SIAM J. CONTROL OPTIM.
Vol. 38, No. 5, pp. 1453–1467

SPECTRAL ANALYSIS OF FOKKER–PLANCK AND RELATED OPERATORS ARISING FROM LINEAR STOCHASTIC DIFFERENTIAL EQUATIONS*

DANIEL LIBERZON† AND ROGER W. BROCKETT‡

Abstract. We study spectral properties of certain families of linear second-order differential operators arising from linear stochastic differential equations. We construct a basis in the Hilbert space of square-integrable functions using modified Hermite polynomials, and obtain a representation for these operators from which their eigenvalues and eigenfunctions can be computed. In particular, we completely describe the spectrum of the Fokker–Planck operator on an appropriate invariant subspace of rapidly decaying functions. The eigenvalues of the Fokker–Planck operator provide information about the speed of convergence of the corresponding probability distribution to steady state, which is important for stochastic estimation and control applications. We show that the operator families under consideration can be realized as solutions of differential equations in the double bracket form on an operator Lie algebra, which leads to a simple expression for the flow of their eigenfunctions.

Key words. modified Hermite polynomials, linear stochastic differential equation, Fokker–Planck operator, double bracket equation

AMS subject classifications. 35P05, 35Q58, 93E03

PII. S0363012998338193

1. Introduction.

Given a system of stochastic differential equations, one can associate with it a (deterministic) partial differential equation which describes the evolution of the probability density with time. This so-called *Fokker–Planck equation* takes the form

$$\frac{\partial \rho}{\partial t} = L\rho, \tag{1}$$

where L is a second-order linear differential operator known as the *Fokker–Planck operator*. If g_0, g_1, g_2, ... are the eigenfunctions of L corresponding to distinct eigenvalues λ_0, λ_1, λ_2, ..., then the solution of (1) with initial condition

$$\rho(0, x) = \sum_{i=0}^{\infty} \alpha_i g_i(x), \qquad \alpha_i \in \mathbb{R},$$

is given by

$$\rho(t, x) = \sum_{i=0}^{\infty} \alpha_i e^{\lambda_i t} g_i(x).$$

*Received by the editors May 4, 1998; accepted for publication (in revised form) October 18, 1999; published electronically May 11, 2000.
http://www.siam.org/journals/sicon/38-5/33819.html

†Department of Electrical Engineering, Yale University, New Haven, CT 06520-8267 (liberzon@sysc.eng.yale.edu). The research of this author was supported in part by ARO grant DAAH04-95-1-0114, NSF grant ECS 9634146, and AFOSR grant F49620-97-1-0108.

‡Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 (brockett@hrl.harvard.edu). The research of this author was supported in part by Army DAAH 04-96-1-0445, Army DAAG 55-97-1-0114, and Army DAAL 03-92-G-0115.

Thus the eigenvalues of the Fokker–Planck operator L, particularly the one with the smallest magnitude, provide information about the speed of convergence of the probability distribution to steady state (when one exists), which is important in stochastic filtering and control applications. For a discussion along these lines and examples, see [3].

In the paper by Holley, Kusuoka, and Stroock [14], and more recently in [7], [8], [9], spectral properties of Fokker–Planck operators associated with certain types of nonlinear stochastic systems were investigated with the view towards applications to function minimization procedures. In this paper we confine our attention to Fokker–Planck operators that correspond to *linear* stochastic differential equations. An understanding of their spectral properties, besides being of interest in its own right, under certain circumstances helps shed some light on the nonlinear case (see [16, p. 88]). As in [14], we consider these operators as acting on a dense subspace of $L^2(\mathbb{R}^n)$ (rather than $L^1(\mathbb{R}^n)$ which might seem more natural from the probabilistic point of view). We apply a standard gauge transformation technique to convert them to self-adjoint operators, which greatly facilitates the analysis.

The paper starts with the one-dimensional case. Motivated by the explicit form of the steady-state probability density, we modify the classical Hermite polynomials by introducing one additional parameter σ (in our context, σ corresponds to the steady-state variance). This construction leads to an orthonormal basis for $L^2(\mathbb{R})$ with respect to which the operators under consideration take a particularly transparent form. The representation thus obtained allows us to compute their eigenvalues and eigenfunctions directly. As a result, we are able to provide a complete description of the spectrum of the Fokker–Planck operator on an appropriate invariant subspace of rapidly decaying functions. We then show that the essential features of this analysis carry over to the multidimensional case and enable us to obtain information about eigenvalues of Fokker–Planck operators in a more general setting.

Moreover, we observe that the operator families parameterized by σ can be described by differential equations on an operator Lie algebra which take the so-called *double bracket* form $\frac{dL}{d\sigma} = [L, [L, M]]$. This leads to a simple expression for the flow of the corresponding eigenfunctions. The study of differential equations in the double bracket form on finite-dimensional Lie algebras was initiated in [2] and [6] in connection with integrable gradient flows and numerical algorithms. It was shown, in particular, that such equations give rise to isospectral flows. In this paper we present what seems to be a new framework in which double bracket equations appear. The corresponding flows on an operator Lie algebra preserve the eigenvalues (actually, the entire spectrum in the self-adjoint case). This property is supported by probabilistic intuition.

The paper is organized as follows. In section 2 we construct an orthonormal basis in $L^2(\mathbb{R})$ using modified Hermite polynomials. In section 3 we study second-order differential operators arising from scalar linear stochastic differential equations. In section 4 we treat the multidimensional case, giving generalizations of the previous results. In section 5 we discuss double bracket differential equations on an operator Lie algebra and indicate connections with some known results on completely integrable gradient flows.

