Week 8. Existence: Vanishing Viscosity, Stability, and the Bernstein Method

Tran: §1.3 (Lemma 1.12, Theorem 1.13), §1.4 (Lemma 1.15, Theorems 1.16–1.17), §1.7 (Theorem 1.25, Remark 1.27)

1. Overview

In Week 6, we introduced viscosity solutions through smooth test functions and the vanishing viscosity argument. In Week 7, we studied generalized differentials and proved directly from the Dynamic Programming Principle that the discounted control value function solves

$$\lambda u + H(x, Du) = 0$$

in the viscosity sense. That proof identified the solution associated with a given control problem.

This week we study a complementary question: how can we construct a viscosity solution by solving regularized PDEs, and why does the equation persist in the limit? We work mainly with the time-dependent Cauchy problem. The compactness and stability arguments do not require a control representation of the Hamiltonian.

The central idea is the **vanishing viscosity method**. For each fixed positive viscosity, we solve a uniformly parabolic equation. Estimates independent of the viscosity parameter then allow us to extract a convergent subsequence. Smooth touching functions allow us to pass the PDE inequalities to its limit.

We consider the Cauchy problem

$$u_t + H(x, Du) = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \qquad u(x, 0) = u_0(x) \quad \text{on } \mathbb{R}^n.$$

Existence and stability are the goals here. Uniqueness requires a comparison principle; the first-order stationary and Cauchy comparison principles are treated in Weeks 9 and 10, respectively.

2. Learning Objectives

By the end of this lecture, students should be able to:

1. explain why a small second-order term regularizes a first-order Hamilton–Jacobi equation;
2. use stability of strict extremal points to pass viscosity inequalities to a limit;
3. distinguish compactness estimates from stability arguments in the vanishing viscosity method;
4. verify consistency of viscosity solutions at differentiability points;
5. derive the differential equation for $\frac{1}{2}|Du^\varepsilon|^2$ in the Bernstein method;
6. explain how the structural condition in Tran (1.27) yields a gradient bound independent of ε.

Notation and function spaces

The operators D, D^2, and Δ act on the spatial variable x; u_t denotes the time derivative. For a Hamiltonian $H(x, p)$, $D_x H$ means differentiation in x with p held fixed, and $D_p H$ means differentiation in p. We use the matrix norm

$$|D^2 w|^2 = \sum_{i,j=1}^{n} |w_{x_i x_j}|^2.$$

Recall that BUC denotes bounded uniformly continuous functions and Lip denotes Lipschitz continuous functions. On the infinite time interval, we use the viscosity definition from Week 6 on each finite strip $\mathbb{R}^n \times [0, T]$. A solution may grow as $t \to \infty$ while remaining bounded on every such strip.

The notation $u^\varepsilon \to u$ **locally uniformly** means that

$$\sup_{z \in K} |u^\varepsilon(z) - u(z)| \to 0$$

for every compact set K in the stated domain. When the initial condition is involved, the domain includes $t = 0$. We assume that H is continuous throughout; the stronger differentiability assumptions needed for the Bernstein method will be stated in §7.

For touching test functions, we use the notation from Week 6:

- φ is an **upper test function**, touching u from above: $u - \varphi$ has a local maximum, and the subsolution inequality is tested.
- ψ is a **lower test function**, touching u from below: $u - \psi$ has a local minimum, and the supersolution inequality is tested.

3. Stability of Extremal Points (Tran Lemma 1.12, p.13)

Let u^ε and u be continuous, with $u^\varepsilon \to u$ locally uniformly, and let $\varphi \in C^1$. Assume that $u - \varphi$ has a **strict local maximum** at an interior point (x_0, t_0), where $t_0 > 0$. Then, for all sufficiently small ε, the function

$$u^\varepsilon - \varphi$$

has a local maximum at a point $(x_\varepsilon, t_\varepsilon)$ that can be chosen so that

$$(x_\varepsilon, t_\varepsilon) \to (x_0, t_0).$$

The same statement holds for strict local minima.

This lemma is the basic mechanism for passing viscosity inequalities to the limit. A viscosity subsolution or supersolution is tested by smooth functions touching from above or below. The lemma says that if a touching point exists for the limit u, then nearby touching points also exist for the approximating functions u^ε. Therefore, one can apply the PDE satisfied by u^ε at those nearby points and then pass to the limit.

