

COMPUTATION OF FOKKER-PLANCK EQUATION

BY

STEPHEN S.-T. YAU

*Department of Mathematics, Statistics and Computer Science, M/C 249, 851 South Morgan Street,
Chicago, Illinois 60607-7045*

Abstract. In plasma physics, the interaction of radio-frequency waves with a plasma is described by a Fokker-Planck equation with an added quasilinear term. In nonlinear filtering with conditional probability density of the state x_t given the observation $\{y(s) : 0 \leq s \leq t\}$ is also described by a Fokker-Planck equation with an added first order term. Method for solving Fokker-Planck equation by means of ordinary differential equations is discussed.

1. Introduction. Fokker-Planck models are most useful for the study of rf -driven currents [Fi] or neutral beam heating in tokamaks on time scales longer than the collisional time (τ_{coll}). In [Ka], Karney took the plasma to be azimuthally symmetric about the magnetic field and homogeneous (representative of the central portion of a tokamak plasma). He presented some numerical methods to solve the Fokker-Planck equation in time and two velocity (or momentum) dimensions only. However, a complete Fokker-Planck treatment of rf or neutral beam heating in tokamaks generally requires solution of an equation which is at least two dimensions in momentum/velocity space and two dimensions in configuration space. A reduction in dimensionality occurs in cases where the bounce/transit time of the particles, τ_b , is shorter when compared to the collision time, i.e. $\tau_b \ll \tau_{\text{coll}}$. The present generation of larger tokamak experimental devices often operate with most of the plasma in this low-collisionality “banana” regime. Moreover, it is usually the case that the non-Maxwellian particles generated by auxiliary heating and current drive are in the low-collisionality regime. In such cases, a “bounce-average” over the bounce or toroidal transit motion of the particle is appropriate, reducing the Fokker-Planck equation to be essentially three-dimensional since the particle distributions as a function of poloidal angle become constant when expressed as a function of the collisionless constants of motion. In [Ke-Mc], Kerbel and McCoy developed a numerical solution scheme for the 3-dimensional Fokker-Planck equation. Therefore in plasma

Received January 27, 2003.

2000 *Mathematics Subject Classification.* Primary 60H10, 82C31.

Research supported partially by ARO, DAAD 19-02-10292, and W911nf 04-1-0085.

E-mail address: yau@uic.edu

physics, it remains very desirable to solve the Fokker-Planck equation for dimensions larger than 3.

On the other hand, nonlinear filtering is concerned with making estimates of quantities associated with a stochastic process $\{x_t\}$ on the basis of information gleaned from a related process $\{y_t\}$. The process $\{x_t\}$ is called the signal or state process and $\{y_t\}$ is called the observation process. The goal is the computation, for each t , of least square estimates of functions of the signal $\{x_t\}$ given the observation history $\{y_s : 0 \leq s \leq t\}$, i.e., the computation of conditional expectations of the form $E[\phi(x_t)|y_s : 0 \leq s \leq t]$, or perhaps even the computation of the entire conditional distribution $\rho(t, x)$ of x_t . It is well known that $\rho(t, x)$ is given by normalizing a function $\sigma(t, x)$ which satisfies a Duncan-Mortensen-Zakai (DMZ) equation. By gauge transformation, one can show that it is sufficient to solve the robust DMZ equation, which is essentially the Fokker-Planck equation with an added first order term. In [Ya-Ya], it is shown that in order to get a numerical solution of the robust DMZ equation, it is sufficient to find numerical solution of the Fokker-Planck equation. Hence from the point of view of nonlinear filtering, it is also very desirable to solve the Fokker-Planck equation for dimensions larger than 3.

The purpose of this paper is to present an ODE method to solve the Fokker-Planck equation. These ODEs are of first order. The total number of ODEs is $n + n^2$, where n is the state space dimension of the Fokker-Planck equation. Since there are many well-known ODE solvers, the Fokker-Planck equation can be solved very efficiently, even for very large n .

In Sec. 2, we recall the background of the Fokker-Planck equation from the viewpoints of plasma physics and nonlinear filtering. In Sec. 3, we show that certain kinds of Fokker-Planck equations can be solved easily by means of solution of system of first order ordinary differential equations.

