Theses and Dissertations

May 2017

# Numerical Methods for Hamilton-Jacobi-Bellman Equations

Constantin Greif
*University of Wisconsin-Milwaukee*

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Mathematics Commons

# Numerical Methods for Hamilton-Jacobi-Bellman Equations

by

Constantin Greif

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
in
Mathematics

at

The University of Wisconsin - Milwaukee
May 2017

# Abstract

# Numerical Methods for Hamilton-Jacobi-Bellman equation
by
Constantin Greif

The University of Wisconsin - Milwaukee, 2017
Under the Supervision of Professor Bruce A. Wade

In this work we considered HJB equations, that arise from stochastic optimal control problems with a finite time interval. If the diffusion is allowed to become degenerate, the solution cannot be understood in the classical sense. Therefore one needs the notion of viscosity solutions. With some stability and consistency assumptions, monotone methods provide the convergence to the viscosity solution. In this thesis we looked at monotone finite difference methods, semi lagragian methods and finite element methods for isotropic diffusion. In the last chapter we introduce the vanishing moment method, a method not based on monotonicity.

# TABLE OF CONTENTS

# 1. Hamilton-Jacobi-Bellman Equations

In this thesis, we are searching for the numerical solution of a class of second-order fully nonlinear partial differential equations (PDE), namely the Hamilton-Jacobi-Bellman (HJB) equations. These PDE are named after Sir William Rowan Hamilton, Carl Gustav Jacobi and Richard Bellman. The equation is a result of the theory of dynamic programming which was pioneered by Bellman. In continuous time, the result can be seen as an extension of earlier work in classical physics on the Hamilton-Jacobi equation. The HJB equations we consider arise from optimal control models for stochastic processes.

## 1.1. Outline

In this Chapter we briefly describe how HJB equations arise from stochastic optimal control problems. Then in Chapter 2 we will introduce the concept of viscosity solutions and we will look at the Barles-Souganidis Argument, which guarantees us the convergence to the viscosity solution for monotone schemes. In Chapter 3 we will explain Howard's Algorithm, which is included in many methods solving the HJB. In Chapter 4 we will look at monotone finite difference methods. In Chapter 5, we will look at Semi-Lagrangian Schemes, which also get the convergence through the monotonicity argument. In Chapter 6 we will look at monotone finite element methods for isotropic problems. In Chapter 7 we will look at a very different concept, which is this of the vanishing moment method.

## 1.2. HJB in Optimal Control Problems

Optimal control problems describe the time evolution of a state vector $X : t \to \Omega \subset \mathbb{R}^d$ related to a control process $\lambda : t \to \Lambda$, where $\Lambda$ is the set of admissible control values. If not other stated, $\Lambda$ will be a compact metric space. If we say $\lambda(\cdot) \in \Lambda$, then we mean a measurable function with $\lambda_t \in \Lambda$ almost everywhere. The state vector $X$ satisfies a given stochastic differential equation (SDE), whose drift $\mu \in \mathbb{R}^d$ and diffusion matrix $\sigma \in \mathbb{R}^{d \times p}$ are functions which depend on the control $\lambda(\cdot) \in \Lambda$:

$$dX_t = \mu^{\lambda_t}(t, X_t)dt + \sigma^{\lambda_t}(t, X_t)dW_t \text{ for } t > 0,$$
$$X_0 = x, \tag{1.1}$$

where $W_t$ is a given $p$-dimensional Wiener process.

The task is to minimize a given cost functional $J^\lambda(t, x) \in \mathbb{R}$ with $(\lambda, t, x) \in \Lambda \times [0, T) \times \Omega$, dependent on functions $f^{\lambda_t}(t, x) \in \mathbb{R}$ and $g(t, x) \in \mathbb{R}$.

Then we can define the value function as

$$u(t, x) := \inf_{\lambda \in \Lambda} J^\lambda(t, x) \text{ ,where } J^\lambda(t, x) := \mathbb{E}\Big[ \int_t^\tau f^{\lambda_s}(s, X_s)ds + g(\tau, X_\tau) \Big], \tag{1.2}$$

where $\tau = \inf\{s \geq t \mid (s, X_s) \notin (0, T) \times \Omega\}$ is the final time of our problem when the state $X_t$ leaves the open domain set $\Omega_T = (0, T) \times \Omega$. The control problem (1.2) with the value function $u$ leads to the following HJB equation:

$$\partial_t u + \inf_{\lambda \in \Lambda}[L^\lambda u + f^\lambda] = 0 \text{ in } \Omega_T,$$
$$u = g \text{ on } \partial^* \Omega_T = \{T\} \times \Omega \cup (0, T) \times \partial\Omega, \tag{1.3}$$

where the linear operator $L^\lambda$ is defined by:

$$L^\lambda v := \text{Tr}[a^\lambda D^2 v] + \mu^\lambda \cdot \nabla v, \ v \in H^2(\Omega), \lambda \in \Lambda,$$

with $a^\lambda := \frac{1}{2}\sigma(\sigma)^\top \in \mathbb{R}^d \times \mathbb{R}^d$, the symmetric positive semidefinite diffusion coefficient matrix and

$D^2v$ denotes the Hessian matrix after $x$. Equation (1.3) is now solved by the value function we just defined in (1.2).

**Example 1.2.1.(Control problem with explicit solution)** If the drift is given by $\mu^\lambda(t, X_t) = -c_1 X_t + c_2 \lambda_t$, with $c_1$ and $c_2$ constants, the diffusion is also just a constant $\sigma^\lambda(t, X_t) = -\sigma$ and the cost function is given by $f^\lambda(t, X_t) = \frac{r(t)\lambda^2}{2} + \frac{l(t)X_t^2}{2}$, where $r(t) > 0$ and $l(t) > 0$ are functions just dependend on time. Then the HJB equation is given by

$$\partial_t u + \inf_{\lambda \in \Lambda}\left[-(c_1 x + c_2\lambda)\partial_x u - \sigma\partial_{xx} u + \frac{r(t)\lambda^2}{2} + \frac{l(t)x^2}{2}\right] = 0$$

Through a basic calculation by derivation after $\lambda$, we see that the unique exact solution at time $t$ is given by

$$\lambda_t = \frac{c_2 \partial_x u}{r(t)},$$

which leads to the following HJB equation without an $\inf(\cdot)$ operator

$$\partial_t u = \left(c_1 x + \frac{c_2^2 \partial_x u}{2r(t)}\right)\partial_x u + \sigma\partial_{xx} u - \frac{l(t)x^2}{2}$$

This is just an example, in this Thesis we have the focus on cases, where it can't be solved analytically.

We will cite the Theorem in [11].

**Theorem 1.2.2.(Krylov)** If the following hold:

- The control set $\Lambda$ is compact,

- $\Omega$ is bounded

- $\partial\Omega$ is of class $C^3$ (roughly speaking, the boundary is locally the graph of a $C^3$ function),

- The functions $a^\lambda, \mu^\lambda, f^\lambda$ are in $C(\bar{\Omega}_T \times \Lambda)$ with their $t-$partial derivative and first and second $x-$partial derivatives for all $\lambda \in \Lambda$,

- $h \in C^3([0, T] \times \mathbb{R}^d)$,

and furthermore, there exists $\gamma > 0$ such that for every $(t, x) \in \bar{\Omega}_T$ and $\lambda \in \Lambda$, $L^\lambda$ is uniformly elliptic, i.e. for $a^\lambda(t, x)$ holds

$$\sum_{i,j=1}^{d} a_{ij}^\lambda(t, x)\xi_i\xi_j \geq \gamma|\xi|^2 \text{ for all } \xi \in \mathbb{R}^d.$$

Then the HJB equation has a unique classical solution $u \in C(\bar{\Omega}_T)$ with continuous $t-$partial derivative and continuous first and second $x-$partial derivatives.

**Remark 1.2.3.** If we allow the HJB equation to become degenerate, a unique classical solution is not guaranteed anymore. This is why we need the concept of viscosity solutions. Under suitable assumptions, which does not include uniformly ellipticy, the HJB equation (1.3) has a unique, bounded, Hölder continuous, viscosity solution $u$.

## 1.2.1. Dynamic Programming Principle

The HJB equation is a result of the dynamic programming principle of Bellman, which allows us to split the value function.

**Theorem 1.2.4.** For every $t_1 \in [t, \tau]$ and $y \in \Omega$ with $X_t = y$, we have

$$u(t, y) = \inf_{\lambda \in \Lambda} \left\{ \mathbb{E}\Big[ \int_t^{t_1} f^{\lambda_s}(s, X_s)ds \Big] + u(t_1, X_{t_1}) \right\} \tag{1.4}$$

By knowing this, we can apply the Dynamic Programming Principle. Means, instead of looking for $u(0, x_0)$, we can go backwards. Therefore we start with $u(t, x) = g(t, x)$ for $(t, x) \in \partial^* \Omega_T$ and then inductively, by knowing $u(t_{k+1}, \cdot)$, we get $u(t_k, x) = \inf_{\lambda \in \Lambda} \left\{ \mathbb{E}\Big[ \int_{t_k}^{t_{k+1}} f^{\lambda_s}(s, X_s)ds \Big] \right\} + u(t_{k+1}, X_{t_{k+1}})$, where most numerical methods then approximate $\mathbb{E}\Big[ \int_{t_k}^{t_{k+1}} f^{\lambda_s}(s, X_s)ds \Big] \approx \Delta t \, L_k^\lambda u + \Delta t \, f^\lambda$.

**Theorem 1.2.5.** Assume that the value function is in $u \in C(\bar{\Omega}_T)$ and $u = g$ on $\partial\Omega_T$. Then with the dynamic programming principle for $u$, we get $u$ to be a solution of the belonging HJB equation.

## 1.3. Fully Nonlinear Second Order PDEs

The second order PDE (1.3) is fully nonlinear, because the dependence on $D^2u$ is not linear, since the Hessian is included in an $\inf(\cdot)$, or often $\sup(\cdot)$, operator in the PDE. The HJB equation would be linear, if the control set $\Lambda$ was a singleton. To solve fully nonlinear PDEs, there are classical and weak solution concepts and theories. It is well known that for a class of fully nonlinear second order PDEs, a $C^{2,\alpha}$ a priori estimate is provided by the celebrated Evans-Krylov Theorem [4].

The nonlinearity of the highest order derivative in (1.3) makes it impossible to use a weak solution concept based on the integration by parts approach, like we would do for linear, quasi-linear or semilinear PDEs. So in fact, there was no weak solution concept for fully nonlinear PDEs until Crandall and Lions [7] introduced the notion of viscosity solutions for first order fully nonlinear PDEs. Then their notion and theory were quickly extended to second order fully nonlinear PDEs, like we use in this work.

### 1.3.1. Challenges

For solving PDEs numerically there are three main classes, or none of the below.

- Methods based on directly approximating derivatives by difference quotients.

- Methods based on variational principles and approximating infinte-dimensional spaces by finite-dimensional spaces.

- Methods based on finite basis expansions and approximating PDEs at sampling points.

Unfortunately none of those methods work right away for fully nonlinear second order PDEs. A naive application of the first and third class can already lead to very bad results. And like mentioned before, the second class can not even be formulated due to the nonlinearity.

## 1.4. Monotone Methods

In [1] Barles and Souganidis provided a general convergence theory for a broad class of possibly degenerate fully nonlinear elliptic and parabolic PDEs. In specific, that the underlying PDE satisfies a certain maximums principle and that the FDM is monotone, consistent and stable in the sense that the sequence of approximations $\{u_h\}_h$ remains bounded in the maximums norm, then it can be shown that $||u - u_h||_{L^\infty} \to \infty$ for stepsize $h \to 0$.

Some of the first computational methods for HJB equations in stochastic control are based on approximating the underlying SDE by a discrete Markov chain. Later it became clear that under some assumptions these are equivalent to monotone finite difference methods [3]. The computational practice of monotone methods has lagged behind their theoretical development, especially for strongly anisotropic problems. Two main problems are the lower order convergence rate in the first place and the necessary choice of wide stencil width. In fact, to achieve monotonicity for strongly anisotropic problems, compact stencils cannot offer consistence and monotonicity. But increasing the stencil width increases the truncation error. In the degenerate case, there are examples, where no finite stencil can yield a monotone discretization, so it needs to increase when the grid is refined. Bonnans and Zidani showed in [3] the number of conditions needed for the diffusion coefficient. Bonnans gave in [2] an fast algorithm for computing monotone schemes in 2-dim with finite stencil and a consistency error depending on the stencil width.

Debraband and Jakobsen gave in [8] a semi-lagrangian framework in which the stencil width also continuously increases as the mesh is refined. One advantage of these methods is, that the mononticity is guaranteed for $h \to 0$.

The argument of Barles and Souganidis can not be applied for finite element schemes right away, since it is made for finite difference schemes. But Ian Smears and Jensen were still able [12] to create a monotone finite element scheme for possible degenerate isotropic HJB equations. It was shown that this method converges to the viscosity solution in the $L^\infty$ norm.

## 1.5. Nonmonotone Methods

To guarantee monotonicity and consistency one needs wide stencils, which causes high truncation errors and therefore reduces the accuracy. In fact, monotone schemes are in practice behind their theoretical development. Therefore many authors proposed different nonmonotone methods in order to avoid the stencil restrictions. One of them is the vanishing moment method, which involves fourth order perturbations to the PDE. Non of the non-monotone methods currently offers a satisfactory convergence analysis. Nevertheless, some methods have offered good computational results.

## 1.6. Different Types of HJB equations

HJB equations can have different forms, like

$$\partial_t v + H = 0 \text{ or}$$

$$\partial_t v - H = 0 \text{ or}$$

$$H = 0,$$

where $H$ is the Hamiltonian, in our case (1.3) $H = \inf_{\lambda \in \Lambda} [L^\lambda v + f^\lambda]$. To show that different types of the HJB equation arise from the same control problem, we will use the following lemma.

**Lemma 1.6.1.** For a compact set $A$ and continuous function $F : A \to \mathbb{R}$, we have

$$\sup_{a \in A}[F(a)] = - \inf_{a \in A}[-F(a)]$$

*Proof.* $\inf_{a \in A}[-F(a)] \leq -F(a)$ for all $a \in A \implies - \inf_{a \in A}[-F(a)] \geq F(a)$ for all $a \in A \implies - \inf_{a \in A}[-F(a)]$ is a upper boundary, i.e. $- \inf_{a \in A}[-F(a)] \geq \sup_{a \in A}[F(a)]$.
And similarly, $\sup_{a \in A}[F(a)] \geq F(a)$ for all $a \in A \implies - \sup_{a \in A}[F(a)] \leq -F(a)$ for all $a \in A \implies$ it is a lower boundary, i.e. $- \sup_{a \in A}[F(a)] \leq \inf_{a \in A}[-F(a)] \implies \sup_{a \in A}[F(a)] \geq - \inf_{a \in A}[-F(a)]$.