Proof

Write $z = (x, t)$, $z_0 = (x_0, t_0)$, and $w = u - \varphi$. After adding a constant to φ, assume $w(z_0) = 0$. Choose $r > 0$ so that the closed ball $\overline{B}_r(z_0)$ lies in the interior of the space-time domain and

$$w(z) < 0 \qquad \text{for } z \in \overline{B}_r(z_0) \setminus \{z_0\}.$$

Continuity and compactness of the boundary give a positive gap

$$\eta = -\max_{\partial B_r(z_0)} w > 0.$$

For small ε, local uniform convergence gives

$$\sup_{\overline{B}_r(z_0)} |u^\varepsilon - u| < \frac{\eta}{3}.$$

Thus $(u^\varepsilon - \varphi)(z_0) > -\eta/3$, whereas $u^\varepsilon - \varphi < -2\eta/3$ on the boundary. A maximizer z_ε on the closed ball must therefore be interior.

If a subsequence of these maximizers converges to z_*, the maximizing inequality and uniform convergence imply

$$w(z_*) \geq w(z_0) = 0.$$

Strictness forces $z_* = z_0$. Hence $z_\varepsilon \to z_0$, and also $u^\varepsilon(z_\varepsilon) \to u(z_0)$. For a lower test function ψ, the minimum case follows by applying the same argument to $-u^\varepsilon$, $-u$, and $-\psi$.

If a normalized touching function is needed at z_ε, replace φ by

$$\varphi_\varepsilon(z) = \varphi(z) + u^\varepsilon(z_\varepsilon) - \varphi(z_\varepsilon).$$

This changes only its value, not its derivatives. The touching point generally moves with ε; it need not remain equal to z_0.

4. Existence by the Vanishing Viscosity Method (Tran Theorem 1.13, p.13)

Tran states Theorem 1.13 for $H = H(p)$. The same limiting argument applies to a fixed continuous Hamiltonian $H = H(x, p)$: if $(x_\varepsilon, p_\varepsilon) \to (x, p)$, then $H(x_\varepsilon, p_\varepsilon) \to H(x, p)$. We use this slightly extended form below.

Assume

$$u_0 \in \mathrm{BUC}(\mathbb{R}^n) \cap \mathrm{Lip}(\mathbb{R}^n).$$

For each sufficiently small $\varepsilon > 0$, suppose that a classical solution u^ε, smooth for $t > 0$ and continuous up to $t = 0$, exists for

$$u^\varepsilon_t + H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon, \qquad u^\varepsilon(x, 0) = u_0(x).$$

Assume further that there exists a constant $C > 0$, independent of these sufficiently small ε, such that

$$\|u^\varepsilon_t\|_{L^\infty(\mathbb{R}^n \times (0,\infty))} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n \times (0,\infty))} \le C.$$

Then there is a sequence $\varepsilon_j \downarrow 0$ for which u^{ε_j} converges locally uniformly on $\mathbb{R}^n \times [0, \infty)$ to a function u. The limit belongs to $\mathrm{BUC}(\mathbb{R}^n \times [0, T])$ for every finite T and is a viscosity solution of

$$u_t + H(x, Du) = 0, \qquad u(x, 0) = u_0.$$

Why this theorem is important

This theorem gives a standard existence mechanism in nonlinear PDE:

$$\text{uniform estimates} + \text{compactness} + \text{stability of viscosity inequalities} \Longrightarrow \text{existence}.$$

The theorem assumes both the existence of the classical approximations and their uniform estimates. It does not establish them for an arbitrary continuous H. Sections 7–9 give sufficient assumptions and estimates. Since the argument takes $\varepsilon \downarrow 0$, it is enough to have these estimates for all sufficiently small positive viscosities.

Step 1. Compactness and the initial condition

For $0 \le s, t \le T$, the derivative bounds and continuity at the initial time give

$$|u^\varepsilon(x, t)| \le \|u_0\|_\infty + CT,$$

$$|u^\varepsilon(x, t) - u^\varepsilon(y, s)| \le C(|x - y| + |t - s|).$$

Thus the approximations are uniformly bounded and equicontinuous on every compact subset of space-time. Starting with any sequence of viscosities tending to zero, the Arzelà–Ascoli theorem gives a uniformly convergent subsequence on each compact set. Applying it successively on $\overline{B}_m(0) \times [0, m]$, $m = 1, 2, ...$, and taking a diagonal subsequence gives local uniform convergence on the entire domain.

