FEM for degenerate isotropic Hamilton–Jacobi–Bellman Equations

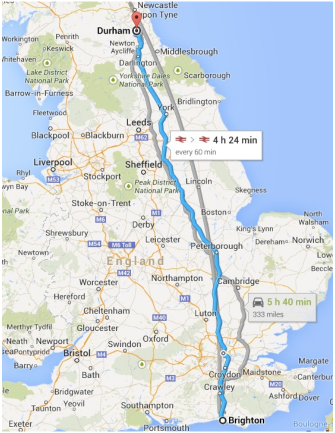
Max Jensen

University of Sussex

with Gabriele Braconi and Iain Smears

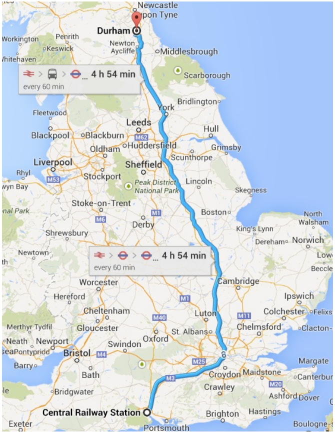
An optimal control problem

- We have **paths**, e.g. all roads from Brighton B to Durham A.
- We have **controls**, e.g. steering wheel + accelerator + brake.
- We have a **cost functional** on the set of paths, e.g. the driving time or petrol cost.
- We denote the minimal cost to get from B to A by $v(B)$.
- We can assign to every C on the map a minimal cost.
- This defines the **value function** v.



An optimal control problem

- We have **paths**, e.g. all roads from Brighton B to Durham A.
- We have **controls**, e.g. steering wheel + accelerator + brake.
- We have a **cost functional** on the set of paths, e.g. the driving time or petrol cost.
- We denote the minimal cost to get from B to A by $v(B)$.
- We can assign to every C on the map a minimal cost.
- This defines the **value function** v.



- Now suppose that the path choice depends on the control through an Itô process.
- If v is smooth it solves the Hamilton–Jacobi–Bellman equation.
- Set

$$\mathcal{H}w := \sup_{\alpha}(\underbrace{-a^{\alpha}\,\Delta w + b^{\alpha}\cdot\nabla w}_{\substack{\text{linear, 2nd order} \\ \text{non-divergence form}}} - r^{\alpha}).$$

Hamilton-Jacobi-Bellman problem

Find the (right kind of) solution of

$$-\partial_t v + \mathcal{H}v = 0$$

with final conditions $v(T,\cdot) = \Psi$ and homogeneous Dirichlet BCs.

- Now suppose that the path choice depends on the control through an Itô process.
- If v is smooth it solves the Hamilton–Jacobi–Bellman equation.
- Set

$$\mathcal{H}w := \sup_{\alpha}(\underbrace{-a^{\alpha}\,\Delta w + b^{\alpha}\cdot\nabla w}_{\substack{\text{linear, 2nd order}\\ \text{non-divergence form}}} - r^{\alpha}).$$

Hamilton-Jacobi-Bellman problem

Find the (right kind of) solution of

$$-\partial_t v + \mathcal{H}v = 0$$

with final conditions $v(T,\cdot) = \Psi$ and homogeneous Dirichlet BCs.

Model Problem: Uncertain volatility

- From (Avellaneda, Levy, Paras; 1995).
- Consider stock prices as geometric Brownian motions

  $$\mathrm{d}S = f\,S\,\mathrm{d}t + \sigma\,S\,\mathrm{d}w(t) \qquad (w \text{ Brownian motion})$$

  An option is a financial product whose value

  $$\mathcal{J}(t, S(t), \sigma)$$

  depends on the stock price $S(t)$ and the volatility $\sigma : [0, T] \to \mathbb{R}_+$. This value is the cost of hedging against the risk associated with the option.
- Suppose the volatility is uncertain, however, guaranteed to be

  $$\sigma : [0, T] \to [\sigma_*, \sigma^*].$$

  To hedge against the uncertain volatility costs at time t and price S

  $$V(t, S) = \sup_{\sigma} \mathcal{J}(t, S, \sigma).$$

Model Problem: Uncertain volatility

- From (Avellaneda, Levy, Paras; 1995).
- Consider stock prices as geometric Brownian motions

  $$\mathrm{d}S = f\,S\,\mathrm{d}t + \sigma\,S\,\mathrm{d}w(t) \qquad (w \text{ Brownian motion})$$

