TECHNICAL RESEARCH REPORT

Group Invariance and Symmetries in Nonlinear Control and Estimation

by John S. Baras

T.R. 2000-30

Group Invariance and Symmetries in Nonlinear Control and Estimation

John S. Baras
Department of Electrical and Computer Engineering,
Department of Computer Science and
the Institute for Systems Research,
University of Maryland, College Park, MD 20742
e-mail: baras@isr.umd.edu

Abstract

We consider nonlinear filtering problems, nonlinear robust control problems and the partial differential equations that characterize their solutions. These include the Zakai equation, and in the robust control case two coupled Dynamic Programming equations. We then characterize equivalence between two such problems when we can compute the solution of one from the solution of the other using change of dependent, independent variables and solving an ordinary differential equation. We characterize the resulting transformation groups via their Lie Algebras. We illustrate the relationship of these results to symmetries and invariances in physics, Noether's theorem, and calculus of variations. We show how using these techniques one can solve nonlinear problems by reduction to linear ones.

1 Introduction

Symmetries have played an important role in mathematical physics as well as in systems and control. Symmetries in mathematical physics [1] are essential. Essentially all physics theories can be based in symmetries and symmetry properties. Some of the more celebrated results are:

(i) Conservation laws; various physics theories.

(ii) Quantum electrodynamics, elementary particles, quarks, strings.

(iii) Quantum field theory, reductions, symmetry braking.

Symmetries have been also fundamental in systems and control. Perhaps the most well known principle has been the unifying role that equivalences of internal and external representations and associated groups of transformations play in system theory. Some of the more celebrated results are:

(a) Electrical networks and realization theory.

(b) Feedback invariants.

(c) Nonlinear Filtering and Estimation Algebra.

(d) Parameterizations of Rational Transfer Functions.

(e) Canonical forms of linear analytic systems (linear in the controls).

(f) Feedback linearization.

(g) Symmetries in multibody mechanical systems and continuum mechanics

Given the rich interplay between mathematical physics and control systems, especially variational problems and optimal control, there remain many unexplored theoretical and applied aspects of symmetries for systems and control. In this paper we describe our research in this direction with focus on stochastic estimation and nonlinear control.

Both mathematical physics and systems and control deal with differential equations (DE). Therefore, symmetry groups of differential equations and systems of differential equations provide a natural starting point for understanding the key methods and concepts. As a simple example consider a scalar ordinary differential equation (ODE):

$$F\left(x, u, \frac{du}{dx}\right) = \frac{du}{dx} - f(x,u) = 0 \quad . \tag{1.1}$$

The left hand side of (1.1) can be viewed as defining a surface in $\mathbb{R}^3$ (three variables: $x, u, \frac{du}{dx}$). The middle term of (1.1) is the ODE and its solutions are scalar valued curves. A *Symmetry Group* of an ODE [1] is a group of transformations on (x,u) (the independent and dependent variables) which maps any solution of the DE to another solution of the DE. Similarly for systems of DEs. Thus finding symmetry groups for (1.1) amounts to finding transformations (diffeomorphisms)

$$\left.\begin{array}{ll} H : \mathbb{R}^2 & \to \mathbb{R}^2 \\ (x,u) & \mapsto (\phi(x,u), \psi(x,u)) \end{array}\right\} \tag{1.2}$$

which permute solution curves. Finding such groups is a celebrated old problem initiated by Lie and later extended by Ovsjannikov and many others. Continuing with this simple example, if we find such a transformation H, we can extend it to derivatives using the simple observation that if a curve passes through (x,u) with slope du/dx, its image (under H) passes through (x', u') with slope du'/dx' where:

$$\begin{aligned} x' &= \phi(x,u) \\ u' &= \psi(x,u) \\ \frac{du'}{dx'} &= \frac{(\psi_x + \psi_u \; du/dx)}{(\phi_x + \phi_u \; du/dx)} \end{aligned} \tag{1.3}$$

The map:

$$\begin{array}{rl} H' : \mathbb{R}^3 & \to \mathbb{R}^3 \\ \left(x, u, \frac{du}{dx}\right) & \mapsto \left(x', u', \frac{du'}{dx'}\right) \end{array} \tag{1.4}$$

is an extension of H.

A key theorem in the investigation of symmetry groups for ODEs is the establishment of the result that H permutes solutions of the ODE (1.1) iff H' leaves the surface in $\mathbb{R}^3$, defined by the left-hand side of (1.1) invariant. This equivalence gives to the problem of constructing symmetry groups of ODEs a very attractive geometric foundation. The so called Lie-Ovsjannikov method [1] for constructing symmetry groups of ODEs is to find all H' that have the surface $s = f(x, u)$ as an invariant manifold.

This idea, as worked out in this simple example, can be extended directly to n^{th} order ODEs and to systems of ODEs. For an n^{th} order ODE

$$\begin{aligned} F(x, u, u_1, u_2, \cdots, u_{n-1}) &= u_n - f(x, u_1, u_2, \cdots, u_{n-1}) = 0 \\ u_i &= \frac{d^i u}{dx^i} \end{aligned} \tag{1.5}$$

one extends a transformation H on (x, u), n-fold to a transformation H' on $(x, u, u_1, \cdots, u_n)$. The Lie-Ovsjannikov method (1.1) extends as well.

More interesting are one-parameter Lie groups which leave the solutions of (1.5) invariant:

$$\left.\begin{aligned} x' &= X(x, u; \epsilon) \\ u' &= U(x, u; \epsilon) \end{aligned}\right\} \tag{1.6}$$

The infinitesimal generator of this group

$$X = \xi(x, u)\ \frac{\partial}{\partial x}\ + n(x, u)\ \frac{\partial}{\partial u} \tag{1.7}$$

plays a fundamental role. One extends the group n-fold to derivatives of all orders to get an n^{th} order group and n^{th} order infinitesimal generator $X^{(n)}$ [1]. Some of the basic results of the theory are:

(i) The one-parameter Lie group leaves the ODE invariant iff its n^{th} order extension leaves the surface $F = 0$ invariant.

(ii) The family of all solution curves is invariant under the Lie group iff it is a symmetry group.

(iii) **Theorem** (Lie): The one-parameter Lie group is an invariance group of $F = 0$, iff

$$\begin{aligned} X^{(n)} &= (u_n - f(x, u, u_1, \cdots, u_{n-1})) = 0, \quad \text{when} \\ u_n &= f(x, u, u_1, \cdots, u_{n-1}) \ . \end{aligned}$$

The consequences of these foundations were pursued by Lie who showed how to construct the Lie group (of invariance), and that the Lie algebra of infinitesimal generators determine the local Lie group.

The subject has attracted many researchers through the years. Some of the more interesting results that have been obtained are:

- The reduction of the intractable nonlinear conditions of group invariance to linear homogeneous equations, which determine the infinitesimal generators.
- Invariance of an ODE under a one-parameter Lie group of point transformations leads to reduction of the order of the ODE by one.
- Invariance of an n^{th} order ODE under an r-parameter Lie group with solvable Lie algebra is reduced to an $(n-r)^{\text{th}}$ order ODE plus r quadratures (integrals).
- Invariance of a linear partial differential equation (PDE) under a Lie group leads to superposition of solutions in terms of transforms.

We are more interested in symmetry and invarance groups of PDEs. These transformation groups are *local* Lie groups. *Point symmetries* are point transformations on the space of independent and dependent variables. *Contact symmetries* are contact transformations acting on the space of independent, dependent variables and derivatives of dependent variables. Ovsjannikov showed that if a system of PDEs is invariant under a Lie group, we can find special solutions of the system, which are called *similarity solutions*.

A further generalization of these concepts with key significance for both systems and control and mathematical physics are the Lie-Bäcklund symmetries (or transformations) [1]. In these transformations the infinitesimal generators of the local Lie groups depend on derivaties of the dependent variables up to any finite order:

$$\left.\begin{aligned} x' &= x + \epsilon\xi(x,u,u_1,u_2,\cdots,u_p) + O(\epsilon^2) \\ u' &= u + \epsilon\eta(x,u,u_1,u_2,\cdots,u_p) + O(\epsilon^2) \end{aligned}\right\} \tag{1.8}$$

It is a basic result in the theory of such transformations that the infinitesimal generators can be computed by a simple extension of Lie's aglorithm. Another key result is that invariance of a PDE under a Lie-Bäcklund symmetry usually leads to invariance under an infinite number of symmetries (connected by recursion operators) [1].

A most celebrated results in variational problems with foundational consequences in mathematical physics is E. Noether's Theorem [1]. Euler-Lagrange equations are the governing equations of many physical systems; they are of fundamental importance in mathematical physics. Euler-Lagrange equations provide the dynamics of systems from a variational formulation (typically energy-based variational formulation). These ideas from mathematical physics have inspired many research efforts in systems and control: from stability theory, to dissipative systems, to communication network routing, to robot path planning (to mention just a few).

In this context a physical system has independent variables x in $\mathbb{R}^n$ and dependent variables u in $\mathbb{R}^m$. The system independent variables x can take values in a domain Ω of $\mathbb{R}^n$. A function is given or constructed which depends on the independent variables x, the dependent variables u, and derivatives of the dependent variables up to order k, $u_1, u_2, \cdots, u_k$. The dynamics of the system evolve so that the paths $u(x)$ correspond to extremals of the integral

$$J(u) = \int_\Omega L(x,u,u_1,u_2,\cdots,u_k)\ dx \tag{1.9}$$

The function L is called a *Lagrangian* and the integral $J(u)$ an *action integral.* The path $u(x) = (u^1(x), u^2(x), \cdots, u^m(x))$ describes the state evolution of the system and typically has to satisfy some boundary conditions on $\partial\Omega$. Such formulations are well known and used by control theorists and practitioners. Clearly, if $u(x)$ is an extremum of (1.9), any infinitesimal change

$$u(x) \mapsto u(x) + \epsilon v(x) \tag{1.10}$$

which also satisfies the boundary conditions, should leave $J(u)$ unchanged to order $O(\epsilon)$.

