Mem. Differential Equations Math. Phys. 19(2000), 142–149

R. Tevzadze

MIXED PROBLEM FOR THE BELLMAN EQUATION WITH MEASURABLE COEFFICIENTS

(Reported on March 15, 1999)

1. Introduction and the Main Results

Consider the Cauchy-Dirichlet mixed problem for Bellman's equation

$$S_t(t,x)+\max_{a\in A}\left[\frac{1}{2}\sigma^2(t,x,a)S_{xx}(t,x)+b(t,x,a)S_x(t,x))\right]=0 \tag{1}$$

$$S(T,x)=g(x),S(t,0)=h_1(t),S(t,l)=h_2(t) \tag{2}$$

under the following conditions on the coefficients b,σ and on the terminal reward functions g,h_1,h_2:

A1) the functions b,σ are measurable and bounded, i.e.,

$$|b(t,x,a)|+|\sigma(t,x,a)|\le C$$

for some $C>0$,

A2) there exists some constant $\lambda>0$ such that

$$\sigma^2(t,x,a)>\lambda$$

for all $t\in[0,T]$, $x\in[0,l]$, $a\in A$,

A3) the functions b,σ are continuous in a for each $t\in[0,T]$, $x\in[0,l]$,

A4) the functions g,h_1,h_2 belong to the Sobolev space W^1 and $g(0)=h_1(T),g(l)=h_2(T)$.

The purpose of this paper is to show the existence of a unique generalized solution of the problem (1),(2).

The novelty (of this paper) is that the question of existence of optimal controls is solved without any regularity assumptions on the coefficients and the use is made of the integral equations. The problem for the full space was studied in [6].

Our method is as follows: For the problem (1),(2) we compose a system of nonlinear integral equations

$$\psi(t,x)=\frac{h_2(t)-h_1(t)}{l}+\int_0^l\rho_x^r(T-t,x,y)g_1(y)dy+$$
$$+\int_t^T\int_0^l\rho_x^r(T-u,x,y)G_1(u,y,\psi(u,y),\hat\psi(u,y))dydu, \tag{3}$$

1991 *Mathematics Subject Classification.* 35K55.

Key words and phrases. Controlled diffusions, Bellman's equation, integral equation.

$$\widetilde{\psi}(t,x) = \int_0^l \rho^r_{xx}(T-t,x,y)g_1(y)dy +$$

$$+ \int_t^T \int_0^l \rho^r_{xx}(u-t,x,y)G_1(u,y,\psi(u,y),\widetilde{\psi}(u,y))dydu, \tag{4}$$

where $\rho(t,x,y) = \frac{2}{l}\sum_{n=1} e^{-\frac{r}{2}(\frac{\pi}{l}n)^2 t}\sin\frac{\pi}{l}nx\sin\frac{\pi}{l}ny, g_1(y) = g(y) - g(0) - \frac{y}{l}(g(l) - g(0))$,

$G(t,x,p,q) = \max_a[\frac{1}{2}(\sigma^2(t,x,a) - r)p + b(t,x,a)q],\ \ G_1 = G - h_1'(t) - \frac{x}{l}(h_2'(t) - h_1'(t))$.

This system can be obtained from the equation

$$S_t(t,x) + \frac{r}{2}S_{xx}(t,x) + G(t,x,S_x(t,x),S_{xx}(t,x)) = 0,$$

equivalent to (1) using the Cauchy formula

$$S(t,x) = h_1(t) + \frac{x}{l}(h_2(t) - h_1(t)) + \int_0^l \rho^r(T-s,x,y)g_1(y)dy$$

$$+ \int_t^T \int_0^l \rho^r(s-t,x,y)G_1(s,y,S_x(s,y),S_{xx}(s,y))dsdy \tag{5}$$

and taking the first and second derivatives in x.

