

AN ADAPTIVE GRID SCHEME FOR THE DISCRETE HAMILTON-JACOBI-BELLMAN EQUATION

revised version

Lars Grüne*
Institut für Mathematik
Universität Augsburg
Universitätsstr. 8
86135 Augsburg
Germany

Abstract. In this paper an adaptive finite difference scheme for the solution of the discrete first order Hamilton-Jacobi-Bellman equation is presented. Local a posteriori error estimates are established and certain properties of these estimates are proved. Based on these estimates an adapting iteration for the discretization of the state space is developed. An implementation of the scheme for two-dimensional grids is given and numerical examples are discussed.

Keywords. discrete Hamilton Jacobi Bellman equation, local error estimates, adaptive grids

Mathematics subject classification. 65N50, 49L25, 49M25

1 Introduction

In this paper an adaptive grid scheme for the solution of the discrete first order Hamilton-Jacobi-Bellman equation

$$\sup_{u \in U} \{v_h(x) - \beta v_h(\Phi_h(x, u)) - hg(x, u)\} = 0$$

on $\Omega \subset \mathbb{R}^n$ with $0 < \beta < 1$ is developed. Here Φ_h is the right hand side of a discrete time control system and g is the so called cost function.

For $\beta = 1 - \delta h$ the solution v_h of this equation is the optimal value function of the *discrete discounted optimal control problem*

$$\min_{u_x(\cdot) \in \mathcal{U}_h} J_h(x, u_x(\cdot)), \quad \text{with } J_h(x, u_x(\cdot)) := h \sum_{i=0}^{\infty} \beta^i g(x_i, u_i) \quad (1.1)$$

*E-Mail: Lars.Gruene@Math.Uni-Augsburg.de

where $x_{i+1} = \Phi_h(x_i, u_i)$, $\mathcal{U}_h := \{u : \mathbb{R} \rightarrow U : u \text{ is constant on } [ih, (i+1)h)\}$, $u_i = u(ih)$ and $U \subset \mathbb{R}^m$ is the *set of admissible control values*. The value $\delta > 0$ is called the *discount rate*.

Control problems of this type arise e.g. by discretization of continuous time discounted optimal control systems (see e.g. [3], [10]) or when time discrete control systems are considered, i.e. $h = 1$ and Φ_h is the right hand side of the time discrete control system. The setup is also related to Markov Chain Approximations of discounted continuous time stochastic control problems (cp. [15, Section 4.2]).

To solve this equation numerically a finite difference scheme is used, for which a discretization of Ω is necessary. The convergence of this method for the case of an equidistant discretization with node distance k of Ω is analyzed e.g. in [10] and [12], algorithms to solve the discretized equation are proposed e.g. in [9] and [11].

For $\delta \rightarrow 0$ the optimal value function of the continuous time discounted optimal control problem converges to the optimal value function of the *average cost optimal control problem* (cp. [4], [19], [13, Section 2]). Therefore in order to approximate this optimal value function it is necessary to solve (1.1) for small discount rates $\delta > 0$. However, it turns out that in this case the convergence rate for $k \rightarrow 0$ is very small. It can only be proved that $\|v_h - v_h^k\|_\infty < Ck^{\frac{\gamma}{2}}$ with $\gamma = \frac{D}{\delta}$ for constants $C, D > 0$. Furthermore for certain optimal control problems and small $\delta > 0$ the value function is almost constant on large parts of the state space with steep areas around certain subsets of Ω (cp. Section 4). This leads to the idea to use a coarse discretization of Ω with refinement on certain "critical regions", which means an adaptive grid for the discretization of Ω to calculate a good approximation of v_h .

In Section 2 we will introduce *local a posteriori error estimates* e_i corresponding to the simplices $S_i, j = i, \dots, P$ of the discretization Ξ of Ω . We will prove the existence of constants $C_1, C_2 > 0$ such that

$$C_1 \sup_{i=1, \dots, P} e_i \leq \|v_h - v_h^\Xi\|_\infty \leq C_2 \sup_{i=1, \dots, P} e_i.$$

Error estimates with similar properties have been investigated for Finite Element Methods (see e.g. [1], [17]) and for Boundary Element Methods (see e.g. [17], [18] and [8]). It will also be proved that $e_i \rightarrow 0$ as the size of the corresponding simplex S_j converges to 0.

In Section 3 first an algorithm to solve the discretized equation is developed, then the adaptive grid scheme is presented. The adapting iteration – based on numerically calculated approximations \tilde{e}_i of the error estimates e_i – is given for the general n -dimensional case and a refinement and coarsening method is described for two-dimensional grids.

Finally, in Section 4 numerical examples are discussed.

2 Local error estimates

We want to solve the discrete Hamilton Jacobi Bellman equation

$$\sup_{u \in U} \{v_h(x) - \beta v_h(\Phi_h(x, u)) - hg(x, u)\} = 0 \tag{2.1}$$

on $\Omega \subset \mathbb{R}^n$ with $\beta := 1 - \delta h$ and $\delta h < 1$. We assume that there exist constants $L_g, L_h, M_g, M_h > 0$ such that

$$|g(x, u) - g(y, u)| \leq L_g \|x - y\| \quad \text{and} \quad \|\Phi_h(x, u) - \Phi_h(y, u)\| \leq L_h \|x - y\|$$

as well as

$$|g(x, u)| < M_g \quad \text{and} \quad \|\Phi_h(x, u) - x\| \leq M_h$$

for all $x, y \in \Omega$ and all $u \in U$ where $U \subset \mathbb{R}^m$ is compact.

The operator T_h related to (2.1) is defined by

$$T_h(v_h)(x) = \inf_{u \in U} \{\beta v_h(\Phi_h(x, u)) + hg(x, u)\}. \quad (2.2)$$

T_h is a contraction and v_h is the unique bounded solution of the fixed point equation $v_h(x) = T_h(v_h)(x)$ for all $x \in \Omega$ (cp. [2]).

In order to reduce this equation to a finite dimensional problem we use a triangulation Ξ of Ω into a finite number P of simplices S_j with N nodes x_i and look for the solution of (2.1) in the space of piecewise affine functions $W := \{w \in C(\Omega) \mid \nabla w(x) = c_j \text{ in } S_j\}$.

Every point $\Phi_h(x_i, u)$ can be written as a convex combination of the nodes or vertices x_j of the simplex containing it with coefficients $\lambda_{ij}(u)$. Let $\Lambda(u) := [\lambda_{ij}(u)]_{i,j=1,\dots,N}$ be the matrix containing these coefficients and $G(u) := [g(x_i, u)]_{i=1,\dots,N}$ an N -dimensional vector. Now we can use (2.2) to obtain a fixed point equation in the nodes of the triangulation

$$V = T_h^\Xi(V), \quad T_h^\Xi(V) := \inf_{u \in U} (\beta \Lambda(u)V + hG(u)) \quad (2.3)$$

T_h^Ξ is a contraction in \mathbb{R}^N with contraction factor $\beta = 1 - \delta h$ and therefore has a unique fixed point V^* . Let v_h^Ξ be the function obtained by linear interpolation of V^* . This function v_h^Ξ can be used to calculate control functions for the optimal control problem (1.1): Let $x \in \Omega$ and $u^x \in U$ such that

$$\beta v_h^\Xi(\Phi_h(x, u^x)) + hg(x, u^x) = \min_{u \in U} \{\beta v_h^\Xi(\Phi_h(x, u)) + hg(x, u)\} \quad (2.4)$$

Then for any initial value $x_0 \in \Omega$ we can define a control function $u_{x_0} \in \mathcal{U}_h$ by $u_{x_0, i} = u^{x_i}$, $x_{i+1} = \Phi_h(x_i, u_{x_0, i})$.

