METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 11, No. 3, pp. 303–316, September 2004

001

A NEW THEORY AND THE EFFICIENT METHODS OF SOLUTION OF STRONG, PATHWISE, STOCHASTIC VARIATIONAL PROBLEMS*

JOHN GREGORY† AND H. R. HUGHES†

Abstract. This paper has two major objectives. The first objective is to give a new, natural extension of the classical calculus of variations to a stochastic setting. Most significantly, it appears that this is the first time in the literature that a random objective functional is used rather than its mean. This is accomplished by using an appropriate class of variations.

Our second major objective is to give an efficient method of solution for these problems. To do this we derive the Euler equation in this setting which is the primary necessary condition for a critical point or extremal solution. We also derive transversality and corner conditions. These results, which are the three basic necessary conditions for an extremal, are sufficient to construct closed form solutions, when they exist. Our results hold for $n \geq 1$ dependent variables.

Of particular interest are several example problems. They illustrate random objective functionals that are functions of a Brownian path, pathwise extremals of these functionals, and the use of necessary conditions to find solutions in closed form. We also illustrate that our theory can generalize the usual stochastic control theory by considering individual paths and not the mean value of the objective functional.

Finally, these results illustrate steps towards a complete stochastic variational theory for random objective functionals by extending deterministic ideas of the first author. The three necessary conditions of this paper allow a general stochastic control theory with general equality and inequality constraints.

Key words. Optimal Control, Stochastic Analysis, Calculus of Variations

AMS subject classifications. Primary 93E20, 49K45, 60H99

1. Introduction. It is surprising that there is currently no natural stochastic extension of the deterministic methods for solving constrained optimization problems in which the calculus of variations is used to obtain Euler equations and other necessary conditions. Such an extension would contain a stochastic control theory that can handle general constraints. Perhaps, a major reason is that the mathematical theory of stochastic differential equations has not been fully exploited in solving these problems.

In the last several years, the mathematical theory of stochastic differential equations has become well understood as a mathematical theory. Textbooks and references are now available to explain the basic theory and the important uses of these equations. In addition, insightful texts such as [9] provide many of the tools and motivation so that we can begin to understand the numerical aspects of these equations. In a sense, the situation corresponds to that of ODE's in the early 1960's when the landmark work of Henrici [8], for the numerical solution of ordinary differential equations, appeared.

These advances appear to have little effect on the important areas of stochastic optimization. Although some progress has been made using deep mathematical methods to extend Hamiltonian-Jacobi-Bellman ideas to specialized control problems, these results are very limited for applied problems when compared with the deterministic situation where (for example) the first author has given a complete numerical solution to general constrained problems in the calculus of variations/optimal control

*Received December 24, 2003; accepted for publication April 28, 2004.
†Department of Mathematics, Southern Illinois University Carbondale, Carbondale IL 62901-4408, USA (jgregory@math.siu.edu; hrhughes@math.siu.edu).

theory with algorithms which have a global error of order $O(h^p)$, $p \geq 2$ (see [5] and [6]). They are also limited in the sense they do not provide strong, stochastic solutions for optimization problems.

Specifically, we believe there is currently a beautiful mathematical theory of stochastic optimal control as exemplified by the text of Chen, Chen and Hsu [2]. However, our approach has several advantages:

a) Our objective functional is not averaged over the probability measure. It depends on individual stochastic paths and not just the probability law. Thus, there is a natural extension from the deterministic calculus of variations/optimal control theory to finding strong solutions in a stochastic sense.
b) Our methods are associated with the increasingly important area of stochastic differential equations. Solutions will be given by these equations and simpler problems will be solved in this way.
c) There is a direct connection to the classical calculus of variations and hence both the theoretical ideas of this most important area of applied mathematics and the motivation provided by many of history's most important applied problems can be easily explored in a stochastic setting.
d) The practical method of solution of the current theory relies on dynamic programming and hence is, at best, very difficult. Our simpler problems will be solved by the use of stochastic differential equations. More complex problems will be solved by efficient numerical procedure with an a priori error estimate of the form $O(h^p)$ where h is the node size and p is at least 1.5. This is what one would expect by analogy with the theory of stochastic [4] or deterministic differential equations [6] ... or as the first author has done for general, deterministic constrained calculus of variations/optimal control problems.
e) Our methods allow us to handle equality and inquality constraints in this setting.

In the deterministic case cited above, a numerical theory and efficient algorithms were first given to find a critical point solution for

$$I(x) = \int_a^b f(t, x, x')\, dt$$

with general boundary/transversality conditions [6] and then a companion theory was given [5] or [7] to efficiently convert general constrained problems in the calculus of variations/optimal control theory to this setting so that the extremal solutions were easily identified.

