	115	137
$\epsilon = 1, \eta = .3, \perp$ slab	13.06	13.48
$\epsilon = 1, \eta = .1, \text{oblique slab}$	162	367

Table 2
Time Scales

Symbol	Name	Meaning
τ_D	Deflection time	Time required for orbit to become unreliable in velocity space
τ_d	Deviation time	Time required for orbit to become unreliable in configuration space
τ_c	Correlation time	Time required for magnetic field to become self-incoherent
τ'_c	Coherence time	Time required for velocity to become self-incoherent

Figure Captions

Figure 1. The diffusion coefficient $D_{\mu\mu}$ for the perpendicular slab model when

$$\epsilon = z_c \cdot m \gamma \langle \mu \rangle / p = 1 \text{ and } \eta = \langle \sigma B^2 \rangle^{1/2} / \langle B \rangle = .05.$$

Figure 2. $D_{\mu\mu}$ vs. μ for the perpendicular slab model when $\epsilon = 1$ and $\eta = .1$

Figure 3. $D_{\mu\mu}$ vs. μ for the perpendicular slab model when $\epsilon = 1$ and $\eta = .3$

Figure 4. The p_x, p_y plane showing regions where particle exits toward $+z$ ($\langle \mu \rangle > 0$), particle exits toward $-z$ ($\langle \mu \rangle < 0$), and where mirroring can possibly occur (indicated by "M?").

Figure 5. $D_{\mu\mu}$ vs. μ for the oblique slab model when $\epsilon = 1$ and $\eta = .1$

References

1. J. R. Jokipii, *Astrophys. J.* 146, 480 (1966).
2. E. C. Roelof, Ph.D. dissertation, University of California, Berkeley, 1966.
3. D. E. Hall and P. A. Sturrock, *Phys. Fluids* 10, 2620 (1967).
4. A. A. Vedenov, E. P. Velikhov, and R. Z. Sagdeev, *Nucl. Fusion* 1, 82 (1961); *Nucl. Fusion Suppl.* 2, 465 (1962); W. E. Drummond and D. Pines, *Nucl. Fusion Suppl.* 3, 1049 (1962).
5. J. R. Jokipii, *Rev. Geophys. and Space Phys.* 9, 27 (1971); *Astrophys. J.* 172, 319 (1972).
6. K. Hasselman and G. Wibberenz, *Zs. f. Geophys.* 34, 353 (1968).
7. C. E. Newman, *J. Math. Phys.* 14, 502 (1973).
8. R. Kulsrud and W. P. Pearce, *Astrophys. J.* 156, 445 (1969).
9. A. N. Kaufman, *Phys. Fluids* 11, 326 (1968).
10. A. J. Klimas and G. Sandri, *Astrophys. J.* 169, 41 (1971).
11. T. B. Kaiser, F. C. Jones, and T. J. Birmingham, *Astrophys. J.* 189, 239 (1973).
12. J. A. Earl, *Astrophys. J.* 188, 379 (1974).
13. P. D. Noerdlinger, *Phys. Rev. Lett.* 20, 1513 (1968).
14. T. H. Dupree, *Phys. Fluids* 9, 1773 (1966); *Phys. Fluids* 10, 1049 (1967); *J. Weinstock, Phys. Fluids.* 11, 1977 (1968).
15. M. L. Goldstein, *Astrophys. J.* 204, 900 (1976).
16. A. J. Owens, *Astrophys. J.* 191, 235 (1974); H. J. Völk, *Astrophys. and Space Sci.* 25, 471 (1973).