2. Preliminaries.

2.1. *The Fokker-Planck equation in plasma physics.* We first recall the Fokker-Planck equation from the point of view of plasma physics. The Fokker-Planck equation for the electron e can be written as

$$(2.1.1) \quad \frac{\partial f_e}{\partial t} - \sum_s C(f_e, f_s) + \nabla \cdot \vec{S}_w + \frac{q_e \vec{E}}{m_e} \cdot \nabla f_e = 0,$$

where q_s and m_s are the charge and mass of species s , $C(f_a, f_b)$ is the collision term for species a colliding off species b , the sum extends over all the species of the plasma (typically electrons and ions), \vec{S}_w is the wave (w)-induced quasilinear flux, and $\vec{E} = E \vec{V}_{\parallel}$ is the electric field (assumed to be parallel to the magnetic field). The quantity q_s carries the sign of the charge, thus $q_e = -e$. The subscript \parallel refers to the direction parallel to the magnetic field. The $\nabla = \partial/\partial \vec{V}$ operator operates in velocity space.

Because collisions in a plasma are primarily due to small-angle scattering, the collision term can be written as the divergence of a flux

$$(2.1.2) \quad C(f_a, f_b) = -\nabla \cdot S_c^{a/b}$$

in which Eq. (2.1.1) can be expressed as

$$(2.1.3) \quad \frac{\partial f_e}{\partial t} + \nabla \cdot \vec{S} = 0,$$

where

$$\vec{S} = \vec{S}_c + \vec{S}_w + \vec{S}_e$$

is the total flux in velocity space, and

$$(2.1.4) \quad \vec{S}_c = \sum_s S_c^{e/s}$$

$$(2.1.5) \quad \vec{S}_e = \frac{q_e \vec{E}}{m_e} f_e$$

are the collisional (*c*)- and electric-field (*e*) - induced electron fluxes.

Typically, two types of terms appear in \vec{S} : a diffusion term and a friction term

$$(2.1.6) \quad \vec{S} = -D \cdot \nabla f_e + \vec{F} f_e.$$

Combining (2.1.3) and (2.1.6), we see that the Fokker-Planck equation in plasma physics looks like:

$$(2.1.7) \quad \frac{\partial f_e}{\partial t} = \nabla \cdot (D \cdot \nabla f_e) + \nabla \cdot (\vec{F} f_e).$$

2.2. *Fokker-Planck equation in nonlinear filtering.* The filtering problem considered here is based on the following signal observation model:

$$(2.2.1) \quad \begin{cases} dx(t) = f(x(t))dt + g(x(t))dv(t) & x(0) = x_0 \\ dy(t) = h(x(t))dt + dw(t) & y(0) = 0 \end{cases}$$

in which x, v, y , and w are respectively $\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^m$, and \mathbb{R}^m valued processes and v and w have components that are independent, standard Brownian processes. We further assume that $n = p$; f, g , and h are vector-valued, matrix-valued, and vector-valued C^∞ smooth functions. We shall refer to $x(t)$ as the state of this system at time t and to $y(t)$ as the observation at time t .

Let $\rho(t, x)$ denote the conditional probability density of the state given the observation $\{y(s) : 0 \leq s \leq t\}$. It is well known that $\rho(t, x)$ is given by normalizing a function $\sigma(t, x)$ that satisfies the following Duncan-Mortensen-Zakai equation:

$$(2.2.2) \quad \begin{cases} d\sigma(t, x) = L_0\sigma(t, x)dt + \sum_{i=1}^m L_i\sigma(t, x)dy_i(t) \\ \sigma(0, x) = \sigma_0(x), \end{cases}$$

where

$$L_0 = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [g(x)g^T(x)]_{ij} - \sum_{i=1}^n f_i \frac{\partial}{\partial x_i} - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} - \frac{1}{2} \sum_{i=1}^m h_i^2$$

for $i = 1, \dots, m$, L_i is the zero degree differential operator of multiplication by h_i , and σ_0 is the probability density of the initial point x_0 . In most of the applications, $[g(x)g^T(x)]_{ij}$

are assumed to be constants G_{ij} , $1 \leq i, j \leq n$. Note that $G_{ij} = G_{ji}$. Then

$$(2.2.3) \quad L_0 = \frac{1}{2} \sum_{i,j=1}^n G_{ij} \frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^n f_i \frac{\partial}{\partial x_i} - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} - \frac{1}{2} \sum_{i=1}^m h_i^2.$$

Equation (2.2.2) is a stochastic partial differential equation in the sense of Stratonovich. Define a new unnormalized density

$$(2.2.4) \quad u(t, x) = \exp \left(- \sum_{i=1}^m h_i(x) y_i(t) \right) \sigma(t, x).$$

Then we can reduce (2.2.2) to the following time varying partial differential equation

$$(2.2.5) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) &= L_0 u(t, x) + \sum_{i=1}^m y_i(t) [L_0, L_i] u(t, x) \\ &+ \frac{1}{2} \sum_{i,j=1}^m y_i(t) y_j(t) [[L_0, L_i], L_j] u(t, x) \\ u(0, x) &= \sigma_0(x), \end{cases}$$

where $[L_0, L_i]$ denotes the commutator of the differential operators.