The most significant relationship of this formulation is with respect to conservation laws of a system. A *conservation law* of a system, is an equation in divergence-free form

$$div\ f = \sum_{i=1}^{n} D_i\ f_i(x, u, u_1, u_2, \cdots, u_k) = 0 \tag{1.11}$$

Equation (1.11) must hold for any extremal function $u(x)$ of (1.9). The vector f is called a conserved *flux* [1] since (1.11) implies that a net flow of f through any closed surface in the space x is zero.

Euler-Lagrange equations are the (often dynamical) equations that need to be satisfied by an extremum of (1.9). We refer to [1, pp. 254-257] for a concise and clear derivation. The Euler-Lagrange equations can be ODEs or PDEs dependent on the problem. As such, we may ask the question if they have symmetry groups (invariance groups). Noether's key idea and result was that in order to find conservation laws it is far more fruitful to investigate the invariance of the action integral (1.9). Noether considered Lie-B*ä*cklund transformations that leave the action integral invariant:

$$\begin{aligned} x' &= x + \epsilon\xi(x, u, u_1, u_2, \cdots, u_p) + O(\epsilon^2) \\ u' &= u + \epsilon\eta(x, u, u_1, u_2, \cdots, u_p) + O(\epsilon^2) \end{aligned} \tag{1.12}$$

Noether showed that the existence of such transformations lead constructively to conservation laws of the corresponding Euler-Lagrange equations. She established the explicit relationship between the infinitesimals ξ, η and the conserved flux f. For a concise proof of this celebrated theorem we refer to [1, p. 257-260]. This celebrated theorem induced fundamental reformulations of mathematical physics, bringing a certain degree of unification. They include:

- Invariance under time translation leads to energy conservation.
- Invariance under translation or rotation in space, leads to conservation of linear or angular momentum.
- Relativity theory formulations.

The relationship with symmetry groups allows the determination of variational symmetries using Lie's algorithm. Noether's theorem resulted in many specific applications (for specific physical systems or phenomena) such as [1]:

- Conservation of the Runge-Lenz vector in Kepler's problem.
- Existence of infinity of conservation laws for the Korteweg-deVries equation and other soliton equations.

We close this brief review of the history of research on symmetry groups for ODEs, PDEs and dynamical systems by listing some more recent results and activities. Symmetry groups allow discovery of related DEs of simpler form. This for instance leads to transformations that map a given equation to a target equation. Comparing the Lie groups of symmetries admitted by each equation, actually allows the construction of the mapping. Such results have significant implications in facilitating the solution of new ODEs and PDEs using solutions of other ODEs and PDEs, known already.

For our subject, it is important to consider transformations beyond local symmetries. These are transformations where the dependence on u and the derivatives of u is global (*i.e.* not just through the instantaneous values $u(x)$). Gauge transformations in mathematical physics and quantum field theory can be such global transformations. In the theory of symmetry groups such transformations are called potentials.

Ideas, techniques and algorithms from symmetry groups have made fundamental contributions to mathematical physics. As it should be clear from this brief exposition there is great potential for similar impact and fundamental new advances by the systematic explowhatitation of symmetry groups in systems and control problems. Many of the advances in mathematical physics came out of application of symmetry groups in fundamental PDEs of mathematical physics. This inspires us to apply similar techniques in the fundamental PDEs of systems and control: dynamic programming, Zakai equations for nonlinear filtering, information state equations for robust control and others. In addition Noether's theorem can lead to significant advances in nonlinear optimization. The results described in the subsequent sections are but a small set of what could be accomplished by such methods in systems and control

2 Constructive Use of Symmetry Groups of PDEs: A Simple Example

An interesting, for systems and control (as we shall see) theory, application of ideas from symmetry groups is the following. Use the symmetry group of a PDE to compute easily solutions to new PDEs. This is a non-conventional use of symmetry groups developed by Rosencrans [14].

To explain the idea clearly, we use the simple example of the heat equation.

$$\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u(t,x)}{\partial x^2}. \tag{2.1}$$

It is well known [7][12] that (2.1) is invariant under the variable transformation

$$\left.\begin{array}{rcl} x & \longmapsto & e^s x \\ t & \longmapsto & e^{2s} t. \end{array}\right\} \tag{2.2}$$

That is to say if $u(t,x)$ is a solution of (2.1), so is $u(e^{2s}t, e^s x)$. Clearly the initial data should be changed appropriately. So if ϕ is the initial data for u, the initial data for the transformed (under 2.2)) solution are $\phi(e^s x)$. This elementary invariance can be written symbolically as

$$e^{tD^2} e^{sxD} \phi = e^{sxD} e^{e^{2s} tD^2} \phi. \tag{2.3}$$

Here

$$D: = \frac{\partial}{\partial x} \tag{2.4}$$
$$D^2: = \frac{\partial^2}{\partial x^2} \quad .$$

Often in this paper we shall give double meaning to exponentials of partial differential operators. Thus while $exp(tD^2)$ in (2.3) denotes the semigroups generated by D^2 [13], $exp(sxD)$ is viewed as an element of the Lie group of transformations generated by xD. It is easy to verify that

$$\phi(e^s x) = [exp(sxD)\phi](x) \quad , \tag{2.5}$$

where we view $exp(sxD)$ as such a transformation, with parameter s. Now the association

$$(t,s) \longmapsto e^{tD^2} e^{sxD} \tag{2.6}$$

defines a two parameter semigroup with product rule

$$(t,s)\cdot(t_1,s_1) := (t_1 exp(-2s) + t, s + s_1) \quad , \tag{2.7}$$

because of the invariance (2.3). A one parameter subgroup is

$$\left.\begin{aligned} t &= a(exp(2cr) - 1) \\ s &= -cr \end{aligned}\right\} \tag{2.8}$$

where a, c are positive constants and $r > 0$ is the group parameter. To this subgroup (2.6) associates the one parameter semigroup of operators

$$\begin{aligned} H(r): &= exp^{a(exp(2cr)-1)D^2 - crxD} \\ &= e^{crxD} e^{a(1-exp(-2cr))D^2} \quad . \end{aligned} \tag{2.9}$$

It is straightforward to compute the infinitesimal generator of H

$$M\phi := lim_{r\to 0}\frac{H(r)\phi - \phi}{r} = 2acD^2\phi - cxD\phi \quad . \tag{2.10}$$

But in view of (2.9) and (2.10) we have the operator identity

$$e^{Mt} = e^{-crtD} e^{a(1-exp(-2ct))D^2} \quad . \tag{2.11}$$

To understand the meaning of (2.11) recall that for appropriate functions ϕ, $exp(Mt)\phi$ is the solution to the initial value problem

$$\begin{aligned} \frac{\partial}{\partial t}w(t,x) &= [Mw](x) = 2ac\frac{\partial^2}{\partial x^2}w(t,x) - cx\frac{\partial}{\partial x}x(t,x) \\ w(0,x) &= \phi(x) \end{aligned} \tag{2.12}$$

Then (2.11) suggests the following indirect procedure for solving (2.12):

Step 1: Solve the simpler initial value problem

$$\begin{aligned} \frac{\partial}{\partial t}u(t,x) &= \frac{\partial^2}{\partial x^2}u(t,x) \\ u(0,x) &= \phi(x) \end{aligned} \tag{2.13}$$

Step 2: Change independent variables in u to obtain w via

$$w(t,x) = u(a(1 - exp(-2ct)), exp(-ct)x) \quad . \tag{2.14}$$

Here we have interpreted the exponential in (2.11) as a transformation of variables.

This simple example illustrates the main point of this particular application of symmetry groups: knowing that a certain partial differential equation (such as (2.1)) is invariant under a group of local transformations (such as (2.8)) can be used to solve a more difficult equation (such as (2.12)) by first solving the simpler equation (such as (2.1)) and then changing variables.

This idea has been developed by S.I. Rosencrans in [8] [14]. It is appropriate to emphasize at this point that this use of a group of invariance of a certain PDE is not traditional. The more traditional use of group invariance is discussed at length in [7] [12], and is to reduce the number of independent variables involved in the PDE. Thus the traditional use of group invariance, is just a manifestation and mathematical development of the classical similarity methods in ODE.

The point of the simple example above is to illustrate a different use of group invariance which goes roughly as follows: given a parabolic PDE

$$u_t = Lu \tag{2.15}$$

and a group of local transformations that leave the solution set of (2.15) invariant, use this group to solve a "perturbed" parabolic PDE

$$w_t = (L + P)w \tag{2.16}$$

by a process of variable changes and the possible solution of an ordinary (not partial) differential equation. The operator P will be referred to as the "perturbation".

One of the contributions in this paper can be viewed as an extension of the results of Rosencrans to the stochastic partial differential equations that play a fundamental role in nonlinear filtering theory.

3 The Invariance Group of a Linear Parabolic PDE.

Consider the general, linear, nondegenerate elliptic partial differential operator

$$L := \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i} + c(x)id \quad . \tag{3.1}$$

and assume that the coefficients a_{ij}, b_i, c are smooth enough, so that $\mathcal{L}$ generates an analytic group [13], denoted by $exp(tL)$, for at least small $t \geq 0$, on some locally convex space X of initial functions ϕ and appropriate domain $Dom(L)$.

Let V be the set of solutions to

$$\begin{aligned}\frac{\partial u}{\partial t} &= Lu \\ u(0,x) &= \phi(x)\end{aligned} \tag{3.2}$$

in X, as we vary ϕ. The aim is to find a local Lie transformation group G which transforms every element of V into another element of V, that is an **invariance group** of (3.2) or of L. Note that G induces a group $\tilde{G}$ acting on the space of functions on M with values in $\mathbb{R}^p$, denoted by $\mathcal{F}(M;\mathbb{R}^p)$. The element $\tilde{g}$ corresponding to a g in G will map the function A into A', *i.e.*

$$A' = \tilde{g}(A) \quad . \tag{3.3}$$

It is easy to show [14] that G and $\tilde{G}$ are isomorphic as groups. We are interested in groups $\tilde{G}$ acting linearly. For that we need:

Definition: $\tilde{G}$ is linear if there exists a Lie group of transformations $\Sigma : M \to M$ such that for each $\tilde{g}\epsilon\tilde{G}$, there exists a $\sigma\epsilon\Sigma$, a $p \times p$ matrix "multiplier" $\nu = \nu(x,\tilde{g})$ and a solution ψ of (3.2) such that

$$\tilde{g}(A)(x) = \nu(x,\tilde{g})A(\sigma(x)) + \psi(x) \quad . \tag{3.4}$$

The meaning of (3.4) is rather obvious. The way $\tilde{G}$ acts on functions is basically via the "coordinate change" group Σ of M. The main result of Rosencrans [14], concerns the case of a single parabolic equation (3.2), *i.e.*.