$\square$

## 1.6.1. Time-dependent Case

If we consider the minimising stochastic control problem like before, where we defined the value function as $u(t, x) := \inf_{\lambda \in \Lambda} J^\lambda(t, x)$, and with $\Omega = \mathbb{R}^d$, then we get the HJB equation with terminal condition:

$$\partial_t u + \inf_{\lambda \in \Lambda} [L^\lambda u + f^\lambda] = 0 \text{ in } \Omega_T,$$

$$u(T, x) = g(T, x) \text{ with } x \in \Omega,$$

If we change the value function to $v(t, x) := -u(t, x)$, then we get the following HJB

$$\partial_t v + \sup_{\lambda \in \Lambda} [L^\lambda v - f^\lambda] = 0 \text{ in } \Omega_T,$$

$$u(T, x) = -g(T, x) \text{ with } x \in \Omega,$$

If we substitute the time by defining $\tilde{u}(t, x) := u(T - t, x)$, then we get the following HJB with initial condition

$$\partial_t \tilde{u} - \inf_{\lambda \in \Lambda} [L^\lambda \tilde{u} + f^\lambda] = 0 \text{ in } \Omega_T,$$

$$\tilde{u}(0, x) = g(T, x) \text{ with } x \in \Omega,$$

which is equivalent to

$$\partial_t \tilde{u} + \sup_{\lambda \in \Lambda} [-L^\lambda \tilde{u} - f^\lambda] = 0 \text{ in } \Omega_T,$$

$$\tilde{u}(0, x) = gT, x) \text{ with } x \in \Omega,$$

If we do both at the same time $w(t, x) := -u(\tau - t, x)$, then we get

$$-\partial_t w + \sup_{\lambda \in \Lambda} [L^\lambda w - f^\lambda] = 0 \text{ in } \Omega_T,$$

$$w(0, x) = -g(T, x) \text{ with } x \in \Omega,$$

which is equivalent to

$$\partial_t w - \sup_{\lambda \in \Lambda}[L^\lambda w - f^\lambda] = 0 \text{ in } \Omega_T,$$

$$w(0, x) = -g(T, x) \text{ with } x \in \Omega.$$

## 1.6.2. Time-independent Case

In this case, the HJB equation is elliptic and we have

$$\inf_{\lambda \in \Lambda}[L^\lambda v + f^\lambda] = 0 \text{ in } \Omega. \tag{1.5}$$

Or if we define $v(x) := -u(x)$, we get

$$\sup_{\lambda \in \Lambda}[L^\lambda u - f^\lambda] = 0 \text{ in } \Omega. \tag{1.6}$$

## 1.6.3. Infinite Time-horizon Case

Case 3 is, when we have a infinite time-horizon, then if we consider the value function:

$$u(x) = \inf_{\lambda \in \Lambda} \mathbb{E}\Big[ \int_0^\infty f^{\lambda_s}(s, X_s) e^{-\gamma s} ds \Big]$$

$$\text{subject to } dX_t = \mu^{\lambda_t}(t, X_t)dt + \sigma^{\lambda_t}(t, X_t)dW_t \text{ for } t > 0, \tag{1.7}$$

$$X_0 = x,$$

we get the HJB equation

$$\gamma u - \inf_{\lambda \in \Lambda}[L^\lambda u + f^\lambda] = 0 \text{ in } \Omega,$$

$$u = 0 \text{ on } \partial\Omega, \tag{1.8}$$

**Remark 1.6.2.** If we have $\gamma = 0$, it is the elliptic case (1.5) of the HJB equation.

### 1.6.4. Maximising Problem

If we are dealing with a maximising problem

$$u(t,x) = \sup_{\lambda \in \Lambda} \mathbb{E}\Big[\int_t^\tau f^{\lambda_s}(s, X_s)ds + h(\tau, X_\tau)\Big], \tag{1.9}$$

we get the HJB

$$\partial_t u + \sup_{\lambda \in \Lambda}[L^\lambda u + f^\lambda] = 0 \text{ in } \Omega_T,$$

$$u = g \text{ on } \partial^*\Omega_T, \tag{1.10}$$

**Remark 1.6.3.** It is possible to rewrite methods for time-dependent HJBs to time-independent or infinite time interval HJBs, since the approximation of the Hamiltonian is the challenging task.

## 1.7. Examples

**Example 1.7.1.** Let's consider the stochastic control problem with value function $u(t,x) = \min_\lambda \mathbb{E}[\int_t^\tau 1ds]$, which means we want to leave the domain $\Omega_T$ as soon as possible to minimize this integral.

Then with $\Lambda = \{-1, 1\}, f^\lambda = 1, \sigma^\lambda = 0, \mu^\lambda = \lambda$, the HJB equation to this is given by

$$\partial_t u(t,x) - \sup_{\lambda \in \Lambda}\{-\lambda \nabla u(t,x) - 1\} = 0 \text{ for } t \in (0,1) \text{ and } x \in (-1, 1)$$

$$u = 0 \text{ on } \{1\} \times (-1, 1) \cup (0, 1) \times \{-1, 1\}.$$

With unique viscosity solution $u(t,x) = \min(1 - t, 1 - |x|)$. See Figure (2.2).

**Example 1.7.2.** If we consider the HJB equation with

$$f^\lambda(t,x) = \sin(x_1)\sin(x_2)[(1 + 2\beta^2)(2 - t) - 1] - 2(2 - t)\cos(x_1)\cos(x_2)\sin(x_1 + x_2)\cos(x_1 + x_2),$$

$$c^\lambda(t,x) = \mu^\lambda(t,x) = 0,$$

$$\sigma^\lambda(t,x) = \sqrt{2}\begin{pmatrix} \sin(x_1 + x_2) & \beta & 0 \\ \sin(x_1 + x_2) & 0 & \beta \end{pmatrix} \text{ with } \beta^2 = 0.1. \text{ In this Case the HJB eqaution is linear and the}$$
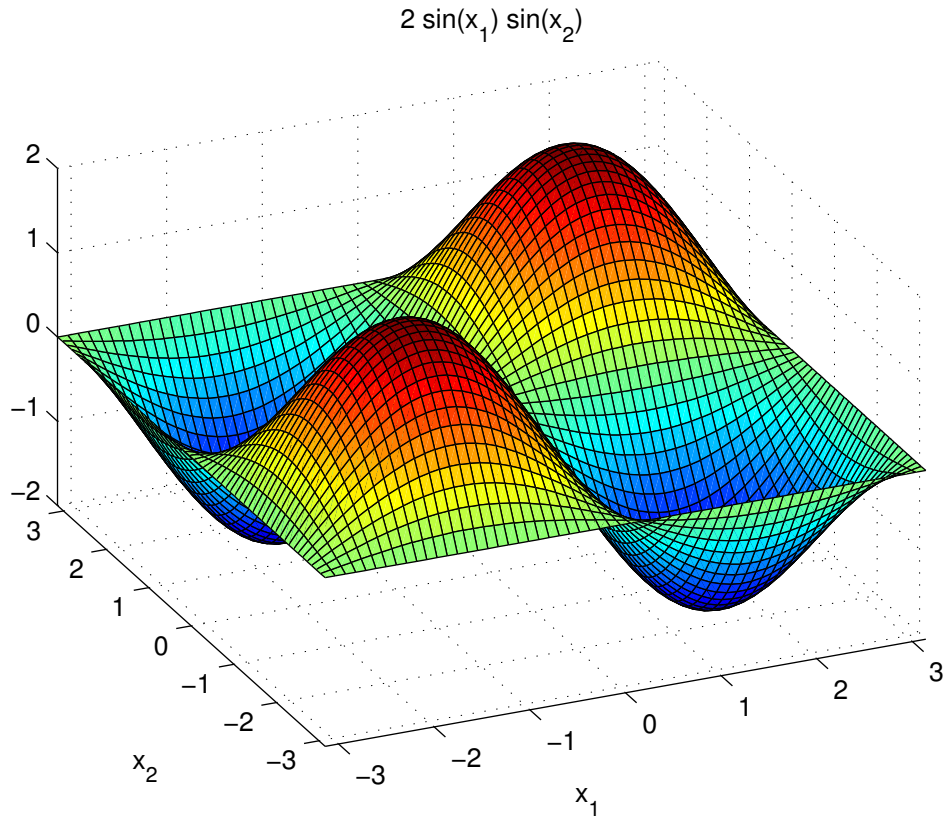
Figure 1.1.: $u(t, x_1, x_2) = (2 - t) \sin(x_1) \sin(x_2)$, plot for $t = 0$

solution of this is

$$u(t, x) = (2 - t) \sin(x_1) \sin(x_2). \tag{1.11}$$

See Figure (1.1)

# 2. Viscosity Solution and the Barles-Souganidis Convergence Argument

**Definition 2.0.1.(degenerate elliptic)** An operator $F : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times S_n(\mathbb{R}) \to \mathbb{R}$ is called degenerate elliptic on $\Omega$ if for all $x \in \Omega$, $r \in \mathbb{R}$ $P, Q \in S_n(\mathbb{R})$ with $P \geq Q$ and $y \in \mathbb{R}^d$ we have

$$F(x, r, y, P) \leq F(x, r, y, Q).$$

We call an operator $-\partial_t + F$ degenerate parabolic, if $F(\cdot, t, \cdot, \cdot, \cdot)$ is degenerate elliptic for all $t \in (0, T)$.

**Definition 2.0.2.(proper)** An operator $F : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times S_n(\mathbb{R}) \to \mathbb{R}$ is called proper on $\Omega$ if for all $x \in \Omega$, $r, l \in \mathbb{R}$ with $r \geq l$, $P \in S_n(\mathbb{R})$ and $y \in \mathbb{R}^d$ we have

$$F(x, r, y, P) \geq F(x, l, y, P).$$

We call an operator $-\partial_t + F$ proper, if $F(\cdot, t, \cdot, \cdot, \cdot)$ is proper for all $t \in (0, T)$.

Recall the HJB operator

$$-\partial_t u + \sup_{\lambda \in \Lambda} [-L^\lambda u - f^\lambda]. \tag{2.1}$$

It can be proven, that this operator is in fact degenerate parabolic and proper.

## 2.1. Viscosity Solution

Without further requirements, general fully nonlinear PDEs second order, like our HJB equation, do not necessarily have a classical solution. Because of the nonlinearity on the highest order derivative

Figure 2.1.: viscosity solution, i.e. $\varphi_t(x_0) + F[\varphi(x_0)] \leq 0$, right: vis. supersolution, i.e. $\varphi_t(x_0) + F[\varphi(x_0)] \geq 0$

in (2.2), we can also not extend a weak solution concept based on the integration by parts approach for fully nonlinear PDEs. So in general there is no variational/weak formulation for fully nonlinear PDEs. In 1983 Crandall and Lions [7] introduced the notion of viscosity solutions and established their theory for the Hamilton-Jacobi equations of first order. For the definition of first order, we refer to Chapter 7. The notion and theory of viscosity solutions to fully nonlinear second order got extended by Jensen, who established the uniqueness of solutions and by Ishii, who proved the existence of solutions. Viscosity solutions are a mathematical concept to select the value function $u$ from the possibly infinite set of weak solutions for the HJB equation.

### 2.1.1. Motivation

To motivate the notion of viscosity solutions, suppose for a moment that $F$ is degenerate elliptic and $u$ is a $C^2$-function satisfying $F(D^2u(x), \nabla u(x), u(x), x) \leq 0$ (resp. $F[u(x)] \geq 0$). Suppose further that $\varphi$ is also a $C^2$-function satisfying $u \leq \varphi$ (resp. $u \geq \varphi$) and $u - \varphi$ has a local maximum (resp. minimum) at $x_0 \in \Omega \subset \mathbb{R}^d$, without loss of generalization the maximum is allocated at zero, so $u(x_0) = \varphi(x_0)$. Then elementary calculus tells us that $\nabla u(x_0) = \nabla \varphi(x_0)$ and $D^2u(x_0) \leq D^2\varphi(x_0)$ (resp. $D^2u(x_0) \geq D^2\varphi(x_0)$). So we get $F(D^2\varphi(x_0), \nabla \varphi(x_0), \varphi(x_0), x_0) \leq F(D^2u(x_0), \nabla u(x_0), u(x_0), x_0) \leq 0$ (resp. $F[\varphi(x_0)] \geq F[u(x_0)] \geq 0$). Consider Figure (2.1)

## 2.1.2. Definition

With the observation above, we give the definition of a viscosity solution for

$$-\partial_t u + F(x, t, u, \nabla u, D^2 u) = 0 \text{ on } \Omega_T \tag{2.2}$$

**Definition 2.1.1.(continuous case)** Consider $F : \mathbb{R}^{d \times d} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a continuous non-linear function.

(i) A function $u \in C^0(\Omega)$ is called a viscosity subsolution of (2.2) if, for every $C^2$ function $\varphi$ such that $u - \varphi$ has a local maximum at $x_0 \in \Omega$, there holds $-\varphi_t + F(D^2\varphi(x_0), \nabla\varphi(x_0), \varphi(x_0), x_0) \leq 0$.

(ii) A function $u \in C^0(\Omega)$ is called a viscosity supersolution of (2.2) if, for every $C^2$ function $\varphi(x)$ such that $u - \varphi$ has a local minimum at $x_0 \in \Omega$, there holds $\partial_t\varphi + F(D^2\phi(x_0), \nabla\phi(x_0), \phi(x_0), x_0) \geq 0$.

A function $u \in C^0(\Omega)$ is called a viscosity solution of (2.2) if it is both, a viscosity subsolution and a viscosity supersolution.

This definition can be generalized to the case when both $F$ and $u$ are just bounded functions, which can be easily done using the lower and upper semi continuous envelopes of $F$ and $u$.

**Definition 2.1.2.(semi continuous envelope)** For $u \in B(\bar{\Omega})$ we define the lower semi-continuous envelope as

$$u_*(x) = \liminf_{y \to x} u(x) \text{ , for } x \in \Omega$$

and similarly the upper semi-continuous envelope as

$$u^*(x) = \limsup_{y \to x} u(x) \text{ , for } x \in \Omega$$

**Definition 2.1.3.** Consider $F : \mathbb{R}^{d \times d} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $u : \Omega \to \mathbb{R}$ are bounded functions.

(i) $u$ is called a viscosity subsolution of (2.2), if for every $C^2$ function $\varphi$ such that $u^* - \varphi$ has a local maximum at $x_0 \in \Omega$, there holds $-\partial_t\varphi + F(D^2\varphi(x_0), \nabla\varphi(x_0), \varphi(x_0), x_0) \leq 0$.

(ii) $u$ is called a viscosity supersolution of (2.2), if for every $C^2$ function $\varphi(x)$ such that $u_* - \varphi$ has a local minimum at $x_0 \in \Omega$, there holds $-\partial_t\varphi + F(D^2\varphi(x_0), \nabla\varphi(x_0), \varphi(x_0), x_0) \geq 0$.

Then $u$ is called a viscosity solution of (2.2) if it is both, a viscosity subsolution and a viscosity supersolution.

Geometrically speaking, $u$ is a viscosity solution if, for every $C^2$ function $\varphi$ that touches $u$ from above at $x_0$, there holds $F[\varphi(x_0)] \leq 0$ and if $\varphi$ touches the graph of $u$ from below at $x_0$, there holds $F[\varphi(x_0)] \geq 0$.

**Example 2.1.4.** If we consider the following PDE

$$-\partial_t u(t,x) + |\nabla u(t,x)| = 1 \text{ with } t \in (0,1] \text{ and } x \in (-1,1)$$
$$u = 0 \text{ on } \{1\} \times (-1,1) \cup (0,1) \times \{-1,1\}. \tag{2.3}$$

Then with $\Lambda = \{-1,1\}, f^\lambda = 1, \sigma^\lambda = 0, \mu^\lambda = \lambda$, we get the HJB equation

$$\partial_t u(t,x) - \sup_{\lambda \in \Lambda}\{-\lambda \nabla u(t,x) - 1\} = 0 \text{ for } t \in (0,1] \text{ and } x \in (-1,1)$$
$$u = 0 \text{ on } \{1\} \times (-1,1) \cup (0,1) \times \{-1,1\}. \tag{2.4}$$

Then the unique viscosity solution for this equation is given by

$$u(t,x) = \min(1-t, 1-|x|). \tag{2.5}$$

**Theorem 2.1.5.** Let $u \in C(\bar{\Omega}_T) \cap C^{2,1}(\Omega_T)$. Then $u$ is a viscosity solution of (2.2) if and only if $u$ is a classical pointwise solution of (2.2).

## 2.1.3. Final Value Problem and Comparison

In viscosity solution theory for second order fully nonlinear equations the comparison principle gives us the uniqueness of a viscosity solution. Generally speaking, the comparison principle asserts that if $F$ is elliptic, $u$, $v$ are, respectively, a viscosity subsolution and a viscosity supersolution, i.e., if $-\partial_t u + F[u] \leq 0$ and $-\partial_t u + F[v] \geq 0$ and $u \leq v$ on $\partial \Omega_T$, then $u \leq v$ in all of $\Omega_T$. Clearly such a result leads to uniqueness of viscosity solutions, namely, if $\partial_t u + F[u] = \partial_t v + F[v]$ in $\Omega$ and $u = v$ on $\partial \Omega_T$, then $u = v$. Indeed, if $u,v$ are two viscosity solutions with $u = v$ on $\partial \Omega_T$, then
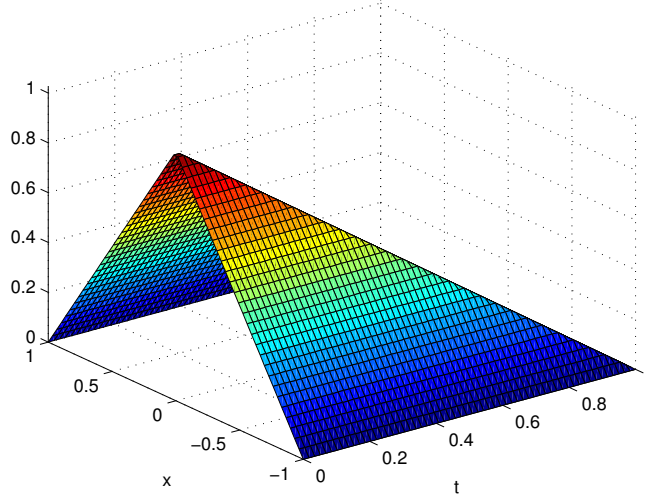
Figure 2.2.: $\min(1 - t, 1 - |x|)$

$-\partial_t u + F[u] \leq 0, -\partial_t v + F[v] \geq 0$, and $u \leq v$, implying $u \leq v$ in $\Omega_T$. Switching the roles of $u$ and $v$ we obtain $u \geq v$ and so $u = v$.