2. Orthonormal bases in $L^2(\mathbb{R})$.

It is well known (see, e.g., [15, p. 121]) that the *Hermite functions*

$$u_k(x) = h_k(x)e^{-x^2/2}, \qquad k = 0, 1, \ldots, \tag{2}$$

where $h_k(x) = \frac{1}{\pi^{1/4}\sqrt{2^k k!}}\, e^{x^2}\, \frac{d^k e^{-x^2}}{dx^k}$ are the *Hermite polynomials*, form an orthonormal basis for $L^2(\mathbb{R})$. We consider here the *modified Hermite polynomials*

$$h_k(x,\sigma) := \frac{\sqrt{\sigma^k}}{(\sigma\pi)^{1/4}\sqrt{2^k k!}} e^{x^2/\sigma} \frac{d^k e^{-x^2/\sigma}}{dx^k}\,, \qquad k = 0, 1, \dots,$$

where $\sigma > 0$ is a real parameter, and introduce the *modified Hermite functions*

$$u_k(x,\sigma) := h_k(x,\sigma) e^{-x^2/2\sigma} = c_k(\sigma) e^{x^2/2\sigma} \frac{d^k e^{-x^2/\sigma}}{dx^k} \tag{3}$$

with constants $c_k(\sigma)$ given by the relations

$$c_k(\sigma) = \frac{\sqrt{\sigma^k}}{(\sigma\pi)^{1/4}\sqrt{2^k k!}}. \tag{4}$$

The functions (3) reduce to those given by (2) for $\sigma = 1$. Various modifications of the classical Hermite polynomials, analogous to (and more general than) the one considered here, can be found in the literature [10], [13].

Lemma 1. *For any $\sigma > 0$, the functions* (3) *form an orthonormal basis for $L^2(\mathbb{R})$.*

Proof. We have

$$\begin{aligned}\langle u_k(x,\sigma), u_l(x,\sigma)\rangle &= c_k(\sigma)c_l(\sigma) \int_{-\infty}^{\infty} e^{x^2/\sigma} \frac{d^k e^{-x^2/\sigma}}{dx^k} \frac{d^l e^{-x^2/\sigma}}{dx^l} dx \\ &= c_k(1)c_l(1) \int_{-\infty}^{\infty} e^{y^2} \frac{d^k e^{-y^2}}{dy^k} \frac{d^l e^{-y^2}}{dy^l} dy,\end{aligned}$$

where we have made the change of variable $x = \sqrt{\sigma}\, y$. The statement of the lemma follows from the fact that the Hermite functions (2) form an orthonormal basis for $L^2(\mathbb{R})$. ∎

3. Fokker–Planck operators in $L^2(\mathbb{R})$. Let us consider the linear stochastic differential equation in the Itô sense

$$dx = -ax\, dt + b\, dw, \qquad a > 0, \tag{5}$$

where $x \in \mathbb{R}$ and w is a standard Wiener process. The reader may consult [11] for basic concepts of the theory of stochastic differential equations. The equation for the steady-state probability density that corresponds to (5) is $L(a,b)\rho(x) = 0$, where

$$L(a,b)\rho := \frac{b^2}{2}\rho_{xx} + ax\rho_x + a\rho \tag{6}$$

and ρ_x and ρ_{xx} denote the first and the second derivatives of ρ, respectively. The operator $L(a,b)$ is the *Fokker–Planck operator* associated with (5). Define

$$\sigma = \frac{b^2}{2a}. \tag{7}$$

The steady-state probability density is then given by $\bar\rho(x) = Ne^{-x^2/2\sigma}$, where $N > 0$ is a normalization constant. Dividing the Fokker–Planck operator $L(a,b)$ by a, we are led to studying a one-parameter family of differential operators, L_σ, defined by

$$L_\sigma\rho := \frac{1}{a}L(a, \sqrt{2a\sigma})\rho = \sigma\rho_{xx} + x\rho_x + \rho\,, \qquad \sigma > 0. \tag{8}$$

Before proceeding, we need to specify the domain of the above operators. It is easy to see that $L_\sigma u_k(x,\sigma) \in L^2(\mathbb{R})$ for each k, and L_σ is well defined by the formula (8) on the dense subspace U of $L^2(\mathbb{R})$ consisting of finite linear combinations of the functions $u_k(x,\sigma)$. We then define L_σ to be the minimal closed linear operator in $L^2(\mathbb{R})$ such that $L_\sigma\rho$ is given by (8) whenever $\rho \in C^2(\mathbb{R}) \cap L^2(\mathbb{R})$ and $\sigma\rho_{xx} + x\rho_x + \rho \in L^2(\mathbb{R})$. We thus obtain an operator $L_\sigma : \mathcal{D}_{L_\sigma} \to L^2(\mathbb{R})$, where $\mathcal{D}_{L_\sigma}$ is a dense subspace of $L^2(\mathbb{R})$ that contains U. Throughout the paper, unless specified otherwise, all differential operators are to be interpreted in the above sense.[1] For details on defining differential operators in this way, see [12].

The analysis of the operators L_σ is complicated by the fact that they are not self-adjoint. There is a standard technique which allows one to convert these operators to self-adjoint ones (this is sometimes referred to as *gauge*, or *ground state, transformation*). In our case, write $\rho = vf$, where the function v is to be fixed. We have

$$L_\sigma(vf) = \sigma v_{xx} f + 2\sigma v_x f_x + \sigma v f_{xx} + x v_x f + x v f_x + vf.$$

We see that in order for the first-order derivatives to disappear, v must satisfy the equation $v_x = -\dfrac{x}{2\sigma} v$. Letting

$$v = e^{-x^2/4\sigma} \tag{9}$$

we obtain $v^{-1}L_\sigma(vf) = \sigma f_{xx} + (\frac{1}{2} - \frac{x^2}{4\sigma})f$.

Motivated by the above discussion, we define a new operator family, T_σ, by the formula

$$T_\sigma\rho := \sigma\rho_{xx} + \left(\frac{1}{2} - \frac{x^2}{4\sigma}\right)\rho, \qquad \sigma > 0. \tag{10}$$

For any positive σ, the operator T_σ is closed and self-adjoint, its domain being a dense subspace $\mathcal{D}_{T_\sigma}$ of $L^2(\mathbb{R})$ (defined as explained before).

We know that $L_\sigma u_0(x,\sigma) = L_\sigma c_0 e^{-x^2/2\sigma} = 0$, i.e., $e^{-x^2/2\sigma}$ is an eigenfunction with the eigenvalue zero. To investigate the spectral properties of the operators L_σ and T_σ, it seems natural to use the basis given by the modified Hermite functions (3) (with the same value of σ). We first carry out direct calculations for the family L_σ, setting the stage for the multidimensional case. We will then see that the analysis of the self-adjoint operators T_σ is more straightforward and allows one to obtain precise information about the spectrum of the original Fokker–Planck operator on an appropriate space of rapidly decaying functions.