Theorem 3.1 [14]: Every transformation $\tilde{g}$ in the invariance group $\tilde{G}$ of a linear parabolic equation is of the form

$$u(t,x) \longmapsto \nu(p(t,x))u(p(t,x)) + \psi(x) \tag{3.5}$$

where p is a transformation acting on the variables (t,x), ψ a fixed solution of the parabolic equation.

Clearly for linear parabolic equations $\tilde{G}$ is always infinite dimensional since it always includes the infinite dimensional subgroup $\tilde{\mathcal{H}}$ consisting of transformations of the form

$$A \longmapsto cA + \phi \tag{3.6}$$

where $A\epsilon\mathcal{F}(M;\mathbb{R})$, c a scalar $\neq 0$, ϕ a fixed solution of (3.2). Because of (3.8) one says that $\tilde{G}$ acts as a **multiplier representation** of Σ upon the space of solutions of (3.2).

We consider now one-parameter subgroups of the invariance group G of a given partial differential equation. That is we consider subgroups of G of the form $\{X_s\}$ where s "parametrizes" the elements. According to standard Lie theory the infinitesimal generators of these one-parameter subgroups form the Lie algebra $\Lambda(G)$ of the local Lie group G [7]. We shall, using standard Lie theory notation, denote X_s by $exp(sX)$ where X is the infinitesimal generator of the one parameter group $\{X_s\}$. Thus $X\epsilon\Lambda(G)$. Clearly the elements of $\Lambda[G]$ can be considered as first order partial differential operators in $\mathbb{R}^{n+1}$

$$X = \gamma(x,u)\frac{\partial}{\partial u} - \sum_{i=1}^{n}\beta_i(x,u)\frac{\partial}{\partial x_i}. \tag{3.7}$$

Indeed this follows from an expansion of $exp(sX)(x,u)$ for small s. Now $\{X_s\}$ induces a one-parameter subgroup $\{\tilde{X}_s\}$ in $\tilde{G}$, acting on functions. Let $\tilde{X}$ be the infinitesimal generator of $\{\tilde{X}_s\}$. Given a function $A\epsilon\mathcal{F}(\mathbb{R}^n;\mathbb{R})$ let

$$A(s,x) := \tilde{X}_s(A)(x). \tag{3.8}$$

If x_i, u are transformed to x_i', u' by a specific one-parameter subgroup $exp(sX)$ of G we can expand

$$\begin{aligned} u^{'} &= A(x) + s\gamma(x, A(x) + 0(s^2) \\ x_i^{'} &= x_i - s\beta_i(x, A(x) + 0(s^2). \end{aligned} \tag{3.9}$$

Thus

$$A(x') = A(x) - s\sum_{i=1}^{n} \beta i(x, A(x))\frac{\partial A(x)}{\partial x_i} + 0(s^2)$$

or

$$\begin{aligned} \tilde{X}(A)(x) &= \lim_{s\longrightarrow 0} \frac{A(s,x) - A(0,x)}{s} \\ &= \lim_{s\longrightarrow 0} \frac{A(s,x) - A(x)}{s} \\ &= \lim_{s\longrightarrow 0} \frac{A(s,x^{'}) - A(x^{'})}{s} \\ &= \lim_{s\longrightarrow 0} \frac{u^{'} - A(x^{'})}{s} \\ &= \gamma(x, A(x)) + \sum_{i=1}^{n} \beta_i(x, A(x))\frac{\partial A(x)}{\partial x_i}. \end{aligned} \tag{3.10}$$

In view of (3.5) the condition for $\tilde{G}$ to be linear is that [14]

$$\beta_{i,u} = \gamma_{uu} = 0 \quad . \tag{3.11}$$

The best way to characterize G (or $\tilde{G}$) is by computing its Lie algebra $\Lambda(G)$ (or $\Lambda(\tilde{G})$). A direct way of doing this is the following. By definition $\tilde{X}\epsilon\Lambda(\tilde{G})$ if

$$\mathcal{D}(A) = 0 \Longrightarrow \mathcal{D}(e^{s\tilde{X}}A) = 0 \quad \text{for small} \quad s. \tag{3.12}$$

When $\mathcal{D}$ is linear this reduces to

$$\mathcal{D}(A) = 0 \Longrightarrow \mathcal{D}(\tilde{X}(A)) = 0, \tag{3.13}$$

since

$$\frac{d}{ds}\mathcal{D}(e^{s\tilde{X}}A) = \mathcal{D}(e^{s\tilde{X}}(A))$$

implies (3.12) if we set $s = 0$. It is not difficult to show that (3.12) leads to a system of partial differential equations for γ and β_i.

We shall consider further the determination of $\Lambda(\tilde{G})$ in the case when $\tilde{G}$ is linear, since it is the only case of importance to our interests in the present paper. Then in view of (3.11)

$$\begin{aligned} \beta_i(x,u) &= \beta_i(x) \\ \gamma(x,u) &= u\delta(x) + \phi(x) \end{aligned} \tag{3.14}$$

for some β_i, δ, ϕ. Let us denote by β the vector $[\beta_1, \beta_2, \cdots, \beta_n]^T$. Then if A is a solution of (3.2), another solution is

$$A(s,x) = exp(s\tilde{X})A,$$

which satisfies

$$\begin{aligned} \frac{\partial}{\partial s}A(s,x) &= \delta(x)A(x) + \sum_{i=1}^{n} \beta_i(x)\frac{\partial A(x)}{\partial x_i} + \phi(x) \\ A(0,x) &= A(x) \end{aligned} \tag{3.15}$$

in view of (3.10) and due to the linearity assumption (3.14). The crucial point is that (3.15) is a first order hyperbolic PDE and thus it can be solved by the method of characteristics. The latter, very briefly, entails the following. Let $\epsilon(t)$ be the flow of the vector field $\sum_{i=1}^{n} \beta_i \frac{\partial}{\partial x_i}$, *i.e.* the solution of the ODE

$$\begin{aligned} &\frac{d}{dt}\epsilon(t,x) = \beta(\epsilon(t,x)) \\ &\epsilon(0,x) = x. \end{aligned} \tag{3.16}$$

Then from (3.15)

$$\begin{aligned} \frac{d}{dt}A(s-t,\epsilon(t,x)) &= -\delta(\epsilon(t,x))A(s-t,\epsilon(t,x)) \\ &+ \phi(\epsilon(t,x)) \end{aligned}$$

and therefore

$$\begin{aligned} A(s,x) &= exp(\int_0^s \delta(\epsilon(r,x))dr)A(\epsilon(s,x)) \\ &+ \int_0^s \Phi(t,x)dt \end{aligned} \tag{3.17}$$

where

$$\Phi(t,x) = exp(\int_0^t \delta(\epsilon(r,x))dr)\phi(\epsilon(t,x)). \tag{3.18}$$

By comparison with (3.5) one can view $exp(\int_0^s \delta(\epsilon(r,x))dr)$ as the "multiplier" ν. (3.17) clearly displays the linearity of $\tilde{G}$ near the identity.

The most widely known example, for which $\Lambda(\tilde{G})$ has been computed explicitly is the heat equation (2.1). The infinitesimal generators in this case are six, as follows

$$\begin{gathered} \frac{\partial}{\partial t},\ 2t\frac{\partial}{\partial t} + x\frac{\partial}{\partial x},\ \frac{\partial}{\partial x} \\ 1,\ 2t\frac{\partial}{\partial x} + x,\ 4t^2\frac{\partial}{\partial t} + 4tx\frac{\partial}{\partial x} + x^2. \end{gathered} \tag{3.19}$$

Let us apply these general results to a linear parabolic equation, like (3.2). From Theorem 3.1, then $\tilde{G}$ is linear. The infinitesimal generators of $\tilde{G}$ are given in view of (3.10) (3.14) (note that $x_1 = t$ here) by

$$Z = \alpha(t,x)\frac{\partial}{\partial t} + \sum_{i=1}^{n} \beta_i(t,x)\frac{\partial}{\partial x_i} + \gamma(t,x)id \quad , \tag{3.20}$$

for some functions α, β_i, γ of t and x. If u solves (3.2) so does

$$v(s) = exp(sZ)u, \quad \text{for small} \ \ s. \tag{3.21}$$

However v is also the solution of

$$\begin{aligned} \frac{\partial}{\partial s}v = \alpha\frac{\partial}{\partial t}v + \sum_{i=1}^{n}\beta_i\frac{\partial}{\partial x_i}v + \gamma v \\ v(0) = u, \end{aligned} \tag{3.22}$$

a first order hyperbolic PDE (solvable by the method of characteristics). Clearly since $\frac{\partial}{\partial t} - L$ is linear (3.12) applies and therefore

$$Zu\epsilon V \ \ \text{if} \ \ u\epsilon V. \tag{3.23}$$

The converse is also true: if (3.23) holds for some first order partial differential operator, Z is a generator of $\tilde{G}$.

Now (3.23) indicates how to compute α, β, γ. Namely

$$(\frac{\partial}{\partial t} - L)u = 0 \tag{3.24}$$

implies

$$(\frac{\partial}{\partial t} - L)(\alpha u_t + \sum_{i=1}^{n}\beta_i\frac{\partial u}{\partial x_i} + \gamma u) = 0. \tag{3.25}$$

For $u\epsilon V$ the second reads

$$\alpha_t u_t + \sum_{i=1}^{n}\beta_{i,t}u_{x_i} + \gamma_t u + \alpha u_{tt} + \sum_{i=1}^{n}\beta_i u_{x_{i,t}} + \gamma u_t = LZu, \tag{3.26}$$

or

$$\frac{d}{dt}Zu = (LZ - ZL)u \quad , \tag{3.27}$$

or

$$\frac{d}{dt}Z = [L,Z] \ \ \text{on} \ \ V. \tag{3.28}$$

In (3.28) [,] denotes commutator and $\frac{d}{dt}Z$ is symbolic of

$$\alpha_t\frac{\partial}{\partial t} + \sum_{i=1}^{n}\beta_{i,t}\frac{\partial}{dx_i} + \gamma_t \ id$$

Thus the elements of $\Lambda(\tilde{G})$ in this case satisfy a Lax equation. It is immediate from (3.28) that Z form a Lie algebra. Furthermore it can be shown [14] that α is independent of x, *i.e.* $\alpha(t,x) = \alpha(t)$ and that every Z satisfies an ODE

$$d_l \frac{d^l Z}{dt^l} + d_{l-1} \frac{d^{l-1} Z}{dt^{l-1}} + \ldots + d_0 Z = 0 \tag{3.29}$$

where $\ell \leq dim\tilde{G}$.