For that, we will consider (2.2) with boundary:

$$-\partial_t u + F(x, t, u, \nabla u, D^2 u) = 0 \text{ in } \Omega_T$$
$$u = g \text{ on } \partial\Omega_T$$
(2.6)

**Definition 2.1.6.** An upper semi-continous function $u \in USC(\bar{\Omega}_T)$ is a viscosity subsolution of (2.6) if $u$ is a viscosity subsolution of the (2.2) in the sense of the above definition and $u \leq g$ on $\partial\Omega_T$.

A lower semi-continous function $u \in LSC(\bar{\Omega}_T)$ is a viscosity supersolution of (2.6) if $u$ is a viscosity supersolution of (2.2) in the sense of definition and $u \geq g$ on $\partial\Omega_T$.

A function $u \in C(\bar{\Omega}_T)$ is a viscosity solution if both holds true.

Now we will cite some theorems in [9].

**Theorem 2.1.7.** Let $\Omega \subset \mathbb{R}^d$ be open and bounded. Let $F \in C([0, T] \times \bar{\Omega} \times \mathbb{R} \times \mathbb{R}^d \times S_n(\mathbb{R}))$ be continuous, proper, and degenerate elliptic with the same function $w$. If $u$ is a subsolution of (2.6) and $v$ is a supersolution of (2.6) then $u \leq v$ on $[0, T] \times \Omega$.

16

**Assumption 2.1.8.** We assume that there exists $C \geq 0$ such that for all $\lambda \in \Lambda, x, y \in \mathbb{R}^d$ and $t, x \in [0, T]$

$$
\begin{aligned}
|\mu^\lambda(t, x) - \mu^\lambda(s, y)| &\leq C(|x - y| + |t - s|) \\
|\sigma^\lambda(t, x) - \sigma^\lambda(s, y)| &\leq C(|x - y| + |t - s|) \\
|\mu^\lambda(t, x)| &\leq C(1 + |x|) \\
|\sigma^\lambda(t, x)| &\leq C(1 + |x|)
\end{aligned}
\tag{2.7}
$$

**Assumption 2.1.9.**

$$
\begin{aligned}
|f^\lambda(t, x) - f^\lambda(s, y)| &\leq C(|x - y| + |t - s|) \\
|f^\lambda(t, x)| &\leq C(1 + |x|) \\
|g^\lambda(t, x)| &\leq C(1 + |x|)
\end{aligned}
\tag{2.8}
$$

**Theorem 2.1.10.** Given assumptions (2.1.8) and (2.1.9) then there is at most one viscosity solution to the HJB final value problem.

**Theorem 2.1.11.** Provided that the value function is uniformly continous up to the boundary. $u$ is a viscosity solution of the HJB equation with no boundary and if furthermore $u = g$ on the boundary, then $u$ is a viscosity solution of the HJB with boundary.

## 2.2. The Barles-Souganidis Convergence Argument

Barles and Souganidis showed in [1] the convergence of a wide class of approximation schemes to the solution of fully nonlinear second order degenerate elliptic or degenerate parabolic PDE's. They proved that any monotone, stable and consistent scheme converges to the unique viscosity solution, provided that there exists a comparison principle, which is the case in our setting.

## 2. Viscosity Solution and the Barles-Souganidis Convergence Argument

Again, consider the fully nonlinear operator

$$-\partial_t u + F(D^2 u, \nabla u, u) = 0 \text{ in } \mathbb{R}^d$$
$$u(0, x) = u_0(x) \text{ in } \mathbb{R}^d, \tag{2.9}$$

while $F$ is continuous in all of it's arguments and degenerate elliptic.

**Definition 2.2.1.** We say (2.9) satisfies the strong comparison principle for a bounded solution, if for all bounded functions $u \in USC$ and $v \in LSC$ it holds:

- $u$ is a viscosity subsolution

- $v$ is a viscosity supersolution

- the boundary condition holds in the viscosity sense

$$\max\{\partial_t u - F[u], u - u_0\} \geq 0 \text{ on } \{0\} \times \mathbb{R}^d$$
$$\min\{\partial_t u - F[u], u - u_0\} \leq 0 \text{ on } \{0\} \times \mathbb{R}^d,$$

then we have $u \leq v$ on $[0, T] \times \mathbb{R}^d$.

Let's consider the general numerical scheme

$$K(h, t, x, u_h(t, x), [u_h]_{t,x}) = 0 \text{ for } (t, x) \in G_h \setminus \{t = 0\},$$

where $h = (\Delta t, \Delta x)$, $G_h = \Delta t\{0, 1, ..., n_d\} \times \Delta x \mathbb{Z}^d$, $[u_h]_{t,x}$ stands for the value of $u_h$ at other points than $(t, x)$.

(i) Monotonicity.
$$K(h, t, x, r, a) \geq K(h, t, x, r, b) \text{ whenever } a \leq b,$$

where this monotonicity assumption can be weakened. We only need it to hold approximately, with an error that vanishes to 0 as $h$ goes to zero.

(ii) Solvability and Stability. For arbitrary $h > 0$, there exists a solution $u_h \in B(\bar{\Omega}_T)$ to

$$K[u_h] = 0, x \in \bar{\Omega} \tag{2.10}$$

also, there also exists a constant $C > 0$ such that

$$||u_h||_{L^\infty} \leq C.$$

(iii) Consistency.

For all $x \in \bar{\Omega}$ and $\varphi \in C^\infty$ there holds

$$\limsup_{(h,t-s,y-x,\xi)\to 0} \frac{K(h,s,y,\varphi(y)+\xi,\varphi+\xi)}{h} \leq -\partial_t\varphi + F(D^2\varphi(x), \nabla\varphi(x), \varphi(x), x)$$

$$\liminf_{(h,t-s,y-x,\xi)\to 0} \frac{K(h,s,y,\varphi(y)+\xi,\varphi+\xi)}{h} \geq -\partial_t\varphi + F(D^2\varphi(x), \nabla\varphi(x), \varphi(x), x)$$

**Remark 2.2.2.** If $F$ is not continuous in all of its arguments, then $F$ has to be replaced by its upper and and lower semi-continuous envelopes $F^*$ and $F_*$, respectively.

To approximate a degenerate parabolic PDE $-\partial_t u + F[u] = 0$, in our case $F[u] := Hu$ with $Hu := \sup_\lambda(L^\lambda u - f^\lambda)$, we consider a not more specified sequence of numerical schemes in the $i$-th refinement level

$$K_i[u_i](s_i^k, x_i^l) = 0,$$

with solutions $u_i$, where $\{x_i^l\}_l$ is the set of grid points and $\{s_i^k\}_k$ the set of time nodes. Under a stability condition, one can define the upper and lower envelope of the sequence by

$$u^*(t,x) := \sup_{(s_i^k,x_i^l)_{i\in\mathbb{N}}\to(t,x)} \limsup_{i\to\infty} u_i(s_i^k, x_i^l)$$

$$u_*(t,x) := \inf_{(s_i^k,x_i^l)_{i\in\mathbb{N}}\to(t,x)} \liminf_{i\to\infty} u_i(s_i^k, x_i^l).$$

We obviously get then $u^* \geq u_*$. Then in [?] it is proven, that if $u^* - w$ has a strict local maximum, for smooth $w$, also $u_i - \mathcal{I}_i w$ has a strict local maximum for a nearby point $(s_i^k, y_i^l)$, where $\mathcal{I}_i$ is a nodal

19

interpolation operator. A certain monotonicity assumption implies

$$0 = K_i[u_i](s_i^k, x_i^l) \geq K_i[\mathcal{I}w](s_i^k, x_i^l),$$

together with the consistency condition

$$K_i[\mathcal{I}_i w](s_i^k, x_i^l) \to -\partial_t w(t, x) + Hw(t, x)$$

we get

$$-\partial_t w(t, x) + Hw(t, x) \leq 0.$$

Therefore $u^*$ is a subsolution. Similar argument leads to $u_*$ being a supersolution. Finally with a comparison principle subsolutions are bounded from above by supersolutions

$$u^* \leq u_*,$$

which gives convergence.

**Theorem 2.2.3.** Assume that the problem (2.9) satisfies the strong comparison principle for bounded functions. Assume further that the scheme satisfies the consistency, monotonicity and stability properties then it's Solution $u_h$ converges locally uniformly to the unique viscosity solution of (2.9).

**Remark 2.2.4.** For example in [13] Oberman described why monotone schemes are necessary and gave an example of a scheme which is stable, but nonmonotone and nonconvergent.

# 3. Using Howard's Algorithm

### 3.0.1. Idea

The HJB equation like in (1.3) is naturally related to linear nondivergence form equations with discontinuous coefficients. The relation between these linear and nonlinear problems is that nondivergence form linear operators can be viewed as linearisations of the fully nonlinear operator. Howard's Algorithm can be interpreted as a Newton method for a nonlinear operator equation. Another name is policy iteration and it is included in several numerical methods.

### 3.0.2. Problem statement

Here we consider the following fully non-linear HJB equation.

$$\partial_t u + \inf_{\lambda \in \Lambda} \left\{ L^\lambda u + f^\lambda \right\} = 0 \text{ in } \Omega_T$$
$$u = g \text{ on } \partial^* \Omega_T$$

(3.1)

with

$$L^\lambda v = \mathrm{Tr}[a^\lambda D^2 v] + \mu^\lambda \cdot \nabla v$$

(3.2)

defined like in Chapter 1.

## 3.1. Howard's Algorithm

We consider the value functional $J^\lambda(t,x) := \mathbb{E}\Big[ \int_t^\tau f^{\lambda_s}(s, X_s) ds + h(\tau, X_\tau) \Big]$. For the optimal control $\lambda^*$ and time $t$, the value functional is the value function $u(t,x) = J^{\lambda^*}(t,x)$.

## 3. Using Howard's Algorithm

**Lemma 3.1.1.** Let $J^\lambda(t, x)$ be the value functional.

Then $J^\lambda$ is also the solution of the boundary value problem corresponding to the arbitrary but fixed control law $\lambda \in \Lambda$:

$$\partial_t J^\lambda + L^\lambda J^\lambda + f^\lambda = 0, \text{ in } \Omega_T$$
$$\text{with boundary data } J^\lambda(t, x) = g(t, x), \text{ for } (t, x) \in \partial^* \Omega_T$$

(3.3)

With this information we can define a successive approximation algorithm. The sequence of control laws is given by

$$\lambda^{k+1} = \arg \min_{\lambda \in \Lambda} \{L^\lambda(t, X_t) J^{\lambda_k} + f^\lambda(t, X_t)\}$$

(3.4)

With that, we get $L^{\lambda^{k+1}}(t, X_t) J^{\lambda^k} + f^{\lambda^{k+1}}(t, X_t) \leq L^{\lambda^k}(t, X_t) J^{\lambda^k} + f^{\lambda^k}(t, X_t)$. Now let $J^{\lambda^{k+1}}$ be defined as the solution of (3.3) corresponding to the new control law $\lambda^{k+1}$:

$$\partial_t J^{\lambda^{k+1}} + L^{\lambda^{k+1}} J^{\lambda^{k+1}} + f^{\lambda^{k+1}} = 0, \text{ on } \Omega_T$$
$$\text{with boundary data } J^{\lambda^{k+1}}(t, x) = g(t, x), \text{ for } (t, x) \in \partial^* \Omega_T$$

(3.5)

If we continue defining the sequences of the control laws $\lambda^k$ and their belonging value functionals $J^{\lambda^k}$ like above, then we get the following Lemma and Theorem.

**Lemma 3.1.2.** The sequence $\{J^{\lambda^k}\}_{k \in \mathbb{N}}$ satisfies:

$$J^{\lambda^{k+1}} \leq J^{\lambda^k}$$

(3.6)

**Theorem 3.1.3.** $\lambda^k$ together with $J^{\lambda^k}$ converge to the optimal feedback control law $\lambda$ and the value function $u(x, t)$ of our optimal control problem :

$$\lim_{k \to \infty} \lambda^k = \lambda$$
$$\lim_{k \to \infty} J^{\lambda^k}(t, x) = u(t, x)$$

This gives us the following Algorithm

**Algorithm 3.1.4.**      • For $k = 0$, Choose initial control law $\lambda^0 \in \Lambda$.

- Get the functional $J^{\lambda^k}(t, x)$ by solving the boundary value problem for the already known control law $\lambda^k$:

  $\partial_t J^{\lambda^k} + L^{\lambda^k} J^{\lambda^k} + f^{\lambda^k} = 0$, on $\Omega_T$

  with boundary $J^{\lambda^k}(t, x) = g(t, x)$, for $(t, x) \in \partial^* \Omega_T$

- Compute the control law $\lambda^{k+1}$, by solving:

  $\lambda^{k+1} = \arg \max_{\lambda \in \Lambda} \{ f^\lambda(t, X_t) + L^\lambda(t, X_t) J^k \}$

- $k = k + 1$ and go back to the second step

## 3.2. Computational Implementation

For using our successive Algorithm, we still need to solve the PDE (3.5) and the Optimization problem (3.4). We can takle (3.5) with standard methods for solving linear parabolic PDEs, like the heat equation. Then this can be solved by finite difference schemes with upwind differences for the first order derivatives and mixed derivatives for the second order derivatives. Now we just have a finite grid, so also the optimization problem (3.4) just needs to be solved for every grid point.

**Remark 3.2.1.** Many full methods have Howard's algorithm included, like for example the finite element method we will look at in Chapter 6.

## 3.3. Application

For the 1-dimensional applications I did use the following scheme. First we approximate the PDE (3.5) to

$$\frac{u(t_{n+1}, x_i) - u(t_n, x_t)}{\Delta t} + \mu^\lambda D^\pm u(t_n, x_i) + \sigma^\lambda \frac{u(t_n, x_{i+1}) - 2u(t_n, x_i) + u(t_n, x_{i-1})}{(\Delta x)^2} + f^\lambda, \qquad (3.7)$$

where $D^\pm$ stands for the upwind operator related to the drift $\mu$

$$D^\pm \varphi(t_n, x_i) = \frac{\varphi(t_n, x_i) - \varphi(t_n, x_{i-1})}{\Delta x} \ , \text{ if } \mu > 0$$
$$D^\pm \varphi(t_n, x_i) = \frac{\varphi(t_n, x_{i+1}) - \varphi(t_n, x_i)}{\Delta x} \ , \text{ if } \mu < 0$$

Then we defined this as $F(y)$, where $y = u(t_n, x_i)$. Then we applied the Newton method

$$y_{k+1} = y_k - \frac{F(y_k)}{F'(y_k)}$$

to get the solution $F(y) = 0$.