The displayed bounds pass to the limit, proving the claimed boundedness and uniform continuity on finite strips. Since the convergence includes $t = 0$,

$$u(x, 0) = \lim_{j \to \infty} u^{\varepsilon_j}(x, 0) = u_0(x).$$

This step is where the uniform estimates are used. Once a locally uniform limit is known to exist, the viscosity argument below only uses that convergence and the PDE.

Step 2. The subsolution inequality

By Week 6, §9.1, it is enough to use C^2 test functions, although the definition uses C^1 tests. Let $\varphi \in C^2$ and suppose that $u - \varphi$ has a strict local maximum at (x_0, t_0), with $t_0 > 0$. By Lemma 1.12, along the convergent subsequence there are local maxima of $u^{\varepsilon_j} - \varphi$ at $(x_j, t_j) \to (x_0, t_0)$. For brevity, write ε for ε_j and $(x_\varepsilon, t_\varepsilon)$ for (x_j, t_j). At such a point,

$$D(u^\varepsilon - \varphi) = 0, \qquad (u^\varepsilon - \varphi)_t = 0, \qquad \Delta(u^\varepsilon - \varphi) \le 0.$$

Hence

$$Du^\varepsilon = D\varphi, \qquad u_t^\varepsilon = \varphi_t, \qquad \Delta u^\varepsilon \le \Delta\varphi.$$

Substituting these relations into the viscous equation yields

$$\varphi_t(x_\varepsilon, t_\varepsilon) + H(x_\varepsilon, D\varphi(x_\varepsilon, t_\varepsilon)) \le \varepsilon\Delta\varphi(x_\varepsilon, t_\varepsilon).$$

Letting $\varepsilon \to 0$, we obtain

$$\varphi_t(x_0, t_0) + H(x_0, D\varphi(x_0, t_0)) \le 0,$$

which is exactly the viscosity subsolution inequality.

Step 3. The supersolution inequality

Let $\psi \in C^2$ touch u from below at an interior point (x_0, t_0), with a strict local minimum of $u - \psi$. At nearby local minima of $u^\varepsilon - \psi$,

$$Du^\varepsilon = D\psi, \qquad u_t^\varepsilon = \psi_t, \qquad \Delta u^\varepsilon \ge \Delta\psi.$$

Consequently,

$$\psi_t(x_\varepsilon, t_\varepsilon) + H(x_\varepsilon, D\psi(x_\varepsilon, t_\varepsilon)) \ge \varepsilon\Delta\psi(x_\varepsilon, t_\varepsilon).$$

The limit gives the supersolution inequality. Together with Step 1, this proves the theorem.

The small quantity is $\varepsilon\Delta\varphi$ or $\varepsilon\Delta\psi$, evaluated near a fixed contact point. We do not need a uniform bound on D^2u^ε, or convergence of Du^ε.

What the theorem establishes

The conclusion is subsequential existence. If a comparison principle gives uniqueness for the same Cauchy problem in the relevant function class, every subsequential limit must coincide. Together with the compactness already proved, this then gives convergence of the whole family. That uniqueness step is separate from the argument above.

An exact stationary illustration

An exact one-dimensional stationary example shows what the limiting process can look like. Define

$$u^\varepsilon(x) = -2\varepsilon \log\cosh\left(\frac{x}{2\varepsilon}\right).$$

Direct differentiation gives

$$(u^\varepsilon)'(x) = -\tanh\left(\frac{x}{2\varepsilon}\right), \qquad (u^\varepsilon)''(x) = -\frac{1}{2\varepsilon\cosh^2(x/(2\varepsilon))}.$$

Using $\tanh^2 s + 1/\cosh^2 s = 1$, we obtain

$$\frac{1}{2}\left|(u^\varepsilon)'(x)\right|^2 - \varepsilon(u^\varepsilon)''(x) = \frac{1}{2}.$$

Moreover,

$$0 \le u^\varepsilon(x) + |x| \le 2\varepsilon \log 2,$$

so $u^\varepsilon \to -|x|$ uniformly. Each positive-viscosity profile is smooth, while its derivative develops an increasingly thin transition layer near the limiting cusp. The limit solves $\frac{1}{2}|u'|^2 - \frac{1}{2} = 0$ in the viscosity sense, consistent with the downward-cusp example from Week 6.

