ON THE RATE OF CONVERGENCE
OF FINITE-DIFFERENCE APPROXIMATIONS
FOR BELLMAN EQUATIONS WITH CONSTANT COEFFICIENTS

HONGJIE DONG AND N. V. KRYLOV

ABSTRACT. Elliptic Bellman equations with coefficients independent of the variable x are considered. Error bounds for certain types of finite-difference schemes are obtained. These estimates are sharper than the earlier results in Krylov's article of 1997.

§1. INTRODUCTION

Our main purpose in this paper is to present some new estimates for the rate of convergence of finite-difference approximations in the problem of finding viscosity or probabilistic solutions of *degenerate* elliptic Bellman equations. Historically, the first estimates were obtained in [8] for the equations with "constant" coefficients. There, the convergence rate of order $h^{1/3}$, where h is the mesh size, was established in the case where the "free" term is Lipschitz continuous and the finite-difference approximations are monotone, translation invariant, and apart from that, almost arbitrary with the order of consistency h. It was also shown in [8] that in the general framework the order of accuracy cannot be better than $h^{1/2}$. After that, in [9] the results were extended to parabolic degenerate Bellman equations with variable coefficients. It was proved that if the data are Hölder $1/2$ continuous in the time variable and Lipschitz continuous in the space variables, then, again, the approximation error for quite arbitrary finite-difference approximations admits an estimate of order of accuracy $h^{1/3}$ from one side and of order of accuracy $h^{1/21}$ from the other. Also in [9], some approximations similar to finite-difference ones were suggested with the order of accuracy not less than $h^{1/3}$; see, e.g., Theorem 5.7 in [9]. This theorem is close to Theorem 3.5 in [1], where in a particular elliptic situation the error bound of order of accuracy $h^{1/2}$ was obtained on the account of a special approximation. The authors of [1] and [2] did a very good job of surveying the literature related to finite-difference approximations for the Hamilton–Jacobi and Bellman equations; instead of copying their comments, we allow ourselves to refer to the interested reader to [1] and [2] for that information.

One of the purposes in the papers [9] and [10] (the latter is the basis for the former) was not only to give the rates of convergence, but also prove the convergence itself. In a subsequent paper, the second author intends to establish such convergence for general approximation schemes in the case where the limit function is *not* a viscosity solution of the corresponding Bellman equation.

One of the main ideas in [8, 9] is that the original Bellman equation and the approximate finite-difference one should play symmetric roles. Another idea is to "shake" the coefficients if they depend on x in both the approximate finite-difference equation and the Bellman equation. Both ideas were also used in the elliptic case in [1, 2], and better results than those in [8, 9] were obtained for special finite-difference schemes. One of the main results of [2] is that the order of approximation is not less than $h^{1/5}$.

In this paper we also investigate what happens if instead of general consistent finite-difference approximations as in [8, 9], where we were partly aimed at proving mere convergence, we take those that better suit finding solutions numerically. In our view, the first natural step in this direction is to investigate equations with "constant" coefficients. We use the standard symmetric approximation for the second-order derivatives and two approximations for the first derivative: the monotone one, which depends on the sign of the coefficient of the derivative, and the symmetric one such as $[u(x + h) - u(x - h)]/(2h)$. We also assume better regularity, namely, the $\mathcal{C}^{1,1}$-regularity of the free terms, and prove that, with monotone approximation of the first-order term, the error bound is of order of accuracy h (Theorem 2.8) and not better (Example 3.2).

Our methods are very different from those in [1, 2], which are based on the theory of viscosity solutions, and it would be very interesting to understand whether our results can be obtained on the basis of that theory. In Example 3.2 the proof of the sharpness of Theorem 2.8 is based on the fact that the first-order coefficient is not zero. We do not know anything about sharpness if it is identically zero; probably yet other methods are needed to clarify the situation.

By assuming more structure on the equation and using the symmetric approximation of the first derivatives we obtain the error bound of order of accuracy h^2 (Theorem 2.11) and show that it is sharp (Example 3.5).

As far as we understand, our results are stronger than those in [1, 2]; however, they are only proved for equations with constant "coefficients". It should also be noted that our Theorem 2.9 may look like Theorem 3.5 in [1]. Yet the latter treats a different kind of approximations, such as Theorem 5.7 in [9].

All our main results (mentioned above) are stated in §2. We discuss them in §3. §4 contains general auxiliary results, part of which shows that the second-order differences of generally not very smooth functions interact with mollification much like the second-order derivatives do. The proof of Theorem 2.8 is given in §5 and Theorems 2.9 and 2.11 are proved in the final §6.

To conclude the Introduction, we set up some notation: $\mathbb{R}^d$ is a d-dimensional Euclidean space with a fixed standard orthonormal basis $e_1, e_2, \ldots, e_d$; $x = (x^1, x^2, \ldots, x^d)$ is a typical point in $\mathbb{R}^d$ and $(x, y) = x^i y^i$ is the inner product for $x, y \in \mathbb{R}^d$ (as usual, the summation convention over repeated indices is enforced). For any $l = (l^1, l^2, \ldots, l^d) \in \mathbb{R}^d$ and any differentiable function u on $\mathbb{R}^d$, we denote $D_l u = u_{x^i} l^i$ and $D_l^2 u = u_{x^i x^j} l^i l^j$, etc. In general, various constants are denoted by N, and the expression $N = N(\cdots)$ means that the constant N in question depends only on what is listed in the parentheses.

§2. The setting and main results

Let A be a separable metric space (the set of all admissible controls), let $d_1, d \geq 1$ be integers, and let $\ell_1, \ldots, \ell_{d_1}$ be nonzero vectors in $\mathbb{R}^d$. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, and let $\{\mathcal{F}_t; t \geq 0\}$ be an increasing filtration of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ that are complete with respect to $\mathcal{F}, P$. We denote by $\mathfrak{A}$ the set of all A-valued $\mathcal{F}_t$-adapted processes measurable with respect to the product of the Borel σ-algebra on $(0, \infty)$ and $\mathcal{F}$. Assume that a d_1-dimensional Wiener process w_t is defined on $(\Omega, \mathcal{F}, P)$ for $t \geq 0$. We suppose that w_t is a Wiener process with respect to $\mathcal{F}_t$.

Suppose that on A we are given continuous real-valued functions $\sigma_k(\alpha)$, $b_k(\alpha)$, $k = 1, \ldots, d_1$, and $c(\alpha)$. We introduce functions $\sigma(\alpha)$, $a(\alpha)$, and $b(\alpha)$, taking values in the sets of $(d \times d_1)$- and $(d \times d)$-matrices and in $\mathbb{R}^d$, respectively, by

$$\sigma^{ik}(\alpha) = \ell_k^i \sigma_k(\alpha), \quad \sigma(\alpha) = (\sigma^{ik}(\alpha)), \quad a = (1/2)\sigma\sigma^*, \quad b(\alpha) = \ell_r b_r(\alpha),$$

with no summation over k. Let $f(\alpha, x)$ be a real-valued function on $A \times \mathbb{R}^d$.

For a number $\gamma \in (0, 1]$ and a function g given on a domain $G \subset \mathbb{R}^d$, we set

$$[g]_{0,\gamma,G} := \sup_{x,y \in G} \frac{|g(x) - g(y)|}{|x - y|^\gamma}, \quad |g|_{0,G} := \sup_{x \in G} |g(x)|,$$

$$|g|_{0,\gamma,G} := |g|_{0,G} + [g]_{0,\gamma,G}$$

and denote by $\mathcal{C}^{0,\gamma}(G)$ the space of functions with finite norm $|\cdot|_{0,\gamma,G}$. For any integer $n \geq 1$, we let $\mathcal{C}^n(G)$ and $\mathcal{C}^{n,\gamma}(G)$ be the spaces of n times continuously differentiable functions g on G with finite norms $|g|_{n,G}$ and $|g|_{n,\gamma,G}$, respectively, where

$$|g|_{n,G} := \sum_{i=0}^n [g]_{i,G}, \quad [g]_{i,G} := |D^i g|_{0,G}, \quad |g|_{n,\gamma,G} := [D^n g]_{0,\gamma,G} + |g|_{n,G}.$$

If $G = \mathbb{R}^d$, we drop the symbol G in the notation of spaces and norms, so that, for instance, $\mathcal{C}^{n,\gamma} := \mathcal{C}^{n,\gamma}(\mathbb{R}^d)$ and $|\cdot|_{n,\gamma} := |\cdot|_{n,\gamma,\mathbb{R}^d}$.

Assumption 2.1. *For some numbers $K \geq 1$ and $\kappa > 0$, for any $\alpha \in A$ we have*

$$(2.1) \qquad \sum_{k=1}^{d_1} (|\ell_k| + |\sigma_k(\alpha)|^2 + |b_k(\alpha)|) + |c(\alpha)| + |f(\alpha, \cdot)|_{1,1} \leq K, \quad c(\alpha) \leq -\kappa.$$

Denote

$$L^\alpha u(x) = a^{ij}(\alpha) u_{x^i x^j}(x) + b^i(\alpha) u_{x^i}(x),$$

$$(2.2)$$

$$F(u_{ij}, u_i, u, x) = \sup_{\alpha \in A} \{a^{ij}(\alpha) u_{ij} + b^i(\alpha) u_i + c(\alpha) u + f(\alpha, x)\}.$$

Observe that if

$$a_k(\alpha) := (1/2)|\sigma_k(\alpha)|^2,$$

then

$$(2.3) \qquad a^{ij}(\alpha) u_{x^i x^j} = a_k(\alpha) D_{\ell_k}^2 u,$$

so that L^α may appear to have a very special form. The fact that, actually, L^α is a more or less general operator is explained in Remark 3.1.