17. F. C. Jones, T. B. Kaiser, and T. J. Birmingham, *Phys. Rev. Lett.* 31, 485 (1973).
18. F. C. Jones (to be published).
19. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd edit. (John Wiley, New York, 1968), p. 244.
20. T. B. Kaiser, F. C. Jones, and T. J. Birmingham
21. J. R. Jokipii and P. J. Coleman, *J. Geophys. Res.* 73, 5495 (1968).
22. J. W. Belcher and L. Davis, Jr., *J. Geophys. Res.* 76, 3534 (1971).
23. H. J. Völk and W. Alpers, *Astrophys. and Space Sci.* 20, 267 (1973).
24. A. Barnes, *Phys. Fluids* 9, 1483 (1966).
25. F. C. Jones, T. J. Birmingham, and T. B. Kaiser, *Astrophys. J.* 180, L139 (1973).
26. M. A. Lee and H. J. Völk, *Astrophys. J.* 198, 485 (1975).
27. A. M. Weinberg and E. P. Wigner, *The Physical Theory of Neutron Chain Reactors* (Univ. of Chicago Press, Chicago, 1958), Chapter 9.
28. J. A. Earl, *Astrophys. J.* 193, 231 (1974).
29. M. L. Goldstein, to be published.
30. Unless otherwise noted, all integrals are to run over the entire range of the indicated variables, i.e., infinite integrals.
31. S. O. Rice, *Bell System Tech. J.* 23, 282 (1944); *Bell System Tech. J.* 24, 46 (1945). Both of these papers are reprinted in *Selected Papers on Noise and Stochastic Processes*, ed. N. Wax (Dover Publications, New York, 1954).