These profiles are unbounded on $\mathbb{R}$. They illustrate the local stationary limiting argument rather than the bounded Cauchy theorem above. Notice also that

$$\varepsilon(u^\varepsilon)''(0) = -\frac{1}{2}.$$

Thus the viscous term need not tend to zero at a singular point of the limit. This is why the proof passes inequalities through smooth test functions.

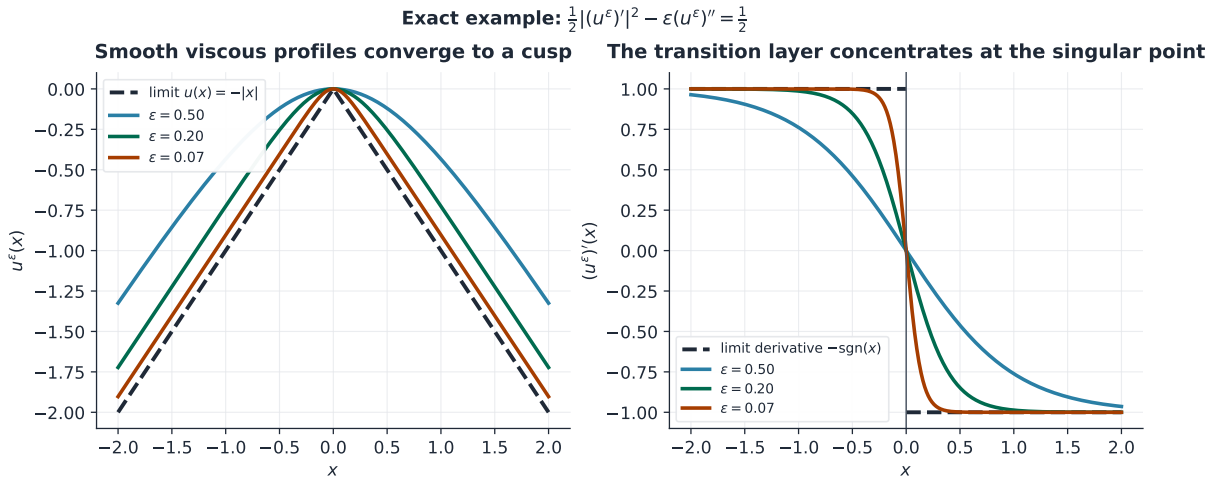


Figure 8.1. The functions u^ε solve the displayed viscous equation exactly. The profiles converge to $-|x|$, and their derivatives converge to $-\mathrm{sgn}(x)$ away from the singular point $x = 0$.

5. Consistency at Differentiability Points (Tran Theorem 1.16, p.15)

A viscosity solution is defined by touching test functions, not by pointwise derivatives of the solution itself. Nevertheless, at points where the solution happens to be differentiable, the viscosity formulation agrees with the classical PDE.

The key auxiliary result is the following.

Lemma 1.15

Let u be continuous and differentiable in all space-time variables at an interior point (x_0, t_0). Then there exist an upper C^1 test function φ and a lower C^1 test function ψ such that

$$\psi(x_0, t_0) = u(x_0, t_0) = \varphi(x_0, t_0),$$

$$\psi < u < \varphi \quad \text{near } (x_0, t_0) \text{ away from the touching point,}$$

and

$$D\psi(x_0, t_0) = Du(x_0, t_0) = D\varphi(x_0, t_0), \qquad \psi_t(x_0, t_0) = u_t(x_0, t_0) = \varphi_t(x_0, t_0).$$

This means that at a differentiability point we can build one test function touching from below and one touching from above, both with the correct first-order derivatives.

Proof of the lemma using Week 7

Set $z = (x, t) \in \mathbb{R}^{n+1}$ and $z_0 = (x_0, t_0)$. Differentiability gives

$$u(z_0 + h) = u(z_0) + q \cdot h + o(|h|), \qquad q = (Du(x_0, t_0), u_t(x_0, t_0)).$$

Hence q belongs to both generalized differential sets with respect to z. Apply Week 7, Theorem 1.8, in $\mathbb{R}^{n+1}$ to obtain a lower and an upper C^1 touching function with derivative q. Subtract $|z - z_0|^2$ from the lower function and add $|z - z_0|^2$ to the upper function. This makes the touching strict without changing their values or first derivatives at z_0. Denote the resulting lower function by ψ and the upper function by φ.