Using the similar technique developed in [Ya-Ya], one can show that it is enough to solve the Fokker-Planck equation of the following form:

$$(2.2.6) \quad \frac{\partial u}{\partial t}(t, x) = \frac{1}{2} \sum_{i,j=1}^m G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n f_i(x) \frac{\partial u}{\partial x_i}(t, x) - \left(\sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2} \sum_{i=1}^m h_i^2(x) \right) u(t, x).$$

2.3. *Fokker-Planck equation in general form.* In view of (2.1.7) and (2.2.6), we shall consider the following general form of the Fokker-Planck equation on \mathbb{R}^n :

$$(2.3.1) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) &= \sum_{i,j=1}^n G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n f_i(x) \frac{\partial u}{\partial x_i}(t, x) + q(x) u(t, x) \\ u(0, x) &= u_0(x) \end{cases}$$

where $G_{ij} = G_{ji}$ are constants.

3. Explicit solution of the Fokker-Planck equation in terms of solutions of ODEs. We shall solve the Fokker-Planck equation (2.3.1) where $f_i(x)$ are degree one polynomials and $Q(x)$ is a degree 2 polynomial:

$$(3.1) \quad f_i(x) = l_i(x) = \sum_{j=1}^n d_{ij} x_j + d_i, \quad 1 \leq i \leq n$$

$$(3.2) \quad q(x) = x^T Q x + p^T x + r,$$

where $Q = (q_{ij})$ is an $n \times n$ symmetric matrix, $p^T = (p_1, \dots, p_n)$ and $x^T = (x_1, \dots, x_n)$ are $1 \times n$ matrices, and r is a scalar.

It is well known that any distribution is well approximated by a finite linear combination of Gaussians of the form $\alpha_1 G_1 + \dots + \alpha_k G_k$, where α_i 's are real numbers and G_i 's

are Gaussian distributions. Let u_i be the solution of (2.3.1) with initial distribution G_i . Since (2.3.1) is a linear partial differential equation, it follows that the solution of (2.3.1) is of the form $\alpha_1 u_1 + \dots + \alpha_k u_k$. Therefore it remains to solve (2.3.1) with Gaussian initial distribution. The following theorem gives an explicit solution of (2.3.1) with Gaussian initial distribution in terms of solutions of ODEs.

Theorem 3.1. *Consider the Fokker-Planck equation on \mathbb{R}^n with Gaussian initial distribution*

$$(3.3) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) = \sum_{i,j=1}^n G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n \ell_i(x) \frac{\partial u}{\partial x_i}(t, x) + q(x)u(t, x) \\ u(0, x) = e^{x^T A(0)x + B^T(0)x + C(0)}, \end{cases}$$

where $G_{ij} = G_{ji}$ are constants, $A(0) = (A_{ij}(0))$ is an $n \times n$ matrix, $B^T(0) = (B_1(0), \dots, B_n(0))$, and $x^T = (x_1, \dots, x_n)$ are $1 \times n$ matrices, and $C(0)$ is a scalar. Suppose conditions (3.1) and (3.2) hold. Then the solution of (3.3) is of the following form:

$$(3.4) \quad u(t, x) = e^{x^T A(t)x + B^T(t)x + C},$$

where $A(t)$ is an $n \times n$ symmetric matrix valued function of t , $B^T(t) = (B_1(t), \dots, B_n(t))$ is a $1 \times n$ matrix valued function of t , and $C(t)$ is a scalar function of t . Moreover, $A(t)$, $B^T(t)$, and $C(t)$ satisfy the following system of ODEs:

$$(3.5) \quad \frac{dA}{dt} = A^T G A + 2AG A + AG^T A^T - (A^T + A)D + Q$$

$$(3.6) \quad \frac{dB^T}{dt} = B^T G A + 2B^T G^T A^T + B^T G^T A - B^T D - d^T A - d^T A^T + p^T$$

$$(3.7) \quad \frac{dC}{dt} = 2tr(GA) + B^T G B - d^T B + r.$$

PROOF: Differentiating (3.4) with respect to t and x_j , we get the following equations:

$$(3.8) \quad \begin{aligned} \frac{\partial u}{\partial t} &= \left(x^T \frac{dA}{dt} x + \frac{dB^T}{dt} x + \frac{dC}{dt} \right) u \\ \frac{\partial u}{\partial x_j} &= \left[\sum_{i,k=1}^n A_{ik} \left(\frac{\partial x_i}{\partial x_j} x_k + x_i \frac{\partial x_k}{\partial x_j} \right) + B_j \right] u \\ &= \left(\sum_{k=1}^n A_{jk} x_k + \sum_{i=1}^n A_{ij} x_i + B_j \right) u \\ \nabla u^T &= [(Ax)^T + x^T A + B^T] u = (x^T A^T + x^T A + B^T) u \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 u}{\partial x_i \partial x_j} &= \left[A_{ji} + A_{ij} + \left(\sum_{k=1}^n A_{jk} x_k + \sum_{k=1}^n A_{kj} x_k + B_j \right) \left(\sum_{k=1}^n A_{ik} x_k + \sum_{k=1}^n A_{ki} x_k + B_i \right) \right] u \\
&= \left[A_{ji} + A_{ij} + \sum_{k,\ell=1}^n A_{jk} A_{i\ell} x_k x_\ell + \sum_{k,\ell=1}^n A_{jk} A_{\ell i} x_k x_\ell + B_i \sum_{k=1}^n A_{jk} x_k + \sum_{k,\ell=1}^n A_{kj} A_{i\ell} x_k x_\ell + \sum_{k,\ell=1}^n A_{kj} A_{\ell i} x_k x_\ell + B_i \sum_{k=1}^n A_{kj} x_k + B_j \sum_{k=1}^n A_{ik} x_k + B_j \sum_{k=1}^n A_{ki} x_k + B_j B_i \right] u \\
\sum_{i,j=1}^n G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} &= \left[\sum_{i,j=1}^n G_{ij} A_{ji} + \sum_{i,j=1}^n G_{ij} A_{ij} + \sum_{i,j,k,\ell=1}^n G_{ij} A_{jk} A_{i\ell} x_k x_\ell + \sum_{i,j,k,\ell=1}^n G_{ij} A_{jk} A_{\ell i} x_k x_\ell + \sum_{i,j,k=1}^n G_{ij} B_i A_{jk} x_k + \sum_{i,j,k,\ell=1}^n G_{ij} A_{kj} A_{i\ell} x_k x_\ell + \sum_{i,j,k=1}^n G_{ij} B_i A_{kj} x_k + \sum_{i,j,k=1}^n G_{ij} B_j A_{ik} x_k + \sum_{i,j,k=1}^n G_{ij} B_j A_{ki} x_k + \sum_{i,j=1}^n G_{ij} B_j B_i \right] u \\
&= \left[\sum_{i,j=1}^n G_{ij} A_{ji} + \sum_{i,j=1}^n G_{ji} A_{ij} + \sum_{i=1}^n \left(\sum_{\ell=1}^n A_{i\ell} x_\ell \right) \left(\sum_{j,k=1}^n G_{ij} A_{jk} x_k \right) + \sum_{i=1}^n \left(\sum_{\ell=1}^n A_{\ell i} x_\ell \right) \left(\sum_{j,k=1}^n G_{ij} A_{jk} x_k \right) + \sum_{i=1}^n B_i \sum_{j,k=1}^n G_{ij} A_{jk} x_k + \sum_{i,j,h,\ell=1}^n x_k A_{kj} G_{ji} A_{i\ell} x_\ell + \sum_{j=1}^n \left(\sum_{k=1}^n x_k A_{kj} \right) \left(\sum_{i,\ell=1}^n x_\ell A_{\ell i} G_{ij} \right) + \sum_{j=1}^n \left(\sum_{k=1}^n x_k A_{kj} \right) \left(\sum_{i=1}^n G_{ji} B_i \right) + \sum_{i=1}^n \left(\sum_{j=1}^n G_{ij} B_j \right) \left(\sum_{k=1}^n A_{ik} x_k \right) + \sum_{i=1}^n \left(\sum_{k=1}^n x_k A_{ki} \right) \left(\sum_{j=1}^n G_{ij} B_j \right) + \sum_{i,j=1}^n G_{ij} B_j B_i \right] u
\end{aligned}$$

$$\begin{aligned}
&= [2\text{tr}(GA) + (Ax)^T(GAx) + (x^T A)(GAx) + B^T GAx + x^T AGAx \\
&\quad + x^T A(x^T AG)^T + (x^T A)GB + (GB)^T Ax + (x^T A)(GB) + B^T GB]u \\
&= [x^T(A^T GA + AGA + AGA + AG^T A^T)x + (B^T GA + B^T G^T A^T \\
&\quad + B^T G^T A + B^T G^T A^T)x + 2\text{tr}(GA) + B^T GB]u \\
&= [x^T(A^T GA + 2AGA + AG^T A^T) + (B^T GA + 2B^T G^T A^T + B^T G^T A)x \\
&\quad + 2\text{tr}(GA) + B^T GB]u.
\end{aligned}$$