4 Using the Invariance Group of a Parabolic PDE in Solving New PDEs.

In this section we use the results of the previous section, to generalize the ideas presented via the example of section 2. We follow Rosencrans [8][14].

Thus we consider a linear parabolic equation like (3.2) and we assume we know the infinitesimal generators Z of the nontrivial part of $\tilde{G}$. Thus if u solves (3.2), so does $v(s) = exp(sZ)u$ **but** with some new initial data, say $R(s)\phi$. That is

$$e^{sZ} e^{tL} = e^{tL} R(s) \quad \text{on} \quad X. \tag{4.1}$$

Now $R(\cdot)$ has the following properties. First

$$\lim_{s \longrightarrow 0} R(s)\phi = \phi. \tag{4.2}$$

Furthermore from (4.1)

$$\begin{aligned} e^{tL} R(r) R(s)\phi &= e^{rZ} e^{tL} R(s)\phi = e^{rZ} e^{sZ} e^{tL}\phi \\ &= e^{(r+s)Z} e^{tL}\phi = e^{tL} R(r+s)\phi. \end{aligned} \tag{4.3}$$

Or

$$R(r)R(s) = R(r+s) \quad \text{for} \quad r,s \geq 0 \tag{4.4}$$

From (4.3), (4.4), $R(\cdot)$ is a semigroup. Let M be its generator:

$$M\phi = \lim_{s \longrightarrow 0} \frac{R(s)\phi - \phi}{s}, \quad \phi \in \ \mathrm{Dom}(M). \tag{4.5}$$

It is straightforward to compute M, given Z as in (3.20). Thus

$$M\phi = \alpha(0) L\phi + \sum_{i=1}^{n} \beta_i(0,x) \frac{\partial \phi}{\partial x_i} + \gamma(0,x)\phi. \tag{4.6}$$

Note that M is uniquely determined by the Z used in (4.1). The most important observation of Rosencrans [8] was that the limit as $t \to 0$ of the transformed solution $v(s) = exp(sZ)u$, call it w, solves the new initial value problem

$$\begin{aligned} \frac{\partial w}{\partial s} &= Mw \\ w(0) &= \phi. \end{aligned} \tag{4.7}$$

That is

$$e^{sZ}e^{tL} = e^{tL}e^{sM} \quad \text{on} \quad X \tag{4.8}$$

or

$$Ze^{tL} = e^{tL}M \quad \text{on Dom}(L).$$

This leads immediately to the following generalization of discussions in section 2: To solve the initial value problem

$$\begin{aligned} \frac{\partial w}{\partial s} &= Mw \\ w(0) &= \phi \end{aligned} \tag{4.9}$$

where

$$M = \alpha(0)L + \sum_{i=1}^{n} \beta_i(0,x)\frac{\partial}{\partial x_i} + \gamma(0,x)id \tag{4.10}$$

follow the steps given below.

Step 1: Solve $u_t = Lu, u(0) = \phi$.

Step 2: Find generator Z of $\tilde{G}$ corresponding to M and solve

$$\left.\begin{array}{l} \frac{\partial}{\partial s}v = \alpha(t)\frac{\partial}{\partial t}v + \sum_{i=1}^{n}\beta_i(t,x)\frac{\partial}{\partial x_i}v + \gamma(t,x)v \\ v(0) = u \end{array}\right\}$$

via the method of characteristics. Note this step requires the solution of ordinary differential equations only.

Step 3: Set $t = 0$ to $v(s,t,x)$.

This procedure allows easy computation of the solution to the "perturbed" problem (4.9)-(4.10) if we know the solution to the "unperturbed" problem (3.2). The "perturbation" which is of degree $\leq 1^{st}$, is given by the part of M:

$$P = \sum_{i=1}^{n} \beta_i(0,x)\frac{\partial}{\partial x_i} + \gamma(0,x). \ \ id. \tag{4.11}$$

We shall denote by $\Lambda(P)$ the set of all perturbations like (4.11), that permit solutions of $u_t = (L+P)u$ to be computed from solutions of $u_t = Lu$, by integrating only an additional ordinary differential equation. We would like to show that $\Lambda(P)$ is a Lie algebra strongly related to the Lie algebra $\Lambda(G)$ of the invariance group of L.

Definition: The Lie algebra $\Lambda(P)$ will be called the **perturbation algebra** of the elliptic operator L.

To see the relation between $\Lambda(\tilde{G})$ and $\Lambda(P)$, observe first that each generator Z in $\Lambda(\tilde{G})$ uniquely specifies an M, via (3.20), (4.6). Conversely suppose M is given. From the Lax equation (3.28) we find that

$$\frac{dZ}{dt}\mid_{t=0}=[L,Z]\mid_{t=0}=[L,M]=\alpha_t(0)L+\sum_{i=1}^{n}\beta_{i,t}(0,x)\frac{\partial}{\partial x_i}+\gamma_t(0,x)id. \tag{4.12}$$

Note that the right hand side of (4.12) is another perturbed operator M'. Thus given an M, by repeated bracketting with L all initial derivatives of Z can be obtained. Since from (3.29) Z satisfies a linear ordinary differential equation, Z **can be determined from** M. So there exists a 1-1 correspondence between $\Lambda(\tilde{G})$ and the set of perturbed operators M, which we denote by $\Lambda(M)$. It is easy to see that $\Lambda(M)$ is a Lie algebra isomorphic to $\Lambda(\tilde{G})$. Indeed let Z_i correspond to M_i, $i=1,2$. Then from (4.8) we have

$$\begin{aligned} e^{tL}[M_1,M_2]\phi &= e^{tL}M_1M_2\phi-e^{tL}M_2M_1\phi \\ &= Z_1e^{tL}M_2\phi-Z_2e^{tL}M_1\phi \\ &= Z_1Z_2e^{tL}\phi-Z_2Z_1e^{tL}\phi=[Z_1,Z_2]e^{tL}\phi. \end{aligned} \tag{4.13}$$

This establishes the claim. Since each perturbation P is obtained from an M by omitting the component of M that involves the unperturbed operator L, it is clear that $\Lambda(P)$ is a Lie subalgebra of $\Lambda(M)$. Moreover the dimension of $\Lambda(P)$ is one less than that of $\Lambda(M)$. In view of the isomorphism of $\Lambda(M)$ and $\Lambda(\tilde{G})$ we have established [8]:

Theorem 4.1: The perturbation algebra $\Lambda(P)$ of an elliptic operator L, is isomorphic to a Lie subalgebra of $\Lambda(\tilde{G})$ (*i.e.* of the Lie algebra of the invariance group of L). Moreover $dim\Lambda(P))=dim(\Lambda(\tilde{G}))-1$.

One significant question is: can we find the perturbation algebra $\Lambda(P)$ without first computing $\Lambda(\tilde{G})$, the invariance Lie algebra? The answer is affirmative and is given by the following result [8].

Theorem 4.2: Assume L has analytic coefficients. An operator P_0 of order one or less (*i.e.* of the form (4.12)) is in the perturbation algebra $\Lambda(P)$ of L if there exist a sequence of scalars $\lambda_1,\lambda_2,\cdots$ and a sequence of operators $P_l,P_2,\cdots$ of order less than or equal to one such that

$$[L,P_n]=\lambda_nL+P_{n+1}\quad,\quad n\geq 0$$

and $\sum\lambda_kt^k/k!,\sum P_kt^k/k!$ converge at least for small t.

It is an easy application of this result to compute the perturbation algebra of the heat equation in one dimension or equivalently of $L=\frac{\partial^2}{\partial x^2}$. It turns out that $\Lambda(P)$ is 5-dimensional and spanned by

$$\Lambda(P)=\mathrm{Span}(1,\ x,\ x^2,\ \frac{\partial}{\partial x},\ x\frac{\partial}{\partial x}). \tag{4.14}$$

So the general perturbation for the heat equation looks like

$$P=(ax+b)\frac{\partial}{\partial x}+(cx^2+dx+e)id \tag{4.15}$$

where a, b, c, d, e are arbitrary constants. Note that the invariance group of the heat equation is 6-dimensional (3.19). It is straightforward to rework the example of section 2, along the lines suggested here.

The implications of these results are rather significant. Indeed consider the class of linear parabolic equations $u_t = Lu$, where L is of the form (3.1). We can define an equivalence relationship on this class by: "L_1 is equivalent to L_2 if $L_2 = L_1 + P$ where P is an element of the perturbation algebra $\Lambda^1(P)$ of L_1". Thus elliptic operators of the form (3.1), or equivalently linear parabolic equations are divided into equivalent classes (orbits); within each class (orbit) $\{L(k)\}$ (k indexes elements in the class) solutions to the initial value problem $u(k)_t = L(k)u(k)$ with fixed data ϕ (independent of k) can be obtained by quadrature (*i.e.* an ODE integration) from any one solution $u(k_0)$.

We close this section by a list of perturbation algebras for certain L, from [8].

Elliptic operator L	Generators of perturbation algebra $\Lambda(P)$
D^2	$1, x, x^2, D, xD$
xD^2	$1, x, xD$
x^2D^2	$x \log xD, xD, \log x, (\log x)^2, 1$
x^3D^2	$1, x^{-1}, xD$
$e^x D^2$	$1, e^{-x}, D$

Table 4.1: Examples of perturbation algebras.

5 Strong Equivalence of Nonlinear Filtering Problems

In this section we will apply the methods described in sections 2-4 for parabolic equations to the fundamental PDEs governing nonlinear filtering problems. As with all symmetry group methods these techniques have a strong geometric flavor.