We applied this method to the following examples

**Example 3.3.1.** For the HJB

$$\partial_t u + \sup_{\lambda \in \Lambda} \{\lambda \partial_{xx} u + (1 - x^2)\} = 0 \text{ in } (0, 1) \times (-1, 1)$$
$$u(t, x) = 0 \text{ for } (t, x) \in \{1\} \times [-1, 1] \cup (0, 1) \times \{-1, 1\},$$

$$(3.8)$$

with $\Lambda = \{0, 0.5, 1\}$. Then the solution of this PDE is $u(t, x) = (1 - t)(1 - x^2)$. The numerical results are in Figure (3.1)

**Example 3.3.2.** For the HJB

$$\partial_t u + \sup_{\lambda \in \Lambda} \{-\lambda x \partial_{xx} u + \lambda \partial_x u - \pi \cos(\pi t) 25 x^2\} = 0 \text{ in } t \in (0, 1) \times (-1, 1)$$
$$u(t, x) = \sin(\pi t)(5x)^2 \text{ for } (t, x) \in \{1\} \times [-1, 1] \cup (0, 1) \times \{-1, 1\},$$

$$(3.9)$$

with $\Lambda = \{0, 0.5, 1\}$. Then the solution of this PDE is $u(t, x) = \sin(\pi t)(5x)^2$. The numerical results for $h = \frac{1}{10}$ are in Figure (3.2)

24

$(1{-}t)\ (1{-}x^2)$
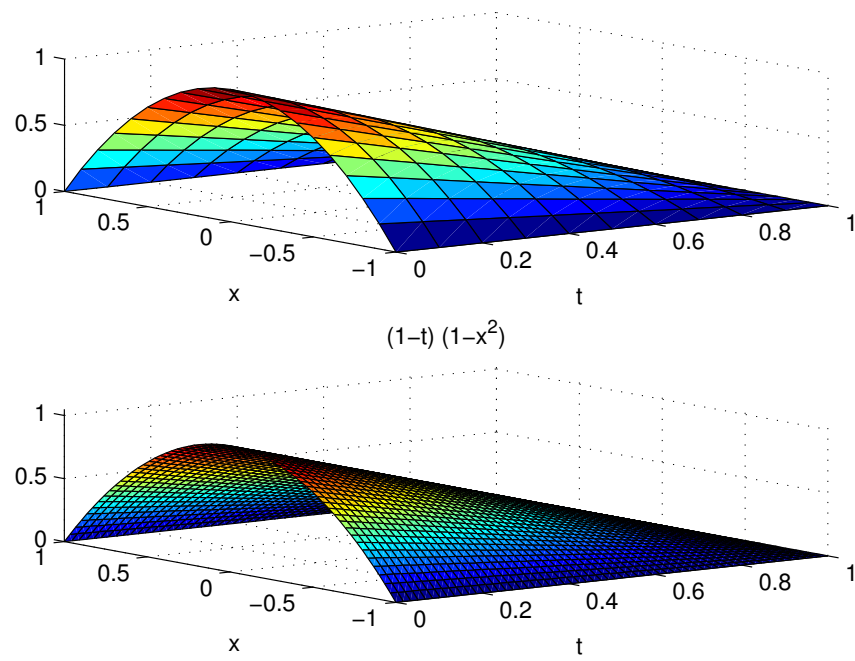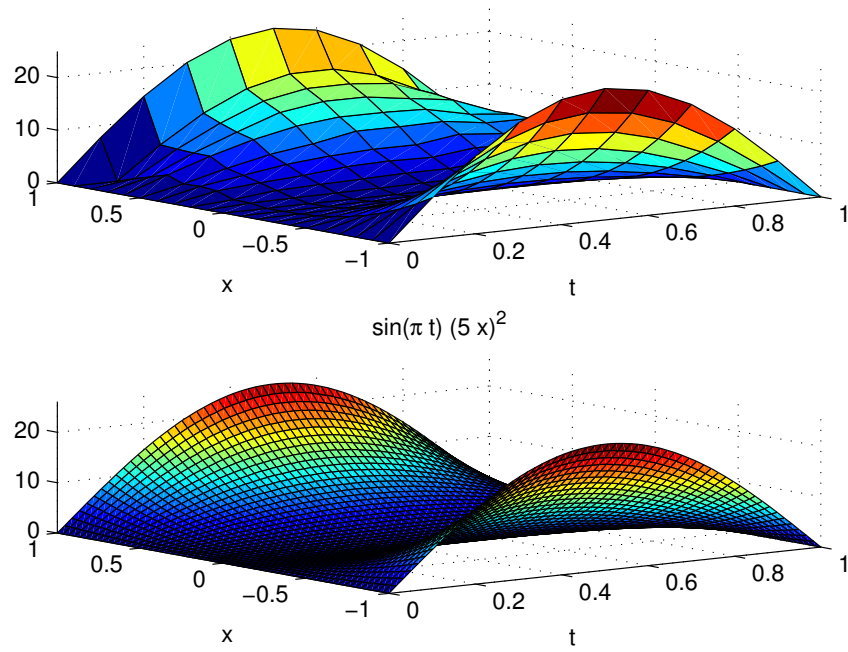


Figure 3.1.:

Figure 3.2.: top: numerical solution, botton: exact solution

# 4. Finite Difference Methods

Here we consider two different finite difference approximations in space and then the $\theta-$method for the approximation in time. For simplicity we take $h = (\Delta t, \Delta x)$ and consider the uniform grid $G_h = \Delta t\{0, 1, 2, ..., K\} \times \Delta x \mathbb{Z}^d$ and $G_h^+ = G_h \setminus \{t = 0\}$

### 4.0.1. Problem Statement

Here we consider the following fully non-linear diffusion equations

$$\partial_t u + \sup_{\lambda \in \Lambda} \left\{ - L^\lambda u - c^\lambda u - f^\lambda \right\} = 0 \text{ in } \Omega_T = \mathbb{R}^d \cup (0, T]$$

$$u(0, x) = u_0(x) \text{ on } \mathbb{R}^d$$

(4.1)

with

$$L^\lambda[u](t, x) = \text{Tr}(a^\lambda(t, x) D^2 u(t, x)) + \mu^\lambda(t, x) \nabla u(t, x)$$

(4.2)

defined like in Chapter 1.

### 4.0.2. Well-Posedness

We will use the following assumptions on the initial value problem (4.1)

**Assumption 4.0.1.** For any $\lambda \in \Lambda$, $a^{\lambda,\beta} = \frac{1}{2}\sigma^\lambda \sigma^{\lambda\top}$ for some $d \times p$ matrix $\sigma^\lambda$. There is a constant $K$ independent of $\lambda$ such that

$$\mid u_0 \mid_1 + \mid \sigma^\lambda \mid_1 + \mid \mu^\lambda \mid_1 + \mid c^\lambda \mid_1 + \mid f^\lambda \mid_1 \leq K$$

This assumption ensures that we get a well-posedness bounded Lipschitz continuous (resp. to the value $x \in \Omega$) value function, which satisfies the comparison principle.

**Proposition 4.0.2.** If assumption (5.0.1) holds. Then there exists a unique solution $u$ of the inital value problem (4.1) and a constant $C$ only depending on $T$ and $K$ from the assumption such that we have

$$\mid u \mid_1 \leq C.$$

Furthermore, if $u_1$ and $u_2$ are sub- and supersolutions of (5.1) satisfing $u_1(0, \cdot) \leq u_2(0, \cdot)$, then it holds $u_1 \leq u_2$.

## 4.1. Approximation in Space

To get the approximation in space we approximate $L^\lambda$ by a finite difference operator $L_h^\lambda$

$$L_h^\lambda \psi(t, x) = \sum_{\eta \in \mathcal{S}} C_h^\lambda(t, x, \eta)(\psi(t, x + \eta \Delta x) - \psi(t, x)) \text{ for } (t, x) \in G_h, \tag{4.3}$$

where the stencil $\mathcal{S}$ is a finite subset of $\mathbb{Z}^d \setminus \{0\}$ and where

$$C_h^\lambda(t, x, \eta) \geq 0 \text{ for all } \eta \in \mathcal{S}, (t, x) \in G_h^+, h = (\Delta t, \Delta x) > 0, \lambda \in \Lambda, \tag{4.4}$$

which gives us a difference approximation of positive type, which is a sufficient assumption for monotonicity in the stationary case.

### 4.1.1. Approximation of Kushner-Dupuis

We denote by $\{e_i\}_i^d$ the standard basis for $\mathbb{R}^d$ and we define

$$L_h^\lambda \psi(t, x) = \sum_{i=1}^d \left( a_{ii}^\lambda(x, t) \Delta_{ii} \psi(x, t) + \sum_{i \neq j} \left[ a_{ij}^{+,\lambda}(x, t) \Delta_{ij}^+ \psi(x, t) - a_{ij}^{-,\lambda}(x, t) \Delta_{ij}^- \psi(x, t) \right] \right)$$
$$\sum_{i=1}^d [\mu_i^{+,\lambda} \delta_i^+ \psi(x, t) - \mu_i^{-,\lambda} \delta_i^- \psi(x, t)], \tag{4.5}$$

where $\mu^+ = \max\{\mu, 0\}, \mu^- = -\min\{\mu, 0\}$ and

$$\delta_i^+ \psi(x,t) := \frac{\psi(x + e_i\Delta x, t) - \psi(x, t)}{\Delta x},$$

$$\delta_i^- \psi(x,t) := \frac{\psi(x, t) - \psi(x - e_i\Delta x, t)}{\Delta x},$$

$$\Delta_{ii}\psi(x,t) := \frac{\psi(x + e_i\Delta x, t) - 2\psi(x, t) + \psi(x - e_i\Delta x, t)}{\Delta x^2},$$

$$\Delta_{ij}^+ v(x,t) := \frac{1}{2\Delta x^2}\big(\psi(x + e_i\Delta x + e_j\Delta x, t) + 2\psi(x, t) + \psi(x - e_i\Delta x - e_j\Delta x, t)\big)$$
$$- \frac{1}{2\Delta x^2}\big(\psi(x + e_i\Delta x, t) + \psi(x - e_i\Delta x, t) + \psi(x + e_j\Delta x, t) + \psi(x - e_j\Delta x, t)\big),$$

$$\Delta_{ij}^- \psi(x,t) := -\frac{1}{2\Delta x^2}\big(\psi(x + e_i\Delta x - e_j\Delta x, t) + 2\psi(x, t) + \psi(x - e_i\Delta x + e_j\Delta x, t)\big)$$
$$+ \frac{1}{2\Delta x^2}\big(\psi(x + e_i\Delta x, t) + \psi(x - e_i\Delta x, t) + \psi(x + e_j\Delta x, t) + \psi(x - e_j\Delta x, t)\big).$$

This approximation is of positive type if and only if $a$ is diagonal dominant, i.e.

$$a_{ii}^\lambda(t, x) - \sum_{i \neq j} |a_{ij}^\lambda(t, x)| \geq 0 \text{ in } \Omega_T, \lambda \in \Lambda, i = 1, 2, ..., d. \tag{4.6}$$

One can prove now, that this scheme is monotone in the stationary case. For the proof we refer to [3].

### 4.1.2. Approximation of Bonnans and Zidani

We assume a finite stencil $\bar{\mathcal{S}}$ and a set of positive coefficients $\bar{a}_\eta : \eta \in \bar{\mathcal{S}}\} \subset \mathbb{R}_+$ such that

$$a^\lambda(t, x) = \sum_{\eta \in \bar{\mathcal{S}}} \bar{a}_\eta^\lambda(t, x)\eta^\top\eta \text{ in } \Omega_T, \lambda \in \Lambda, \tag{4.7}$$

which also ensures the approximation to be of positive type. The approximation of Bonnans and Zidani is then given by

$$L_h^\lambda\psi = \sum_{\eta \in \bar{\mathcal{S}}} \bar{a}_\eta^\lambda\Delta_\eta\psi + \sum_{i=1}^d [\mu_i^{+,\lambda}\delta_i^+ - \mu_i^{-,\lambda}\delta_i^-]\psi, \tag{4.8}$$

where $\Delta_\eta$ is an approximation of $\text{Tr}[\eta\eta^\top D^2]$

$$\Delta_\eta\psi(x,t) = \frac{\psi(x+\eta\Delta x,t) - 2\psi(x,t) + \psi(x-\eta\Delta x,t)}{|\eta|^2\Delta x^2}$$

This approximation is of positive type per definition, and so monotone in the stationary case.

For both approximations there is a constant $C > 0$ such that for every $\psi \in C^4(\mathbb{R}^d)$ and $(t,x) \in G_h^+$

$$|L^\lambda\psi - L_h^\lambda\psi| \le C(|\mu^\lambda|_0|D^2\psi|_0\Delta x + |a^\lambda|_0|D^4\psi|_0\Delta x^2)$$

## 4.2. Fully Discrete Scheme

For $\theta \in [0,1]$, we set the fully discrete scheme then as

$$
\begin{aligned}
u(t,x) = u(t-\Delta t, x) &- (1-\theta)\Delta t \sup_\lambda\{-L_h^\lambda u - c^\lambda u - f^\lambda\}(t-\Delta t, x) \\
&- \theta\Delta t \sup_\lambda\{-L_h^\lambda u - c^\lambda u - f^\lambda\}(t,x) \text{ in } G_h^+
\end{aligned}
\tag{4.9}
$$

Under assumption (4.4) this this fully discrete scheme is monotone if also the following CFL condition holds

**Assumption 4.2.1.**

$$\Delta t(1-\theta)(-c^\lambda(t,x) + \sum_\eta C_h^\lambda(t,x\eta)) \le 1,$$

$$\Delta t\theta(c^\lambda(t,x) + \sum_\eta C_h^\lambda(t,x\eta)) \le 1,$$

## 4.3. Fast Algorithm for 2 Dimensions

In [2] Bonnans proposed an algorithm for computing monotone discretisations of two-dimensional HJB problems with finite stencils, with a consistency error depending on the stencil width. This is achieved by approximating the diffusion coefficient $a$ by another coefficient $\tilde{a}$ for which a monotone
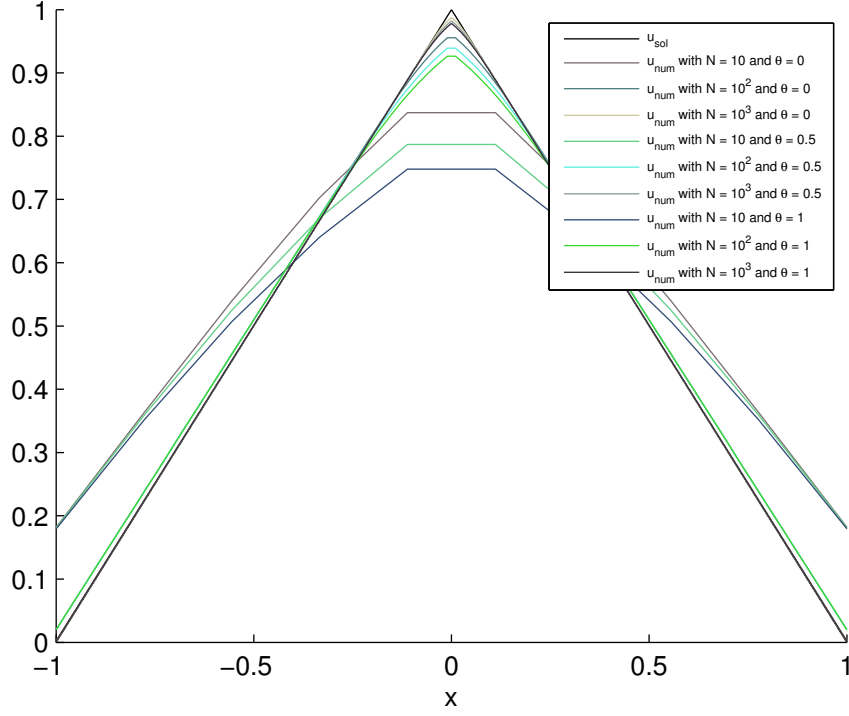
Figure 4.1.: Numerical solutions and the exact solution $u(t,x) = \min(1-t, 1-|x|)$, plotted for $t = 0$

discretisation is available on a user-specified stencil. To compute the coefficients, one needs to solve a linear programming problem at each point of the grid. They also show that this can be solved in $\mathcal{O}(p)$, if $p$ is the stencil size. The method uses the Stern-Brocot tree and on the related filling of the set of positive semidefinite matrices. Convergence is then achieved by increasing the stencil size along with mesh refinement.

## 4.4. Application

**Example 4.4.1.** Here we assumed the HJB in Example (2.1.4) and applied the Kushner-Dupuis Scheme on it, see Figure (4.1). The plot is for $t = 0$. $N$ stands for the number of grid points, i.e. $x_1, x_2, ..., x_N$.

$(2-\pi)\, \sin(x)\, \sin(y)$
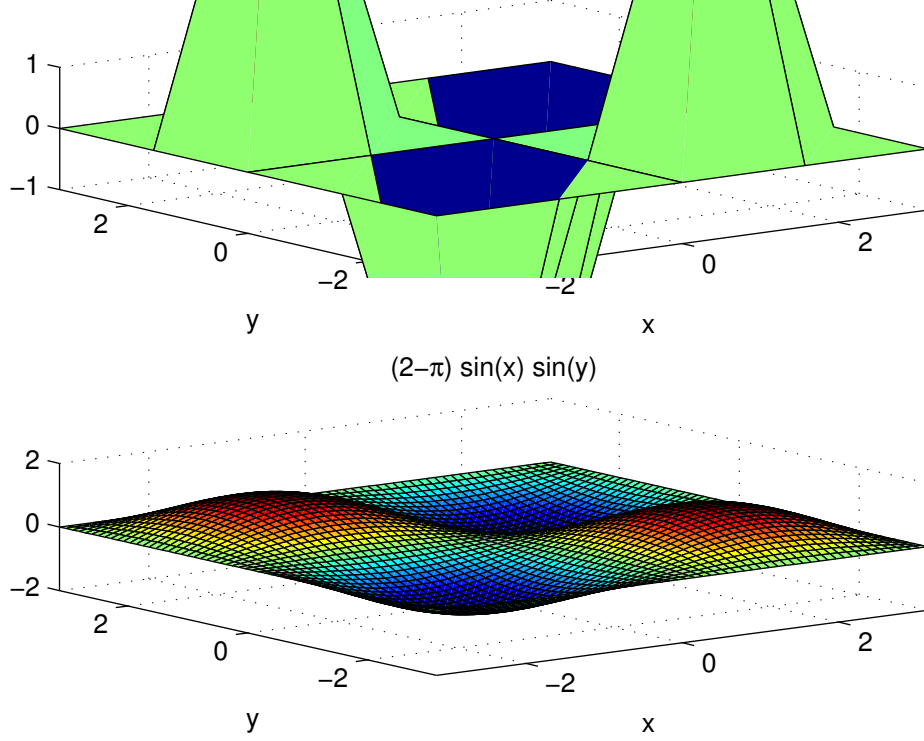


Figure 4.2.: Numerical solutions and the exact solution $u(t, x_1, x_2) = (2 - t)\sin(x_1)\sin(x_2)$, plotted for $t = \pi$

**Example 4.4.2.** If we consider the HJB equation with

$$f^\lambda(t,x) = \sin(x_1)\sin(x_2)[(1+2\beta^2)(2-t)-1] - 2(2-t)\cos(x_1)\cos(x_2)\sin(x_1+x_2)\cos(x_1+x_2),$$

$$c^\lambda(t,x) = \mu^\lambda(t,x) = 0,$$

$$\sigma^\lambda(t,x) = \sqrt{2}\begin{pmatrix} \sin(x_1+x_2) & \beta & 0 \\ \sin(x_1+x_2) & 0 & \beta \end{pmatrix}$$ with $\beta^2 = 0.1$. In this Case the HJB equation is linear and the solution of this is

$$u(t,x) = (2-t)\sin(x_1)\sin(x_2). \tag{4.10}$$

See Figure (4.2), where I applied the Kushner-Dupuis Scheme for stepsize $\pi$.