  An option is a financial product whose value

  $$\mathcal{J}(t, S(t), \sigma)$$

  depends on the stock price $S(t)$ and the volatility $\sigma : [0, T] \to \mathbb{R}_+$. This value is the cost of hedging against the risk associated with the option.
- Suppose the volatility is uncertain, however, guaranteed to be

  $$\sigma : [0, T] \to [\sigma_*, \sigma^*].$$

  To hedge against the uncertain volatility costs at time t and price S

  $$V(t, S) = \sup_{\sigma} \mathcal{J}(t, S, \sigma).$$

Model Problem: Uncertain volatility

- One can derive directly or in the stochastic HJB framework that

$$0 = \frac{\partial}{\partial_t} V + \min_{\sigma} \left(\frac{1}{2} \sigma^2 S^2 \frac{\partial^2}{\partial_{SS}} V + r\, S \frac{\partial}{\partial_S} V - r\, V \right)$$

$$= \frac{\partial}{\partial_t} V + \frac{1}{2} \varsigma \left(\frac{\partial^2}{\partial_{SS}} V \right)^2 S^2 \frac{\partial^2}{\partial_{SS}} V + r\, S \frac{\partial}{\partial_S} V - r\, V$$

  with

$$\varsigma \left(\frac{\partial^2}{\partial_{SS}} V \right) = \begin{cases} \sigma^* : & \partial^2_{SS} V < 0, \\ \sigma_* : & \partial^2_{SS} V \geq 0. \end{cases}$$

- Three features we want to address:
  - **fully-nonlinear** (= nonlinear in highest derivative)
  - **discontinuous** in nonlinearity ς
  - **degenerate** at $S = 0$

Model Problem: Uncertain volatility

- One can derive directly or in the stochastic HJB framework that

$$0 = \frac{\partial}{\partial_t} V + \min_{\sigma} \left(\tfrac{1}{2} \sigma^2 S^2 \frac{\partial^2}{\partial_{SS}} V + r\, S \frac{\partial}{\partial_S} V - r\, V \right)$$

$$= \frac{\partial}{\partial_t} V + \tfrac{1}{2} \varsigma \left(\frac{\partial^2}{\partial_{SS}} V \right)^2 S^2 \frac{\partial^2}{\partial_{SS}} V + r\, S \frac{\partial}{\partial_S} V - r\, V$$

with

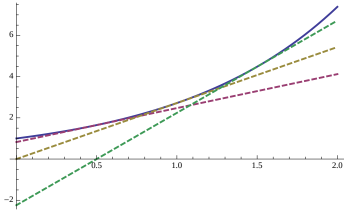
$$\varsigma \left(\frac{\partial^2}{\partial_{SS}} V \right) = \begin{cases} \sigma^* : & \partial^2_{SS} V < 0, \\ \sigma_* : & \partial^2_{SS} V \geq 0. \end{cases}$$

- Three features we want to address:
  - **fully-nonlinear** (= nonlinear in highest derivative)
  - **discontinuous** in nonlinearity ς
  - **degenerate** at $S = 0$

There is a wide range of Optimal Control applications, e.g. in Finance:

- Merton portfolio problem
- Irreversible and reversible abandonment and investment
- Valuation of natural resources
- ...

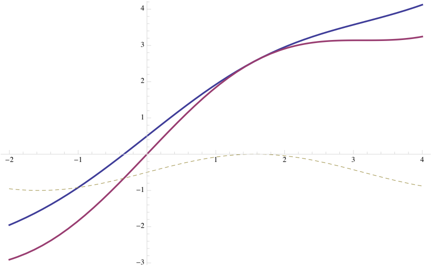
Moreover, many nonlinear equations with a convex structure can be modelled in the HJB setting (often requiring **anisotropic** diffusions):



- Monge-Ampère equation
- Pucci's equations

Viscosity solutions

- **Weak solutions of these problems are usually not unique.**



- Suppose v was smooth. If $v - \psi$ has a maximum, then there

$$\nabla v = \nabla \psi, \quad D^2 v \leq D^2 \psi \qquad (D^2 \text{ Hessian})$$

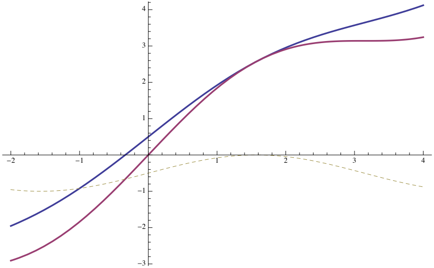
  where

$$A \leq B :\Leftrightarrow B - A \text{ is positive semi-definite.}$$

- Instead of plugging the derivatives of v into HJB, we want to work with the derivatives of ψ. This is the trick!