In order to estimate how good v_h^Ξ approximates v_h we define *a posteriori* error estimates that are based on local values of the residual.

Definition 2.1 For every point $x \in \Omega$ and every simplex $S_i \in \Xi$ we define

$$e(x) := |v_h^\Xi(x) - T_h(v_h^\Xi)(x)| \quad \text{and} \quad e_i := \max_{x \in S_i} e(x).$$

These error estimates have local and global properties as stated in the following two theorems. The global properties of these estimates are given in the first theorem:

Theorem 2.2 Let e_i be the error estimates as defined in Definition 2.1. Then the following inequality holds:

$$\frac{1}{2} \max_{i=1,\dots,P} e_i \leq \sup_{x \in \Omega} |v_h(x) - v_h^{\Xi}(x)| \leq \frac{1}{\delta h} \max_{i=1,\dots,P} e_i.$$

Proof: Observe that for any two functions $v, \tilde{v} : \Omega \rightarrow \mathbb{R}$ and any $x \in \Omega$ the operator T_h^{Ξ} satisfies

$$|T_h(v)(x) - T_h(\tilde{v})(x)| \leq \beta \sup_{y \in \Phi_h(x,U)} |v(y) - \tilde{v}(y)|. \quad (2.5)$$

Since $T_h(v_h) = v_h$ it follows for all $x \in \Omega$ that

$$\begin{aligned} |v_h^{\Xi}(x) - T_h(v_h^{\Xi})(x)| &= |v_h^{\Xi}(x) - v_h(x) + T_h(v_h)(x) - T_h(v_h^{\Xi})(x)| \\ &\leq |v_h^{\Xi}(x) - v_h(x)| + |T_h(v_h)(x) - T_h(v_h^{\Xi})(x)| \\ &\leq 2 \sup_{y \in \Omega} |v_h^{\Xi}(y) - v_h(y)| \end{aligned}$$

where the last inequality follows by (2.5). This shows the first inequality.

Conversely for all $x \in \Omega$ we have

$$\begin{aligned} |v_h(x) - v_h^{\Xi}(x)| &= |T_h(v_h^{\Xi})(x) - v_h^{\Xi}(x) + T_h(v_h)(x) - T_h(v_h^{\Xi})(x)| \\ &\leq |T_h(v_h^{\Xi})(x) - v_h^{\Xi}(x)| + |T_h(v_h)(x) - T_h(v_h^{\Xi})(x)| \\ &\leq e(x) + \beta \sup_{y \in \Omega} |v_h(y) - v_h^{\Xi}(y)| \end{aligned}$$

where the last inequality again follows by (2.5). This implies the second inequality since $\beta = 1 - \delta h$. \square

The construction of the error estimates also allows some local results as stated in the following theorem.

Theorem 2.3 Let $e(x)$ be the error estimates as defined in Definition 2.1. Then

$$\frac{1}{2} e(x) \leq \sup_{y \in U_{M_h}(x)} |v_h(y) - v_h^{\Xi}(y)|$$

where $U_{M_h}(x)$ denotes the ball around x with radius M_h .

Conversely if $K \subset \Omega$ satisfies $\Phi_h(x, u^x) \in K$ for all $x \in K$ and the control values u^x from (2.4) then

$$v_h(x) \leq v_h^{\Xi}(x) + \frac{1}{\delta h} \sup_{y \in K} e(y) \text{ for all } x \in K,$$

which means that in this case $v_h^{\Xi} + \frac{1}{\delta h} \sup_{y \in K} e(y)$ is an upper bound for v_h .

Proof: With the same arguments as in the proof of Theorem 2.2 where the suprema are taken over $U_{M_h}(x)$ and K , respectively. \square

Remark 2.4 The opposite inequality for the second estimate can be obtained if we assume invariance of K with respect to the control values where (2.1) attains its minimum. However, this property is difficult to check.

The local error estimates can also be used to give an estimate for the functional along the calculated trajectory as defined in (2.4).

Theorem 2.5 Let $u_x(\cdot)$ be the discrete control function as defined in (2.4). Let $(x_i)_{i \in \mathbb{N}}$ be the corresponding discrete trajectory. Then

$$|J_h(x, u_x(\cdot)) - v_h^{\Xi}(x)| \leq \sum_{i=0}^{\infty} \beta^i e(x_i).$$

In particular if (x_i) stays inside some set $K \subset \Omega$ such that $e(x) < \varepsilon$ for all $x \in K$ it follows that

$$|J_h(x, u_x(\cdot)) - v_h^{\Xi}(x)| \leq \frac{\varepsilon}{\delta h}.$$

Proof: By definition of $e(x)$ we know that $|hg(x_i, u_i) - (v_h^{\Xi}(x_i) - \beta v_h^{\Xi}(x_{i+1}))| \leq e(x_i)$. Hence

$$\begin{aligned} J_h(x, u_x(\cdot)) &= h \sum_{i=0}^{\infty} \beta^i g(x_i, u_i) \\ &\leq \sum_{i=0}^{\infty} \beta^i (v_h^{\Xi}(x_i) - \beta v_h^{\Xi}(x_{i+1}) + e(x_i)) \\ &= v_h^{\Xi}(x) + \sum_{i=0}^{\infty} \beta^i e(x_i) \end{aligned}$$

where $\sum_{i=0}^{\infty} \beta^i = \frac{1}{\delta h}$. Since the opposite inequality can be obtained in the same way the assertion follows. \square

Remark 2.6 In this theorem we assume the worst case, i.e. that all errors e_i along x_i accumulate. Since this is not necessarily true, the value $J_h(x, u_x)$ may be smaller than $v_h^{\Xi}(x)$. In this case $J_h(x, u_x)$ can be used as an approximation for $v_h(x)$ and the inequality

$$v_h(x) \leq J_h(x, u_x) < v_h^{\Xi}(x)$$

holds (cp. second example in Section 4).

We will now discuss how the local error estimate e_i depends on the size of the corresponding simplex S_i . To do this we need a result about the continuity of v_h^{Ξ} which can be obtained by looking at the geometry of the grid. For this purpose the grid has to satisfy a *regularity condition*.

Definition 2.7 A grid Ξ satisfies the *regularity condition* with regularity constant $c_\Xi > 0$ if for each simplex S_m with nodes $x_{p_1}, \dots, x_{p_{n+1}}$ the inequality

$$\|x_{p_i} - x_{p_j}\| \leq c_\Xi \|x_{p_l} - \sum_{\substack{k=1, \dots, n+1 \\ k \neq l}} \lambda_k x_{p_k}\|$$

holds for all $i, j, l = 1, \dots, n+1$ and all $\lambda_k \geq 0$ with $\sum_{\substack{k=1 \\ k \neq l}}^{n+1} \lambda_k = 1$.

This condition guarantees that the distance between a node and the opposite edge cannot become arbitrarily small in relation to the distance of two nodes of this simplex. In the two-dimensional case this means that no triangles with arbitrarily small angles appear in the triangulation.

