SINGULAR STOCHASTIC CONTROL FOR DIFFUSIONS AND SDE WITH DISCONTINUOUS PATHS AND REFLECTING BOUNDARY CONDITIONS

JIN MA*

Abstract. In this paper, we continue to study a diffusion-type, finite-fuel singular stochastic control problem and the related stochastic differential equations with discontinuous paths and reflecting boundary conditions as defined in the previous work of the author [18]. The measurable dependence of the solution with respect to the initial state and the underlying probability measures (with Prohorov metric) is derived. For the application in the control problem, we study more thoroughly the "control" term in the stochastic differential equation which causes the discontinuity of the paths of the solutions. The approximation of certain complete class of controls by the continuous paths ones is proved to be possible in a weak sense. With the help of these results, we prove the Dynamic Programming Principle (Bellman principle) on a rigorous base and that the value function is a viscosity solution of certain Hamilton-Jacobi-Bellman equation.

keywords: discontinuous reflecting problem, S.D.E. with discontinuous paths and reflecting boundary condition, singular stochastic control, Bellman principle, viscosity solution.

1. Introduction. This paper is the continuation of the previous work of the author [18]. We consider the following reflected, diffusion type, finite-fuel singular stochastic control problem. Let $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be a filtered probability space satisfying the *usual conditions* on which is defined an $\{\mathcal{F}_t\}$-Brownian motion $B = \{B_t : t \geq 0\}$. For each *control* process $\xi = \{\xi_t : t \geq 0\}$ which is assumed to be left-continuous, $\mathcal{F}_t$-adapted, with paths of locally bounded variation, the control system is given by

$$X_t = x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t, \qquad t \geq 0, \tag{1.1}$$

where a, σ are real-valued functions satisfying Itô's conditions; $\{K_t : t \geq 0\}$ is the *local-time-like* process which prevents X_t from becoming negative; $x \in [0, \infty)$ is the initial state, which will be assumed to be deterministic throughout the paper; and $T < \infty$ is some fixed time duration. The objective of a decision maker is to find a suitable probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$, together with a Brownian motion B and a control process ξ so as to minimize the cost function:

$$J(P, B, \xi, r, x) = E^P \left\{ \int_0^{T-r} h(t + r, X_t)dt + \int_{[0,T-r)} f(t + r)d\check{\xi}_t + g(X_{T-r}) \right\}, \tag{1.2}$$

where $r \in [0, T]$, and h, f, g are certain smooth functions.

In the sequel, we will call the six-tuple $(\Omega, \mathcal{F}, P, \mathcal{F}_t, B, \xi)$ a *set-up* and denote it by S. Thus the cost function can be denoted as $J(S, r, x)$. Further, the term "finite-fuel" reflects the fact that the control processes is subject to the constraint:

$$P\{\check{\xi}_T \leq y\} = 1, \tag{1.3}$$

* Department of Mathematics, Purdue University, West Lafayette, IN 47907.

for some $y \geq 0$, which stands for the total initial fuel available. If we denote the set of all set-up's satisfying (1.3) by $\mathcal{S}_y$, then the value function is defined by

$$Q(r,x,y) = \inf_{S \in \mathcal{S}_y} J(S,r,x). \tag{1.4}$$

The precise formulation and the notations will be given in §2.

To be consistent with the previous paper, we will call equation (1.1) a *Stochastic Differential Equation with Discontinuous Paths and Reflecting Boundary Conditions* (not to fuss with the name, we shall always use the abbreviation SDEDR as in [18], even when the paths are actually continuous). We note here that the SDEDR in this paper should always be understood as SDEDR(I) defined in [18], which is based on the *Discountinuous Reflecting Problem* (DRP for short) defined by Chaleyat-Maurel et al. [4]. The solution of (1.1) will often be denoted by $(X^{x,\xi}, K^{x,\xi})$ as usual.

In [18] we established some basic properties of SDEDR, especially some comparison results. As applications to the control problem, we also proved the completeness of certain class of admissible controls (see (2.11), (2.12)) and the existence of the optimal control. We shall continue developing the results in both aspects in this paper. Our primary goal is to characterize the value function and, if possible, to describe the optimal control process. One of the possible ways of executing this in the singular control literature is to find a classical solution of the heuristically derived optimality (or H-J-B) equation by using, for instance, the principle of smooth fit; and to prove via a verification theorem that the solution of the H-J-B equation coincides with the value function. In the mean time, it is often possible to discover what a optimal control process should look like. This methodology has worked very well in the case when the system is of linear or constant drift and diffusion coefficients (cf. [2], [9],[12], [14], [21] and the references therein); and some special cases for nonlinear diffusions type systems (cf. [17], [20]). However, in general, finding the classical solution of a second order H-J-B equation or a variational inequality is a very difficult problem, and sometimes impossible.

In this paper, we use another way to investigate the control problem. Namely, we prove rigorously the dynamical programming principle (Bellman Principle) and that the value function is a viscosity solution of certain H-J-B equation. For this purpose, some more thorough studies of the SDEDR are required. For example, we would like to know that the solution of an SDEDR depends at least measurably on those elements that determine the cost function, this will enable us to prove that the complete class of admissible controls that we are interested in is a measurable subset in certain Borel space, so that the Jankov-von Newmann measurable selection theorem can be applied to derive a measurable selector. This motivation leads to §3, §5 and §6 in which we present several measurability results that may be of independent interest. On the other hand, it is often convenient to know whether the value function can be determined by the class of admissible controls with continuous paths. The result of this kind is somewhat traditional in the control theory literature; and in the present context, it will facilitate the proof of uniform continuity of the value function as was done in [11]. The proof of such result basically requires that the solution of the SDEDR depend "continuously" on the the control term in a certain sense; one should note that the issue of this kind is no longer trivial under our setting because the solution mapping of the DRP is at most Lipschitz under the uniform topology (cf. [4], [7], or [18]); therefore, the usual estimate via the Gronwall inequality does not apply (because one will be facing a dilemma to approximate a process which has, in general, discontinuous paths, by continuous processes under the uniform topology). On the other hand, an approximation merely in distribution is not sufficient either, since it does not provide us enough pathwise information to prove the convergence of the cost functions. We shall see in §4, however, that an approximation in a weaker sense but with sufficient pathwise convergence properties is possible, so that the task can be accomplished in a satisfactory way.

With the help of these results, we devote the rest of the paper, §6 and §7, to the proof of the Bellman principle of optimality, and of the fact that the value function is a viscosity solution of a second order variational inequality of the H-J-B type, together with some terminal and boundary properties. The H-J-B variational inequality is derived on a heuristic base in [14] when $a \equiv 0$ and $\sigma \equiv 1$; in a deterministic setting, a similar (but of first order) H-J-B equation was derived recently by Barron, Jenson and Menaldi [1]. We will adapt their ideas to get our result, which is compatible with both cases above and provides a general version for the nonlinear diffusion setting.

2. Definitions, Preliminaries and Formulations. In this paper, we will inherit most of the notations from the previous paper [18] but make some necessary adjustments. First, Recall the following spaces from [18]:

(1) $W \stackrel{\Delta}{=} C[0,\infty)$ is the space of all real-valued continuous functions defined on $[0,\infty)$, with the usual norm;

(2) $D \stackrel{\Delta}{=} \{$all real càglàd functions defined on $[0,\infty)\}$ with the Skorohod topology as was defined in [18], where càglàd means *left continuous with right-limit*;

(3) $A \stackrel{\Delta}{=} \{\xi \in D : \xi$ is nondecreasing$\}$;

(4) $\hat{D} \stackrel{\Delta}{=} \{\xi \in D : \xi$ is of locally bounded variation$\}$;

(5) for $V = W, D, A, \hat{D}...$, we denote $V_0 \stackrel{\Delta}{=} \{v \in V : v(0) = 0\}$.

For each $\xi \in \hat{D}$, we decompose it in a standard way as $\xi_t = \xi_t^+ - \xi_t^-$, $t \geq 0$, where $\xi^\pm \in A$, so that the *total variation* of ξ up to time t can be writen as $\check{\xi}_t = \xi_t^+ + \xi_t^-$. Moreover, for each $0 < \tau < \infty$ and $y > 0$, we define

$$A_\tau(y) \stackrel{\Delta}{=} \{\xi \in A_0 : \xi_{\tau+} \leq y;\ \xi_t = \xi_{\tau+},\ t > \tau\}; \tag{2.1}$$

$$\hat{D}_\tau(y) \stackrel{\Delta}{=} \{\xi \in \hat{D}_0 : \check{\xi}_{\tau+} \leq y;\ \xi_t = \xi_{\tau+},\ t > \tau\}. \tag{2.2}$$

In what follows, when we say "DRP" we mean the DRP(I) defined in [18], *i.e.*, the one defined by Chaleyat-Maurel et al. [4]. If on some probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$, two

D-valued semimartingales Y and $\hat{Y}$ are given with the decompositions $Y = Y_0 + M + A$ and $\hat{Y} = \hat{Y}_0 + \hat{M} + \hat{A}$ respectively, then by an almost identical proof of Proposition 8 in [4] but using Proposition 7 there more prudently, one can easily improve the "Lipschitz continuity" for the solution mapping Γ of DRP (see [4, Proposition 8] or [18, Proposition 3.6]) to the case when $Y_0 \neq \hat{Y}_0$. Namely, if the pair $(\Gamma(\cdot), K(\cdot))$ is such that $\Gamma(Y) = Y + K(Y)$ solves DRP(Y) for any $Y \in D$, then there exists a constant $C > 0$ such that for any stopping time $0 \le \tau < \infty$, one has

$$
\begin{aligned}
(2.3)\qquad & E\left[\sup_{0\le t\le\tau} |\Gamma(Y)_t - \Gamma(\hat{Y})_t|^2\right] + E\left[\sup_{0\le t\le\tau} |K(Y)_t - K(\hat{Y})_t|^2\right] \\
&\le\; CE\left[|Y_0 - \hat{Y}_0|^2 + [M - \hat{M}, M - \hat{M}]_\tau + \left(\int_{[0,\tau)} |d(A - \hat{A})_t|\right)^2\right].
\end{aligned}
$$

In this paper, we also use the notional convention that if $(U_1, \mathcal{B}_{U_1})$, $(U_2, \mathcal{B}_{U_2})$ are two measurable spaces and $f : U_1 \to U_2$ is a mapping, then we say "f is $\mathcal{B}_{U_1}/\mathcal{B}_{U_2}$" if f is $\mathcal{B}_{U_1}/\mathcal{B}_{U_2}$-measurable; we sometimes denote this by "$f \in \mathcal{B}_{U_1}/\mathcal{B}_{U_2}$".

Finally, if U is a metrizable space, then we denote the Borel σ-field on U by $\mathcal{B}_U$ and the totality of the probability measures on the measurable space $(U, \mathcal{B}_U)$ by $\mathcal{P}(U)$. We endow the space $\mathcal{P}(U)$ with the Prohorov metric. It is known that $\mathcal{P}(U)$ is a Polish space if U is so (see, for example, [8]). Moreover, if U is separable, then the σ-field $\mathcal{B}_{\mathcal{P}(U)}$ has the following structure (cf. [3, Proposition 7.25]):

$$
(2.4)\qquad \mathcal{B}_{\mathcal{P}(U)} \;=\; \sigma\left[\bigcup_{E\in\Xi} \theta_E^{-1}(\mathcal{B}_{\mathbf{R}})\right],
$$

where Ξ is any π-system generating $\mathcal{B}_U$ and $\theta_E : \mathcal{P}(U) \mapsto \mathbf{R}$ is defined by $\theta_E(P) = P(E)$, for $E \in \Xi$, $P \in \mathcal{P}(U)$. Also recall that if $U = D$ or W and $U_t = U|_{[0,t]}$, then

$$
(2.5)\qquad \mathcal{B}_{U_t} = \sigma\{\pi_s; 0 \le s \le t\},
$$

where $\pi_t : D \mapsto \mathbf{R}$ is the projection mapping defined by $\pi_t(\eta) = \eta_t$, $t \ge 0$, $\eta \in U$ (cf. [8]). Thus in (2.4) we may take Ξ to be the collection of all the cylinder sets $E^{t_1,\cdots,t_n}_{A_1,\cdots,A_n} \triangleq \cap_{i=1}^n \pi_{t_i}^{-1}(A_i)$, $0 \le t_1 < t_2 < \cdots < t_n < \infty$, $A_i \in \mathcal{B}_{\mathbf{R}}$, $n = 1, 2, \cdots$.

2.1. Set-up, Admissible set-up and Canonical set-up. As was already mentioned in §1, we call a six-tuple $(\Omega, \mathcal{F}, P, \mathcal{F}_t, B, \xi)$ a *set-up* if

(i) $(\Omega, \mathcal{F}, P)$; $\{\mathcal{F}_t\}_{t\ge 0}$ is a filtered probability space satisfying the usual conditions;

(ii) $\{B_t : t \ge 0\}$ is an $\{\mathcal{F}_t\}$-Brownian motion defined on $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$;

(iii) $\{\xi_t : t \ge 0\}$ is an $\{\mathcal{F}_t\}$-adapted processes such that $P(\xi \in \hat{D}_0) = 1$.

The totality of the such set-up's is denoted by $\mathcal{S}_{ad}$. The generic element of $\mathcal{S}_{ad}$ will be denoted by S. For any $\tau > 0$ and $y > 0$, a set-up $S \in \mathcal{S}_{ad}$ is called *(τ, y)-admissible* if $P(\xi \in \hat{D}_\tau(y)) = 1$. We denote the totality of (τ, y)-admissible set-up's by $\mathcal{S}_{ad}(\tau, y)$.

Slightly different from [18], in this paper we will always assume that the initial state $X_0 \equiv x \ge 0$ is deterministic, so that the joint distribution of the triple (X_0, B, ξ)

is completely determined by that of (B,ξ). Thus, the canonical space defined in [18] can be simplified as follows.

We say a set-up $(\Omega,\mathcal{F},P,\mathcal{F}_t,B,\xi)$ is *canonical* if $\Omega=W_0\times D_0$; (B,ξ) is the *canonical process* defined by $(B_t,\xi_t)(\omega)=(w(t),\zeta(t)),\ t\ge 0,\ \omega=(w,\zeta)\in\Omega$; and if $\{\mathcal{F}_t^0\}$ is the smallest filtration which measures (B,ξ), $\mathcal{F}^0=\mathcal{F}^0_\infty$, then $\mathcal{F}$, $\{\mathcal{F}_t\}$ are the usual P-augmentation of $\mathcal{F}^0$, $\{\mathcal{F}_t^0\}$ respectively. The totality of all canonical set-up's will be denoted by $\mathcal{S}_{can}$.

It is easily seen that a canonical set-up $S=(\Omega,\mathcal{F},P,\mathcal{F}_t,B,\xi)$ is determined completely by the probability P, thus we may write $S=S(P)$ and often say that $P\in\mathcal{S}_{can}$ if $S(P)\in\mathcal{S}_{can}$ when there is no danger of confusion. Therefore, we might as well view $\mathcal{S}_{can}$ as a set of probability measures on the canonical space $(\Omega,\mathcal{F})$, which actually corresponds to the set $\mathcal{M}$ defined in [18]; more precisely, $\mathcal{M}\stackrel{\Delta}{=}\{P: S(P)\in\mathcal{S}_{can}\}$. We can also define $\mathcal{M}(\tau,y)$ to be the set of all (τ,y)-admissible canonical set-up's, and define $\mathcal{M}^-(\tau,y)\subset\mathcal{M}(\tau,y)$ to be all the elements in $\mathcal{M}(\tau,y)$ such that $P\{\xi^+\equiv 0\}=1$. If $P\in\mathcal{M}^-(\tau,y)$, the second component of the canonical process is often denoted by $-\xi$, where $\xi\in A_\tau(y)$.

It is worth noting that, since $\mathcal{P}(\Omega)$ is obviously a Polish space, and $\hat{D}_0$ is a Borel subset of D (cf. [18]), it follows that $\mathcal{M}$ is a Borel set in $\mathcal{P}(\Omega)$ and hence a standard space (cf. [10]). In particular, $\mathcal{M}$ is a seperable metric space.

2.2. Problem Formulation. Consider the control system described in §1 on a given set-up:

$$X_t=X_0+\int_0^t a(X_s)ds+\int_0^t\sigma(X_s)dB_s+\xi_t+K_t,\qquad t\ge 0; \tag{2.6}$$

namely, (X,K) is the solution to SDEDR(x,ξ). We will make use of the following basic assumptions throughout.

(C.1) The functions a and σ are bounded, continuous on $\mathbf{R}$, such that

$$\inf_{x\in\mathbf{R}}\sigma(x)\stackrel{\Delta}{=}\sigma_0>0; \tag{2.7}$$

(C.2) There exists a constant $C_1>0$ such that

$$|a(x)-a(y)|+|\sigma(x)-\sigma(y)|\le C_1|x-y|,\qquad \text{for all } x,y\in\mathbf{R}. \tag{2.8}$$

Let $T>0$ be fixed, and denote in the sequel $E=[0,T]\times[0,\infty)\times(0,\infty)$. Recall that, for each $(r,x,y)\in E$ and $S\in\mathcal{S}_{ad}(r,y)$, the cost function is defined by

$$J(S,r,x)=E^P\left\{\int_0^{T-r}h(t+r,X_t^{x,\xi})dt+\int_{[0,T-r)}f(t+r)d\check{\xi}_t+g(X_{T-r}^{x,\xi})\right\}. \tag{2.9}$$

We pose the following conditions on the functions h,f,g appeared in (2.9).

(C.3) The functions $h:[0,T]\times[0,\infty)\to[0,\infty)$, $f:[0,T]\to[0,\infty)$ and $g:[0,\infty)\to[0,\infty)$ are all continuously differentiable, such that

(i) $h(t,\cdot)$ is nondecreasing and $h_x(t,0) \geq 0$ for all $t \geq 0$;
(ii) g is convex, nondeceasing;
(iii) there exist $C > 0$, $m \geq 1$ such that for any $(t,x) \in [0,T] \times [0,\infty)$,

$$0 \leq h_x(t,x) + |h_t(t,x)| + g'(x) \leq C(1 + x^m);$$

(iv) $\sup_{x\geq 0} g'(x) \leq \inf_{0\leq t\leq T} f(t)$. □

From now on, we will denote all the constants depending only on T, f, g, h, a, σ by a generic one C, which may vary line by line.

It is fairly easy to prove that, for given $(r,x,y) \in E$,

$$Q(r,x,y) = \inf_{S\in\mathcal{S}_{ad}(T-r,y)} J(S,r,x) = \inf_{P\in\mathcal{M}(T-r,y)} J(P,r,x). \tag{2.10}$$

Next let $(r,x,y) \in E$ and $P \in \mathcal{M}(T-r,y)$ be given, by the argument analogous to [18], we can find a pair of functions $(F_P^x, G_P^x) : \Omega \to D \times D$ such that $X = F_P^x,\ K = G_P^x$ solves SDEDR(x,ξ) (1.1) on the canonical space. We define a subclass $\mathcal{D}_{can}(r,x,y)$ of $\mathcal{M}(T-r,y)$ by

$$\begin{aligned}(2.11)\ \mathcal{D}_{can}(r,x,y) &\triangleq \{P \in \mathcal{M}^-(T-r,y) : G_P^x \text{ is continuous},\ a.s.P\}\\ &= \{P \in \mathcal{M}^-(T-r,y) : 0 \leq \xi_t \leq (F_P^x)_t,\ 0 \leq t \leq T,\ a.s.P\}.\end{aligned}$$

It was shown in [11], [13] and [18, Proposition 6.4] that for each $(r,x,y) \in E$, there exists a $P^* \in \mathcal{D}_{can}(r,x,y)$, such that

$$J(P^*,r,x) = Q(r,x,y) = \inf_{P\in\mathcal{M}(T-r,y)} J(P,r,x). \tag{2.12}$$

In other words, the subclass $\mathcal{D}_{can}(r,x,y)$ is "complete".