It is well known the that equation (1), (2) is closely connected to a stochastic control problem for a system whose dynamics is discribed by the stochastic differential equation (SDE)

$$dX_t = b(t,X_t,u_t)dt + \sigma(t,X_t,u_t)dW_t,\ \ X_0 = x_0 \in (0,l). \tag{6}$$

Here $(W_t, t \geq 0)$ is a standard Wiener process defined on some complete probability space $(\Omega,\mathcal{F},P)$ and the control $u = (u_t, t \in [0,T])$ is a feedback of the current state, i.e., $u_t = u(t,X_t)$ for some given function $u(t,x)$ taking values in a decision set A which is assumed to be a separable metric space. To each control u we associate one (fixed) solution of SDE (1) (the conditions $A1) - A3)$ imply the existence of a weak solution of SDE (1) ([3])) and the notation $P^u_{t,x}$ is used for the distributon of this solution starting at $X_t = x$, before the exit time τ from the set $(0,l)$. The problem is to maximize the expected cost $E^u\widetilde{g}(T\wedge\tau, X_{T\wedge\tau})$, where $\widetilde{g}$ is a function on $\{(t,0), t\in[0,T]\}\cup\{(T,x), x\in[0,l]\}\cup\{(t,l), t\in[0,T]\}$ and coincides with g, h_1, h_2 on $\{(T,x), x\in[0,l]\}, \{(t,0), t\in[0,T]\}, \{(t,l), t\in[0,T]\}$, respectively, by a suitable choice of feedback controls.

The formal application of Bellman's "dynamic programming" idea leads to the Bellman equation (1), (2) whose solution, if it exists, is easily shown to be the value function

$$S(t,x) = E^u_{t,x}\widetilde{g}(T\wedge\tau, X_{T\wedge\tau}), \tag{7}$$

of the control problem ($E^u_{t,x}$ is the expectation relative to the measure $P^u_{t,x}$). Moreover, if S solves (1),(2) then the optimal control u^* may be constructed by the pointwise maximization of the hamiltonian

$$H(t,x,a) = \frac{1}{2}\sigma^2(t,x,a)S_{xx}(t,x) + b(t,x,a)S_x(t,x). \tag{8}$$

Therefore the main problem consists of finding conditions under which the solution of Bellman's equation exists.

We use the notation $L^p_\beta([0,T]\times[0,l])$, (resp. $L^p_\beta([0,T])$) for the space of p-integrable functions under the weight $e^{-p\beta(T-t)}dtdx$, (resp. $e^{-p\beta(T-t)}dt$) and the notation $W^{1,2}_\beta$ for the Sobolev space with the norm

$$||f||_{W^{1,2}_\beta} = \sup_{(t,x)\in[0,T]\times[0,l]} |f(t,x)| + ||f_t||_{L^2_\beta} + ||f_x||_{L^2_\beta} + ||f_{xx}||_{L^2_\beta}.$$

$W^1[0,l]$ denotes the space with the norm

$$||f||_{W^1} = \Big(\int_0^l (|f(x)|^2 + |f_x(x)|^2)dx\Big)^{\frac{1}{2}}$$

and $W^1_0[0,l]$ denotes its subspace $\{f \in W[0,l], f(0) = f(l)\}$. The following statements are proved in section 3 of this paper.

Theorem 1. *Let the conditions* A1) *and* A4) *be satisfied. Then*

a) *If V is a solution of Bellman's equation* (1), (2) *from the class $W^{1,2}_{\beta_0}$ for some β_0, then the pair (V_x, V_{xx}) of generalized derivatives will be a solution of the system* (3) *for each $\beta, r > 0$.*

b) *If for some $r > 0, \beta$ there exists a pair $(\psi, \widetilde{\psi})$ from the class $L^2_\beta([0,T]\times[0,l]) \times L^2_\beta([0,T]\times[0,l])$ which solves the system* (3), (4) *then the function*

$$V(t,x) = h_1(t) + \frac{x}{l}(h_2(t) - h_1(t)) + \int_0^l g_1(y)\rho^r(T-t,x,y)dy +$$

$$+ \int_t^T \int_0^l G_1(v,y,\psi(v,y),\widetilde{\psi}(v,y))\rho^r(v-t,x,y)dydv \tag{9}$$

will be a solution of the problem (1), (2).