To formulate the consequences of this condition we introduce the following notation:

For each simplex S_l of Ξ denote by $\nabla v_h^\Xi|_{S_l}$ the gradient of v_h^Ξ on S_l (which is constant) and let $\|\nabla v_h^\Xi\|_\infty$ be the maximum over all simplices S_l of $\|\nabla v_h^\Xi|_{S_l}\|$. Furthermore for any two nodes $x_i \neq x_j$ belonging to one simplex S_l of Ξ we define the directional derivative along the edge connecting x_i and x_j by

$$d_{ij} := \frac{v_h^\Xi(x_i) - v_h^\Xi(x_j)}{\|x_i - x_j\|}.$$

Using this notation the following lemma shows a consequence of the regularity condition.

Lemma 2.8 Let v_h^Ξ be the continuous and piecewise affine solution of (2.3) corresponding to a grid Ξ satisfying the regularity condition with constant c_Ξ . Then

$$\|\nabla v_h^\Xi\|_\infty \leq c_\Xi \max |d_{ij}|$$

where the maximum is taken over all simplices S_l of Ξ and all $i \neq j$ with x_i, x_j belonging to S_l .

Proof: For any simplex S_l there exist points x, y in S_l such that $\|\nabla v_h^\Xi|_{S_l}\| = \frac{|v_h^\Xi(x) - v_h^\Xi(y)|}{\|x - y\|}$. Now consider the line from $(x, v_h^\Xi(x))$ to $(y, v_h^\Xi(y))$ in \mathbb{R}^{n+1} . This line can be shifted parallelly and scaled to a line from $(x_{p_j}, v_h^\Xi(x_{p_j}))$ to $(z, v_h^\Xi(z))$ where x_{p_j} is a node of S_l and $z = \sum_{\substack{i=1, \dots, n+1 \\ i \neq j}} \lambda_i x_{p_i}$ is a convex combination of the other nodes of S_l .

Since $v_h^\Xi|_{S_l}$ is affine (i.e. it can be written as $v_h^\Xi(x) = c^T x + b$ where $c = \nabla v_h^\Xi|_{S_l}$) these operations do not change the difference quotient:

$$\frac{|v_h^\Xi(x) - v_h^\Xi(y)|}{\|x - y\|} = \frac{|v_h^\Xi(x_{p_j}) - v_h^\Xi(z)|}{\|x_{p_j} - z\|}.$$

We now pick nodes x_k and x_m from S_l such that $|v_h^\Xi(x_k) - v_h^\Xi(x_m)|$ becomes maximal. Then

$$|v_h^\Xi(x_{p_j}) - v_h^\Xi(z)| \leq |v_h^\Xi(x_k) - v_h^\Xi(x_m)|$$

and the regularity condition (Definition 2.7) yields

$$c_{\Xi} \|x_{p_j} - z\| \geq \|x_k - x_m\|.$$

Hence

$$\frac{|v_h^{\Xi}(x) - v_h^{\Xi}(y)|}{\|x - y\|} = \frac{|v_h^{\Xi}(x_{p_j}) - v_h^{\Xi}(z)|}{\|x_{p_j} - z\|} \leq c_{\Xi} \frac{|v_h^{\Xi}(x_k) - v_h^{\Xi}(x_m)|}{\|x_k - x_m\|} = c_{\Xi} |d_{km}|$$

and the assertion follows. \square

Using this lemma we can prove that the function v_h^{Ξ} is *Hölder continuous*.

Theorem 2.9 For any grid Ξ with $c_{\Xi} < C$ and any two points $x, y \in \Omega$ the inequality

$$|v_h^{\Xi}(x) - v_h^{\Xi}(y)| \leq L \|x - y\|^{\gamma}$$

holds for constants $L, \gamma > 0$ independent of Ξ .

Proof: We use the contraction operator T_h^{Ξ} to prove this property. The function v_h^{Ξ} can be obtained by the iteration procedure

$$v_0 \equiv 0 \quad v_{k+1} := T_h^{\Xi}(v_k) \quad (2.6)$$

Since the contraction factor β of T_h^{Ξ} satisfies $\beta < e^{-\delta h}$ and v_h^{Ξ} is bounded by M_g/δ it follows that

$$\|v_k - v_h^{\Xi}\|_{\infty} \leq \beta^k \|v_0 - v_h^{\Xi}\|_{\infty} \leq e^{-\delta h k} \frac{M_g}{\delta}$$

which can be written as

$$\|v_k - v_h^{\Xi}\|_{\infty} \leq \int_{hk}^{\infty} e^{-\delta t} M_g dt. \quad (2.7)$$

By induction we will now prove that v_k is Lipschitz continuous: Let L_k denote the corresponding Lipschitz constant. Clearly $L_0 = 0$ holds. For any two vertices $x_i \neq x_j$ belonging to one simplex of the grid we have

$$|v_k(x_i) - v_k(x_j)| = |T_h^{\Xi}(v_{k-1}(x_i)) - T_h^{\Xi}(v_{k-1}(x_j))| \leq \beta L_{k-1} L_h \|x_i - x_j\| + h L_g \|x_i - x_j\|$$

hence $|d_{ij}| \leq \beta L_{k-1} L_h + h L_g$. The Mean Value Theorem and Lemma 2.8 now imply

$$L_k \leq \|\nabla v_h^{\Xi}\|_{\infty} \leq C(\beta L_{k-1} L_h + h L_g)$$

and hence with $\alpha := \frac{\ln(CL_h)}{h}$ and $b := \max\{e^{h(\delta-\alpha)}, 1\}$ it follows

$$L_k \leq h C L_g \sum_{j=0}^{k-1} e^{jh(\alpha-\delta)} \leq b \int_0^{h(k-1)} C L_g e^{(\alpha-\delta)t} dt \quad (2.8)$$

Putting (2.7) and (2.8) together one obtains

$$|v_h^{\Xi}(x) - v_h^{\Xi}(y)| \leq b \int_0^{\infty} e^{-\delta t} \Phi(t) dt$$

where $\Phi(t) \leq \min\{C L_g e^{\alpha t} \|x - y\|, 2M_g\}$. Now [3, Lemma 4.1] can be used to give an estimate for this integral and the assertion follows. \square

Corollary 2.10 The local error estimates $e(x)$ and e_i from Definition 2.1 satisfy

$$|e(x) - e(y)| \leq L_e \|x - y\|^\gamma \quad \text{and} \quad e_i \leq L_e \max\{\|x - y\| : x, y \in S_i\}^\gamma.$$

Proof: The first assertion follows directly from Theorem 2.9, and the second from the first since $e(x_i) = 0$ for all nodes x_i of Ξ . \square

Remark 2.11 The rate of convergence of $e_i \rightarrow 0$ for small discount rates δ is very slow, since in this case we have $\gamma = D/\delta$ for some constant $D > 0$ (which again follows from the estimate in [3, Lemma 4.1]). An example of a problem with small discount rate is shown in Section 4.

Since $\alpha \rightarrow \infty$ as $h \rightarrow 0$ in the proof of Theorem 2.9 the convergence may also become slow for small parameters $h > 0$. However, the estimate $|v_h^\Xi - v_{\tilde{h}}^\Xi| \leq M_g(h + \tilde{h} + |h - \tilde{h}|)$ that holds for sufficiently small $h, \tilde{h} > 0$ (cp. [12, Section 3.4.3]) guarantees that the error estimates do not converge to infinity for $h \rightarrow 0$.