# 5. Semi-Lagrangian Schemes

### 5.0.1. Idea

The results in this Chapter are based on the work of K. Debrabant and E. R. Jakobsen [8]. Semi-Lagrangian schemes are a type difference-interpolation schemes and arise as time-discretizations of the following semi-discrete equation

$$\partial_t u - \inf_{\lambda \in \Lambda} \left\{ L_k^\lambda [\mathcal{I}_{\Delta x} u](t,x) + f^\lambda(t,x) \right\},$$

where $\mathcal{I}$ is a monotone interpolation operator of the grid and $L_k^\lambda$ is a monotone difference approximation of the operator $L^\lambda$. One advantage of these methods is the guaranteed monotonicity of the discretisations, with consistency achieved for the step-size $h \to 0$. But to achieve this the stencil size width continually increases as the mesh is refined.

Here we treat HJB equations especially posed on the entire space $\mathbb{R}^d$, rather than on bounded domains. Using a boundary, the formula used needs to be modified for points close to the boundary, for example by a one-sided asymmetric formula.

### 5.0.2. Problem Statement

Here we consider the following fully non-linear diffusion equations

$$\partial_t u - \inf_{\lambda \in \Lambda} \left\{ L^\lambda u + c^\lambda u + f^\lambda \right\} = 0 \text{ in } \Omega_T = (0,T) \times \mathbb{R}^d$$

$$u(0,x) = u_0(x) \text{ in } \mathbb{R}^d$$

(5.1)

with

$$L^\lambda[u](t,x) = \text{Tr}[a^\lambda(t,x)D^2 u(t,x)] + \mu^\lambda(t,x)\nabla u(t,x), \tag{5.2}$$

where the functions are defined like in Chapter 1. By setting $c^\lambda = 0$, we get the HJB equation in Chapter 1.

### 5.0.3. Well-Posedness

We will use the following assumptions on the initial value problem (5.1)

**Assumption 5.0.1.** For any $\lambda \in \Lambda$, $a^\lambda = \frac{1}{2}\sigma^\lambda\sigma^{\lambda\top}$ for some $d \times p$ matrix $\sigma^\lambda$. There is a constant $K$ independent of $\lambda$ such that

$$\mid u_0 \mid_1 + \mid \sigma^\lambda \mid_1 + \mid \mu^\lambda \mid_1 + \mid c^\lambda \mid_1 + \mid f^\lambda \mid_1 \leq K$$

This assumption ensures that we get a well-posedness bounded Lipschitz continuous (resp. to the value $x \in \Omega$) value function, which satisfies the comparison principle.

**Proposition 5.0.2.** If assumption (5.0.1) holds. Then there exists a unique solution $u$ of the inital value problem (5.1) and a constant $C$ only depending on $T$ and $K$ from the assumption such that we have

$$\mid u \mid_1 \leq C.$$

Furthermore, if $u_1$ and $u_2$ are sub- and supersolutions of (5.1) satisfing $u_1(0,\cdot) \leq u_2(0,\cdot)$, then it holds $u_1 \leq u_2$.

The norms are defined in the last Chapter.

## 5.1. Definition of SL-Schemes

Let $G_{\Delta t, \Delta x}$ be a not necessarily structured family of grids with

$$G = G_{\Delta t, \Delta x} = \{(t_n, x_i)\}_{n\in\mathbb{N}_0, i\in\mathbb{N}} = \{t_n\}_{n\in\mathbb{N}_0} \times X_{\Delta x},$$

for $\Delta t, \Delta x > 0$. Here $0 = t_0 < t_1 < ... < t_n < t_{n+1} < ... < T$ satisfy $\max_n \Delta t_n \le \Delta t$, where $\Delta t_n = t_n - t_{n-1}$ and $X_{\Delta x} = \{x_i\}_{i \in \mathbb{N}}$ is the set of vertices of nodes for a non-degenerate polyhedral subdivision $\mathcal{T}^{\Delta x} = \{T_j^{\Delta x}\}_{j \in \mathbb{N}}$ of $\mathbb{R}^d$. For some $\rho \in (0,1)$ the polyhedrons $T_j = T_j^{\Delta x}$ satisfy

$$\text{int}(T_j \cap T_i) \stackrel{i \ne j}{=} \emptyset, \quad \bigcup_{j \in \mathbb{N}} T_j = \mathbb{R}^d, \quad \rho \Delta x \le \sup_{j \in \mathbb{N}} \{\text{diam} B_{T_j}\} \le \sup_{j \in \mathbb{N}} \{\text{diam} T_j\} \le \Delta x,$$

where $\text{diam}(\cdot)$ is the diameter of the set and $B_{T_j}$ the greatest ball contained in $T_j$.

Let's say the matrix $\sigma = (\sigma_1, \sigma_2, ..., \sigma_P)$ with $\sigma_m$ is the $m$-th column of $\sigma$, $\psi$ is a smooth function and $k > 0$. In the second equation we replace $\psi$ by its interpolant $\mathcal{I}\psi$ on the grid $G$. We get the approximation

$$\frac{1}{2}\text{Tr}[\sigma\sigma^\top D^2 \psi(x)] = \sum_{m=1}^{P} \frac{1}{2} \frac{\psi(x + k\sigma_m) - 2\psi(x) + \psi(x - k\sigma_m)}{k^2} + \mathcal{O}(k^2)$$

$$\approx \sum_{m=1}^{P} \frac{1}{2} \frac{(\mathcal{I}\psi)(x + k\sigma_m) - 2(\mathcal{I}\psi)(x) + (\mathcal{I}\psi)(x - k\sigma_m)}{k^2}$$

$$\mu\nabla\psi(x) = \frac{1}{2} \frac{\psi(x + k^2\mu) - 2\psi(x) + \psi(x + k^2\mu)}{k^2} + \mathcal{O}(k^2)$$

$$\approx \frac{1}{2} \frac{(\mathcal{I}\psi)(x + k\mu) - 2(\mathcal{I}\psi)(x) + (\mathcal{I}\psi)(x - k\mu)}{k^2}$$

These approximations are positive monotone. If we also consider the interpolation to be monotone, we get a fully monotone discretization. The general finite difference operator has the form

$$L_k^\lambda[\psi](t, x) = \sum_{i=1}^{M} \frac{1}{2} \frac{\psi(t, x + y_{k,i}^{\lambda,+}(t,x)) - 2\psi(t,x) + \psi(t, x + y_{k,i}^{\lambda,-}(t,x))}{k^2}, \tag{5.3}$$

where for smooth functions $\psi$ holds

$$|L_k^\lambda[\psi] - L^\lambda[\psi]| \le C(|D\psi| + ... + |D^4\psi|_0)k^2 \tag{5.4}$$

Then we get the final scheme by

$$\delta_{\Delta t_n} U_i^n = \inf_{\lambda \in \Lambda} \left\{ L_k^{\lambda}[\mathcal{I}\bar{U}.^{\theta,n}]_i^{n-1+\theta} + c_i^{\lambda,n-1+\theta}\bar{U}_i^{\theta,n} + f_i^{\lambda,n-1+\theta} \right\} \text{ in } G$$

$$U_i^0 = h(x_i) \text{ in } X_{\Delta x},$$

(5.5)

where $\theta \in [0,1]$ and $U_i^n = U(t_n, x_i)$, $f_i^{\lambda,n-1+\theta} = f^{\lambda}(t_{n-1} + \Delta t \theta t_n, x_i)$ with $(t_n, x_i) \in G$,
$\delta_{\Delta t}\psi(t,x) := \frac{\psi(t,x)-\psi(t-\Delta t,x)}{\Delta t}$ and $\bar{\psi}.^{\theta,n} := (1-\theta)\psi.^{n-1} + \theta\psi.^n$

**Example 5.1.1.** For $\theta = 0$ we get the Explicit Euler.

$$\frac{U(t_n, x_i) - U(t_{n-1}, x_i)}{\Delta t} = \inf_{\lambda} \left\{ L_k^{\lambda}[\mathcal{I}U(t_{n-1}, \cdot)](t_{n-1}, x_i) + c^{\lambda}(t_{n-1}, x_i)U(t_{n-1}, x_i) + f^{\lambda}(t_{n-1}, x_i) \right\}$$

If we choose $\theta = 1$ we get implicit Euler and for $\theta = \frac{1}{2}$ we get the midpoint rule.

### 5.1.1. Collocation Method

We can also interpret the scheme (5.5) as a collocation method for a derivative free equation. If

$$W^{\Delta x}(Q_T) = \{v : v \text{ is a function on } Q_T \text{ satisfying } v = \mathcal{T}v \text{ in } Q_T\}$$

denotes the interpolant space, equation (5.5) can be stated in a equivalent way:
Find $U \in W^{\Delta x}(Q_T)$ solving

$$\delta_{\Delta t_n} U_i^n = \inf_{\lambda \in \Lambda} \left\{ L_k^{\lambda}[\bar{U}^{\theta,n}]_i^{n-1+\theta} + c_i^{\lambda,n-1+\theta}\bar{U}_i^{\theta,n} + f_i^{\lambda,n-1+\theta} \right\} \text{ in } G$$

(5.6)

## 5.2. Analysis

In this section, we give assumptions under which the SL scheme (5.5) is monotone and consistent, and we also present $L^{\infty}$-stability, existence, uniqueness, and convergence results for these schemes.

## 5. Semi-Lagrangian Schemes

**Assumption 5.2.1.** For the operator $L_k^\lambda$ we will assume that

$$\sum_{i=1}^{M} [y_{k,i}^{\lambda,+} + y_{k,i}^{\lambda,-}] = 2k^2\mu^\lambda + \mathcal{O}(k^4),$$

$$\sum_{i=1}^{M} [y_{k,i}^{\lambda,+} \otimes y_{k,i}^{\lambda,+} + y_{k,i}^{\lambda,-} \otimes y_{k,i}^{\lambda,-}] = k^2\sigma^\lambda\sigma^{\lambda\top} + \mathcal{O}(k^4),$$

$$\sum_{i=1}^{M} [\otimes_{j=1}^{3} y_{k,i}^{\lambda,+} + \otimes_{j=1}^{3} y_{k,i}^{\lambda,-}] = \mathcal{O}(k^4),$$

$$\sum_{i=1}^{M} [\otimes_{j=1}^{4} y_{k,i}^{\lambda,+} + \otimes_{j=1}^{4} y_{k,i}^{\lambda,-}] = \mathcal{O}(k^4),$$

(5.7)

where $\otimes$ stands for the Outer-product. And for the interpolation operator $\mathcal{I}$ we will assume that there are $K \geq 0, r \in \mathbb{N}$ such that

$$|(\mathcal{I}\psi) - \psi|_0 \leq K|D^p\psi|_0\Delta x^p \tag{5.8}$$

for all $p \leq r$ and smooth functions $\psi$.

Further we will assume that there is a non-negative basis of functions $\{w_j(x)\}_j$ such that

$$(\mathcal{I}\psi)(x) = \sum_j \psi(x_j)w_j(x), \ w_i(x_j) = \delta_{ij}, \ \text{and} \ w_j(x) \geq 0 \ \text{for all} \ i,j \in \mathbb{N}, \tag{5.9}$$

where $\delta_{ij}$ stands for the Kronecker Delta.

**Lemma 5.2.2.** Assume that all three assumptions in (5.2.1) hold, Then we get

- The consistency error of the scheme (5.5) is bounded by

$$\frac{|1-2\theta|}{2}|\psi_{ll}|_0\Delta t + C\Big(\Delta t^2(|\psi_{ul}|_0 + |\psi_{lll}|_0 + |\nabla\psi_{ll}|_0 + |D^2\psi_{ll}|_0)$$
$$+ |D^r\psi|_0\frac{\Delta x^r}{k^2} + (|\nabla\psi|_0 + ... + |D^4\psi|_0)k^2\Big)$$

- The scheme (5.5) is monotone if the following CFL condition holds

$$(1-\theta)\Delta t\Big[\frac{M}{k^2} - c_i^{\lambda,n-1+\theta}\Big] \leq 1 \ \text{and} \ \theta\Delta t c_i^{\lambda,n-1+\theta} \leq 1 \ \text{for all} \ \lambda, n \ \text{and} \ i. \tag{5.10}$$

**Theorem 5.2.3.** If we assume that the Assumptions (5.0.1) and (5.2.1) and (5.10) holds, then

- There exists a unique bounded solution $U_h$ of (5.5).

- If $2\theta \Delta t \sup_\lambda |c^{\lambda,+}|_0 \leq 1$ :

$$|U^n|_0 \leq e^{2\sup_\lambda, |c^{\lambda,+}|_0 t_n}[|h|_0 + t_n \sup_\lambda |f^\lambda|_0],$$

  then the solution $U_h$ of (5.5) is $L^\infty$ stable

- $U_h$ converges uniformly to the solution $u$ of (5.1) for $\Delta t, k, \frac{\Delta x^r}{k^2} \to 0$.

**Remark 5.2.4.** When solving PDEs on bounded domains, the SL schemes may exceed the domain and therefore needs to be modified.

- For Dirichlet conditions, the scheme must be modified or the boundary conditions must be extrapolated.

- For Homogenous Neumann conditions, the scheme can be implemented exactly by extending in the normal direction.

- If the boundary has no regular points, no boundary condition may be imposed.

## 5.3. Specific SL Schemes

- The approximation of Falcone,

$$\mu^\lambda D\psi \approx \frac{\mathcal{I}\psi(x + h\mu^\lambda) - \mathcal{I}\psi(x)}{h}$$

corresponding to our $L_k^\lambda$ for $y_k^{\lambda,\pm} = k^2 \mu^\lambda$ and $k = \sqrt{h}$.

- The approximation of Crandall-Lions,

$$\frac{1}{2}\text{Tr}[\sigma^\lambda \sigma^{\lambda\top} D^2\psi] \approx \sum_{j=1}^{p} \frac{\mathcal{I}\psi(x + k\sigma_j^\lambda) - 2\mathcal{I}\psi(x) + \mathcal{I}\psi(x - k\sigma_j^\lambda)}{2k^2}$$

corresponding to our $L_k^\lambda$ for $y_k^{\lambda,\pm} = \pm k\sigma_j^\lambda$ and $M = p$.

- If we combine the first two approximations, we get

$$
\frac{1}{2}\mathrm{Tr}[\sigma^\lambda \sigma^{\lambda\top} D^2\psi] + \mu^\lambda D\psi \approx
$$
$$
\sum_{j=1}^{p} \frac{\mathcal{I}\psi(x + k\sigma_j^\lambda) - 2\mathcal{I}\psi(x) + \mathcal{I}\psi(x - k\sigma_j^\lambda)}{2k^2} + \frac{\mathcal{I}\psi(x + h\mu^\lambda) - \mathcal{I}\psi(x)}{k^2} \tag{5.11}
$$

- The approximation of Camilli-Falcone,

$$
\frac{1}{2}\mathrm{Tr}[\sigma^\lambda \sigma^{\lambda\top} D^2\psi] + \mu^\lambda \nabla\psi \approx \sum_{j=1}^{p} \frac{\mathcal{I}\psi(x + \sqrt{h}\sigma_j^\lambda + \frac{h}{p}\mu^\lambda) - 2\mathcal{I}\psi(x) + \mathcal{I}\psi(x - \sqrt{h}\sigma_j^\lambda + \frac{h}{p}\mu^\lambda)}{2h}
$$

- And last but not least a modified more efficient version of the last one,

$$
\frac{1}{2}\mathrm{Tr}[\sigma^\lambda \sigma^{\lambda\top} D^2\psi] + \mu^\lambda D\psi \approx \sum_{j=1}^{p-1} \frac{\mathcal{I}\psi(x + k\sigma_j^\lambda) - 2\mathcal{I}\psi(x) + \mathcal{I}\psi(x - k\sigma_j^\lambda)}{2k^2}
$$
$$
+ \frac{\mathcal{I}\psi(x + k\sigma_p^\lambda + k^2\mu^\lambda) - 2\mathcal{I}\psi(x) + \mathcal{I}\psi(x - k\sigma_p^\lambda + k^2\mu^\lambda)}{2k^2}
$$

## 5.4. Linear Interpolation SL Schemes

A natural choice to keep our scheme (5.5) monotone, is to use linear or multi-linear interpolation. In this case we call the scheme (5.5) the LISL scheme. If we apply the results of Section 6.3. to this special case we get

**Corollary 5.4.1.** Let's assume that Assumptions (5.0.1) and (5.2.1) hold, then

- The LISL scheme is monotone if the CFL conditions (5.10) hold.