Theorem 2. *Let* A1)-A4) *be satisfied. Then there exists (r^*, β^*) such that for any $r > r^*, \beta > \beta^*$ the operator defined by* (3), (4) *is a contraction. Consequently the system* (3), (4) *has a unique solution which belongs to the class $L^2_\beta([0,T]\times[0,l]) \times L^2_\beta([0,T]\times[0,l])$.*

As a corollary of Theorems 1 and 2 we obtain an existence of a generalized solution of Bellman's equation. Moreover, it is shown in Section 3 that this solution coincides with the value function of the optimal control problem under consideration.

Theorem 3. *The value function* (7) *uniquely solves the problem* (1), (2) *in the class $W^{1,2}_\beta$. If the decision set A is a compact subset of a metric space, then there exists an optimal control in the class of Markovian strategies. The optimal control u^* is constructed from the maximizing of the Hamiltonian* (8)

$$H(t,x,u^*(t,x)) = \max_{a\in A} H(t,x,a)$$

for each $(t,x) \in [0,T]\times[0,l]$.

Moreover, $u^*(t,x)$ can be found from the maximizing of the expression

$$\frac{1}{2}\sigma^2(t,x,a)\widetilde{\psi}(t,x) + b(t,x,a)\psi(t,x),$$

for each $(t,x) \in [0,T]\times[0,l]$, where $\widetilde{\psi}$ and ψ are solutions of equation (3),(4).

The dynamic programming method to prove the existence of an optimal control in the case of diffusion processes for the first time was applied in [5] (Rishel) and [2] (Davis, Varaiya). Theorem 3 was proved in [3] (Krylov) under the Lipschitz condition on the coefficients b and σ.

2. Estimates of the Norms of Some Integral Operators

Consider the operators

$$\psi(t,x)=\sum_{n=1}^{\infty} e_n'(x)\int_t^T e^{-r/2n^2(s-t)}ds\int_0^l e_n(y)dy=$$
$$=\int_t^T\int_0^l \rho_x(s-t,x,y)\varphi(s,y)dsdy \tag{10}$$
$$\widetilde{\psi}(t,x)=\sum_{n=1}^{\infty} e_n''(x)\int_t^T e^{-r/2n^2(s-t)}ds\int_0^l e_n(y)dy=$$
$$=\int_t^T\int_0^l \rho_{xx}(s-t,x,y)\varphi(s,y)dsdy \tag{11}$$

in $L^2_\beta([0,T]\times[0,l])$, where $e_n(x)=\sqrt{2/l}\,\sin\frac{\pi}{l}nx$ is an orthonormal system in $L^2[0,l]$. Using the expansion $\varphi(t,x)=\sum c_n(t)\varphi_n(x), \varphi\in L^2_\beta$ with $\sum|c_n|^2_{L^2_\beta}<\infty$, the system (10),(11) may be rewritten as operators

$$K:\sum_{n=1}^{\infty}e_n(x)c_n(t)\to\sum_{n=1}^{\infty}e_n'(x)\int_t^T c_n(s)e^{-r/2n^2(s-t)}ds, \tag{12}$$

$$\widetilde{K}:\sum_{n=1}^{\infty}e_n(x)c_n(t)\to\sum_{n=1}^{\infty}e_n''(x)\int_t^T c_n(s)e^{-r/2n^2(s-t)}ds. \tag{13}$$

Lemma 1. *The norm of the operator $Q_\lambda: c(t)\to\int_t^T c(s)e^{-\lambda(s-t)}ds$ in the space $L^p_\beta[0,T]$ is estimated by $\frac{1}{c+\lambda}$.*

Proof. At First we consider the case $\beta=0$. We have

$$\Big(\int_0^T dt\Big|\int_t^T e^{-\lambda(s-t)}c(s)ds\Big|^p\Big)^{1/p}\le\Big(\int_0^T dt\Big|\int_0^T e^{-\lambda s}c(s+t)I_{(s<T-t)}ds\Big|^p\Big)^{1/p}.$$

By generalized Hölder inequality [4,134 p.] the second expression is less than

$$\int_0^T e^{-s\lambda}\Big(\int_0^T\big|c(s+t)\big|^pI_{(s<T-t)}dt\Big)^{1/p}ds,$$

which is estimated by $\frac{1}{\lambda}|c|_{L^p_\beta[0,T]}$.