The next lemma gives a property that will be used to coarsen the grid.

Lemma 2.12 Let $V \in \mathbb{R}^N$ be a vector with $\|T_h^\Xi(V) - V\|_\infty < \varepsilon$. Then the solution $V^* \in \mathbb{R}^N$ of (2.3) satisfies

$$\|V - V^*\|_\infty < \frac{\varepsilon}{\delta h}.$$

Proof: For some $u \in U$ we have

$$\begin{aligned} |[V - V^*]_i| &\leq \beta |[\Lambda(u)(V - V^*)]_i| + \varepsilon \\ &\leq \beta \|V - V^*\|_\infty + \varepsilon \end{aligned}$$

and hence

$$\|V - V^*\|_\infty \leq \beta \|V - V^*\|_\infty + \varepsilon$$

which implies the assertion since $\beta = 1 - \delta h$. \square

3 Adaptive Grids

The previous section gives us all properties to construct adaptive grids based on local values of the residual. We will now present an algorithm that solves the equation (2.3) taking advantage of a solution of (2.3) corresponding to a different grid.

To describe the algorithm we have to introduce some notation: Let $V \in \mathbb{R}^N$ be an N -dimensional vector. Let $[V]_i$ denote the i -th component of this vector and $F_i(V, w) \in \mathbb{R}^N$ denote the vector that is obtained by replacing the i -th component of V by w , i.e.

$$F_i(V, w) := ([V]_1, \dots, [V]_{i-1}, w, [V]_{i+1}, \dots, [V]_N)^T$$

With this notation we can define an operator S_i as follows.

Definition 3.1 Let $V \in \mathbb{R}^N$ be a vector and $i \in \{1, \dots, N\}$. Let $w \in \mathbb{R}$ such that

$$[T_h^\Xi(F_i(V, w))]_i = w \quad \text{and define} \quad S_i(V) := F_i(V, w).$$

The uniqueness of w follows from the fact that T_h^Ξ is a contraction. Using this operator we can perform either a total step (Step 2a) or a single step (Step 2b) iteration:

Step 1: Take $V_0 \in \mathbb{R}^N$ arbitrary, let $n = 0$

Step 2a: Let $[V_{n+1}]_i := [S_i(V_n)]_i$ for $i = 1, \dots, N$

or

Step 2b: Let $V_n^0 := V_n$, compute sequentially $V_n^i = S_i(V_n^{i-1})$ for $i = 1, \dots, N$ and let $V_{n+1} := V_n^N$

Step 3: If $\|V_n - V_{n+1}\| < \delta_{end}$ stop, else let $n := n + 1$ and continue with Step 2 (a or b)

Remark 3.2 The value w in Definition 3.1 can be determined as follows: We are looking for a $w \in \mathbb{R}$ such that

$$w = \min_{u \in U} \left\{ \beta \sum_{\substack{j=1, \dots, N \\ j \neq i}} \lambda_{ij}(u) [V]_j + \beta \lambda_{ii}(u) w + h G_i(u) \right\}.$$

It is easily seen that this is equivalent to

$$w = \min_{u \in U} \left\{ \frac{\beta \sum_{\substack{j=1, \dots, N \\ j \neq i}} \lambda_{ij}(u) [V]_j + h G_i(u)}{1 - \beta \lambda_{ii}(u)} \right\}$$

cp. [13, Section 4]

If the initial vector V_0 lies in the set \mathcal{V} of monotone convergence of T_h^Ξ (as defined in [13, Section 4]) the single step iteration is exactly the increasing coordinate algorithm described there. The main difference lies in the fact, that here no assumptions on V_0 are imposed.

The convergence of the algorithm is proved in the following lemma.

Lemma 3.3 Let $S(V)$ denote the vector obtained by applying Step 2 (a or b) for $i = 1, \dots, N$ to a vector $V \in \mathbb{R}^N$. Then for any two vectors $V_1, V_2 \in \mathbb{R}^N$ one has

$$\|S(V_1) - S(V_2)\|_\infty \leq \beta \|V_1 - V_2\|_\infty$$

and hence $S : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a contraction and the algorithm converges to a unique fixed point of S which then is also a fixed point of T_h^Ξ .

Proof: We first show an estimate for the components. Fix $i \in \{1, \dots, N\}$. Then for some $\bar{u} \in U$ we have

$$\begin{aligned} |[S_i(V_1)]_i - [S_i(V_2)]_i| &= |[T_h^{\Xi}(S_i(V_1))]_i - [T_h^{\Xi}(S_i(V_2))]_i| \\ &\leq \beta \left| \sum_{j=1, \dots, N} \lambda_{ij}(\bar{u}) [S_i(V_1) - S_i(V_2)]_j \right| \\ &\leq \beta \max_{j=1, \dots, N} |[S_i(V_1) - S_i(V_2)]_j|. \end{aligned}$$

If this maximum is attained for $j = i$ then it follows that $[S_i(V_1)]_i - [S_i(V_2)]_i = 0$ and hence $\max_{j=1, \dots, N} |[S_i(V_1) - S_i(V_2)]_j| = 0$. This implies

$$|[S_i(V_1)]_i - [S_i(V_2)]_i| \leq \beta \max_{\substack{j=1, \dots, N \\ j \neq i}} |[S_i(V_1) - S_i(V_2)]_j| = \beta \max_{\substack{j=1, \dots, N \\ j \neq i}} |[V_1 - V_2]_j| \leq \beta \|V_1 - V_2\|_{\infty}$$

Now the assertion follows directly for both iterations. \square

Remark 3.4 Except for the case of monotone convergence (cp. [13, Lemma 4.3]) it is not clear that this iteration is faster than the iteration using T_h^{Ξ} . However, the estimate for the operator S_i from Step 2 (a or b) as shown in the proof of Lemma 3.3

$$|[S_i(V_1)]_i - [S_i(V_2)]_i| \leq \beta \max_{\substack{j=1, \dots, N \\ j \neq i}} |[V_1 - V_2]_j|$$

is better than the corresponding estimate for T_h^{Ξ}

$$|[T_h^{\Xi}(V_1)]_i - [T_h^{\Xi}(V_2)]_i| \leq \beta \max_{j=1, \dots, N} |[V_1 - V_2]_j|.$$

Numerical tests have shown that the iterations using S_i are considerably faster.

Remark 3.5 The acceleration method developed by R.L.V. Gonzáles and C.A. Sagastizábal [11] can also be applied to this iteration and shows good results.

With this algorithm it is possible to use a solution $v_h^{\Xi_1}$ corresponding to a grid Ξ_1 as an initial value for the iteration using a new grid Ξ_2 . An initial vector V_0 is obtained by interpolation: $[V_0]_i = v_h^{\Xi_1}(x_i)$ for all nodes x_i of Ξ_2 .

Corollary 2.10 implies that $e_i = \max_{x \in S_i} e(x)$ can be approximated by evaluating $\tilde{e}_i := \max_{y_k \in S_i} e(y_k)$ for finitely many points $y_k \in S_i$. Clearly, these points y_i should not coincide with the nodes x_i of the grid since the identity $e(x_i) = 0$ is obvious. The points y_k used in the examples discussed in Section 4 are described there.