2.3. Localization. In this paper, We shall restrict all the processes to a finite time interval $[0,T]$, the time duration for our control problem. The main advantage of such a restriction occurs in §4 when we deal with the Girsanov-Cameron-Martin transformation. But on the other hand, it also causes some other technical difficulties when we apply the "path-shifting" mathod, for instance, in §6. Therefore it is useful to introduce the notion of *localization*, which we now describe.

We begin with an arbitrary set-up $(\Omega, \mathcal{F}, P, \mathcal{F}_t, B, \xi)$. Let $T > 0$ be given. By "localizing" the set-up we simply mean to restrict the processes B and ξ to the finite interval $[0,T]$. Let us denote the class of "local" set-up's by $\mathcal{S}_{ad}^{loc}$. It is clear that for any $S \in \mathcal{S}_{ad}$, there is a local set-up $S^{loc} \in \mathcal{S}_{ad}^{loc}$ such that $J(S,r,x) = J(S^{loc},r,x)$. Thus we may "imbed" $\mathcal{S}_{ad}$ into $\mathcal{S}_{ad}^{loc}$ in a obvious way. If we define, for given $(r,x,y) \in E$,

$$Q^{loc}(r,x,y) = \inf_{S\in\mathcal{S}_{ad}^{loc}(T-r,y)} J(S,r,x); \qquad Q(r,x,y) = \inf_{S\in\mathcal{S}_{ad}(T-r,y)} J(S,r,x),$$

then it is evident that

$$Q^{loc}(r,x,y) \leq Q(r,x,y). \tag{2.13}$$

We wish to prove that *equality* actually prevails in (2.13).

Let us consider the canonical set-up's. Denote $\Omega_T = C[0,T] \times D_T$ where $D_T \triangleq D|_{[0,T]}$; let $\mathcal{G}$ and $\mathcal{G}_t$ be the smallest σ-field and filtration respectively measuring the canonical process. By an analogy with $\mathcal{S}^{loc}_{ad}(T-r,y)$, we can define $\mathcal{M}^{loc}(T-r,y)$ (similarly, $\mathcal{D}^{loc}_{can}(r,x,y)$). We have the following proposition.

Proposition 2.1. *For any $x \geq 0$ and $P \in \mathcal{M}^{loc}(T-r,y)$, there exists a $\tilde{P} \in \mathcal{M}(T-r,y)$, such that $\tilde{P}^{loc} \triangleq \tilde{P}|_{\mathcal{G}_T} = P$ and $J(\tilde{P},r,x) = J(P,r,x)$. Consequently, we have*

$$Q(r,x,y) = Q^{loc}(r,x,y). \tag{2.14}$$

Proof. Let $P \in \mathcal{M}^{loc}(T-r,y)$ be given. Define a probability measure Q on $(\Omega,\mathcal{F})$ by $Q = P^W \times \delta_{\{0\}}$, where P^W is the Wiener measure on W_0 and $\delta_{\{0\}}$ is the point mass at the zero-function in D_0. Consider the new space $\hat{\Omega} = \Omega_T \times \Omega$ with the corresponding product σ-field $\hat{\mathcal{F}}$. Define the probability measure on $(\hat{\Omega},\hat{\mathcal{F}})$ by

$$\hat{P}(A \times B) = P(A)Q(B), \qquad A \in \Omega_T;\ B \in \Omega.$$

Let the generic element in $\hat{\Omega}$ be $\hat{\omega} = (\omega^1,\omega^2)$ where $\omega^1 = (w^1,\zeta^1) \in \Omega_T$ and $\omega^2 = (w^2,\zeta^2) \in \Omega$. Define a process on $(\hat{\Omega},\hat{\mathcal{F}})$ by $(\hat{B},\hat{\xi})$, where

$$\hat{B}_t(\hat{\omega}) = \left\{ \begin{array}{ll} w^1_t & 0 \leq t \leq T \\ & \\ w^2_t - w^2_T + w^1_T & t > T \end{array} \right\}; \quad \hat{\xi}_t(\hat{\omega}) = \left\{ \begin{array}{ll} \zeta^1_t & 0 \leq t \leq T \\ & \\ \zeta^1_T & t > T \end{array} \right\}.$$

Let $\{\hat{\mathcal{F}}_t\}$ be the smallest filtration measuring $(\hat{B},\hat{\xi})$, satisfying the usual conditions, then it is easy to check that $\hat{B}$ is an $(\hat{\mathcal{F}}_t,\hat{P})$-Brownian motion and $\hat{\xi}_t = \hat{\xi}_T = \hat{\xi}_{T-r},\ t > T,\ a.s.\hat{P}$. Furthermore, if we let $\tilde{P} = \hat{P} \circ (\hat{B},\hat{\xi})^{-1}$, then we have $\tilde{P} \in \mathcal{M}(T-r,y)$ and $\tilde{P}|_{\mathcal{G}_T} = \hat{P} \circ ((\hat{B},\hat{\xi})|_{[0,T]})^{-1} = P$, proving the first assertion. That (2.14) follows from the first assertion and the facts (2.10) and (2.13) is obvious, the proof is complete. □

3. The measurable dependence of the solution to SDEDR. Recall from [18] that for any given $(r,x,y) \in E = [0,T] \times [0,\infty) \times (0,\infty)$ and $P \in \mathcal{S}_{can} = \mathcal{M}$, there exists a pair of progressively measurable functions (F^x_P, G^x_P) from $\Omega (\triangleq W_0 \times D_0)$ to $D \times D$ which solves the SDEDR (1.1). For the same reason, there also exists a version $Y^x_P \in \mathcal{F}_\infty / \mathcal{B}_D$ of the solution Y^x to the *unrestricted* equation:

$$Y^x_t = x + \int_0^t a(X^x_s)ds + \int_0^t \sigma(X^x_s)dB_s + \xi_t, \quad t \geq 0. \tag{3.1}$$

We would like to investigate in this section the dependence of the solutions F^x_P, G^x_P, Y^x_P on the parameters x and P. Moreover, if we define three mappings $\Phi,\ \Psi,\ \Theta : [0,\infty) \times \mathcal{M} \to \mathcal{P}(D)$ by

$$\Phi(x,P) \triangleq P \circ (F^x_P)^{-1}; \quad \Psi(x,P) \triangleq P \circ (G^x_P)^{-1}, \quad \Theta(x,P) \triangleq P \circ (Y^x_P)^{-1}, \tag{3.2}$$

then we wish to establish the following theorem.

Theorem 3.1. *The functions* Φ, Ψ, Θ *are* $\mathcal{B}_{[0,\infty)} \times \mathcal{B}_{\mathcal{M}}/\mathcal{B}_{\mathcal{P}(D)}$.

We shall split the proof of the Theorem 3.1 into two main propositions.

Proposition 3.2. *For fixed* $P \in \mathcal{M}$*, the mappings* $\Phi(\cdot, P)$, $\Psi(\cdot, P)$, $\Theta(\cdot, P)$: $[0, \infty) \mapsto \mathcal{P}(D)$ *are continuous.*

Proof. Since convergence in probability implies convergence in distribution and in the space D, the uniform topology is stronger than the Skorohod topology, we need only show that for fixed $P \in \mathcal{M}$ and $t \geq 0$, with Z^x denoting F^x_P, G^x_P, Y^x_P respectively, we have $E^P|Z^x - Z^{x'}|^{*,2}_t \to 0$, whenever $x \to x'$; where for $t \in [0, \infty)$, $|Z|^*_t \triangleq \sup_{0\leq s\leq t}|Z_s|$ and $|Z|^{*,2}_t \triangleq (|Z|^*_t)^2$.

To this end, let $P \in \mathcal{M}$ be fixed; define $X^x = F^x_P$, $K^x = G^x_P$ and $Y^x = Y^x_P$, then it is known (cf. [18]) that

$$X^x = \Gamma(Y^x); \quad K^x = X^x - Y^x. \tag{3.3}$$

where Γ is the solution mapping of DRP. Further, by (2.3) and Doob's inequality, we have for any $t \geq 0$,

$$\begin{aligned} E^P|X^x - X^{x'}|^{*,2}_t &= E^P|\Gamma(Y^x) - \Gamma(Y^{x'})|^{*,2}_t \\ \leq \ C_T &\left\{ |x - x'|^2 + E^P\left[\int_0^T [|a(X^x_s) - a(X^{x'}_s)|^2 + |\sigma(X^x_s) - \sigma(X^{x'})|^2]ds\right]\right\}. \end{aligned} \tag{3.4}$$

Thus, the conditions on the coefficients a and σ and the Gronwall inequality yield that

$$E^P\left\{|X^x - X^{x'}|^{*,2}_T\right\} \leq C|x - x'|^2. \tag{3.5}$$

It is readily seen by (3.1) and (3.3) that (3.5) is also true for Y and K. The lemma is proved. □

It is now easy to see that the Theorem 3.1 will follow from our second proposition:

Proposition 3.3. *For fixed* $x \geq 0$*, the mapplings* $\Phi(x, \cdot), \Psi(x, \cdot)$ *and* $\Theta(x, \cdot)$ *are all* $\mathcal{B}_{\mathcal{M}}/\mathcal{B}_{\mathcal{P}(D)}$*-measurable.*

The proof of Proposition 3.3 is more involved, and therefore we shall split it into several lemmas. Our first step is to prove that for fixed $x \geq 0$, the mappings $(P, \omega) \mapsto F^x_P(\omega)$ (resp. $G^x_P(\omega)$, $Y^x_P(\omega)$) is "jointly" measurable in certain sense; and then the Proposition 3.3 shall follow easily. Since now the probability P is itself a "variable", we need a new device to handle those subsets $A \subset \mathcal{M} \times \Omega$ such that for each $P \in \mathcal{M}$, the "P-section" of A is a P-null set. More precisely, we shall introduce a notion which we call the "$\mathcal{M}$-augmentation" of the σ-fields on $\mathcal{M} \times \Omega$ in the sequel. Let $\{\mathcal{F}^\circ_t\}$ be the smallest filtration on the canonical space Ω that measures the canonical process and we may modify it to be right-continuous. Define for any set $A \subset \mathcal{M} \times \Omega$ and $P \in \mathcal{M}$ the P-section of A by $A^P = \{\omega : (P, \omega) \in A\}$; a set $A \subset \mathcal{M} \times \Omega$ is called an $\mathcal{M}$-null set if for any $P \in \mathcal{M}$, $P(A^P) = 0$. Let us denote

$$\mathcal{N} = \{F \subseteq \mathcal{M} \times \Omega : \exists G \in \mathcal{B}_{\mathcal{M}} \times \mathcal{F}^\circ_\infty \text{ such that } G \text{ is an } \mathcal{M}\text{-null set, } F \subseteq G\}. \tag{3.6}$$

Then the $\mathcal{M}$-augmentation of any sub-σ-algebra $\mathcal{G} \subseteq \mathcal{B}_{\mathcal{M}} \times \mathcal{F}^{\circ}_{\infty}$ is defined by

$$\overline{\mathcal{G}}^{\mathcal{M}} \triangleq \sigma(\mathcal{G} \cup \mathcal{N}).$$

It is then easily shown as usual (see, for example, [15, problem 2.7.3]) that

$$\begin{aligned} \overline{\mathcal{G}}^{\mathcal{M}} &= \{F \subseteq \mathcal{M} \times \Omega : \exists G \in \mathcal{G} \text{ such that } F \Delta G \in \mathcal{N}\} \\ &= \{F \subseteq \mathcal{M} \times \Omega : \exists G \in \mathcal{G}, N \in \mathcal{N} \text{ such that } F = G \Delta N\}, \end{aligned} \tag{3.7}$$

where $A \Delta B \triangleq (A \setminus B) \cup (B \setminus A)$. We will be interested in the following σ-fields:

$$\mathcal{Y}_t \triangleq \overline{\mathcal{B}_{\mathcal{M}} \times \mathcal{F}^{\circ}_t}^{\mathcal{M}}, \quad t \geq 0; \qquad \mathcal{Y}_{\infty} \triangleq \overline{\mathcal{B}_{\mathcal{M}} \times \mathcal{F}^{\circ}_{\infty}}^{\mathcal{M}}. \tag{3.8}$$

Now denote by $\mathcal{H}$ the class of all real-valued functions $f_t(P, \omega)$ defined on $[0, \infty) \times \mathcal{M} \times \Omega$. In $\mathcal{H}$ we say that f and g are $\mathcal{M}$-*equivalent* if the set

$$\{(P, \omega) : \exists t \geq 0, \; f_t(P, \omega) \neq g_t(P, \omega)\}$$

is an $\mathcal{M}$-null set; and that f and g are $\mathcal{M}$-*versions* to each other if they are $\mathcal{M}$-equivalent. We will consider the ($\mathcal{M}$-)equivalent classes in $\mathcal{H}$. Denote

$$\mathcal{H}_{meas} = \{f \in \mathcal{H} : f_t \in \mathcal{Y}_t / \mathcal{B}_{\mathbf{R}}, \text{ for all } t \geq 0\}; \tag{3.9}$$

$$\mathcal{H}^D_{meas} = \{f \in \mathcal{H}_{meas} : \{(P, \omega) : f(P, \omega) \notin D\} \in \mathcal{N}\}. \tag{3.10}$$

The following lemma gives the main properties of the spaces $\mathcal{H}_{meas}$ and $\mathcal{H}^D_{meas}$.

Lemma 3.4. *(1) Both $\mathcal{H}_{meas}$ and $\mathcal{H}^D_{meas}$ are algebras;*

(2) For any $f \in \mathcal{H}^D_{meas}$, there exists an $\mathcal{M}$-version $\hat{f}$ of f such that $\hat{f}(P, \omega) \in D$ for all $(P, \omega) \in \mathcal{M} \times \Omega$ and that $\hat{f}$ is $\mathcal{Y}_t / \mathcal{B}_{D_t}$, for all $t \geq 0$;

(3) Let $\{F^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{H}_{meas}$ (resp. $\mathcal{H}^D_{meas}$) and $F \in \mathcal{H}$. Suppose that for any $P \in \mathcal{M}$ there exists a set N_P such that $P(N_P) = 0$ and that for all $\omega \notin N_P$,

$$\lim_{n \to \infty} |F^{(n)}_t(P, \omega) - F_t(P, \omega)| = 0, \quad t \geq 0 \tag{3.11}$$

(respectively,

$$\lim_{n \to \infty} |F^{(n)}(P, \omega) - F(P, \omega)|^*_t = 0, \quad t \geq 0), \tag{3.12}$$

then $F \in \mathcal{H}_{meas}$ (resp. $\mathcal{H}^D_{meas}$).

Proof. (1) is obvious. To see (2), first note that for any $f \in \mathcal{H}^D_{meas}$, the complement of the set $\mathsf{A}_f \triangleq \{(P, \omega) : f(P, \omega) \in D\}$ is in $\mathcal{N}$, whence $\mathsf{A}_f \in \mathcal{Y}_t$, for all $t \geq 0$. Thus if we define

$$\hat{f}_t(P, \omega) = 1_{\mathsf{A}_f}(P, \omega) f_t(P, \omega), \quad (t, P, \omega) \in [0, \infty) \times \mathcal{M} \times \Omega,$$

then $\hat{f}$ is an $\mathcal{M}$-version of f, an element in $\mathcal{H}^D_{meas}$; and $\hat{f}(P, \omega) \in D$ for all $(P, \omega) \in \mathcal{M} \times \Omega$. Therefore, we can view $\hat{f}$ as a mapping from $\mathcal{M} \times \Omega$ to D and it is easily checked by the definition of $\mathcal{B}_{D_t}$ (recall (2.5)) that $\hat{f}$ is $\mathcal{Y}_t / \mathcal{B}_{D_t}$ for all $t \geq 0$.

We now prove (3). Since the "$\mathcal{H}_{meas}$"-case follows from an easy analogue of the "$\mathcal{H}^D_{meas}$"-case, we only prove the latter. By definition, the complement of the set

$$\mathsf{A}_\infty \triangleq \{(P,\omega) : \lim_{n,m\to\infty} |F^{(n)}(P,\omega) - F^{(m)}(P,\omega)|^*_t = 0,\ t \geq 0\}$$

is in $\mathcal{Y}_\infty$ and is $\mathcal{M}$-null, so it is easy to construct, by using (3.8), some other $\mathcal{M}$-null set $G \in \mathcal{B}_\mathcal{M} \times \mathcal{F}^\circ_\infty$, such that $\mathsf{A}^c_\infty \subseteq G$, whence $\mathsf{A}^c_\infty \in \mathcal{N}$ and $\mathsf{A}_\infty \in \mathcal{Y}_t$ for all $t \geq 0$. Therefore, by defining

$$\bar{F}^{(n)}_t(P,\omega) = 1_{\mathsf{A}_\infty}(P,\omega)\cdot F^{(n)}_t(P,\omega), \tag{3.13}$$

$$\bar{F}_t(P,\omega) = 1_{\mathsf{A}_\infty}(P,\omega)\cdot F_t(P,\omega), \tag{3.14}$$

we have $\lim_{n\to\infty} \bar{F}^{(n)}_t(P,\omega) = \bar{F}_t(P,\omega)$ for all $(P,\omega) \in \mathcal{M}\times\Omega$, uniformly in t on compacts, hence $\bar{F}(P,\omega) \in D$ for all $(P,\omega) \in \mathcal{M}\times\Omega$ and is $\mathcal{Y}_t/\mathcal{B}_{D_t}$ for all $t \geq 0$. Finally, since

$$\{(P,\omega) : \bar{F}(P,\omega) \neq F(P,\omega)\} \subseteq \mathsf{A}^c_\infty \in \mathcal{N},$$

we see that $\bar{F} \in \mathcal{H}^D_{meas}$ is an $\mathcal{M}$-version of F, the proof is complete. □

We now denote by $\mathcal{H}^C_{meas}$ the subspace of $\mathcal{H}^D_{meas}$ satisfying the definition (3.10) with D being replaced by W.

Lemma 3.5. *Suppose $f \in \mathcal{H}_{meas}$. Let a, σ be the functions satisfying (C.1), (C.2) in §2. Define for $(t,P,\omega) \in [0,\infty)\times\mathcal{M}\times\Omega$,*

$$[I^1(f)]_t(P,\omega) = \int_0^t a(f_s(P,\omega))ds; \tag{3.15}$$

$$[I^2(f)]_t(P,\omega) = \left[\int_0^t \sigma(f_s(P,\cdot))dB_s(\cdot)\right](\omega). \tag{3.16}$$

Then $I^1(f), I^2(f) \in \mathcal{H}^C_{meas}$.

Proof. (i) Set, for each n and $(P,\omega) \in \mathcal{M}\times\Omega$,

$$[I^{1;n}(f)]_t(P,\omega) = \sum_{i=0}^\infty a(f_{t\wedge(i/2^n)}(P,\omega))[t\wedge((i+1)/2^n) - t\wedge(i/2^n)],$$

then it is clear that $I^{1;n}(f) \in \mathcal{H}_{meas}$, $n = 1,2,\cdots$ by Lemma 3.4-(1). Since for each $P \in \mathcal{M}$,

$$\lim_{n\to\infty} |I^{1;n}(f)(P,\omega) - I^1(f)(P,\omega)|^{*,2}_t = 0, \quad t \geq 0, \text{ for } P-a.e.\ \omega \in \Omega,$$

$I^1(f) \in \mathcal{H}_{meas}$ by Lemma 3.4-(3). That I^1 has continuous paths is clear, so $I^1(f) \in \mathcal{H}^C_{meas}$.