PROPOSITION 2. *The spectrum of the operator $L_\sigma : \mathcal{D}_{L_\sigma} \to L^2(\mathbb{R})$ is independent of σ. For any $\sigma > 0$, the eigenvalues of L_σ are all numbers in the half-plane $\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < 1/2\}$.*

---

[1] Alternatively, $C^2(\mathbb{R})$ here could be replaced by the space of functions $\rho : \mathbb{R} \to \mathbb{R}$ such that ρ_x exists and is absolutely continuous, i.e., the space of twice weakly differentiable functions for which the differential expression (8) is defined almost everywhere.

Proof.[2] Straightforward computations give

$$L_\sigma u_k(x,\sigma) = c_k(\sigma)\Bigg[\left(2+\frac{2x^2}{\sigma}\right)e^{x^2/2\sigma}\frac{d^k e^{-x^2/\sigma}}{dx^k} \tag{11}$$
$$+3xe^{x^2/2\sigma}\frac{d^{k+1}e^{-x^2/\sigma}}{dx^{k+1}} + \sigma e^{x^2/2\sigma}\frac{d^{k+2}e^{-x^2/\sigma}}{dx^{k+2}}\Bigg].$$

We introduce the notation $d_k(x,\sigma) = e^{x^2/2\sigma}\dfrac{d^k e^{-x^2/\sigma}}{dx^k}$, so that (11) becomes

$$L_\sigma u_k = c_k\left(\left(2+\frac{2x^2}{\sigma}\right)d_k + 3xd_{k+1} + \sigma d_{k+2}\right). \tag{12}$$

To obtain recurrence relations on d_k, notice that by Newton's binomial formula we have

$$d_{k+1} = e^{x^2/2\sigma}\left(-\frac{2x}{\sigma}\frac{d^k e^{-x^2/\sigma}}{dx^k} - \frac{2k}{\sigma}\frac{d^{k-1}e^{-x^2/\sigma}}{dx^{k-1}}\right)$$

which in the new notation becomes

$$d_{k+1} = -\frac{2x}{\sigma}d_k - \frac{2k}{\sigma}d_{k-1}\,. \tag{13}$$

From (13) we obtain

$$xd_k = -\frac{\sigma}{2}d_{k+1} - kd_{k-1}, \qquad k=1,2,\dots, \tag{14}$$

and also (multiplying both sides of (13) by x and then using (14))

$$\frac{2x^2}{\sigma}d_k = (2k+1)d_k + \frac{\sigma}{2}d_{k+2} + \frac{2k(k-1)}{\sigma}d_{k-2}. \tag{15}$$

Combining (12)–(15) gives

$$L_\sigma u_k = c_k\left(-kd_k + \frac{2k(k-1)}{\sigma}d_{k-2}\right) \tag{16}$$

and we see that the terms containing d_{k+2} disappear. Moreover, notice that we have

$$\frac{2k(k-1)}{\sigma}\,c_k = \sqrt{k(k-1)}\,c_{k-2}. \tag{17}$$

The formulas (16) and (17) imply that with respect to the basis (3) the operator L_σ takes the upper triangular form as given by

$$L_\sigma u_k(x,\sigma) = -ku_k(x,\sigma) + \sqrt{k(k-1)}u_{k-2}(x,\sigma). \tag{18}$$

From (18) it immediately follows that the spectrum of L_σ is independent of σ. Moreover, it is easy to see that the nonpositive integers are eigenvalues of L_σ. The

[2]The proofs in this section are given in sketched form; full details can be found in [16].

corresponding eigenfunctions are finite linear combinations of the basis elements u_k and thus belong to $C^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$. They do not, however, form a complete set of eigenfunctions. The formula (18) implies that the existence of an eigenfunction of L_σ with an eigenvalue λ is equivalent to the convergence of at least one of the series

$$\sum_{n=1}^{\infty} \frac{\lambda^2(\lambda+2)^2 \cdots (\lambda+2n-2)^2}{(2n)!}$$

and

$$\sum_{n=1}^{\infty} \frac{(\lambda+1)^2(\lambda+3)^2 \cdots (\lambda+2n-1)^2}{(2n+1)!}.$$

Using Gauss' test for convergence (see, e.g., [18]), one can show in a straightforward manner that each series converges if $\mathrm{Re}\lambda < 1/2$ and diverges if $\mathrm{Re}\lambda \geq 1/2$. □

We can gain more insight into the spectral properties of the operator L_σ from its probabilistic interpretation. Recall that L_σ was defined in terms of the Fokker–Planck operator $L(a, b)$ via the formula (8). It follows from Proposition 2 that the eigenvalues of $L(a, b)$ are all numbers in the half-plane $\{\lambda \in \mathbb{C} : \mathrm{Re}\,\lambda < a/2\}$. (However, it can be deduced from (1) that any eigenfunction of $L(a, b)$ that is nonnegative and belongs to $L^1(\mathbb{R})$ must be proportional to the steady-state probability density, which corresponds to the eigenvalue zero.) The fact that the spectrum of $L(a, b)$ does not depend on the noise coefficient b should not be surprising if we notice that we can change b by simply rescaling x; i.e., substituting $y = px$ in (5) for an arbitrary $p \in \mathbb{R}$ gives $\dot{y} = -ay + pb\dot{w}$. It is easy to check that the spectrum of the Fokker–Planck operator associated with (5) is not affected by such changes of variable.

We now turn our attention to the family of self-adjoint operators T_σ defined by (10). It is well known that Hermite polynomials appear frequently in expressions for eigenfunctions of self-adjoint linear second-order differential operators. The next proposition shows that the eigenfunctions of T_σ are given by the modified Hermite functions (3) and is to be considered as a preparation for a more general result to be presented in the next section. For $\sigma = 1/2$, the statement reduces to a standard result involving the classical Hermite functions (see, e.g., [1, p. 256]).