Let $D = (d_{ij})$ be an $n \times n$ matrix and $d^T = (d_1, \dots, d_n)$ be a $1 \times n$ matrix. Then

$$\begin{aligned}
\sum_{i=1}^n \ell_i(x) \frac{\partial u}{\partial x_i} &= \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_j \frac{\partial u}{\partial x_i} + \sum_{i=1}^n d_i \frac{\partial u}{\partial x_i} \\
&= \nabla u^T D x + d^T \nabla u \\
&= (x^T A^T + x^T A + B^T)u D x + d^T (Ax + A^T x + B)u \\
&= [x^T(A^T + A)D x + (B^T D + d^T A + d^T A^T)x + d^T B]u.
\end{aligned}$$

Thus the L.H.S. of (3.3) is given by

$$\begin{aligned}
(3.9) \quad \sum_{i,j=1}^n G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n \ell_i(x) \frac{\partial u}{\partial x_i}(t, x) + q(x)u(t, x) \\
= [x^T(A^T GA + 2AGA + AG^T A^T - (A^T + A)D + Q)x + (B^T GA + 2B^T G^T A^T \\
+ B^T G^T A - B^T D - d^T A - d^T A^T + p^T)x + 2\text{tr}(GA) + B^T GB - d^T B + r]u.
\end{aligned}$$

Equating (3.8) and (3.9) and comparing terms, we get equations (3.5), (3.6), and (3.7).
Q.E.D.

Theorem 3.2. *The solution of the following Fokker-Planck equation on \mathbb{R}^n can be found by the method described above.*

(3.10)

$$\frac{\partial u}{\partial t}(t, x) = \sum_{i,j=1}^n G_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n \left(\sum_{j=1}^n G_{ij} \frac{\partial F}{\partial x_j}(x) + \ell_i(x) \right) \frac{\partial u}{\partial x_i}(t, x) + q(x)u(t, x)$$

where $G_{ij} = G_{ji}$ are constants; $\ell_i(x)$ are degree one polynomials; $F(x)$ and $q(x)$ are C^∞ functions on \mathbb{R}^n such that

$$\begin{aligned}
(3.11) \quad \tilde{q}(x) := \frac{1}{2} \sum_{i,j=1}^n G_{ij} \frac{\partial^2 F}{\partial x_i \partial x_j}(x) + \frac{1}{4} \sum_{i,j=1}^n \frac{\partial F}{\partial x_i}(x) \frac{\partial F}{\partial x_j}(x) \\
- \frac{1}{2} \sum_{i=1}^n \frac{\partial F}{\partial x_i}(x) \left(\sum_{j=1}^n G_{ij} \frac{\partial F}{\partial x_j}(x) + \ell_i(x) \right) + q(x)
\end{aligned}$$

is a polynomial of degree two. In fact, let

$$(3.12) \quad u(t, x) = e^{\frac{1}{2}F(x)} \tilde{u}(t, x).$$

Then (3.10) can be reduced to the following equation:

(3.13)

$$\frac{\partial \tilde{u}}{\partial t}(t, x) = \sum_{i,j} G_{ij} \frac{\partial^2 \tilde{u}}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^n \ell_i(x) \frac{\partial \tilde{u}}{\partial x_i}(t, x) + \tilde{q}(x) \tilde{u}(t, x).$$

PROOF: By putting (3.12) into (3.10), one deduces that (3.13) holds.

Q.E.D.

REFERENCES

- [Fi] N. J. Fisch, Confining a Tokamak Plasma with *rf*-Driven currents, *Phys. Rev. Lett.* 41 (1978), 873–876.
- [Ka] C. F. F. Karney, Fokker-Planck and Quasilinear Codes, *Computer Physics Reports* 4 (1983), 183.
- [Ke-Mc] G. D. Kerbel and M. G. McCoy, Kinetic theory and simulation of multispecies plasmas in tokamaks excited with electromagnetic waves in the ion-cyclotron range of frequencies, *Phys. Fluids*, vol. 28 (1985), 3629–3649.
- [Hu-Ya] G. Q. Hu and S. S.-T. Yau, Finite dimensional filters with nonlinear drift XV: New direct method for construction of universal finite dimensional filter, (to appear) *IEEE Transactions on Aerospace and Electronic Systems*.
- [Ya-Ya] S. T. Yau and S. S.-T. Yau, Real time solution of nonlinear filtering problem without memory I, *Mathematical Research Letters* 7 (2000), 671–693.