The tangent affine function alone need not touch u from either side. The touching-function theorem supplies the required small corrections.

Theorem 1.16

Tran states this theorem for the viscosity solution constructed by vanishing viscosity. The touching-function argument below in fact applies to any viscosity solution: if u is differentiable at (x_0, t_0), then

$$u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) = 0.$$

Proof

By Lemma 1.15, there exists an upper test function φ and a lower test function ψ touching u at (x_0, t_0), both with the same first-order derivatives as u at that point. Applying the viscosity subsolution inequality to the upper test φ gives

$$u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) \leq 0,$$

while applying the viscosity supersolution inequality to the lower test ψ gives

$$u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) \geq 0.$$

Therefore equality holds:

$$u_t(x_0, t_0) + H(x_0, Du(x_0, t_0)) = 0.$$

Thus the viscosity notion is fully consistent with the classical one at differentiability points. The converse statement, that an almost everywhere solution must be a viscosity solution, is false: the sawtooth example from Week 6 fails the lower test at a corner.

6. Stability of Viscosity Solutions (Tran Theorem 1.17, p.15)

Let H_k and H be continuous Hamiltonians. Work with functions u_k and u in $\mathrm{BUC}(\mathbb{R}^n \times [0, T])$ for each finite T, and assume

$$H_k \to H \quad \text{locally uniformly in } \mathbb{R}^n \times \mathbb{R}^n,$$

$$u_{0,k} \to u_0 \quad \text{locally uniformly on } \mathbb{R}^n,$$

and

$$u_k \to u \quad \text{locally uniformly on } \mathbb{R}^n \times [0, \infty).$$

Suppose that for each k, u_k is a viscosity solution of

$$(u_k)_t + H_k(x, Du_k) = 0, \qquad u_k(x, 0) = u_{0,k}(x).$$

Then u is a viscosity solution of

$$u_t + H(x, Du) = 0, \qquad u(x, 0) = u_0(x).$$

The function-class assumption on u is separate from local uniform convergence. Local uniform limits of bounded uniformly continuous functions need not be globally bounded or uniformly continuous. In §4, the uniform estimates guarantee both properties on every finite strip.

Proof

Take a test function $\varphi \in C^1$ such that $u - \varphi$ has a strict local maximum at (x_0, t_0), with $t_0 > 0$. By the stability of extremal points, for all large k, the function $u_k - \varphi$ has a local maximum at some $(x_k, t_k) \to (x_0, t_0)$. Normalize the test at each touching point by adding a constant, as in §3; its derivatives remain unchanged. Since u_k is a viscosity subsolution of the k-th equation,

$$\varphi_t(x_k, t_k) + H_k(x_k, D\varphi(x_k, t_k)) \leq 0.$$

The points $(x_k, D\varphi(x_k, t_k))$ stay in a fixed compact subset of $\mathbb{R}^n \times \mathbb{R}^n$. Thus local uniform convergence of H_k, together with continuity of H, gives

$$H_k(x_k, D\varphi(x_k, t_k)) \longrightarrow H(x_0, D\varphi(x_0, t_0)).$$

Passing to the limit, we obtain

$$\varphi_t(x_0, t_0) + H(x_0, D\varphi(x_0, t_0)) \leq 0.$$

This proves the subsolution property. For a lower test ψ, nearby local minima instead give

$$\psi_t(x_k, t_k) + H_k(x_k, D\psi(x_k, t_k)) \geq 0.$$

The same limiting argument proves the supersolution property. Finally,

$$u(x, 0) = \lim_{k\to\infty} u_k(x, 0) = \lim_{k\to\infty} u_{0,k}(x) = u_0(x).$$

This completes the proof.

Interpretation

This theorem is a closure property: a locally uniform limit retains the viscosity inequalities when the Hamiltonians converge locally uniformly as well. It does not supply compactness or prove convergence of a chosen approximation. Those are separate steps.

No derivatives of u_k are taken in this proof. All differentiations are applied to the test function.