We will only briefly discuss the focal points of our current understanding of the nonlinear filtering problem and we will refer the reader to [9] or the references [2]-[6] for details. Thus the "nonlinear filtering problem for diffusion processes" consists of a model for a "signal process" $x(t)$ via a stochastic differential equation

$$dx(t) = f(x(t))dt + g(x(t))dw(t) \tag{5.1}$$

which is assumed to have unique solutions in an appropriate sense (strong or weak, see [9]). In addition we are given "noisy" observations of the process $x(t)$ described by

$$dy(t) = h(x(t))dt + dv(t). \tag{5.2}$$

Here $w(t), v(t)$ are independent standard Wiener processes and h is such that y is a semimartingale. The problem is to compute conditional statistics of functions of the signal process $\phi(x(t))$ at time t given the data observed up to time t, *i.e.* the σ-algebra

$$\mathcal{F}_t^y = \sigma\{y(s), 0 \le s \le t\} \tag{5.3}$$

Clearly the maximum information about conditional statistics is obtained once we find ways to compute the conditional probability density of $x(t)$ given $\mathcal{F}_t^y$. Let us denote this conditional density by $p(t,x)$. It is more convenient to use a different function, so called unnormalized conditional density, $u(t,x)$ which produces p after normalization

$$p(t,x) = \frac{u(t,x)}{\int u(t,z)dz} \tag{5.4}$$

The reason for the emphasis put on u is that it satisfies a **linear** stochastic PDE driven directly by the observations. This is the so called Mortensen-Zakai stochastic PDE, which in Itô's form is

$$du(t,x) = \mathcal{L}u(t,x)dt + h^T(x)u(t,x)dy(t) \tag{5.5}$$

Here $\mathcal{L}$ is the adjoint of the infinitesimal generator of the diffusion process $x(\cdot)$

$$[\mathcal{L}\phi](x) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}[\sigma_{ij}(x)\phi] - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}[f_i(x)\phi(x)] \tag{5.6}$$

which is also called the Fokker-Planck operator associated with $x(\cdot)$. In (5.6) the matrix σ is given by

$$\sigma(x) = g(x)g(x)^T \quad , \tag{5.7}$$

and we shall assume that σ is positive definite, *i.e.* the elliptic operator $\mathcal{L}$ is nondegenerate. When applying geometric ideas to (5.5) it is more convenient to consider the Stratonovich version

$$\frac{\partial u(t,x)}{\partial t} = (\mathcal{L} - \frac{1}{2}h(x)^T h(x))u(t,x) + h^T(x)u(t,x)\frac{dy(t)}{dt}. \tag{5.8}$$

We shall primarily work with (5.8) in the present paper. Letting

$$A : = \mathcal{L} - \frac{1}{2}h^T h \tag{5.9}$$

$$B_j : = \text{ Mult. by } h_j (j^{th} \text{ comp. of } h) \tag{5.10}$$

we can rewrite (5.8) as an infinite dimensional bilinear equation

$$\frac{du(t)}{dt} = (A + \sum_{j=1}^{p} B_j \dot{y}_j(t))u(t). \tag{5.11}$$

We shall assume that every equation of the form (5.8) considered has a complete existence and uniqueness theory established on a space X. Furthermore we shall assume that continuous dependence of solutions on $y(\cdot)$ has been established.

The estimation Lie algebra introduced by Brockett [2] and analyzed in [2]-[6] is the Lie algebra

$$\Lambda(E) = \text{ Lie algebra generated by } A \text{ and } B_j,\ j = 1, \cdots p. \tag{5.12}$$

Again we shall assume that for problems considered the operators A, B_j have a common, dense invariant set of analytic vectors in X [10] and that the mathematical relationship between $\Lambda(E)$ and the existence-uniqueness theory of (5.8) is well understood.

We develop a methodology for recognizing mathematically "equivalent" problems. Equivalence here carries the following meaning: two nonlinear filtering problems should be equivalent when knowing the solution of one, the solution of the other can be obtained by relatively simple additional computations. Examples discovered by Beneš [11], created certain excitement for the possibility of a complete classification theory. We shall see how transparent Beneš' examples become from the point of view developed in this paper.

To make things precise consider two nonlinear filtering problems (vector)

$$\begin{aligned} dx^i(t) &= f^i(x^i(t)dt + g^i(x^i(t))dw^i(t) \\ dy^i(t) &= h^i(x^i(t))dt + dv^i(t) \quad ; \quad i = 1, 2 \end{aligned} \tag{5.13}$$

and the corresponding Mortensen-Zakai equations in Stratonovich form

$$\frac{\partial u_i(t,x)}{\partial t} = (\mathcal{L}^i - \frac{1}{2} \parallel h^i(x) \parallel^2) u_i(t,x) + h^{iT}(x) u_i(t,x) y^i(t); \quad i = 1, 2 \tag{5.14}$$

Definition: The two nonlinear filtering problems above are **strongly equivalent** if u_2 can be computed from u_l, and vice versa, via the following types of operations:

Type 1: $(t, x^2) = \alpha(t, x^1)$, where α is a diffeomorphism.

Type 2: $u_2(t,x) = \psi(t,x)u_1(t,x)$, where $\psi(t,x) \geq 0$ and $\psi^{-1}(t,x) \geq 0$.

Type 3: Solving a set of ordinary (finite dimensional) differential equations (i.e. quadrature).

Brockett [2], has analyzed the effects of diffeomorphisms in x-space and he and Mitter [4] the effects of so called "*gauge*" transformations (a special case of our type 2 operations) on (5.8). Type 3 operations are introduced here for the first time, and will be seen to be the key in linking this problem with mathematical work on group invariance methods in ODE and PDE's.

Our approach starts from the abstract version of (5.14)(*i.e.* (5.11)):

$$\frac{\partial u_i}{\partial t} = (A^i + \sum_{j=1}^{p} B^i_j \cdot y_j(t)) u_i \quad ; \quad i = 1, 2 \tag{5.15}$$

where A^i, B^i_j are given by (5.9)-(5.10). We are thus dealing with two parabolic equations. We will first examine whether the evolutions of the time invariant parts can be computed from one another. This is a classical problem and the methods of section 3, 4 apply. In this section we give an extension to the full equation (5.15) under certain conditions on B^i_j. We shall then apply this result to the examples studied by Beneš and recover the Riccati equations as a consequence of strong equivalence.

Our main result concerning equivalence (in a computational sense) of two nonlinear filtering problems is the following.

Theorem 5.1: Given two nonlinear filtering problems (see (5.13)), such that the corresponding Mortensen-Zakai equations (see (5.14)) have unique solutions, continuously dependent on $y(\cdot)$. Assume that using operations of type 1 and 2 (see definition just above) these stochastic PDEs can be transformed in bilinear form

$$\frac{\partial u_i}{\partial t} = (A^i + \sum_{j=1}^{p} B_j^i \xi_j^i(t))u_i; \quad i = 1, 2$$

such that:

(i) A^i, $i = 1, 2$, are nondegenerate elliptic, belonging to the same equivalence class (see end of section 4)

(ii) B_j^i, $j = l, \cdots p$, $i = 1, 2$ belong to the perturbation algebra $\Lambda(P)$ of (i).

Then the two filtering problems are strongly equivalent.

Proof: Only a sketch will be given here. One first establishes that is enough to show computability of solutions for piecewise constant ξ, from one another, by the additional computation of solutions of an ODE. For piecewise constant ξ the solution to any one of the PDEs in bilinear form is given by

$$u_i = e^{(A^i + B^i_{j_m}\xi^i_{j_m})(t_m - t_{m-1})} \cdot e^{(A^i + B^i_{j_{m-1}}\xi^i_{j_{m-1}})(t_{m-1} - t_{m-2})} \cdots e^{(A^i + B^i_{j_1}\xi^i_{j_1})t_1}\phi \quad ; \quad i = 1, 2 \tag{5.16}$$

Since A^1, A^2, belong to the same equivalence class there exist $Z^{12}\epsilon\Lambda(G)$, (where $\Lambda(G)$ is the Lie algebra of the invariance group for the class) and $P^{12}\epsilon\Lambda(P)$ (where $\Lambda(P)$ is the perturbation algebra of the class) such that (see (4.8)):

$$A^2 = A^1 + P^{12} \tag{5.17}$$

$$e^{sZ^{12}}e^{tA^1} = e^{tA^1}e^{sA^2}; \quad t, s \geq 0. \tag{5.18}$$

That is consider A^2 as a "perturbation" of A^1. We know by now what (5.17) means: to compute the semigroup generated by A^2, we first compute the semigroup generated by A^1, we then solve the ODE associated with the characteristics of the hyperbolic PDE

$$\frac{\partial v}{\partial s} = Z^{12}v \tag{5.19}$$

and we have

$$e^{sA_2} = [e^{sZ^{12}}e^{tA^1}]|_{t=0}. \tag{5.20}$$

More generally since $A^1 + B_j^1$, $A^2 + B_k^2$ belong to the same class there exist $Z_{jk}^{12}\epsilon\Lambda(G), P_{jk}^{12}\epsilon\Lambda(P)$ such that

$$\begin{aligned} A^2 + B_k^2 &= A^1 + B_j^1 + P_{jk}^{12} \\ e^{sZ_{jk}^{12}}e^{t(A^1+B_j^1)} &= e^{t(A^1+B_j^1)}e^{s(A^2+B_k^2)}. \end{aligned} \tag{5.21}$$

It is now apparent that if we know (5.16) explicitly for $i = 1$, we obtain u_2 from (5.21) with the only additional computations being the integration of the ODEs associated with the characteristics of the hyperbolic PDEs

$$\frac{\partial v}{\partial s} = Z_{jk}^{12} v,\ k, j = 1, ..., p. \tag{5.22}$$

This completes the proof.