Now consider the case $\beta > 0$. It is clear that

$$\int_0^T e^{-p\beta(T-t)}\Big|\int_t^T e^{-\beta(s-t)}c(s)ds\Big|^p dt = \int_0^T \Big|\int_t^T e^{-\beta(T-s)-\beta(s-t)-\lambda(s-t)}c(s)ds\Big|^p dt =$$

$$= \int_0^T \Big|\int_t^T e^{-(\lambda+\beta)(s-t)}c(s)e^{-(\beta+\lambda)(T-s)}ds\Big|^p dt.$$

By the obtained result for the case $\beta = 0$ we have the estimation

$$\int_0^T \Big|\int_t^T e^{-(\lambda+\beta)(s-t)}c(s)e^{-(\beta+\lambda)(T-s)}ds\Big|^p dt \le$$

$$\le \left(\frac{1}{\beta+\lambda}\right)^p \int_0^T |e^{-\beta(T-s)}c(s)|^p ds = \left(\frac{1}{\beta+\lambda}\right)^p |c|^p_{L^p_\beta}. \quad \square$$

Lemma 2. *The norms of K and $\widetilde{K}$ are estimated by $\frac{1}{\sqrt{\beta r}}$ and $\frac{2}{r}$ respectively.*

Proof. For simplicity we consider the case $a = \pi$. By the Parseval identity we have

$$||\psi||^2_{L^2_\beta} = \sum_{n=1}^{\infty} n^2 \int_0^T e^{-2\beta(T-t)}\Big|\int_t^T\int_0^\pi e_n(y)\varphi(s,y)dye^{-n^2(s-t)r/2}ds\Big|^2 dt.$$

By Lemma 1 we have $||\psi||^2_{L^2_\beta} = \sum_{n=1}^\infty n^2/(\frac{r^2}{4}n^4 + r\beta n^2 + \beta^2) \le \max_{n\ge1}\frac{n^2}{r\beta n^2} = \frac{1}{r\beta}$. Similarly, for the operator $\widetilde{K}$ we have $||\widetilde{K}|| \le \max_{n\ge1}\frac{n^2}{\beta+n^2\frac{r}{2}}$. $\square$

Lemma 3. *A mapping $\psi(t,x) = \int_0^T \rho_x(T-t,x,y)\varphi(y)dy$ is a bounded operator from $L^2[0,l]$ into $L^2_\beta([0,T]\times[0,l])$ and a mapping $\widetilde{\psi}(t,x) = \int_0^T \rho_{xx}(T-t,x,y)\widetilde{\varphi}(y)dy$ is bounded as an operator $W^1[0,l] \to L^2_\beta$.*

Proof. It is clear that $\psi(t,x) = \sum ne_n'(x)e^{-\frac{r}{2}n^2(T-t)}\int_0^l e_n(y)\varphi(y)dy$. We have

$$||\psi||_{L^2_\beta} = \sum_{n\ge1} n^2 c_n^2 \frac{1-e^{-(rn^2+\beta)T}}{n^2r+\beta}, \quad \text{where} \quad c_n = \int_0^l e_n(y)\varphi(y)dy$$

Since a sequence $\{\frac{n^2}{rn^2+\beta}(1-e^{-(rn^2+\beta)})\}_{n\ge1}$ is bounded and $\sum c_n^2 = ||\varphi||^2$, then we have $||\psi||_{L^2_\beta} \le const.||\varphi||$.