(ii) To prove the result for I^2, we first fix $P \in \mathcal{M}$ and observe that the first coordinate of the canonical process $B_t(\omega) = w(t),\ t \geq 0,\ \omega = (w,\eta) \in \Omega$ is an $(\mathcal{F}_t, P)$-Brownian motion. Define a mapping $\varphi : [0,\infty) \to [0,\infty)$ by

$$\varphi(x) = \int_0^x \frac{du}{\sigma(u)}. \tag{3.17}$$

By the condition on the function σ, φ is C^2, strictly increasing and non-negative. We now consider an element $F \in \mathcal{H}$ defined by

$$F_t(P,\omega) \triangleq w(t) + \frac{1}{2}\int_0^t \varphi''(f_s(P,\omega))ds, \tag{3.18}$$

where $\{w(t) : t \geq 0\}$ is the first component of the canonical process, which is obviously an element of $\mathcal{H}^C_{meas}$. Furthermore, the same argument as (i) shows that the integral in (3.18) is in $\mathcal{H}^C_{meas}$, whence $F \in \mathcal{H}^C_{meas}$.

It is now clear by applying Itô's Formula (for each fixed P) that

$$\varphi(I^2(f)_t(P,\omega)) = F_t(P,\omega), \quad \text{for } P-a.e.\ \omega \in \Omega,$$

thus $\varphi \circ I^2(f)$ is a $\mathcal{M}$-version of F, and therefore in $\mathcal{H}_{meas}$. Noting that φ is actually a homeomorphism, we get that $I^2 \in \mathcal{H}_{meas}$.

Finally, it is clear that $I^2(f)$ also has continuous paths, the proof is complete. □

Now let $f \in \mathcal{H}_{meas}$. By Lemma 3.5-(2), we may assume that f is $\mathcal{Y}_\infty/\mathcal{B}_D$. It is then well known (by using the property (3.7) and the Polishness of D) that there exists a mapping $\hat{f}$ from $\mathcal{M}$ to D which is $\mathcal{B}_\mathcal{M} \times \mathcal{F}^\circ_\infty/\mathcal{B}_D$ and differs from f only on a set $N \in \mathcal{N}$, *i.e.*, $\hat{f}$ is an $\mathcal{M}$-version of f. Furthermore, it is evident that for each $P \in \mathcal{M}$, we have

$$P \circ f(P,\cdot)^{-1} = P \circ \hat{f}(P,\cdot)^{-1} \in \mathcal{P}(D).$$

Namely, the mapping $P \mapsto P \circ \hat{f}(P,\cdot)^{-1}$ is independent of the choice of such $\mathcal{M}$-versions of f. Hence the function

$$F(P) \triangleq P \circ f(P,\cdot)^{-1} \in \mathcal{P}(D) \tag{3.19}$$

is well-defined; and we may always assume that the corresponding f is $\mathcal{B}_\mathcal{M} \times \mathcal{F}^\circ_\infty/\mathcal{B}_D$ when we mention the function F defined by (3.19) in the sequel. We have the following lemma.

Lemma 3.6. *Suppose $f \in \mathcal{H}^D_{meas}$. Then the mapping $F : \mathcal{M} \mapsto \mathcal{P}(D)$ is $\mathcal{B}_\mathcal{M}/\mathcal{B}_{\mathcal{P}(D)}$.*

Proof. By the convention preceding the lemma, f is $\mathcal{B}_\mathcal{M} \times \mathcal{F}^\circ_\infty/\mathcal{B}_D$. So by (2.4), it suffices to show that for every $\hat{A} = E^{t_1,\cdots,t_n}_{A_1,\cdots,A_n} \in \mathcal{B}_D$ and $B \in \mathcal{B}_\mathbf{R}$, the set $F^{-1}(\hat{A}) = \{P : P \circ f(P,\cdot)^{-1}(\hat{A}) \in B\}$ is in $\mathcal{B}_\mathcal{M}$. We claim that this can be reduced to the following simpler assertion: for any bounded function $H \in \mathcal{B}_\mathcal{M} \times \mathcal{F}^\circ_\infty/\mathcal{B}_\mathbf{R}$, if $L_H(P) \triangleq \int_\Omega H(P,\omega)P(d\omega)$, then L_H is $\mathcal{B}_\mathcal{M}/\mathcal{B}_\mathbf{R}$.

To substantiate the claim, note that

$$\begin{aligned}\hat{L}(P) &\triangleq P \circ f(P,\cdot)^{-1}(\hat{A}) = \int_\Omega 1_{\hat{A}}(f(P,\omega))P(d\omega)\\ &= \int_\Omega \Pi_{i=1}^n 1_{A_i}(f_{t_i}(P,\omega))P(d\omega),\end{aligned}$$

so if we let $H \triangleq \Pi_{i=1}^n (1_{A_i} \circ \pi_{t_i} \circ f)$ which is obviously $\mathcal{B}_{\mathcal{M}} \times \mathcal{F}_\infty^\circ / \mathcal{B}_{\mathbf{R}}$, then $\hat{L} = L_H$ and the lemma follows from the assertion.

We now prove the assertion. By a standard Monotone-Class argument, it suffices to prove the assertion for $H = 1_{A\times B}$ where $A \in \mathcal{B}_{\mathcal{M}}$, $B \in \mathcal{F}_\infty^\circ$. But in this case, $L_H(P) = 1_A(P) \cdot P(B) = 1_A(P) \cdot \theta_B(P)$, where θ is the same as that was defined in (2.4). So replacing D by Ω in (2.4), we see that L_H is $\mathcal{B}_{\mathcal{M}}/\mathcal{B}_{\mathbf{R}}$, proving the lemma. □

Proof of Proposition 3.3:

Let $x \geq 0$ be fixed. Define $\mathsf{Y}^x(P,\omega) = Y_P^x(\omega)$; $\mathsf{X}^x(P,\omega) = X_P^x(\omega)$; $\mathsf{K}^x(P,\omega) = K_P^x(\omega)$; thus $\Theta(x,P) = P \circ [\mathsf{Y}^x(P,\cdot)]^{-1}$; $\Phi(x,P) = P \circ [\mathsf{X}^x(P,\cdot)]^{-1}$; $\Psi(x,P) = P \circ [\mathsf{K}^x(P,\cdot)]^{-1}$. So by Lemma 3.6, it suffices to show that $\mathsf{Y}^x, \mathsf{X}^x, \mathsf{K}^x \in \mathcal{H}_{meas}^D$. To this end, denote $I_t^0(P,\omega) = x - \xi_t(\omega) = x - \zeta(t)$, $t \geq 0$, where ξ is the second coordinate of the canonical process, namely, $\xi_t(\omega) = \zeta(t)$ for $\omega = (w,\zeta) \in \Omega$. It is evident that $I^0 \in \mathcal{H}_{meas}^D$.

Let $\Gamma : D \to D$ be the solution mapping of DRP. It can be checked from the construction of the functions $(\Gamma(Y), K)$, $Y \in D$ (cf. [4, Proposition 2, Theorem 5]) that Γ is $\mathcal{B}_{D_t}/\mathcal{B}_{D_t}$, $t \geq 0$. Therefore, if $f \in \mathcal{H}_{meas}^D$, then by Lemma 3.4-(2), we can choose an $\mathcal{M}$-version $\hat{f}$ of f such that $\hat{f}(P,\omega) \in D$ for all $(P,\omega) \in \mathcal{M}\times\Omega$ and is $\mathcal{Y}_t/\mathcal{B}_{D_t}$ for all $t \geq 0$, thus so is the composition $\Gamma(\hat{f})$. Since $\Gamma(\hat{f})$ is a $\mathcal{M}$-version of $\Gamma(f)$, we have $\Gamma(f) \in \mathcal{H}_{meas}^D$.

We now recall from Lemma 3.5 the functions $I^1(\cdot)$, $I^2(\cdot)$ and define for $n = 0, 1, 2, \cdots$, $\mathsf{X}^{(n)} = \Gamma(\mathsf{Y}^{(n)})$; $\mathsf{K}^{(n)} = \mathsf{X}^{(n)} - \mathsf{Y}^{(n)}$, where

$$\mathsf{Y}^{(0)} = I^0; \quad \mathsf{Y}^{(n)} = I^0 + I^1(\mathsf{X}^{(n-1)}) + I^2(\mathsf{X}^{(n-1)}), \quad n = 1, 2, \cdots. \tag{3.20}$$

It follows from the previous argument that $\mathsf{Y}^{(0)} = I^0 \in \mathcal{H}_{meas}^D$ implies that $\mathsf{X}^{(0)}, \mathsf{K}^{(0)} \in \mathcal{H}_{meas}^D$. Furthermore, if $\mathsf{Y}^{(n-1)}, \mathsf{X}^{(n-1)}, \mathsf{K}^{(n-1)} \in \mathcal{H}_{meas}^D$, then by Lemma 3.5 and Lemma 3.4-(1), we have $\mathsf{Y}^{(n)} \in \mathcal{H}_{meas}^D$; thus so is $\mathsf{X}^{(n)}$ and then $\mathsf{K}^{(n)}, \cdots$. So by induction, we have $\mathsf{Y}^{(n)}, \mathsf{X}^{(n)}, \mathsf{K}^{(n)} \in \mathcal{H}_{meas}^D$ for all $n = 1, 2, \cdots$.

Moreover, a similar estimate as (3.3) gives that for fixed $P \in \mathcal{M}$,

$$\begin{aligned} (3.21)\quad & E^P \left| \mathsf{X}^{(n+1)}(P,\cdot) - \mathsf{X}^{(n)}(P,\cdot) \right|_t^{*,2} \\ = \quad & E^P \left| \Gamma(\mathsf{Y}^{(n+1)}(P,\cdot)) - \Gamma(\mathsf{Y}^{(n)}(P,\cdot)) \right|_t^{*,2} \leq C \int_0^t E^P |\mathsf{X}^{(n)}(P,\cdot) - \mathsf{X}^{(n-1)}(P,\cdot)|_s^{*,2} ds. \end{aligned}$$

By the boundedness of a and σ, one easily check, by using (2.3) that for each $t \geq 0$, there exists a constant $C_t > 0$ such that for all $P \in \mathcal{M}$,

$$E^P |X^{(1)}(P,\cdot) - X^{(0)}(P,\cdot)|_s^{*,2} \leq C_t \cdot s, \quad 0 \leq s \leq t.$$

Therefore, a simple iteration shows that

$$\sup_{P\in\mathcal{M}} E^P |X^{(n+1)}(P,\cdot) - X^{(n)}(P,\cdot)|_t^{*,2} \leq \frac{C}{n!} \cdot C_t t^n,$$

whence there exists an $\tilde{\mathsf{X}} \in \mathcal{H}$, such that

$$\lim_{n\to\infty} \sup_{P\in\mathcal{M}} E^P|\mathsf{X}^{(n)}(P,\cdot) - \tilde{\mathsf{X}}(P,\cdot)|_t^* = 0,$$

for all $t \geq 0$. By a fairly easy analogy with F. Riesz's Theorem, we can extract a subsequence that is independent of P, and hence may assume to be $\{X^{(n)}\}_{n=0}^\infty$ itself, such that for all $P \in \mathcal{M}$,

$$\lim_{n\to\infty} |\mathsf{X}^{(n)}(P,\omega) - \tilde{\mathsf{X}}(P,\omega)|_t^{*,2} = 0, \quad t \geq 0, \; P - a.e. \; \omega \in \Omega.$$

By Lemma 3.4-(3), $\tilde{\mathsf{X}} \in \mathcal{H}^D_{meas}$. It is then not hard to check, by using (3.2), that $\{\mathsf{Y}^{(n)}\}$, and then $\{\mathsf{K}^{(n)}\}$ also converge in a same manner to some limits $\tilde{\mathsf{Y}}, \tilde{\mathsf{K}} \in \mathcal{H}^D_{meas}$, respectively; moreover, it holds that

$$\tilde{\mathsf{Y}} = I^0 + I^1(\tilde{\mathsf{X}}) + I^2(\tilde{\mathsf{X}}). \tag{3.22}$$

It remains to show that $\tilde{\mathsf{X}} = \Gamma(\tilde{\mathsf{Y}})$ so that for each $P \in \mathcal{M}$, $(\tilde{\mathsf{X}}(P,\cdot), \tilde{\mathsf{K}}(P,\cdot))$ solves SDEDR$(\tilde{\mathsf{Y}}(P,\cdot))$ on $(\Omega, \mathcal{F}^\circ, P)$. To see this, define $\hat{\mathsf{X}} = \Gamma(\tilde{\mathsf{Y}})$. Since $E^P|\tilde{\mathsf{X}}(P,\cdot) - \mathsf{X}^{(n)}(P,\cdot)|_t^{*,2} \to 0$ as $n \to \infty$ for all $t \geq 0$ by definition, we wish to show that

$$E^P|\mathsf{X}^{(n)}(P,\cdot) - \hat{\mathsf{X}}(P,\cdot)|_t^{*,2} \to 0, \quad \text{for all } t \geq 0, \tag{3.23}$$

as $n \to \infty$. Once (3.23) is proved, then $\hat{\mathsf{X}} = \tilde{\mathsf{X}}$ in $\mathcal{H}$ and the uniqueness of the solution of SDEDR will yield that $\mathsf{Y}^x = \tilde{\mathsf{Y}}$; $\mathsf{X}^x = \hat{\mathsf{X}} = \tilde{\mathsf{X}}$; $\mathsf{K}^x = \tilde{\mathsf{K}}$ in $\mathcal{H}$, *i.e.*, $\mathsf{Y}^x, \mathsf{X}^x, \mathsf{K}^x \in \mathcal{H}^D_{meas}$ and the conlusion follows.

Finally, note that $\mathsf{X}^{(n)} - \hat{\mathsf{X}} = \Gamma(\mathsf{Y}^{(n)}) - \Gamma(\tilde{\mathsf{Y}})$. So by using (3.20) and (3.22); the similar estimate as (3.21) and the fact that $E^P|\mathsf{X}^{(n-1)}(P,\cdot) - \tilde{\mathsf{X}}(P,\cdot)|_t^{*,2} \to 0$, $t \geq 0$, we can derive (3.23) easily. The proof is now complete. □

4. An approximation theorem. In this section we validate the fact that there exists a "dense" sub-class of admissible controls with continuous paths. To be more specific, let $\mathcal{C}_{can}(T-r,y)$ by definition be the subclass of admissible controls consisting of all the elements $P \in \mathcal{M}$ such that $P\{\xi \in C[0,\infty) \cap \hat{D}_{T-r}(y), \; \xi^+ = 0\} = 1$, then it is clear that $\mathcal{C}_{can}(r,y) \subseteq \mathcal{D}_{can}(r,x,y)$ for any $x \geq 0$. Our main result of this section is to strengthen (2.12):

Theorem 4.1. *It holds that*

$$Q(r,x,y) = \inf_{P\in\mathcal{C}_{can}(T-r,y)} J(P,r,x). \tag{4.1}$$

□

First, it should be clear by Proposition 2.1 that we must only deal with the "local" counterpart of (4.1). Namely, it suffices to prove that

$$Q^{loc}(r,x,y) = \inf_{P\in\mathcal{C}^{loc}_{can}(T-r,y)} J(P,r,x).$$

Therefore in this section we will always restrict ourselves to the finite time interval $[0,T]$ and the local canonical space $(\Omega_T, \mathcal{G})$ and $\{\mathcal{G}_t\}$.

We start from some lemmas concerning the SDEDR. Let us fix $(r,x,y) \in E$ and $P \in \mathcal{D}^{loc}_{can}(r,x,y)$, the system (1.1) can be writen as

$$X_t^{x,-\xi} = x + \int_0^t a(X_s^{x,-\xi})ds + \int_0^t \sigma(X_s^{x,-\xi})dB_s - \xi_t + K_t, \quad 0 \le t \le T. \tag{4.2}$$

We recall the function $\phi(x)$ of (3.17), which is of class C^2, non-negative and strictly increasing. So by setting $Z_t = \phi(X_t^{x,-\xi})$, $z = \phi(x)$, $\psi = \phi^{-1}$ and applying Itô's Formula (in a general form, see [19]) to $\phi(X_t^{x,-\xi})$, we get from a little computation that

$$Z_t = z + \int_0^t \hat{a}(Z_s)ds + B_t - \hat{\xi}_t + \hat{K}_t, \tag{4.3}$$

where

$$\hat{a}(x) = \phi'(\psi(x))a(\psi(x)) + \frac{1}{2}\phi''(\psi(x))\sigma^2(\psi(x)); \tag{4.4}$$

$$\hat{\xi}_t = \int_{[0,t)} \phi'(X_s^{x,-\xi})d\xi_s - \sum_{0\le s<t} \left[\phi(X_{s+}^{x,-\xi}) - \phi(X_s^{x,-\xi}) + \phi'(X_s^{x,-\xi})\Delta X_s^{x,-\xi}\right]; \tag{4.5}$$

$$\hat{K}_t = \int_0^t \phi'(X_s^{x,-\xi})dK_s. \tag{4.6}$$

Lemma 4.2. *(1) The process $\hat{\xi}$ is non-decreasing, a.s. P;*

(2) The process $\hat{K}$ is continuous, and $(Z,\hat{K})$ is the solution of SDEDR (4.3) on $[0,T]$.

Proof. (i) Since $P \in \mathcal{D}^{loc}_{can}(r,x,y)$, K is continuous; whence $\Delta X_t^{x,-\xi} = -\Delta\xi_t$ and (4.5) can be writen as

$$\hat{\xi}_t = \int_{[0,t)} \phi'(X_s^{x,-\xi})d\xi^c_s - \sum_{0\le s<t} \left[\phi(X_{s+}^{x,-\xi}) - \phi(X_s^{x,-\xi})\right], \quad t \in [0,T].$$

Because ϕ is strictly increasing, it is easily seen that both terms on the right hand side above are nondecreasing since ξ^c is nondecreasing and $\Delta X_t^{x,-\xi} = -\Delta\xi_t \le 0$ for all $t \ge 0$, (1) thus follows.

(ii) It is obvious that the process $\hat{K}$ is continuous since K is so. The definition of Z_t gives that $Z_t \ge 0,\ 0 \le t \le T,\ a.s.P$. It remains to show that $\int_0^T Z_t d\hat{K}_t = 0,\ a.s.\ P$. But this is clear since

$$\int_0^T Z_t d\hat{K}_t = \int_0^T \phi(X_s^{x,-\xi})\phi'(X_s^{x,-\xi})1_{\{X_s^{x,-\xi}=0\}}dK_s = 0, \quad a.s.\ P,$$

proving the lemma. □

Our next step is to remove the "drift" term in equation (4.3) via the change of the probability measure. To this end, define

$$\theta_t = \exp\left\{-\int_0^t \hat{a}(Z_s)dB_s - \frac{1}{2}\int_0^t |\hat{a}(Z_s)|^2 ds\right\}. \tag{4.7}$$

The basic assumption (C.1), (C.2) on the functions a, σ leads to the boundedness of the function $\hat{a}$, which renders the process $\{\theta_t : 0 \le t \le T\}$ a $(\mathcal{G}_t, P)$-martingale. Thus if we define a probability measure P° on the canonical space $(\Omega_T, \mathcal{G})$ by

$$\frac{dP^\circ}{dP} = \theta_T, \tag{4.8}$$

then the Girsanov-Carmeron-Martin Theorem shows that the process $W_t \triangleq \int_0^t \hat{a}(Z_s)ds + B_s$ is a $(\mathcal{G}_t, P^\circ)$-Brownian motion for $0 \le t \le T$; and on the new probability space $(\Omega_T, \mathcal{G}, P^\circ; \mathcal{G}_t)$, we have

$$Z_t = z + W_t - \hat{\xi}_t + \hat{K}_t, \quad 0 \le t \le T. \tag{4.9}$$

Since $P^\circ \ll P$, it can be easily checked that $(Z, \hat{K})$ solves DRP$(z + W - \hat{\xi})$ on $(\Omega_T, \mathcal{G}, P^\circ; \mathcal{G}_t)$.