The problem we will consider in this paper involves the random cost functional

$$J(x, u) = \int_a^b f(t, x, u)\, dt + k(x(b)) \tag{1a}$$

and the trajectory equation

$$dx_t = u\, dt + \sigma(t)\, dW_t. \tag{1b}$$

We also extend our results to the more general trajectory equation

$$dx_t = g(t, x, u)\, dt + \sigma(t)\, dW_t. \tag{1c}$$

This setting involves a multitude of important physical problems in engineering and the sciences. It is also the first step from the deterministic case toward a general, pathwise, stochastic control theory.

Our major theoretical result in this paper will be to obtain stochastic Euler equations for pathwise critical points of variations of the functional J subject to (1b) or (1c). In addition, we obtain the corner conditions and transversality conditions so that feasible methods to obtain solutions can be implemented. We expect to use this variational approach to obtain efficient numerical algorithms for the solutions of stochastic calculus of variations and constrained stochastic optimization problems similar to the deterministic case (described above).

The remainder of this paper is as follows. In Section 2, we will define our basic problem using trajectory (1b) and derive the Euler equation, transversality conditions, and corner condition for this problem. We also consider examples with closed form solutions. Example 2 is of particular interest. It illustrates that our "extremal"/critical point solutions are random and lead to an expected cost which is smaller than that in the adapted case [11], which we also derive in an alternate way. In Section 3, we extend these results to the vector case for the dependent variable $x(t)$ with $n \geq 1$ components subject to variation and $m \geq 0$ components that are not varied. These latter components are the solution of a stochastic differential equation. In Section 4, we generalize our results for trajectory equations of the form (1c).

2. Pathwise solutions. In this section we define a new stochastic calculus of variations problem. Like the Malliavin calculus, we limit variations to a Cameron-Martin space [10, pp. 24–25]. Even with this limitation, directional derivatives of the cost functional yield a stochastic Euler equation (2.6) and other necessary conditions.

Let W_t be standard Brownian motion and $\mathcal{F}_t$ the corresponding filtration of σ-algebras. In this section, we seek extremals for well-defined problems associated with the random cost functional

$$J(x,u) = \int_a^b f(t,x,u)\,dt + k(x(b)), \tag{2.1a}$$

where

$$dx_t = u\,dt + \sigma(t)\,dW_t, \tag{2.1b}$$

and $u(t)$ is piecewise continuous in t. Conditions on f and its derivatives are assumed as needed.

Let $H_{a,b} = \{z\colon [a,b] \to \mathbf{R} \mid z(t)$ is absolutely continuous with respect to t and $\int_a^b (z'(t))^2\,dt < \infty\}$. Also let $H^0_{a,b} = \{z \in H_{a,b} \mid z(a) = 0 = z(b)\}$. We consider the variation of J for $z \in H_{a,b}$,

$$I(x,u,z,\epsilon) = J(x+\epsilon z, u+\epsilon z') \tag{2.2}$$

and seek critical point solutions to optimize J:

$$\frac{\partial}{\partial \epsilon} I(x,u,z,\epsilon)\bigg|_{\epsilon=0} = \int_a^b \big(z f_x(t,x,u) + z' f_u(t,x,u)\big)\,dt + k'(x(b))\,z(b) = 0, \tag{2.3}$$

for all $z \in H_{a,b}$, almost surely. Note that

$$d(x+\epsilon z)_t = (u+\epsilon z')\,dt + \sigma(t)\,dW_t \tag{2.4}$$

so that if $x(t,\epsilon) = x(t) + \epsilon z(t)$ and $u(t,\epsilon) = u(t) + \epsilon v(t)$ are respectively families of admissible arcs such that $x(t,0) = x(t)$, $u(t,0) = u(t)$, and

$$d(x + \epsilon z)_t = (u + \epsilon v)\,dt + \sigma(t)\,dW_t, \tag{2.5}$$

then $v(t) = z'(t)$.