Viscosity solutions

- **Weak solutions of these problems are usually not unique.**



- Suppose v was smooth. If $v - \psi$ has a maximum, then there

$$\nabla v = \nabla \psi, \quad D^2 v \leq D^2 \psi \qquad (D^2 \text{ Hessian})$$

where

$$A \leq B :\Leftrightarrow B - A \text{ is positive semi-definite.}$$

- Instead of plugging the derivatives of v into HJB, we want to work with the derivatives of ψ. This is the trick!

The trick works because of the following:

If $A \leq B$ then with $q_0 \in \mathbb{R}, \vec{q} \in \mathbb{R}^n$

$$-q_0 + \mathcal{H}(x, t, \vec{q}, A) \geq -q_0 + \mathcal{H}(x, t, \vec{q}, B).$$

Theorem

- *v is C^2 solution of HJB. Then for all smooth ψ*

$$-\partial_t \psi(t,x) + \mathcal{H}(x, t, \nabla\psi(t,x), D^2\psi(t,x)) \leq 0$$

at every (t,x) which maximises $v - \psi$ with $v(t,x) = \psi(t,x)$.

Viscosity Solution (Crandall, Lions)

- $v \in C^0$ is a subsolution of HJB if for all smooth ψ

$$-\partial_t \psi(t,x) + \mathcal{H}(x, t, \nabla\psi(t,x), D^2\psi(t,x)) \leq 0$$

at every (t,x) which maximises $v - \psi$ with $v(t,x) = \psi(t,x)$.
- Supersolution similar. Subsolution + supersolution =: solution.

The trick works because of the following:

If $A \leq B$ then with $q_0 \in \mathbb{R}, \vec{q} \in \mathbb{R}^n$

$$-q_0 + \mathcal{H}(x, t, \vec{q}, A) \geq -q_0 + \mathcal{H}(x, t, \vec{q}, B).$$

Theorem

- *v is C^2 solution of HJB. Then for all smooth ψ*

$$-\partial_t \psi(t, x) + \mathcal{H}(x, t, \nabla \psi(t, x), D^2 \psi(t, x)) \leq 0$$

at every (t, x) which maximises $v - \psi$ with $v(t, x) = \psi(t, x)$.

Viscosity Solution (Crandall, Lions)

- $v \in C^0$ is a subsolution of HJB if for all smooth ψ

$$-\partial_t \psi(t, x) + \mathcal{H}(x, t, \nabla \psi(t, x), D^2 \psi(t, x)) \leq 0$$

at every (t, x) which maximises $v - \psi$ with $v(t, x) = \psi(t, x)$.

- Supersolution similar. Subsolution + supersolution =: solution.

Under Reasonable Assumptions ...

Theorem (Comparison principle)

Let w^* be a viscosity subsolution and w_* be a viscosity supersolution then

$$\sup_{\Omega_T}(w^* - w_*) = \sup_{\partial_{\mathrm{p}}\Omega_T}(w^* - w_*).$$

Here $\partial_{\mathrm{p}}\Omega_T$ parabolic boundary.

Corollary (Uniqueness)

The HJB viscosity solution is unique and equal to the value function.

Theorem (v viscosity solution $\rightsquigarrow$ existence)

The value function of the optimal control problem is a viscosity solution of the HJB equation.

For proofs see for instance (Fleming,Soner; Chapter V).

Summary of the PDE Theory . . .

- The value function is defined in terms of the original optimal control problem.
- We now have a very nice theory of the HJB equation, which gives us
  - in a general context the value function as solution of the HJB eqn
  - and no other wrong solution
  - without need to solve path integrals of the underlying optimal control problem.
- It is not really a differential equation anymore (in a narrow sense).
- Instead it is a problem posed on the set of continuous functions using monotonicity properties of the equation.
- **If we use continuous functions, what about the Greeks (finance speak for partial derivatives)?**
- The research I present today addresses this question in the context of computational methods.