PROPOSITION 3. *For any $\sigma > 0$, the spectrum of the operator $T_\sigma : \mathcal{D}_{T_\sigma} \to L^2(\mathbb{R})$ consists of the nonpositive integers, all of which are eigenvalues. The corresponding eigenfunctions are the functions $u_k(x, 2\sigma)$, i.e., $T_\sigma u_k(x, 2\sigma) = -k\, u_k(x, 2\sigma)$.*

Proof. For $\rho = e^{x^2/4\sigma} \frac{d^k e^{-x^2/2\sigma}}{dx^k}$ one can verify that

$$T_\sigma \rho = e^{x^2/4\sigma} \frac{d^k e^{-x^2/2\sigma}}{dx^k} + x e^{x^2/4\sigma} \frac{d^{k+1} e^{-x^2/2\sigma}}{dx^{k+1}} + \sigma e^{x^2/4\sigma} \frac{d^{k+2} e^{-x^2/2\sigma}}{dx^{k+2}}$$

which in our previous notation becomes

$$T_\sigma d_k(x, 2\sigma) = d_k(x, 2\sigma) + x d_{k+1}(x, 2\sigma) + \sigma d_{k+2}(x, 2\sigma). \tag{19}$$

Replacing σ by 2σ in (14) and substituting into (19), we arrive at

$$T_\sigma d_k(x, 2\sigma) = -k\, d_k(x, 2\sigma).$$

This immediately implies the second part of the statement. The first part of the statement follows from this, since we have found an orthonormal basis in $L^2(\mathbb{R})$ consisting of eigenfunctions of T_σ. □

As a consequence, the eigenvalues of the original operator L_σ restricted to the space of functions of the form $\rho = vf$, where v is given by (9) and $f \in \mathcal{D}_{T_\sigma}$, are the nonpositive integers. The eigenfunction that corresponds to the eigenvalue $-k$ is given by $e^{-x^2/4\sigma}u_k(x, 2\sigma) = c_k(2\sigma)\frac{d^k e^{-x^2/2\sigma}}{dx^k}$. This leads us to a complete characterization of the spectrum of the Fokker–Planck operator $L(a, b)$ restricted to an appropriate space of rapidly decaying functions. Namely, let us denote by $\mathcal{L}_\sigma$ the space of functions that can be represented by finite linear combinations of the form $\sum_{k=1}^m \alpha_k \frac{d^k e^{-x^2/2\sigma}}{dx^k}$, $\alpha_k \in \mathbb{R}$. From the definitions of L_σ and T_σ and from Proposition 3 we immediately obtain the following result.

COROLLARY 4. *The space $\mathcal{L}_\sigma$ is invariant with respect to the Fokker–Planck operator $L(a, b)$ associated with* (5). *The spectrum of the restriction of $L(a, b)$ to $\mathcal{L}_\sigma$ consists of the numbers* $0, -a, -2a, -3a, \dots$, *all of which are eigenvalues.*

Remark 1. The eigenfunctions of L_σ on $\mathcal{L}_\sigma$ found above form an orthonormal basis for the space $L^2(\mathbb{R}, e^{x^2/2\sigma}dx)$, on which the operator L_σ can be shown to be self-adjoint. If instead of $\mathcal{L}_\sigma$ we consider a dense subspace of $L^2(\mathbb{R}, e^{x^2/2\sigma}dx)$ containing $\mathcal{L}_\sigma$, which can be constructed as explained at the beginning of the section, the statement about the spectrum still applies. Clearly, this larger subspace is no longer invariant under the action of $L(a, b)$. The operator T_σ is convenient because it is self-adjoint with respect to the standard inner product on $L^2(\mathbb{R})$.

We see in view of (7) that as the value of a increases while the noise coefficient b stays constant, the rate of decay of functions in $\mathcal{L}_\sigma$ becomes more rapid and so does the convergence to steady state. If we fix one member of the family $\{T_\sigma : \sigma > 0\}$, say, $T_{1/2}$, then for any value of σ the operator T_σ can be expressed as $T_\sigma = \Theta_\sigma^{-1} T_{1/2} \Theta_\sigma$, where Θ_σ is the unitary operator defined by $\Theta_\sigma u_k(x, 2\sigma) = u_k(x, 1) = u_k(x)$. We will use this observation in section 5.

4. Fokker–Planck operators in $L^2(\mathbb{R}^n)$.

Consider the system of linear stochastic differential equations

$$dx = Ax\,dt + B\,dw\,, \qquad x \in \mathbb{R}^n, \tag{20}$$

where w is a standard m-dimensional Wiener process and A and B are matrices of suitable dimensions. Recall that *separable functions*, i.e., functions that can be expressed as products $\rho_1(x_1)\cdots\rho_n(x_n)$, span a dense subspace of $L^2(\mathbb{R}^n)$. Thus we can construct an orthonormal basis for $L^2(\mathbb{R}^n)$ by taking products of the modified Hermite functions (3) for each variable. The analysis of the previous section now directly generalizes to those linear stochastic systems in $\mathbb{R}^n$ whose equations are completely decoupled. In this case, the Fokker–Planck operator decomposes into a sum of Fokker–Planck operators of the kind considered above for each variable. Our earlier results then imply, in particular, that the sums of the eigenvalues of the matrix A are eigenvalues of the corresponding Fokker–Planck operator, and that the corresponding eigenfunctions belong to the space $C^\infty(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and can be explicitly constructed.

Although the analysis for the general multidimensional system (20) is more complicated than in the scalar case, results that parallel most of our earlier developments can be obtained. Let us denote the Fokker–Planck operator associated with (20) by L_n and consider it as being a closed operator defined on a dense subspace $\mathcal{D}_{L_n}$ of

$L^2(\mathbb{R}^n)$ (cf. section 3). We have the following expression for L_n:

$$L_n\rho = \frac{1}{2}\sum_{i,j=1}^{n}(BB^T)_{ij}\,\rho_{x_ix_j} - \sum_{i,j=1}^{n}A_{ij}x_j\rho_{x_i} - \mathrm{tr}A\cdot\rho. \tag{21}$$

From this point on, let us make the following two assumptions with regard to the system (20):

(a) The eigenvalues of A have negative real parts.

(b) (A,B) is a controllable pair (i.e., $\mathrm{rank}(B, AB, \ldots, A^{n-1}B) = n$).

Under these assumptions, the steady-state variance equation

$$AQ + QA^T + BB^T = 0 \tag{22}$$

associated with (20) has a positive definite symmetric solution Q. After an appropriate change of coordinates in $\mathbb{R}^n$ we can have $Q = \frac{1}{2}I$, so that $A = \Omega - BB^T$ with Ω skew-symmetric. Such a coordinate transformation does not change the eigenvalues of the Fokker–Planck operator L_n. The steady-state probability density then becomes $\bar{\rho}(x) = Ne^{-x^Tx}$, $N > 0$, and this is an eigenfunction that corresponds to the eigenvalue zero of the Fokker–Planck operator.