THEOREM 1. *A critical point solution to* (2.1) *satisfies the stochastic differential equation,*

$$d(f_u(t,x,u))_t = f_x(t,x,u)\,dt, \tag{2.6a}$$
$$dx_t = u\,dt + \sigma(t)\,dW_t. \tag{2.6b}$$

In addition, the critical point solution satisfies the transversality conditions:

$$x(a) \text{ not specified implies } f_u(a,x(a),u(a)) = 0; \tag{2.7a}$$
$$x(b) \text{ not specified implies } f_u(b,x(b),u(b)) + k'(x(b)) = 0. \tag{2.7b}$$

Finally, the critical point solution satisfies the corner condition,

$$f_u(t,x(t),u(t)) \text{ is continuous on the interval } [a,b]. \tag{2.8}$$

Proof. Suppose $\int_a^b (zf_x + z'f_u)\,dt + k'(x(b))\,z(b) = 0$ for all $z \in H_{a,b}$, a.s. Let $y(t) = \int_a^t f_x(s,x(s),u(s))\,ds$. Then, integrating by parts,

$$\int_a^b \big(f_u(t,x(t),u(t)) - y(t)\big)z'\,dt + \big(y(b) + k'(x(b))\big)z(b) = 0, \tag{2.9}$$

for all $z \in H_{a,b}$, a.s. Restricting z to $H^0_{a,b}$ and using the Lemma of Dubois-Reymond in [12, p. 50] or [5, p. 36], it follows that

$$f_u(t,x(t),u(t)) - \int_a^t f_x(s,x(s),u(s))\,ds = f_u(a,x(a),u(a)) \quad \text{a.s.}, \tag{2.10}$$

which gives (2.6a) in integral form. The continuity of $f_u(t,x(t),u(t))$ follows immediately from (2.10).

To derive the transversality conditions, first note that (2.9) and (2.10) imply that

$$\begin{aligned} 0 &= \int_a^b f_u(a,x(a),u(a))\,z'\,dt + \big(y(b) + k'(x(b))\big)z(b) \\ &= \big(f_u(a,x(a),u(a)) + y(b) + k'(x(b))\big)z(b) - f_u(a,x(a),u(a))\,z(a) \\ &= \big(f_u(b,x(b),u(b)) + k'(x(b))\big)z(b) - f_u(a,x(a),u(a))\,z(a). \end{aligned} \tag{2.11}$$

Now consider $z \in H_{a,b}$ with different boundary conditions for z. For $z(a) \neq 0$ and $z(b) = 0$, it follows that $f_u(a,x(a),u(a)) = 0$. For $z(a) = 0$ and $z(b) \neq 0$, it follows that $f_u(b,x(b),u(b)) + k'(x(b)) = 0$. ∎

Note that, equation (2.10) shows that $f_u(t,x(t),u(t))$ is absolutely continuous and hence (2.6a) involves the standard differential. In the case that u is an Itô process, if

$f \in C^3$ and $f_{uu} > 0$, (2.6) can also be written explicitly as the system of stochastic differential equations,

$$du_t = \left(\frac{f_x - f_{ut} - u f_{ux}}{f_{uu}} - \frac{\sigma^2}{2} \frac{(f_{ux})^2 f_{uuu} + 2 f_{ux} f_{uu} f_{uux} + (f_{uu})^2 f_{uxx}}{(f_{uu})^3} \right) dt - \sigma \frac{f_{ux}}{f_{uu}} \, dW_t, \tag{2.12a}$$

$$dx_t = u\, dt + \sigma\, dW_t, \tag{2.12b}$$

where the arguments of the derivatives of f are always (t, x, u) and $\sigma = \sigma(t)$.

To illustrate our ideas and the results in Theorem 1, above, we consider the following two examples. Example 1, where we calculate the solution explicitly, illustrates that our theory is quite different from the more classical stochastic control theory. For example, our functionals in (2.1a) are random, whereas the classical theory deals only with the probabilistic mean of these functionals.

EXAMPLE 1. Consider the following problem.

$$\min \ \frac{1}{2} \int_0^1 u^2 \, dt \tag{2.13a}$$

$$\text{s.t.} \quad dx_t = u\, dt + \sigma\, dW_t \tag{2.13b}$$

$$\text{and} \quad x(0) = 0, \quad x(1) = 1. \tag{2.13c}$$

In the deterministic case where $\sigma \equiv 0$ and $u = x'$ the solution can be found by the Euler equation

$$x'' = \frac{d}{dt} f_{x'} = f_x = 0 \tag{2.14}$$

which implies $x(t) = t$ because of the boundary conditions.

In the stochastic case, if we assume $\sigma \neq 0$ and, for convenience, $W(0) = 0$, the necessary conditions (from Theorem 1) are

$$\begin{aligned} du_t &= df_u = f_x \, dt = 0, \\ dx_t &= u\, dt + \sigma\, dW_t, \\ x(0) &= 0, \quad x(1) = 1 \end{aligned} \tag{2.15}$$

which implies $u(t) \equiv c$ and hence

$$x(t) = x(0) + \int_0^t u(s)\, ds + \int_0^t \sigma\, dW_s = ct + \sigma W_t. \tag{2.16}$$

Now $x(1) = 1 = c + \sigma W_1$ or $c = 1 - \sigma W_1$ and $x(t) = (1 - \sigma W_1)t + \sigma W_t$ which is a stochastic process.

