A Balance Law Interpretation of the Fokker-Planck Equation

Author: *R. Scott McIntire*

Date: Nov 22, 2002

Overview

In this paper we use a balance law to derive the heat equation in one and several dimensions. We also use it to study the propagation of a probability density function driven by a differential equation. Finally, we combine the two derivations and provide a derivation of the Fokker-Planck equation.

Derivation of the Heat Equation

We examine the problem of heat flow in R^n. First we look at the problem in one dimension and then see how it generalizes to arbitrary dimension. The strategy is to write a balance law for the heat in an arbitrary region in R^n based on the conservation of energy. This will lead to a local expression of heat flow; namely, a partial differential equation. By relating heat to temperature, an equation for the temperature is obtained. The resulting partial differential equation for temperature is called *the heat equation*.

The Heat Equation in One Dimension

We derive a partial differential equation for the temperature distribution in a one dimensional material. To do this we derive a balance law for the heat contained in an arbitrary interval $[x_1, x_2]$ over an arbitrary time interval $[t_1, t_2]$. Let $h(t, x)$ be the heat per unit length at time t and position x. Let $f(t, x)$ be the heat flux at time t and position x. Where heat flux is defined as the rate at which heat flows †. The increase (or decrease) in the heat of the interval $[x_1, x_2]$ over $[t_1, t_2]$ is equal to the integral of the rate at which heat has moved into (or out of) the region through the boundaries. Mathematically, this balance may be expressed as:

$$\int_{x_1}^{x_2} h(t_2, x) - h(t_1, x)\, dx = \int_{t_1}^{t_2} f(t, x_1) - f(t, x_2)\, dt$$

Now, divide this expression by $1/(t_2 - t_1)$ and take the limit as $t_2 \to t_1$. This gives:

$$\int_{x_1}^{x_2} h_t(t_1, x)\, dx = f(t_1, x_1) - f(t_1, x_2) \tag{1}$$

This can be written as

$$\frac{d}{dt} \int_{x_1}^{x_2} h(t, x)\, dx \Big|_{t=t_1} = f(t_1, x_1) - f(t_1, x_2) \tag{2}$$

Equation (2) can be interpreted as an instantaneous balance law for the heat contained in the interval $[x_1, x_2]$. It states that the rate of change of the heat in the interval $[x_1, x_2]$ is equal to the rate at which heat enters through the boundary.

† positive means increasing x.

By the *the Fundamental Theorem of Calculus*, $\int_{x_1}^{x_2} f_x(t_1, x)\, dx = f(t_1, x_2) - f(t_1, x_1)$. Therefore, bringing the right hand side to the left hand side in equation (1) we may write:

$$\int_{x_1}^{x_2} h_t(t_1, x) + f_x(t_1, x)\, dx = 0$$

If we *assume* that the integrand is continuous, then since the interval $[x_1, x_2]$ is arbitrary, the integrand must vanish for all x; that is,

$$h_t(t_1, x) + f_x(t_1, x) = 0$$

Finally, since t_1 is arbitrary the above must be true for all times t; so that,

$$h_t(t, x) + f_x(t, x) = 0$$

To get an equation for the temperature, u, we must relate it to the heat and the heat flux. To do so, we must make *assumptions* about properties of the material. The simplest assumption about the heat is that it is proportional to temperature: $h(t, x) = c\, u(t, x)$. This proportionality constant, $c > 0$, is called the specific heat. The simplest assumption about the heat flux is that it is proportional to the temperature gradient: $f(t, x) = -k u_x(t, x)$ with $k > 0$. Therefore, we may write:

$$c\, u_t - k u_{xx} = 0$$

Letting $\kappa = k/c$ we obtain the one dimensional heat equation with diffusion constant κ:

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial^2 x}$$

The larger the value of κ, the more rapid the diffusion. Since κ is made from two constants, we give an interpretation to understand their role. Consider a bar made of different materials and different thicknesses. Let us assume that the differing materials all have the same heat capacity, but different heat flux rate constants. Then we have two independent controls: the thickness of the bar controls the value of the specific heat, c; and, the material controls the heat flux rate, k.

Under these conditions, we can get a better physical understanding of the influence of c and k on κ. If the bar is thicker at a certain spot, it can absorb more heat and therefore, is less diffusive there. If the bar is made of a material that is more insular at a certain spot, the heat flux rate constant goes down and therefore, the material is less diffusive there.