Summary of the PDE Theory . . .

- The value function is defined in terms of the original optimal control problem.
- We now have a very nice theory of the HJB equation, which gives us
  - in a general context the value function as solution of the HJB eqn
  - and no other wrong solution
  - without need to solve path integrals of the underlying optimal control problem.
- It is not really a differential equation anymore (in a narrow sense).
- Instead it is a problem posed on the set of continuous functions using monotonicity properties of the equation.
- **If we use continuous functions, what about the Greeks (finance speak for partial derivatives)?**
- The research I present today addresses this question in the context of computational methods.

Summary of the PDE Theory . . .

- The value function is defined in terms of the original optimal control problem.
- We now have a very nice theory of the HJB equation, which gives us
  - in a general context the value function as solution of the HJB eqn
  - and no other wrong solution
  - without need to solve path integrals of the underlying optimal control problem.
- It is not really a differential equation anymore (in a narrow sense).
- Instead it is a problem posed on the set of continuous functions using monotonicity properties of the equation.
- **If we use continuous functions, what about the Greeks (finance speak for partial derivatives)?**
- The research I present today addresses this question in the context of computational methods.

Task:

- Discretise HJB equations with a Finite Element Method.

Main Results:

- First proof of **uniform convergence** of Galerkin approximations with
  - non-smooth viscosity solutions (e.g. no classical solutions),
  - non-smooth HJB operators (no linearisation of $\mathcal{H}$).

  (To my knowledge this has been an open problem for a long time.)
- Novel variational argument for **gradient convergence**.
- Unstructured meshes permitted and **Newton solvers** globally convergent from above.
- HJB operators may be **degenerate**, but assumed **isotropic**.

Notation:

- y^ℓ is ℓth node
- s^k is the kth time step
- h time step size
- ϕ^ℓ is the hat function at y^ℓ with volume $\|\phi^\ell\|_{L^1(\Omega)} = 1$

Add artificial diffusion:

$$\bar{a}_h^\alpha \geq a^\alpha$$

Numerical scheme/framework

- Let $v_h(T, \cdot) = \Psi$.
- Find $v_h(s^k, \cdot)$ such that

$$-\frac{v_h(s^{k+1}, y^\ell) - v_h(s^k, y^\ell)}{h} + \sup_\alpha \Big(\bar{a}_h^\alpha(y^\ell) \underbrace{\langle \nabla v_h, \nabla \phi^\ell \rangle}_{\substack{\text{conceptually} \\ -\langle \Delta v_h, \phi_\ell \rangle \approx -\Delta v_h(y^\ell)}} + \langle b^\alpha \cdot \nabla v_h - r^\alpha, \phi^\ell \rangle \Big) = 0.$$

Consistency

As $h \to 0$

$$\|a^\alpha - \bar{a}_h^\alpha\|_{L^\infty(\Omega)} \to 0$$

Weak discrete maximum principle (wDMP)

For fixed α, discrete linear operators have M-matrix property.

Elliptic projection converging in W_1^∞

Finite element solutions to the Laplace equation converge in $W_1^\infty(\Omega)$

E.g. see recent work by Demlow, Schatz, Wahlbin, etc.

Selected literature

- **Finite element methods**
  - (Smears,Süli; 2013) DG method, Cordes theory, *hp* convergence proof under H^2 regularity.
  - (Lakkis, Pryer; 2011) similar to our method, also anisotropic-diffusion numerical experiments; a convergence result by Neilan for 2D Monge-Ampère for *smooth* operator + solution
  - (Feng, Neilan; 2011) primarily for Monge-Ampère equations, special case of HJB, biharmonic regularisation '$\epsilon\Delta^2 v + Fv = 0$', convergence for *smooth* operator + solution, see also Brenner et al.
  - (Böhmer; 2008) approximation with smooth functions (C^1 approximation space), convergence for *smooth* operator + solution
  - (Cortey-Dumont; 1987) $\rightsquigarrow$ (Boulbrachene; 2001, 2004) QVI approach: convergence rates, but expensive!
- **Finite difference methods**
  - method design: (Kushner; 1977), (Lions, Mercier; 1980), (Lions, Souganidis; 1995), (Froese, Oberman; 2012)
  - convergence: (Barles, Souganidis; 1991)
  - convergence rates: (Barles, Jakobsen; 2002, 2005, 2007), (Caffarelli, 2008), (Krylov; 2005), (Dong, Krylov; 2007)