We are interested in the following Bellman equation:

$$(2.4) \qquad F(u_{x^i x^j}(x), u_{x^i}(x), u(x), x) = 0.$$

One of the revelations of (2.4) is the following system characterizing an obstacle problem:

$$\Delta u - u \leq 0, \quad u \geq g, \quad \Delta u - u = 0 \text{ on the set } \{u > g\}.$$

Indeed, the above system is equivalent to one equation:

$$\sup_{\alpha \in [0,1]} \{\alpha \Delta u - u + (1 - \alpha)g\} = 0.$$

Therefore, if $|g|_{1,1} < \infty$, then all the results below are applicable.

We know (see, e.g., [3]) that under the above conditions there is a unique bounded viscosity solution v of (2.4), which coincides with the probabilistic one given by

$$(2.5) \qquad v(x) = \sup_{\alpha \in \mathfrak{A}} E \int_0^\infty f(\alpha_t, x_t^\alpha(x)) \exp\left(\int_0^t c(\alpha_s)\, ds\right) dt,$$

where

$$x_t^\alpha(x) = x + \int_0^t \sigma_k(\alpha_s)\ell_k \, dw_s^k + \int_0^t b_k(\alpha_s)\ell_k \, ds.$$

Finding v is one of the optimization problems in the theory of controlled diffusion processes.

By Assumption 2.1, we have

$$\Big| \int_0^\infty f(\alpha_t, x_t^\alpha(x)) \exp\Big(\int_0^t c(\alpha) \, ds \Big) \, dt \Big| \le \int_0^\infty K e^{-\kappa t} \, dt = \kappa^{-1} K.$$

Thus, $|v|_0 \le K/\kappa$.

We do not want to allow our controlled process to drift in some directions without diffusion along the same directions. Therefore, we make the following assumption, which is automatically satisfied if $b \equiv 0$.

Assumption 2.2. *If $k \in \{1, \ldots, d_1\}$, $\alpha \in A$, and $b_k(\alpha) \ne 0$, then $\sup_\beta |\sigma_k(\beta)| > 0$.*

In Theorems 2.8 and 2.9 we use the following finite-difference approximations. For every $h > 0$ and $l \in \mathbb{R}^d$, we define the first-order and the second-order differences in the direction l with step size h:

$$\Delta_{h,l}^+ u(x) = h^{-1}(u(x+hl) - u(x)),$$
$$\Delta_{h,l}^- u(x) = h^{-1}(u(x) - u(x-hl)),$$
$$\Delta_{h,l}^2 u(x) = h^{-2}(u(x+hl) - 2u(x) + u(x-hl)).$$

Let $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$ be the set of all bounded functions on $\mathbb{R}^d$. For every $h > 0$ and every $\alpha \in A$, we introduce a bounded linear operator $L_h^\alpha : \mathcal{B} \to \mathcal{B}$ as follows. Let $a^\pm = (|a| \pm a)/2$; we put

$$(2.6) \qquad L_h^\alpha u = a_k(\alpha)\Delta_{h,\ell_k}^2 u + b_k^+(\alpha)\Delta_{h,\ell_k}^+ u - b_k^-(\alpha)\Delta_{h,\ell_k}^- u.$$

The finite-difference approximations of v which we have in mind will be introduced by means of the equation

$$(2.7) \qquad F_h[u](x) := \sup_{\alpha \in A}\{L_h^\alpha u(x) + c(\alpha)u(x) + f(\alpha, x)\} = 0, \quad x \in \mathbb{R}^d.$$

Few straightforward properties of the above objects will be used.

Lemma 2.3. *For any $h > 0$ and any $\alpha \in A$ we have $L_h^\alpha 1 = 0$. Furthermore, denote $p_h = K(h^{-2} + h^{-1})$. Then for any $h > 0$ and any $\alpha \in A$ the operator $u \to L_h^\alpha u + p_h u$ is monotone, by which we mean that if $u_1, u_2 \in \mathcal{B}$ and $u_1 \le u_2$, then*

$$L_h^\alpha u_1 + p_h u_1 \le L_h^\alpha u_2 + p_h u_2.$$

Next, we restate some results of [8] in the following three lemmas based on Lemma 2.3. The first one gives the existence and uniqueness of solutions of (2.7). The second plays the role of the comparison principle for finite-difference schemes.

Lemma 2.4. *For any $h > 0$ and any bounded function $g(\alpha, x)$ on $A \times \mathbb{R}^d$ there exists a unique bounded solution of equation (2.7) with g in place of f. Moreover, if we denote by $v_h = v_h(x)$ the solution of (2.7), then*

$$(2.8) \qquad [v_h]_1 \le \kappa^{-1}[f]_1 \le K\kappa^{-1}.$$

Lemma 2.5. (i) *If $\varphi(x)$ is a bounded function and*

$$(2.9) \qquad F_h[\varphi](x) \le (\ge)\, 0, \quad x \in \mathbb{R}^d,$$

then $\varphi(x) \ge (\le)\, v_h$ on $\mathbb{R}^d$.

(ii) *If $g_1(\alpha, x)$, $g_2(\alpha, x)$, u_1, and u_2 are bounded functions, C is a constant, and*

(2.10)
$$\sup_{\alpha \in A}\{L_h^\alpha[u_1](x) + c(\alpha)u_1(x) + g_1(\alpha, x)\} + C$$
$$\geq \sup_{\alpha \in A}\{L_h^\alpha[u_2](x) + c(\alpha)u_2(x) + g_2(\alpha, x)\}$$

for all $x \in \mathbb{R}^d$, then

(2.11)
$$u_1 - u_2 \leq \kappa^{-1}\sup_{\alpha,x}[g_1(\alpha, x) - g_2(\alpha, x)]^+ + \kappa^{-1}C^+.$$

Remark 2.6. Obviously, we have $F_h(-\kappa^{-1}K) \geq 0$ and $F_h(\kappa^{-1}K) \leq 0$. Therefore, $|v_h|_0 \leq \kappa^{-1}K$.

Lemma 2.7. *The same estimate as in (2.8) is valid for v in place of v_h, i.e.,*

(2.12)
$$[v]_1 \leq \kappa^{-1}[f]_1 \leq K\kappa^{-1}.$$

Here are two main results of this paper.

Theorem 2.8. *There is a constant N depending only on d, d_1, K, κ, ℓ_k, and $a_k(\alpha)$ and such that for any $x \in \mathbb{R}^d$ and any $h > 0$ we have*

(2.13)
$$|v_h(x) - v(x)| \leq Nh.$$

Theorem 2.9. *Suppose that Assumption 2.1 is fulfilled with $|f|_{0,1}$ in place of $|f|_{1,1}$. Then there is a constant N depending only on d, d_1, K, κ, ℓ_k, and $a_k(\alpha)$ and such that for any $x \in \mathbb{R}^d$ and any $h > 0$ we have*

(2.14)
$$|v_h(x) - v(x)| \leq Nh^{1/2}.$$

In the case where F is smooth we have a better estimate. To state the result we need an assumption stronger than Assumption 2.2, and we change our notation somewhat.

Assumption 2.10. *We have $|b_k(\alpha)| \leq Ka_k(\alpha)$ for all $\alpha \in A$ and $k = 1, \ldots, d_1$.*

This time we introduce the operators L_h^α by a formula different from (2.6):

(2.15)
$$L_h^\alpha u(x) = a_k(\alpha)\Delta_{h,\ell_k}^2 u(x) + b_k(\alpha)\Delta_{2h,\ell_k}^+ u(x - h\ell_k).$$

As above, Lemmas 2.3, 2.4, 2.5, and 2.7 remain true but this time only if $Kh \leq 2$. Of course, in Lemmas 2.4 and 2.7 and in Theorem 2.11 below, by v_h we mean a unique bounded solution of (2.7) with the new operators L_h^α.

Theorem 2.11. *Let $\gamma \in (0, 1)$ be a number. Suppose there is a function $\Phi(u_{ij}, u_i, u, x)$ on $\mathbb{R}^{d_0}$, where $d_0 = d^2 + 2d + 1$, such that*

(i) $\Phi(u_{ij}, u_i, u, x) = \Phi(u_{ji}, u_i, u, x)$;

(ii) $\Phi(u_{ij}, u_i, u, x) = 0$ *on the surface*

$$\Gamma := \{(u_{ij}, u_i, u, x) : F(u_{ij}, u_i, u, x) = 0\} \subset \mathbb{R}^{d_0};$$

(iii) *for each $R \in [0, \infty)$ we have $\Phi \in \mathcal{C}^{2,\gamma}(G_R)$, where $G_R = U_R \times \mathbb{R}^d$,*

(2.16)
$$U_R = \Big\{(u_{ij}, u_i, u) \in \mathbb{R}^{d_0'} : \sum_{ij}|u_{ij}| + \sum_i|u_i| + |u| \leq R\Big\}, \quad d_0' = d_0 - d;$$

(iv) *for each $R \in [0, \infty)$ there is a constant $\mu(R) > 0$ such that for any $(u_{ij}, u_i, u, x) \in \Gamma \cap G_R$ and any unit $\lambda \in \mathbb{R}^d$ we have*

(2.17)
$$\Phi_{u_{kr}}(u_{ij}, u_i, u, x)\lambda^k\lambda^r \geq \mu(R).$$

Then there is a constant N such that

$$|v - v_h|_0 \leq Nh^2$$

for all $0 < h \leq 2/K$.