We note that c and k may depend on x. In this case we obtain a generalized heat equation:

$$\frac{\partial u}{\partial t} = \frac{1}{c(x)} \frac{\partial}{\partial x} \left(k(x) \frac{\partial u}{\partial x} \right)$$

Technically, to have similar properties of the heat equation, $c(x)$ and $k(x)$ should satisfy(for all x):

1. $c(x) > c_0$ for some $c_0 > 0$.
2. $k(x) > k_0$ for some $k_0 > 0$.
3. $c(x)$ is continuous and $k(x)$ is continuously differentiable with respect to x.

The Heat Equation in Higher Dimensions

We derive a partial differential equation for the temperature distribution of a material in R^n. To do this we will mimic each of the steps in the last section. We start by deriving a balance law for the heat contained in an arbitrary region Ω over an arbitrary time interval $[t_1, t_2]$. Let $h(t, \mathbf{x})$ be the heat per unit volume at time t and position $\mathbf{x}$. Let $\mathbf{f}(t, \mathbf{x})$ be the heat flux at time t and position $\mathbf{x}$. Where heat flux is defined as a

vector describing the rate at which heat flows per area†. The increase (or decrease) in the heat of the region Ω over $[t_1, t_2]$ is equal to the integral of the rate at which heat has moved into (or out of) the region through the boundaries. Mathematically, this may be expressed as:

$$\int_\Omega h(t_2, \mathbf{x}) - h(t_1, \mathbf{x})\, d\mathbf{x} = -\int_{t_1}^{t_2} \underbrace{\oint_{\partial\Omega} \overbrace{\mathbf{f}(t, \mathbf{x}) \cdot \mathbf{n}}^{\text{Heat flow rate per unit area leaving } \Omega}\, d\mathbf{S}}_{\text{Heat flow rate leaving } \Omega}\, dt$$

As in the one dimensional case, divide by $1/(t_2 - t_1)$ and take the limit as $t_2 \to t_1$. This gives:

$$\int_\Omega h_t(t_1, \mathbf{x})\, d\mathbf{x} = -\oint_{\partial\Omega} \mathbf{f}(t_1, \mathbf{x}) \cdot \mathbf{n}\, d\mathbf{S} \tag{3}$$

This can be written as

$$\frac{d}{dt} \int_\Omega h(t, \mathbf{x})\, d\mathbf{x}\Big|_{t=t_1} = -\oint_{\partial\Omega} \mathbf{f}(t_1, \mathbf{x}) \cdot \mathbf{n}\, d\mathbf{S} \tag{4}$$

Equation (4) can be interpreted as an instantaneous balance law for the heat contained in the region Ω. It states that the rate of change of the heat in the region Ω is equal to the rate at which heat enters through the boundary.

By the *the Fundamental Theorem of Calculus*, $\int_\Omega \nabla \cdot \mathbf{f}(t_1, \mathbf{x})\, d\mathbf{x} = \oint_{\partial\Omega} \mathbf{f}(t_1, \mathbf{x}) \cdot \mathbf{n}\, d\mathbf{S}$. Therefore, bringing the right hand side to the left hand side in equation (3) we may write:

$$\int_\Omega h_t(t_1, \mathbf{x}) + \nabla \cdot \mathbf{f}(t_1, \mathbf{x})\, d\mathbf{x} = 0$$

If we *assume* that the integrand is continuous, then since the region Ω is arbitrary, the integrand must vanish for all x; that is,

$$h_t(t_1, \mathbf{x}) + \nabla \cdot \mathbf{f}(t_1, \mathbf{x}) = 0$$

Finally, since t_1 is arbitrary the above must be true for all times t; so that,

$$h_t(t, \mathbf{x}) + \nabla \cdot \mathbf{f}(t, \mathbf{x}) = 0$$

To get an equation for the temperature, u, we must relate it to the heat and the heat flux. To do so, we must make *assumptions* about the properties of the material. The simplest assumption about the heat is that it is proportional to temperature: $h(t, \mathbf{x}) = c\, u(t, \mathbf{x})$. As in the one dimensional case, the proportionality constant, $c > 0$, is called the specific heat. The simplest assumption about the heat flux is that it is proportional to the temperature gradient:‡ $\mathbf{f}(t, \mathbf{x}) = -K \nabla u(t, \mathbf{x})$ with $K > 0$. Therefore, we may write:

$$c\, u_t(t, \mathbf{x}) - K \nabla \cdot \nabla u(t, \mathbf{x}) = 0$$

Letting $\kappa = K/c$ and noting that $\Delta = \nabla \cdot \nabla$, we obtain the n-dimensional heat equation with diffusion constant κ.

$$\frac{\partial u}{\partial t} = \kappa\, \Delta u$$

† By area we mean the n-1 dimensional area.