The second operator is bounded by the equality

$$||\widetilde{\psi}|| = \Big\|\sum_{n\ge1} e^{-\frac{r}{2}n^2(T-t)-\beta(T-t)}e_n(x)\widetilde{c}_n\Big\|^2 = \sum_{n\ge1} n^4\frac{1-e^{-(2n^2-\beta)T}}{n^2r/2+\beta}$$

if $\sum n^2\widetilde{c}_n^2 < \infty$, which is equivalent to $\widetilde{\varphi} \in W^1_0[0,l]$. $\square$

Proposition 1. *The norm of the operator* $\varphi \to (K\varphi, \widetilde{K}\varphi)$ *from* L^2_β *into* $L^2_\beta \times L^2_\beta$ *is estimated by* $\frac{2}{r} + \frac{1}{\sqrt{r\beta}} = \frac{1}{r}(2 + \sqrt{\frac{r}{\beta}})$.

Proof. It is sufficient to see that $||(K\varphi, \widetilde{K}\varphi)|| = (||K\varphi||^2 + ||\widetilde{K}\varphi||^2)^{1/2} \le (\frac{2}{r} + \frac{1}{\sqrt{r\beta}})||\varphi||$. □

3. Contraction Property of Integral Equations and Proofs of Main Results

Now consider the nonlinear part of the operators (3),(4). The function $G_1(t,x,p,q)$ defines the nonlinear operator

$$(\varphi, \widetilde{\varphi}) \to \widetilde{G}(\varphi, \widetilde{\varphi}) \equiv \{G_1(t,x,\varphi(t,x), \widetilde{\varphi}(t,x))\}_{(t,x)\in[0,T]\times[0,l]}$$

from $L^2_\beta([0,T]\times[0,l])^2$ into $L^2_\beta([0,T]\times[0,l])$.

Lemma 4. *For each* $r > f^* = \max((c+1)^2, \frac{1}{\lambda})$ *the function* G_1 *and the operator* $\widetilde{G}$ *satisfy the Lipschitz condition with the constant* $\frac{1}{2}(r - \frac{1}{r})$.

Proof. See [6]. □

Proposition 2. *The system* (3), (4) *defines the contractive operator in the space* $L^2_\beta([0,T]\times[0,l])^2$ *for some constants* r, β.

Proof. By Lemma 2 and Lemma 4 the Lipschitz constant for the mapping (3.1) is equal to $\frac{1}{2}(r - \frac{1}{r})(\frac{2}{r} + \frac{1}{\sqrt{\beta r}}) = 1 - \frac{1}{r^2} + \frac{1}{2}(\sqrt{\frac{r}{\beta}} - \frac{1}{\sqrt{\beta r^3}})$.

If $\frac{1}{\sqrt{\beta}} < \frac{2}{r^3}$ then $\frac{1}{2}(r - \frac{1}{r})(\frac{2}{r} + \frac{1}{\sqrt{\beta r}}) < 1$. □

Proof of Theorem 1. If the pair $(\psi, \widetilde{\psi})$ belongs to L^2_β, then the function $\widetilde{G}$ also belongs to the same class and therefore the function $V(t,x)$ defined by (9) is a solution of the problem

$$V_t(t,x) + r/2 V_{xx}(t,x) = G(t,x,\psi(t,x), \widetilde{\psi}(t,x)),$$
$$V(T,x) = g(x), V(t,0) = h_1(t), V(t,l) = h_2(t).$$

Since the pair $(\psi, \widetilde{\psi})$ is a solution of (3),(4) taking the first and second derivatives (at x) in (9) we obtain that $V_x = \psi,\ \ V_{xx} = \widetilde{\psi}\ \ dtdx - a.e.$ Therefore, (3) and (4) imply that

$$V_t(t,x) + r/2 V_{xx}(t,x) = G(t,x,V_x(t,x), V_{xx}(t,x)) \tag{14}$$

which gives that the function V satisfies the Bellman equation (1),(2).

Now suppose that there exists a solution of the problem (1),(2), which belongs to the class $W^{1,2}_{\beta_0}$. Let r be a strictly positive constant. Then S is a solution of (12). Clearly $G(t,x,S_x,S_{xx})$ belongs to the class L^2_β. By the Cauchy formula

$$S(t,x) = h_1(t) + \frac{x}{l}(h_2(t) - h_1(t)) + \int_0^l g_1(y)\rho(T-t,x,y)dy +$$
$$\int_t^T \int_0^l G_1(s,y,S_x(s,y),S_{xx}(s,y))\rho(s-t,x,y)dsdy.$$

The differentiation of this equation in x implies that the pair (S_x, S_{xx}) satisfies the system (3),(4). □

As a corollary of this theorem and Proposition 2 we obtain Theorem 2.