We now construct as in [11] a sequence of continuous processes

$$\hat{\xi}_t^{(n)} = \left\{ \begin{array}{ll} 0; & 0 \le t \le \frac{1}{n} \\ n \int_{t-\frac{1}{n}}^t \hat{\xi}_s ds; & \frac{1}{n} < t \le T - r \\ \hat{\xi}_{T-r}^{(n)}; & t > T - r \end{array} \right\}; \quad n \ge 1,$$

on the probability space $(\Omega_T, \mathcal{G}, P^\circ; \mathcal{G}_t)$. Then thanks to the Lemma 5.4 and Proposition 5.6 in [11], we have that $P^\circ\{\hat{\xi}_t^{(n)} \nearrow \hat{\xi}_t, \text{ as } n \to \infty, 0 \le t < \infty\} = 1$ and that

$$P^\circ \left\{ \lim_{n\to\infty} (Z_t^{(n)}, \hat{K}^{(n)}) = (Z_t, \hat{K}_t), \quad 0 \le t \le T \right\} = 1, \tag{4.10}$$

where $(Z^{(n)}, \hat{K}^{(n)})$ are the solutions of DRP$(z + W - \hat{\xi}^{(n)})$, $n = 1, 2, \cdots$.

Define for each n a probability measure $P^{(n)}$ on $(\Omega_T, \mathcal{G})$ by

$$\frac{dP^{(n)}}{dP^\circ} = \theta_T^{(n)}, \tag{4.11}$$

where

$$\theta_t^{(n)} = \exp\left\{ \int_0^t \hat{a}(Z_s^{(n)})dW_s - \frac{1}{2}\int_0^t |\hat{a}(Z_s^{(n)})|^2 ds \right\}, \tag{4.12}$$

It is readily seen that $\theta_T^{(n)} \to \theta_T^{-1}$, $a.s.P^\circ$, hence $P^{(n)} \Rightarrow P$. Thus $\{P^{(n)}\}$ should be the right candidates for our purpose as soon as we can modify them to satisfy $P^{(n)} \in \mathcal{C}_{can}^{loc}(T - r, y)$ and $\lim_{n\to\infty} J(P^{(n)}, r, x) = J(P, r, x)$.

To this end, we consider the probability space $(\Omega_T, \mathcal{G}, P^\circ; \mathcal{G}_t)$ on which $(Z, \hat{K})$ satisfies (4.9), and $(Z^{(n)}, \hat{K}^{(n)})$ satisfies

$$Z_t^{(n)} = z + W_t - \hat{\xi}_t^{(n)} + \hat{K}_t^{(n)}. \tag{4.13}$$

Define $X_t^{(n)} \triangleq \psi(Z_t^{(n)})$ and

$$K_t^{(n)} = \sigma(0)\hat{K}_t^{(n)};\ \xi^{(n)} = \int_0^t \psi'(Z_s^{(n)})d\hat{\xi}_s^{(n)};\ B_t^{(n)} = W_t - \int_0^t \hat{a}(Z_s^{(n)})ds. \tag{4.14}$$

Then it is easily checked that $(\Omega_T, \mathcal{G}, P^{(n)}, \mathcal{G}_t, B^{(n)}, -\xi^{(n)}) \in \mathcal{S}_{ad}^{loc}$. Furthermore, we have the following lemma.

Lemma 4.3. *(1) For each n, $(X^{(n)}, K^{(n)})$ solves SDEDR($x, -\xi^{(n)}$) on the set-up $(\Omega_T, \mathcal{G}, P^{(n)}, \mathcal{G}_t, B^{(n)}, -\xi^{(n)})$.*

(2) P-almost surely, probably along a subsequence, one has

$$\lim_{n\to\infty} \xi_t^{(n)} = \xi_t, \quad 0 \le t \le T; \tag{4.15}$$

$$\lim_{n\to\infty} (X_t^{(n)}, K_t^{(n)}) = (X_t, K_t), \quad 0 \le t \le T. \tag{4.16}$$

Remark 4.1. Unlike the Brownian motion case (see [11, Lemma 5.4]), here $(X^{(n)}, K^{(n)})$ may not even be the solution of SDEDR$(x, -\xi^{(n)})$ on the original space $(\Omega_T, \mathcal{G}, P)$. However, (4.15) and (4.16) do provide sufficient information for our approximation scheme. □

Proof. (i) First note that $P^{(n)} \ll P^\circ$ for all $n \ge 1$ and that on each probability space $(\Omega_T, \mathcal{G}, P^{(n)})$, the process $Z^{(n)}$ satisfies

$$Z_t^{(n)} = z + \int_0^t \hat{a}(Z_s^{(n)})ds + dB_s^{(n)} - \hat{\xi}_t^{(n)} + \hat{K}_t^{(n)}, \quad t \in [0,T], \tag{4.17}$$

and has continuous paths. Therefore by Itô's Formula we have

$$\begin{aligned} \psi(Z_t^{(n)}) \;&=\; x + \int_0^t [\psi'(Z_s^{(n)})\hat{a}(Z_s^{(n)}) + \frac{1}{2}\psi''(Z_s^{(n)})]ds + \int_0^t \psi'(Z_s^{(n)})dB_s^{(n)} \\ &\quad - \int_0^t \psi'(Z_s^{(n)})d\hat{\xi}_s^{(n)} + \int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)}. \end{aligned} \tag{4.18}$$

A direct computation shows that $a(x) = \psi'(\phi(x))\hat{a}(\phi(x)) + \frac{1}{2}\psi''(\phi(x))$, whence (4.18) can be rewriten as

$$X_t^{(n)} = x + \int_0^t a(X_s^{(n)})ds + \int_0^t \sigma(X_s^{(n)})dB_s^{(n)} - \xi_t^{(n)} + K_t^{(n)}, \tag{4.19}$$

where we used (4.14) and the fact that

$$\int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)} = \int_0^t \psi'(Z_s^{(n)})1_{\{Z_s^{(n)}=0\}}d\hat{K}_s^{(n)} = \sigma(0)\hat{K}_t^{(n)} = K_t^{(n)}. \tag{4.20}$$

The conclusion (1) now follows from a direct varification of the definition of a solution to SDEDR (see [18]).

(ii) The proof of (2) is a little bit more involved. We go back to the probability space $(\Omega_T, \mathcal{G}, P^\circ; \mathcal{G}_t)$ and apply Itô's Formula there to the C^2 function ψ by using Z

and $Z^{(n)}$ respectively to get

$$\begin{aligned}(4.21)\ \psi(Z_s^{(n)}) &= \psi(z) + \int_0^t \psi'(Z_s^{(n)})dW_s + \int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)} + \frac{1}{2}\int_0^t \psi''(Z_s^{(n)})ds \\ &\quad - \int_0^t \psi'(Z_s^{(n)})d\hat{\xi}_s^{(n)} \\ (4.22)\quad \psi(Z_s) &= \psi(z) + \int_0^t \psi'(Z_s)dW_s + \int_0^t \psi'(Z_s)d\hat{K}_s + \frac{1}{2}\int_0^t \psi''(Z_s)ds \\ &\quad - \int_0^t \psi'(Z_s)d\hat{\xi}_s + \sum_{0\le s<t}\left[\psi(Z_{s+}) - \psi(Z_s) - \psi'(Z_s)\Delta Z_s\right].\end{aligned}$$

Define

$$\Lambda_t = \int_0^t \psi'(Z_s)d\hat{\xi}_s - \sum_{0\le s<t}\left[\psi(Z_{s+}) - \psi(Z_s) - \psi'(Z_s)\Delta Z_s\right]. \tag{4.23}$$

then by (4.21), (4.22), (4.20) and note that $\psi'(\phi(x)) = \sigma(x)$, we have

$$\begin{aligned}(4.24)\qquad \xi_t^{(n)} &= -\Big\{\psi(Z_s^{(n)}) - \psi(z) - \int_0^t \psi'(Z_s^{(n)})dW_s \\ &\quad - \int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)} - \frac{1}{2}\int_0^t \psi''(Z_s^{(n)})ds\Big\}; \\ (4.25)\qquad \Lambda_t &= -\Big\{\psi(Z_s) - \psi(z) - \int_0^t \psi'(Z_s)dW_s \\ &\quad - \int_0^t \psi'(Z_s)d\hat{K}_s - \frac{1}{2}\int_0^t \psi''(Z_s)ds\Big\}.\end{aligned}$$

Note that the functions ψ' and ψ'' are bounded because of the assumptions (C.1), (C.2), thus by virtue of (4.24) and (4.25), the Bounded Convergence Theorem and the continuity of ϕ, we see that for each $0 \le t \le T$,

$$\xi_t^{(n)} \to \Lambda_t \Longleftrightarrow \int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)} \to \int_0^t \psi'(Z_s)d\hat{K}_s, \quad \text{as } n \to \infty.$$

However, by the same reason as (4.20), we have for P°-almost surely,

$$\int_0^t \psi'(Z_s^{(n)})d\hat{K}_s^{(n)} = \sigma(0)\hat{K}_t^{(n)}; \qquad \int_0^t \psi'(Z_s)d\hat{K}_s = \sigma(0)\hat{K}_t, \quad 0 \le t \le T.$$

So it follows from (4.10) that $P^\circ\{\lim_{n\to\infty}\xi_t^{(n)} = \Lambda_t, \text{ for all } t \in [0,T]\} = 1$. It then remains to show that $\Lambda = \xi$. Observe that $\phi(X_s) - \phi(X_{s+}) = -\Delta Z_s = \Delta\hat{\xi}_s$ and similarly, $\psi(Z_{s+}) - \psi(Z_s) = \Delta X_s = -\Delta\xi_s$, we have

$$\begin{aligned}(4.26)\qquad \int_0^t \psi'(Z_s)d\hat{\xi}_s &= \int_0^t \psi'(Z_s)\phi'(X_s)d\xi_s \\ &\quad + \sum_{0\le s<t}\psi'(Z_s)[\phi(X_s) - \phi(X_{s+}) - \phi'(X_s)\Delta\xi_s] \\ &= \xi_t + \sum_{0\le s<t}\left\{\psi'(Z_s)[\phi(X_s) - \phi(X_{s+})] - \Delta\xi_s\right\}. \\ &= \xi_t + \sum_{0\le s<t}\left\{\psi(Z_{s+}) - \psi(Z_s) - \psi'(Z_s)\Delta Z_s\right\}.\end{aligned}$$

This, together with (4.23), proves (4.15). Since $(X^{(n)}_t, K^{(n)}_t) = (\psi(Z^{(n)}_t), \sigma(0)\hat{K}^{(n)}_t)$ and $(\psi(Z_t), \sigma(0)\hat{K}_t) = (X_t, K_t)$, the conlusion (4.16) follows easily from (4.10), the continuity of the function ψ, and the fact that $P \ll P^\circ$. The proof is now complete. □

Our final step is to get a modification of $\xi^{(n)}$, say $\tilde{\xi}^{(n)}$, which satisfies the condition $P^{(n)}\{\tilde{\xi}^{(n)}_{T-r} \le y\} = 1$. To this end, define for each $0 \le r \le T$, $y > 0$ and $n = 1, 2, \cdots$ the stopping times

$$\tau^{(n)}_{r,y} = \begin{cases} \inf\{t \in [0, T-r];\ \xi^{(n)}_t > \xi_{T-r}\}; \\ T-r, \qquad\qquad \text{if}\{\cdots\} = \emptyset; \end{cases} \tag{4.27}$$

and define $\tilde{\xi}^{(n)}_t = \xi^{(n)}_{t\wedge\tau^{(n)}_{r,y}}$, $0 \le t \le T$, where $\xi^{(n)}$ is defined by (4.14). Then it is clear that $P^{(n)}$-almost surely, $\tilde{\xi}^{(n)}_t \le \xi_{T-r} \le y$, $0 \le t \le T$; and $\tilde{\xi}^{(n)}_{\tau^{(n)}_{r,y}} = \xi^{(n)}_{\tau^{(n)}_{r,y}} = \xi_{T-r}$ since $\xi^{(n)}$ is continuous. Therefore, if we denote $\tilde{P}^{(n)} = P^{(n)} \circ (B^{(n)}, \tilde{\xi}^{(n)})^{-1}$, then $\tilde{P}^{(n)} \in \mathcal{C}^{loc}_{can}(T-r, y)$.

Theorem 4.1 will now follow from the following lemma.

Lemma 4.4. *Let $\bar{P}^{(n)} = P^{(n)} \circ (B^{(n)}, \xi^{(n)})^{-1}$, then it holds that*

$$\lim_{n\to\infty} J(\bar{P}^{(n)}, r, x) = J(P, r, x); \tag{4.28}$$

$$\lim_{n\to\infty} [J(\tilde{P}^{(n)}, r, x) - J(\bar{P}^{(n)}, r, x)] = 0. \tag{4.29}$$

Proof. (i) First, by using the basic assumption (C.1),(C.2), it is easy to check that the sequence $\{E^{P^\circ}[\theta^{(n)}_T]^2\}$ is uniformly bounded in n; and by definition,

$$\begin{aligned} J(\bar{P}^{(n)}, r, x) &= E^{P^{(n)}}\left\{\int_0^{T-r} h(t+r, X^{(n)}_t)dt + \int_{[0,T-r)} f(r+t)d\xi^{(n)}_t + g(X^{(n)}_{T-r})\right\} \\ &= E^{P^\circ}\left\{\theta^{(n)}_T\left[\int_0^{T-r} h(t+r, X^{(n)}_t)dt + \int_{[0,T-r)} f(r+t)d\xi^{(n)}_t + g(X^{(n)}_{T-r})\right]\right\}. \end{aligned}$$

Thus a standard argument using Schwartz inequality, Dominated Convergence Theorem, Lemma 4.3, and the argument analogous to Lemma 5.5 of [11] will yield that

$$\begin{aligned} &\lim_{n\to\infty} J(\bar{P}^{(n)}, r, x) \\ &= E^{P^\circ}\left\{\theta^{-1}_T\left[\int_0^{T-r} h(t+r, X_t)dt + \int_{[0,T-r)} f(r+t)d\xi_t + g(X_{T-r})\right]\right\} \\ &= E^{P}\left\{\int_0^{T-r} h(t+r, X_t)dt + \int_{[0,T-r)} f(r+t)d\xi_t + g(X_{T-r})\right\} \\ &= J(P, r, x), \end{aligned}$$

thus (4.28) is proved. To prove (4.29), we first claim that

$$\lim_{n\to\infty}\int_0^{T-r}|d(\xi^{(n)}-\tilde{\xi}^{(n)})_s|\to 0,\qquad a.s.\ P^\circ. \tag{4.30}$$

Indeed, note that $\xi_t^{(n)}-\tilde{\xi}_t^{(n)}=1_{\{t>\tau_{r,y}^{(n)}\}}[\xi_t^{(n)}-\xi_{\tau_{r,y}^{(n)}}^{(n)}]$ so that $\xi^{(n)}-\tilde{\xi}^{(n)}$ is continuous, nondecreasing, we have by Lemma 4.3 that

$$\int_0^{T-r}|d(\xi^{(n)}-\tilde{\xi}^{(n)})_s|=\xi_{T-r}^{(n)}-\xi_{\tau_{r,y}^{(n)}}^{(n)}=\xi_{T-r}^{(n)}-\xi_{T-r}\to 0,\qquad a.s.\ P^\circ,$$

as $n\to\infty$, this proves the claim. Next, if we denote $(\tilde{X}^{(n)},\tilde{K}^{(n)})$ to be the solution of SDEDR$(x,\ -\tilde{\xi}^{(n)})$ on $(\Omega_T,\mathcal{G},P^{(n)})$, and

$$\begin{aligned}\tilde{Y}^{(n)} &= x+\int_0^t a(\tilde{X}_s^{(n)})ds+\int_0^t\sigma(\tilde{X}_s^{(n)})dB_s^{(n)}-\tilde{\xi}_t^{(n)};\\ Y^{(n)} &= x+\int_0^t a(X_s^{(n)})ds+\int_0^t\sigma(X_s^{(n)})dB_s^{(n)}-\xi_t^{(n)},\end{aligned}$$

Then under $P^{(n)}$, $X^{(n)}=\Gamma(Y^{(n)})$; $\tilde{X}^{(n)}=\Gamma(\tilde{Y}^{(n)})$. By (2.3), (C.1) and (C.2), we get

$$\begin{aligned} & E^{P^{(n)}}|X^{(n)}-\tilde{X}^{(n)}|_t^{*,2}=E^{P^{(n)}}|\Gamma(Y^{(n)})-\Gamma(\tilde{Y}^{(n)})|_{T-r}^{*,2} \\ \le\ & CE^{P^{(n)}}\left\{\left(\int_0^{T-r}|d(\xi^{(n)}-\tilde{\xi}^{(n)})_s|\right)^2+\int_0^t|X^{(n)}-\tilde{X}^{(n)}|_s^{*,2}ds\right\}.\end{aligned} \tag{4.31}$$

The Gronwall inequality and (4.30) then yield that $E^{P^{(n)}}|X^{(n)}-\tilde{X}^{(n)}|_{T-r}^{*,2}\to 0$ as $n\to\infty$. It is now not hard to check that (4.29) will follow from the Dominated Convergence Theorem and the fact (4.30), so Lemma 4.4 (hence Theorem 4.1) is proved. □

5. Basic properties of the cost and value functions. In the first half of this section we shall prove that the cost function $J:\mathcal{M}\times[0,T]\times(0,\infty)\to\infty$ is jointly measurable. The second half will be devoted to the study of the properties of the value function Q.

We begin with a lemma which may be of independent interest. Recall the space D and $\mathcal{P}(D)$ defined in §2.

Lemma 5.1. *Suppose that U is a seperable metrizable space with Borel σ-algebra $\mathcal{B}_U$, and that the mapping $\phi:U\mapsto\mathcal{P}(D)$ is $\mathcal{B}_U/\mathcal{B}_{\mathcal{P}(D)}$. Suppose $F:\mathbf{R}\times\mathbf{R}\mapsto\mathbf{R}$ is Borel measurable. Define, for each $(u,v,t)\in U\times\mathbf{R}\times[0,\infty)$,*

$$I(u,v,t)=\int_D F(v,\eta_t)\phi(u)(d\eta). \tag{5.1}$$

Then the function I is $(\mathcal{B}_U\times\mathcal{B}_{\mathbf{R}}\times\mathcal{B}_{[0,\infty)})/\mathcal{B}_{\mathbf{R}}$, provided the integrals exist for all $(u,v,t)\in U\times\mathbf{R}\times[0,\infty)$.

Proof. For each $u\in U$, consider the probability space $(D,\mathcal{B}_D,\phi(u))$ and the coordinate process $\{\pi_t:t\ge 0\}$ defined by $\pi_t(\eta)=\eta_t,\ t\ge 0$. Clearly, for fixed $u\in U$, the function $\hat{\phi}(u,t)\triangleq\phi(u)\circ\pi_t^{-1}$ is left-continuous in t.

On the other hand, for fixed $t \geq 0$, the function $\hat{\phi}(\cdot, t) = \phi(\cdot) \circ \pi_t^{-1}$ is $\mathcal{B}_U/\mathcal{B}_{\mathcal{P}(\mathbf{R})}$. Indeed, by (2.4) and the discussion related to it, we need only observe that for any $A, B \in \mathcal{B}_{\mathbf{R}}$ and $t \geq 0$, $\{p \in \mathcal{P}(D) : p(\pi_t^{-1}(A)) \in B\} = \theta^{-1}_{\pi_t^{-1}(A)}(B) \in \mathcal{B}_{\mathcal{P}(D)}$, hence by the measurability of ϕ, we have

$$\{u : \hat{\phi}(u,t)(A) \in B\} = \{u : \phi(u) \circ \pi_t^{-1}(A) \in B\} = \phi^{-1}(\theta^{-1}_{\pi_t^{-1}(A)}(B)) \in \mathcal{B}_U,$$

Consequently, $\hat{\phi} \in \mathcal{B}_U \times \mathcal{B}_{[0,\infty)}/\mathcal{B}_{\mathcal{P}(\mathbf{R})}$.