‡ That is, the *direction* of heat flow should be in the opposite direction of maximum change of temperature; and, the *magnitude* of the heat flow should be proportional to the magnitude of the maximum change in the temperature.

We note that c and K may depend on $\mathbf{x}$ and we note K could be a matrix. In this case we obtain a generalized heat equation:

$$\frac{\partial u}{\partial t} = \frac{1}{c(\mathbf{x})}\nabla \cdot \Big(K(\mathbf{x})\nabla u \Big)$$

Technically, to have similar properties of the heat equation, $c(\mathbf{x})$ and $K(\mathbf{x})$ should satisfy(for all $\mathbf{x}$):

1. $c(\mathbf{x}) > c_0$ for some $c_0 > 0$;
2. $K(\mathbf{x})$ should be a positive definite symmetric matrix;
3. The smallest eigenvalue of $K(\mathbf{x})$ is greater than k_0 for some $k_0 > 0$;
4. $c(\mathbf{x})$ is continuous and $K(\mathbf{x})$ is continuously differentiable with respect to $\mathbf{x}$.

When $K(\mathbf{x})$ is a matrix, the heat flux vector $f(t, \mathbf{x})$ will not, in general, point in the direction opposite of the maximum temperature increase. This may be interpreted physically as having a material that has a different heat flow rate in different directions. In this case, the heat flux vector, representing the heat flow path, will be modified by these material differences and will not necessarily flow along the temperature gradient.

Propagation of Probability Density Functions

In this section, we wish to derive equations for the propagation of probability density functions driven by systems of differential equations.

A first Order System

Given a first order system of differential equations of the form:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \tag{5a}$$

$$\mathbf{x}(0) = \mathbf{x_0} \tag{5b}$$

let $\mathcal{X}(t; \mathbf{x_0})$ be the solution; which is assumed to be invertible on the interval $[0, T)$. That is, for each $t \in [0, T)$, $\mathcal{X}(t; \cdot) : R^n \mapsto R^n$ is invertible. If X is a random variable with probability density function $P_X(\mathbf{x})$, then we wish to compute the probability density function of the random variable induced by applying the transformation $\mathcal{X}(t; \cdot)$ to the random variable X. This will be done by writing down a balance law for a suitable probabilistic expression. As in previous sections, we will not obtain a formula for the induced probability density function; rather, we will produce a partial differential equation describing its evolution.

A Balance Law Computation

Given an arbitrary region, $\widehat{\Omega}$, and an arbitrary time $t \in [0, T)$, the probability contained in $\widehat{\Omega}$ will remain the same as the region is transformed either forward or backward in time under the differential equation (5a-b). Therefore, there is a one parameter family of regions Ω_τ such that $\Omega_t = \widehat{\Omega}$ and $\Omega_0 = \Omega$. That is, Ω morphs into $\widehat{\Omega}$ under the action of the differential equation (5a-b). Let $p(t, \mathbf{x}) = P_{\mathcal{X}(t;X)}(\mathbf{x})$ be the probability density function of the induced random variable $\mathcal{X}(t; X)$. The conservation of probability of $P_{\mathcal{X}(t,X)}(\Omega_t)$ is expressed mathematically as an instantaneous balance law:

$$\frac{d}{d\tau}\int_{\Omega_\tau} p(\tau, \mathbf{x})\, d\mathbf{x} = 0 \tag{6}$$

To compute the derivative we change the integration variable so as to keep the region constant. To this end, let $\mathbf{x} = \mathcal{X}(\tau, \mathbf{z})$. The integral becomes:

$$\frac{d}{d\tau}\int_{\Omega} p(\tau, \mathcal{X}(t, \mathbf{z}))\Big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\Big|\, d\mathbf{z} = 0$$

which is

$$\int_{\Omega} \big(p_\tau(\tau, \mathcal{X}(\tau, \mathbf{z})) + \nabla p(\tau, \mathcal{X}(\tau, \mathbf{z})) \cdot \dot{\mathcal{X}}(\tau, \mathbf{z})\big) \big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| + p(\tau, \mathcal{X}(\tau, \mathbf{z})) \frac{d}{d\tau}\big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| \, d\mathbf{z} = 0 \tag{7}$$