- The consistency error of the LISL scheme is $\mathcal{O}(|1 - 2\theta|\Delta t + \Delta t^2 + k^2 + \frac{\Delta x^2}{k^2})$, and hence it is first order accurate when $k = \mathcal{O}(\Delta x^{\frac{1}{2}})$ and $\Delta t = \mathcal{O}(\Delta x)$ for $\theta \neq \frac{1}{2}$ or $\Delta t = \mathcal{O}(\Delta x^{\frac{1}{2}})$ for $\theta = \frac{1}{2}$.

- If $2\theta\Delta t \sup_\lambda |c^{\lambda,+}|_0 \leq 1$ and (5.10) hold, then there exists a unique bounded and $L^\infty$-stable solutiuon $U_h$ of the LISL scheme converging uniformly to the solution $u$ of (5.1) as $\Delta t, k, \frac{\Delta x}{k} \to \infty$.

So this scheme is at most first order accurate, has wide and increasing stencil and a good CFL condition.

## 5.5. **Stochastic Control**

Let's assume that $\mathcal{B}$ is a singleton such that the equation simplifies to the HJB equation for a optimal stochastic control problem. Then we can apply the dynamic programming principle.

We will assume that Assumption (5.0.1) holds and that $c^\lambda(t, x) = 0$ and all coefficients are independent of time $t$. Then we know that the viscosity solution $u$ of (5.1) is

$$u(T - t, x) = \min_{\lambda(\cdot) \in \Lambda} \mathbb{E}\Big[ \int_t^T f^{\lambda_s}(X_s)ds + g(X_T) \Big] \tag{5.12}$$

constrained to the SDE

$$X_t = x \text{ and } dX_s = \sigma^{\lambda_s}(X_s)dW_s + \mu^{\lambda_s}ds \text{ for } s > t \tag{5.13}$$

Now we will write a SL scheme like in the collocation form (5.6). Let $\{t_0 = 0, t_1, ..., t_M = T\}$ be the discrete time steps and consider the discretization of (5.12), $A_M \subset \Lambda$ is an appropriate subset of piecewise constant controls and $k_n = \sqrt{(p+1)\Delta t_n}$.

$$\tilde{u}(T - t_m, x) = \min_{\lambda \in A_M} \mathbb{E}\Big[ \sum_{k=m}^{M-1} f^{\lambda_k}(\tilde{X}_k)\Delta t_{k+1} + g(\tilde{X}_M) \Big]$$

$$\tilde{X}_m = x, \tilde{X}_n = \tilde{X}_{n-1} + \sigma^{\lambda_n}(\tilde{X}_{n-1})k_n\xi_n + \mu^{\lambda_n}(\tilde{X}_{n-1})k_n^2\eta_n, n > m, \tag{5.14}$$

where $\xi = (\xi_{n,1}, ..., \xi_{n,p})^\top$ and $\eta_n$ are independent sequences of identically distributed random variables with

$$\Pr\big((\xi_{n,1}, ..., \xi_{n,p}, \eta_n) = \pm e_j\big) = \frac{1}{2(p+1)} \text{ , for } j \in \{1, 2, ..., p\}$$

$$\Pr\big((\xi_{n,1}, ..., \xi_{n,p}, \eta_n) = e_{p+1}\big) = \frac{1}{(p+1)}$$

Now we get

$$\tilde{u}(T - t_m, x) = \min_{\lambda \in A_M} \mathbb{E}\Big[\Delta t_{m+1} f^\lambda(x) + \Delta t_{m+1} L^\lambda_{k_{m+1}}[\tilde{u}](T - t_{m+1}, x) + \tilde{u}(T - t_{m+1}, x)\Big], \qquad (5.15)$$

where we used $L^\lambda_k$ like in Chapter 5.4 the third approximation.

## 5.6. Application

We applied the above scheme with an equidistant step size $\Delta t$, where we got the following scheme then for $k = \sqrt{(p + 1)\Delta t}$, $0 = t_0, t_1, ..., T$ and the piecwise constant control space $A_M$

$$\tilde{u}(t_j, x) = \min_{\lambda \in A_M} \Big[\Delta t f^\lambda(x) + \Delta t L^\lambda_k[\tilde{u}](t_{j-1}, x) + \tilde{u}(t_{j-1}, x)\Big]$$

**Example 5.6.1.** Here we applied the scheme to the HJB equation

$$\partial_t u - \inf_\lambda[\lambda \partial_{xx} u + (1 - x^2)] = 0,$$

$$u = 0 \text{ on } [0, 1) \times \{-1, 1\} \cup \{1\} \times (-1, 1),$$

with exact solution $u(t, x) = t(1 - x^2)$. In Figure (5.2) we see the solution with stepsize $\Delta t = \frac{1}{4}$ and $\Delta x = \frac{1}{10}$ and in Figure (5.1) with stepsize $\Delta t = 0.12$ and $\Delta x = \frac{1}{10}$.

**Example 5.6.2.** Here we applied the scheme for $x \in \mathbb{R}^2$ with the distance function as a solution $u(t, x_1, x_2) = \min(t, 1 - |x_1|, 1 - |x_2|)$ to the HJB equation, where the diffusion is zero, $\mu^\lambda = (\lambda, \lambda)^\top$ and the cost function is given by $f^\lambda = 1$, $\Lambda = \{-1, 1\}$ and $\Omega_T = (0, 1) \times (-1, 1)^2$. We can see the numerical result at time $T = 1$ for $\Delta t = \Delta x = \frac{1}{25}$ in Figure (5.3).

**Example 5.6.3.** Here we applied the scheme for $(t, x_1, x_2) \in \Omega_T = (0, 1) \times (-1, 1)^2$ with the solution $u(t, x_1, x_2) = \exp(t) \exp(\frac{x_1}{\sqrt{2}}) \exp(\frac{x_2}{\sqrt{2}})$ to the HJB equation, where the drift and the cost function are both zero and the diffusion is given by $\sigma^\lambda = \begin{pmatrix} \sqrt[4]{2} & 0 \\ 0 & \sqrt[4]{2} \end{pmatrix}$. We can see the numerical result at the terminal time $T = 1$ for $\Delta t = \Delta x = \frac{1}{4}$ in Figure (5.4) and for $\Delta t = \Delta x = \frac{1}{20}$ in Figure (5.5).

Figure 5.1.: top: numerical solution with $\Delta t = \frac{1}{4}$ and $\Delta x = \frac{1}{10}$, botton: exact solution
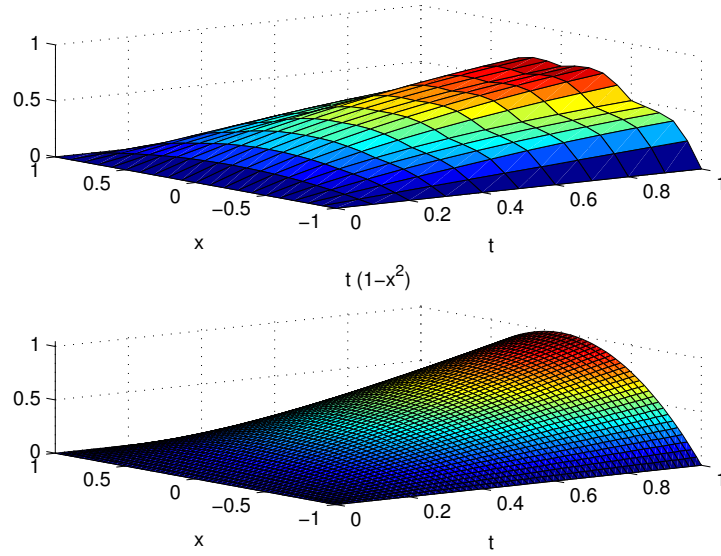


Figure 5.2.: top: numerical solution with $\Delta t = 0.12$ and $\Delta x = \frac{1}{10}$, botton: exact solution

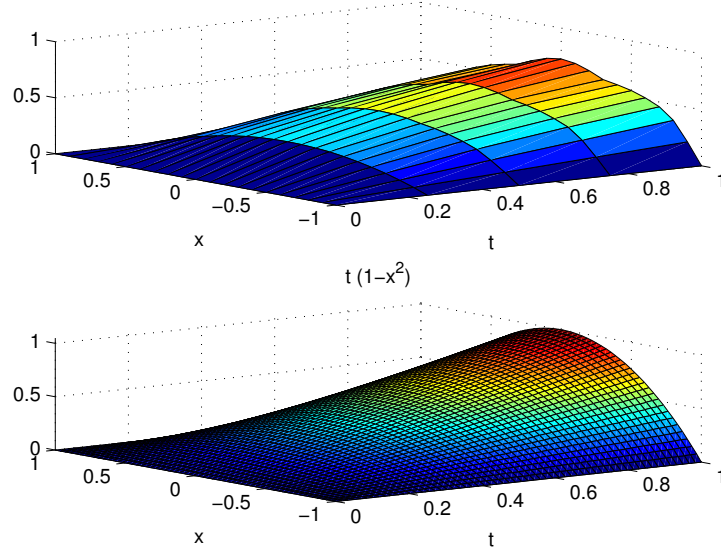Figure 5.3.: top: numerical solution with $\Delta t = \Delta x = \frac{1}{25}$, botton: exact solution



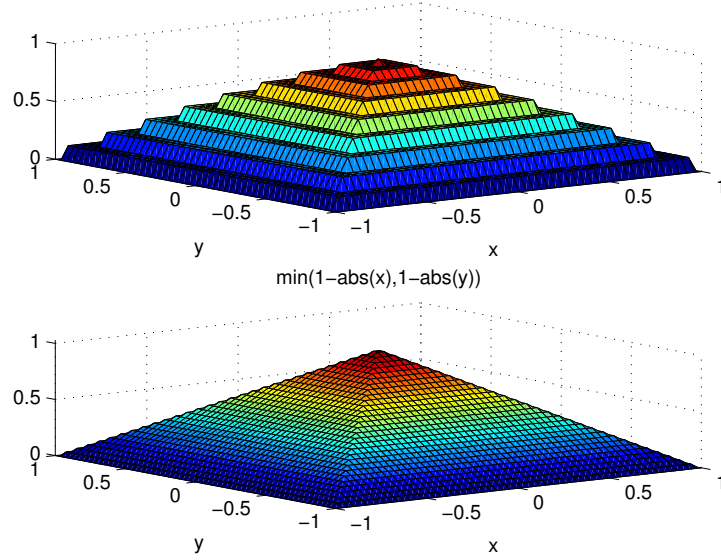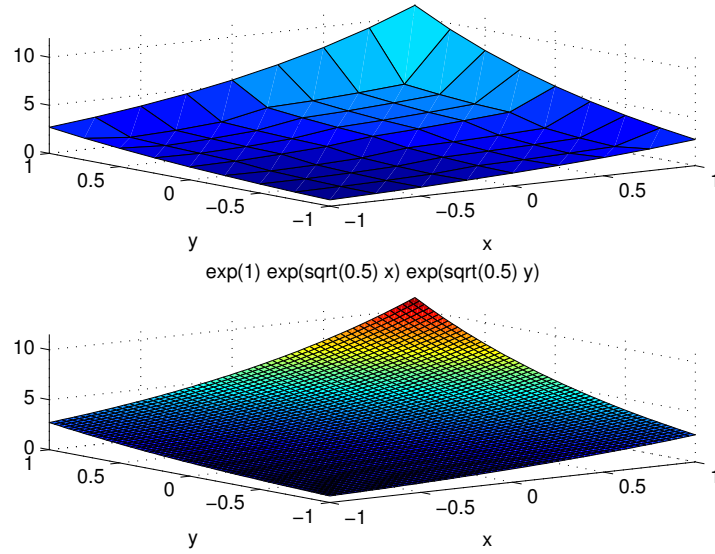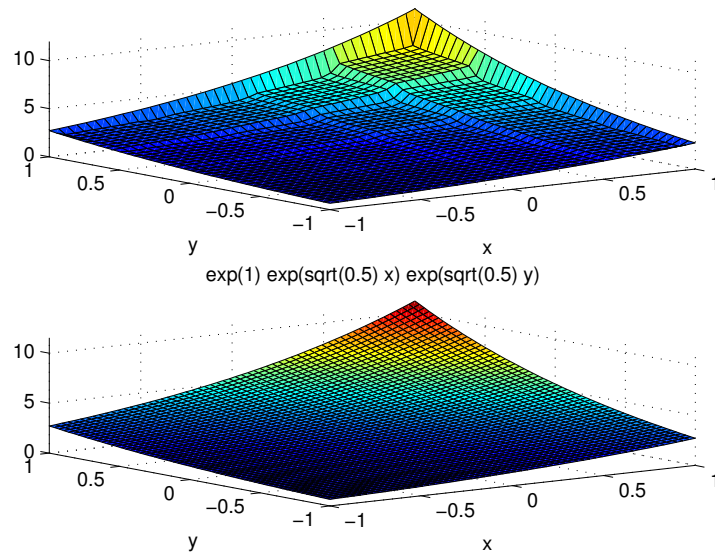Figure 5.4.: top: numerical solution with $\Delta t = \Delta x = \frac{1}{4}$, botton: exact solution

Figure 5.5.: top: numerical solution with $\Delta t = \Delta x = \frac{1}{20}$, botton: exact solution

# 6. Finite Element Method for Isotropic Diffusion

### 6.0.1. Idea

In this Chapter we will discuss the work of Max Jensen and Iain Smears [12]. Here we look at monotone P1 finite element methods on unstructered meshes for fully nonlinear HJB equations. We study stochastic control problems with isotropic diffusions, i.e. the diffusion matrix is given by a constant

$$a^\lambda = \text{const.} Id \in \mathbb{R}^d \times \mathbb{R}^d.$$

Like for the Finite Difference Methods and the Semi-Lagrangian Schemes, in this method we will use the monotonicity properties of the operator, rather than using the underlying optimal control structure. We will also use the montonicty argument of Barles and Souganidis to guarantee the convergence to the viscosity solution, but we have to modify some assumptions, since we are not dealing with finite differences.

To get rid of the Laplacian, we will use the following finite element approach in this paper

$$a(y)\Delta u(y) = a(y) \int \Delta u(y)\hat{\phi}(x)dx \approx -a(y^l) \int \nabla u(x) \cdot \nabla \hat{\phi}(x)dx = -a(y^l) \int \nabla Pu(x) \cdot \nabla \hat{\phi}(x)dx, \ (6.1)$$

where $\hat{\phi}$ is a hat test function and in the last step we used an orthogonal protection $Pu$ with respect to $\int \nabla v \cdot \nabla w \ dx$.

## 6.0.2. Problem Statement

Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$, $\Lambda$ be a compact metric space and we assume $\{a^\lambda\}_{\lambda\in\Lambda}$, $\{\mu^\lambda\}_{\lambda\in\Lambda}$, $\{c^\lambda\}_{\lambda\in\Lambda}$, $\{f^\lambda\}_{\lambda\in\Lambda}$ to be equicontinuous. The bounded linear operator is defined by:

$$L^\lambda w := -a^\lambda \Delta w + \mu^\lambda \nabla w + c^\lambda w, \text{ with } w \in H_0^1(\Omega), \lambda \in \Lambda,$$

with $a^\lambda \geq 0$. The HJB considered is then:

$$-\partial_t v + \sup_\lambda (L^\lambda v - f^\lambda) = 0 \text{ in } (0,T) \times \Omega,$$
$$v = 0 \text{ on } (0,T) \times \partial\Omega, \tag{6.2}$$
$$v = v_T \text{ on } \{T\} \times \bar{\Omega},$$

where we assume $f^\lambda \geq 0$ pointwise, $v_T \in C(\bar{\Omega})$ with $v_T \geq 0$ on $\bar{\Omega}$. Furthermore, we will assume:

$$\sup_{\lambda\in\Lambda} ||(a^\lambda, \mu^\lambda, c^\lambda, f^\lambda)||_{C(\bar{\Omega})\times C(\bar{\Omega},\mathbb{R}^d)\times C(\bar{\Omega})\times C(\bar{\Omega})} < \infty, \quad \sup_{\lambda\in\Lambda} ||L^\lambda||_{C^2(\bar{\Omega})\to C(\bar{\Omega})} < \infty.$$

For the definition of a viscosity solution, we refer to Chapter 2.