The rest of the proof is standard. We first check the result for those F's of the form $F(z_1, z_2) = 1_{A\times B}(z_1, z_2)$, $A, B \in \mathcal{B}_{\mathbf{R}}$. In that case, $I(u, v, t) = 1_A(v) \cdot \phi(u)(\pi_t^{-1}(B)) = 1_A(v) \cdot \hat{\phi}(u, t)(B)$, which is obviously $\mathcal{B}_U \times \mathcal{B}_{\mathbf{R}} \times \mathcal{B}_{[0,\infty)}/\mathcal{B}_{\mathbf{R}}$. The conclusion then follows from a standard Monotone-Class argument. □

We now define $I_1(u, v, z) \triangleq \int_D [\int_0^z F(v + t, \eta_t)dt]\, \phi(u)(d\eta)$, then by Fubini's Theorem, we have $I_1(u, v, z) = \int_0^z I(u, v + t, t)dt$. The following Corollary is obvious.

Corollary 5.2. *With the assumptions and the notations of Lemma 5.1, the function* $I_1(u, v, z)$ *is* $\mathcal{B}_U \times \mathcal{B}_{[0,\infty)} \times \mathcal{B}_{\mathbf{R}}/\mathcal{B}_{\mathbf{R}}$. □

We can now prove the following

Theorem 5.3. *The cost function* $J(\cdot, \cdot, \cdot)$ *is* $(\mathcal{B}_{\mathcal{M}} \times \mathcal{B}_{[0,T]} \times \mathcal{B}_{[0,\infty)})/\mathcal{B}_{\mathbf{R}}$.

Proof. Recall from Theorem 3.1 that the mapping $(x, P) \mapsto \Phi(x, P) \triangleq P \circ (F_P^x)^{-1}$ is $\mathcal{B}_{[0,\infty)} \times \mathcal{B}_{\mathcal{M}}/\mathcal{B}_{\mathcal{P}(D)}$. Write $J(P, r, x) = J_1(P, r, x) + J_2(P, r) + J_3(P, r, x)$, where

$$\begin{aligned}
J_1(P, r, x) &\triangleq E^P \int_0^{T-r} h(r + t, (F_P^x)_t)dt = \int_D \left[\int_0^{T-r} h(r + t, \eta_t)dt\right] \Phi(x, P)(d\eta);\\
J_2(P, r) &\triangleq E^P \int_{[0,T-r)} f(r + t)d\check{\xi}_t;\\
J_3(P, r, x) &\triangleq E^P g((F_P^x)_{T-r}) = \int_D g(\eta_{T-r})\Phi(x, P)(d\eta).
\end{aligned}$$

It is easily seen, by applying Lemma 5.1 and Corollary 5.2 with U being replaced by $[0, \infty)\times\mathcal{M}$ and $\phi(u)$ by $\Phi(x, P)$, that J_1 and J_3 are $\mathcal{B}_{\mathcal{P}(\Omega)}\times\mathcal{B}_{[0,T]}\times\mathcal{B}_{[0,\infty)}/\mathcal{B}_{\mathbf{R}}$. Therefore, We need only show the measurability of J_2.

Recall from §2 that $\mathcal{M}$ is the collection of all the element of a closed subset of $\mathcal{P}(\Omega)$ (*i.e.*, $\mathcal{M}_1$) restricted on $\hat{\Omega}$, we may content ourselves with $\hat{\Omega}$. Consider the mapping $\varphi : \hat{\Omega} \to \hat{D}$ defined by $\varphi(\omega) = \check{\zeta}$, where $\omega = (w, \zeta) \in \hat{\Omega}$. It is fairly easily to show that φ is $\hat{\mathcal{F}}/\mathcal{B}_{\hat{D}}$, where $\hat{\mathcal{F}}$ is the restriction of $\mathcal{F}$ on $\hat{\Omega}$. Thus $\phi : \mathcal{M} \to \mathcal{P}(\hat{D})$ defined by $\phi(P) = P \circ \varphi^{-1}$ is $\mathcal{B}_{\mathcal{M}}/\mathcal{B}_{\mathcal{P}(\hat{D})}$. Replacing the space D by $\hat{D}$ in Lemma 5.1, we see that for any $0 \leq s \leq t$, the function

$$J_{s,t}(P, r) \triangleq E^P[f(r + s)\check{\xi}_t] = \int_{\hat{D}} f(r + s)\eta_t\phi(P)(\eta)$$

is $\mathcal{B}_{\mathcal{M}} \times \mathcal{B}_{[0,T]}/\mathcal{B}_{\mathbf{R}}$. Thus for any partition $0 = t_0 < t_1 < \cdots < t_n = T - r$,

$$J_2^{(n)}(P, r) \triangleq E^P \left\{\sum_{i=1}^n f(r + t_i)(\check{\xi}_{t_{i+1}} - \check{\xi}_{t_i})\right\} = \sum_{i=1}^n [J_{t_i,t_{i+1}}(P, r) - J_{t_i,t_i}(P, r)]$$

is also $\mathcal{B}_{\mathcal{M}} \times \mathcal{B}_{[0,T]}/\mathcal{B}_{\mathbf{R}}$. Since $J_2(P,r)$ is the (pointwise) limit of the function of the form $J_2^{(n)}(P,r)$, it is $\mathcal{B}_{\mathcal{M}} \times \mathcal{B}_{[0,T]}/\mathcal{B}_{\mathbf{R}}$ as well. □

We now turn to the value function. First we claim that, by using Theorem 4.1, the moment estimates and comparison theorems for the solutions of SDEDR's that we derived in [18], and some well known facts of a reflected diffusion, we can "duplicate" most of the assertions in [11, §6] under our setting. The following theorem states the basic properties of Q that can be derived quite easily by a line to line analogy with those results in [11, §6]; thus we omit the details.

Theorem 5.4. *Suppose that the conditions (C.1) and (C.2) hold. Then the value function $Q : [0,T] \times [0,\infty) \times (0,\infty) \to \mathbf{R}$ is locally uniformly continuous. Moreover, for fixed $r \in [0,T]$, $y \geq 0$, the function $Q(r,\cdot,y)$ is nondecreasing; while for fixed $r \in [0,T]$, $x \geq 0$, the function $Q(r,x,\cdot)$ is nonincreasing. Finally, the value function satisfies the boundary condition:*

$$\lim_{\delta\searrow 0} \frac{Q(r,\delta,y) - Q(r,0,y)}{\delta} = 0. \tag{5.2}$$

□

Remark 5.1. The only results in [11, §6] that cannot be easily adapted to the nonlinear diffusion case are Lemma 6.1 (the convexity of Q in the variables x and y); a part of Lemma 6.7 and Corollary 6.8. For the latter, the basic difficulty seems to be that the inequality (6.6) there, which is essential for deriving (6.13) in Lemma 6.7 and (6.14) in Corollary 6.8 in [11], is no longer true when the drift and diffusion coefficients are non-constant. Nevertheless, by using a "cruder" estimate, namely, allowing the right hand sides of (6.13) and (6.14) there be multiplied by some constant $C > 0$, one can still derive the results stated in Theorem 5.4 above, which is sufficient for our purpose in this paper. □

Next, we shall establish some further properties of the value function concerning its behavier near the boundary $y = 0$ and the terminal time $r = T$. These properties will serve as the boundary and terminal conditions of our H-J-B equation in §8. Let $P \in \mathcal{M}$ be any canonical set-up. We donote

$$v(r,x) = E^P\left\{\int_0^{T-r} h(r+t, X_t^x)dt + g(X_{T-r}^x)\right\}, \tag{5.3}$$

where X^x is the reflected diffusion defined by

$$X_t^x = x + \int_0^t a(X_s^x)ds + \int_0^t \sigma(X_s^x)dB_s + K_t^x, \qquad t \geq 0. \tag{5.4}$$

It is clear that $v(r,x)$ is in dependent of the choice of the set-up P.

Proposition 5.5. *The value function Q satisfies the terminal condition:*

$$\lim_{r\nearrow T} Q(r,x,y) = \inf_{0\leq u\leq x\wedge y}\{f(T)u + g(x-u)\}, \quad (r,y) \in (0,T) \times (0,\infty); \tag{5.5}$$

and the boundary condition:

$$\lim_{y\searrow 0} Q(r,x,y) = v(r,x), \quad (r,x)\in(0,T)\times(0,\infty). \tag{5.6}$$

Proof. (i) For any $u\in[0,x\wedge y]$, define $P=P^W\times\delta_{\{\xi^u\}}$ on the canonical space $(\Omega,\mathcal{F})$, where P^W is the Wiener measure on $(W_0,\mathcal{B}_{W_0})$ and $\xi^u\in D$ is defined by $\xi^u_t = u\cdot 1_{(0,\infty)}(t),\ t\geq 0$. Then it is easily checked that

$$Q(r,x,y)\leq E^P\left\{\int_0^{T-r} h(r+t,X^{x-u}_t)dt + f(r)u + g(X^{x-u}_{T-r})\right\},$$

where X^{x-u} is the reflected diffusion starting from $x-u$. Thus it follows that

$$\limsup_{r\nearrow T} Q(r,x,y)\leq \inf_{0\leq u\leq x\wedge y}\{f(T)u+g(x-u)\}. \tag{5.7}$$

To prove the reverse inequality, we denote $b=\inf_{0\leq u\leq x\wedge y}\{f(T)u+g(x-u)\}$, and assert that for any sequence $r_n\nearrow T$, there exists a subsequence $\{r_{n'}\}$ such that $\lim_{n'\to\infty}Q(r_{n'},x,y)\geq b$. To prove our assertion, let $\{r_n\}$ be any sequence such that $r_n\nearrow T$. We choose for each n a $P^{(n)}_*\in\mathcal{D}_{can}(r_n,x,y)$ such that

$$J(P^{(n)}_*,r_n,x) = Q(r_n,x,y), \qquad n=1,2,3,\cdots.$$

Then by using the moment estimate of the solution to SDEDR (cf. [18]), the conditions on h and f, we can find a constant $C>0$ depending on T,x,y,h,f so that for any $\delta>0$ and n large enough, we have

$$Q(r_n,x,y)\geq -C\delta + E^{P^{(n)}_*}\{f(T)\cdot\xi_{T-r_n} + g(X^{x,-\xi}_{T-r_n})\} \tag{5.8}$$

Since $\{P^{(n)}_*\}$ is tight by [11, Corollary 12.3], the Skorohod Theorem enables us to find a probability space $(\hat\Omega,\hat{\mathcal{F}},\hat P)$ on which is defined a sequence $(\hat B^{(n)},\hat\xi^{(n)})$ with the distribution $P^{(n)}_*$, $n=1,2,\cdots$, such that for $\hat P$-almost surely (probably along a subsequence), $\lim_{n\to\infty}(\hat B^{(n)},-\hat\xi^{(n)})=(B,-\theta)$ for some Brownian motion B, and some nondecreasing process θ. By the property of the Skorohod topology (see for example, [8, Proposition 3.6.5]), $\lim_{n\to\infty}\hat\xi^{(n)}_{T-r_n}$ exists $\hat P$-almost surely, and if we denote $\lim_{n\to\infty}\hat\xi^{(n)}_{T-r_n}=\lambda$, then λ equals either $\theta_0=0$ or $\theta_{0+}\geq 0$. Observe that on $(\hat\Omega,\hat{\mathcal{F}},\hat P)$, if we denote $X^{(n)}=X^{x,-\hat\xi^{(n)}}$, then

$$X^{(n)}_{T-r_n} = x+\int_0^{T-r_n} a(X^{(n)}_s)ds + \int_0^{T-r_n}\sigma(X^{(n)}_s)d\hat B^{(n)}_s - \hat\xi^{(n)}_{T-r_n} + K^{(n)}_{T-r_n}\geq 0.$$

Hence upon taking a further subsequence one can show that for $\hat P$-almost surely,

$$\liminf_{n\to\infty} X^{(n)}_{T-r_n}\geq (x-\lambda)\vee 0 = x-(\lambda\wedge x)\geq 0.$$

It is obvious that $0 \le \lambda \le y$ so that $0 \le \lambda \wedge x \le x \wedge y$. So by Fatou's lemma and the monotonicity of g ((C.3)-(ii)), and non-negativity of f, we get

$$\begin{aligned}\liminf_{n\to\infty} E^{P_*^{(n)}}\{f(T)\xi_{T-r_n} + g(X^{x,-\xi}_{T-r_n})\} &= \liminf_{n\to\infty} E^{\hat P}\{f(T)\xi^{(n)}_{T-r_n} + g(X^{(n)}_{T-r_n})\} \\ &\ge E^{\hat P}\{f(T)(\lambda\wedge x) + g(x-(\lambda\wedge x))\} \ge b.\end{aligned}$$

Applying this to (5.8) gives us that $\liminf_{r\nearrow T} Q(r,x,y) \ge b$ since δ in (5.8) is arbitrary. Along with (5.7), we obtain (5.5).

(ii) Define $P = P^W \times \delta_{\{0\}}$ on the canonical space, where $\delta_{\{0\}}$ is the point mass at the zero function in D. Clearly, $P \in \mathcal{C}_{can}(T-r,y)$ for any (r,x,y), hence

$$v(r,x) = E^P\left\{\int_0^{T-r} h(r+t, X^{x,0}_t)dt + g(X^{x,0}_{T-r})\right\} \ge Q(r,x,y),$$

It follows that $\limsup_{y\searrow 0} Q(r,x,y) \le v(r,x)$.

Conversely, for any sequence $\{y^{(n)}\}$ such that $y^{(n)} \to 0$, let $P^{(n)}_* \in \mathcal{D}_{can}(r,x,y^{(n)})$ be the optimal controls with respect to $(r,x,y^{(n)})$, then

$$Q(r,x,y^{(n)}) \ge E^{P_*^{(n)}}\left\{\int_0^{T-r} h(r+t, X^{x,-\xi}_t)dt + g(X^{x,-\xi}_{T-r})\right\}.$$

By a similar argument as in part (i), we can find a probability space $(\hat\Omega, \hat{\mathcal{F}}, \hat P)$ and a sequence $(\hat B^{(n)}, \hat\xi^{(n)})$ with the distribution $P^{(n)}_*$, such that $\lim_{n\to\infty}(\hat B^{(n)}, -\hat\xi^{(n)}) = (B,-\theta)$, $\hat P$-almost surely. Since $\hat P(\hat\xi^{(n)}_{T-r} \le y^{(n)}) = P^{(n)}_*(\xi_{T-r} \le y^{(n)}) = 1$, we have $\hat\xi^{(n)}_t \le y^{(n)} \to 0$, $t \ge 0$, $\hat P$-a.s., whence $X^{x,-\hat\xi^{(n)}}_t \to X^{x,0}_t$ as $n\to\infty$, uniformly in t in compacts, $\hat P$-almost surely. Thus, $\liminf_{n\to\infty} Q(r,x,y^{(n)}) \ge v(r,x)$. The proof in now complete. □

6. A measurable selector. In [18], we proved that for any fixed $(r,x,y) \in E \triangleq [0,T]\times[0,\infty)\times(0,\infty)$, there exists a probability measure $P^* \in \mathcal{D}_{can}(r,x,y)$, such that

$$J(P^*, r, x) = Q(r,x,y).$$

In this section we shall find a *universally measurable* selector $R : E \mapsto \mathcal{M}$, such that $R(r,x,y) \in \mathcal{D}_{can}(r,x,y)$ and $J(R(r,x,y),r,x) = Q(r,x,y)$ for all $(r,x,y) \in E$.

To begin with, we look at the set $\mathcal{K} \triangleq \{(P,r,x,y) : P \in \mathcal{D}_{can}(r,x,y),\ (r,x,y) \in E\}$. We shall prove that $\mathcal{K}$ is a Borel measurable set in the space $\mathcal{M}\times E$. In the case when $a = 0$, $\sigma = 1$, this is quite clear (see [11, §13]). However, when the system is non-linear as ours, it takes a little bit more work. In fact, the results of the previous sections are mainly motivated by this consideration.

Recall the set $\mathcal{M}(T-r,y)$ defined in §2, where $r \in [0,T]$ and $y > 0$; and define

$$\mathcal{C} = \{p \in \mathcal{P}(D) : p(C[0,\infty)) = 1\}. \tag{6.1}$$

It is not hard to check that $\mathcal{C}$ is a Borel subset of $\mathcal{P}(D)$.

For each $T > 0$ and $N \geq 0$, let $K_N \triangleq [0, T] \times [0, N] \times [\frac{1}{N}, N]$ and

$$\mathcal{K}_N \triangleq \{(P, r, x, y) \in \mathcal{K} : (r, x, y) \in K_N\}.$$

Then $\mathcal{K} = \cup_{N=1}^{\infty} \mathcal{K}_N$, hence it suffices to show that each $\mathcal{K}_N$ is Borel measurable. But this follows immediately from the following lemma.

Lemma 6.1. *(1)* $\mathcal{K}_N = \mathcal{K}_N^1 \cap \mathcal{K}_N^2$, *where*

$$\mathcal{K}_N^1 \triangleq \{(P, r, x, y) : P \in \mathcal{M}^-(T - r, y)\};$$
$$\mathcal{K}_N^2 \triangleq \{(P, r, x, y) : P \in \Psi(x, \cdot)^{-1}(\mathcal{C}),\ (r, x, y) \in K_N\};$$

where $\Psi(x, P)$ *is defined by (3.2).*

(2) $\mathcal{K}_N^1$ *is a closed set in* $\mathcal{M} \times E$*;*

(3) $\mathcal{K}_N^2$ *is a Borel set in* $\mathcal{M} \times E$*.*

Proof. (1) is obvious by the definition and the fact that

$$\Psi(x, P) \in \mathcal{C} \Longleftrightarrow P\{G_P^x \in C[0, \infty)\} = 1.$$

To see (2), let $(P_n, r_n, x_n, y_n) \in \mathcal{K}_N^1$ such that $(P_n, r_n, x_n, y_n) \to (P, r, x, y)$ as $n \to \infty$. Then $(r, x, y) \in K_N$ since K_N is closed. To check that $P \in \mathcal{M}^-(T-r, y)$, we again apply the Skorohod Theorem to obtain on some probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ a sequence of Ω-valued processes $\{(B^{(n)}, \xi^{(n)})\}_{n=0}^{\infty}$ with distributions P and $\{P_n\}_{n=1}^{\infty}$ respectively, such that $(B^{(n)}, \xi^{(n)}) \to (B^{(0)}, \xi^{(0)}),\ n \to \infty,\ a.s.\hat{P}$ for some $(\hat{\mathcal{F}}_t, \hat{P})$-Brownian motion $B^{(0)}$. Moreover, it is easy to check that for each $t > T - r$, we have $\hat{P}\{\xi_t^{(0)} = \lim_{n\to\infty} \xi_t^{(n)} \leq y\} = 1$. So by letting $t \searrow T - r$, we get $\hat{P}\{\xi_{(T-r)+}^{(0)} \leq y\} = 1$, whence $P \in \mathcal{M}^-(T-r, y)$.

The proof of (3) is based on the results derived in §3. It is easily seen by changing the order of variables that

$$\mathcal{K}_N^2 = \{(P, r, x, y) \in \mathcal{M} \times E : (r, y, x, P) \in (\hat{K}_N \times \mathcal{M}) \cap [0, T] \times [\frac{1}{N}, N] \times \Psi^{-1}(\mathcal{C})\},$$

where $\hat{K}_N = \{(r, y, x) : (r, x, y) \in K_N\}$. Therefore (3) follows from Theorem 3.1, and the proof is complete. □

We can now prove the following

Theorem 6.2. *There exists a analytically measurable function* $R : E \mapsto \mathcal{P}(\Omega)$ *such that for each* $(r, x, y) \in E$*,* $R(r, x, y) \in \mathcal{D}_{can}(r, x, y)$ *and* $J(R(r, x, y), r, x) = Q(r, x, y)$*.*
□

Proof. By Theorem 5.3 and 5.4, $J(P, r, x)$ is $\mathcal{B}_{\mathcal{P}(\Omega)} \times \mathcal{B}_E / \mathcal{B}_{\mathbf{R}}$ and $Q(r, x, y)$ is continuous; thus Lemma 6.1 and the argument preceding it tell us that the set

$$\mathcal{L} \triangleq \{(P, r, x, y) : P \in \mathcal{D}_{can}(r, x, y),\ J(P, r, x) = Q(r, x, y)\}$$

is a Borel measurable set in the Borel space $\mathcal{M} \times E$. Moreover, the existence of the optimal control (cf. [18, §7]) implies that $Proj_E(\mathcal{L}) = E$. Therefore, by the Jankov-von Neumann Theorem (cf. *e.g.*, [3]), there exists a analytically measurable (therefore universally measurable) selector $R : E = Proj_E(\mathcal{L}) \mapsto \mathcal{P}(\Omega)$ such that $Graph(R) \subseteq \mathcal{L}$.
□

7. The Bellman Principle. In this section, we put the results that we derived in the previous paper [18] and the previous sections in this paper together to prove the Equation of Dynamic Programming (*i.e.*, the Bellman Principle). We begin with some notations.