It turns out that the derivative, $\frac{d}{d\tau}\big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big|$, may be written as†

$$\frac{d}{d\tau}\big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| = \big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| \, \mathrm{tr}\big(\nabla \mathbf{f}(\tau, \mathcal{X}(\tau, \mathbf{z}))\big) = \big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| \, \nabla \cdot \mathbf{f}(\tau, \mathcal{X}(\tau, \mathbf{z}))$$

Therefore, (7) becomes

$$\int_{\Omega} \Big(p_\tau(\tau, \mathcal{X}(t, \mathbf{z})) + \nabla p(\tau, \mathcal{X}(\tau, \mathbf{z})) \cdot \mathbf{f}(\tau, \mathcal{X}(\tau, \mathbf{z})) + p(\tau, \mathcal{X}(\tau, \mathbf{z})) \nabla \cdot \mathbf{f}(\tau, \mathcal{X}(\tau, \mathbf{z}))\Big) \big|\frac{\partial \mathcal{X}(\tau, \mathbf{z})}{\partial \mathbf{z}}\big| \, d\mathbf{z} = 0 \tag{8}$$

Changing coordinates back, (8) becomes

$$\int_{\Omega_\tau} p_\tau(\tau, \mathbf{x}) + \nabla p(\tau, \mathbf{x}) \cdot \mathbf{f}(\tau, \mathbf{x}) + p(\tau, \mathbf{x}) \nabla \cdot \mathbf{f}(\tau, \mathbf{x}) \, d\mathbf{x} = 0$$

This may be written as

$$\int_{\Omega_\tau} p_\tau(\tau, \mathbf{x}) + \nabla \cdot \big(p(\tau, \mathbf{x}) \mathbf{f}(\tau, \mathbf{x})\big) \, d\mathbf{x} = 0 \tag{9}$$

Equation (9) is true for all $\tau \in [0, T)$, in particular it is true for $\tau = t$. If we assume that the integrand in (9) is continuous and since $\widehat{\Omega} = \Omega_t$ and t are both arbitrary, we have for all $t \in [0, T)$ and for all $\mathbf{x}$

$$p_t(t, \mathbf{x}) + \nabla \cdot \big(p(t, \mathbf{x}) \mathbf{f}(t, \mathbf{x})\big) = 0 \tag{10a}$$

$$p(0, \mathbf{x}) = P_X(\mathbf{x}) \tag{10b}$$

In coordinates this is written as

$$\frac{\partial p(t, \mathbf{x})}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \big(p(t, \mathbf{x}) \mathbf{f}(t, \mathbf{x})\big) = 0$$

$$p(0, \mathbf{x}) = P_X(\mathbf{x})$$

Equations (10a-b) constitute a first order partial differential equation describing the evolution of the probability density function $P_X(\mathbf{x})$ with respect to the differential equation (5a-b).

A Second Order System

Many physical systems are more naturally described by a system of second order differential equations than a first order system. In this section, we derive an equation governing the evolution of a probability density function driven by a second order system. Consider the second order system of differential equations of the form:

$$\ddot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \dot{\mathbf{x}}) \tag{11a}$$

$$\mathbf{x}(0) = \mathbf{x_0} \tag{11b}$$

$$\dot{\mathbf{x}}(0) = \mathbf{v_0} \tag{11c}$$

† This is a consequence of Louisville's Theorem.

This may be written as a first order system as follows: Let $\mathbf{v} = \dot{\mathbf{x}}$, then (11a-c) may be written as

$$\dot{\mathbf{x}} = \mathbf{v} \tag{12a}$$
$$\dot{\mathbf{v}} = \mathbf{f}(t, \mathbf{x}, \mathbf{v}) \tag{12b}$$
$$\mathbf{x}(0) = \mathbf{x_0} \tag{12c}$$
$$\mathbf{v}(0) = \mathbf{v_0} \tag{12d}$$