Remark 2.12. Similar results for the Bellman equations with "inf" in place of "sup" can be obtained by replacing u with $-u$, with the use of the relation $\sup\{-\cdots\} = -\inf\{\cdots\}$.

§3. Discussion and examples

Remark 3.1. One may think that considering the operators L^α written in the form $a_k(\alpha)D^2_{\ell_k} + b_k(\alpha)D_{\ell_k}$ is a severe restriction. However, it turns out that if a finite subset $B \subset \mathbb{Z}^d$ such that $\operatorname{Span} B = \mathbb{R}^d$ is fixed, and if an operator $Lu = a^{ij}u_{x^i x^j} + b^i u_{x^i}$ admits a monotone finite-difference approximation

$$L_h u(x) = \sum_{y \in B} p_h(y)u(x + hy),$$

then automatically

$$L = \sum_{\substack{l \in B \\ l \neq 0}} a_l D_l^2 + \sum_{\substack{l \in B \\ l \neq 0}} b_l D_l$$

for some $a_l \geq 0$ and $b_l \in \mathbb{R}$.

To be more precise, assume that for any smooth $u(x)$ we have

(3.1) $$Lu(0) = \lim_{h \downarrow 0} L_h u(0)$$

and $L_h u(0) < 0$ if u has a strict maximum at 0. The latter condition simply means that $0 \in B$, $p_h(0) < 0$, and $p_h(y) \geq 0$ for $y \in B \setminus \{0\}$.

The term $b^i u_{x^i}$ admits the above representation because $\operatorname{Span} B = \mathbb{R}^d$. Therefore, it suffices to deal with $a^{ij}u_{x^i x^j}$, assuming that $a^{ij} \not\equiv 0$, so that

$$\bar{a} := \operatorname{Trace} a > 0.$$

Let $B = \{0, l_1, \ldots, l_m\}$, where $|l_k| > 0$, and observe that for $v(x) = u(x) + u(-x)$ we have

$$L_h v(0) = \sum_{k=1}^m p_h(l_k)[u(hl_k) + u(-hl_k)] + 2p_h(0)u(0).$$

Furthermore,

$$0 = L1 = \lim_{h \downarrow 0} L_h 1 = \lim_{h \downarrow 0}\left[\sum_{k=1}^m p_h(l_k) + p_h(0)\right],$$

so that

$$L_h v(0) = \sum_{k=1}^m p_h(l_k)[u(hl_k) - 2u(0) + u(-hl_k)] + o(1)$$

as $h \downarrow 0$. It follows that

$$2a^{ij}u_{x^i x^j}(0) = \lim_{h \downarrow 0}\sum_{k=1}^m h^2 p_h(l_k)\Delta^2_{h, l_k} u(0).$$

In particular, for $u(x) = |x|^2$,

$$\frac{1}{\bar{p}_h} := \sum_{k=1}^m h^2 p_h(l_k)|l_k|^2 \to 2\bar{a} > 0.$$

Also, obviously, the terms $h^2 p_h(l_k)\bar{p}_h$ are bounded in h, and along a sequence $h_n \downarrow 0$ they converge to some $\bar{a}_k \geq 0$. Thus, we see that

$$a^{ij}u_{x^i x^j}(0) = \bar{a}\sum_{k=1}^m \bar{a}_k D^2_{l_k} u(0),$$

which proves the claim.

Example 3.2. We show that the result of Theorem 2.8 is sharp even if A is a singleton, so that F is linear. Let $d = 1$, $Lu = L^\alpha u = u_{xx} + u_x$, $c = -1$, and $f(\alpha, x) = \cos x$. Then $v(x) = \text{Re}(\beta \exp(ix))$ is the solution of $Lu - u + f = 0$ for

$$\beta = \frac{1}{2 - i}.$$

To find the solution v_h of the equation $L_h v_h - v_h + \cos x = 0$, we represent v_h as the real part of the solution of $L_h u - u + \exp(ix) = 0$. The latter equation can be solved if we try $u(x) = \beta \exp(ix)$. It turns out that $v_h(x) = \text{Re}(\beta_h \exp(ix))$, where

$$\beta_h = \left(1 + 2\frac{1 - \cos h}{h^2} + \frac{1 - e^{ih}}{h}\right)^{-1}.$$

Not surprisingly, $\beta_h \to \beta$ as $h \downarrow 0$. It follows that

$$\lim_{h \downarrow 0} \frac{v_h(0) - v(0)}{h} = \frac{d}{dh} \text{Re}\,\beta_h\big|_{h=0} = \text{Re}\,\frac{1}{2(2 - i)^2} = \frac{3}{50},$$

whence $|v(0) - v_h(0)| \sim 3h/50$ as $h \downarrow 0$. Therefore, Theorem 2.8 is sharp.

Remark 3.3. In Example 3.2, the reason why we only have the first order of convergence is that $b \neq 0$. Admittedly, we do not know anything about the sharpness of Theorem 2.8 if $b \equiv 0$.

Remark 3.4. In Theorem 2.11 we can take $\Phi = F$ if F is smooth and the controlled process is uniformly nondegenerate, that is, if there exists a constant $\delta > 0$ such that

(3.2) $$(a(\alpha)\lambda, \lambda) \geq \delta|\lambda|^2 \quad \text{for any } \alpha \in A \text{ and any } \lambda \in \mathbb{R}^d.$$

Indeed,

$$F_{u_{ij}}(u_{ij}, u_i, u, x)\lambda^i\lambda^j = \lim_{t \downarrow 0} \frac{1}{t}[F(u_{ij} + t\lambda^i\lambda^j, u_i, u, x) - F(u_{ij}, u_i, u, x)]$$
$$\geq \inf_{\alpha \in A}(a(\alpha)\lambda, \lambda).$$

In Example 3.8 we shall see that, sometimes, Theorem 2.11 is still applicable even if condition (3.2) is violated.

Example 3.5. We show that the result of Theorem 2.11 is sharp, and again even for linear equations. Let $d = 1$; we take $L^\alpha u = u_{xx}$, $c(\alpha) = -1$, and $f(\alpha, x) = f(x) = \cos x$. The corresponding linear second-order equation can be solved easily: $v(x) = (1/2)\cos x$. The solution v_h of the finite-difference equation can be found in the same way as in Example 3.2. It turns out that $v_h(x) = \beta_h \cos x$, where

$$\beta_h = \left(1 + 2\frac{1 - \cos h}{h^2}\right)^{-1}.$$

We immediately get $|v(0) - v_h(0)| \sim h^2/48$ as $h \downarrow 0$, which shows that Theorem 2.11 is sharp indeed.

Remark 3.6. Example 3.5 and Remark 3.1 basically show that approximations of order higher than h^2 cannot be obtained by using monotone schemes.

Before passing to the next example, we state and prove the following lemma, in which assertion (i) is a very particular case of some results in [7].

Lemma 3.7. (i) *Let* f, g, u_1, u_2 *be real numbers such that* $f \geq 0$ *and* $g \geq 0$. *Then the set of conditions*

(3.3) $$u_1 u_2 = f^2 + g(u_1 + u_2), \quad u_1, u_2 \geq g$$

is equivalent to

$$(3.4) \qquad \inf_{\alpha \in A} (\alpha_1 u_1 + \alpha_2 u_2 - 2\sqrt{\alpha_1 \alpha_2}\sqrt{f^2 + g^2} - g) = 0,$$

where $A = \{\alpha = (\alpha_1, \alpha_2) : \alpha_i \geq 0, \alpha_1 + \alpha_2 = 1\}$.
  (ii) *If $f, g \in \mathcal{C}^{1,1}$ and $f, g \geq 0$, then $\sqrt{f^2 + g^2} \in \mathcal{C}^{1,1}$.*

Proof. (i) Relation (3.3) can be written as

$$v_1 v_2 = f^2 + g^2, \quad v_i \geq 0,$$

where $v_i = u_i - g$. It is well known that the latter equation is equivalent to

$$\inf_{\alpha \in A} (\alpha_1 v_1 + \alpha_2 v_2 - 2\sqrt{\alpha_1 \alpha_2}\sqrt{f^2 + g^2}) = 0,$$

which is (3.4).

(ii) Statement (ii) is proved by approximating f and g with positive functions $f_n, g_n \in \mathcal{C}^2$ and by obtaining uniform estimates of the $\mathcal{C}^{1,1}$-norm of $\sqrt{f_n^2 + g_n^2}$. The latter is done by straightforward differentiating, with the use of the well-known pointwise inequality $|u_{x^i}(x)| \leq 2\sqrt{u(x)}\sqrt{|u_{x^i x^i}|_0}$, which is valid for nonnegative functions in $\mathcal{C}^2$.
  The lemma is proved. $\qquad \square$

Example 3.8. We give an example of using Theorem 2.11, which shows the advantage of imposing condition (2.17) only on $\Gamma \cap G_R$. For $d = 2$, consider the following equation, which is similar to one in a series of Monge–Ampère equations:

$$(3.5) \qquad (v_{x^1 x^1} - v)(v_{x^2 x^2} - v) = f^2(x) + g(x)(\Delta v - 2v).$$

We seek a solution of (3.5) in the class of functions $v \in \mathcal{C}^{1,1}$ such that $v_{x^i x^i} - v \geq g$ a.e., $i = 1, 2$, assuming that $f, g \in \mathcal{C}^{1,1}$, $f, g \geq 0$.