The error estimates \tilde{e}_i can now be calculated numerically and can be used as a criterion to refine all simplices S_i that "belong" to large \tilde{e}_i .

To refine and to coarsen simplices in the two-dimensional case we use a grid scheme proposed by R. Löhner [16] for the Finite Element Method. To refine a simplex we proceed

as follows:

For any two nodes x_j, x_k belonging to S_i we calculate the new node

$$x_{new}(x_j, x_k) := \frac{1}{2}(x_j + x_k).$$

Thus every edge of S_i is divided into two parts of the same length. Using this new nodes S_i can be divided into 4 new simplices. This refinement is called *regular*.

To make sure that v_h^{Ξ} is continuous it is necessary that every simplex has exactly one neighbour on any edge. To guarantee this property without introducing arbitrarily "flat" simplices (which would cause arbitrarily large regularity constants c_{Ξ} , cp. Definition 2.7) we use the following neighbourhood condition:

Any simplex may have at most one edge with more than one neighbour; at this edge there may be at most two neighbours.

All simplices not satisfying this condition are refined regularly as described above and we end up with a grid that only consists of regularly refined simplices satisfying this condition. In the next step all simplices with two neighbours on one edge are divided into two simplices at this edge. Thus any of the new simplices becomes a unique neighbour of one of the old neighbouring simplices. This partition into two new simplices is called *irregular* refinement. Figure 3.1 illustrates this principle: in the first step all regular refinements are inserted, in the second step the irregular refinements are added.

Figure 3.1: Neighbour refinement

An irregular subsimplex will not be refined further. If this is necessary – i.e. if the corresponding error estimate is large – then the irregular refinement will be reversed and the upper simplex will be refined regularly. This guarantees that the grid constant c_{Ξ} will not increase during the refinement procedure.

Lemma 2.12 is used to give a criterion to coarsen a previously refined simplex S_i : If

$$|v_h^{\Xi}(x_{new}(x_j, x_k)) - \frac{1}{2}(v_h^{\Xi}(x_j) + v_h^{\Xi}(x_k))| < \varepsilon \quad (3.1)$$

for all nodes x_j, x_k of S_i then this simplex will be coarsened. From Lemma 2.12 it follows that the solution corresponding to the coarsened grid Ξ_1 and the solution corresponding to the refined grid Ξ_2 satisfy

$$\|v_h^{\Xi_1} - v_h^{\Xi_2}\|_{\infty} < \frac{\varepsilon}{\delta h}.$$

Using this techniques the grid adapting iteration can be described as follows

Step 1: Take an initial grid Ξ_0 covering Ω and refinement parameters $tol > 0, \Theta \in (0, 1)$; let $n=0$

Step 2: Calculate the solution $v_h^{\Xi_n}$ corresponding to the grid Ξ_n and the error estimates \tilde{e}_i , $i = 1, \dots, P$. If $\sup_{i \in \{1, \dots, P\}} \tilde{e}_i < tol$ stop, else refine all simplices S_i corresponding to error estimates $\tilde{e}_i > \Theta \sup_{i \in \{1, \dots, P\}} \tilde{e}_i$

Step 3: Coarsen all simplices S_i that satisfy (3.1) (with $\varepsilon = tol$) and have not been refined in Step 2

Step 4: Calculate a new grid Ξ_n by inserting all necessary neighbour refinements, let $n := n + 1$ and continue with Step 2

By virtue of Corollary 2.10 the refining iteration (i.e. the iteration without Step 3) will terminate with a grid Ξ and a solution satisfying $\|v_h^{\Xi} - v_h\|_{\infty} < \frac{tol}{\delta h}$. The coarsening (Step 3) destroys the monotonicity of the adapting procedure and therefore convergence is no longer guaranteed. However, Lemma 2.12 yields that the difference between the coarsened and the refined solution is small. The numerical examples in the next section show that - using a suitable parameter tol - this step helps to reduce the number of simplices and vertices during the adapting iteration without increasing the error too much.

As already mentioned in Remark 2.11 the convergence of $e_i \rightarrow 0$ may be very slow for small discount rates δ . In the next section we will also discuss an example where this is the case.

4 Numerical examples

The system discussed here is a three-dimensional linear oscillator given by

$$\dot{y} = A(u)y \quad \text{with} \quad A(u) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -(c+u) & -b & -a \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad (4.1)$$

The solutions of this - so called bilinear - control system are projected onto the projective space \mathbb{P}^2 (cp. [7]) which is identified with one half of the unit sphere \mathbb{S}^2 . The system is discretized by an Euler method on \mathbb{S}^2 with time step h , i.e. performing an Euler step in \mathbb{R}^3 and projecting the solution back to \mathbb{S}^2 . Hence we have

$$\Phi_h^{\mathbb{S}^2}(s, u) := \frac{s + hA(u)s}{\|s + hA(u)s\|}.$$

The cost function g is given by

$$g(s, u) = (-c - u)s_1s_3 + s_1s_2 + (1 - b)s_2s_3 - as_3^2 \quad (4.2)$$

with $s = (s_1, s_2, s_3)^T \in \mathbb{S}^2$. Using similar techniques as in [3] on \mathbb{S}^2 it can be shown that the value function v_h of the discretized system converges to the value function v_{δ} of the original system for $h \rightarrow 0$.

To obtain a region $\Omega \subset \mathbb{R}^2$ that can be divided into simplices S_i we use the stereographic projection to map a part of \mathbb{S}^2 containing the upper half onto $[-1, 1] \times [-1, 1] =: \Omega$. With this projection we obtain the function $\Phi_h(x, u)$ on Ω from the discretization $\Phi_h^{\mathbb{S}^2}(s, u)$ on \mathbb{S}^2 defined above.

For $\delta \rightarrow 0$ the value function δv_δ converges to a discontinuous function λ^* . For suitable parameters $a, b, c \in \mathbb{R}$ there exists a set $D_1 \subset \mathbb{P}^2$ (a so called open control set) such that λ^* is constant on D_1 , constant outside clD_1 with a different value and hence jumps at ∂D_1 . $\lambda^*(s)$ is the minimal Lyapunov exponent of (4.1) with initial value s , cp. [13, Section 2]. (For more information about projected bilinear control systems and Lyapunov exponents see e.g. [6], [7], [5].) Clearly this convergence cannot be uniform for the whole state space but it can be shown that δv_δ converges uniformly on compact subsets of $intD_1$ and on compact subsets of $int(\mathbb{P}^2 \setminus D_1)$ (cp. [13]). This implies that for small $\delta > 0$ δv_δ will be steep around ∂D_1 . Figure 4.1 shows the boundary of D_1 , it was computed using the program CS2DIM by G. Häckl [14].