We can use the result of the last section to write a formula for the propagation of the probability density function. Let $P_{X,V}(\mathbf{x}, \mathbf{v})$ be the joint probability density function for random variables X and V. Let $p(t, \mathbf{x}, \mathbf{v}) = P_{\mathcal{X}(t;X,Y),\mathcal{V}(t;X,Y)}(\mathbf{x}, \mathbf{v})$ denote the probability density function of the pair, (X, Y), transformed by the solution of (12a-d) at time t. From the previous section, p satisfies†

$$p_t(t, \mathbf{x}, \mathbf{v}) + (\, \nabla_{\mathbf{x}} p(t, \mathbf{x}, \mathbf{v}) \;\; , \nabla_{\mathbf{v}} p(t, \mathbf{x}, \mathbf{v}) \,) \cdot \begin{pmatrix} \mathbf{v} \\ \mathbf{f}(t, \mathbf{x}, \mathbf{v}) \end{pmatrix} + p(t, \mathbf{x}, \mathbf{v}) \operatorname{tr} \begin{pmatrix} 0 & I \\ \nabla_{\mathbf{x}} \mathbf{f}(t, \mathbf{x}, \mathbf{v}) & \nabla_{\mathbf{v}} \mathbf{f}(t, \mathbf{x}, \mathbf{v}) \end{pmatrix} = 0$$

Simplifying and adding the initial condition this becomes

$$p_t(t, \mathbf{x}, \mathbf{v}) + \nabla_{\mathbf{x}} p(t, \mathbf{x}, \mathbf{v}) \cdot \mathbf{v} + \nabla_{\mathbf{v}} p(t, \mathbf{x}, \mathbf{v}) \cdot \mathbf{f}(t, \mathbf{x}, \mathbf{v}) + p(t, \mathbf{x}, \mathbf{v}) \nabla_{\mathbf{v}} \cdot \mathbf{f}(t, \mathbf{x}, \mathbf{v}) = 0 \tag{13a}$$
$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{13b}$$

Or,

$$p_t(t, \mathbf{x}, \mathbf{v}) + \nabla_{\mathbf{x}} p(t, \mathbf{x}, \mathbf{v}) \cdot \mathbf{v} + \nabla_{\mathbf{v}} \cdot \big(p(t, \mathbf{x}, \mathbf{v}) \mathbf{f}(t, \mathbf{x}, \mathbf{v})\big) = 0 \tag{14a}$$
$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{14b}$$

In coordinates this is written as

$$\frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial t} + \sum_{i=1}^{n} \frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial x_i} v_i + \sum_{i=1}^{n} \frac{\partial}{\partial v_i} \big(p(t, \mathbf{x}, \mathbf{v}) f_i(t, \mathbf{x}, \mathbf{v})\big) = 0 \tag{15a}$$
$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{15b}$$

The Fokker-Planck Equation I

We now consider the evolution of a probability density function driven primarily by the differential system (5a-b), but which incorporates the dynamics of diffusion as well. Such a governing equation is called a Fokker-Planck equation. If X is a random variable and $\mathcal{X}(t, \mathbf{x})$ is the solution of (5a-b), then let $p(t, \mathbf{x}) = P_{\mathcal{X}(t,X)}(\mathbf{x})$ be the probability density function for the random variable X transformed by $\mathcal{X}(t, \cdot)$. As before, let $\widehat{\Omega}$ be an arbitrary region in R^n and let Ω_τ parameterize a family of regions such that $\Omega_0 = \Omega$ and $\Omega_t = \widehat{\Omega}$.

Instead of equation (6) we assume that

$$\frac{d}{d\tau} \int_{\Omega_\tau} p(\tau, \mathbf{x}) \, d\mathbf{x} = \oint_{\partial \Omega_\tau} \big(A(\mathbf{x}) \nabla p(\tau, \mathbf{x})\big) \cdot \mathbf{n} \, d\mathbf{S} \tag{16}$$

In the simplest case, $A(\mathbf{x}) = I$, equation (16) says that the changing regions, Ω_τ, allow diffusion based on the probability gradient on the boundary. This is the same mechanism as the heat equation. In the general case, the diffusion behaves as the diffusion in the generalized heat equation.

† This follows from the identity $\nabla \cdot (\varphi \mathbf{w}) = \nabla \varphi \cdot \mathbf{w} + \varphi \operatorname{tr} \nabla \mathbf{w}$.