EPSILON = 1.00
 ETA = 0.05
 ALPHA = 0°

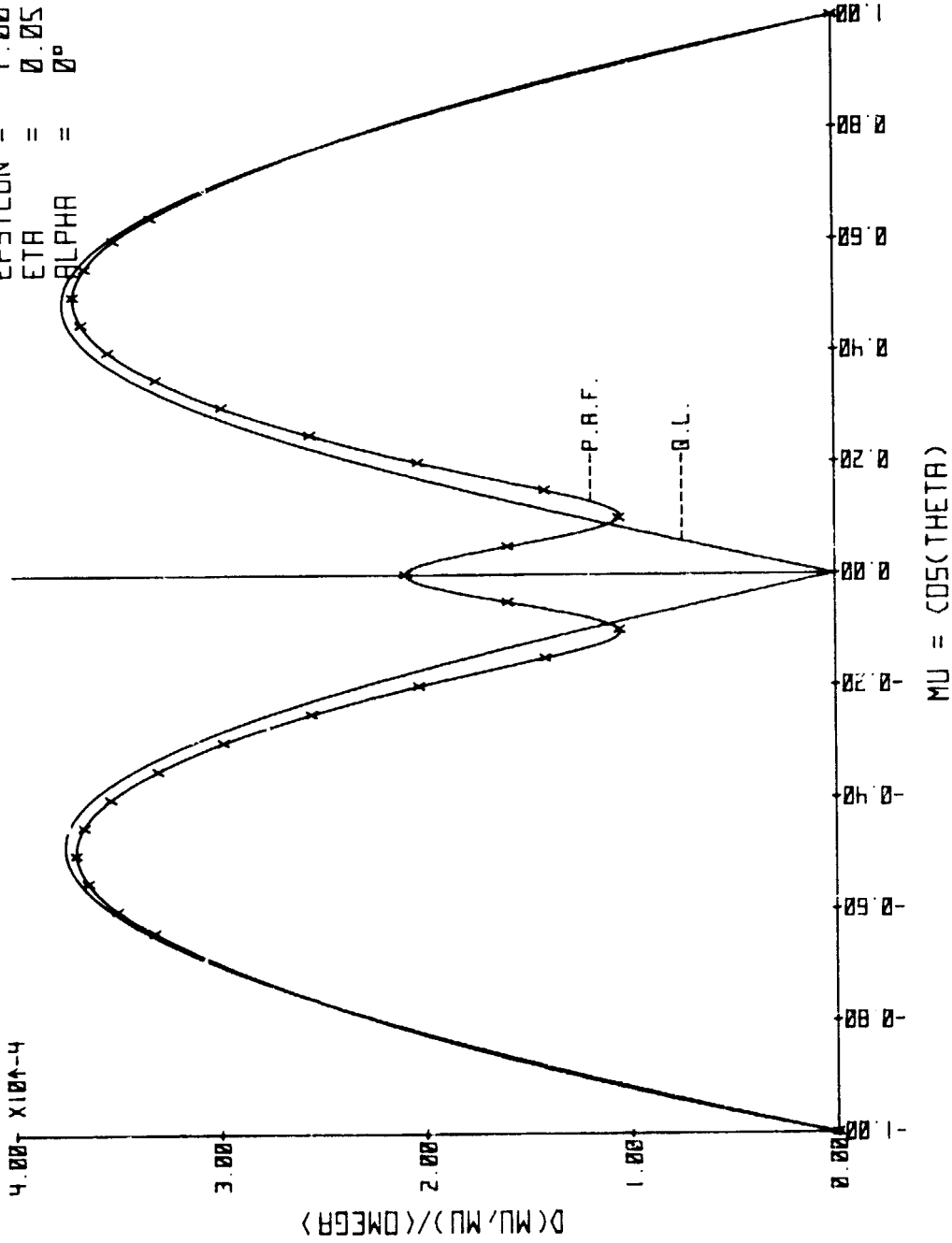


Figure 1. The diffusion coefficient $D_{\mu\mu}$ for the perpendicular slab model when $\epsilon = z_c$, $m\gamma \langle \omega \rangle / p = 1$ and $\eta = \langle \delta B^2 \rangle^{1/2} / \langle B \rangle = .05$