Next let us determine all eigenfunctions of L_n that take the form

$$\rho(x) = (h_1x_1 + \cdots + h_nx_n)\bar{\rho}(x) = h^Tx\bar{\rho}(x), \qquad h \in \mathbb{R}^n. \tag{23}$$

LEMMA 5. *Suppose that $A = \Omega - BB^T$, where $\Omega = -\Omega^T$. Then the function* (23) *is an eigenfunction of the operator L_n with eigenvalue λ if and only if h is an eigenvector of the matrix A with the same eigenvalue λ.*

Proof. Let ρ be of the form (23). Taking into account that $L_n\bar{\rho} = 0$, we have

$$\begin{aligned} L_n\rho &= \sum_{i,j=1}^{n}(BB^T)_{ij}(h)_i(-2x_j)\bar{\rho} - \sum_{i,j}A_{ij}x_j(h)_i\bar{\rho} \\ &= -\bar{\rho}\sum_{i,j=1}^{n}(A_{ji} + 2(BB^T)_{ji})(h)_jx_i \\ &= \bar{\rho}\sum_{i,j=1}^{n}(\Omega_{ij} - (BB^T)_{ij})(h)_jx_i = \sum_{i=1}^{n}(Ah)_ix_i\bar{\rho} \end{aligned}$$

and this obviously equals $\lambda\rho = \lambda\sum_i h_ix_i\bar{\rho}$ if and only if $Ah = \lambda h$. □

Denote by $h_1, \ldots, h_k$ the eigenvectors of A and by $\lambda_1, \ldots, \lambda_k$ the corresponding eigenvalues ($k \le n$). Now let us see how L_n acts on functions of the form

$$\rho(x) = \bar{\rho}(x)\prod_{m\in J}h_m^Tx, \tag{24}$$

where the product is taken over some index set J whose elements are (not necessarily distinct) positive integers no greater than n. Using Lemma 5 and the fact that

$L_n\bar\rho = 0$, we have

$$\begin{aligned} L_n(\bar\rho \textstyle\prod_{m\in J} h_m^T x) &= \displaystyle\sum_{i,j=1}^n \sum_{m,l\in J} (BB^T)_{ij}(h_m)_i(h_l)_j\bar\rho \prod_{p\in J\setminus\{m,l\}} h_p^T x \\ &+ \sum_{i,j=1}^n \sum_{m\in J} (BB^T)_{ij}(h_m)_i\bar\rho_{x_j} \prod_{p\in J\setminus\{m\}} h_p^T x \\ &- \sum_{i,j=1}^n \sum_{m\in J} A_{ij}x_j(h_m)_i\bar\rho \prod_{p\in J\setminus\{m\}} h_p^T x \\ &= \sum_{i,j=1}^n \sum_{m,l\in J} (BB^T)_{ij}(h_m)_i(h_l)_j\bar\rho \prod_{p\in J\setminus\{m,l\}} h_p^T x + \Big(\sum_{m\in J}\lambda_m\Big)\bar\rho \prod_{m\in J} h_m^T x. \end{aligned}$$

Thus functions of the form (24) for various index sets J form an invariant subspace under the action of L_n. It is not hard to see that $\sum_{m\in J}\lambda_m$ are eigenvalues of L_n. The corresponding eigenfunctions are finite linear combinations of functions of the form (24). Summarizing, we have the following theorem.

Theorem 6. *The sums of the eigenvalues of the matrix A are eigenvalues of the Fokker–Planck operator $L_n : \mathcal{D}_{L_n} \to L^2(\mathbb{R}^n)$.*

Theorem 6 can probably be best appreciated in the following context. It is well known and easy to show that there are $N_n^p = \binom{n+p-1}{p}$ linearly independent monomials of degree p in n variables of the form $x_1^{p_1}\dots x_n^{p_n}$, where $\sum_{i=1}^n p_i = p$ and $p_i \geq 0$. The linear differential equation

$$\dot x = Ax, \qquad x \in \mathbb{R}^n,$$

gives rise to the equation

$$\frac{d}{dt}x^{[p]} = A_{[p]}x^{[p]}, \qquad x^{[p]} \in \mathbb{R}^{N_n^p}.$$

One of the basic properties of the matrix $A_{[p]}$ defined in this way is that its eigenvalues are the p-term sums of the eigenvalues of A. As is shown in [4], the matrices $A_{[p]}$ are directly related to the pth moment equations for the system (20).

Theorem 6 shows that the situation in the infinite-dimensional case is consistent with the one described in the previous paragraph in the following sense. Associated with the system (20) we have the Fokker–Planck equation for the probability density

$$\frac{\partial\rho(t,x)}{\partial t} = L_n\rho(t,x).$$

The operator L_n is well defined on a dense subspace of $L^2(\mathbb{R}^n)$. We know that the basis elements in $L^2(\mathbb{R}^n)$ can be taken to be polynomials of an arbitrary degree multiplied by Gaussians, and we have shown that the sums (with an arbitrary number of terms) of the eigenvalues of A are eigenvalues of the operator L_n.

In view of the results of section 3, it would be interesting to obtain conditions under which it is possible to convert the Fokker–Planck operator L_n to a self-adjoint operator by means of an appropriate gauge transformation. The following result provides such conditions, as well as an explicit formula for the function v to be used.

Proposition 7. *Suppose that the matrix B is nondegenerate and that we have*

$$ABB^T = BB^TA^T. \tag{25}$$

If the function v is defined by the formula

$$v = e^{x^T (BB^T)^{-1} Ax/2}, \tag{26}$$

then the operator T_n given by

$$T_n \rho = v^{-1} L_n(v\rho)$$

is self-adjoint.