Current numerical methods

- My personal view is that current methods struggle to combine all of the following features in a single framework:
  - **Uniqueness**: monotonicity $\leftrightarrow$ smoothness
  - **Anisotropy**
  - **Consistency** (of second derivative)
  - **Efficiency** (large stencils, ...)
  - Geometry
- We don't do anisotropy.
- Other methods balance the challenge in other ways.

- Set

$$v^*(t,x) = \underbrace{\limsup_{\substack{(s_h^k, y_h^\ell) \to (t,x) \\ h \to 0}} v_h(s_h^k, y_h^\ell)}_{\text{upper semi-continuous envelope}} \quad , \qquad v_*(t,x) = \underbrace{\liminf_{\substack{(s_h^k, y_h^\ell) \to (t,x) \\ h \to 0}} v_h(s_h^k, y_h^\ell)}_{\text{lower semi-continuous envelope}} \; .$$

Theorem

v^* is a HJB viscosity subsolution and v_* is a HJB viscosity supersolution.

- The proof is based on (Barles, Souganidis).
- However, finite element methods not pointwise consistent.

$$\langle \nabla \mathcal{I} w, \nabla \phi_i \rangle = -\tfrac{3}{2} \Delta w(x) + \mathcal{O}\big(\Delta x_i^2\big)$$

- But FEM for HJB and Laplace are inconsistent in the same way: $\rightsquigarrow$ use cancellation of consistency error with elliptic projection.
- The above theorem and a comparison principle give uniform convergence.

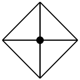
- Set

$$v^*(t,x) = \underbrace{\limsup_{\substack{(s_h^k, y_h^\ell)\to(t,x)\\ h\to 0}} v_h(s_h^k, y_h^\ell)}_{\text{upper semi-continuous envelope}} \quad , \qquad v_*(t,x) = \underbrace{\liminf_{\substack{(s_h^k, y_h^\ell)\to(t,x)\\ h\to 0}} v_h(s_h^k, y_h^\ell)}_{\text{lower semi-continuous envelope}} \; .$$

Theorem

v^* is a HJB viscosity subsolution and v_* is a HJB viscosity supersolution.

- The proof is based on (Barles, Souganidis).
- However, finite element methods not pointwise consistent.



$$\langle \nabla \mathcal{I} w, \nabla \phi_i \rangle = -\tfrac{3}{2} \Delta w(x) + \mathcal{O}\big(\Delta x_i^2\big)$$

- But FEM for HJB and Laplace are inconsistent in the same way:
  $\rightsquigarrow$ use cancellation of consistency error with elliptic projection.
- The above theorem and a comparison principle give uniform convergence.

Convergence of Derivatives

- The partial derivatives of the value function contain crucial information

  Uncertain volatility model: $\partial V/\partial S$ amount of stock in hedging portfolio.
- It has been our aim to design a scheme which naturally provides control of gradient.

Variational bound

Since $v \geq 0$,

$$\partial_t v + \sup_{\alpha}(L^\alpha v - r^\alpha) = 0 \implies \partial_t v + L^\alpha v \leq r^\alpha \quad \text{for any } \alpha$$
$$\implies \langle \partial_t v, v\rangle + \langle L^\alpha v, v\rangle \leq \langle r^\alpha, v\rangle$$

- Variational structure not good enough to identify 'correct solution'. However, once convergence to viscosity solution guaranteed, variational structure can be very useful!

Convergence of Derivatives

- The partial derivatives of the value function contain crucial information

  Uncertain volatility model: $\partial V / \partial S$ amount of stock in hedging portfolio.
- It has been our aim to design a scheme which naturally provides control of gradient.