Then Lemma 3.7 and our previous results (see, in particular, Remark 2.12) guarantee that such a solution exists and is unique; furthermore, for any $h > 0$ the finite-difference equation

$$(3.6) \qquad (\Delta^2_{h,e_1} v_h - v_h)(\Delta^2_{h,e_2} v_h - v_h) = f^2(x) + g(x)(\Delta^2_{h,e_1} v_h + \Delta^2_{h,e_2} v_h - 2v_h)$$

has a unique solution v_h in the class of bounded functions such that $\Delta^2_{h,e_i} v_h - v_h \geq g$, $i = 1, 2$. Theorem 2.8 implies that

$$(3.7) \qquad |v - v_h|_0 \leq Nh$$

for any $h > 0$. Now assume that for some constants $\gamma, \varepsilon \in (0, 1)$ we have

$$f, g \in \mathcal{C}^{2,\gamma}, \quad f^2 + g^2 \geq \varepsilon.$$

We introduce

$$\Phi(u_{ij}, u_i, u, x) = (u_{11} - u)(u_{22} - u) - f^2(x) - g(x)(u_{11} + u_{22} - 2u).$$

Then, obviously, assumptions (i)–(iii) of Theorem 2.11 are satisfied. To check the remaining assumption (iv), observe that

$$\Phi_{u_{kr}}(u_{ij}, u_i, u, x)\lambda^k \lambda^r = (u_{22} - u - g)(\lambda^1)^2 + (u_{11} - u - g)(\lambda^2)^2.$$

As was mentioned above, on Γ we have

$$u_{11} - u \geq g, \quad u_{22} - u \geq g, \quad (u_{22} - u)(u_{22} - u) = f^2 + g(u_{11} + u_{22} - 2u),$$

$$(u_{22} - u - g)(u_{22} - u - g) = f^2 + g^2 \geq \varepsilon.$$

Moreover, if $(u_{ij}, u_i, u) \in U_R$, then $|u_{ii} - u - g| \leq 2R + |g|_0$, and the above relations imply that

$$u_{11} - u - g \geq \varepsilon/(2R + |g|_0), \quad u_{22} - u - g \geq \varepsilon/(2R + |g|_0).$$

It follows that

$$\Phi_{u_{kr}}(u_{ij}, u_i, u, x)\lambda^k\lambda^r \geq |\lambda|^2\varepsilon/(2R + |g|_0), \tag{3.8}$$

so that assumption (iv) of Theorem 2.11 is satisfied with $\mu(R) = \varepsilon/(2R + |g|_0)$, and this theorem yields

$$|v - v_h|_0 \leq Nh^2.$$

We emphasize that (3.8) is false for very many (u_{ij}, u_i, u, x) outside of $\Gamma \cap G_R$.

Remark 3.9. Example 3.12 in [7] treats equation with $g \equiv 0$ for $d \geq 2$, and estimate (3.7) is asserted for $f \in \mathcal{C}^{0,1}$, $f \geq 0$. However, there is an arithmetical error in the argument in that example, so that, actually, h in the corresponding version of (3.7) in [7] should be replaced with $h^{1/2}$. It is still better than the typical $h^{1/3}$ of [7] and agrees well with Theorem 2.9.

§4. SOME AUXILIARY RESULTS

Let $\mathcal{S}$ denote the set of real-valued infinitely differentiable functions $\zeta(x)$ on $\mathbb{R}^d$ that tend to zero together with each derivative as $|x| \to \infty$ faster than $|x|^{-n}$ for any $n \geq 0$.

For $\zeta \in \mathcal{S}$, a bounded function u, and $\varepsilon > 0$, we define the mollification of u by

$$u^{(\varepsilon)} = \zeta_\varepsilon * u$$

where $\zeta_\varepsilon := \varepsilon^{-d}\zeta(x/\varepsilon)$. It is well known that $u^{(\varepsilon)}$ is a smooth function on $\mathbb{R}^d$, and that the differential (difference) operators commute with the operation of mollification. Furthermore, the function $u^{(\varepsilon)}$ is represented as the "sum" over $y \in \mathbb{R}^d$ of $u(\cdot - \varepsilon y)\zeta(y)$. In any Banach space the norm of a sum does not exceed the sum of the norms. This proves the first inequality in (4.1) below.

Lemma 4.1. *Let* $\gamma \in (0, 1]$, $n, m = 0, 1, 2, \ldots$, $u \in \mathcal{C}^{n,\gamma}$. *Then*

$$|u^{(\varepsilon)}|_{n,\gamma} \leq |u|_{n,\gamma} \int_{\mathbb{R}^d} |\zeta|\, dx, \quad |u^{(\varepsilon)}|_{n+m,\gamma} \leq N(d, n, k, \zeta)\varepsilon^{-m}|u|_{n,\gamma}. \tag{4.1}$$

Estimates (4.1) are also true if we drop γ and assume that $u \in \mathcal{C}^n$.

Furthermore, if ζ has unit integral and $u \in \mathcal{C}^{0,1}$, then

$$|u - u^{(\varepsilon)}|_0 \leq N(\zeta, d)\varepsilon|u|_{0,1}. \tag{4.2}$$

Finally, if ζ is even and has unit integral, and $u \in \mathcal{C}^{1,1}$, then

$$|u - u^{(\varepsilon)}|_0 \leq N(\zeta, d)\varepsilon^2|u|_{1,1}. \tag{4.3}$$

The second inequality in (4.1) follows from the first once we bring all derivatives up to the order m to ζ_ε. To prove (4.3) for $u \in \mathcal{C}^2$, it suffices to observe that, by the symmetry of ζ,

$$u^{(\varepsilon)}(x) - u(x) = \int_{\mathbb{R}^d} [u(x - \varepsilon y) - u(x) - \varepsilon y^i u_{x^i}(x)]\zeta(y)\, dy,$$

where the integrand is dominated by $N|y|^2\varepsilon^2|u|_{1,1}$. After that, (4.3) is carried over to all $u \in \mathcal{C}^{1,1}$ by using an obvious extra mollification. Estimate (4.2) is proved similarly by using the inequality $|u(x - \varepsilon y) - u(x)| \leq \varepsilon|y|$.

Lemma 4.2. *Suppose* $u \in \mathcal{C}^{0,1}$, $\eta \in \mathcal{S}$, $\varepsilon, h > 0$, $l \in \mathbb{R}^d$, $|l| = 1$. *Introduce*

$$I(\eta, x) := \int_{\mathbb{R}^d} u(x - \varepsilon y)D_l^2\eta(y)\, dy.$$

Then for any $r > 0$ we have

$$I(\eta, x) = \int_{\mathbb{R}^d} u(x - \varepsilon y)\Delta_{r,l}^2\eta(y)\, dy - \frac{1}{6r^2}\int_{-r}^r (r - |s|)^3 I(D_l^2\eta(\cdot + sl), x)\, ds. \tag{4.4}$$

Moreover,

$$(4.5) \qquad |I(\eta, x)| \leq N\varepsilon^2(|\Delta_{h,l}^2 u|_0 + [u]_{0,1} h^2/\varepsilon^3),$$

$$(4.6) \qquad |D_l^2 u^{(\varepsilon)}|_0 \leq N(|\Delta_{h,l}^2 u|_0 + [u]_{0,1} h^2/\varepsilon^3),$$

where $N = N(d, \eta, \zeta)$.

Proof. Formula (4.4) is an immediate consequence of the following result of application of Taylor's formula:

$$(4.7) \qquad D_l^2 \eta(y) = \Delta_{r,l}^2 \eta(y) - \frac{1}{6r^2} \int_{-r}^{r} D_l^4 \eta(y + sl)(r - |s|)^3 \, ds.$$

Next,

$$|I(\eta, x)| = \left| \int_{\mathbb{R}^d} [u(x - \varepsilon y) - u(x)] D_l^2 \eta(y) \, dy \right|$$

$$\leq \varepsilon[u]_{0,1} \int_{\mathbb{R}^d} |y| \, |D_l^2 \eta(y)| \, dy = N\varepsilon[u]_{0,1}.$$

Also, observe that changing variables shows that the operator $\Delta_{h/\varepsilon, \ell_k}^2$ is selfadjoint. It can easily be checked that, for $r = h/\varepsilon$, the result of application of the operator $\Delta_{r,l}^2$ with respect to y to the function $u(x - \varepsilon y)$ is simply $\varepsilon^2(\Delta_{h,l}^2 u)(x - \varepsilon y)$. Therefore,

$$\left| \int_{\mathbb{R}^d} u(x - \varepsilon y) \Delta_{r,l}^2 \eta(y) \, dy \right| = \varepsilon^2 \left| \int_{\mathbb{R}^d} (\Delta_{h,l}^2 u)(x - \varepsilon y) \eta(y) \, dy \right|$$

$$\leq \varepsilon^2 |\Delta_{h,l}^2 u|_0 \int_{\mathbb{R}^d} |\eta(y)| \, dy = N\varepsilon^2 |\Delta_{h,l}^2 u|_0.$$

Now (4.4) implies that

$$|I(\eta, x)| \leq N\varepsilon^2 |\Delta_{h,l}^2 u|_0 + N\varepsilon[u]_{0,1} \frac{1}{r^2} \int_{-r}^{r} (r - |s|)^3 \, ds = N\varepsilon^2 |\Delta_{h,l}^2 u|_0 + Nr^2 \varepsilon[u]_{0,1}.$$

This proves (4.5). To prove (4.6), it suffices to note that $D_l^2 u^{(\varepsilon)}(x) = \varepsilon^{-2} I(\zeta, x)$. The lemma is proved. $\qquad\square$

Remark 4.3. Applying (4.5) to the estimation of $I(D_l^2 \eta(\cdot + sl), x)$, and using (4.4) and induction on m, one can prove easily that the following estimates are true for any $m, n = 0, 1, \ldots$, and $\varepsilon \geq h > 0$:

$$|I(\eta, x)| \leq N\varepsilon^2(|\Delta_{h,l}^2 u|_0 + (h/\varepsilon)^m \varepsilon^{-1}[u]_{0,1}),$$

$$|D_l^{n+2} u^{(\varepsilon)}|_0 \leq N\varepsilon^{-n}(|\Delta_{h,l}^2 u|_0 + (h/\varepsilon)^m \varepsilon^{-1}[u]_{0,1}),$$

where $N = N(d, m, n, \eta, \zeta)$.