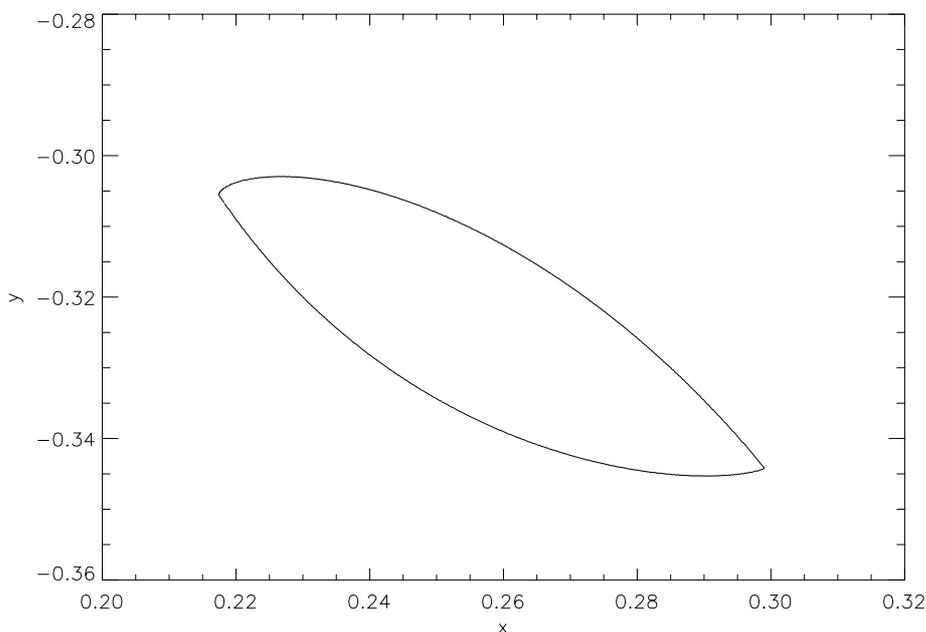


Figure 4.1: Control set D_1

All examples were calculated with $a = 1, b = 0, c = 0.5, U = \{-0.3, -0.25, \dots, 0.3\}, h = 0.05$ and $\Theta = 0.1$. The first example is calculated with discount rate $\delta = 1$. The initial grid was chosen with $N = 169$ vertices. Table 4.1 shows the number of vertices for the adaptive iteration with coarsening (N_1) ($tol = 0.003$) and without coarsening (N_2) and the corresponding error estimates err_1, err_2 i.e. the maximum of the error estimates \tilde{e}_i of the solutions. Here \tilde{e}_i is calculated using the following four points $y_k, k = 1, \dots, 4$ in each simplex: If $x_{l_1}, x_{l_2}, x_{l_3}$ denote the vertices of the simplex then

$$\begin{aligned} y_1 &= \frac{1}{10}x_{l_1} + \frac{9}{20}(x_{l_2} + x_{l_3}), & y_2 &= \frac{1}{10}x_{l_2} + \frac{9}{20}(x_{l_1} + x_{l_3}) \\ y_3 &= \frac{1}{10}x_{l_3} + \frac{9}{20}(x_{l_1} + x_{l_2}), & y_4 &= \frac{1}{3}(x_{l_1} + x_{l_2} + x_{l_3}) \end{aligned}$$

Figure 4.2 shows the value functions for the initial grid, Figure 4.3 for the adapted grid (with coarsening).

$\#iter$	N_1	err_1	N_2	err_2
0	169	0.027048	169	0.027048
1	576	0.021922	576	0.021922
2	1589	0.018499	1597	0.018499
3	2641	0.023188	2907	0.023188
4	2501	0.020628	3291	0.020701
5	2958	0.019959	3636	0.019937
6	3280	0.006341	4016	0.006337
7	10167	0.006775	11401	0.006815
8	9257	0.003953	14542	0.002795
9	18849	0.003176	33322	0.003256

Table 4.1: Vertices and errors for $\delta = 1$

A typical behaviour of the error estimate is that its convergence to zero is not monotone. This happens because the smaller a simplex S becomes the steeper the function $v_h^{\Xi}|_S$ may be. Although Corollary 2.10 guarantees that eventually the error will become smaller if the size of the simplices converges to zero it may happen that one refinement step causes an increasing of the error estimates corresponding to simplices in a neighbourhood of the refined simplices.

The effect of the coarsening of the simplices strongly depends on the choice of the parameter tol . If tol is too small no simplex will be coarsened; if tol is too large, simplices will be coarsened in one iteration step and refined again in the next step which reduces the speed of convergence. Several numerical examples showed that good results can be expected if tol is slightly smaller than the desired accuracy, e.g. in the example above the choice of $tol = 0.003$ is suitable for a desired accuracy 0.004.

The values of $v_h^{\Xi^k}$ lie between -1.37 and 0.1 . Using the solution corresponding to the grid Ξ_9 trajectories with $J_h(x, u_x(\cdot)) < -1.4$ can be calculated. The minimal value of the value function $v_h^{\Xi_0}$ is -1.02 and therefore the error of this solution can be estimated as $|v_h^{\Xi_0} - v_h| > 0.38$. Using the error estimates \tilde{e}_i one obtains $|v_h^{\Xi_0} - v_h| < 0.027/0.05 = 0.54$. Hence here the error estimate from Theorem 2.2 gives a good approximation for the global error of the solution.