# 6.1. Definition of the Numerical Method

First we will define the numerical scheme and then give a numerical method.

## 6.1.1. Numerical Scheme

Let $V_i$ be a sequence of piecewise linear shape-regular finite element spaces with nodes $y_i^l$ and hat functions $\phi_i^l$. Here $l$ is the index ranging over the nodes of the finite element mesh. Let $V_i^0$ be the subspace of functions which satisfy homogeneous Dirichlet conditions. Let $N := \dim V_i^0$ and the nodes $y_i^l \in \Omega$ for $l \leq N$. Now normalize the hat functions by $\hat{\phi}_i^l := \phi_i^l / ||\phi_i^l||_{L^1(\Omega)}$. The largest diameter of an element in each $i$-mesh is denoted by $(\Delta x)_i$, where we will assume convergence to zero for $i \to \infty$.

For every sequence step $i$, we will cut the interval $[0, T]$ in $T/h_i \in \mathbb{N}$ parts, where $h_i$ is the uniform time step. The set of time steps is then $S_i := \{s_i^k : k = 0, 1, ..., \frac{T}{h_i}\}$.

The derivation after time will be discretized by

$$(d_i w(s_i^k, \cdot))_l := \frac{w(s_i^{k+1}, y_i^l) - w(s_i^k, y_i^l)}{h_i}$$

We will approximate the fully nonlinear operator $L^\lambda$ by two linear operators $E_i^\lambda, I_i^\lambda : H_0^1(\Omega) \to H^{-1}$, so that we get $L^\lambda \approx E_i^\lambda + I_i^\lambda$. Where later in the algorithm the first one will stand for the explicit part and the second one for the implicit

$$E_i^\lambda w = -\bar{a}_i^\lambda \Delta w + \bar{b}_i^\lambda \cdot \nabla w + \bar{c}_i^\lambda w,$$

$$I_i^\lambda w = -\bar{\bar{a}}_i^\lambda \Delta w + \bar{\bar{b}}_i^\lambda \cdot \nabla w + \bar{\bar{c}}_i^\lambda w.$$

We also need $\bar{c}_i^\lambda, \bar{\bar{c}}_i^\lambda \geq 0$ and $||\bar{c}_i^\lambda||_{L^\infty} + ||\bar{\bar{c}}_i^\lambda||_{L^\infty} \leq \alpha$, for some $\alpha \in \mathbb{R}$ and all $\lambda \in \Lambda$,

The non-negative cost function $f^\lambda$ will be approximated by a non-negative $f_i^\lambda \approx f^\lambda$. Our first assumption is then

**Assumption 6.1.1.** For all sequences of nodes $(y_i^l)_{i \in \mathbb{N}}$:

$$\lim_{i \to \infty} \sup_{\lambda \in \Lambda} \left( ||a^\lambda - (\bar{a}_i^\lambda(y_i^l) + \bar{\bar{a}}_i^\lambda(y_i^l))||_{L^\infty(\text{supp } \hat{\phi}_i^l)} + ||b^\lambda - (\bar{b}_i^\lambda + \bar{\bar{b}}_i^\lambda)||_{L^\infty(\Omega)} + ||c^\lambda - (\bar{c}_i^\lambda + \bar{\bar{c}}_i^\lambda)||_{L^\infty(\Omega)} + ||f^\lambda - f_i^\lambda||_{L^\infty(\Omega)} \right) = 0$$

## 6.1.2. Numerical Method

Now we will use the finite element approach (6.1)

$$-a(x)\Delta w(x) \approx -a(y_i^l)\langle \Delta w, \hat{\phi}_i^l \rangle_{L^2} = a(y_i^l)\langle \nabla w, \nabla \hat{\phi}_i^l \rangle_{L^2}$$

to discretize the operators $E_i^\lambda, I_i^\lambda$ by operators $\mathrm{E}_i^\lambda, \mathrm{I}_i^\lambda$. So that we get a discretization that is consistent in the sense needed for the analysis by approximating the diffusion.

$$(\mathrm{E}_i^\lambda w)_l := \bar{a}_i^\lambda(y_i^l)\langle \nabla w, \nabla \hat{\phi}_i^l \rangle_{L^2} + \langle \bar{\mu}_i^\lambda \nabla w + \bar{c}_i^\lambda w, \hat{\phi}_i^l \rangle_{L^2},$$

$$(\mathrm{I}_i^\lambda w)_l := \bar{\bar{a}}_i^\lambda(y_i^l)\langle \nabla w, \nabla \hat{\phi}_i^l \rangle_{L^2} + \langle \bar{\bar{\mu}}_i^\lambda \nabla w + \bar{\bar{c}}_i^\lambda w, \hat{\phi}_i^l \rangle_{L^2}, \qquad (6.3)$$

$$(\mathrm{C}_i^\lambda w)_l := \langle f_i^\lambda, \hat{\phi}_i^l \rangle_{L^2}.$$

Obtain the numerical solution $v_i(T, \cdot) \in V_i$ by interpolation of the boundary function $v_T$. Then $v_i(s_i^k, \cdot) \in V_i^0$ at time $s_i^k$ is defined by the implicit (if $I_i^\lambda \equiv 0$ explicit) equation

$$-d_i v_i(s_i^k, \cdot) + \sup_\lambda (E_i^\lambda v_i(s_i^{k+1}, \cdot) + I_i^\lambda v_i(s_i^k, \cdot) - C_i^\lambda) = 0 \qquad (6.4)$$

**Assumption 6.1.2.** Assume for each $\lambda \in \Lambda$ that $E_i^\lambda$ has non-positive off-diagonal entries on $V_i$. Let the time step $h_i$ be small enough so that all operators $h_i E_i^\lambda - Id$ have just non-positive entries. Assume that for each $\lambda \in \Lambda$ that $I_i^\lambda$ satisfies the local monotonicity property, i.e. for all $v \in V_i$ with nonpositive local minimum at $y_i^l$ we have $(I_i^\lambda v)_l \leq 0$.

We will use the short-notation

$$\mathrm{I}_i^{k,w} = \mathrm{I}_i^{\lambda_i(w)^{l,k}} \text{ and } \mathrm{E}_i^{k,w} = \mathrm{E}_i^{\lambda_i(w)^{l,k}},$$

where

$$\lambda_i^{l,k}(w) = \arg\sup_\lambda \left( \mathrm{E}_i^\lambda w(s_i^{k+1}, \cdot) + \mathrm{I}_i^\lambda w(s_i^k, \cdot) - \mathrm{F}_i^\lambda \right)_l$$

### 6.1.3. Solution Algorithm

If $\lambda_i^{l,k}(w)$ maximises $\sup_\lambda \left( \mathrm{E}_i^\lambda w(s_i^{k+1}, \cdot) + \mathrm{I}_i^\lambda w(s_i^k, \cdot) - \mathrm{C}_i^\lambda \right)_l$, we will use the shorter notation $\mathrm{E}_i^{k,w}, \mathrm{I}_i^{k,w}, \mathrm{C}_i^{k,w}$ instead of $\mathrm{E}_i^{\lambda_i^{l,k}(w)}, \mathrm{I}_i^{\lambda_i^{l,k}(w)}, \mathrm{C}_i^{\lambda_i^{l,k}(w)}$.

We can solve the optimization problem (6.4) by the following algorithm, which uses Howard's Algorithm in Chapter 3 as a tool.

**Algorithm 6.1.3.** Given $k \in \mathbb{N}$ and $v_i(s_i^{k+1}, \cdot) \in V_i^0$ for $k = \{0, ..., \frac{T}{h_i} - 1\}$, choose $\lambda_0 \in \Lambda$ to start with and then inductively for $m = \{1, 2, 3, ...\}$ find $w_m \in V_i^0$ such that

$$(h_i \mathrm{I}_i^{w_m} + Id)w_{m+1} = -(h_i \mathrm{E}_i^\lambda - Id)v_i(s_i^{k+1}, \cdot) + h_i \mathrm{F}_i^{w_m}$$

## 6.2. Analysis

### 6.2.1. Well-Posedness

**Theorem 6.2.1.** There exists a unique numerical solution $v_i : S_i \to V_i^0$ that solves (?) and (??). Moreover, $0 \le v_i \le v_i^\lambda$ for each $\lambda \in \Lambda$. Given that $v_i(s_i^{k+1}, \cdot) \in V_i^0$ for $k \in \{0, 1, ..., \frac{T}{h_i} - 1\}$, the iterates of Algorithm (6.1.3) converge superlinearly to the unique solution $v_i(s_i^k, \cdot)$ of (?).

The monotonicity and mass-lumping helps us to get $L^\infty$ bounds of parabolic Galerkin methods, like in our case.

**Lemma 6.2.2.** For all $i \in \mathbb{N}$ one has $||(h_i I_i^\lambda + Id)^{-1}||_\infty \le 1$ and $||(h_i E_i^\lambda + Id)^{-1}||_\infty \le 1$, where the norms are the matrix $\infty-$norms.

**Corollary 6.2.3.** The numerical solutions $v_i$ are uniformly bounded in the $L^\infty$ norm, i.e. there is a $C > 0$ such that for all $i \in \mathbb{N}$ and $\lambda \in \Lambda$

$$||v_i||_{L^\infty(S_i \times \Omega)} \le ||v_i^\lambda||_{L^\infty(S_i \times \Omega)} \le ||v_T||_{L^\infty(\Omega)} + T||f_i^\lambda||_{L^\infty(\Omega)} \le C$$

### 6.2.2. Consistency properties

Barles and Souganidis used in their Framework a nodal interpolation, but this does not give us strongly consistence in this the FEM case. Therefore we use a orthogonal projection with respect to $\int \nabla v \nabla w \, dx$.

**Assumption 6.2.4.** There are linear mappings $P_i$ satisfying

$$\langle \nabla P_i w(T, \cdot), \nabla \hat\phi_i^l \rangle = \langle \nabla w(T, \cdot), \nabla \hat\phi_i^l \rangle \ \forall t \in [0, T] \text{ and } \forall \hat\phi_i^l \in V_i^0 \text{ and a given } w \in C([0, T], H^1(\Omega)).$$

(6.5)

And there is a constant $C \geq 0$ such that for every $w \in C^\infty(\mathbb{R}^d)$ and $i \in \mathbb{N}$

$$||P_i w||_{W^{1,\infty}(\Omega)} \leq C||w||_{W^{1,\infty}(\Omega)} \text{ and } \lim_{i \to \infty} ||P_i w - w||_{W^{1,\infty}(\Omega)} \tag{6.6}$$

**Lemma 6.2.5.** Let $w \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$ and let $s_i^k \to t \in [0,T)$ for $i \to \infty$. Then

$$\lim_{i \to \infty} d_i P_i w(s_i^k, \cdot) = \partial_t w(t, \cdot) \text{ in } W^{1,\infty}(\Omega) \tag{6.7}$$

**Lemma 6.2.6.** Let $w \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$ and let $s_i^k \to t \in [0,T], y_i^k \to x \in \Omega$ for $i \to \infty$. Then

$$\lim_{i \to \infty} \left( E_i^\lambda P_i w(s_i^{k+1}, \cdot) + I_i^\lambda P_i w(s_i^k, \cdot) - F_i^\lambda \right)_l = L^\lambda w(t,x) - f^\lambda(x), \tag{6.8}$$

with uniform convergence over all $\lambda \in \Lambda$.

## 6.2.3. Super- and Subsolution

Like in Chapter 2, we will define.

$$v^*(t,x) = \sup_{(s_i^k, y_i^l)_{i \in \mathbb{N}} \to (t,x)} \limsup_{i \to \infty} v_i(s_i^k, y_i^l)$$

$$v_*(t,x) = \inf_{(s_i^k, y_i^l)_{i \in \mathbb{N}} \to (t,x)} \liminf_{i \to \infty} v_i(s_i^k, y_i^l)$$

**Theorem 6.2.7.** The function $v^*$ is a viscosity subsolution of (?) and $v_*$ is a viscosity supersolution of (?).

## 6.2.4. Uniform Convergence

Now we take a look at the initial and boundary conditions. Then we use the sub- and supersolution property to get a comparision principle to obtain uniform convergence of the numerical solutions.

$$v^{\lambda,*}(t,x) = \sup_{(s_i^k, y_i^l)_{i \in \mathbb{N}} \to (t,x)} \limsup_{i \to \infty} v_i^\lambda(s_i^k, y_i^l) \text{ for } \lambda \in \Lambda$$

**Assumption 6.2.8.** For each $(t, x) \in [0, T] \times \partial\Omega$

$$\inf_{\lambda \in \Lambda} v^{\lambda,*}(t, x) = 0$$

With Theorem (6.2.1) we get $0 \leq v_* \leq v^* \leq v^{\lambda,*}$ and with Assumption (6.2.8) we get that $v_*|_{[0,T] \times \partial\Omega} = v^*|_{[0,T] \times \partial\Omega} = 0$.

**Lemma 6.2.9.** The sub- and supersolutions $v^*$ and $v_*$ satisfy

$$v^*(T, \cdot) = v_*(T, \cdot) = v_T \text{ on } \bar{\Omega} \tag{6.9}$$

**Assumption 6.2.10.** Let $\bar{v}$ be a lower semicontinuous supersolution with $\bar{v}(T, \cdot) = v_T$ and $\bar{v}|_{[0,T] \times \partial\Omega} = 0$. Similarly, Let $\underline{v}$ be a lower semicontinuous supersolution with $\underline{v}(T, \cdot) = v_T$ and $\underline{v}|_{[0,T] \times \partial\Omega} = 0$. Then $\underline{v} \leq \bar{v}$.

If $t \in [s_i^k, s_i^{k+1}]$ we write

$$v_i(t, \cdot) = \theta v_i(s_i^k, \cdot) + (1 - \theta) v_i^{k+1}(s_i^{k+1}, \cdot),$$

for $t = \theta s_i^k + (1 - \theta) s_i^{k+1}$

**Theorem 6.2.11.** For the viscosity solution $v$ of (?), we have $v_* = v^* = v$, with $v(T, \cdot) = v_T$ and $v|_{[0,T] \times \partial\Omega} = 0$. With

$$\lim_{i \to \infty} ||v_i - v||_{L^\infty} = 0$$

## 6.3. Method of Artificial Diffusion

To guarantee monotonicity, for $a^\lambda = \tilde{a}_i^\lambda + \tilde{\tilde{a}}_i^\lambda$ we select certain artificial diffusion parameters $\bar{\eta}_i^{\lambda,l}$ and $\bar{\bar{\eta}}_i^{\lambda,l}$ so that we get $\bar{a}_i^\lambda(y_i^l) \geq \max(\tilde{a}_i^\lambda, \bar{\eta}_i^{\lambda,l})$ and $\bar{a}_i^\lambda(y_i^l) \geq \max(\tilde{\tilde{a}}_i^\lambda, \bar{\bar{\eta}}_i^{\lambda,l})$. In (6.3) for $\bar{\bar{a}}_i^\lambda, \bar{a}_i^\lambda$ we can use a artificial diffusion parameter so that we satisfy assumtions (6.1.1) and (6.1.2) holds.

In [12] they described a possible way, by using strictly acute meshes, how to achieve the local mon-

tonicity

$$(E_i^\lambda w)_l \leq 0, \ (I_i^\lambda w)_l \leq 0,$$

where $w \in V_i$ has a nonpositive local minimum at an inner node $y_i^l$.