Definition 7.1. *For any given numbers $0 \le u < v < \infty$, we denote by $\Im_{u,v}$ the collection of all $\{\mathcal{F}_t\}$-stopping times τ with values in $[u,v]$. Define $\Im = \Im_{0,T}$.* □

For any $\tau \in \Im_{0,T-r}$, $(r,x,y) \in E$, and $P \in \mathcal{M}(T-r,y)$, let $\mathsf{Q}^\omega(A) : \Omega \times \mathcal{F} \mapsto [0,1]$ be the regular conditional probability of P given $\mathcal{F}_\tau$. Such $\mathsf{Q}^\cdot(\cdot)$ always exists since Ω is Polish (cf. for example, [10]). Furthermore, let $(X^{x,-\xi}, K^{x,-\xi})$ be the solution to SDEDR$(x,-\xi)$ on the probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$, consider the $\mathcal{F}_\tau$-measurable random vector $\mathcal{X} \triangleq (\tau, B_\tau, \xi_\tau, X^{x,-\xi}_\tau, K^{x,-\xi}_\tau)$. It is well known that (cf. *e.g.*, [22]), for $P-a.e.$ $\omega \in \Omega$,

$$\mathsf{Q}^\omega\{\omega' : \mathcal{X}(\omega') = \mathcal{X}(\omega)\} = 1, \tag{7.1}$$

and that $\tilde{B}_t \triangleq B_{\tau+t} - B_\tau$ is an $(\mathcal{F}_{\tau+t}, \mathsf{Q}^\omega)$-Brownian motion. The following proposition will be useful.

Proposition 7.2. *Let $(r,x,y) \in E$, $\tau \in \Im_{0,T-r}$, and $P \in \mathcal{D}_{can}(r,x,y)$ be given. Let $\mathsf{Q}^\cdot(\cdot) : \Omega \times \mathcal{F} \mapsto [0,1]$ be the regular conditional probability of P given $\mathcal{F}_\tau$. Suppose that the pair $(X^{x,-\xi}, K^{x,-\xi})$ solves SDEDR($x,-\xi$) on the probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$. Denote, $\tilde{X}^\omega_t = X^{x,-\xi}_{\tau(\omega)+t}$, $\tilde{x}(\omega) = X^{x,-\xi}_{\tau(\omega)}(\omega)$, $\omega \in \Omega$; and with Z denoting $K^{x,-\xi}, B, \xi$ respectively, set $\tilde{Z}_t = Z_{\tau+t} - Z_\tau$. Then for $P-a.e.$ $\omega \in \Omega$, the pair $(\tilde{X}, \tilde{K})$ solves SDEDR$(\tilde{x}(\omega),\ -\tilde{\xi}^\omega)$ on the set-up $(\Omega, \mathcal{F}, \mathsf{Q}^\omega, \tilde{\mathcal{F}}_t, \tilde{B}, -\xi)$, where $\tilde{\mathcal{F}}_t = \mathcal{F}_{\tau+t}$.*

Moreoever, if $\tilde{\mathsf{Q}}^\omega = \mathsf{Q}^\omega \circ (\tilde{B}, -\tilde{\xi})^{-1}$, then $\tilde{\mathsf{Q}}^\omega \in \mathcal{D}_{can}(r+\tau(\omega), \tilde{x}(\omega), y - \xi_{\tau(\omega)}(\omega))$ for $P-a.e.$ $\omega \in \Omega$.

Proof. Let (r,x,y), τ, P be given as assumed, consider the original probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$. Since P-almost surely, we have

$$\begin{aligned} X^{x,-\xi}_{\tau+t} &= \tilde{x} + \int_\tau^{\tau+t} a(X^{x,-\xi}_s)ds + \int_\tau^{\tau+t} \sigma(X^{x,-\xi}_s)dB_s - (\xi_{\tau+t} - \xi_\tau) \\ &\quad + (K^{x,-\xi}_{\tau+t} - K^{x,-\xi}_\tau), \qquad t \ge 0. \end{aligned} \tag{7.2}$$

A standard Monotone-Class argument and the fact (7.1) enable us to rewrite (7.2) as

$$\tilde{X}_t = \tilde{x}(\omega) + \int_0^t a(\tilde{X}_s)ds + \int_0^t \sigma(\tilde{X}_s)d\tilde{B}_s - \tilde{\xi}_t + \tilde{K}_t, \quad t \ge 0,\ a.s.\mathsf{Q}^\omega, \tag{7.3}$$

for $P-a.e.$ $\omega \in \Omega$. It is clear that $\mathsf{Q}^\omega\{\tilde{X}_t \ge 0, t \ge 0\} = 1$; $\mathsf{Q}^\omega\{\tilde{K} \in C[0,\infty)\} = 1$, and

$$\mathsf{Q}^\omega\left\{\int_0^\infty \tilde{X}_t d\tilde{K}_t = 0\right\} = \mathsf{Q}^\omega\left\{\int_{\tau(\omega)}^\infty X^{x,-\xi}_t dK^{x,-\xi}_t = 0\right\} = 1,$$

for $P-a.e.$ $\omega \in \Omega$, so the pair $(\tilde{X}, \tilde{K})$ solves SDEDR$(\tilde{x}(\omega), -\tilde{\xi})$ on $(\Omega, \mathcal{F}, \mathsf{Q}^\omega; \tilde{\mathcal{F}}_t)$ for $P-a.e.$ $\omega \in \Omega$, proving the first assertion.

To see the second part, let $\tilde{\mathsf{Q}}^\omega = \mathsf{Q}^\omega \circ (\tilde{B}, -\tilde{\xi})^{-1}$. By Theorem 4.4 in [18], for each $\omega \in \Omega$, there exists a pair $(F^\omega, G^\omega) : \Omega \to D^2$ which solves SDEDR$(\tilde{x}(\omega), -\bar{\xi})$ on the canonical space $(\Omega, \mathcal{F}, \mathsf{Q}^\omega)$, where $F^\omega \triangleq F^{\tilde{x}(\omega)}_{\tilde{\mathsf{Q}}^\omega}, G^\omega \triangleq G^{\tilde{x}(\omega)}_{\tilde{\mathsf{Q}}^\omega}$, and $-\bar{\xi}$ now denotes the second component of the canonical process; such that for $P - a.e.\ \omega \in \Omega$, Q^ω-almost surely,

$$F^\omega(\tilde{B}, -\tilde{\xi})_t = \tilde{X}_t; \quad G^\omega(\tilde{B}, -\tilde{\xi})_t = \tilde{K}_t, \quad t \geq 0.$$

It is easy to check (recall (7.1)) that for $P - a.e.\ \omega \in \Omega$,

$$\tilde{\mathsf{Q}}^\omega\{\omega' : \bar{\xi}_t(\omega') = \bar{\xi}_{T-r-\tau(\omega)}(\omega) \leq y - \xi_{\tau(\omega)}(\omega), \quad t > T - r - \tau(\omega)\} = 1.$$

Therefore $\tilde{\mathsf{Q}}^\omega \in \mathcal{M}^-(T - r - \tau(\omega), y - \xi_{\tau(\omega)}(\omega))$, for $P - a.e.\ \omega \in \Omega$.

Finally, since $P\{K^{x,-\xi} \in C[0,\infty)\} = 1$, we have for $P - a.e.\ \omega \in \Omega$,

$$\tilde{\mathsf{Q}}^\omega\{G^\omega \in C[0,\infty)\} = \mathsf{Q}^\omega\{G^\omega(\tilde{B}, -\tilde{\xi}) \in C[0,\infty)\} = \mathsf{Q}^\omega\{\tilde{K} \in C[0,\infty)\} = 1,$$

whence $\tilde{\mathsf{Q}}^\omega \in \mathcal{D}_{can}(r + \tau(\omega), \tilde{x}(\omega), y - \xi_{\tau(\omega)}(\omega))$ for $P - a.e.\ \omega \in \Omega$, proving the proposition. □

The main result of this section is the following system of equations of dynamic programming (Bellman Principle).

Theorem 7.3. (Bellman Principle) *For every $(r, x, y) \in E$, and every $\tau \in \Im_{0,T-r}$,*

$$Q(r,x,y) = \inf_{P \in \mathcal{D}_{can}(r,x,y)} E^P \left\{ \int_0^\tau h(r+t, X_t^{x,-\xi})dt + \int_{[0,\tau)} f(r+t)d\xi_t + Q(r+\tau, X_\tau^{x,-\xi}, y - \xi_\tau) \right\}, \tag{7.4}$$

$$Q(r,x,y) = \inf_{0 \leq u \leq x \wedge y} \{f(r)u + Q(r, x-u, y-u)\}. \tag{7.5}$$

Proof. We first assume (7.4) to prove (7.5), since the latter is relatively easier. Define $I(r,x,y;u) = f(r)u + Q(r, x-u, y-u)$. Because $Q(r,x,y) = I(r,x,y,0)$, we have $\inf_{0 \leq u \leq x \wedge y} I(r,x,y;u) \leq Q(r,x,y)$.

Conversely, for any fixed $u \in [0, x \wedge y]$, we define $\xi^u_t = u \cdot 1_{(0,\infty)}(t)$ and $P = P^W \times \delta_{\{\xi^u\}}$, where P^W is the Wiener measure on $(W_0, \mathcal{B}_{W_0})$ and $\delta_{\{\xi^u\}} \in \mathcal{P}(D)$ is the point mass at ξ^u. Clearly, $P \in \mathcal{D}_{can}(r,x,y)$. Now for any $\epsilon > 0$, by (7.4), we have

$$\begin{aligned} &Q(r,x,y) \\ &\leq E^P \left\{ \int_0^\epsilon h(r+t, X_t^{x,-\xi^u})dt + \int_{[0,\epsilon)} f(r+t)d\xi^u_t + Q(r+\epsilon, X_\epsilon^{x,-\xi^u}, y - \xi^u_\epsilon) \right\} \\ &= E^P \left\{ \int_0^\epsilon h(r+t, X_t^{x,-\xi^u})dt + f(r)u + Q(r+\epsilon, X_\epsilon^{x,-\xi^u}, y-u) \right\}. \end{aligned} \tag{7.6}$$

Note that $\lim_{\epsilon\searrow 0} X_\epsilon^{x,-\xi^u} = X_{0+}^{x,-\xi^u} = x-u,\ a.s.P$; and that the value function Q is locally uniformly continuous (see Theorem 5.4), we obtain that

$$
\begin{aligned}
Q(r,x,y) &\le \lim_{\epsilon\searrow 0} E^P\left\{\int_0^\epsilon h(r+t,X_t^{x-u,0})dt + f(r)u + Q(r+\epsilon, X_\epsilon^{x-u,0}, y-u)\right\} \\
&= I(r,x,y;u), \qquad (7.7)
\end{aligned}
$$

whence (7.5) follows from taking the infimum in (7.7).

Proof of (7.4): We split the proof of into two lemmas which will take care of two directions of inequalities respectively.

Lemma 7.4. *For any $(r,x,y)\in E$ and $\tau\in\Im_{0,T-r}$, it holds that*

$$
\begin{aligned}
Q(r,x,y) \ \ge\ \inf_{P\in\mathcal{D}_{can}(r,x,y)} E^P\Big\{&\int_0^\tau h(r+t,X_t^{x,-\xi})dt + \int_{[0,\tau)} f(r+t)d\xi_t \\
&+Q(r+\tau, X_\tau^{x,-\xi}, y-\xi_\tau)\Big\}. \qquad (7.8)
\end{aligned}
$$

Proof. Let $(r,x,y)\in E$ and $\tau\in\Im_{0,T-r}$ be given; pick any $P\in\mathcal{D}_{can}(r,x,y)$. Again denote by Q^ω the regular conditional probability of P given $\mathcal{F}_\tau$. By Proposition 7.2 (with the same notations there), we have $\tilde{\mathsf{Q}}^\omega\in\mathcal{D}_{can}(r+\tau(\omega),\tilde{x}(\omega),y-\xi_{\tau(\omega)}(\omega))$, whence

$$
\begin{aligned}
&Q(r+\tau(\omega),\tilde{x}(\omega),y-\xi_{\tau(\omega)}(\omega)) \\
\le\ & E^{\tilde{\mathsf{Q}}^\omega}\left\{\int_0^{T-r-\tau} h(r+\tau+t,\tilde{X}_t^\omega)dt + \int_{[0,T-r-\tau)} f(r+\tau+t)d\tilde{\xi}_t + g(\tilde{X}^\omega_{T-r-\tau})\right\} \\
=\ & E^P\left\{\int_\tau^{T-r} h(r+t,X_t^{x,-\xi})dt + \int_{[\tau,T-r)} f(r+t)d\xi_t + g(X_{T-r}^{x,-\xi})\,\middle|\,\mathcal{F}_\tau\right\}(\omega),
\end{aligned}
$$

for $P-a.e.\ \omega\in\Omega$. It follows that

$$
\begin{aligned}
J(P,r,x) \ \ge\ E^P\Big\{&\int_0^\tau h(r+t,X_t^{x,-\xi})dt + \int_{[0,\tau)} f(r+t)d\xi_t \\
&+Q(r+\tau, X_\tau^{x,-\xi}, y-\xi_\tau)\Big\}. \qquad (7.9)
\end{aligned}
$$

By taking the infimum over $P\in\mathcal{D}_{can}(r,x,y)$ on both sides of (7.9), we derive (7.8). □

The reverse inequality is the direct consequence of the following lemma.

Lemma 7.5. *For each $(r,x,y)\in E$, $\tau\in\Im_{0,T-r}$ and $P\in\mathcal{D}_{can}(r,x,y)$, there exists a $P^*\in\mathcal{D}_{can}(r,x,y)$ such that $P^*=P$ on $\mathcal{F}_\tau$, and*

$$
\begin{aligned}
J(P^*,r,x) \ =\ E^{P^*}\Big\{&\int_0^\tau h(r+t,X^{x,-\xi})_t)dt + \int_{[0,\tau)} f(r+t)d\xi_t \\
&+Q(r+\tau, X_\tau^{x,-\xi}, y-\xi_\tau)\Big\}. \qquad (7.10)
\end{aligned}
$$

Proof. Suppose that $(r,x,y) \in E$, $\tau \in \Im_{0,T-r}$ and $P \in \mathcal{D}_{can}(r,x,y)$ are given. Let $R : (r,x,y) \mapsto \mathcal{D}_{can}(r,x,y)$ be the (universally) measurable selector obtained in Theorem 6.2. Let $Z(\omega) = (r+\tau(\omega), \tilde{x}(\omega), y - \xi_{\tau(\omega)}(\omega))$ for $\omega \in \Omega$ (recall that $\tilde{x}(\omega) = X^{x,-\xi}_{\tau(\omega)}(\omega)$); and let $\mu = P \circ Z^{-1} \in \mathcal{P}(E)$. By the universal measurability of R, there exists a function $R_\mu : (r,x,y) \mapsto \mathcal{D}_{can}(r,x,y)$ which is a version of R (under μ) and is $\mathcal{B}_E/\mathcal{B}_{\mathcal{M}}$, (see, for example, [3, §7]). We now define $\mathsf{Q}_\omega \triangleq R_\mu(Z(\omega)) \in \mathcal{M}$, $\omega \in \Omega$ (note the notational difference between this Q_ω and the Q^ω defined before). It is easy to check that for each $A \in \mathcal{F}$, $\mathsf{Q}_\cdot(A)$ is $\mathcal{F}_\tau$-measurable; moreover, we have by definition that the value function at $Z(\omega)$ is $Q(Z(\omega)) = J(\mathsf{Q}_\omega, r+\tau(\omega), \tilde{x}(\omega))$.

To construct the desired P^*, we follow the idea of [11] and [22]. First, for each $t \geq 0$, we define a mapping $\gamma : \Omega \times \Omega \to \Omega$ by

$$\gamma_t(\omega,\omega') = \omega'(t) - \omega'(t \wedge \tau(\omega)) + \omega(t \wedge \tau(\omega)), \qquad t \geq 0. \tag{7.11}$$

We claim that γ is $\mathcal{F}_\tau \times \mathcal{F}/\mathcal{F}$. Indeed, observe that $\gamma = \sum_{i=1}^3 \gamma^i$, where for $t \geq 0$,

$$\gamma_t^1(\omega,\omega') = \omega'(t);\ \gamma_t^2(\omega,\omega') = \omega'(t \wedge \tau(\omega));\ \gamma_t^3(\omega,\omega') = \omega(t \wedge \tau(\omega)).$$

It is readily seen that γ^1 and γ^3 are $\mathcal{F}_\tau \times \mathcal{F}/\mathcal{F}$ since they depend actually only on one variable. As for γ^2, we first fix $t \geq 0$, consider the following two mappings: $\chi : [0,\infty) \times \Omega \to \mathbf{R}$ and $\tilde{\tau}_t : \Omega \times \Omega \to [0,\infty) \times \Omega$ defined by

$$\chi(t',\omega') = \omega'(t'); \qquad \tilde{\tau}_t(\omega,\omega') = (t \wedge \tau(\omega), \omega').$$

It is then obvious that χ is $\mathcal{B}_{[0,\infty)} \times \mathcal{F}/\mathcal{B}_{\mathbf{R}}$ and $\tilde{\tau}_t$ is $\mathcal{F}_{t\wedge\tau} \times \mathcal{F}/\mathcal{B}_{[0,\infty)} \times \mathcal{F}$ for every $t \geq 0$. Since $\gamma_t^2 = \chi \circ \tilde{\tau}_t$, it is $\mathcal{F}_{t\wedge\tau} \times \mathcal{F}/\mathcal{B}_{\mathbf{R}}$, whence $\mathcal{F}_\tau \times \mathcal{F}/\mathcal{B}_{\mathbf{R}}$, for every $t \geq 0$. Because $\mathcal{F}$ is generated by the cylinder sets, we obtain that γ^2 is $\mathcal{F}_\tau \times \mathcal{F}/\mathcal{F}$. This substantiates the claim.