Let us apply this result to the Beneš case. We consider the linear filtering problem (scalar x, y)

$$\left.\begin{aligned} dx(t) &= dw(t) \\ dy(t) &= x(t)dt + dv(t). \end{aligned}\right\}$$

and the nonlinear filtering problem (scalar x, y)

$$\left.\begin{aligned} dx(t) &= f(x(t))dt + dw(t) \\ dy(t) &= x(t)dt + dv(t). \end{aligned}\right\}$$

The corresponding Mortensen-Zakai equations in Stratonovich form are: for the linear

$$\frac{\partial u_1(t,x)}{\partial t} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} - x^2)u_1(t,x) + x\dot{y}(t)u_1(t,x); \tag{5.23}$$

for the nonlinear

$$\frac{\partial u_2}{\partial t} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} - x^2)u_2(t,x) - \frac{\partial}{\partial x}(fu_2) + x\dot{y}(t)u_2(t,x). \tag{5.24}$$

We wish to show that (5.23)(5.24) are strongly equivalent only if f (the drift) is a global solution of the Riccati equation

$$f_x + f^2 = ax^2 + bx + c. \tag{5.25}$$

First let us apply to (5.23)(5.24) an operation of type 2. That is let (defines v_2)

$$u_2(t,x) = v_2(t,x)exp(\int_0^x f(u)du). \tag{5.26}$$

The transformation (5.26) is global, and is an example of such more general transformations needed for systems and control problems and discussed at the end of section 1. This is like a gauge transformation from mathematical physics, or a potential transformation in symmetry group theory [1].

Then the new function v_2 satisfies

$$\frac{\partial v_2(t,x)}{\partial t} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} - x^2 - V(x))v_2(t,x) + x\dot{y}(t)v_2(t,x), \tag{5.27}$$

where

$$V(x) = f_x + f^2. \tag{5.28}$$

Existence, uniqueness and continuous dependence on $y(\cdot)$ for (5.23)(5.24) have been established using classical p.d.e. results. We apply Theorem 5.1 to (5.23) (5.27). So

$$\begin{aligned} A^1 &= \frac{1}{2}(\frac{\partial^2}{\partial x^2} - x^2) \\ A^2 &= \frac{1}{2}(\frac{\partial^2}{\partial x^2} - x^2 - V) \end{aligned} \tag{5.29}$$

while

$$B^1 = B^2 = \text{Mult. by } x. \tag{5.30}$$

From the results of section 4, the only possible equivalence class is that of the heat equation. Clearly from (4.15) or Table 4.1, $A^1, B^1, B^2 \epsilon \Lambda(P)$ for this class. For A^2 to belong to $\Lambda(P)$ it is necessary that V be quadratic, which is the same as f satisfying the Riccati equation (5.25), in view of (5.28).

Recall that the solution of (5.23) is

$$u_1(t,x) = exp(-\frac{(x-\mu(t))^2}{2\sigma(t)}) \tag{5.31}$$

where

$$\begin{aligned} d\mu(t) &= \sigma(t)(dy(t) - \mu(t)dt); \quad \mu(0) = \xi \\ d\sigma(t) &= 1 - \sigma^2(t); \quad \sigma(0) = 0 \end{aligned} \tag{5.32}$$

Beneš [11], using a path integral computation showed that the solution of (5.27), when (5.25) is satisfied is given by

$$v_2(t,x) = exp(-\frac{(x-\mu(t))^2}{2\sigma(t)}) \tag{5.33}$$

where

$$\begin{aligned} d\mu(t) &= -(a+1)\sigma(t)\mu(t)dt - \frac{1}{2}\sigma(t)bdt + \sigma(t)dy(t) \\ d\sigma(t) &= 1 - (a+1)\sigma^2(t). \end{aligned} \tag{5.34}$$

What we have shown here is a converse, from the point of view that strong equivalence of the linear and nonlinear filtering examples implies the Riccati equation. We also maintain that knowledge of group invariance theory makes the result immediate at the level of comparing (5.23) with (5.27).

6 Reduction of Nonlinear Output Robust Control Problems

In this section we will apply methods from symmetry groups to a problem of recent interest in nonlinear control: output robust control. As has been developed fully in [24], the nonlinear output robust control problem (in an H^∞ sense) is equivalent to a risk-sensitive partially observed stochastic control problem and to a dynamic partially observed game [20][26]. A key result in establishing these equivalences was the introduction of the information state and the nonlinear PDE that it satisfies. In this section we apply systematic methods from symmetry groups to this fundamental PDE of nonlinear robust control.

The dynamic game representation of the equivalent nonlinear robust control problem is as follows. Given the dynamical system

$$\left.\begin{array}{rcl} \dot{x}(t) & = & b(x(t),u(t)) + w(t), \quad x(0) = X_0 \\ y(t) & = & Cx(t) + v(t) \end{array}\right\} \tag{6.1}$$

where w, v are L^2-type disturbances. We want to find a control $u(\cdot)$, which is a non-anticipating functional of $y(\cdot)$, to minimize ($\mu > 0$)

$$\begin{aligned} J(u) = \sup_{w \epsilon L^2} \sup_{v \epsilon L^2} \sup_{X_0 \epsilon L^2} \{\bar{p}(x_0) \; &+ \; \int_0^T [L(x(s),u(s)) - \frac{1}{2}\mu(\mid w(s) \mid^2 + \mid v(s) \mid^2)]ds \\ &+ \; \Phi(x(T))\} \end{aligned} \tag{6.2}$$

One of the questions we want to answer, is when can we reduce this nonlinear problem to a linear one? Group invariance methods can be applied to this problem. Let us make the following structural assumptions:

$$\left.\begin{array}{lcl} b(x,u) & = & f(x) + A(u)x + B(u) \\ f(x) & = & DF(x) \text{ for some } F \\ \frac{1}{2} \mid f(x) \mid^2 & + & f(x) \cdot (A(u)x + B(u)) = \frac{1}{2}x^T \sum(u)x + \Lambda(u)x + \frac{1}{2}\Gamma(u) \\ \bar{p}(x) & = & -\frac{1}{2}(x - \bar{x}^T \bar{Y}^{-1}(x - \bar{x}) + \bar{\phi} + \frac{1}{\mu}F(x) \\ L(x,u) & = & \frac{1}{2}R(u) + \frac{1}{2}x^T Q(u)x \end{array}\right\} \tag{6.3}$$

Let us next consider the information state PDE for (6.1)-(6.2):

$$\frac{\partial p}{\partial t} = -Dp.b(x,u) + \frac{\mu}{2} \mid Dp \mid^2 + L(x,u) - \frac{1}{2\mu}(y(t) - Cx)^2 \tag{6.4}$$

The optimal control is a memoryless function of the information state since the cost (6.2), can be expressed using the information state as follows [23][24]:

$$J(u) = \sup_{y \epsilon L^2}\{(P_T, \Phi) : p_0 = \bar{p}\}; \;\; (p,q) = \sup_{x \epsilon \Re^n}\{p(x) + q(x)\} \;\; . \tag{6.5}$$

That is why the information state PDE (6.4) is so fundamental. Under the structural assumptions (6.3), it is not hard to show [21][27][28] that the PDE (6.4) has a finitely parameterizable solution

$$p_t(x) = -\frac{1}{2\mu}(x - \hat{x}(t))^T Y(t)^{-1}(x - \hat{x}(t)) + \phi(t) + \frac{1}{\mu}F(x) \tag{6.6}$$

where

$$\left.\begin{array}{lcl} \dot{\hat{x}}(t) & = & (A(u(t)) + \mu Y(t)Q(u(t)) - Y(t)\sum(u(t)))\hat{x}(t) + B(u(t)) \\ & & -Y(t)\Lambda(u(t)) + Y(t)C^T(y(t) - C\hat{x}(t)) \\ \hat{x}(0) & = & \bar{x} \\ \dot{Y}(t) & = & Y(t)A(u(t))^T + A(u(t))Y(t) - Y(t)(C^TC - \mu Q(u(t)) + \sum(u(t)))Y(t) + I \\ Y(0) & = & \bar{Y} \\ \dot{\phi}(t) & = & \frac{1}{2}R(u(t)) + \frac{1}{2}\hat{x}^T(t)Q(u(t))\hat{x}(t) - \frac{1}{\mu}(\frac{1}{2}\hat{x}(t)^T\sum(u(t))\hat{x}(t) + \Lambda(u(t))\hat{x}(t) \\ & & +\frac{1}{2}\Gamma(u(t))) - \frac{1}{2\mu} \mid y(t) - (C\hat{x}(t)) \mid^2 \\ \phi(0) & = & \bar{\phi} \end{array}\right\} \quad (6.7)$$

A consequence of this is that the robust output control for the nonlinear system (6.1) is finite dimensional and easily implementable. The explanation for these specific results becomes clear from an invariance group perspective. Under an appropriate transformation the dynamic game (6.1)-(6.2) becomes linear, quadratic (and therefore has a well known finite dimensional controller) with state

$$\rho(t) = (\hat{x}(t), Y(t), \phi(t)) \quad (6.8)$$

and cost

$$J(u) = \sup_{y\epsilon L^2}\{\hat{\Phi}(\rho(T)); \rho(0) = \bar{\rho}\} \quad (6.9)$$

where

$$\left.\begin{array}{lcl} \hat{\Phi}(\rho) & = & (p_\rho, \Phi) \\ p_\rho & = & -\frac{1}{2\mu}(x-\hat{x})^TY^{-1}(x-\hat{x}) + \phi + \frac{1}{\mu}F(x) \end{array}\right\} \quad (6.10)$$

Similar reductions can be obtained, under similar assumptions, for the associated partially observed, risk sensitive stochastic control problem. These results can be obtained, understood and generalized by studying the group invariance of the information state PDE, and in particular using the methods of sections 3 and 4. Specifically the structural assumptions (6.3) are completely analogous to the Beneš structural assumptions. Requiring problem (6.1)-(6.2) to be the nonlinear equivalent to a linear quadratic problem (from the perspective of equivalence of the corresponding information state PDEs, as in sections 3, 4,) implies the structural assumptions (6.3). Thus it is important to study the symmetry groups (invariance properties) of the information state PDE for the linear control problem.