Proof of Theorem 3. Let V be a solution of the problem (1),(2) from the class $W_\beta^{1,2}$. Let us show that it coincides with the value function of the optimal control problem. Applying the generalized Itô formula ([2],[1]) for the function V and the controlled process X^u we have

$$V(t\wedge\tau, X^u_{t\wedge\tau}) = V(0,X_0) + \int_0^{t\wedge\tau} V_x(s,X^u_s)\sigma(s,X^u_s,u_s)dW_s + \\ = \int_0^{t\wedge\tau}(L^uV)(s,X^u_s)ds, \qquad (15)$$

where τ is a first exit time of X^u_t from the open set $(0,l)$ and

$$(L^uf)(t,X^u_t) = f_t(t,X^u_t) + b(t,X^u_t,u_t)f_x(t,X^u_t) + \frac{1}{2}\sigma^2(t,X^u_t,u_t)f_{xx}(t,X^u_t).$$

Since the process $V(t,X^u_t)$ is bounded and $E\int_0^{T\wedge\tau}|(L^uV)(s,X^u_s)|ds<\infty$, the stochastic integral in the right-hand side of (15) is a uniformly integrable martingale. On the other hand we have from (1) that $L^uV(s,X^u_s)\le 0$ and taking expectations in (15) we obtain from the boundary condition (2) that

$$V(t\wedge\tau, X^u_{t\wedge\tau}) \ge E^u(V(T\wedge\tau, X^u_{T\wedge\tau})/\mathcal{F}_t) = E^u(\widetilde{g}(T\wedge\tau, X^u_{T\wedge\tau})/\mathcal{F}_t).$$

Therefore,

$$V(t,x)\ge \sup_u E^u_{t,x}(\widetilde{g}(T\wedge\tau X^u_{T\wedge\tau}) = S(t,x). \qquad (16)$$

Let us prove the inverse inequality. Since the function H defined by (8) is continuous in a for each (t,x) and the decision set A is compact, by Philippov's lemma a measurable function $u^* = (u^*(t,x), t\in[0,T], x\in[0,l])$ exists such that

$$H(t,x,u^*(t,x)) = \max_{a\in A} H(t,x,a).$$

Therefore $(L^uV)(s,X^{u^*}_s)=0$ and using again the Itô formula we obtain that

$$V(t,x) = E_{t,x}V(T\wedge\tau, X^{u^*}_{T\wedge\tau}) = E_{t,x}\widetilde{g}(T\wedge\tau, X^{u^*}_{T\wedge\tau}),$$

hence $V(t,x)=S(t,x)$.

Acknowledgement

Research was supported by Georgian Acad. Sci. Grant 1.26.

References

1. R. J. Chitashvili and M. G. Mania, Generalized Ito's formula and derivation of Bellman's equation. *Stochastic Processes and Related Topics, Stochastic Monographs, Gordon & Breach Science Publishers, London,* **10**(1996), 7–27.
2. M. H. A. Davis and P. P. Varaiya, Dynamic programming conditions for partially observable stochastic systems. *SIAM J. Control* **2**(1973), 226–261.
3. N. V. Krylov, Controlled diffusion processes. *Springer-Verlag, New York,* 1980.
4. N. Dunford and T. Schwartz, Linear operators. (Russian) part 1, 1962.
5. R. Rishel, Necessary and sufficient dynamic programming conditions for continuous-time stochastic optimal control. *SIAM J. Control* **8**(1970), 559–571.

6. M. Mania and R. Tevzadze, Solution of Bellman equation by means of a system of nonlinear singular integral equations. *Mem. Differential Equations Math. Phys.* **13**(1998), 121–129.

Author's address:
Institute of Cybernetics
Georgian Academy of Sciences
5, Euli St., Tbilisi 380086
Georgia