We now define for each $\omega \in \Omega$,

$$\tilde{\mathsf{Q}}_\omega = \mathsf{Q}_\omega \circ \gamma(\omega,\cdot)^{-1}. \tag{7.12}$$

By an easy Monotone-Class argument, one shows that for any $f \in \mathcal{F}_\tau \times \mathcal{F}/\mathcal{B}_{\mathbf{R}}$, the mapping $\omega \mapsto \int_\Omega f(\omega,\omega')\mathsf{Q}_\omega(d\omega')$ is $\mathcal{F}_\tau/\mathcal{B}_{\mathbf{R}}$. Consequently, the mapping

$$\omega \mapsto \tilde{\mathsf{Q}}_\omega(A) = \int_\Omega 1_A(\gamma(\omega,\omega'))\mathsf{Q}_\omega(d\omega') = \int_\Omega 1_{\gamma_A^{-1}}(\omega,\omega')\mathsf{Q}_\omega(d\omega')$$

is $\mathcal{F}_\tau/\mathcal{B}_{\mathbf{R}}$. Furthermore, it is clear by the definition (7.11) and (7.12) that $\tilde{\mathsf{Q}}_\omega\{\omega' : \omega'(\tau(\omega)) = \omega(\tau(\omega))\} = 1$ for every $\omega \in \Omega$. Therefore, we can apply Theorem 6.1.2 in [22] to obtain for each $P \in \mathcal{D}_{can}(r,x,y)$ a $P^* \in \mathcal{P}(\Omega)$ such that

(1) $P^*|_{\mathcal{F}_\tau} = P|_{\mathcal{F}_\tau}$;

(2) $P^*[\ \cdot\ |\mathcal{F}_\tau](\omega) = \delta_\omega \otimes_{\tau(\omega)} \tilde{\mathsf{Q}}_\omega$, for $P-a.e.\ \omega \in \Omega$,

where for $s \geq 0$, $\mathsf{Q} \in \mathcal{P}(\Omega|_{[s,\infty)})$ and $\omega \in \Omega$, $\delta_\omega \otimes_s \mathsf{Q}$ denotes the probablity measure in $\mathcal{P}(\Omega)$ satisfying (1) $\delta_\omega \otimes_s \mathsf{Q}\{\omega' : \omega'(t) = \omega(t),\ t \leq s\} = 1$ and (2) $\delta_\omega \otimes_s \mathsf{Q}(A) = \tilde{\mathsf{Q}}(A)$, for $A \in \sigma\{\omega'(t);\ t \geq s\}$ (see [22, Lemma 6.1.1]). It is then intuitively clear by the construction of P^* and the properties of P and $\tilde{\mathsf{Q}}$ that P^* belongs to $\mathcal{D}_{can}(r, x, y)$. The justification is straight forward and similar to [11, Proposition 13.5], with the help of [22, Theorem 1.2.10], we therefore omit it.

Finally, baring in mind that $P^* = P$ on $\mathcal{F}_\tau$, we can apply Proposition 7.2 and use the similar argument as in Lemma 7.4 to obtain that

$$
\begin{aligned}
(7.13)\quad & J(P^*, r, x) \\
= \ & E^{P^*}\left\{\int_0^{T-r} h(r+t, (F^x_{P^*})_t)dt + \int_{[0,T-r)} f(r+t)d\xi_t + g((F^x_{P^*})_{T-r})\right\} \\
= \ & E^{P^*}\left\{\int_0^{\tau} h(r+t, (F^x_{P^*})_t)dt + \int_{[0,\tau)} f(r+t)d\xi_t \right.\\
& \left. + E^{P^*}\left[\int_\tau^{T-r} h(r+t, (F^x_{P^*})_t)dt + \int_{[0,T-r)} f(r+t)d\xi_t + g((F^x_{P^*})_{T-r}) \middle| \mathcal{F}_\tau\right]\right\} \\
= \ & E^{P^*}\left\{\int_0^{\tau} h(r+t, (F^x_{P^*})_t)dt + \int_{[0,\tau)} f(r+t)d\xi_t + Q(T-r-\tau, X^{x,-\xi}_\tau, y-\xi_\tau)\right\}.
\end{aligned}
$$

Namely, (7.10) holds. □

8. The Hamilton-Jacobi-Bellman equation. Throughout this section, we denote $z \triangleq (r, x, y)$ and $E^\circ \triangleq (0, T) \times (0, \infty) \times (0, \infty)$. For any $u \in C^2(E^\circ)$, let Du be the gradient and D^2u the Hessian of u. Define the following differential operators:

$$
(8.1)\qquad (Lu)(z) = u_r(z) + a(x)u_x(z) + \frac{1}{2}\sigma^2(x)u_{xx}(z);
$$

$$
(8.2)\qquad (Gu)(z) = u_x(z) + u_y(z),
$$

where u_r, u_x, u_y are the first derivatives of u with respect to r, x, y respectively; and u_{xx} is the second derivative of u with respect to x. We construct the Hamiltonians by first letting

$$
(8.3)\qquad H_1(z, Du, D^2u) = h(r, x) + (Lu)(z); \quad H_2(z, Du) = f(r) - (Gu)(z);
$$

and then setting $H = -(H_1 \wedge H_2)$. The H-J-B equation is of the following form:

$$
(8.4)\qquad H(z, Du, D^2u) = 0.
$$

Next we introduce the definition of a viscosity solution due to Crandall-Lions [6] or Lions [16].

Definition 8.1. *A continuous function $u : E^\circ \to \mathbf{R}$ is called a viscosity subsolution (resp. supersolution) of the equation (8.4), if for all $\psi \in C^{1,2}([0, T] \times \mathbf{R}^2)$ for which $u - \psi$ has a maximum (resp. minimum) point at $z \in E^\circ$, we have*

$$
H(z, D\psi, D^2\psi) \leq 0; \quad (resp.\ \ H(z, D\psi, D^2\psi) \geq 0).
$$

u is called a viscosity solution of (8.4) if it is both a viscosity subsolution and a viscosity supersolution.

We first prove that the value function $Q(r, x, y)$ is a viscosity subsolution of (8.4), which comes from the following proposition.

Proposition 8.2. *Let $\psi \in C^{1,2}([0, T] \times \mathbf{R}^2)$ be such that $z \in E^\circ$ is a maximum point of $Q - \psi$ and such that $Q(z) = \psi(z)$. Then*

$$h(r, x) + (L\psi)(z) \geq 0; \qquad f(r) - (G\psi)(z) \geq 0. \tag{8.5}$$

Proof. By a simple stopping-time-argument, we may assume that $z \in E^\circ$ is a global maximum of $Q - \psi$. Thus $Q(z') \leq \psi(z')$, for all $z' \in E^\circ$. Let $\delta > 0$ be given; choose $P \in \mathcal{D}_{can}(z)$ so that $P\{\xi_t = 0,\ 0 \leq t \leq \delta\} = 1$. Define

$$\tau = \inf\{t \geq 0 : X_t^{x,0} = 0\}.$$

Since $x > 0$, $P(\tau > 0) = 1$. By the first equation of Bellman Principle (7.4) and the definition of ψ, one can easily derive that

$$\psi(z) \leq E^P \left\{ \int_0^{\delta\wedge\tau} h(r + t, X_t^{x,0})dt + \psi(r + \delta, X_{\delta\wedge\tau}^{x,0}, y) \right\}. \tag{8.6}$$

Denoting $Z_t = (r + t, X_t^{x,0}, y)$ and applying Itô's formula to $\psi(Z_t)$, we get

$$\begin{aligned} \psi(Z_{\delta\wedge\tau}) &= \psi(z) + \int_0^{\delta\wedge\tau} (L\psi)(Z_s)ds \\ &\quad + \int_0^{\delta\wedge\tau} \psi_x(Z_s)\sigma(X_s^{x,0})dB_s + \int_0^{\delta\wedge\tau} \psi_x(Z_s)dK_s^{x,0}. \end{aligned}$$

It is readily seen that $K_t^{x,0} = 0$ for $t \leq \tau$, $a.s.P$ and the stochastic integral on the right hand side above is a true martingale, thus upon taking expectation, we derive from (8.6) that $0 \leq [h(r, x) + (L\psi)(z)]\delta + o(\delta)$. Deviding δ on both sides and then letting $\delta \to 0$, we obtain the first half of (8.5).

To show the second half of (8.5), we use the second equation of Bellman Principle (7.5). Namely, for any $\delta < x \wedge y$, we have

$$\psi(r, x, y) = Q(r, x, y) \leq f(r)\delta + \psi(r, x - \delta, y - \delta). \tag{8.7}$$

Now applying Taylor's formula to the second term on the right hand side of (8.7) and then using the similar argument as before, we derive the second inequality of (8.5). □

We now turn to prove that Q is also a viscosity supersolution of (8.4). We fisrt give an assumption on the optimal control which is technically important for the rest of the section.

(C.5): There exist a constant $C > 0$, a stopping time $\tau \in \Im_{0,T-r}$ and an optimal control $P^* \in \mathcal{D}_{can}(r, x, y)$ such that $P^*\{\tau > 0\} = 1$, and that for all stopping time

$\tau' \in \Im_{0,T-r}$ such that $P^*\{0 < \tau' \le \tau\} = 1$, it holds that $P^*\{\xi^d_{\tau'} - \xi_{0+} < C\tau'\} = 1$, where $\xi^d_{\tau'} = \sum_{0 \le s < \tau'} \Delta\xi_s$. □

Remark. The condition (C.5) essentially says that $(\xi^d_t - \xi_{0+})/t = O(1)$, as $t \to 0$, uniformly for $P^* - a.e.\ \omega \in \Omega$. Clearly, this is trivially true if ξ_t is continuous for small $t > 0$. As a matter of fact, in almost all known solvable problems of this kind, the optimal controls are continuous except for an initial jump (see for example, [2], [14], [17], [20], among others), both for the linear systems and the nonlinear systems, so the condition (C.5) is always satisfied. However, under the present context, we need this condition *a priori*. □

The following Proposition leads to the conclusion that the value function Q is also a viscosity supersolution of (8.4).

Proposition 8.3. *Suppose that condition (C.5) holds. Let $\psi \in C^{1,2}([0,T] \times \mathbf{R}^2)$ be such that $(r,x,y) \in E^\circ$ is a minimum point of $Q - \psi$ and such that $Q(r,x,y) = \psi(r,x,y)$. Then one of the following must hold:*

$$h(r,x) + (L\psi)(z) \le 0; \qquad f(r) - (G\psi)(z) \le 0. \tag{8.8}$$

Proof. Suppose that the conclusion is not true. Then there must exist a constant $\epsilon > 0$, such that

$$h(r,x) + (L\psi)(z) > \epsilon; \quad f(r) - (G\psi)(z) > \epsilon. \tag{8.9}$$

Let us denote $F_\psi(z) = (h(r,x) + (L\psi)(z)) \wedge (f(r) - (G\psi)(z))$ and set

$$\mathcal{E} = \{z' \in E^\circ : F_\psi(z') > \frac{\epsilon}{2};\ |(G\psi)(z') - (G\psi)(z)| < \frac{\epsilon}{2C};\ x > 0\}, \tag{8.10}$$

where C is the constant given in Condition (C.5). Pick an optimal control $P^* \in \mathcal{D}_{can}(r,x,y)$ satisfying the condition (C.5); denote $Z_t = (r+t, X^{x,-\xi}_t, y - \xi_t),\ t \ge 0$, and define

$$\hat\tau = \inf\{t \ge 0 : Z_t \in \mathcal{E}^c\}, \tag{8.11}$$

we have the following lemma.

Lemma 8.4. *Suppose (8.9) holds. Then we have that*

(1) $P^*\{\Delta\xi_0 = 0\} = 1$, where ξ is the second component of the canonical process.

(2) $P^*\{\hat\tau > 0\} = 1$.

(3) P^*-almost surely, one has

$$\sum_{0 \le s < \hat\tau} [\psi(Z_{s+}) - \psi(Z_s) - \psi_x(Z_s)\Delta X_s - \psi_y(Z_s)\Delta Y_s] \ge -\frac{\epsilon}{C} \sum_{0 \le s < \hat\tau} \Delta\xi_s.$$

Proof. (i) First observe that, if we denote $z = (r,x,y)$, $z^u = (r, x-u, y-u)$, and $\theta^Q_z(u) \triangleq f(r)u + Q(z^u)$ for $0 \le u \le x \wedge y$, then the function $\theta^Q_z(\cdot)$ is nondecreasing.

Indeed, for $0 \le u < u' \le x \wedge y$ and $\Delta u = u' - u > 0$, we have from the second equation of Bellman principle (7.5) that

$$\begin{aligned} \theta_z^Q(u) &= uf(r) + Q(z^u) \le uf(r) + [\Delta u \cdot f(r) + Q(z^{u+\Delta u})] \\ &= u'f(r) + Q(z^{u'}) = \theta_z^Q(u'). \end{aligned}$$

Note that under $P^* \in \mathcal{D}_{can}(r,x,y)$, we have by definition that $P^*\{\Delta\xi_0 \le x \wedge y\} = 1$; moreover, if we let $\tau \equiv \delta > 0$ in the first equation of Bellman principle (7.4) and note that the "inf" sign on the right hand can be removed since P^* is optimal, then by letting $\delta \to 0$ in (7.4), we get that $Q(z) = E^{P^*}\{\theta_z^Q(\Delta\xi_0)\}$.

On the other hand, because the set $\mathcal{E}$ is open, we can find a $d > 0$ such that for all $0 \le u \le d$, $z^u \in \mathcal{E}$. Thus by the monotonicity of $\theta_z^Q(\cdot)$, the definition of the function ψ, and a little computation, we get

$$\begin{aligned} Q(z) &\ge E^{P^*}\{\theta_z^Q(\Delta\xi_0 \wedge d)\} \ge E^{P^*}\{\theta_z^\psi(\Delta\xi_0 \wedge d)\} \\ &\ge E^{P^*}\{\int_0^{\Delta\xi_0 \wedge d}[f(r) - (G\psi)(z^u)]du + \psi(r,x,y) \ge \frac{\epsilon}{2}E^{P^*}[\Delta\xi_0 \wedge d] + Q(z). \end{aligned}$$

The part (1) follows immediately.

(ii) Since ξ does not have any initial jump and $x > 0$, we must have $\hat{\tau} > 0$, $a.s.P^*$, proving part (2).

(iii) Observe that for any $0 \le s < \tau$, we must have $K_s^{x,-\xi} = 0$, thus $\Delta X_s = \Delta Y_s = -\Delta\xi_s$, so that $\psi_x(Z_s)\Delta X_s + \psi_y(Z_s)\Delta Y_s = -(G\psi)(Z_s)\Delta\xi_s$. On the other hand, if we denote $Z_t^\theta = Z_t + \theta(Z_{t+} - Z_t)$, $t \ge 0$, then it is readily seen that

$$\psi(Z_{s+}) - \psi(Z_s) = \int_0^1 \frac{d}{d\theta}\psi(Z_s^\theta)d\theta = \int_0^1 (G\psi)(Z_s^\theta)d\theta \cdot (-\Delta\xi_s).$$

Therefore, for each $0 \le s < \hat{\tau}$, we have

$$\psi(Z_{s+}) - \psi(Z_s) - \psi_x(Z_s)\Delta X_s - \psi_y(Z_s)\Delta Y_s = \Delta\xi_s \cdot \int_0^1 [(G\psi)(Z_s) - (G\psi)(Z_s^\theta)]d\theta.$$

Since $s < \hat{\tau}$, we have $Z_s, Z_s^\theta \in \mathcal{E}$ for all $\theta \in (0,1)$. Therefore, P^*-almost surely,

$$(G\psi)(Z_s) - (G\psi)(Z_s^\theta) \ \ \ge\ \ -|(G\psi)(Z_s) - (G\psi)(z)| - |(G\psi)(z) - (G\psi)(Z_s^\theta)| \ge -\frac{\epsilon}{C},$$

whence the result follows since $\Delta\xi_s \ge 0$, for all $s \ge 0$. □

We can now finish the proof of Proposition 8.3. Without loss of generality, we may assume that $\hat{\tau} = \tau$, where τ is the stopping time defined in Condition (C.5); for otherwise we can always consider $\tau' = \tau \wedge \hat{\tau}$.

Applying the generalized Itô's Formula to $\psi(Z_t)$:

$$\begin{aligned} \psi(Z_\tau) &= \psi(z) + \int_0^\tau (L\psi)(Z_s)ds - \int_0^\tau (G\psi)(Z_s)d\xi_t \\ &\quad + \int_0^\tau \psi_x(Z_s)\sigma(X_s)dB_s + \int_0^\tau \psi_x(Z_s)dK_s^{x,0} \\ &\quad + \sum_{0 \le s < \tau}[\psi(Z_{s+}) - \psi(Z_s) - \psi_x(Z_s)\Delta X_s - \psi_y(Z_s)\Delta Y_s]. \end{aligned} \tag{8.12}$$

By a similar argument as in Propositon 8.2, together with the definition of τ $(=\hat{\tau})$ and (8.10), (8.11), we see that in the right hand side of (8.12): P^*-almost surely, the first integral is no less than $-\int_0^\tau h(r+s, X_s^{x,-\xi})ds + \frac{\epsilon}{2}\tau$; the second integral is no less than $-\int_0^\tau f(r+s)d\xi_s + \frac{\epsilon}{2}\tau$; moreover, the fourth term is a true martingale and the fifth term should vanish. Thus, upon taking expectation and doing a little algebra, we obtain that

$$
\begin{aligned}
(8.13) \qquad & \psi(z) + E^{P^*}[\epsilon\tau + \Delta_\tau] \\
& \leq \; E^{P^*}\left\{\int_0^\tau h(r+s, X_s^{x,-\xi})ds + \int_0^\tau f(r+s)d\xi_s + \psi(Z_\tau)\right\},
\end{aligned}
$$

where $\Delta_\tau = \sum_{0\leq s<\tau}[\psi(Z_{s+}) - \psi(Z_s) - \psi_x(Z_s)\Delta X_s - \psi_y(Z_s)\Delta Y_s]$. Now recall that $z=(r,x,y)$ is the zero minimum point of $Q-\psi$, (8.13) will remain true if we replace ψ by Q and then the right hand side of (8.13) equals $Q(z)$ $(=\psi(z))$ by the first equation of Bellman principle and the fact that P^* is optimal. It follows that $E^{P^*}[\epsilon\tau+\Delta_\tau]\leq 0$. However, by Lemma 8.4 and the assumption (C.5), we should have $\Delta_\tau \geq -\frac{\epsilon}{C}\xi_\tau^d > -\epsilon\tau$, $a.s.P^*$, whence $E^{P^*}[\epsilon\tau+\Delta_\tau] > 0$. The contradiction yields the proposition. □

Remark. (1) From the proof of Proposition 8.3, we see that the proposition will remain true as long as the last term in (8.12) is nonnegative. One of the sufficient conditions for this is that *the value function Q is convex in the variables x and y* so that ψ can be modified to be convex near z so that the last term of (8.12) is nonnegative. In the linear case (a and σ are constant or linear in x), this condition is easy to be verified, but in our case, this is again not known *a priori*.

(2) One may have already noticed that we didn't mention the uniqueness of the viscosity of the H-J-B equation (8.4) (with terminal condition (5.5) and mixed boundary conditions (5.2) and (5.6)). In fact, because the domain $(0,\infty)\times(0,\infty)$ in (x,y)-space is not of smooth boundary, we have not found any existing result concerning this. Therefore we would like to raise this question to those who might be interested. □

Acknowledgments: I would like to thank Professor Naresh Jain for his valuable help during the time when most of this paper was accomplished. My thanks are also due to Professor Nicolai V. Krylov for many useful discussions.