The cost functional is minimized at the value

$$I_1 = \frac{1}{2} \int_0^1 (1 - \sigma W_1)^2 \, dt = \frac{1}{2} (1 - \sigma W_1)^2 \tag{2.17}$$

in the sense that

$$\min \frac{1}{2} \int_0^1 (c + \eta'(t))^2 \, dt = \frac{1}{2} \int_0^1 c^2 \, dt \tag{2.18}$$

for deterministic variation η s.t. $\eta(0) = \eta(1) = 0$ and η is absolutely continuous on $[0, 1]$. We note again a difference between our problem formulation and the classical stochastic control formulation such as [2], where the functional value is deterministic, and not a random variable, as I_1 is in our example. We have that

$$E(I_1) = \frac{1}{2}\left[1 - 2\sigma E(W_1) + \sigma^2 E(W_1^2)\right] = \frac{1}{2} + \frac{\sigma^2}{2} \tag{2.19}$$

differs from I_1 for $\sigma \neq 0$. Finally, we note that I_1 is a function of the Brownian path and we refer to the solution as a strong solution.

EXAMPLE 2. In this example, consider the trajectory

$$dx_t = u\,dt + \sigma\,dW_t, \quad x_0 = \xi, \tag{2.20}$$

where σ and ξ are constants. Consider also the random functional,

$$J(x, u) = \int_0^b \frac{1}{2}u^2\,dt + \frac{r}{2}(x_b)^2 \tag{2.21}$$

where r is a constant.

The Euler-Lagrange equations in this case are

$$\begin{aligned} dx_t &= u\,dt + \sigma\,dW_t \\ du_t &= 0\,dt \end{aligned} \tag{2.22}$$

and the transversality condition

$$u_b + r\,x_b = 0, \tag{2.23}$$

along with the initial condition, $x_0 = \xi$, give a system of SDE's. However we must determine how to interpret the end conditions.

One option is to allow anticipating solutions. In this case, we have $u \equiv c$, constant with respect to time, but possibly only measurable with respect to $\mathcal{F}_b$. It follows then that

$$x_t = \xi + c\,t + \sigma\,W_t \tag{2.24}$$

and using the end condition,

$$c + r(\xi + c\,b + \sigma\,W_b) = 0, \tag{2.25}$$

we have critical solution

$$\begin{aligned} u_t^* &= c = \frac{-r(\xi + \sigma\,W_b)}{1 + r\,b}\,, \\ x_t^* &= \xi - \frac{r(\xi + \sigma\,W_b)t}{1 + r\,b} + \sigma\,W_t. \end{aligned} \tag{2.26}$$

The value of the functional is random and given by

$$J^* = \frac{r(\xi + \sigma\,W_b)^2}{2(1 + r\,b)}\,, \tag{2.27}$$

which depends on W_b. The processes x_t and u_t also depend on the end value W_b.

Thus our pathwise solution is anticipating and critical with respect to variation in the specified directions in the Cameron-Martin space. This problem differs from the treatment of anticipative stochastic control in Davis [3] where mean cost functionals are considered for arbitrary anticipative controls. Allowing anticipative controls in LQG control problems, Davis obtains lower mean cost when compared to the adapted case. We obtain similar results as noted below.

A second approach to interpret the end conditions is to treat the system of stochastic differential equations as a coupled forward-backwards system. In this case we seek the projection onto the space of adapted solutions by conditioning on $\mathcal{F}_t$ while satisfying $u_b + r x_b = 0$. In particular, since in the original system, $u_t = u_b$ is constant, we have $\tilde{u}_t = E(u_t|\mathcal{F}_t) = E(u_b|\mathcal{F}_t)$ is a martingale. Thus there exists an auxiliary adapted process V such that $d\tilde{u}_t = V_t\, dW_t$. Replacing $\tilde{u}$ with u, we have the system:

$$\begin{aligned} dx_t &= u\, dt + \sigma\, dW_t \\ du_t &= V_t\, dW_t \\ x_0 &= \xi \\ u_b &= -r x_b. \end{aligned} \tag{2.28}$$