Stationary version and the connection with Week 7

The same argument applies to the discounted stationary equation. Let $u_k, u \in \mathrm{BUC}(\mathbb{R}^n)$. If $\lambda > 0$ is fixed, $u_k \to u$ and $H_k \to H$ locally uniformly, and

$$\lambda u_k + H_k(x, Du_k) = 0$$

in the viscosity sense, then u satisfies $\lambda u + H(x, Du) = 0$ in the viscosity sense. At an upper touching point for u_k, the inequality is

$$\lambda u_k(x_k) + H_k(x_k, D\varphi(x_k)) \leq 0.$$

Since $u_k(x_k) \to u(x_0)$, the zeroth-order term also passes to the limit. This is the form relevant to approximations of the control value functions studied in Week 7.

7. The Bernstein Method (Tran Theorem 1.25, p.25)

We now explain how to obtain the uniform estimates required in Theorem 1.13. These are a priori estimates: we first work with a classical solution and then show that its bounds do not deteriorate as $\varepsilon \downarrow 0$.

Consider

$$u^\varepsilon_t + H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon, \qquad u^\varepsilon(x, 0) = u_0(x).$$

The goal is to bound **both the time derivative and the spatial gradient**, uniformly in the viscosity parameter. The bound must also be independent of the length of the time interval. We first state the assumptions and the precise estimate, then follow Tran's two-step argument: bound u^ε_t, and use that bound to control Du^ε.

Assumptions for the Bernstein estimate

(1.26)

$$u_0 \in C^2(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n) \quad \text{and} \quad \|u_0\|_{C^2(\mathbb{R}^n)} < \infty.$$

Here the bounded C^2-norm means that u_0 and all its derivatives through order two are bounded.

(1.27)

- $H \in C^2(\mathbb{R}^n \times \mathbb{R}^n)$,
- $H, D_p H \in \mathrm{BUC}(\mathbb{R}^n \times B(0,R))$ for each $R > 0$,
- and the following growth condition holds:

$$\lim_{|p|\to\infty} \inf_{x\in\mathbb{R}^n} \left\{ \frac{1}{2} H(x,p)^2 + D_x H(x,p) \cdot p \right\} = +\infty.$$

This condition means that the expression in braces eventually exceeds any prescribed constant, uniformly in x. It is a condition on the whole expression, not just on H.

Under (1.26)–(1.27), we use the classical parabolic existence result cited by Tran: for each fixed $\varepsilon \in (0,1)$, the regularized problem has a unique solution smooth enough for the calculations below, with finite bounds on u_t^ε and Du^ε on finite time strips. These preliminary bounds may depend on ε. Proving this classical existence result is outside the scope of the present lecture.

Statement to be proved: the uniform Bernstein estimate

Under (1.26)–(1.27), there exists a constant $C > 0$, depending on u_0, H, and n, but independent of ε and T, such that for every $T > 0$ and all sufficiently small $\varepsilon > 0$,

$$\boxed{\|u_t^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,T])} + \|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,T])} \le C.}$$

Because the same C works for every T, the estimate holds on $\mathbb{R}^n \times [0,\infty)$. It also implies

$$|u^\varepsilon(x,t)| \le \|u_0\|_\infty + Ct,$$

so the functions themselves are uniformly bounded on each finite time strip. These are exactly the estimates needed for the compactness step in §4. We do not claim that u^ε is bounded uniformly over all $t \ge 0$, or that its Hessian has a bound independent of ε.

Relation to Tran Theorem 1.25. Tran states the theorem for $\varepsilon \in (0,1)$. His proof on p.26 takes $\varepsilon < 1/n$ in the gradient calculation. We follow that calculation to prove the uniform estimate as $\varepsilon \downarrow 0$, which is the part used in §4. The smallness condition will be written only where it enters the proof; no separate cutoff parameter is needed.

These stronger assumptions apply only to this Bernstein argument. They are not additional assumptions on the DPP theorem proved in Week 7. A control Hamiltonian defined as a supremum need not be C^2, so that theorem cannot automatically be combined with the present estimate without further approximation arguments.

Comparison used before the later comparison lectures

We have not yet developed comparison theorems systematically. The **first-order viscosity comparison principles** will be proved in Week 9 for the stationary equation and Week 10 for the Cauchy problem. In the Bernstein proof, however, $\varepsilon > 0$ is fixed and the solutions are classical. We use the standard **classical parabolic comparison principle** for the regularized equations. Thus the existence argument here does not presuppose the first-order viscosity comparison theorems from the following weeks.

The parabolic fact being used is stated below, with a short maximum-principle argument. In Step 1, it will be applied first to the linear equation for u_t^ε, and then to the comparison barriers $u_0 \pm C_0 t$.