We shall omit the control parameter u from the notation, since it plays no part in the following calculations; in other words, instead of writing $A(u(t))$, we abbreviate to $A(t)$. The information state $p(t,x) \equiv p(t,x_1,\cdots,x_n)$ for the linear control problem satisfies the scalar PDE

$$\begin{gathered} F(t,x,p_t,\nabla p) \equiv p_t + \nabla p\cdot(Ax+b) \\ -|\nabla p|^2/2 + x^TGx/2 + h\cdot x + l = 0\ , \end{gathered} \quad (6.11)$$

where $G \equiv G(t)$ is a symmetric matrix, $A \equiv A(t)$ is a square matrix, $b \equiv b(t)$ and $h \equiv h(t)$ are n-vectors, and $l \equiv l(t)$ is a scalar function. James and Yuliar [21] point out that there is a solution of the form

$$p(t,x) = -(x-r(t))^TW(t)(x-r(t))/2 + \phi(t), \quad (6.12)$$

with W symmetric. Taking the gradient of (6.12), substituting in (6.11) and equating coefficients of terms quadratic, linear, and constant in x we obtain the ODEs

$$\begin{aligned}
\dot{W} &= -WA - A^T W - W^2 + G \ , \\
\dot{r} &= W^{-1}(-\dot{W}r + Wb - A^T Wr - W^2 r - h) \ , \\
\dot{\phi} &= r^T(W\dot{r} + \dot{W}r/2 - Wb + W^2 r/2) - l \ .
\end{aligned} \tag{6.13}$$

The last two equations can be rewritten as

$$\dot{r} = Ar + b - W^{-1}(Gr + h) \ , \tag{6.14}$$

$$\dot{\phi} = r^T(WA - A^T W - G)r/2 - r^T h - l \ . \tag{6.15}$$

We now turn to the Lie transformation theory for the Information state PDE. We consider the invariance of (6.11) under an infinitesimal transformation given by a vector field of the form (note that we are not including a $\partial/\partial t$ term)

$$X \equiv \sum_{i=l}^{n} \xi^i(t,x)\frac{\partial}{\partial x_i} + \eta(t,x,p)\frac{\partial}{\partial p} \ . \tag{6.16}$$

According to Bluman and Kumei [1], Theorem 4.1.1-1, the criterion for invariance is that

$$\begin{aligned}
&X^{(1)}F(t,x,p_t,\nabla p) = 0 \text{ whenever} \\
&F(t,x,p_t,\nabla p) = 0 \ ,
\end{aligned} \tag{6.17}$$

where $X^{(1)}$ is the first extended infinitesimal generator, namely

$$X^{(1)} \equiv \sum_{i=l}^{n} \xi^i \frac{\partial}{\partial x_i} + \eta \frac{\partial}{\partial p} + \eta_0^{(1)} \frac{\partial}{\partial p_t} + \sum_{i=1}^{n} \eta_i^{(1)} \frac{\partial}{\partial p_i}$$

where

$$p_i \equiv \frac{\partial}{\partial x_i} \ , \tag{6.18}$$

$$\eta_0^{(1)}(t,x,p,p_t,\nabla p) \equiv D_t\eta - \sum_{j=1}^{n} (D_t\xi^j)p_j \ , \tag{6.19}$$

$$\eta_i^{(1)}(t,x,p,p_t,\nabla p) \equiv D_i\eta - \sum_{j=1}^{n} (D_i\xi^j)p_j, \quad i = 1,\ldots, n \ , \tag{6.20}$$

$$D_t \equiv \frac{\partial}{\partial t} + p_t \frac{\partial}{\partial p}, \ D_i \equiv \frac{\partial}{\partial x_i} + p_i \frac{\partial}{\partial p}, \quad i = 1,\ldots n \ . \tag{6.21}$$

Evaluating the entries of (6.16) term by term,

$$\sum_{i=1}^{n} \xi^i \frac{\partial F}{\partial x_i} = (A^T(\nabla p) + Gx + h)\cdot \xi \ ,$$

$$\eta \frac{\partial F}{\partial p} = 0,$$

$$\eta_0^{(1)} \ \frac{\partial F}{\partial p_t} \ = \eta_t + p_t \eta_p - (\nabla p \cdot \xi_t) \ ,$$

$$\begin{aligned}\sum_{i=1}^n \ \eta_i^{(1)} \ \frac{\partial F}{\partial p_i} \ &= \ \sum_{i=1}^n \left(D_i \eta - \ \sum_{j=1}^n \ (D_i \xi^j) p_j \right) \\ &\qquad\qquad (Ax + b - \nabla p)^i \\ &= (Ax + b - \nabla p) \cdot (\nabla \eta + \eta_p \nabla p - (\nabla p \cdot \nabla) \xi) \ .\end{aligned}$$

Adding up all these terms shows that (6.16) gives

$$\begin{aligned}&(A^T(\nabla p) + Gx + h) \cdot \xi + \eta_t + p_t \eta_p - (\nabla p \cdot \xi_t) \\ &= -(Ax + b - \nabla p) \cdot (\nabla \eta + \eta_p \nabla p - (\nabla p \cdot \nabla) \xi) \ .\end{aligned}$$

Grouping terms, we obtain:

Theorem 6.1 (Fundamental Transformation Relation): The vector fields ξ and η of the infinitesimal generator of a symmetry group of (6.11) must satisfy:

$$\begin{aligned}&\eta_t + (Ax + b - \nabla p) \cdot \nabla \eta + (p_t + (Ax + b - \nabla p) \cdot \nabla p) \eta_p \\ = \ &\nabla p \cdot \xi_t - (A^T \nabla p + Gx + h) \cdot \xi + \\ &(Ax + b - \nabla p) \cdot (\nabla p \cdot \nabla) \xi \ ,\end{aligned} \tag{6.22}$$

where p is given by (6.12).

Note that only **linear** differential operators acting on ξ and η are involved, and ∇p and p_t are quadratic in x. Hence for any choice of ξ we may solve for η by the method of characteristics. That is given ξ, solving for η involves only the solution of an ODE.

We describe next how to use the fundamental transformation relation. Let

$$\varphi(\varepsilon; t, x, p)) \equiv (t, \bar{x}(\varepsilon, t, x), \bar{p}(\varepsilon, t, x, p))$$

denote the flow of the vector field

$$X \equiv \ \sum_{i=1}^n \ \xi^i(t, x) \ \frac{\partial}{\partial x_i} \ + \eta \ \frac{\partial}{\partial p} \tag{6.23}$$

where $\bar{x}(\varepsilon, t, x)$ are the transformed state space coordinates and $\bar{p}(\varepsilon, t, x, p) \equiv \bar{p}(t, \bar{x})$ is the information state for the transformed problem. By definition of φ,

$$\frac{d\varphi}{d\varepsilon} \ (\varepsilon) = X \varphi(t, \bar{x}, \bar{p}), \ \ \varphi(0; (t, x, p)) = (t, x, p) \cdot \tag{6.24}$$

This breaks down into the system of ODEs

$$\frac{\partial \bar{x}}{\partial \varepsilon} \ = \ \xi(t, \bar{x}), \ \ \bar{x}(0, t, x) = x \ ; \tag{6.25}$$

together with the scalar ODE

$$\frac{\partial \bar{p}}{\partial \varepsilon} = \eta(t, \bar{x}, \bar{p}), \ \bar{p}(0, t, x, p) = p \ . \tag{6.26}$$

Therefore we have established the following:

Theorem 6.2: *Assume p satisfies (6.11). Suppose $\varepsilon \to \bar{x}\ (\varepsilon, t, x)$ is a one-parameter family of transformations of the space variable x satisfying the system of ODEs (6.25), for some choice of $\xi \equiv \xi(t, x)$, and that $\eta \equiv \eta(t, x, p)$ is chosen to satisfy the Fundamental Transformation Relation (6.22) in terms of ξ. Then the solution $\varepsilon \to \bar{p}(\varepsilon, t, x, p)$ to the ODE (6.26), if unique, is a one-parameter family of transformations of the information state variable p, so that (6.11) holds with (x, p) replaced by $(\bar{x}, \bar{p})$.*

7 A Case Explicitly Computable

The drawback of Theorem 6.2 is that it is too abstract to be of immediate practical use. Therefore we consider a more specialized situation admitting explicit computations. We shall constrain the choice of η so as to satisfy

$$\eta = -x^T W \xi \ . \tag{7.1}$$

This implies

$$\begin{aligned} \eta_t &= -x^T \dot{W} \xi - x^T W \xi_t, \ \eta_p = 0, \\ \nabla \eta &= -(Wx \cdot \nabla)\xi - W\xi \ . \end{aligned} \tag{7.2}$$

Now (6.24)implies

$$\begin{aligned} &-x^T \dot{W} \xi - x^T W \xi_t - (Ax + b - \nabla p) \cdot ((Wx \cdot \nabla)\xi + W\xi) \\ &= \nabla p \cdot \xi_t - (A^T \nabla p + Gx + h) \cdot \xi \\ &+(Ax + b - \nabla p) \cdot (\nabla p \cdot \nabla)\xi \ . \end{aligned}$$

Rearranging terms gives

$$\begin{aligned} (A^T \nabla p + Gx + h - \dot{W}x - W(Ax + b - \nabla p)) \cdot \xi \\ -(\nabla p + Wx) \cdot \xi_t \end{aligned}$$

$$= (Ax + b - \nabla p) \cdot ((\nabla p + Wx) \cdot \nabla)\xi \ ,$$

$$\begin{aligned} (-A^T W(x - r) + Gx + h - \dot{W}x - \\ W(Ax + b + W(x - r))) \cdot \xi - (Wr) \cdot \xi_t \end{aligned}$$

$$= (Ax + b + W(x - r)) \cdot ((Wr) \cdot \nabla)\xi \ .$$

The coefficient of x in the first bracket is $-A^T W + G - \dot{W} - W)(A+W) = 0$, by (6.13). Define the following vector functions in terms of quantities determined above:

$$\beta(t) \equiv Wr,\ \Gamma(t) \equiv A + W,\ \gamma(t) \equiv b - \beta\ , \tag{7.3}$$

$$\alpha(t) \equiv (A^T + W)Wr + h - Wb = -\beta_t\ , \tag{7.4}$$

where the last identity follows from (6.13) and (6.14)-(6.15), since

$$\begin{aligned} W\dot{r} + \dot{W}r &= WAr + Wb - Gr - h + \\ &\quad (-WA - A^T W - W^2 + G)r \\ &= -(A^T W)r - W^2 r - h + Wb\ . \end{aligned}$$