Given the form of the last condition, we guess that the process u_t may be of the form $u_t = \theta(t) x_t$, where $\theta(b) = -r$. Then

$$\begin{aligned} du_t &= \theta'(t) x_t\, dt + \theta(t)\, dx_t \\ &= (\theta' + \theta^2) x_t\, dt + \sigma\theta\, dW_t. \end{aligned} \tag{2.29}$$

If u_t is of this form, then

$$\theta' + \theta^2 = 0, \qquad \theta(b) = -r, \tag{2.30}$$

and

$$V_t = \sigma\theta(t). \tag{2.31}$$

Solving for $\theta(t)$ we have

$$\begin{aligned} \theta(t) &= \frac{-r}{1 + r(b-t)}\ , \\ du_t &= \frac{-r\,\sigma}{1 + r(b-t)}\, dW_t, \\ dx_t &= \frac{-r\, x_t}{1 + r(b-t)}\, dt + \sigma\, dW_t. \end{aligned} \tag{2.32}$$

Substituting,

$$y_t = \frac{1 + r\,b}{\xi[1 + r(b-t)]}\, x_t, \tag{2.33}$$

we have

$$\begin{aligned} dy_t &= \frac{1 + r\,b}{\xi[1 + r(b-t)]}\, \sigma\, dW_t, \\ y_0 &= 1. \end{aligned} \tag{2.34}$$

Hence the optimal solutions are

$$
\begin{aligned}
y_t^* &= 1 + \int_0^t \frac{1 + r\,b}{\xi[1 + r(b - s)]}\,\sigma\,dW_s \\
x_t^* &= \frac{\xi(1 + r(b - t))}{1 + r\,b} + [1 + r(b - t)] \int_0^t \frac{\sigma}{1 + r(b - s)}\,dW_s \qquad (2.35)\\
u_t^* &= \frac{-r\xi}{1 + r\,b} - r \int_0^t \frac{\sigma}{1 + r(b - s)}\,dW_s.
\end{aligned}
$$

The minimum random functional $J(x^*, u^*)$ can then be computed from these solutions. In this case, the mean of the functional is

$$E[J] = \frac{r\,\xi^2}{2(1 + r\,b)} + \frac{\sigma^2}{2}\ln(1 + r\,b). \tag{2.36}$$

We note that here u_t is an explicit function of time, t, and x_t for all t. We also note that the mean $E[J]$ is greater than the mean in the anticipating case,

$$E\left[\frac{r(\xi + \sigma\,W_b)^2}{1 + r\,b}\right] = \frac{r\,\xi^2}{2(1 + r\,b)} + \frac{\sigma^2 r\,b}{2(1 + r\,b)}. \tag{2.37}$$

There is an extra cost for the adapted case.

Finally, we note that the mean cost obtained in the adapted (FBSDE) case is identical to that obtained by the HBJ methods for the *traditional* stochastic control problem [11, pp. 220–222].

In our third example, we consider the following stochastic version of the harmonic oscillator problem. Our purpose is to illustrate that there are physically interesting problems and that nontrivial problems can be solved explicitly by extending the usual deterministic techniques (i.e., when $\sigma = 0$).

EXAMPLE 3. We start with the cost functional

$$J(x, u) = \frac{1}{2}\int_0^b (u^2 - \alpha^2 x^2)\,dt \tag{2.38a}$$

where

$$dx_t = u\,dt + \sigma\,dW_t. \tag{2.38b}$$

The Euler equations are then

$$du_t = -\alpha^2 x\,dt, \tag{2.39a}$$

$$dx_t = u\,dt + \sigma\,dW_t, \tag{2.39b}$$

which have a general solution of the form

(2.40a)

$$x_t = x_0 \cos\alpha t + \frac{u_0}{\alpha}\sin\alpha t + \cos\alpha t \int_0^t \sigma\cos\alpha s\,dW_s + \sin\alpha t\int_0^t \sigma\sin\alpha s\,dW_s,$$

(2.40b)

$$u_t = -\alpha x_0 \sin\alpha t + u_0\cos\alpha t - \sin\alpha t\int_0^t \alpha\sigma\cos\alpha s\,dW_s + \cos\alpha t\int_0^t \alpha\sigma\sin\alpha s\,dW_s.$$

Considering the anticipating solution, boundary conditions on $x(0)$ and $x(b)$ determine x_0 and u_0 and the latter are in general only $\mathcal{F}_b$-measurable. For example, if $x(0) = 0 = x(b)$, then when αb is not an integral multiple of π,

$$x_0 = 0, \tag{2.41a}$$

$$u_0 = -\alpha \cot \alpha b \int_0^b \cos \alpha s \, dW_s - \alpha \int_0^b \sin \alpha s \, dW_s. \tag{2.41b}$$

