ARTICLE IN PRESS

Journal of Atmospheric and Solar-Terrestrial Physics xxx (2017) 1–6

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar-Terrestrial Physics

journal homepage: www.elsevier.com/locate/jastp

Diagonalization of diffusion equations in two and three dimensions

J.M. Albert

Air Force Research Laboratory, Space Vehicles Directorate, Kirtland Air Force Base, New Mexico, USA

ARTICLE INFO

Keywords:
Radiation belts
Multidimensional diffusion

ABSTRACT

Multidimensional diffusion is central to radiation belt and ring current dynamics, but obtaining numerical solutions reliably is subject to difficulties related to "cross terms" in the diffusion equation. Eliminating them, by constructing new coordinates which diagonalize the diffusion matrix, has been found to be effective in two dimensions. Here this approach is reformulated to be both clearer and more robust, with analytical expressions replacing numerical solutions of differential equations. An approach to extending the method to three dimensions is presented and discussed.

1. Introduction

Multidimensional diffusion processes form the core of most current models of radiation belt dynamics (Albert et al., 2009, 2016; Subbotin and Shprits, 2009; Drozdov et al., 2015; Tu et al., 2013), and are also a key ingredient in ring current models (Zheng et al., 2011; Kang et al., 2016; Yu et al., 2016; Jordanova et al., 2016). These processes include Landau and cyclotron-resonant interactions with various classes of waves, driving pitch angle and energy scattering, and drift-resonant interactions with ULF waves, leading to radial transport.

Wave-induced pitch angle and energy diffusion are coupled, so the diffusion coefficients, e.g., $D_{\alpha_0\alpha_0}$ and D_{pp}, are accompanied by "cross" diffusion coefficients $D_{\alpha_0 p}$. The corresponding terms in the diffusion equation can lead to numerical difficulties, and much effort has gone into dealing with them effectively (Camporeale et al., 2013a,b,c; Albert, 2013). Nevertheless, just guaranteeing positivity of the solution from traditional, grid-based finite differencing remains surprisingly difficult (Albert et al., 2009; Albert, 2013). One approach that does so is the "diagonalization" procedure of Albert and Young (2005), which uses the values of the diffusion coefficients to construct new variables (Q_1, Q_2), in which the cross terms vanish. Other, rather different approaches which also guarantee positivity include Monte Carlo solutions of the equivalent stochastic differential equations or SDEs (Tao et al., 2008), and the related (but deterministic) "layer" method (Tao et al., 2009, 2016).

In the original implementation of the diagonalization method, Q_1 was simply set to α_0 for simplicity. More generally, as mentioned by Albert and Young (2005), it may be better to require that the two new coordinate curves be orthogonal, which implies they are aligned with the eigenvectors of the diffusion matrix. This is worked out in Section 2.

The diffusion coefficients are generally available only on a tabulated grid, which is interpolated. Albert and Young (2005) did this in the course of integrating differential equations for the coordinate curves throughout the (α_0, p) plane. Here, the limited availability of the diffusion coefficients is used to advantage: with bilinear interpolation, the coordinate curves can be integrated and the new variables constructed analytically, within each cell of the table. This is carried out in Section 3. Advancing the diffusion equation for $f(t)$ requires f to be evaluated at points where the constant-Q curves meet cell edges; this is done with more sophisticated interpolation techniques as discussed in Section 4.

If the geomagnetic field is approximated as a dipole, radial diffusion decouples from energy and pitch angle diffusion and the two-dimensional diagonalization method can be used at each L (Albert et al., 2009). However, in more general magnetic field models, azimuthal asymmetry couples pitch angle diffusion to radial diffusion, giving a non-decomposable 3×3 diffusion matrix (Schulz, 1991; O'Brien, 2014, 2015). While both the SDE method and the layer method can readily be generalized to 3D (Zheng et al., 2014, 2016; Selesnick, 2016; Wang et al., 2017), doing so for the diagonalization method is not as straightforward. One approach is explored in Section 5.