The classical parabolic maximum principle

We will use the following elementary parabolic fact. Let $a(x,t)$ be a bounded vector field, let $\varepsilon > 0$ be fixed, and let z be bounded on $\mathbb{R}^n \times [0,T]$, continuous up to $t = 0$, and smooth for $t > 0$. If

$$z_t + a(x,t) \cdot Dz - \varepsilon \Delta z \le 0,$$

then

$$\sup_{x \in \mathbb{R}^n} z(x,t) \le \sup_{x \in \mathbb{R}^n} z(x,0) \qquad (0 \le t \le T).$$

The reversed inequality gives the corresponding lower bound. This is a statement about smooth functions with a fixed positive diffusion coefficient.

To see the mechanism, set $M = \sup_x z(x,0)$ and consider

$$z(x,t) - M - \eta t - \delta \chi(x), \qquad \chi(x) = \sqrt{1 + |x|^2}, \qquad \eta, \delta > 0.$$

Because z is bounded and $\chi(x) \to \infty$, a positive maximum, if it exists, is attained at a point with $t > 0$. At that point,

$$z_t \ge \eta, \qquad Dz = \delta D\chi, \qquad \Delta z \le \delta \Delta \chi.$$

Using $|D\chi| \le 1$ and $0 \le \Delta\chi \le n$, the left-hand side of the differential inequality is at least

$$\eta - \delta(\|a\|_\infty + n\varepsilon) > 0$$

for sufficiently small δ, a contradiction. Let $\delta \downarrow 0$ and then $\eta \downarrow 0$.

For completeness, let r be a classical subsolution and s a classical supersolution of the viscous Hamilton–Jacobi equation on a finite strip, with bounded values and gradients and $r(x,0) \le s(x,0)$. Their difference $z = r - s$ satisfies

$$z_t + a(x,t) \cdot Dz - \varepsilon \Delta z \le 0,$$

where

$$a(x,t) = \int_0^1 D_p H(x, Ds + \theta(Dr - Ds))\, d\theta.$$

The fundamental theorem of calculus gives this formula. Boundedness of $D_p H$ on bounded gradient regions makes a bounded for each fixed strip. The maximum principle yields $z \le 0$, or $r \le s$. This is the precise classical comparison statement used for the barriers below.

8. Two Steps in the Bernstein Proof (Tran pp.25–27)

We retain $w^\varepsilon = u_t^\varepsilon$ and $\Psi = \frac{1}{2}|Du^\varepsilon|^2$ for the auxiliary quantities, keeping φ and ψ reserved for upper and lower test functions elsewhere in the notes. All preliminary estimates below are taken on a finite strip $\mathbb{R}^n \times [0,T]$; the final constants will not depend on T.

Step 1. Uniform bound for u_t^ε

Write

$$C_0 = \sup_{x \in \mathbb{R}^n} |H(x, Du_0(x))| + \|\Delta u_0\|_\infty.$$

This is finite because Du_0 is bounded and H is bounded on each set with bounded p. The constant C_0 is independent of ε.

Set

$$w^\varepsilon = u_t^\varepsilon.$$

Differentiating the viscous equation in time gives

$$w_t^\varepsilon + D_p H(x, Du^\varepsilon) \cdot Dw^\varepsilon = \varepsilon \Delta w^\varepsilon.$$

This is a linear parabolic equation for w^ε. By the classical parabolic maximum principle stated in §7, applied to both w^ε and $-w^\varepsilon$,

$$\inf_{x\in\mathbb{R}^n} w^\varepsilon(x,0) \le w^\varepsilon(x,t) \le \sup_{x\in\mathbb{R}^n} w^\varepsilon(x,0).$$

Therefore, to control u_t^ε for all time, it is enough to control $u_t^\varepsilon(\cdot,0)$.