PRECEDING PAGE BLANK NOT FILMED

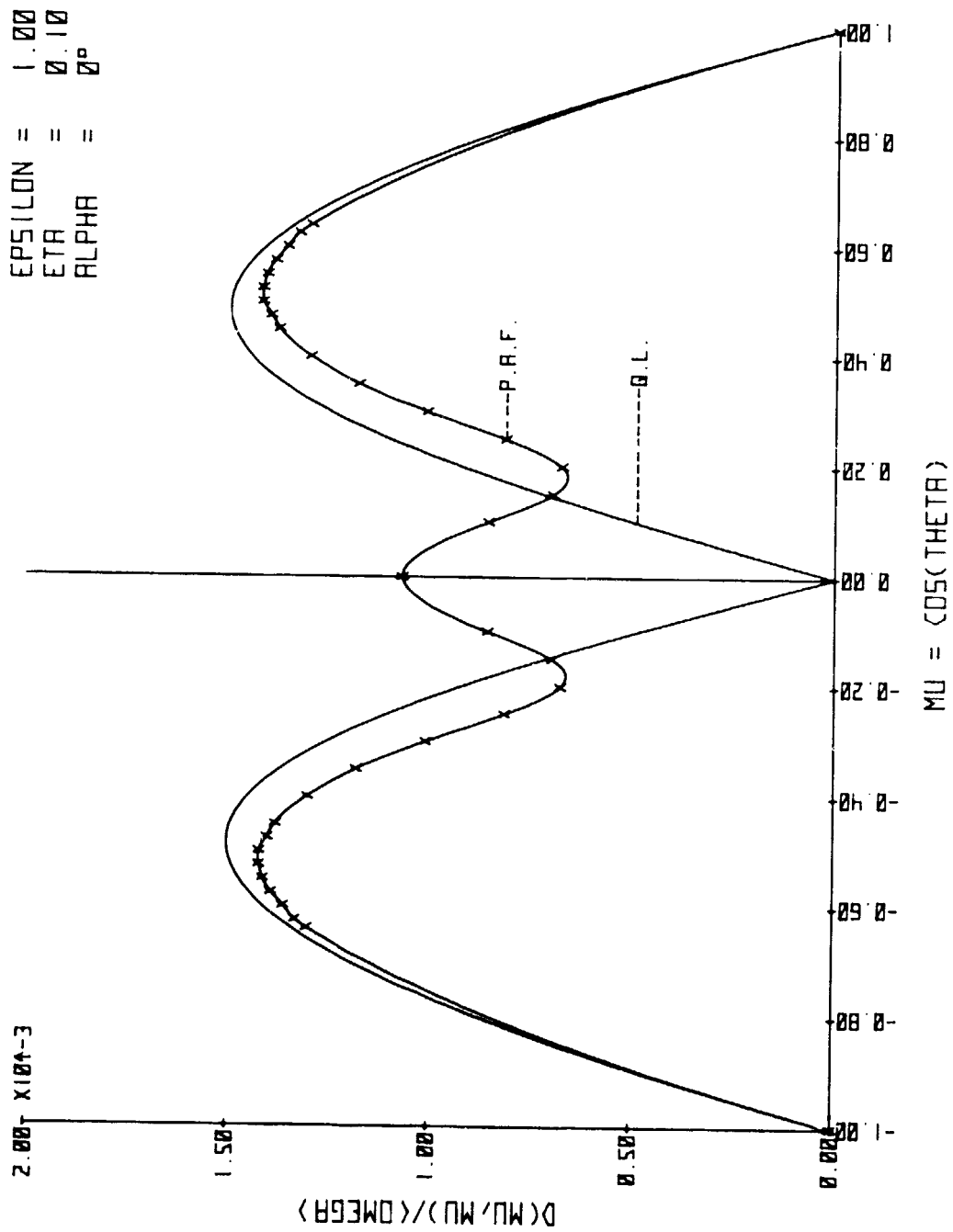


Figure 2. D_{\perp} vs. μ for the perpendicular slab model when $\epsilon = 1$ and $\eta = .1$