2. Variables

The two-dimensional bounce-averaged diffusion equation for phase space density $f(x_1, x_2)$ at a fixed value of L may be written as

$$\frac{\partial f}{\partial t} = \frac{1}{G}\nabla\cdot G\begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{bmatrix}\nabla f. \tag{1}$$

Here (x_1, x_2) are any two independent functions of the adiabatic

E-mail address: jay.albert@us.af.mil.

http://dx.doi.org/10.1016/j.jastp.2017.08.008
Received 26 April 2017; Received in revised form 6 July 2017; Accepted 9 August 2017
Available online xxxx
1364-6826/Published by Elsevier Ltd.

invariants (J_1, J_2), such as (α_0, p) or $(\sin\alpha_0, \log E)$, and ∇ means $(\partial/\partial x_1, \partial/\partial x_2)$, while G is the Jacobian determinant $|\partial(J_1,J_2)/\partial(x_1,x_2)|$. The matrix in Eq. (1) will also be referred to as D_{XX}. It is important that in this notation $D_{ij}\equiv D_{x_ix_j}$ has dimensions of x_ix_j/t so that, for example, $D_{\alpha_0\alpha_0}$ has dimensions of $1/t$ while D_{pp} has dimensions of p^2/t. However, the variables may be nondimensionalized (normalized) without changing the form of the equation. It is important that in any set of variables, the form of bounce-averaged quasi-linear diffusion coefficients mathematically guarantees the condition $D_{11}D_{22} > D_{12}^2$ (Albert, 2004).

2.1. Variable transformations

The relationship between the diffusion matrices $[D_{JJ}]$ and $[D_{QQ}]$ for any other set of variables, (Q_1, Q_2), may be written as a matrix equation (Haerendel, 1968; Schulz, 1991). This is just due to the general rules for partial derivatives, not any special properties of (J_1, J_2). In fact, repeating it for $[D_{JJ}]$ and $[D_{XX}]$ and eliminating $[D_{JJ}]$ gives the natural generalization

$$[D_{QQ}] = \begin{bmatrix} \partial_1 Q_1 & \partial_2 Q_1 \\ \partial_1 Q_2 & \partial_2 Q_2 \end{bmatrix} [D_{XX}] \begin{bmatrix} \partial_1 Q_1 & \partial_1 Q_2 \\ \partial_2 Q_1 & \partial_2 Q_2 \end{bmatrix} \tag{2}$$

for any two sets of variables, where $\partial_j Q_i$ means $\partial Q_i/\partial x_j$.

As in Albert and Young (2005), these partial derivatives also appear when considering curves on which Q_1 and Q_2 are constant:

$$\begin{aligned} dQ_1 = 0 \Rightarrow S_1 &= \left.\frac{dx_2}{dx_1}\right|_{Q_1} = -\frac{\partial Q_1/\partial x_1}{\partial Q_1/\partial x_2}, \\ dQ_2 = 0 \Rightarrow S_2 &= \left.\frac{dx_2}{dx_1}\right|_{Q_2} = -\frac{\partial Q_2/\partial x_1}{\partial Q_2/\partial x_2}, \end{aligned} \tag{3}$$

where the total derivatives can be interpreted as slopes S_1 and S_2, respectively, in the (x_1, x_2) plane. Then the condition $D_{Q_1Q_2} = 0$ can be written as

$$D_{11}S_1S_2 - D_{12}(S_1 + S_2) + D_{22} = 0. \tag{4}$$

This single requirement allows a wide range of choice in determining Q_1 and Q_2.

With any such choice, the transformation between (x_1, x_2) and (Q_1, Q_2) also defines a transformation between (J_1, J_2) and (Q_1, Q_2), and the diffusion equation becomes

$$\frac{\partial f}{\partial t} = \frac{1}{\Gamma}\left(\frac{\partial}{\partial Q_1}\Gamma D_{Q_1Q_1}\frac{\partial f}{\partial Q_1} + \frac{\partial}{\partial Q_2}\Gamma D_{Q_2Q_2}\frac{\partial f}{\partial Q_2}\right) \tag{5}$$

(Schulz, 1991), where the Jacobian Γ is given by

$$\Gamma = \left|\frac{\partial(J_1,J_2)}{\partial(Q_1,Q_2)}\right| = \left|\frac{\partial(J_1,J_2)}{\partial(x_1,x_2)}\right| \left|\frac{\partial(Q_1,Q_2)}{\partial(x_1,x_2)}\right|^{-1}. \tag{6}$$

If the maximum values of x_1 and Q_1 correspond to equatorially mirroring particles, the associated condition of no transport across the boundary is $D_{11}\partial f/\partial x_1 + D_{12}\partial f/\partial x_2 = 0$ or the simpler but analogous form $\partial f/\partial Q_1 = 0$.