We have already done the computation for the left hand side of (16), so we may write

$$\int_{\Omega_\tau} p_\tau(\tau,\mathbf{x}) + \nabla\cdot\big(p(\tau,\mathbf{x})\mathbf{f}(\tau,\mathbf{x})\big)\,d\mathbf{x} = \oint_{\partial\Omega_\tau}\big(A(\mathbf{x})\nabla p(\tau,\mathbf{x})\big)\cdot\mathbf{n}\,d\mathbf{S} \tag{17}$$

Using the *Fundamental Theorem of Calculus* on the right hand side of (17) and combining with the left hand side, we have

$$\int_{\Omega_\tau} p_\tau(\tau,\mathbf{x}) + \nabla\cdot\big(p(\tau,\mathbf{x})\mathbf{f}(\tau,\mathbf{x})\big) - \nabla\cdot\big(A(\mathbf{x})\nabla p(\tau,\mathbf{x})\big)\,d\mathbf{x} = 0 \tag{18}$$

Since this is true for all $\tau\in[0,t]$; and, t and $\Omega_t=\widehat{\Omega}$ are arbitrary, we may write the Fokker-Planck equation governing the evolution of $P_X(\mathbf{x})$†.

$$p_t(t,\mathbf{x}) + \nabla\cdot\big(p(t,\mathbf{x})\mathbf{f}(t,\mathbf{x})\big) = \nabla\cdot\big(A(\mathbf{x})\nabla p(t,\mathbf{x})\big) \tag{19a}$$

$$p(0,\mathbf{x}) = P_X(\mathbf{x}) \tag{19b}$$

In coordinates this is written as

$$\frac{\partial p(t,\mathbf{x})}{\partial t} + \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\bigg(p(t,\mathbf{x})f_i(t,\mathbf{x})\bigg) = \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\bigg(\sum_{j=1}^{n}A(\mathbf{x})_{ij}\frac{\partial p(t,\mathbf{x})}{\partial x_j}\bigg) \tag{20a}$$

$$p(0,\mathbf{x}) = P_X(\mathbf{x}) \tag{20b}$$

The Fokker-Planck Equation II

We now consider the evolution of a probability density function driven primarily by the second order differential system (12a-d), but which incorporates the dynamics of diffusion as well. If X and V are random variables and $\{\mathcal{X}(t,\mathbf{x},\mathbf{v})\ ,\ \mathcal{V}(t,\mathbf{x},\mathbf{v})\}$ is the pair of functions representing the solution of (11a-b), then let $p(t,\mathbf{x},\mathbf{v}) = P_{\mathcal{X}(t,X),\mathcal{V}(t,V)}(\mathbf{x},\mathbf{v})$ be the joint probability density function for the random variables $\{X,V\}$ transformed by $\{\mathcal{X}(t,\cdot)\ ,\ \mathcal{V}(t,\cdot)\}$. As before, let $\widehat{\Omega}$ be an arbitrary region in $R^n\text{x}R^n$, and let Ω_τ parameterize a family of regions in $R^n\text{x}R^n$ such that $\Omega_0=\Omega$ and $\Omega_t=\widehat{\Omega}$.

We assume, as in the previous section that the primary dynamics of the evolution of $p(t,\mathbf{x},\mathbf{v})$ is the differential system (12a-d). Additionally, we assume that probability diffuses through the $\mathbf{v}$ component of the regions Ω_τ based *only* on the gradient of p with respect to $\mathbf{v}$. That is, the instantaneous balance law (6) becomes

$$\frac{d}{d\tau}\int_{\Omega_\tau} p(\tau,\mathbf{x},\mathbf{v})\,d(\mathbf{x}\text{x}\mathbf{v}) = \oint_{\partial\Omega_\tau}(\,0\ \ ,\ A(\mathbf{x},\mathbf{v})\nabla_{\mathbf{v}}p(\tau,\mathbf{x},\mathbf{v})\,)\cdot\begin{pmatrix}\mathbf{n}_{\mathbf{x}}\\ \mathbf{n}_{\mathbf{v}}\end{pmatrix}\,d(\mathbf{S}_{\mathbf{x}}\text{x}\mathbf{S}_{\mathbf{v}}) \tag{21}$$

The left hand side has been previously calculated; so, using the *Fundamental Theorem of Calculus* on the right hand side, we may write

$$\int_{\Omega_\tau} p_\tau(\tau,\mathbf{x},\mathbf{v}) + \nabla_{\mathbf{x}}p(\tau,\mathbf{x},\mathbf{v})\cdot\mathbf{v} + \nabla_{\mathbf{v}}\cdot\big(p(\tau,\mathbf{x},\mathbf{v})\mathbf{f}(\tau,\mathbf{x},\mathbf{v})\big) - \nabla_{\mathbf{v}}\cdot\big(A(\mathbf{x},\mathbf{v})\nabla_{\mathbf{v}}p(\tau,\mathbf{x},\mathbf{v})\big)\,d(\mathbf{x}\text{x}\mathbf{v}) = 0 \tag{22}$$