9. Appendix 1. We now sketch the proof of Lemma 5.1.

(i) The proof of (1) is a modification of [11, Lemma 6.3]. Let $0<\theta<r<T$ and $P\in\mathcal{C}_{can}(r-\theta, y)$ be given. Let $X \stackrel{\Delta}{=} F_P^x$, $K \stackrel{\Delta}{=} G_P^x$ be the solution to SDEDR$(x,-\xi)$. Then

$$
Q(r,x,y) \leq E^P\left[\int_0^{T-r} h(r+t, X_t)dt + \int_0^{T-r} f(r+t)d\xi_t + g(X_{T-r})\right].
$$

whence a similar estimates as in [11, Lemma 6.3] leads to

$$
(9.1) \quad J(P, r-\theta, x, y) - Q(r,x,y) \geq E^P\left\{\int_{T-r}^{T-r+\theta} h(r-\theta+t, X_t)dt\right.
$$

$$
\begin{aligned}
&+\int_0^{T-r}[h(r-\theta+t,X_t)-h(r+t,X_t)]dt+\int_0^{T-r}[f(r-\theta+t)-f(r+t)]d\xi_t\Big\}\\
\geq\ & E^P\Big\{\int_{T-r}^{T-r+\theta}g'(X_s)[a(X_s)ds+\sigma(X_s)dB_s-d\xi_s+dK_s]\\
&+\int_{T-r}^{T-r+\theta}f(r-\theta+t)d\xi_t+\frac{1}{2}\int_0^{\infty}[\Lambda_{T-r+\theta}(a)-\Lambda_{T-r}(a)]dg'(a)\Big\}.
\end{aligned}
$$

where $\{\Lambda_t(a) : t\geq 0\}$ is the local time of the continuous semimartingale X at level a. The moment estimate of the solution X and the conditions on the functions h, f, g and a give that $E^P\int_{T-r}^{T-r+\theta}g'(X_s)\sigma(X_s)dB_s=0$; $\int_{T-r}^{T-r+\theta}[f(r-\theta+t)-g'(X_s)]d\xi_s\geq 0,\ \ a.s.$; $\int_{T-r}^{T-r+\theta}g'(X_s)dK_s=g'(0)[K_{T-r+\theta}-K_{T-r}]\geq 0,\ \ a.s.$; $\int_0^{\infty}[\Lambda_{T-r+\theta}(a)-\Lambda_{T-r}(a)]dg'(a)\geq 0,\ \ a.s.$. Thus

$$J(P,r-\theta,x,y)-Q(r,x,y)\geq -\theta\cdot C_M,$$

and whence

$$Q(r-\theta,x,y)-Q(r.x.y)\geq -\theta\cdot C_M. \tag{9.2}$$

Conversely, let $P\in\mathcal{C}_{can}(T-r,y)$ and note that $a.s.P$, the coordinate process ξ satisfies that $\xi_t=\xi_{T-r}\leq y$, thus $P\in\mathcal{C}_{can}(T-r+\theta,y)$. Following the same estimate as that in [11, Lemma 6.4], we can easily obtain that

$$Q(r-\theta,x,y)-J(P,r,x,y)\leq \theta\cdot C_M+E^P[g(X_{T-r+\theta})-g(X_{T-r})]. \tag{9.3}$$

Moreover, it is easily checked that for $t\in[T-r,T]$, (X,K) satisfies

$$X_t=X_{T-r}+\int_{T-r}^t a(X_s)ds+\int_{T-r}^t\sigma(X_s)dB_s+K_t-K_{T-r},$$

namely, X is a reflected diffusion process starting from X_{T-r}. It then easily follows that $E^P|X_{T-r+\theta}-X_{T-r}|^2$ is of order of θ, the conclusion of part (1) follows from the same argument as [11, Lemma 6.4, Corollary 6.5]

(ii) First note that for fixed $P\in\mathcal{C}_{can}(r,y)$, we have

$$J(P,r,x+\delta,y)\geq J(P,r,x,y),$$

for all $x\geq 0$, it follows that the function $Q(r,\cdot,y)$ is nondecreasing on $[0,\infty)$, the proof of part (2) is essentially the same as [11, Lemma 6.7, Corollary 6.9]. However, one should note here that the result of [11, Corollary 6.8] *can not* be easily adapted to nonlinear diffusion case, the basic difficulty is that the inequality (6.6) in [11], which is essential to derive the result of Corollary 6.9 in [11], is no longer true when the drift and diffusion coefficients are non-constant. Nevertheless, the rest of the proof is not hard to be adjusted to our case.

For an arbitrary $\delta > 0$ and $P \in \mathcal{D}_{can}(r,x,y) \subseteq \mathcal{D}_{can}(r,x+\delta,y)$, denote (X,K), $(X(\delta),K(\delta))$ to be the solution to the SDEDR$(x,-\xi)$ and SDEDR$(x+\delta,-\xi)$ respectively, and utilize all the notations in [11]. A similar estimate as Lemma 6.7 in [11], together with the comparison theorem of the SDEDR (see [18]), gives that

$$\begin{aligned} & Q(r,x+\delta,y) - J(P,r,x,y) \\ & \quad \le \quad E^P\left[\int_0^{(T-r)\wedge T_\delta} (X_t(\delta)-X_t)h_x(r+t,X_t(\delta))dt + (X_{T-r}(\delta)-X_{T-r})g'(X_{T-r}(\delta))1_{\{T-r<T_\delta\}}\right], \end{aligned}$$

where $T_\delta \triangleq \inf\{t \ge 0 : X_t \ge X_t(\delta)\}$. It is readily seen that the moment estimate

$$E^P|X(\delta)-X|_T^{*,2m} \le C_{m,T}\cdot\delta,$$

holds for all $m > 0$ and $\delta \in (0,1)$, where $C_{m,T}$ is a constant depending only on m, T. Thus part (2) follows right away.

(iii) We use the same method as in Lemma 6.10. Again, we note that the function $Q(r,x,\cdot)$ is non-increasing. Indeed, it is obvious that $\mathcal{D}_{can}(r,x,y) \subseteq \mathcal{D}_{can}(r,x,y')$, for any $y \le y'$, thus $Q(r,x,y) \ge Q(r,x,y')$, for $y \le y'$. We now prove the second inequality.

To begin with, let P be an arbitrary element in $\mathcal{D}_{can}(r,x,y+\epsilon)$ for some $\epsilon > 0$. Define the stopping time

$$\tau_y \triangleq (T-r)\wedge\inf\{t \ge 0 : \xi_t \ge y\},$$

where ξ is the second component of the canonical process. Set

$$\zeta_t = \begin{cases} \xi_t & 0 \le t \le \tau_y; \\ & \\ y & t > \tau_y. \end{cases}$$

It is obvious that ζ is nondecreasing and $\zeta_{T-r} \le y$, $a.s.P$. Further, the pathwise uniqueness of the solution of SDEDR implies that for any $0 \le t < \tau_y$, one has $X_t^{x,-\xi} = X_t^{x,-\zeta}$, $a.s.P$. Thus

$$\Delta\zeta_t = \Delta\xi_t \le X_t^{x,-\xi} = X_t^{x,-\zeta},$$

and for $t > \tau_y$, one has $\Delta\zeta_t \equiv 0$. As for $t = \tau_y$, note that $\Delta\zeta_{\tau_y} = y - \xi_{\tau_y}$, and $\xi_{\tau_y+} \ge y$ by the definiton. Hence, $\Delta\xi_{\tau_y} \le X_{\tau_y}^{x,-\xi}$ implies that

$$\Delta\zeta_{\tau_y} \le \Delta\xi_{\tau_y} \le X_{\tau_y}^{x,-\xi} = X_{\tau_y}^{x,-\zeta},$$

where the last equality is due to the càglàd property of the paths. Therefore, it holds that for any $t \ge 0$,

$$\Delta\zeta_t \le X_t^{x,-\zeta}, \qquad a.s.P. \tag{9.4}$$

Observe that, on the probability space $(\Omega, \mathcal{F}, P)$, we have

$$-\zeta_t - (-\xi_t) = \xi_t - \zeta_t = \begin{cases} 0, & 0 \le t \le \tau_y; \\ \\ \xi_t - y & t > \tau_y, \end{cases}$$

which is non-negative, non-decreasing. Note that the reflecting process $K^{x,-\xi}$ is continuous, the Comparison Theorem [18, Corollary 5.5], gives $X_t^{x,-\zeta} \ge X_t^{x,-\xi} \ge 0$, for all $t \ge 0$. Now let $\tilde{P} = P \circ (B, -\zeta)^{-1}$, we have $\tilde{P} \in \mathcal{D}_{can}(r,x,y)$, the conditions on the functions h, f, g lead to that

(9.5)
$$\begin{aligned} & Q(r,x,y) \le J(\tilde{P}, r, x, y) \\ = \;& E^P \Bigg[\int_0^{\tau_y} h(r+t, X_t^{x,-\xi}) dt + \int_{\tau_y}^{T-r} h(r+t, X_t^{x,-\zeta}) dt \\ & + \int_{[0,\tau_y]} f(r+t) d\xi_t + g(X_{T-r}^{x,-\xi}) 1_{\{\tau_y \ge T-r\}} + g(X_{T-r}^{x,-\zeta}) 1_{\{\tau_y < T-r\}} \Bigg]. \end{aligned}$$

(9.6)

In conjunction with the similar estimate as in [11, Lemma 6.10] and the Schwartz inequality, one can easily derive that

(9.7)
$$\begin{aligned} & Q(r,x,y) - J(P, r, x, y+\epsilon) \\ \le \;& \left\{ E^P |X^{x,-\xi} - X^{x,\zeta}|_{T-r}^{*,2} \right\}^{1/2} \cdot \left\{ E^P \left[\int_{\tau_y}^{T-r} \int_0^1 h_x(r+t, Z_t(\theta)) d\theta dt + \int_0^1 g'(Z_t(\theta)) d\theta \right]^2 \right\}^{1/2}, \end{aligned}$$

where $Z(\theta) = \theta X^{x,-\zeta} + (1-\theta) X^{x,-\xi}$ Let

$$\begin{aligned} Y_t^{x,-\xi} &= x + \int_0^t a(X_s^{x,-\xi}) ds + \int_0^t \sigma(X_s^{x,-\xi}) dB_s - \xi_t \\ Y_t^{x,-\zeta} &= x + \int_0^t a(X_s^{x,-\zeta}) ds + \int_0^t \sigma(X_s^{x,-\zeta}) dB_s - \zeta_t. \end{aligned}$$

Then

$$X^{x,-\xi} = \Gamma(Y^{x,-\xi}); \quad X^{x,\zeta} = \Gamma(Y^{x,-\zeta}),$$

where Γ is the solution mapping of DRP(I), by definition. However, the continuity of the reflecting process $K^{x,-\xi}$ and $K^{x,-\zeta}$ implies that Γ coincides with the solution mapping of DRP(I), so that it is Lipschitz continuous under the sup-norm. Therefore, a simple Gronwall- argument yields that

$$|X^{x,-\xi} - X^{x,-\zeta}|_{T-r}^{*,2} \le C_M |\xi - \zeta|_{T-r}^{*,2} \le C_M \cdot \epsilon, \quad t \ge 0, \quad a.s.P.$$

The consequence then follows easily from the moment estimate of the solutions to SDEDR, the conditions on the functions h, f, g, a, σ, and the inequlity (9.5). The proof is now completed. $\square$

10. Appendix 2. In this section we prove Lemma **??**. *i.e.*, the measure $\mathsf{P}^* \triangleq P\otimes_{\tau(\cdot)}\mathsf{Q}^*_\omega \in \mathcal{D}^{P,\tau}_{can}(r,x,y)$. Since $P\in\mathcal{D}_{can}(r,x,y)$ and $\mathsf{P}^*=P$ on $\mathcal{F}_\tau$ by the construction, $\mathsf{P}^*\in\mathcal{D}^{P,\tau}_{can}(r,x,y)$ provided $\mathsf{P}^*\in\mathcal{D}_{can}(r,x,y)$.

(i) We prove that the first component of the canonical process, $B_t(w,\zeta)=w(t),\ t\geq 0,\ \omega=(w,\zeta)\in\Omega$, is an $(\{\mathcal{F}_t\},\mathsf{P}^*)$-Brownian motion by checking that for any $0\leq s<t$ and $\theta\in\mathbf{R}$,

$$E^{P^*}[\exp[i\theta(B_t-B_s)]|\mathcal{F}_s]=\exp[-(t-s)\theta^2/2],\ a.s.P^* \tag{10.1}$$

It suffices to check that for any cylinder set $A=A^{t_1,\cdots,t_n}_{\gamma_1,\cdot,\gamma_n},\ 0\leq t_1\leq\cdots\leq t_n\leq s$, $\gamma_1,\cdots,\gamma_n\in\mathcal{B}_{\mathbf{R}}$, one has

$$E^{P^*}[\exp[i\theta(B_t-B_s):A]=\exp[-(t-s)\theta^2/2]P^*(A). \tag{10.2}$$

In the sequel, we will denote $P^\omega=\delta_\omega\otimes_{\tau(\omega)}\mathsf{Q}^*_\omega$ for simplicity. By definition of the probability P^*, we have

$$\begin{aligned} (10.3)\quad & E^{P^*}[\exp[i\theta(B_t(\omega')-B_s(\omega')]:A]\\ &=\int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_A(\omega')P^\omega(d\omega')P(d\omega)\\ &=\int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_A(\omega')[1_{\{\tau(\omega)<s\}}+1_{\{s\leq\tau(\omega)<t\}}+1_{\{t\leq\tau(\omega)\}}]P^\omega(d\omega')P(d\omega)\\ (10.4)\quad &= I_1+I_2+I_3, \end{aligned}$$

where

$$\begin{aligned} I_1&=\int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_A(\omega')1_{\{\tau(\omega)<s\}}P^\omega(d\omega')P(d\omega);\\ I_2&=\int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_A(\omega')1_{\{s\leq\tau(\omega)<t\}}P^\omega(d\omega')P(d\omega);\\ I_3&=\int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_A(\omega')1_{\{t\leq\tau(\omega)\}}P^\omega(d\omega')P(d\omega). \end{aligned}$$

For each $\omega\in\Omega$, there exists a $k(\omega)\in\{t_1,\cdots,t_n\}$, such that $\tau(\omega)\in(k(\omega),k(\omega)+1]$, so that $A=A_1(\omega)\cap A_2(\omega)$, where

$$\begin{aligned} A_1(\omega)&\triangleq\{\omega':\omega'(t_1)\in\gamma_1,\cdots,\omega'(t_{k(\omega)})\in\gamma_{k(\omega)}\};\\ A_2(\omega)&\triangleq\{\omega':\omega'(t_{k(\omega)+1})\in\gamma_{k(\omega)+1},\cdots,\omega'(t_n)\in\gamma_n\}. \end{aligned}$$

Thus

$$\begin{aligned} I_1 &= \int_\Omega\int_\Omega \exp[i\theta(B_t(\omega')-B_s(\omega')]1_{A_1(\omega)}(\omega')1_{A_2(\omega)}(\omega')1_{\{\tau(\omega)<s\}}P^\omega(d\omega')\\ &= \int_\Omega 1_{A_1(\omega)}(\omega')1_{\{\tau(\omega)<s\}}(\omega')\left[\int_{A_2(\omega)}\exp[i\theta(B_t(\omega')-B_s(\omega')]\mathsf{Q}^*_\omega(d\omega')\right]P(d\omega)\\ &= \exp[-\frac{\theta^2}{2}(t-s)]\int_\Omega 1_{A_1(\omega)}(\omega')1_{\{\tau(\omega)<s\}}(\omega')\mathsf{Q}^*_\omega(A_2(\omega))P(d\omega)\\ &= P^*(A\cap\{\tau<s\}), \end{aligned}$$

because $A_2 \cap \{\tau < s\} \in \mathcal{F}_\tau$. As for I_2, we have

$$\begin{aligned} I_2 &= \int_\Omega\int_\Omega \exp[i\theta(B_t(\omega') - B_s(\omega')]1_A(\omega')1_{\{s\le\tau(\omega)<t\}}P^\omega(d\omega') \\ &= \int_\Omega \left[\int_\Omega \exp[i\theta(B_t(\omega') - B_{\tau(\omega)}(\omega)]\mathsf{Q}^*_\omega(d\omega')\right] 1_A(\omega)1_{\{s\le\tau(\omega)<t\}} \exp[i\theta(B_{\tau(\omega)}(\omega) - B_s(\omega)]P(d\omega) \\ &= \int_\Omega \exp[-\frac{\theta^2}{2}(t-\tau(\omega))]1_A(\omega)1_{\{s\le\tau(\omega)<t\}} \exp[i\theta(B_{\tau(\omega)}(\omega) - B_s(\omega)]P(d\omega) \\ &= \exp[-\frac{\theta^2}{2}t]E^P\left[M_\tau \cdot \exp[-i\theta B_s : A\cap\{s<\tau\le t\}\right], \end{aligned}$$

where $M_t \triangleq \exp[i\theta B_t + \frac{\theta^2}{2}t]$ is an $(\{\mathcal{F}_t\}, P)$-martingale. By noting that both A and $\{\tau \le t\} \cap \{\tau > s\}$ are in $\mathcal{F}_s$, and applying the Optional Sampling Theorem, we get

$$\begin{aligned} I_2 &= \exp[-\frac{\theta^2}{2}t]E^P\left[M_\tau \cdot \exp[-i\theta B_s : A\cap\{s<\tau\le t\}\right] \\ &= \exp[-\frac{\theta^2}{2}t]E^P\left[M_s \cdot \exp[-i\theta B_s : A\cap\{s<\tau\le t\}\right] \\ &= \exp[-\frac{\theta^2}{2}t]\exp[\frac{\theta^2}{2}s]P\{A\cap\{s<\tau\le t\}\} \\ &= \exp[-\frac{\theta^2}{2}(t-s)]P\{A\cap\{s<\tau\le t\}\} \end{aligned}$$

Finally,

$$I_3 \;=$$

$\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega\omega$

REFERENCES

[1] E.N. Barron, R. Jensen and J.L. Menaldi, *Optimal Control and Differential Games with Measures*, Preprint, (1990).

[2] V. E. Beněs, L. A. Shepp and H. S. Witsenhausen, *Some Solvable Stochastic Control Problems*, Stochastics, **4** (1980), pp.39–83.

[3] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control, The Discrete Time Case,* Academic Press (1978).

[4] M. Chaleyat-Maurel, N. El Karoui and B. Marchal, *Réflexion discontinue et systèmes stochastiques,* Ann. Probab. **8** (1980), pp. 1049-1067.

[5] P. L. Chow, J. L. Menaldi and M. Robin, *Additive Control of Stochastic Linear Systems with Finite Horizon*, SIAM. J. Contr. and Optim., **23**, 6 (1985), pp. 858–899.

[6] M.G. Crandall and P.-L. Lions, *Viscosity Solutions of Hamilton-Jacobi Equations*, Transactions Amer. Math. Soc. **277** (1983), pp. 1-42.

[7] P. Dupuis and H. Ishii, *On Lipschitz Continuity of the Solution Mapping to the Skorohod Problem, with Applications,* Stochastics and Stochastic Reports, **35** (1991), pp. 31-62.

[8] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence,* John Wiley & Sons. Inc. (1986)

[9] J. M. Harrison and M. I. Taksar, *Instantaneous Control of Brownian Motion*, Math. of Oper. Res., **8**, 3 (1983), pp.439–453.

[10] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes,* North-Holland (1981).

[11] N. El Karoui and I. Karatzas, *Probabilistic Aspects of Finite-Fuel, Reflected Follower Problems,* Acta Applicandae Mathematicae **11** (1988), pp. 223-258.

[12] I. Karatzas, *A class of Singular Stochastic Control Problems*, Adv. Appl. Prob., **15** (1983), pp.225–254.

[13] I. Karatzas and S.E. Shreve, *Connections Between Optimal Stopping and Singular Stochastic Control II. Reflected Follower Problems*, SIAM J. Contr. and Optim., **23**, 3 (1985), pp.433–451.

[14] I. Karatzas and S.E. Shreve, *Equivalent Models for Finite-Fuel Stochastic Control,* Stochastics **18** (1986), pp. 245-276.

[15] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-verlag, New York, 1987.

[16] P.-L. Lions, *Optimal Control of Diffusion Procisses and Hamilton Jacobi Equations, part 2: viscosity solutions and uniqueness*, Comm. Partial. Diff. Equ., **11** (1983), 1229-1276.

[17] J. Ma, *On the Principle of Smooth Fit for A Class of Singular Stochastic Control Problems for Diffusions,* SIAM. J. Contr. and Optim., **30**, July (1992).

[18] J. Ma, *Discontinuous reflection, and a class of singular stochastic control problems for diffusions,* preprint (to appear in Stochastics), (1992).

[19] P. A. Meyer, *Un Cours Sur Les Intégrales Stochastiques*, Séminaire de Probabilités, X., Lecture Notes in Math. **511** (1976), Springer.

[20] S. E. Shreve, J. P. Lehoczky and D. P. Gaver, *Optimal Consumption for General Diffusions with Absorbing and Reflecting Barriers*, SIAM. J. Contr. and Optim., **22**, 1 (1984), pp.55–75.

[21] H. M. Soner and S. E. Shreve, *Regularity of the Value Function for a Two-Dimensional Singular Stochastic Control Problem*, SIAM. J. Contr. and Optim., **27**, 4 (1989), pp.876–907.

[22] D.W. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin-Heidelberg-New York (1979).