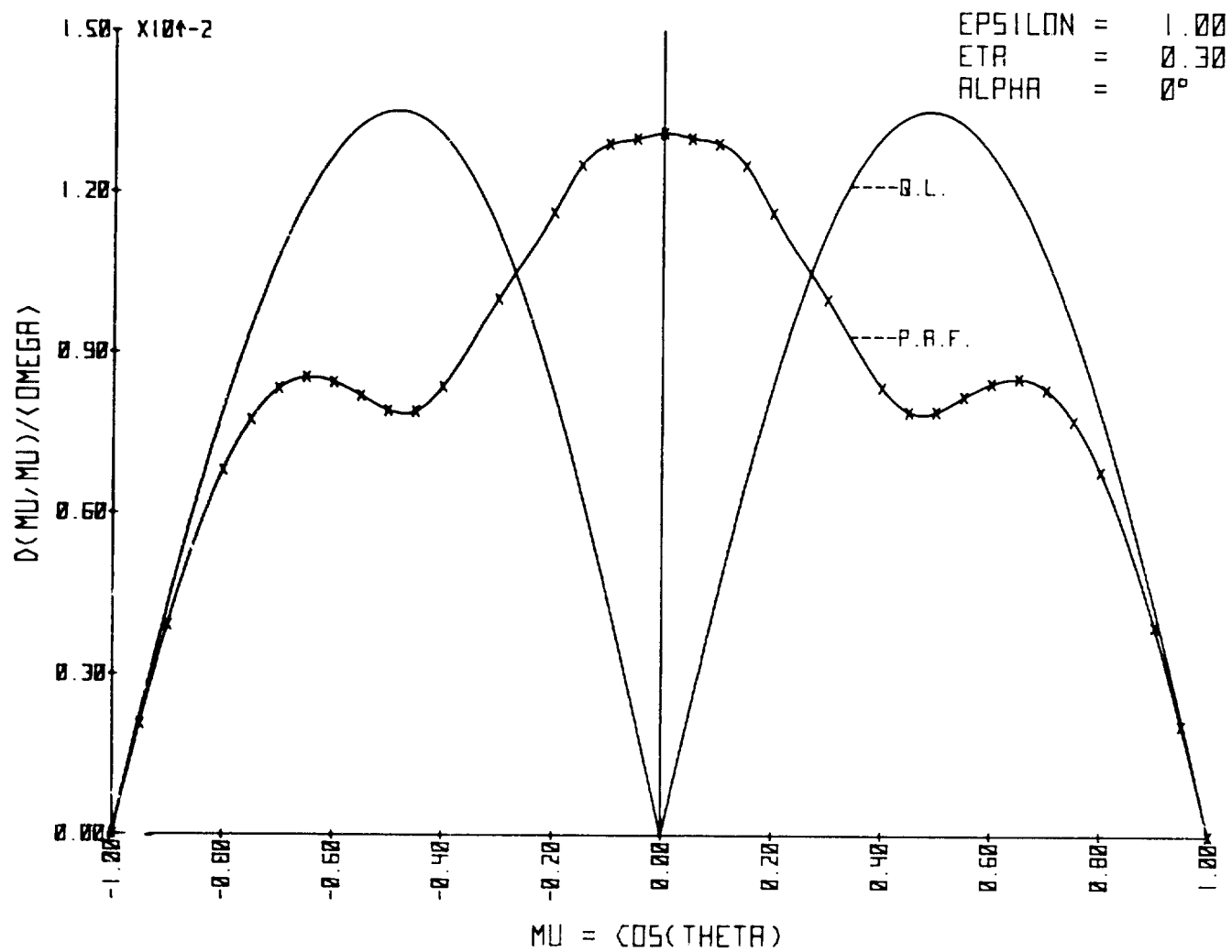


Figure 3. D_{\perp} vs. μ for the perpendicular slab model when $\epsilon = 1$ and $\eta = .3$

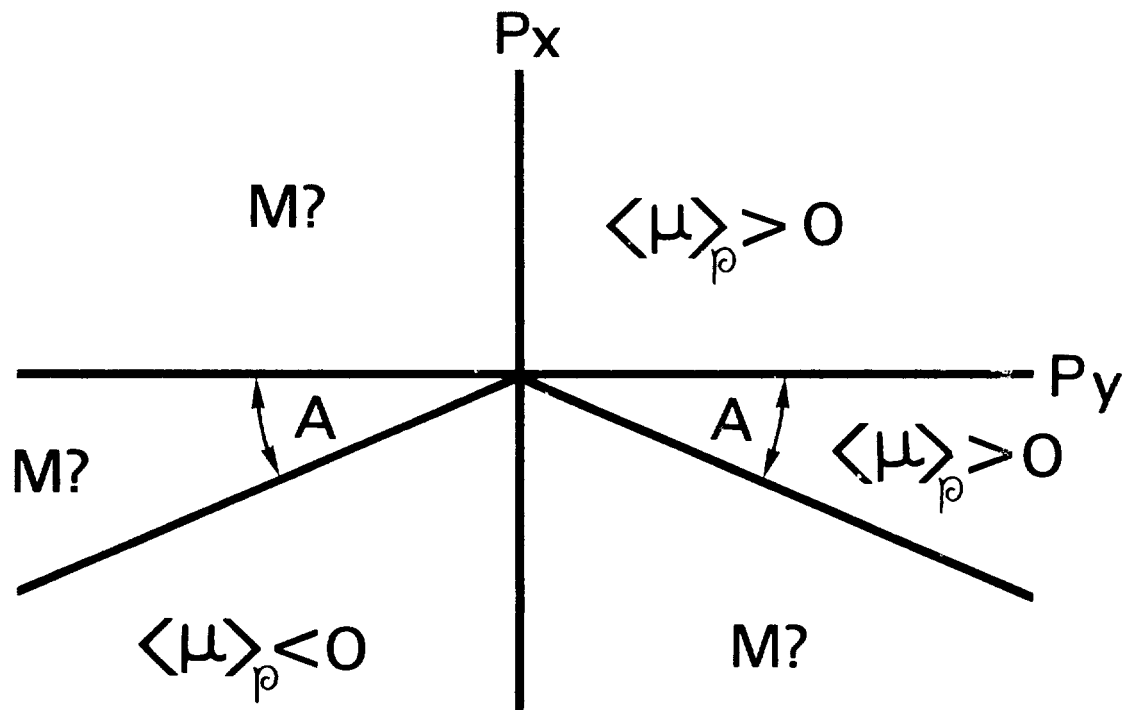


Figure 4. The p_x , p_y plane showing regions where particle exits toward $+z$ ($\langle \mu \rangle_\rho > 0$), particle exits toward $-z$ ($\langle \mu \rangle_\rho < 0$), and where mirroring can possibly occur (indicated by "M?").