Variational bound

Since $v \geq 0$,

$$\partial_t v + \sup_\alpha (L^\alpha v - r^\alpha) = 0 \implies \partial_t v + L^\alpha v \leq r^\alpha \quad \text{for any } \alpha$$

$$\implies \langle \partial_t v, v \rangle + \langle L^\alpha v, v \rangle \leq \langle r^\alpha, v \rangle$$

- Variational structure not good enough to identify 'correct solution'. However, once convergence to viscosity solution guaranteed, variational structure can be very useful!

- While HJB operator not smooth, in above applications there is at least one semi-definite L^α with smooth coefficients!
- This gives (in spirit) a ‘discrete variational bound’

  $$\text{“}\|w_h\|_{L^2(H^1_\gamma)} \lesssim \langle \partial_t w_h, w_h\rangle + \langle L_h^\alpha w_h, w_h\rangle \le \langle r^\alpha, w_h\rangle\text{”} \quad (w_h = Q_h v - v_h)$$

  if $Q_h v$ is
  - approximates v in $H^1((0,T)\times\Omega)$,
  - non-negative,
  - satisfies boundary conditions,
  - is bounded from above by v_h.

  H^1_γ is a weighted Sobolev space.
- Precisely

  $$\begin{aligned}\|w_h\|_{L^2(H^1_\gamma)} &\lesssim \textstyle\sum_k (\langle\langle (hL_h^\alpha + \mathrm{Id})w_h^k, \cdot) - w_h^{k+1}, w_h^k\rangle\rangle) + \frac{1}{2}\langle\langle w_h(T,\cdot), w_h(T,\cdot)\rangle\rangle \\ &\lesssim \textstyle\sum_k \langle\langle h r^\alpha, w_h^k\rangle\rangle - \sum_k \langle\langle Q_h v(s^{k+1},\cdot) + (hL_h^\alpha + \mathrm{Id}) Q_h v(s^k,\cdot), w_h^k\rangle\rangle \\ &\quad + \tfrac{1}{2}\langle\langle w_h^{T/h}, w_h^{T/h}\rangle\rangle\end{aligned}$$

  using for non-constant coefficients a super-approximation result **(Demlow, Guzmán, Schatz; 2011)** and a ‘quadrature’ L^2 scalar product $\langle\langle\cdot,\cdot\rangle\rangle$.

- While HJB operator not smooth, in above applications there is at least one semi-definite L^α with smooth coefficients!
- This gives (in spirit) a 'discrete variational bound'

$$\text{``}\|w_h\|_{L^2(H^1_\gamma)} \lesssim \langle \partial_t w_h, w_h \rangle + \langle L^\alpha_h w_h, w_h \rangle \leq \langle r^\alpha, w_h \rangle\text{''} \quad (w_h = Q_h v - v_h)$$

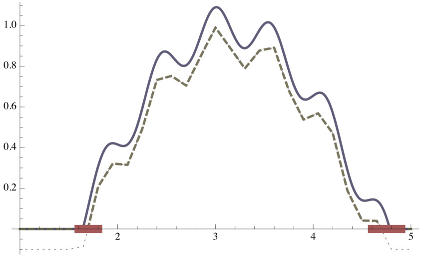
  if $Q_h v$ is
  - approximates v in $H^1((0,T)\times\Omega)$,
  - non-negative,
  - satisfies boundary conditions,
  - is bounded from above by v_h.

  H^1_γ is a weighted Sobolev space.
- Precisely

$$\begin{aligned}\|w_h\|_{L^2(H^1_\gamma)} &\lesssim \textstyle\sum_k (\langle\langle (hL^\alpha_h + \mathrm{Id}) w^k_h, \cdot) - w^{k+1}_h, w^k_h \rangle\rangle) + \frac{1}{2}\langle\langle w_h(T,\cdot), w_h(T,\cdot)\rangle\rangle \\ &\lesssim \textstyle\sum_k \langle\langle h r^\alpha, w^k_h \rangle\rangle - \sum_k \langle\langle Q_h v(s^{k+1},\cdot) + (hL^\alpha_h + \mathrm{Id}) Q_h v(s^k,\cdot), w^k_h \rangle\rangle \\ &\quad + \tfrac{1}{2}\langle\langle w_h^{T/h}, w_h^{T/h} \rangle\rangle\end{aligned}$$

  using for non-constant coefficients a super-approximation result **(Demlow, Guzmán, Schatz; 2011)** and a 'quadrature' L^2 scalar product $\langle\langle \cdot,\cdot \rangle\rangle$.