Proof. The first-order terms in the expression for T_n are

$$\frac{1}{2} \sum_{j,k=1}^{n} (BB^T)_{jk} \left(v_{x_j} \frac{\partial}{\partial x_k} + v_{x_k} \frac{\partial}{\partial x_j} \right) - \sum_{i,j=1}^{n} A_{ij} x_j v \frac{\partial}{\partial x_i}.$$

We see that the coefficient of $\dfrac{\partial}{\partial x_i}$ is

$$\sum_{j=1}^{n} (BB^T)_{ij} v_{x_j} - \sum_{j=1}^{n} A_{ij} x_j v$$

and we need this to be zero for each i. This is equivalent to having

$$(BB^T) \operatorname{grad} v = Ax\, v$$

or

$$\operatorname{grad} v = (BB^T)^{-1} Ax\, v. \tag{27}$$

Therefore, we must have

$$\begin{aligned} v_{x_i x_j} &= \frac{\partial}{\partial x_i} \left[\sum_{k=1}^{n} ((BB^T)^{-1} A)_{jk} x_k\, v \right] \\ &= ((BB^T)^{-1} A)_{ji}\, v + \sum_{k=1}^{n} ((BB^T)^{-1} A)_{jk} x_k\, v_{x_i} \\ &= ((BB^T)^{-1} A)_{ji}\, v + \sum_{k=1}^{n} ((BB^T)^{-1} A)_{jk} x_k \sum_{l=1}^{n} ((BB^T)^{-1} A)_{il} x_l\, v. \end{aligned}$$

The compatibility conditions $v_{x_i x_j} = v_{x_j x_i}$ now imply that the matrix $(BB^T)^{-1}A$ has to be symmetric:

$$(BB^T)^{-1} A = A^T (BB^T)^{-1}.$$

Multiplying both sides of this formula by BB^T, we arrive at (25). It is straightforward to show that the function given by (26) satisfies (27). ∎

Let us switch to coordinates in which $Q = \sigma I$ for some $\sigma > 0$ (in the language of statistical thermodynamics, these are coordinates in which the *equipartition of energy* property holds, and σ is the steady-state *temperature* of the system). Then $(A + A^T)\sigma = -BB^T$, and (25) can be rewritten as $A^2 = (A^T)^2$. This last condition is satisfied, for example, if A is symmetric. In this case (26) becomes

$$v = e^{-x^T x/4\sigma} \tag{28}$$

which is a constant multiple of the square root of the steady-state probability density. This is in accordance with our earlier results for the one-dimensional case.

Denote by ∇ the gradient with respect to the metric on $\mathbb{R}^n$ given by $G = (BB^T)^{-1}$. In other words, given a function $\phi : \mathbb{R}^n \to \mathbb{R}$, we define the vector $\nabla\phi$ by $(\nabla\phi)_i = \sum_{j=1}^n (BB^T)_{ij}\phi_{x_j}$. Assume that $A = A^T$ and that the positive definite solution of (22) is $Q = \sigma I$. Let $\phi(x) = \frac{1}{4\sigma}x^T x$. Then we have

$$(\nabla\phi)_i = 2\sigma \sum_{j=1}^n A_{ij}\phi_{x_j} = \sum_{j=1}^n A_{ij}x_j$$

so that the system (20) can be rewritten as

$$dx = -\nabla\phi(x)\,dt + B\,dw. \tag{29}$$

Systems of the general form (29) and the corresponding steady-state probability densities were studied in [8].

Under the present assumptions, the Fokker–Planck operator takes the form

$$L_{n,\sigma}\rho = -\sigma \sum_{i,j=1}^n A_{ij}\rho_{x_i x_j} - \sum_{i,j=1}^n A_{ij}x_j\rho_{x_i} - \mathrm{tr}A \cdot \rho.$$

Using Proposition 7, we can also construct a self-adjoint operator which in this case is given by

$$T_{n,\sigma}\rho = -\sigma \sum_{i,j=1}^n A_{ij}\rho_{x_i x_j} - \left(\frac{1}{2}\,\mathrm{tr}A - \frac{1}{4\sigma}\sum_{i,j=1}^n A_{ij}x_i x_j \right)\rho.$$

As we have done throughout the paper, we consider the above expression as defining a closed operator acting on a dense subspace of $L^2(\mathbb{R}^n)$ which we denote by $\mathcal{D}_{T_{n,\sigma}}$. The following result is to be viewed as a generalization of Proposition 3 to the case of the multidimensional system (20) written in equipartition coordinates as explained earlier ($Q = \sigma I$), under the assumption that in these coordinates the nonrandom part of the system is symmetric ($A = A^T$). As shown above, this system is of the gradient form (29) for an appropriate quadratic function ϕ and a suitable constant metric.

THEOREM 8. *For any $\sigma > 0$, the spectrum of the operator $T_{n,\sigma} : \mathcal{D}_{T_{n,\sigma}} \to L^2(\mathbb{R}^n)$ consists of eigenvalues which are the sums of the eigenvalues of the matrix A.*

Proof. The matrix A has n real negative eigenvalues $\lambda_1, \dots, \lambda_n$. There exists an orthogonal matrix R such that $RAR^T = D$, where $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Making the change of variables $y = Rx$, we obtain an operator $\overline{T}_{n,\sigma}$ given by

$$\overline{T}_{n,\sigma}\rho = -\sigma\sum_{i=1}^n \lambda_i \rho_{y_i y_i} - \left(\frac{1}{2}\sum_{i=1}^n \lambda_i - \frac{1}{4\sigma}\sum_{i=1}^n \lambda_i y_i^2 \right)\rho.$$

The spectrum of $\overline{T}_{n,\sigma}$ is the same as that of $T_{n,\sigma}$. We have $\overline{T}_{n,\sigma} = -\sum_{i=1}^n \lambda_i T_{\sigma,y_i}$, where T_{σ,y_i} are the operators considered in section 3 for each variable (cf. remarks made at the beginning of this section). To complete the proof, recall that by Proposition 3 the eigenvalues of T_{σ,y_i} are the nonpositive integers. The eigenfunctions of $\overline{T}_{n,\sigma}$ are given by the products of the functions $u_k(y_i, 2\sigma)$ for each variable; they form an orthonormal basis for $L^2(\mathbb{R}^n)$. $\square$

As before, we conclude that the eigenvalues of the original operator $L_{n,\sigma}$ restricted to the space of functions of the form $\rho = vf$, where v is given by (28) and $f \in \mathcal{D}_{T_{n,\sigma}}$, are the sums of the eigenvalues of A. The corresponding eigenfunctions take the form $vg_k(x,\sigma)$, where $g_k(x,\sigma)$, $k = 0, 1, \ldots$ are the eigenfunctions of $T_{n,\sigma}$ described in the proof of Theorem 8. Let $\mathcal{L}_{n,\sigma}$ denote the space of functions $\{\rho : v^{-1}\rho = \sum_{k=1}^{m} \alpha_k g_k(x,\sigma)\}$, $\alpha_k \in \mathbb{R}$. As a generalization of Corollary 4 we have the following statement.