To obtain these results we proceed as follows. From the stochastic Euler equations with $\mathrm{f}_u = u$, $\mathrm{f}_x = -\alpha^2 x$ we have

$$\begin{aligned} du_t &= -\alpha^2 x_t \, dt \\ dx_t &= u_t \, dt + \sigma \, dW_t \end{aligned} \tag{2.42}$$

and setting $X = (x, u)^T$, where "T" denotes transpose, this becomes

$$\begin{aligned} dX &= MX \, dt + \Sigma \, dW_t \\ &= \begin{pmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{pmatrix} X \, dt + \begin{pmatrix} \sigma \\ 0 \end{pmatrix} dW_t. \end{aligned} \tag{2.43}$$

Proceeding as in the deterministic case we note that

$$d(e^{-Mt} X) = e^{-Mt} \Sigma \, dW_t \tag{2.44}$$

and hence

$$X_t = e^{Mt} X(0) + e^{Mt} \int_0^t e^{-Ms} \Sigma \, dW_s. \tag{2.45}$$

In particular, with $e^{-Mt} = \begin{pmatrix} \cos \alpha t & -\frac{1}{\alpha} \sin \alpha t \\ \alpha \sin \alpha t & \cos \alpha t \end{pmatrix}$, we have

$$d\left(e^{-Mt} \begin{pmatrix} x \\ u \end{pmatrix}\right) = \begin{pmatrix} \cos \alpha t & -\frac{1}{\alpha} \sin \alpha t \\ \alpha \sin \alpha t & \cos \alpha t \end{pmatrix} \begin{pmatrix} \sigma \\ 0 \end{pmatrix} dW_t = \begin{pmatrix} \sigma \cos \alpha t \, dW_t \\ \sigma \alpha \sin \alpha t \, dW_t \end{pmatrix}. \tag{2.46}$$

Integrating and multiplying by e^{Mt} gives the general solution in (2.40).

3. Multidimensional problems. The purpose of this section is to extend the results in Section 2. We consider multidimensional X_t and U_t with a more general trajectory equation of the form

$$dX_t = F_0(t, X, U) \, dt + \sum_{i=1}^{r} F_i(t, X, U) \, dW_t^i, \tag{3.1}$$

where W_t is an r-dimensional Brownian motion. In order for the method of Section 2 to carry over directly, we must put restrictions on the functions F_i, $i = 0, 1, 2, \ldots, r$. We allow for X_t to be composed of n components that are varied in the cost functional and m components that are not. Suppose

$$X(t) = \big(\bar{X}(t), Y(t)\big), \tag{3.2}$$

where $\bar{X}(t)$ and $Y(t)$ are, respectively, n- and m-dimensional processes for integers $n \geq 1$ and $m \geq 0$,

$$Z(t) = \big(\bar{Z}(t), 0\big) \tag{3.3}$$

where $\bar{Z}(t) = \big(z_1(t), z_2(t), \ldots, z_n(t)\big)$ and $z_i \in H_{a,b}$ for each i,

$$U(t) = \big(\bar{U}(t), 0\big) \tag{3.4}$$

where $\bar{U}(t)$ is an n-dimensional process, piecewise continuous in t, and

$$F_i(t, X, U) = \big(\bar{F}_i(t, X, U), G_i(t, X, U)\big), \tag{3.5}$$

where $\bar{F}_i$ and G_i are n- and m-dimensional vector-valued functions.

The method of Section 2 carries over directly if

$$\begin{aligned} dX_t + \epsilon Z'_t\,dt &= d(X + \epsilon Z)_t \\ &= F_0(t, X + \epsilon Z, U + \epsilon Z')\,dt + \sum_{i=1}^{r} F_i(t, X + \epsilon Z, U + \epsilon Z')\,dW_t^i. \end{aligned} \tag{3.6}$$

To guarantee that (3.6) holds we restrict F so that $\bar{F}_0(t, X, U) = U + \hat{F}_0(t, Y)$, $\bar{F}_i(t, X, U) = \hat{F}_i(t, Y)$ is independent of $\bar{X}$ and U for $i = 1, 2, \ldots, r$, and $G_i(t, X, U) = \hat{G}_i(t, Y)$ is independent of $\bar{X}$ and U for $i = 0, 1, \ldots, r$. The trajectory equation therefore becomes

$$d\bar{X}_t = \big(\bar{U} + \hat{F}_0(t, Y)\big)\,dt + \sum_{i=1}^{r} \hat{F}_i(t, Y)\,dW_t^i, \tag{3.7a}$$

$$dY_t = \hat{G}_0(t, Y)\,dt + \sum_{i=1}^{r} \hat{G}_i(t, Y)\,dW_t^i. \tag{3.7b}$$

For simplicity, we illustrate the extensions described above with a pair of two-dimensional examples.