- Here $Q_h : \ w \mapsto \mathcal{I}_h \max\{w - \|v - v_h\|_{L^\infty}, 0\}$.



- A crucial step in proof is to bound projection error in red region.

Theorem

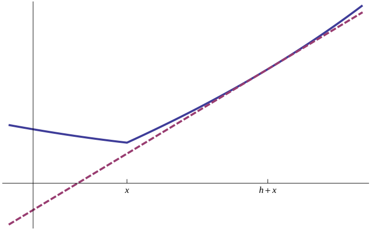
If

- *value function belongs to* $W^{1,d+1+\varepsilon}((0,T)\times\Omega)$, $\varepsilon > 0$,
- *for an* α *the* L^α *is semi-definite with diffusion coefficient* $a^\alpha \in W^{2,\infty}$,
- $\mathcal{O}(h)$ *stabilisation with artificial diffusion,*

then $\|v - v_h\|_{L^2(H^1_\gamma)} \to 0$ *with* $\gamma = \sqrt{a^\alpha}$.

Semi-smooth Newton

- Semi-smooth Newton methods use a weakened concept of differentiability:



- Analysis of Newton methods for discrete HJB equations: (Howard; 1960), (Lions, Mercier; 1980), (Bokanowski, Maroso, Zidani; 2009), (Lakkis, Pryer; 2011) . . .

Theorem

Our scheme has non-negative uniformly bounded unique solutions. Semi-smooth Newton methods converge 'globally from above', monotone and superlinearly.

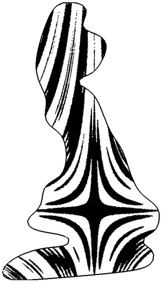
Numerical Experiment

- We examine the equation

$$-v_t + \sup_{\alpha \in [\alpha_0, \alpha_1]} \{-\alpha \Delta v\} + |\nabla v| = f,$$

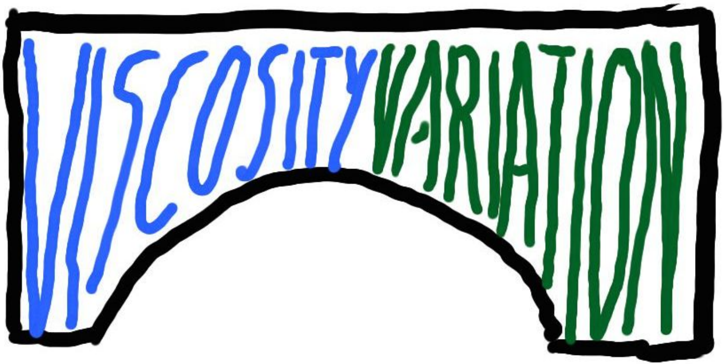
- DoF v. Newton iterations

DoF	aver. no. Newton it.
674	3
2858	3.67
11759	4.04
47693	4.22
192089	4.86



Remark

For smooth problems we observe optimal rates in L^2, L^∞ and H^1.



Thank you for the attention!

Proof – HJB sub- and supersolutions.

- Suppose $v^* - \psi$ has (strict) maximum at (t, x).
- Let P_h be the elliptic projection:

$$\langle \nabla P_h \psi, \nabla \phi_h^\ell \rangle = \langle \nabla \psi, \nabla \phi_h^\ell \rangle. \tag{1}$$

- Show there are nodes (s_h^k, y_h^k)
  - converging to (t, x)
  - maximise $v_h - P_h \psi$ locally
- Then

$$0 \overset{(a)}{=} F_h(v_h)(s_h^k, y_h^k) \overset{(b)}{\geq} F_h(P_h \psi)(s_h^k, y_h^k) + \underset{\text{perturbation}}{\text{small}} \overset{(c)}{\to} F(\psi)(t, x).$$

  using
  (a) definition of scheme
  (b) monotonicity property
  (c) orthogonality (1) for elliptic term, otherwise approximation bounds of elliptic projection

$$a^\alpha(y^\ell) \langle \nabla P_h \psi, \nabla \phi_h^\ell \rangle = a^\alpha(y^\ell) \langle \nabla \psi, \nabla \phi_h^\ell \rangle$$

- Corresponding argument if $v_* - \psi$ minimum at (t, x).