COROLLARY 9. *The space $\mathcal{L}_{n,\sigma}$ is invariant with respect to the Fokker–Planck operator $L_{n,\sigma}$ associated with* (29). *The spectrum of the restriction of $L_{n,\sigma}$ to $\mathcal{L}_{n,\sigma}$ consists of eigenvalues which are the sums of the eigenvalues of the matrix A.*

Remark 2. Under the change of variables described in the proof of Theorem 8 the operator $L_{n,\sigma}$ becomes a Fokker–Planck operator associated with a decoupled system. This makes the statement of Corollary 9 obvious in view of Corollary 4 and the discussion at the beginning of this section. In the case when A is a scalar multiple of the identity matrix, the spectrum (but not the eigenfunctions) of the operator $L_{n,\sigma}$ on $\mathcal{L}_{n,\sigma}$ can be obtained from the analysis of its adjoint presented in [19, section 7.5]. The eigenfunctions of $L_{n,\sigma}$ on $\mathcal{L}_{n,\sigma}$ found above form an orthonormal basis for the space $L^2(\mathbb{R}^n, e^{x^T x/4\sigma}dx)$. We could also consider $L_{n,\sigma}$ as acting on a larger dense subspace of $L^2(\mathbb{R}^n, e^{x^T x/4\sigma}dx)$, which would not change the spectrum—cf. Remark 1 in section 3.

It is interesting to notice that, given the original system (20), we can always find a basis in which $Q = \sigma I$ satisfies (22) and $A = A^T$ if A is allowed to depend on time. First, switch to an equipartition basis in which we have $Q = \sigma I$. Note that the last equality is preserved under orthogonal coordinate transformations. Let $\Omega = \frac{1}{2}(A - A^T)$. Making the change of variable $y = e^{-\Omega t}x$ in (20), we obtain

$$dy = \frac{1}{2}e^{-\Omega t}(A + A^T)e^{\Omega t}y\,dt + e^{-\Omega t}B\,dw \tag{30}$$

and the first term features a symmetric matrix as needed.

5. Double bracket equations. Consider the operators P_1, P_2, P_3, and P_4 acting on the space

$$\mathcal{D} = \{\rho \in C^2(\mathbb{R}) \cap L^2(\mathbb{R}) : \rho_{xx}, x\rho_x, x^2\rho \in L^2(\mathbb{R})\}$$

that are defined as follows:

$$P_1\rho = \rho_{xx}, \qquad P_2\rho = x\rho_x, \qquad P_3\rho = x^2\rho, \qquad P_4\rho = \rho.$$

It is easy to verify that the linear span of the above operators is closed under commutation with respect to the usual Lie bracket $[P_i, P_j] = P_iP_j - P_jP_i$. We will let $\mathfrak{g}$ denote the operator Lie algebra spanned by P_i, $i = 1, 2, 3, 4$. Such Lie algebras and their representations have been studied in the context of quantum mechanics and, more recently, estimation theory [5].

Observe that L_σ and T_σ can be realized as operators in $\mathfrak{g}$ because $\mathcal{D} \subset \mathcal{D}_{L_\sigma}$ for each $\sigma > 0$. More precisely, let us denote by $L(\sigma)$ and $T(\sigma)$ the restrictions of L_σ and T_σ to $\mathcal{D}$. Proposition 3 implies that $T(\sigma)$, $0 < \sigma < \infty$, is an isospectral family of operators in $\mathfrak{g}$. In fact, for any $\sigma > 0$ the spectrum of $T(\sigma)$ consists of eigenvalues which are the nonpositive integers. As we know from Proposition 2, the nonpositive integers are also eigenvalues of $L(\sigma)$ for each $\sigma > 0$ (because the corresponding eigenfunctions of L_σ belong to $\mathcal{D}$).

In this section we show that the families of operators $L(\sigma)$ and $T(\sigma)$ correspond to integral curves of differential equations in the double bracket form on $\mathfrak{g}$. We also obtain the corresponding dynamical representation for the family of modified Hermite functions defined by (3). The proofs are completely straightforward calculations and will not be given.

PROPOSITION 10. *Let M be an operator in $\mathfrak{g}$ defined by $M\rho = \frac{1}{4}\rho_{xx}$. Then $L(\sigma)$, $0 < \sigma < \infty$, is a solution of the differential equation*

$$\frac{dL}{d\sigma} = [L, [L, M]]. \tag{31}$$

The Fokker–Planck operator associated with (5) is uniquely determined by two parameters: σ, which corresponds to the steady-state variance, and a, which describes the speed of convergence to steady state. In making the transition to the operators L_σ we factored out the dependence on a. Thus the flow (31) can be thought of as evolving on the "slice" of Fokker–Planck operators with the same spectral properties but different steady states. As we will see, in the multidimensional case σ corresponds to the steady-state temperature of the system (defined in section 4).

To each of the operators $L(\sigma)$ there corresponds the steady-state probability density ρ_σ which satisfies the equation $L(\sigma)\rho_\sigma(x) = 0$. The flow (31) on the operator Lie algebra $\mathfrak{g}$ thus induces a flow on the manifold of Gaussian probability densities. For example, making the change of variable $\sigma = e^t$, we obtain a particular case of the gradient flow of Gaussians described by Nakamura in [17].

PROPOSITION 11. *Let N be an operator in $\mathfrak{g}$ defined by $N\rho = \frac{1}{2}\rho_{xx}$. Then $T(\sigma)$, $0 < \sigma < \infty$, is a solution of the differential equation*

$$\frac{dT}{d\sigma} = [T, [T, N]]. \tag{32}$$

In view of the remarks made at the end of section 3, we can write $T(\sigma) = \Theta^{-1}(\sigma)T(1/2)\Theta(\sigma)$, with the domain of $\Theta(\sigma)$ properly defined. Using the fact that for all $\sigma > 0$ the operator $\Theta(\sigma)$ is unitary and the operators $T(\sigma)$ and $N(\sigma)$ are self-adjoint, we arrive at the equation

$$\frac{d\Theta}{d\sigma} = T(1/2)\Theta N - \Theta N \Theta^{-1} T(1/2)\Theta = \Theta[T, N] \tag{33}$$

which describes the evolution of the eigenbasis for $T(\sigma)$ induced by the flow (32). This is the same equation as the one obtained in [6] for the finite-dimensional case.