EXAMPLE 4. Let $X = (x_1, x_2)^T$, $U = (u_1, u_2)^T$ and consider the cost functional

$$J(x, u) = \int_a^b f(t, x_1, x_2, u_1, u_2)\,dt, \tag{3.8a}$$

where

$$dX_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} U\,dt + \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} dW_t. \tag{3.8b}$$

Letting $Z = (z_1, z_2)^T$ for $z_1, z_2 \in H_{a,b}$, we have

$$\frac{d}{d\epsilon} I(X, U, Z, \epsilon)\bigg|_{\epsilon=0} = \int_a^b \big(Z^T f_X(t, X, U) + {Z'}^T f_U(t, X, U)\big)\,dt = 0. \tag{3.9}$$

Integrating by parts and applying the Dubois-Reymond Lemma componentwise, we obtain

$$d(f_U(t, X, U))_t = f_X(t, X, U)\,dt \tag{3.10}$$

along with (3.8b).

Example 5. Similarly, we can consider the cost functional

$$J(x,u) = \int_a^b f(t,x,W,u)\,dt, \tag{3.11a}$$

where W is a Brownian motion. Letting $X = (x_1, x_2)^T = (x, W)^T$ and $U = (u, 0)^T$, we consider the trajectory equation

$$dX_t = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U\,dt + \begin{pmatrix} \sigma \\ 1 \end{pmatrix} dW_t. \tag{3.11b}$$

Letting $Z = (z, 0)^T$ for $z \in H_{a,b}$, we have

$$\begin{aligned} \frac{d}{d\epsilon} I(X,U,Z,\epsilon)\bigg|_{\epsilon=0} &= \int_a^b \left(Z^T f_X(t,X,U) + Z'^T f_u(t,X,U)\right) dt \\ &= \int_a^b \left(z f_x(t,x,W,u) + z' f_u(t,x,W,u)\right) dt = 0. \end{aligned} \tag{3.12}$$

Integrating by parts and applying the Dubois-Reymond Lemma, we obtain

$$d(f_u(t,x,W,u))_t = f_x(t,x,W,u)\,dt \tag{3.13}$$

along with (3.11b).

4. More general trajectory equations and constraint problems. The purpose of this section is to show that more general problems can be reduced to the problem (2.1) in vector form. Thus, the trajectory equation (2.1b) initially looks innocuous or simplistic but we can change a variety of complex problems into this form. Specifically, for convenience of exposition, to the form

$$J(X,U) = \int_a^b F(t,X,U)\,dt + K(X(b)), \tag{4.1a}$$

where

$$dX_t = U\,dt + \Sigma(t)\,dW_t. \tag{4.1b}$$

In fact this reduction and the ideas of this section, hold for general F.

Our first example is where (1c) replaces (1b). We include the discussion since it shows how strong Theorem 1 is.

In order to do this we form the multidimensional problem where

$$F(t,Y,V) = f(t,y_1,v_2)\,dt + v_3[g(t,y_1,v_2) - v_1]. \tag{4.2}$$

Thus we consider finding a critical point solution for

$$\int_a^b F(t,Y,V)\,dt + k(y_1(b)) \tag{4.3a}$$

where

$$dY = \begin{pmatrix} dy_1 \\ dy_2 \\ dy_3 \end{pmatrix} = \begin{pmatrix} v_1\,dt + \sigma(t)\,dW_t \\ v_2\,dt \\ v_3\,dt \end{pmatrix} = V\,dt + \begin{pmatrix} \sigma(t) \\ 0 \\ 0 \end{pmatrix} dW_t. \tag{4.3b}$$

We note, applying our earlier results on multidimensional problems, that a critical point solution for this latter problem satisfies $d(F_V) = F_Y\,dt$. Thus, in component form we have

$$d\begin{pmatrix} -v_3 \\ f_{v_2} + v_3 g_{v_2} \\ g - v_1 \end{pmatrix} = \begin{pmatrix} f_{y_1} + v_3 g_{y_1} \\ 0 \\ 0 \end{pmatrix} dt. \tag{4.4}$$