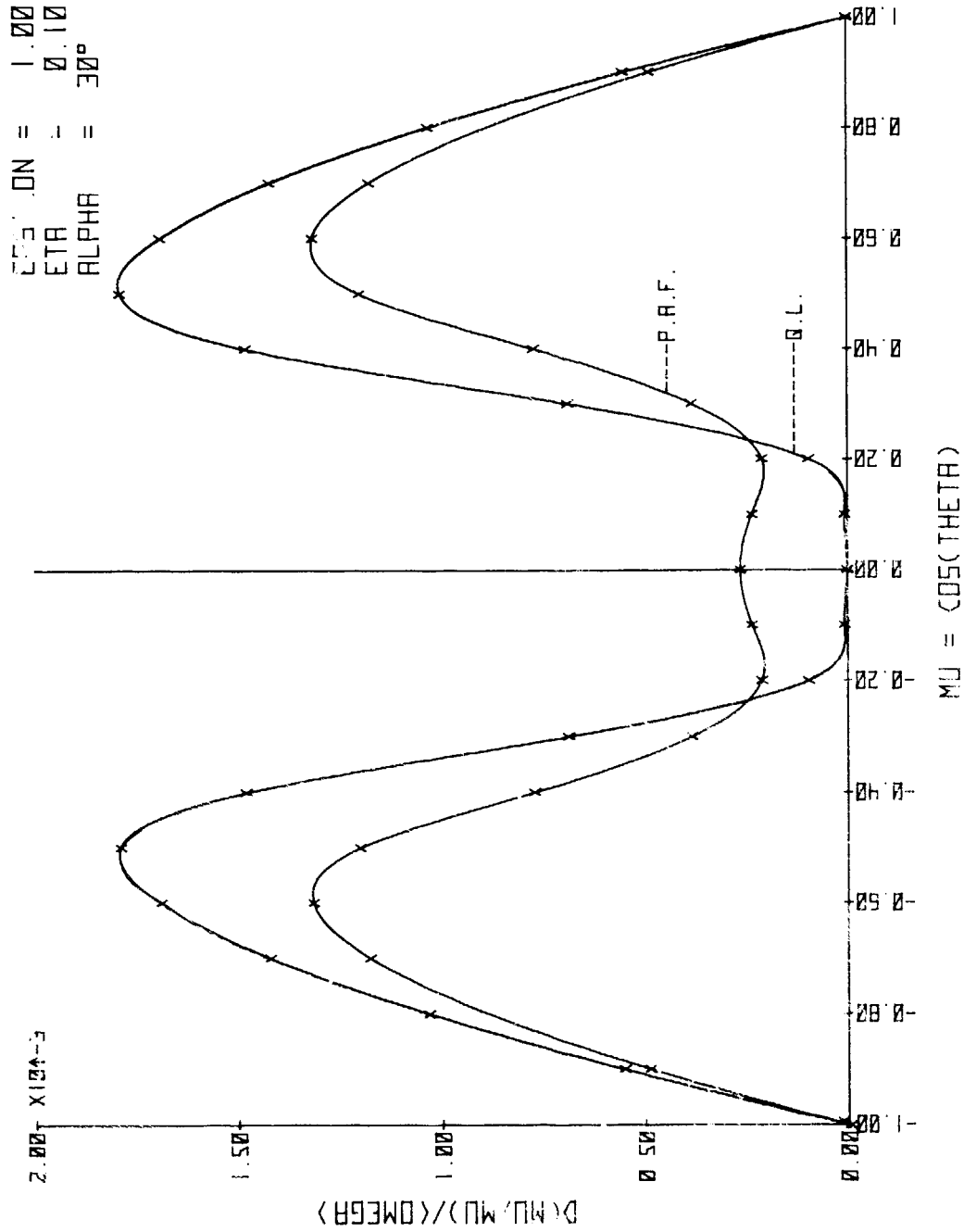


Figure 5. D vs. μ for the oblique slab model when $\zeta = 1$ and $\eta = .1$