2.2. Diagonalization: previous approach

The simplest approach is to directly specify one of the new variables. The choice $Q_1 \equiv x_1$, as in Albert and Young (2005), with $x_1 = \alpha_0$, leads to $1/S_1 = 0$,

$$S_2 = \left.\frac{dx_2}{dx_1}\right|_{Q_2} = \frac{D_{12}}{D_{11}}, \tag{7}$$

and

$$\begin{aligned} D_{Q_1Q_1} &= D_{11}, \\ D_{Q_2Q_2} &= (\partial Q_2/\partial x_2)^2 \left(D_{22} - D_{12}^2/D_{11}\right). \end{aligned} \tag{8}$$

Thus curves with constant Q_1 are vertical in the (x_1, x_2) plane, and constant-Q_2 curves are roughly horizontal when their slope D_{12}/D_{11} is small, giving a viable grid in the new coordinates. Alternatively, the choice $Q_2 \equiv x_2$, with, e.g., $x_2 = p/mc$, leads to $S_2 = 0$,

$$\frac{1}{S_1} = \left.\frac{dx_1}{dx_2}\right|_{Q_1} = \frac{D_{12}}{D_{22}}, \tag{9}$$

and

$$\begin{aligned} D_{Q_1Q_1} &= (\partial Q_1/\partial x_1)^2 \left(D_{11} - D_{12}^2/D_{22}\right), \\ D_{Q_2Q_2} &= D_{22}. \end{aligned} \tag{10}$$

These constant-Q_2 curves are horizontal in the (x_1, x_2) plane, and constant-Q_1 curves have slope D_{22}/D_{12} which should be large to give roughly vertical curves and a good grid.

2.3. Eigenvalues and eigenvectors

Since $[D_{XX}]$ is symmetric, it can be diagonalized by its eigenvalues $\lambda_\pm$ and orthonormal eigenvectors $(\mathbf{V}, \mathbf{W})$:

$$\begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} = \begin{bmatrix} V_1 & V_2 \\ W_1 & W_2 \end{bmatrix} [D_{XX}] \begin{bmatrix} V_1 & W_1 \\ V_2 & W_2 \end{bmatrix}, \tag{11}$$

where

$$\lambda_+ = D_{11} + D_0, \quad \lambda_- = D_{22} - D_0 \tag{12}$$

and

$$\begin{bmatrix} V_1 & W_1 \\ V_2 & W_2 \end{bmatrix} = \frac{1}{\sqrt{D_{12}^2 + D_0^2}} \begin{bmatrix} D_{12} & -D_0 \\ D_0 & D_{12} \end{bmatrix}, \tag{13}$$

with

$$D_0 \equiv \sqrt{\left(\frac{D_{11} - D_{22}}{2}\right)^2 + D_{12}^2} - \frac{D_{11} - D_{22}}{2}. \tag{14}$$

D_0 is positive, and the diffusion coefficients can be normalized to avoid incompatible physical dimensions. As D_{12}^2 takes on values from 0 to its maximum permissible value $D_{11}D_{22}$, λ_+ increases and λ_- decreases within the ranges

$$0 \le \lambda_- \le \min(D_{11},\ D_{22})$$

$$\le \max(D_{11}, D_{22}) \le \lambda_+ \le D_{11} + D_{22}.$$

The ratio $|D_0/D_{12}|$ can be either large or small. If $D_{11} > D_{22}$ then $D_0 < |D_{12}|$, and letting $D_{12} \to 0$ (with fixed D_{11} and D_{22}) gives $D_0/D_{12} \to 0$. Conversely, if $D_{11} < D_{22}$, letting $D_{12} \to 0$ gives $D_{12}/D_0 \to 0$.

2.4. Diagonalization by eigenvectors

Eq. (11) has an obvious resemblance to Eq. (2) with $D_{Q_1Q_2} = 0$. The orthogonality of eigenvectors suggests imposing

$$S_1S_2 = -1 \tag{15}$$

on (Q_1, Q_2), along with Eq. (4). This leads to quadratic equations for S_1 and S_2, with solutions