Now the linear PDE which $\xi(t,x)$ must satisfy is:

$$\beta_t \cdot \xi - \beta \cdot \xi_t - (\Gamma x + \gamma) \cdot (\beta \cdot \nabla)\xi = 0\ . \tag{7.5}$$

This can be put in an even more concise form:

Theorem 7.1: If we assume $\eta = -x^T W\xi$, then $\zeta(t,x) \equiv \beta \cdot \xi = r^T W \xi = \nabla p \cdot \xi - \eta$ must satisfy the linear first order PDE

$$\zeta_t + ((\Gamma x + \gamma) \cdot \nabla)\zeta = 0\ . \tag{7.6}$$

Suppose ζ is a polynomial of order N in x, i.e.

$$\zeta(t,x) \equiv \sum_{k=0}^{N} \Xi^{(k)}(t)(x^{\otimes k})\ , \tag{7.7}$$

where $x^{\otimes k} \equiv x \otimes \cdots \otimes x$ (k factors), and $\Xi^{(k)}(t)$ is a symmetric $(0,k)$-tensor. Then

$$\nabla\zeta(t,x) = \sum_{k=1}^{N} k\Xi^{(k)}(t)(\cdot \otimes x^{\otimes(k-1)})\ .$$

Now (7.6) becomes

$$\sum_{0}^{N} \Xi_t^{(k)} x^{\otimes k} + \sum_{1}^{N} k\Xi^{(k)}((\Gamma x + \gamma) \otimes x^{\otimes(k-1)}) = 0\ .$$

Equating coefficients for each power of x forces the $\{\Xi^{(k)}(t)\}$ to satisfy the following system of ODEs:

$$\Xi_t^{(N)} + N\Xi^{(N)}(\Gamma(\cdot) \otimes \cdot) = 0\ ; \tag{7.8}$$

$$\Xi_t^{(k)} + k\Xi^{(k)}(\Gamma(\cdot) \otimes \cdot) = -(k+1)\Xi^{(k+1)}(\gamma \otimes \cdot) \quad \text{for } k = 1, 2, \ldots,\ N-1\ ; \tag{7.9}$$

$$\Xi_t^{(0)} = \Xi^{(1)}(\gamma) \ . \tag{7.10}$$

Notice the structure of this system of ODEs. Suppose $\Xi^{(0)}(0), \cdots, \Xi^{(N)}(0)$ have been chosen. First we solve (7.8) for $\Xi^{(N)}(t)$; insert this solution in the right side of (7.9); solve (7.9) for $\Xi^{(N-1)}(t)$; and so on, down to $\Xi^{(0)}(t)$. Thus we have established.

Theorem 7.2: In the case when $\eta = -x^T W \xi = \nabla p \cdot \xi - \zeta$, let us assume that

$$\zeta(t,x) \equiv r^T W \xi \equiv \sum_{k=0}^{N} \Xi^{(k)}(t)(x^{\otimes k}) \ . \tag{7.11}$$

Then $\zeta(t,x)$ is completely determined by the initial conditions $\Xi^{(0)}(0), \cdots, \Xi^{(N)}(0)$ and ODEs (7.8)-(7.10). In particular, when $n = 1$ and Wr is never zero, $\xi(t,x)$ is uniquely determined by $\xi(0,x)$, assuming $\xi(t,x)$ is a polynomial of arbitrary degree in x with coefficients depending on t.

Finally we describe the procedure for computation of the transformed information state. The starting-point is the solution p given by (6.12) to the linear control problem. Pick an initial condition $\Xi^{(0)}(0), \cdots \Xi^{(N)}(0)$, and solve for $\zeta(t,x)$ using (7.8)-(7.10) by solving for each of the $\{\Xi^{(k)}(t)\}$. Now pick

$$\xi(t,x) \equiv \sum_{k=1}^{N} \Theta^{(k)}(t)(x^{\otimes x}) \tag{7.12}$$

so that $r^T W \xi = \zeta$, in other words so that

$$\Theta^{(k)}(t) = W(t) r(t) \cdot \Xi^{(k)}(t) \ . \tag{7.13}$$

Now we repeat the steps described at the end of section 6, (6.23)-(6.26), under the assumption $\eta = -x^T W \xi = \nabla p \cdot \xi - \zeta$. As before, we solve the system of ODEs

$$\frac{\partial \bar{x}}{\partial \varepsilon} = \xi(t, \bar{x}), \ \bar{x}(0,t,x) = x \ ; \tag{7.14}$$

(derived from (6.25)) to determine $\bar{x}(\varepsilon, t, x, p)$. Thus $\xi(t,\bar{x})$ and $\zeta(t,\bar{x})$ are now explicitly computable. Finally we determine $\bar{p}(\varepsilon, t, x, p)$ by solving the following first order PDE (derived from (6.25)-(6.26)) by the method of characteristics (see Abraham et al. [18], p. 287):

$$\frac{\partial \bar{p}}{\partial \varepsilon} = \nabla \bar{p} \cdot \xi(t,\bar{x}) - \zeta(t,\bar{x}), \ \bar{p}(0,t,x,p) = p \ , \tag{7.15}$$

which can be written out in full as

$$\frac{\partial \bar{p}}{\partial \varepsilon} = \sum_{k=0}^{N} (\nabla \bar{p} \cdot \Theta^{(k)}(t) - \Xi^{(k)}(t))(\bar{x}^{\otimes k}) \ . \tag{7.16}$$

Additional examples-cases with explicit computations can be found in [29].

References

[1] G.W. Bluman and S. Kumei (1989): *Symmetries and Differential Equations*, Springer, New York.

[2] R.W. Brockett and J.M.C. Clark, "Geometry of the Conditional Density Equation", *Proc. Int. Conf. on An. and Opt. of Stoch. Syst.*, Oxford, England, 1978.

[3] R.W. Brockett, "Classification and Equivalence in Estimation Theory", *Proc. 18th IEEE Conf. on Dec. and Control*, 1979, pp. 172-175.

[4] S.K. Mitter, "Filtering Theory and Quantum Fields" presented in the **Conf. on Alg. and Geom. Methods in System Theory**, Bordeaux, France, 1978; in *Asterisque*, 1980.

[5] S.K. Mitter, "On the Analogy Between Mathematical Problems of Non-linear Filtering and Quantum Physics", in *Richerche di Automatica*, 1980.

[6] M. Hazewinkel and J. Willems (Edts), *Proc. of NATO ASI on Stoch. Syst.:The Mathematics of Filtering and Identification*, Les Arcs, France, 1980.

[7] G.W. Bluman and J.C. Cole, *Similarity Methods for 'Differential Equations*, Springer-Verlag, 1974.

[8] S.I. Rosencrans, "Perturbation Algebra of an Elliptic Operator", *J. of Math. Anal. and Appl.*, 56, 1976, pp. 317-329.

[9] R.S. Liptser and A.N. Shiryayev, *Statistics of Random Processes I*, Springer-Verlag, 1977.

[10] E. Nelson, "Analytic Vectors", *Ann. Math.*, 70, 1959, pp .572-615.

[11] V. Beneš, "Exact Finite Dimensional Filters for Certain Diffusions with Nonlinear Drift", in *Stochastics*, 1980.

[12] G.A. Nariboli, *Stoch. Proc. and their Appl.*, 5, 1977, pp. 157-171.

[13] E. Hille and R. Phillips, **Funct. An. and Semigroups**, AMS Coll. Publ., 1957.

[14] S.I. Rosencrans, *J. of Math. An. and Appl.*, 61, pp. 537-551.

[15] Sophus Lie, *Arch. for Math.*, Vol VI, 1881, No. 3, Kristiana, p. 328-368.

[16] L.V. Ovsjannikov, *Group Properties of Diff. Equations*, Sib. Branch of the USSR Ac. of Sci. 1962.

[17] R.L. Anderson and N.H. Ibragimov. *Lie-Backlund Transformations in Applications*, SIAM, 1979.

[18] R. Abraham, J.E. Marsden, and T. Ratiu (1988): *Manifolds, Tensor Analysis, and Applications*, 2nd ed., Springer, New York.

[19] R.W.R. Darling (1994): *Differential Forms and Connections*, Cambridge University Press, New York.

[**20**] M.R. James, J.S. Baras and R.J. Elliott, “Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems”, *32nd IEEE CDC*, San Antonio, 1993.

[**21**] M.R. James and S. Yuliar, “A nonlinear partially observed differential game with a finite-dimensional information state”, *Systems and Control Letters*, 1996.

[**22**] M.R. James, J.S. Baras and R.J. Elliott, “Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems”, *IEEE Trans. on Automatic Control*, pp. 780-792, Vol. 39, No. 4, April 1994.

[**23**] M.R. James and J.S. Baras, “Robust $H\infty$ output feedback control for nonlinear systems”, *IEEE Trans. on Automatic Control*, pp. 1007-1017, Vol. 40, No. 6, June 1995.

[**24**] M.R. James and J.S. Baras, “Partially observed differential games, infinite dimensional HJI Equations, and Nonlinear H_∞ Control”, *SIAM Journal on Control and Optimization*, Vol. 34, No. 4, pp. 1342-1364, July 1996.

[**25**] J.S. Baras and M.R. James, “Robust and risk-sensitive output feedback control for finite state machines and hidden markov models”, *Journal of Mathematical Systems, Estimation, and Control*, Vol. 7, No. 3, pp. 371-374, 1997.

[**26**] J.S. Baras and N. Patel, “Robust control of set-valued discrete time dynamical systems”, *IEEE Transactions on Automatic Control*, Vol. 43, No. 1, pp. 61-75, January 1998.

[**27**] J.B. Moore and J.S. Baras, “Finite-dimensional optimal controllers for nonlinear plants”, *Systems and Control Letters*, 26 (1995), pp. 223-230.

[**28**] R. Elliott and A. Bensoussan: “General Finite-Dimensional Risk-Sensitive Problems and Small Noise Limits”, *IEEE Trans. on Automatic Control*, Vol. 41, pp. 210-215, January 1996.

[**29**] J.S. Baras and R. Darling, "Finite-Dimensional Methods for Computing the Informatin State in Nonlinear Robust Control”, *Proc. of 1998 IEEE Conference on Decision and Control*, pp. 343-348, Tampa, Florida, 1998.