† We make the same assumptions about $A(\mathbf{x})$ (in this section and the next) that were made when deriving the heat equation in n dimensions:

1. $A(\mathbf{x})$ should be a positive definite symmetric matrix;
2. The smallest eigenvalue of $A(\mathbf{x})$ is greater than a_0 for some $a_0>0$;
3. $A(\mathbf{x})$ is continuously differentiable with respect to $\mathbf{x}$.

Again, $\widehat{\Omega} = \Omega_t$ and t arbitrary imply that the Fokker-Planck equation governing the evolution of $P_{X,V}(\mathbf{x}, \mathbf{v})$ is

$$p_t(t, \mathbf{x}, \mathbf{v}) + \nabla_{\mathbf{x}} p(t, \mathbf{x}, \mathbf{v}) \cdot \mathbf{v} + \nabla_{\mathbf{v}} \cdot \big(p(t, \mathbf{x}, \mathbf{v}) \mathbf{f}(t, \mathbf{x}, \mathbf{v})\big) = \nabla_{\mathbf{v}} \cdot \big(A(\mathbf{x}, \mathbf{v}) \nabla_{\mathbf{v}} p(t, \mathbf{x}, \mathbf{v})\big) \tag{23a}$$

$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{23b}$$

Recalling that the coordinates for the right hand side have been written in the last section we may write the coordinate version of this Fokker-Planck equation as

$$\frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial t} + \sum_{i=1}^{n} \frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial x_i} v_i + \sum_{i=1}^{n} \frac{\partial}{\partial v_i} \big(p(t, \mathbf{x}, \mathbf{v}) f_i(t, \mathbf{x}, \mathbf{v})\big) = \sum_{i=1}^{n} \frac{\partial}{\partial v_i} \Big(\sum_{j=1}^{n} A(\mathbf{x}, \mathbf{v})_{ij} \frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial v_j} \Big) \tag{24a}$$

$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{24b}$$

In the case when $A(\mathbf{x}, \mathbf{v}) = \kappa \mathcal{A}$, $\mathcal{A}$ a constant positive definite symmetric matrix, (24a-b) becomes

$$\frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial t} + \sum_{i=1}^{n} \frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial x_i} v_i + \sum_{i=1}^{n} \frac{\partial}{\partial v_i} \big(p(t, \mathbf{x}, \mathbf{v}) f_i(t, \mathbf{x}, \mathbf{v})\big) = \kappa \sum_{i=1}^{n} \sum_{j=1}^{n} \mathcal{A}_{ij} \frac{\partial^2 p(t, \mathbf{x}, \mathbf{v})}{\partial v_i \, \partial v_j} \tag{25a}$$

$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{25b}$$

If $A(\mathbf{x}, \mathbf{v})$ is specialized further so that $\mathcal{A}$ is the identity matrix, then (25a-b) becomes

$$\frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial t} + \sum_{i=1}^{n} \frac{\partial p(t, \mathbf{x}, \mathbf{v})}{\partial x_i} v_i + \sum_{i=1}^{n} \frac{\partial}{\partial v_i} \big(p(t, \mathbf{x}, \mathbf{v}) f_i(t, \mathbf{x}, \mathbf{v})\big) = \kappa \sum_{i=1}^{n} \frac{\partial^2 p(t, \mathbf{x}, \mathbf{v})}{\partial^2 v_i} \tag{26a}$$

$$p(0, \mathbf{x}, \mathbf{v}) = P_{X,V}(\mathbf{x}, \mathbf{v}) \tag{26b}$$

In the one dimensional case, this is further simplified to

$$\frac{\partial p(t, x, v)}{\partial t} + \frac{\partial p(t, x, v)}{\partial x} v + \frac{\partial}{\partial v} \big(p(t, x, v) f(t, x, v)\big) = \kappa \frac{\partial^2 p(t, x, v)}{\partial^2 v} \tag{27a}$$

$$p(0, x, v) = P_{X,V}(x, v) \tag{27b}$$