Lemma 4.4. *Suppose $u \in \mathcal{C}^{0,1}$, $\varepsilon \geq h > 0$, and $\zeta(x) = (4\pi)^{-d/2} \exp(-|x|^2/4)$. Then*

$$(4.8) \qquad |u^{(\varepsilon)} - u|_0 \leq N(d)\Big(\varepsilon^2 \sum_i |\Delta_{h,e_i}^2 u|_0 + h[u]_{0,1}\Big).$$

Proof. Our particular ζ is well related to the heat semigroup T_t, namely, $u^{(\varepsilon)} = T_{\varepsilon^2} u$. For any $t > s > 0$ we have

$$T_t u - T_s u = \int_s^t \Delta T_r u \, dr,$$

where, by (4.6) with $n = 0$ and $\sqrt{r}$ in place of ε,

$$(4.9) \qquad |\Delta T_r u|_0 \leq N\Big(\sum_i |\Delta_{h,e_i}^2 u|_0 + [u]_{0,1} h^2 r^{-3/2}\Big).$$

Consequently,

$$|T_t u - T_s u|_0 \le N(t-s) \sum_i |\Delta_{h,e_i}^2 u|_0 + N[u]_{0,1} h^2 s^{-1/2}.$$

Furthermore, (4.2) implies that $|T_s u - u|_0 \le N\sqrt{s}[u]_{0,1}$, so that

$$|T_t u - u|_0 \le Nt \sum_i |\Delta_{h,e_i}^2 u|_0 + N[u]_{0,1} h^2 s^{-1/2} + N\sqrt{s}[u]_{0,1}.$$

It only remains to take $t = \varepsilon^2$ and $s = h^2$. The lemma is proved. $\square$

Remark 4.5. In fact, (4.8) with $N(\zeta)$ in place of $N(d)$ is true for any even $\zeta \in \mathcal{S}$ with unit integral.

Indeed, let $P_t u = u^{(\sqrt{t})}$. Then for smooth u we have

$$\frac{\partial}{\partial r} P_r u(x) = -(1/2) r^{-1/2} \int_{\mathbb{R}^d} \zeta(y) y^i u_{x^i}(x - y\sqrt{r}) \, dy$$

(4.10)
$$= \int_{\mathbb{R}^d} \zeta_i(y) u_{x^i x^i}(x - y\sqrt{r}) \, dy$$

$$= \sum_i D_{e_i}^2 \int_{\mathbb{R}^d} \zeta_i(y) u(x - y\sqrt{r}) \, dy,$$

where

$$\zeta_i(y) = -(1/2) \int_{-\infty}^{y^i} \zeta(y - y^i e_i + s e_i) s \, ds.$$

The expression on the right in (4.10) is estimated by the right-hand side of (4.9) as before, and we get the result by following the same lines as in the proof of the lemma.

The last auxiliary result we need is obtained by using (2.3), the mean value theorem, and straightforward inspection.

Lemma 4.6. *For any $u \in \mathcal{C}^4$ and $h > 0$ we have*

$$|L_h^\alpha u - L^\alpha u| \le N(K)\left(h^2 \sum_{k=1}^{d_1} |D_{\ell_k}^4 u|_0 + h \sum_{k=1}^{d_1} |D_{\ell_k}^2 u|_0 \right)$$

if the L_h^α are as in (2.6), and

$$|L_h^\alpha u - L^\alpha u| \le N(K) h^2 \left(\sum_{k=1}^{d_1} |D_{\ell_k}^4 u|_0 + \sum_{k=1}^{d_1} |D_{\ell_k}^3 u|_0 \right)$$

if the L_h^α are as in (2.15).

§5. Proof of Theorem 2.8

Due to Assumption 2.2, without losing generality we may assume that

$$\sup_{\alpha \in A} a_k(\alpha) > 0, \quad k = 1, \dots, d_1.$$

Next, let $\mathcal{L} = \mathrm{Span}\{\ell_1, \dots, \ell_{d_1}\}$, and let $\mathcal{M}$ be the orthogonal complement of $\mathcal{L}$ in $\mathbb{R}^d$. Then we can represent each point $x \in \mathbb{R}^d$ as (x', x''), where $x' \in \mathcal{L}$ and $x'' \in \mathcal{M}$, and since the operators L and L_h only affect the variable x', we can regard x'' as a parameter appearing in $f(\alpha, x)$ and prove (2.13) for each particular value of this parameter.

Also, we note that

$$\inf_{\substack{|l|=1 \\ l \in \mathcal{L}}} \sup_{\alpha \in A} (a(\alpha)l, l) > 0.$$

Indeed, otherwise (cf. (2.3)) we can find a unit $l \in \mathcal{L}$ such that

$$\sup_{\alpha \in A} \sum_{k=1}^{d_1} a_k(\alpha)(l, l_k)^2 = 0,$$

and for each k,

$$(l, l_k)^2 \sup_{\alpha \in A} a_k(\alpha) = 0, \quad (l, l_k) = 0,$$

which is impossible for a unit $l \in \mathcal{L}$.

Taking all these into account and passing from $\mathbb{R}^d$ to $\mathcal{L}$ if necessary, we see that without losing generality we may assume that the following is true.

Assumption 5.1. *There is a constant $\delta > 0$ such that equation (2.4) is weakly δ-nondegenerate; that is, for any unit $l \in \mathbb{R}^d$ we have*

$$(5.1) \qquad\qquad \sup_{\alpha \in A}(a(\alpha)l, l) \geq \delta.$$

We shall see that under this assumption the constant N in (2.13) depends only on $d, d_1, K, \kappa, \delta,$ and ℓ_k. By the way, (5.1) is a condition on $F(u_{ij}, u_i, u, x)$ rather than on its particular representation (2.2), because

$$\sup_{\alpha \in A}(a(\alpha)l, l) = \lim_{t \to \infty} t^{-1}F(l^i l^j t, 0, 0, x).$$

Lemma 5.2. *We have $d_1 \geq d$, and there exist vectors $\ell_{k_1}, \ldots, \ell_{k_d}$ such that they form a basis in $\mathbb{R}^d$ and*

$$(5.2) \qquad\qquad \sup_{\alpha \in A} a_{k_i}(\alpha) \geq \delta/(2d_1 K^2) =: \delta_1, \quad i = 1, \ldots, d.$$

Proof. Since $\mathbb{R}^d$ is spanned by $\ell_1, \ldots, \ell_{d_1}$, we can find $\ell_{k_1}, \ldots \ell_{k_d}$ forming a basis in $\mathbb{R}^d$. If (5.2) is satisfied for all $i = 1, \ldots, d$, we are done. If not, then there is no loss of generality in assuming that (5.2) is violated for $i = 1$. By (2.3), inequality (5.1) can be rewritten as

$$\sup_{\alpha \in A} \sum_{k=1}^{d_1} a_k(\alpha)(l, \ell_k)^2 \geq \delta.$$

Let l_1 be a unit in $\mathbb{R}^d$ perpendicular to $\ell_{k_2}, \ldots, \ell_{k_d}$, and let $\alpha_1 \in A$ be such that

$$\sum_{k=1}^{d_1} a_k(\alpha_1)(l_1, \ell_k)^2 \geq \delta/2.$$

Then we can find $\tilde{k}_1$ satisfying

$$\delta/(2d_1) \leq a_{\tilde{k}_1}(\alpha_1)(l_1, \ell_{\tilde{k}_1})^2 \leq K^2 a_{\tilde{k}_1}(\alpha_1).$$

This and the fact that $l_1 \perp \ell_{k_2}, \ldots, \ell_{k_d}$ imply that $\ell_{\tilde{k}_1} \notin \operatorname{Span}\{\ell_{k_2}, \ldots, \ell_{k_d}\}$. Hence, $\ell_{\tilde{k}_1}, \ell_{k_2}, \ldots, \ell_{k_d}$ is still a basis in $\mathbb{R}^d$. We see that we can replace any ℓ_{k_i} for which (5.2) is violated with a different element of $\{\ell_1, \ldots, \ell_{d_1}\}$ for which (5.2) is satisfied and preserve the linear independence of new vectors. The lemma is proved. $\qquad\square$