We point out an interesting analogy between the results of Propositions 10 and 11 and the sorting algorithms described in [6]. If N is a real diagonal matrix with unrepeated eigenvalues, and if $H(0)$ is a suitably chosen symmetric matrix, then the solution of the double bracket equation $\dot{H} = [H, [H, N]]$ approaches a diagonal matrix $H(\infty)$ such that the diagonal elements of $H(\infty)$ and N are similarly ordered; since $H(\infty)$ is diagonal, it commutes with N. For large positive values of σ, the "principal term" of the operators $L(\sigma)$ and $T(\sigma)$ is $\sigma\frac{d^2}{dx^2}$, which is proportional to both M and N and thus commutes with them. Thus the double bracket equations (31) and (32) can be thought of as performing a task of "operator sorting."

We would like to generalize the above results to the multidimensional case. Consider the operators $P_{1,i,j}$, $P_{2,i,j}$, $P_{3,i,j}$, and P_4 acting on the space

$$\mathcal{D}_n = \{\rho \in C^2(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) : \rho_{x_i x_j}, x_i\rho_{x_j}, x_i x_j \rho \in L^2(\mathbb{R}^n)\ \forall i, j = 1, 2, \dots, n\}$$

that are defined as follows:

$$P_{1,i,j}\rho = \rho_{x_i x_j}, \qquad P_{2,i,j}\rho = x_i \rho_{x_j}, \qquad P_{3,i,j}\rho = x_i x_j \rho, \qquad P_4\rho = \rho.$$

These operators span a Lie algebra which we denote by $\mathfrak{g}_n$. For each $\sigma > 0$, let $L_n(\sigma)$ and $T_n(\sigma)$ denote the restrictions of $L_{n,\sigma}$ and $T_{n,\sigma}$ to $\mathcal{D}_n$. Theorem 8 implies that $T_n(\sigma)$, $0 < \sigma < \infty$, is an isospectral family of operators in $\mathfrak{g}_n$.

PROPOSITION 12. *Let M be an operator in $\mathfrak{g}_n$ defined by*

$$M\rho = -\frac{1}{4}\sum_{i,j=1}^{n} (A^{-1})_{ij}\rho_{x_i x_j}.$$

Then $L_n(\sigma)$, $0 < \sigma < \infty$, is a solution of the differential equation

$$\frac{dL}{d\sigma} = [L, [L, M]].$$

PROPOSITION 13. *Let N be an operator in $\mathfrak{g}_n$ defined by*

$$N\rho = -\frac{1}{2}\sum_{i,j=1}^{n} (A^{-1})_{ij}\rho_{x_i x_j}.$$

Then $T_n(\sigma)$, $0 < \sigma < \infty$, is a solution of the differential equation

$$\frac{dT}{d\sigma} = [T, [T, N]].$$

As in the scalar case, we can define a unitary operator Θ_σ by $\Theta_\sigma g_k(x,\sigma) = g_k(x,1)$, where $g_k(x,\sigma)$ are the eigenfunctions of $T_n(\sigma)$. This operator will then satisfy (33), which describes the flow of these eigenfunctions. Finally, note that Propositions 12 and 13 apply to the system (30) without any changes (except that now $T_n(\sigma)$ will also depend on t).

Acknowledgments. The first author would like to thank Mark Adler for helpful discussions and Daniel Stroock for calling his attention to the book [19].

REFERENCES

[1] G. BIRKHOFF AND G. C. ROTA, *Ordinary Differential Equations. Introductions to Higher Mathematics*, Ginn, Boston, 1962.

[2] A. M. BLOCH, R. W. BROCKETT, AND T. S. RATIU, *Completely integrable gradient flows*, Comm. Math. Phys., 147 (1992), pp. 57–74.

[3] R. W. BROCKETT, *Lie algebras and Lie groups in control theory*, in Geometric Methods in System Theory, D. Q. Mayne and R. W. Brockett, eds., Reidel Publishing Co., Dordrecht, The Netherlands, 1973, pp. 43–82.

[4] R. W. BROCKETT, *Parametrically stochastic linear differential equations*, Math. Programming Stud., 5 (1976), pp. 8–21.

[5] R. W. BROCKETT, *Nonlinear systems and nonlinear estimation theory*, in Stochastic Systems: The Mathematics of Filtering and Identification and Applications, M. Hazewinkel and J. C. Willems, eds., NATO Adv. Study Inst. Ser. C: Math. Phys. Sci. 78, Reidel, Dordrecht, Boston, 1981, pp. 441–477.

[6] R. W. BROCKETT, *Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems*, Linear Algebra Appl., 146 (1991), pp. 79–91.

[7] R. W. BROCKETT, *Oscillatory descent for function minimization*, in Current and Future Directions in Applied Mathematics, M. Alber et al., eds., Birkhäuser, Boston, 1997, pp. 65–82.

[8] R. W. Brockett, *Notes on stochastic processes on manifolds*, in Systems and Control in the Twenty-First Century, C. I. Byrnes et al, eds., Progr. Systems Control Theory 22 Birkhäuser, Boston, 1997, pp. 75–100.

[9] R. W. Brockett and D. Liberzon, *On explicit steady-state solutions of Fokker–Planck equations for a class of nonlinear feedback systems*, in Proceedings of the American Control Conference, Philadelphia, PA, 1998, pp. 264–268.

[10] R. Y. Chang and M. L. Wang, *The properties of modified Hermite polynomials and their applications to functional differential equations*, J. Chinese Inst. Engrs., 9 (1986), pp. 75–81.

[11] I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations*, Springer-Verlag, New York, 1972.

[12] S. Goldberg, *Unbounded Linear Operators*, McGraw-Hill, New York, 1966.

[13] H. W. Gould and A. T. Hopper, *Operational formulas connected with two generalizations of Hermite polynomials*, Duke Math. J., 29 (1962), pp. 51–63.

[14] R. A. Holley, S. Kusuoka, and D. W. Stroock, *Asymptotics of the spectral gap with applications to the theory of simulated annealing*, J. Funct. Anal., 83 (1989), pp. 333–347.

[15] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, UK, 1982.

[16] D. Liberzon, *Asymptotic Properties of Nonlinear Feedback Control Systems*, Ph.D. Thesis, Brandeis University, Waltham, MA, 1998.

[17] Y. Nakamura, *Completely integrable gradient flows on the manifolds of Gaussian and multinomial distributions*, Japan J. Indust. Appl. Math., 10 (1993), pp. 179–189.

[18] E. D. Rainville, *Infinite Series*, Macmillan, New York, 1967 .

[19] D. W. Stroock, *Probability Theory, an Analytic View*, Cambridge University Press, Cambridge, UK, 1993 .