## 6.4. Application

**Example 6.4.1.** Consider again the following HJB equation, like in Chapter 2.

$$-\partial_t v + \sup_{\lambda \in \{-1,1\}} (\lambda \partial_x v - 1) = 0 \text{ in } (0,1) \times (-1,1)$$

$$v = 0 \text{ on } [0,1] \times \{-1,1\} \cup \{1\} \times [-1,1],$$

with the viscosity solution $v(t,x) = \min(1-t, 1-|x|)$. Then we can choose the $(E_i^\lambda w)_l = \epsilon \langle \partial_x w, \partial_x \hat{\phi}_i^l \rangle + \lambda \langle \partial_x w, \hat{\phi}_i^l \rangle$, where $\epsilon$ is the artificial diffusion operator here. Then the matrix has the form

$$(E_i^\lambda w)_{lj} = \begin{cases} -\lambda/2\Delta x_i - \epsilon/\Delta x_i^2 & \text{, if j = l-1} \\ 2\epsilon/\Delta x_i^2 & \text{, if j = l} \\ \lambda/2\Delta x_i - \epsilon/\Delta x_i^2 & \text{, if j = l+1} \\ 0 & \text{otherwise,} \end{cases}$$

where for $\epsilon \geq \Delta x_i/2$, the off-diagonal terms are non-positive. This FEM applied to this specific problem is very effective. In fact method also converges with stepsize $\Delta t = \Delta x = 1$ like we can see in Figure (6.1).
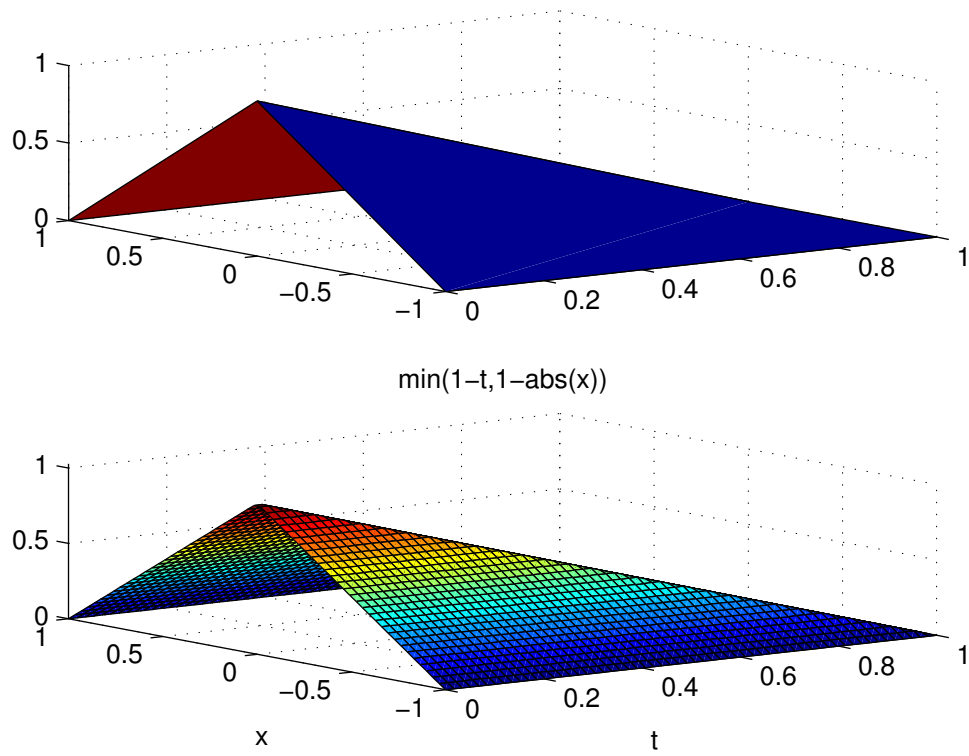
Figure 6.1.: top: numerical solution with stepsize $\Delta t = \Delta x = 1$, botton: exact solution

# 7. Vanishing Moment Method

## 7.1. Idea

The theory in this Chapter is based on the work of Feng and Neilan [10]. The general fully nonlinear second order PDE takes the form:

$$F(D^2 u(x), \nabla u(x), u(x), x) = 0, \text{ with } x \in \Omega, \tag{7.1}$$

where in our HJB time-independent case $F(D^2 u(x), \nabla u(x), u(x), x) = \sup_{\lambda \in \Lambda}[L^\lambda u - f^\lambda]$. Computing viscosity solutions of fully nonlinear second order PDEs has been impracticable. There are several reasons for that. Firstly, the strong nonlinearity is an obvious one. Secondly, the conditional uniqueness of solutions is difficult to handle numerically. Lastly and most important, the notion of viscosity solutions, which is not variational, has no equivalence at the discrete level.

To introduce the notion of viscosity solutions, Crandall and Lions [5] used the vanishing viscosity method to show existence of a solution for the first order Hamilton-Jacobi equation:

$$\partial_t u + H(\nabla u, u, x) = 0$$

The vanishing viscosity method approximates the first order Hamilton-Jacobi equation by the following regularized, second order quasilinear PDE:

$$\partial_t u^\epsilon + H(\nabla u^\epsilon, u^\epsilon, x) = \epsilon \Delta u^\epsilon \tag{7.2}$$

It was shown in [5] that there exists a unique solution $u^\epsilon$ to the regularized Cauchy problem that converges locally and uniformly to a continuous function $u$ which is defined to be a viscosity solution of the first order Hamilton-Jacobi equation. However, to establish uniqueness, the following definition of viscosity solutions for general fully nonlinear first order PDE was also proposed.

**Definition 7.1.1.(Viscosity solution first order)** To make things easier we assume $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \to \mathbb{R}$ to be continuous.

- A function $u \in C^0(\Omega)$ is called a viscosity subsolution of (7.2) if, for every $C^1-$function $\varphi(x)$ such that $u - \varphi$ has a local maximum at $x_0 \in \Omega$, there holds

$$F(\nabla\varphi(x_0), \varphi(x_0), x_0) \leq 0.$$

- A function $u \in C^0(\Omega)$ is called a viscosity supersolution of (7.2) if, for every $C^1-$function $\varphi(x)$ such that $u - \varphi$ has a local minimum at $x_0 \in \Omega$, there holds

$$F(\nabla\varphi(x_0), \varphi(x_0), x_0) \geq 0.$$

- A function $u \in C^0(\Omega)$ is called a viscosity solution of (7.2) if it is both a viscosity subsolution and a viscosity supersolution.

It was shown that every viscosity solution constructed by the vanishing viscosity method is an viscosity solution in the sense of Definition (7.1.1). Besides the uniqueness issue, another reason to favor the second definition is that it can be extended to fully nonlinear second order PDEs like seen in Chapter 2.

## 7.2. General Framework

Because it is not constructive nor variational, the notion of a viscosity solution is not good from a computational point of view. By looking for a better notion for a weak solution, we change the original

second order nonlinear PDE to the following higher order quasi-linear PDE:

$$G_\epsilon(D^r u^\epsilon) + F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x) = 0, \text{ with } r \geq 3, \epsilon > 0, \tag{7.3}$$

where $\{G_\epsilon\}$ is a family of suitably chosen linear or quasilinear operators or order $r$. With that we get the definition of a moment solution.

**Definition 7.2.1.(moment solution)** If the limit $u := \lim_{\epsilon \to 0^+} u^\epsilon$ exisits, with $u^\epsilon$ solves (7.3) for each $\epsilon > 0$, we call $u$ the moment solution of (7.2).

In order to choose the operator $G_\epsilon$, there are some reasonable assumptions, like mentioned in [10].

- $G_\epsilon$ is linear or at least quasiliniear.

- $G_\epsilon \to 0$ in some reasonable sense for $\epsilon \to 0^+$.

- If (7.2) is elliptic, $G_\epsilon(D^r u)$ should also be elliptic.

The last point implies that the operator must be of even order, so we get $r \geq 4$. So for example we can take the biharmonic operator $G_\epsilon(D^4 v) := -\epsilon \Delta^2 v$ and (7.3) becomes

$$-\epsilon \Delta^2 + F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x) = 0 \tag{7.4}$$

Now we want to take care of the boundary conditions. Here we will only consider the Dirichlet problem:

$$F(D^2 u(x), \nabla u(x), u(x), x) = 0, \ x \in \Omega$$
$$u = g, \ x \in \partial\Omega \tag{7.5}$$

We also need this boundary condition for our moment solution. Moreover, since (7.4) is a fourth order PDE, we need to impose an additional boundary condition for $u^\epsilon$. We will suggest three different types of boundary conditions for (7.3):

$$-\epsilon\Delta^2 + F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x) = 0, \ x \in \Omega$$

$$u^\epsilon \approx g, \ x \in \partial\Omega \tag{7.6}$$

$$\Delta u^\epsilon = c_\epsilon \text{ or } \frac{\partial \Delta u^\epsilon}{\partial \mathbf{n}} = c_\epsilon \text{ or } D^2 u^\epsilon \mathbf{n} \cdot \mathbf{n} = c_\epsilon, \ x \in \partial\Omega$$

Since (7.6) is a nonlinear biharmonic equation, one can use any numerical method for biharmonic problems to discretize the equation. In (**??**) they did choose Galerkin methods, where they studied conforming finite element approximations and mixed finite element approximations extensively. Here we will just mention the approach of finite element approximations in two dimensions.

**Remark 7.2.2.** If we pick the first boundary condition with $\Delta u^\epsilon = \epsilon$ on $\partial\Omega$, then with $u^\epsilon = g$ on $\partial\Omega$ we get an additional boundary condition

$$\epsilon^m \Delta u^\epsilon + u^\epsilon = \epsilon^m \epsilon + g = g + \epsilon^{m+1} \text{ on } \partial\Omega$$

**Definition 7.2.3.** We define $u^\epsilon \in H^2(\Omega)$ with $u|_{\partial\Omega} = g$ to be a solution (7.6) if for all $v \in H^2(\Omega) \cap H_0^1(\Omega)$

$$\epsilon(\Delta u^\epsilon, \Delta v) + (F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x), v) = \langle \epsilon^2, \frac{\partial v}{\partial n} \rangle_{\partial\Omega} \tag{7.7}$$

**Definition 7.2.4.** Suppose that $u^\epsilon$ solves problem (7.6). Then $\lim\limits_{\epsilon \to 0^+} u^\epsilon$ is called a weak moment solution to problem (**??**) if the convergence holds in $H^1$-weak (resp. $H^2$-weak) topology.

## 7.3. Finite Element Method in 2d

The variational formulation is then defined as:

Find $u^\epsilon \in H^2(\Omega)$ with $u^\epsilon = g$ on $\partial\Omega$ such that for any test function $\phi \in H^2(\Omega) \cap H_0^1(\Omega)$

$$-\epsilon \langle \Delta u^\epsilon, \Delta \phi \rangle_{L^2} + \langle F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x), \phi \rangle_{L^2} = -\langle \epsilon c_\epsilon, \frac{\partial \phi}{\partial \mathbf{n}} \rangle_{L^2(\partial\Omega)} \tag{7.8}$$

If we now choose $\mathcal{T}_h$ to be a quasiuniform triangular or rectangular mesh with mesh size $h \in (0,1)$ for the domain $\mathbb{R}^2$ and $U_g^h \subset H^2(\Omega)$ denote one of the conforming finite element spaces whose functions take the boundary $g$. Then our finite element method is defined as:

Find $u_h^\epsilon \in U_g^h$ such that

$$-\epsilon \langle \Delta u_h^\epsilon, \Delta \phi_h \rangle_{L^2} + \langle F(D^2 u_h^\epsilon, \nabla u_h^\epsilon, u_h^\epsilon, x), \phi_h \rangle_{L^2} = -\langle \epsilon c_\epsilon, \frac{\partial \phi_h}{\partial \mathbf{n}} \rangle_{L^2(\partial\Omega)} \text{ for all } \phi_h \in U_0^h \tag{7.9}$$

## 7.4. Parabolic Case

If we are dealing with fully nonlinear PDEs of parabolic type, like our HJB equation:

$$-\partial_t u + F(D^2 u, \nabla u, u, x, t) = 0 \text{ in } \Omega_T \tag{7.10}$$

Then the following may be one possible vanishing moment approximation to this

$$-\epsilon \Delta^2 u^\epsilon - \partial_t u^\epsilon + F(D^2 u^\epsilon, \nabla u^\epsilon, u^\epsilon, x, t) = 0 \text{ in } \Omega_T. \tag{7.11}$$

Which gives us

$$-\epsilon \Delta^2 u^\epsilon - \partial_t u^\epsilon + \sup_{\lambda \in \Lambda}[L^\lambda u^\epsilon - f^\lambda] = 0 \text{ in } \Omega_T.$$

# A. Notations

## A.1. Sobolev Spaces

$$W^{k,p}(\Omega) := \{f \in L^p(\Omega) \mid D^\alpha f \in L^p(\Omega), |\alpha| \le k\}$$

For $\Omega \in \mathbb{R}^N$ open set and the boundary $\partial\Omega$ smooth enough, we define

$$H^1(\Omega) := W^{1,2}$$
$$H^1_0(\Omega) := \{f \in H^1(\Omega) : u|_{\partial\Omega} = 0\}$$
$$H^2(\Omega) := W^{2,2}$$

## A.2. Norms

$$\|f\|_{H^1} := \left( \int_\Omega \left( \mid Df \mid^2 + \mid f \mid^2 \right) \right)^{\frac{1}{2}}$$

For a bounded function $\omega$ on $Q'$ we set

$$\mid \omega \mid_0 := \sup_{(t,y) \in Q'} \mid \omega(t,y) \mid$$

And for $\delta \in (0,1]$ we set

$$[\omega]_\delta := \sup_{(t,x) \neq (s,y)} \frac{\mid \omega(t,x) - \omega(s,y) \mid}{(\mid x - y \mid + \mid t - s \mid^{1/2})^\delta}$$

And

$$\mid \omega \mid_\delta = \mid \omega \mid_0 + [\omega]_\delta$$

## A.3. Inner products

For $v, w \in L^2(\Omega)$ with a measure $\zeta$ we define

$$\langle v, w \rangle_{L^2} := \int_\Omega vw d\zeta$$

For $v, w \in L^2(\partial\Omega)$ with a measure $\zeta$ we define

$$\langle v, w \rangle := \int_{\partial\Omega} vw d\zeta$$

For $v, w \in H^1(\Omega)$ we define

$$\langle v, w \rangle_{H^1} := \langle v, w \rangle_{L^2} + \langle Dv, Dw \rangle_{L^2}$$

And for $v, w \in H^2(\Omega)$ we define

$$\langle v, w \rangle_{H^2} := \langle v, w \rangle_{H^1} + \langle D^2 v, D^2 w \rangle_{L^2} = \langle v, w \rangle_{L^2} + \langle Dv, Dw \rangle_{L^2} + \langle D^2 v, D^2 w \rangle_{L^2}$$

$H^1$ and $H^2$ become Hilbert spaces with these inner products.

## A.4. Vectors, Matrices and Functions

We use $S_d(\mathbb{R})$ for the notation of symmetric $d \times d$ matrices in $\mathbb{R}$.
$B(\bar{\Omega})$ is the set of bounded functions with the domain $\bar{\Omega}$.

# Bibliography

[1] BARLES, G. ; SOUGANIDIS, P.: Convergence of approximation schemes for fully nonlinear second-order equations. (1991)

[2] BONNANS, J. F. ; OTTENWAELTER, E. ; ZIDANI, H.: A fast algorithm for the two dimensional HJB equation of stochastic control. 25 (2004)

[3] BONNANS, J. F. ; ZIDANI, H.: Consistency of generalized finite difference schemes for the stochastic HJB equation. (2003)

[4] CAFFARELLI, L. A. ; SILVESTRE, L.: On the EvansKrylov Theorem. (2009)

[5] CRANDALL, M. G. ; EVANS, L. C. ; P.L.LIONS: Some Properties of Viscosity Solutions of Hamilton-Jacobi Equations. (1984)

[6] CRANDALL, Michael G. ; ISHII, Hitoshi ; LIONS, Pierre-Louis: User's guide to viscosity solutions of second order partial differential equations. (1992)

[7] CRANDALL, Michael G. ; LIONS, Pierre-Louis: Viscosity solutions of Hamilton-Jacobi equations. (1983)

[8] DEBRABANT, K. ; JAKOBSEN, E. R.: Semi-Lagrangian schemes for linear and fully nonlinear diffusion equations. (2013)

[9] EVANS, Lawrence C.: Partial differential equations. 19 (1998)

[10] FENG, X. ; NEILAN, M.: Vanishing Moment Method and Moment Solutions for Fully Nonlinear Second Order Partial Differential Equations. (2007)

[11] FLEMING, Wendell H. ; SONER, H. M.: Controlled Markov processes and viscosity solutions. 25 (2006)

[12] JENSEN, Max ; SMEARS, Iain: On the convergence of finite element methods for Hamilton-Jacobi-Bellman Equations. (2011)

[13] OBERMAN, ADAM M.: Convergent Difference Schemes For Degenerate Elliptic And Parabolic Equations: HamiltonJacobi Equations And Free Boundary Problems. In: *SIAM J. NU-MER.ANAL* 44 (2006)

[14] PEYRL, HELFRIED ; HERZOG, FLORIAN ; P.GEERING, HANS: Numerical Solution of the Hamilton-Jacobi-Bellman Equation for Stochastic Optimal Control Problems. (2005)

[15] SMEARS, Iain: Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients (Doctoral dissertation). (2015)