By relabeling if needed, we may assume that

$$\ell_{k_i} = \ell_i, \quad i = 1, \ldots, d.$$

It turns out that a linear transformation allows us to further reduce the general situation to that with

$$(5.3) \qquad\qquad \ell_i = e_i, \quad i = 1, \ldots, d.$$

Indeed, observe that if $Q : \mathbb{R}^d \to \mathbb{R}^d$ is a linear one-to-one transformation, then $v(Qx) = \tilde{v}(x)$, where

$$\tilde{v}(x) = \sup_{\alpha \in \mathfrak{A}} E \int_0^\infty \tilde{f}(\alpha_t, \tilde{x}_t^\alpha(x)) \exp\left(\int_0^t c(\alpha_s)\, ds\right) dt,$$

$$\tilde{x}_t^\alpha(x) = x + \int_0^t \sigma_k(\alpha_s)\tilde{\ell}_k\, dw_s^k + \int_0^t b_k(\alpha_s)\tilde{\ell}_k\, ds,$$

$$\tilde{f}(\alpha, x) = f(\alpha, Qx), \quad \tilde{\ell}_k = Q^{-1}\ell_k.$$

Consequently, the function $\tilde{v}$ is associated with the Bellman equation

$$\sup_{\alpha \in A}[a_k(\alpha)D_{\tilde{\ell}_k}^2 u(x) + b_k(\alpha)D_{\tilde{\ell}_k} u(x) + c(\alpha)u(x) + \tilde{f}(\alpha, x)] = 0.$$

Moreover, for the function $\tilde{v}_h(x) := v_h(Qx)$ we have

$$\Delta_{h,\tilde{\ell}_k}^\pm \tilde{v}_h(x) = (\Delta_{h,\ell_k}^\pm v_h)(Qx), \quad \Delta_{h,\tilde{\ell}_k}^2 \tilde{v}_h(x) = (\Delta_{h,\ell_k}^2 v_h)(Qx),$$

so that

$$\sup_{\alpha \in A} \big[a_k(\alpha)\Delta_{h,\tilde{\ell}_k}^2 \tilde{v}_h(x) + b_k^+(\alpha)\Delta_{h,\tilde{\ell}_k}^+ \tilde{v}_h(x)$$

$$+ b_k^-(\alpha)\Delta_{h,\tilde{\ell}_k}^- \tilde{v}_h(x) + c(\alpha)\tilde{v}_h(x) + \tilde{f}(\alpha, x) \big] = 0.$$

Since $|v_h - v|_0 = |\tilde{v}_h - \tilde{v}|_0$, it suffices to estimate $|\tilde{v}_h - \tilde{v}|_0$, and since always there exists Q such that $\tilde{\ell}_i = Q^{-1}\ell_i = e_i$, $i = 1, \ldots, d$, we see that indeed assuming (5.3) does not restrict generality. Actually, the chain rule and the fact that the ith column of Q is exactly ℓ_k show that Assumption 2.1 is satisfied for our new equation with $\tilde{\ell}_k$, $\tilde{f}$, and $\tilde{K} = \tilde{K}(K, \ell_1, \ldots, \ell_d)$ in place of ℓ_k, f, and K, respectively. All other assumptions are satisfied automatically.

Everywhere below in this section we assume that (5.3) is fulfilled. First, we prove that v and v_h possess some regularity.

Lemma 5.3. *There is a constant $N = N(d, K, \kappa, \delta)$ such that for all $h > 0$, $k = 1, \ldots, d$, we have*

$$(5.4) \qquad\qquad |\Delta_{h,\ell_k}^2 v_h|_0 \le N.$$

Moreover, $v_h(x) + N|x|^2$ is convex and, for any $\zeta \in \mathcal{S}$ and any unit $l \in \mathbb{R}^d$ and $\varepsilon > 0$, we have

$$(5.5) \qquad\qquad |D_l^2 v_h^{(\varepsilon)}|_0 \le N(1 + h^2/\varepsilon^3),$$

where $N = N(d, K, \kappa, \delta, \zeta)$.

Proof. To prove (5.4) we follow the same ideas as in [5]. First, observe that

$$\sup_{\alpha \in A}\{L_h^\alpha v_h + c(\alpha)v_h + f(\alpha, x)\} = 0$$

by the definition of v_h. Let $\gamma > 0$ be a number, and let l be a unit vector in $\mathbb{R}^d$. Taking the symmetric second-order difference $\Delta_{\gamma,l}^2$ of both sides and using the inequality

$$\inf_\alpha \Delta_{\gamma,l}^2 \varphi^\alpha(x) \le \Delta_{\gamma,l}^2 \sup_\alpha \varphi^\alpha(x),$$

we obtain

$$(5.6) \qquad \begin{aligned} &\inf_{\alpha \in A}\{L_h^\alpha \Delta_{\gamma,l}^2 v_h(x) + c(\alpha)\Delta_{\gamma,l}^2 v_h(x) + \Delta_{\gamma,l}^2 f(\alpha, x)\} \\ &\le 0 = \inf_{\alpha \in A}\{L_h^\alpha 0 + c(\alpha)0 + \Delta_{\gamma,l}^2 f(\alpha, x)\} - \inf_{\alpha \in A} \Delta_{\gamma,l}^2 f(\alpha, x). \end{aligned}$$

Here

(5.7)
$$|\Delta^2_{\gamma,l} f(\alpha, x)| \leq K.$$

Indeed, if $f \in \mathcal{C}^2$, then, by Taylor's formula applied to the function

$$y(\gamma) = f(x + \gamma l) - 2f(x) + f(x - \gamma l),$$

we have

$$y(\gamma) = (1/2)\gamma^2 [D^2_l f(x + \theta\gamma l) + D^2_l f(x - \theta\gamma l)],$$

where $\theta \in (0,1)$. In this case (5.7) follows from Assumption 2.1. To get rid of the extra smoothness assumption on f, we can use mollification, easily obtaining (5.7) in full generality.

After that, coming back to (5.6) and using Lemma 2.5 and Remark 2.12, we conclude that

(5.8)
$$\Delta^2_{\gamma,l} v_h(x) \geq -\kappa^{-1} |\inf_{\alpha,x} \Delta^2_{\gamma,l} f(\alpha, x)| \geq -\kappa^{-1} K.$$

To derive an upper bound for $\Delta^2_{h,\ell_k} v_h$, we note that

$$\delta_1 \Delta^2_{h,\ell_1} v_h \leq \sup_{\alpha \in A} \{a_1(\alpha) \Delta^2_{h,\ell_1} v_h\}$$

$$\leq F_h[v_h] + \sup_{\alpha \in A} \left\{ -\sum_{k=2}^{d_1} a_k(\alpha) \Delta^2_{h,\ell_k} v_h - b_k^+(\alpha) \Delta^+_{h,\ell_k} v_h \right.$$

$$\left. - b_k^-(\alpha) \Delta^-_{h,\ell_k} v_h - c(\alpha) v_h - f(\alpha, x) \right\}$$

$$\leq N(d, k, \kappa),$$

where we have used (5.8) with $\gamma = h$, Lemma 2.4, Assumption 2.1, and the positiveness of $a_k(\alpha)$. The quantity $\delta_1 \Delta^2_{h,\ell_i} v_h$ for $i = 2, \ldots, d$ is estimated similarly, which finishes the proof of (5.4).

The arbitrariness of γ and l in estimate (5.8) implies that $v_h(x) + \kappa^{-1} K|x|^2$ is convex. Therefore, if $\zeta \geq 0$, then $u_h := v_h^{(\varepsilon)}(x) + \kappa^{-1} K|x|^2$ is also convex and smooth. In that case, by (5.3), for any unit $l \in \mathbb{R}^d$ we have

$$-N \leq D^2_l v_h^{(\varepsilon)}(x) \leq D^2_l u_h(x) \leq N \sum_{k=1}^d D^2_{\ell_k} u_h(x) = N \sum_{k=1}^d D^2_{\ell_k} v_h^{(\varepsilon)}(x) + N.$$

Now, to prove (5.5) for $\zeta \geq 0$, it only remains to use (4.6). The general case is reduced to the above one by representing an arbitrary ζ as the difference of two nonnegative elements of $\mathcal{S}$. The lemma is proved. $\qquad\square$

Lemma 5.4. *There is a constant $N = N(d, K, \kappa, \delta)$ such that $|v|_{1,1} \leq N$. Moreover, v is a unique $\mathcal{C}^{1,1}$-function satisfying equation (2.4) almost everywhere in $\mathbb{R}^d$.*

Proof. Actually, the lemma is a very particular case of the results in [5]. However, a direct reference may be harder to follow than a sequence of references to easier facts presented in [5]. We prefer to give more details. First, assume that $f(\alpha, \cdot) \in \mathcal{C}^2$.