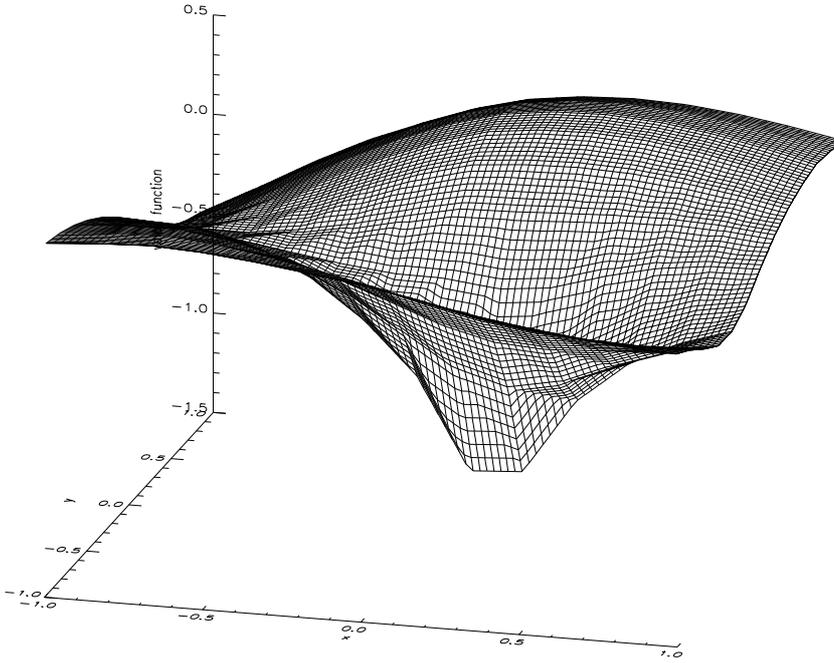


Figure 4.2: Value function for $\delta = 1$, initial grid

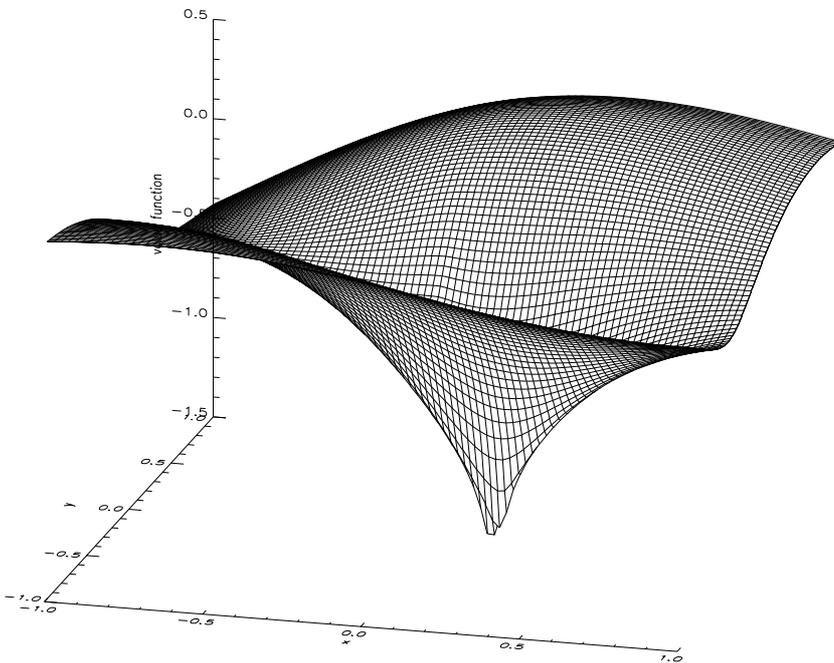


Figure 4.3: Value function for $\delta = 1$, adapted grid

The drop of the refined value function corresponds to the control set D_1 . Figure 4.4 shows the refined grid around D_1 .

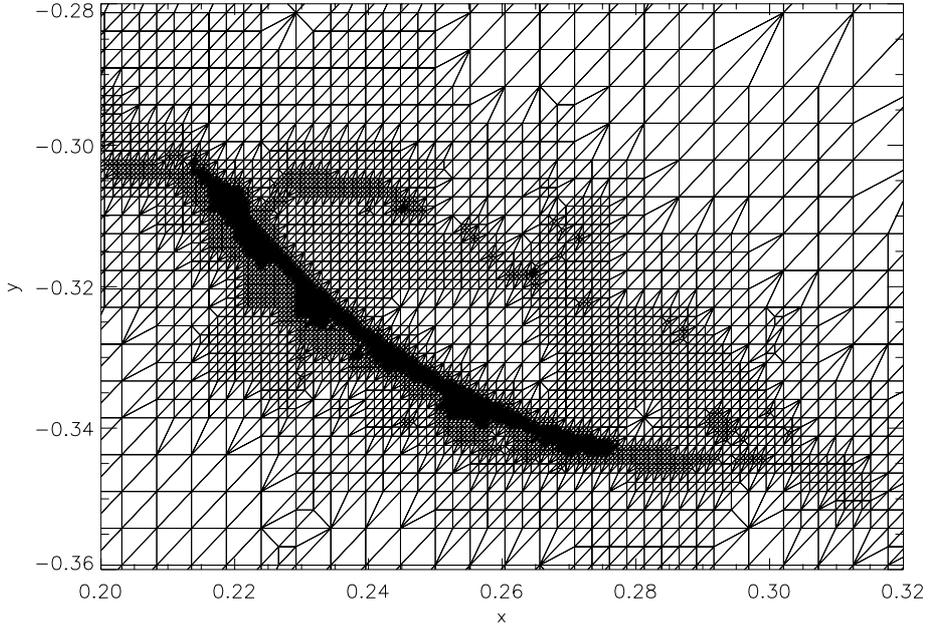


Figure 4.4: Adapted grid for $\delta = 1$ around D_1

The second example is the same optimal control system with lower discount rate $\delta = 0.01$. Table 4.2 shows the number of vertices N and the maximal error estimate err for each iteration step using the algorithm with coarsening and $tol = 0.1$. In addition we pick the initial value $x_0 = (0.255, -0.335)$ inside (and near the boundary of) D_1 and the corresponding trajectory $(x_i)_{i \in \mathbb{N}}$ with the control function from (2.4). The table shows the error $err_{tra} := \sum_{i=0}^{\infty} \beta^i e(x_i)$ along (x_i) as discussed in Theorem 2.5 and the corresponding functional $J_h(x_0, u_{x_0})$. (Note that the values err and err_{tra} are not directly comparable, cp. Theorem 2.2 and Theorem 2.5.)

$\#iter$	N	err	err_{tra}	$J_h(x_0, u_{x_0})$
0	169	0.23	8.638	-0.908
1	231	0.47	12.209	-1.133
2	228	0.48	9.219	-1.346
3	281	1.60	11.828	-1.254
4	306	2.03	14.616	-3.302
5	388	1.89	17.531	-4.337
6	735	7.18	96.307	-128.257
7	1488	15.58	33.255	-131.910
8	2705	21.45	17.164	-137.768
9	4558	19.11	4.681	-138.185
10	8323	31.03	2.098	-138.277
11	12820	30.55	1.952	-139.255
12	20900	30.08	2.376	-139.485

Table 4.2: Vertices and errors for $\delta = 0.01$

In this example the values of $v_h^{\Xi^k}$ lie between -137.5 and 10 , hence the relation between the maximum of the error estimates err and the values of the solutions is worse than in the first example. In addition the supremum of the error estimates increases considerably until the tenth iteration and then decreases very slowly. This slow convergence is exactly what one would expect for low discount rates since the exponent $\gamma > 0$ from Corollary 2.10 is close to 0 , cp. Remark 2.11.

Nevertheless the error err_{tra} along the trajectory decreases and the calculated functional along the trajectory decreases monotonically. Here even in the case when the convergence of e_i to zero is very slow the adaptive grid yields good results for the calculation of optimal controls.

The jump in the functional from iteration step five to step six corresponds to a change in the behaviour of the trajectory x_i : From this step on the trajectory x_i stays inside D_1 ; for all grids before the trajectory leaves D_1 after a short time. After this iteration the value for $J_h(x_0, u_{x_0}(\cdot))$ is smaller than the minimum of the corresponding solution $v_h^{\Xi^k}$, which is about -137.5 for Ξ_{12} (cp. Remark 2.6).

Again for Ξ_0 the estimate from Theorem 2.2 gives a good approximation of the global error of the solution.

The following figures show the value function corresponding to the adapted grid and the corresponding grid around D_1 .