In addition, since $y_2(b)$ and $y_3(b)$ are unspecified, the transversality conditions (2.7) imply that $F_{v_2}(b, Y(b), V(b)) = 0$ and $F_{v_3}(b, Y(b), V(b)) = 0$. From equation (4.4) and the corner conditions it follows that $g - v_1 \equiv 0$ on $[a, b]$, which yields

$$dy_1 = g(t, y_1, v_2) + \sigma(t)\,dW_t, \tag{4.5}$$

and $f_{v_2} + v_3 g_{v_2} \equiv 0$ on $[a, b]$. Thus $v_3 = -f_{v_2}/g_{v_2}$ for $g_{v_2} \neq 0$. To specify a particular soluton, we may assume $y_2(a) = 0$ and $y_3(a) = 0$. However, these values do not effect the other variables. Thus, for $g_u \neq 0$, the system defined by (4.3b) and (4.4) can be reduced to a system in $x = y_1$ and $u = v_2$ alone.

THEOREM 2. *A critical point solution to* (1), (1c) *satisfies the stochastic differential equation,*

$$d\left(\frac{f_u}{g_u}\right)_t = \left(f_x - \frac{g_x f_u}{g_u}\right) dt, \tag{4.6a}$$

$$dx_t = g(t, x, u)\,dt + \sigma(t)\,dW_t. \tag{4.6b}$$

In addition, the critical point solution satisfies the transversality conditions:

$$x(a) \textit{ not specified implies } f_u(a, x(a), u(a)) = 0; \tag{4.7a}$$

$$x(b) \textit{ not specified implies } \frac{f_u(b, x(b), u(b))}{g_u(b, x(b), u(b))} + k'(x(b)) = 0. \tag{4.7b}$$

Finally, the critical point solution satisfies the corner condition,

$$\frac{f_u(t, x(t), u(t))}{g_u(t, x(t), u(t))} \textit{ is continuous on the interval } [a, b]. \tag{4.8}$$

We note that the last example could be thought of as a constrained problem with constraint $h(t, Y, V) = v_1 - g(t, y_1, v_2) = 0$. Following the ideas in this last example, necessary conditions can be derived for critical points of functional

$$\int_a^b f(t, x, u)\,dt + k(x(b)), \tag{4.9a}$$

subject to the trajectory equation,

$$dx_t = u\,dt + \sigma(t)\,dW_t, \tag{4.9b}$$

and subject to the constraint,

$$h(t, x, u) = 0. \tag{4.9c}$$

Thus a quite general problem can be "solved" in the sense that, if it has unique solution we will find a closed form solution in the simplest cases using the necessary conditions derived above. In more complicated cases, we expect to solve the problem numerically using an algorithm similar to the deterministic algorithm in [5] and [6].

Finally, we would like to thank the referees for their insightful comments.

REFERENCES

[1] Bellman, R., *Dynamic Programming*, Princeton University Press, 1957.

[2] Chen, G., Chen, G., and Hsu, S-H., *Linear Stochastic Control Systems*, CRC Press, Boca Raton, 1995.

[3] Davis, M. H. A., *Anticipative LQG control,* IMA J. Math. Control Inform., 6 (1989), pp. 259–265.

[4] Gregory, J. and Hughes, H. R., *New general methods for numerical stochastic differential equations,* Utilitas Mathematica, *to appear*.

[5] Gregory, J. and Lin, C., *Constrained Optimization in the Calculus of Variations and Optimal Control Theory,* Van Nostrand Reinhold, New York, 1992.

[6] Gregory, J. and Lin, C., *Discrete variable methods for the m-dependent variable, nonlinear extremal problem in the calculus of variations, II,* SIAM Journal of Numerical Analysis, 30 (1993), pp. 871–883.

[7] Gregory, J. and Lin, C., *An unconstrained calculus of variations formulation for generalized optimal control problems and for the constrained problem of Bolza,* Journal of Mathematical Analysis and Applications, 187 (1994), pp. 826–841.

[8] Henrici, P. *Discrete Variable Methods in Ordinary Differential Equations*, John Wiley and Sons, Inc., New York, 1962.

[9] Kloeden, P. E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics, Stochastic Modelling, and Applied Probability, vol. 23, Springer, New York, Heidelberg, Berlin, 1992.

[10] Nualart, D., *The Malliavin Calculus and Related Topics*, Probability and its Applications, Springer, New York, Berlin, Heidelberg, 1995.

[11] Øksendal, B., *Stochastic Differential Equations: An Introduction with Applications*, 4th ed. , Springer, Berlin, Heidelberg, 1995.

[12] Sagan, H., *Introduction to the Calculus of Variations*, Dover Publications, New York, 1969.