For any $T \in (0, \infty)$, $s \in [0, T]$, and $x \in \mathbb{R}^d$, we introduce

$$v_T(s, x) = \sup_{\alpha \in \mathfrak{A}} E \int_0^{T-s} f(\alpha_t, x_t^\alpha(x)) \exp\left(\int_0^t c(\alpha_r)\, dr\right) dt.$$

By [5, Theorem 4.7.4] the generalized derivatives $D_s v_T(s, x)$, $D_x v_T(s, x)$, and $D^2_x v_T(s, x)$ are bounded in $[0, T] \times \mathbb{R}^d$ by a constant depending only on d, K, κ, and δ, times

$(1 + |x|^m) \exp(N(T - s))$, where $m = 0$, and N measures the rate with which the magnitudes of the derivatives of $x_t^\alpha(x)$ with respect to x go to infinity as $t \to \infty$. Since the latter derivatives are constant, we have $N = 0$, and the generalized derivatives $D_s v_T(s, x)$ $D_x v_T(s, x)$, and $D_x^2 v_T(s, x)$ are bounded in $[0, T] \times \mathbb{R}^d$ by a constant independent of T. Since, obviously, $|v_T - v|_0 \to 0$ as $T \to \infty$, we have the same bounds on the derivatives of v. Also we recall that for a function to be Lipschitz continuous is the same as to have bounded generalized derivatives. This yields the first assertion of the lemma.

Next, by [5, Theorem 4.7.7] we have

$$D_s v_T(s, x) + F(v_{Tx^i x^j}(s, x), v_{Tx^i}(s, x), v_T(s, x), x) = 0$$

a.e. in $[0, T] \times \mathbb{R}^d$. Letting $T \to \infty$ and using [5, Theorem 4.5.1], we get

$$D_s v(x) + F(v_{x^i x^j}(x), v_{x^i}(x), v(x), x) = 0$$

a.e. in $[0, T] \times \mathbb{R}^d$ for each $T \in [0, \infty)$. Since v is independent of s, we see that v satisfies (2.4) a.e. in $\mathbb{R}^d$.

We assumed above that $f(\alpha, \cdot) \in \mathcal{C}^2$. However, in all estimates only the norm $|f|_{1,1}$ is involved. Therefore, mollifying f with respect to x and passing to the limit, we easily obtain the result in the general case.

If we have another function u in $\mathcal{C}^{1,1}$ (and even in a much wider class) satisfying equation (2.4) a.e. in $\mathbb{R}^d$, then, by [5, Theorem 5.3.14],

$$u(x) = \sup_{\alpha \in \mathfrak{A}} E\bigg[\int_0^T f(\alpha_t, x_t^\alpha(x)) \exp\bigg(\int_0^t c(\alpha_s)\, ds \bigg) dt$$
$$+ u(x_T^\alpha(x)) \exp\bigg(\int_0^T c(\alpha_s)\, ds \bigg) \bigg]$$

for any $T > 0$. Letting $T \to \infty$, we see that $u = v$. The lemma is proved. $\qquad\square$

We combine Lemmas 4.1, 5.3, and 5.4 and the observation that

$$D_l^{m+2} u^{(\varepsilon)} = \varepsilon^{-m} D_l^2(u * \eta),$$

where $\eta(x) = \varepsilon^{-d}(D_l^m \zeta)(x/\varepsilon)$, to get the following statement.

Corollary 5.5. *For any nonnegative integer m and any $\varepsilon, h > 0$, we have*

$$|Dv^{(\varepsilon)}| \le N, \quad |D^{m+2} v^{(\varepsilon)}| \le N\varepsilon^{-m},$$
$$|Dv_h^{(\varepsilon)}| \le N, \quad |D^{m+2} v_h^{(\varepsilon)}| \le N\varepsilon^{-m}(1 + h^2/\varepsilon^3),$$

where $N = N(d, d_1, K, m, \kappa, \delta, \zeta)$.

Now we are ready to prove the theorem. By Lemma 5.4, if $\alpha \in A$ and $\varepsilon > 0$, then on $\mathbb{R}^d$ we have the inequality

$$L^\alpha v^{(\varepsilon)} + c(\alpha)v^{(\varepsilon)} + f^{(\varepsilon)}(\alpha, x) \le 0.$$

By Lemma 4.6, Corollary 5.5, and Lemma 4.1, for nonnegative symmetric ζ with unit integral we obtain

$$\sup_{\alpha \in A}\{L_h^\alpha v^{(\varepsilon)} + c(\alpha)v^{(\varepsilon)} + f(\alpha, x)\}$$

$$\leq \sup_{\alpha \in A}\{L^\alpha v^{(\varepsilon)} + c(\alpha)v^{(\varepsilon)} + f(\alpha, x)\}$$

$$+ N(K)\left[h^2 \sum_{k=1}^{d_1} |D_{\ell_k}^4 v^{(\varepsilon)}|_0 + h \sum_{k=1}^{d_1} |D_{\ell_k}^2 v^{(\varepsilon)}|_0\right]$$

$$\leq \sup_{\alpha \in A}\{L^\alpha v^{(\varepsilon)} + c(\alpha)v^{(\varepsilon)} + f(\alpha, x)\} + Nh^2\varepsilon^{-2} + Nh$$

$$\leq \sup_{\alpha \in A}\{L^\alpha v^{(\varepsilon)} + c(\alpha)v^{(\varepsilon)} + f^{(\varepsilon)}(\alpha, x)\} + N(h^2\varepsilon^{-2} + h + \varepsilon^2)$$

$$\leq N(h^2\varepsilon^{-2} + h + \varepsilon^2).$$

Lemma 2.5(ii) shows that

$$v_h(x) \leq v^{(\varepsilon)}(x) + N(h^2\varepsilon^{-2} + h + \varepsilon^2) \leq v(x) + N(h^2\varepsilon^{-2} + h + \varepsilon^2),$$

for any $x \in \mathbb{R}^d$, where N depends only on κ, K, d, and δ. Letting $\varepsilon = h^{1/2}$, we get the required upper estimate for $v_h - v$.

To estimate the same function from below, we use almost the same argument, but start with the equation defining v_h. Again we take a nonnegative symmetric $\zeta \in \mathcal{S}$ with unit integral. Observe that

$$(5.9) \qquad L_h^\alpha v_h(y) + c(\alpha)v_h(y) + f(\alpha, y) \leq 0 \quad \text{for any } \alpha \in A \text{ and } y \in \mathbb{R}^d.$$

If we multiply (5.9) by $\zeta_\varepsilon(x - y)$ and integrate with respect to y, we obtain

$$(5.10) \qquad \sup_{\alpha \in A}\{L_h^\alpha v_h^{(\varepsilon)}(x) + c(\alpha)v_h^{(\varepsilon)}(x) + f^{(\varepsilon)}(\alpha, x)\} \leq 0.$$

By Lemma 4.6 and Corollary 5.5, this yields

$$\sup_{\alpha \in A}\{L^\alpha v_h^{(\varepsilon)} + c(\alpha)v_h^{(\varepsilon)} + f(\alpha, x)\}$$

$$\leq \sup_{\alpha \in A}\{L_h^\alpha v_h^{(\varepsilon)} + c(\alpha)v_h^{(\varepsilon)} + f^{(\varepsilon)}(\alpha, x)\}$$

$$+ N\left[h^2 \sum_{k=1}^{d_1} |D_{\ell_k}^4 v_h^{(\varepsilon)}|_0 + h \sum_{k=1}^{d_1} |D_{\ell_k}^2 v_h^{(\varepsilon)}|_0\right] + N\varepsilon^2$$

$$\leq N[h^2\varepsilon^{-2}(1 + h^2\varepsilon^{-3}) + h(1 + h^2\varepsilon^{-3})] + N\varepsilon^2.$$

Using this estimate and applying Itô's formula to

$$v_h^{(\varepsilon)}(x_t^\alpha(x)) \exp\left(\int_0^t c(\alpha_s)\,ds\right),$$

we easily see that, for any $x \in \mathbb{R}^d$,

$$v(x) \leq v_h^{(\varepsilon)}(x) + N[h^2\varepsilon^{-2}(1 + h^2\varepsilon^{-3}) + h(1 + h^2\varepsilon^{-3})] + N\varepsilon^2.$$

Relations (5.3) allow us to apply Remark 4.5, and

$$v(x) \leq v_h(x) + N[h^2\varepsilon^{-2}(1 + h^2\varepsilon^{-3}) + h(1 + h^2\varepsilon^{-3})] + N\varepsilon^2,$$

where $N = N(\kappa, K, d, d_1, \delta)$. Putting $\varepsilon = h^{1/2}$, we complete the proof of Theorem 2.8.

§6. Proof of Theorems 2.9 and 2.11

Proof of Theorem 2.9. We use the idea of interpolation with respect to the smoothness of f. To reflect the dependence of v and v_h on f we write $v = v[f]$ and $v_h = v_h[f]$. Notice that, obviously, for any constant $c \geq 0$, we have $v[cf] = cv[f]$ and $v_h[cf] = cv_h[f]$.

We fix a nonnegative $\zeta \in \mathcal{S}$ with unit integral and observe that $|\varepsilon f^{(\varepsilon)}|_{1,1} \leq N(\zeta)|f|_{0,1}$. Therefore, by Theorem 2.8,

$$|v[\varepsilon f^{(\varepsilon)}] - v_h[\varepsilon f^{(\varepsilon)}]|_0 \leq Nh, \quad |v[f^{(\varepsilon)}] - v_h[f^{(\varepsilon)}]|_0 \leq Nh/\varepsilon.$$

Next, we have

$$|v[f^{(\varepsilon)}] - v[f]|_0 \leq \kappa^{-1}|f^{(\varepsilon)} - f|_0 \leq N\varepsilon,$$
$$|v_h[f^{(\varepsilon)}] - v_h[f]|_0 \leq \kappa^{-1}|f^{(\varepsilon)} - f|_0 \leq N\varepsilon,$$

so that

$$|v - v_h|_0 \leq Nh/\varepsilon + N\varepsilon.$$

Now, it only remains to take $\varepsilon = h^{1/2}$. The theorem is proved. $\qquad\square$

Proof of Theorem 2.11. As was explained in Section 5, without loss of generality we may suppose that Assumption 5.1 is satisfied and (5.3) is fulfilled.