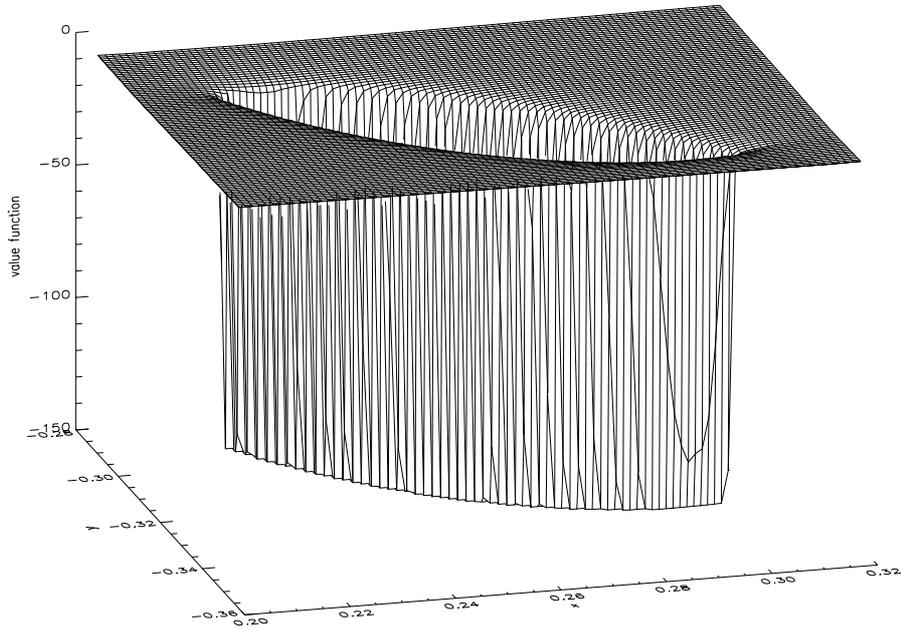


Figure 4.5: Value function for $\delta = 0.01$ around D_1 with adapted grid

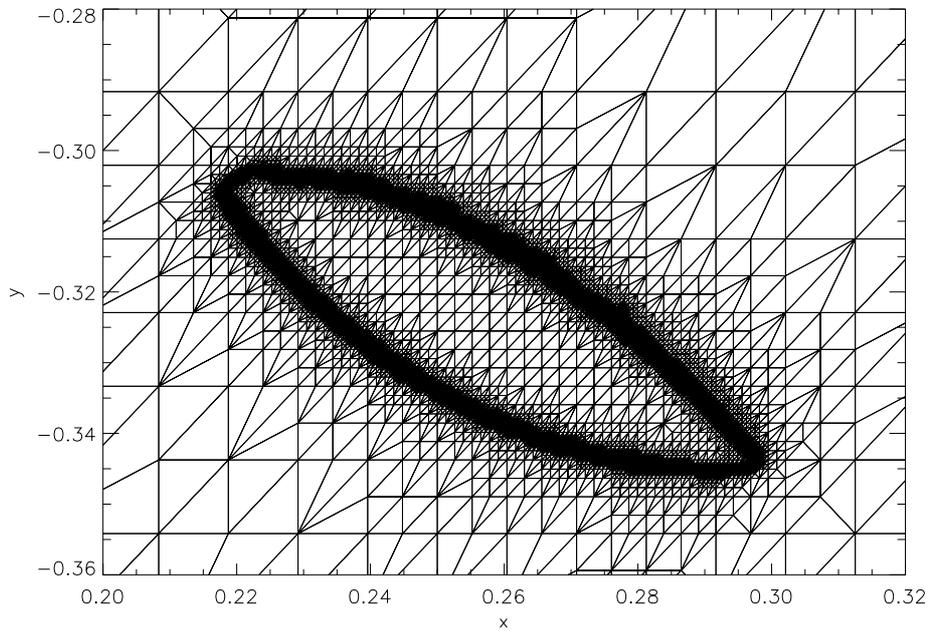


Figure 4.6: Adapted grid for $\delta = 0.01$ around D_1

The only region that is refined in this example is the neighbourhood of ∂D_1 . The steepness of the value function in this region is the main reason for the large local errors.

Conclusions: The adaptive grid strategy based on the local values of the residual e_i turns out to be a good tool to calculate grids for the solution of (2.1). Without any knowledge of the "critical regions" i.e. without further analysis of the optimal control system a good discretization for the problem can be obtained.

In the case of small discount rates δ the strategy helps to develop grids such that good control functions and trajectories can be calculated and hence gives at least a good upper bound for v_h . For large discount rates one can expect that the error estimates e_i converge to zero during the adapting iteration. In this case a good approximation for v_h can be calculated and furthermore the quality of this approximation can be controlled using the error estimates e_i .

Acknowledgement: I would like to thank Fritz Colonius for his constant advice and many useful discussions.

References

- [1] I. BABUŠKA AND A. MILLER, *A feedback finite element method with a posteriori error estimation, Part I*, Computer Methods in applied Mechanics and Engineering, 61 (1987), pp. 1–40.
- [2] I. CAPUZZO DOLCETTA, *On a discrete approximation of the Hamilton-Jacobi equation of dynamic programming*, Applied Mathematics and Optimization, 10 (1983), pp. 367–377.
- [3] I. CAPUZZO DOLCETTA AND H. ISHII, *Approximate solutions of the Bellman equation of deterministic control theory*, Applied Mathematics and Optimization, 11 (1984), pp. 161–181.
- [4] F. COLONIUS, *Asymptotic behaviour of optimal control systems with low discount rates*, Mathematics of Operations Research, 14 (1989), pp. 309–316.
- [5] F. COLONIUS AND W. KLIEMANN, *Asymptotic null controllability of bilinear systems*, in Proceedings of the Workshop "Geometry and Nonlinear Control Theory", Warsaw, June 1993. To appear.
- [6] ———, *Maximal and minimal Lyapunov exponents of bilinear control systems*, Journal of Differential Equations, 101 (1993), pp. 232–275.
- [7] ———, *The Lyapunov spectrum of families of time varying matrices*. Schwerpunktprogramm der Deutschen Forschungsgemeinschaft "Anwendungsbezogene Optimierung und Steuerung", Report No. 504, Augsburg, 1994. To appear in: Transactions of the AMS.
- [8] B. FAERMANN, *Lokale a-posteriori Fehlerschätzer bei der Diskretisierung von Randintegralgleichungen*. Dissertation, Universität Kiel, 1993.
- [9] M. FALCONE, *Numerical solution of deterministic control problems*, in Proceedings of the International Symposium on Numerical Analysis, Madrid, 1985.
- [10] ———, *A numerical approach to the infinite horizon problem of deterministic control theory*, Applied Mathematics and Optimization, 15 (1987), pp. 1–13. *Corrigenda*, ibid. 23 (1991), 213–214.
- [11] R. L. V. GONZÁLES AND C. A. SAGASTIZÁBAL, *Un algorithme pour la résolution rapide d'équations discrètes de Hamilton-Jacobi-Bellman*, Comptes Rendus Acad. Sc. Paris, Serie I, 311 (1990), pp. 45–50.
- [12] R. L. V. GONZÁLES AND M. M. TIDBALL, *On the rates of convergence of fully discrete solutions of Hamilton-Jacobi equations*. INRIA Rapports de Recherche Nr. 1379, 1991.
- [13] L. GRÜNE, *Numerical stabilization of bilinear control systems*, to appear in: SIAM Journal on Control and Optimization, (1996).

- [14] G. HÄCKL, *Numerical approximation of reachable sets and control sets*, Random & Computational Dynamics, 1 (1992–1993), pp. 371–394.
- [15] H. J. KUSHNER AND P. G. DUPUIS, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York, 1992.
- [16] R. LÖHNER, *An adaptive finite element scheme for transient problems in CFD*, Computer Methods in applied Mechanics and Engineering, 61 (1987), pp. 323–338.
- [17] E. RANK, *Adaptivity and accuracy estimation for FEM and BEM*, in Accuracy Estimates and Adaptive Refinements in FE Computations, I. Babuška, O. C. Zienkiewicz, J. Gago, and E. R. Oliveira, eds., Wiley, New York, 1986.
- [18] W. L. WENDLAND AND D. YU, *Adaptive BEM for strongly elliptic integral equations*, Numerische Mathematik, 53 (1988), pp. 539–558.
- [19] F. WIRTH, *Convergence of the value functions of discounted infinite horizon optimal control problems with low discount rates*, Mathematics of Operations Research, 18 (1993), pp. 1006–1019.