Lemma 6.1. *For $\varepsilon \geq 0$, define*

$$A(\varepsilon) = \{\alpha \in A : (a(\alpha)\lambda, \lambda) \geq \varepsilon|\lambda|^2 \text{ for all } \lambda \in \mathbb{R}^d\}.$$

Then for each $R \in [0, \infty)$ there is $\varepsilon > 0$ such that

$$\sup_{\alpha \in A(\varepsilon)} [a^{ij}(\alpha)u_{ij} + b^i(\alpha)u_i + c(\alpha)u + f(\alpha, x)] = 0$$

whenever $(u_{ij}, u_i, u, x) \in \Gamma \cap G_R$.

Proof. Take a point $(u_{ij}, u_i, u, x) \in \Gamma \cap G_R$ and observe that, as a function of $(v_{ij}, v_i, v) \in \mathbb{R}^{d'_0}$, $F(v_{ij}, v_i, v, x)$ is a Lipschitz continuous convex function as the supremum of linear functions, and $F(v_{ij}, v_i, v, x)$ is not constant in balls in $\mathbb{R}^{d'_0}$ because $c(\alpha) \leq -\kappa$. Also, $F(v_{ij}, v_i, s, x) \to -\infty$ as $s \to -\infty$. Therefore,

$$\Theta_x := \{(v_{ij}, v_i, v) \in \mathbb{R}^{d'_0} : F(v_{ij}, v_i, v, x) < 0\}$$

is a nonempty open convex set, and its boundary is

$$\Gamma_x := \{(v_{ij}, v_i, v) \in \mathbb{R}^{d'_0} : F(v_{ij}, v_i, v, x) = 0\}.$$

Next, observe that there is a sequence of $\alpha_n \in A$ such that

$$a^{ij}(\alpha_n)u_{ij} + b^i(\alpha_n)u_i + c(\alpha_n)u + f(\alpha_n, x) \to F(u_{ij}, u_i, u, x) = 0,$$

and for all $(v_{ij}, v_i, v) \in \mathbb{R}^{d_2}$ and all n we have

$$a^{ij}(\alpha_n)v_{ij} + b^i(\alpha_n)v_i + c(\alpha_n)v + f(\alpha_n, x) \leq F(v_{ij}, v_i, v, x).$$

It follows that, for a subsequence $\{n'\} \subset \{n\}$, the limit

$$(\bar{a}^{ij}, \bar{b}^i, \bar{c}, \bar{f}(x)) := \lim_{n' \to \infty} (a^{ij}(\alpha_{n'}), b^i(\alpha_{n'}), c(\alpha_{n'}), f(\alpha_{n'}, x))$$

exists, and

$$\bar{a}^{ij}u_{ij} + \bar{b}^i u_i + \bar{c}u + \bar{f}(x) = 0,$$
$$\bar{a}^{ij}v_{ij} + \bar{b}^i v_i + \bar{c}v + \bar{f}(x) < 0, \quad (v_{ij}, v_i, v) \in \Theta_x.$$

In other words, the equation $\bar{a}^{ij}v_{ij} + \bar{b}^i v_i + \bar{c}v + \bar{f}(x) = 0$ describes a tangent hyperplane to Γ_x at (u_{ij}, u_i, u). Furthermore, we have $\bar{c} \leq -\kappa$, so that

$$\sum_{ij} |\bar{a}^{ij}|^2 + \sum_i |\bar{b}^i|^2 + |\bar{c}|^2 \geq \kappa^2.$$

By assumption, for our smooth function $\Phi(v_{ij}, v_i, v, x)$ we have $\Phi = 0$ on Γ_x, and the gradient of Φ with respect to (v_{ij}, v_i, v) at (u_{ij}, u_i, u) is not zero. It follows that Γ_x is smooth near (u_{ij}, u_i, u), and there is a constant $\beta \neq 0$ such that

$$(\bar{a}^{ij}, \bar{b}^i, \bar{c}, \bar{f}(x)) = \beta(\Phi_{u_{ij}}, \Phi_{u_i}, \Phi_u, \Phi)(u_{ij}, u_i, u, x).$$

Since $(\bar{a}^{ij}) \geq 0$ and $(\Phi_{u_{ij}}) > 0$ (see (2.17)), we have $\beta > 0$. Moreover,

$$\beta = \frac{\bar{c}}{\Phi_u(u_{ij}, u_i, u, x)}, \quad \beta = |\beta| \geq \frac{\kappa}{[\Phi]_{0,1,G_R}} := \rho(R).$$

Consequently,

$$\bar{a}^{ij}\lambda^i\lambda^j = \beta\Phi_{u_{kr}}(u_{ij}, u_i, u, x)\lambda^k\lambda^r \geq \mu(R)\rho(R)|\lambda|^2$$

for all $\lambda \in \mathbb{R}^d$. Finally, it is obvious that $\mu(R)\rho(R)/2$ can be taken as the ε we are looking for. The lemma is proved. $\qquad\square$

This lemma and Lemma 5.4 yield the following.

Corollary 6.2. *There exists $\varepsilon > 0$ such that a.e. in $\mathbb{R}^d$ the function v satisfies the equation*

$$(6.1) \qquad \sup_{\alpha \in A(\varepsilon)} \{L^\alpha u(x) + c(\alpha)v(x) + f(\alpha, x)\} = 0.$$

Now we basically repeat the argument in [7, Remark 5.17], giving more detail. It is well known (see, e.g., [11] or [6, Theorem 6.4.5]) that, since equation (6.1) is uniformly nondegenerate, there exists $\gamma \in (0,1)$ and a unique $u \in \mathcal{C}^{2,\gamma}$ satisfying (6.1) in $\mathbb{R}^d$. By the unique solvability of (6.1) in $\mathcal{C}^{1,1}$, we have $u = v \in \mathcal{C}^{2,\gamma}$.

It follows that v is a $\mathcal{C}^{2,\gamma}$-solution of

$$\Phi(v_{x^i x^j}(x), v_{x^i}(x), v(x), x) = 0.$$

Finally, a classical result (see, e.g., [4, Theorem 3.13] or [6, Lemma 1.3.2]) implies that $v \in \mathcal{C}^{4,\gamma}$, due to the smoothness of Φ. Hence, by Lemma 4.6,

$$|L^\alpha v - L_h^\alpha v| \leq Nh^2, \quad |F_h[v]| \leq Nh^2,$$

and to finish the proof of Theorem 2.11, it only remains to apply Lemma 2.5. The theorem is proved. $\qquad\square$

References

[1] G. Barles and E. R. Jakobsen, *On the convergence rate of approximation schemes for Hamilton–Jacobi–Bellman equations*, M2AN Math. Model. Numer. Anal. **36** (2002), no. 1, 33–54. MR1916291 (2003h:65142)

[2] ———, *Error bounds for monotone approximation schemes for Hamilton–Jacobi–Bellman equations*, Preprint, December 2003; SIAM J. Numer. Anal. **43** (2005), 540–558. MR2177879

[3] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Appl. Math. (New York), vol. 25, Springer-Verlag, New York, 1993. MR1199811 (94e:93004)

[4] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1964. MR0181836 (31:6062)

[5] N. V. Krylov, *Controlled diffusion processes*, "Nauka", Moscow, 1977; English transl., Appl. Math., vol. 14, Springer-Verlag, New York–Berlin, 1980. MR0508417 (80f:60046); MR0601776 (82a:60062)

[6] ———, *Nonlinear elliptic and parabolic equations of the second order*, "Nauka", Moscow, 1985; English transl., Math. Appl. (Soviet Ser.), vol. 7, D. Reidel Publishing Co., Dordrecht, 1987. MR0815513 (87h:35002); MR0901759 (88d:35005)

[7] _____, *On the general notion of fully nonlinear second-order elliptic equations*, Trans. Amer. Math. Soc. **347** (1995), no. 3, 857–895. MR1284912 (95f:35075)

[8] _____, *On the rate of convergence of finite-difference approximations for Bellman's equations*, Algebra i Analiz **9** (1997), no. 3, 245–256; English transl., St. Petersburg Math. J. **9** (1998), no. 3, 639–650. MR1466804 (98h:49033)

[9] _____, *Approximating value functions for controlled degenerate diffusion processes by using piecewise constant policies*, Electron. J. Probab. **4** (1999), no. 2, 19 pp.; http://www.math.washington.edu/~ejpecp/EjpVol4/paper2.abs.html. MR1668597 (2000b:49056)

[10] _____, *Mean value theorems for stochastic integrals*, Ann. Probab. **29** (2001), no. 1, 385–410. MR1825155 (2002c:60090)

[11] M. V. Safonov, *The classical solution of the elliptic Bellman equation*, Dokl. Akad. Nauk SSSR **278** (1984), no. 4, 810–813; English transl., Soviet Math. Dokl. **30** (1984), 482–485. MR0765302 (86f:35081)

127 Vincent Hall, University of Minnesota, Minneapolis, Minnesota 55455
E-mail address: `hjdong@math.umn.edu`

127 Vincent Hall, University of Minnesota, Minneapolis, Minnesota 55455
E-mail address: `krylov@math.umn.edu`