To do this, introduce the lower and upper comparison barriers

$$\underline{u}(x,t) = u_0(x) - C_0 t, \qquad \overline{u}(x,t) = u_0(x) + C_0 t.$$

For $0 < \varepsilon < 1$,

$$|H(x, Du_0) - \varepsilon \Delta u_0| \le C_0.$$

Consequently,

$$\underline{u}_t + H(x, D\underline{u}) - \varepsilon \Delta \underline{u} \le 0,$$

and

$$\overline{u}_t + H(x, D\overline{u}) - \varepsilon \Delta \overline{u} \ge 0.$$

Thus $\underline{u}$ is a classical subsolution and $\overline{u}$ is a classical supersolution of the viscous problem, with the same initial values as u^ε. By the classical parabolic comparison statement in §7,

$$\underline{u}(x,t) \le u^\varepsilon(x,t) \le \overline{u}(x,t).$$

Hence

$$u_0(x) - C_0 t \le u^\varepsilon(x,t) \le u_0(x) + C_0 t,$$

which implies

$$|u_t^\varepsilon(x,0)| \le C_0.$$

Therefore

$$\boxed{\|u_t^\varepsilon\|_{L^\infty} \le C_0.}$$

Equivalently, when the equation is evaluated at the initial time, $u_t^\varepsilon(x,0) = \varepsilon \Delta u_0(x) - H(x, Du_0(x))$. The bounded C^2-norm is used to control both terms uniformly in ε. The bound for later times comes from the maximum principle for w^ε, not from differentiating the barrier inequalities at an arbitrary positive time.

The absence of explicit time dependence in H is also used here. If $H = H(x,t,p)$, differentiating the PDE introduces an additional H_t term.

Step 2. Uniform bound for Du^ε

Define

$$\Psi = \frac{1}{2}|Du^\varepsilon|^2.$$

Differentiating the viscous equation with respect to x_k gives

$$(u^\varepsilon_{x_k})_t + H_{x_k}(x, Du^\varepsilon) + \sum_{j=1}^n H_{p_j}(x, Du^\varepsilon) u^\varepsilon_{x_j x_k} = \varepsilon \Delta u^\varepsilon_{x_k}.$$

Multiply by $u^\varepsilon_{x_k}$ and sum over k. The identities

$$D\Psi = \sum_{k=1}^n u^\varepsilon_{x_k} Du^\varepsilon_{x_k}, \qquad \Delta\Psi = \sum_{k=1}^n u^\varepsilon_{x_k} \Delta u^\varepsilon_{x_k} + |D^2 u^\varepsilon|^2$$

yield

$$\Psi_t + D_p H(x, Du^\varepsilon) \cdot D\Psi + D_x H(x, Du^\varepsilon) \cdot Du^\varepsilon = \varepsilon \Delta\Psi - \varepsilon |D^2 u^\varepsilon|^2.$$

The last term on the right-hand side is nonpositive. To connect it with the structural condition, use

$$|D^2 u^\varepsilon|^2 \geq \frac{(\Delta u^\varepsilon)^2}{n}, \qquad \varepsilon \Delta u^\varepsilon = u^\varepsilon_t + H(x, Du^\varepsilon).$$

Now follow the calculation in Tran's proof on p.26. For $0 < \varepsilon < 1/n$, we have $\varepsilon/n \geq \varepsilon^2$, so

$$\varepsilon |D^2 u^\varepsilon|^2 \geq \frac{\varepsilon}{n} (\Delta u^\varepsilon)^2 \geq \varepsilon^2 (\Delta u^\varepsilon)^2 = \left(u^\varepsilon_t + H(x, Du^\varepsilon)\right)^2.$$

Using $(a+b)^2 \geq \frac{1}{2} b^2 - a^2$ and Step 1, we obtain

$$\varepsilon |D^2 u^\varepsilon|^2 \geq \frac{1}{2} H(x, Du^\varepsilon)^2 - C_0^2.$$

Define

$$Q(x, p) = \frac{1}{2} H(x, p)^2 + D_x H(x, p) \cdot p.$$

The equation for Ψ therefore implies

$$\Psi_t + D_p H(x, Du^\varepsilon) \cdot D\Psi - \varepsilon \Delta\Psi + Q(x, Du^\varepsilon) - C_0^2 \leq 0.$$

By (1.27), choose $R_0 > 0$, independent of ε and T, such that

$$Q(x, p) > C_0^2 + 1 \qquad \text{whenever } |p| \geq R_0, \quad x \in \mathbb{R}^n.$$

Fix $T > 0$ and suppose first that Ψ attains its maximum over $\mathbb{R}^n \times [0, T]$ at (x_0, t_0). If $t_0 = 0$, then

$$\Psi(x_0, 0) \leq \frac{1}{2} \|Du_0\|^2_{L^\infty}.$$

If $t_0 > 0$, then

$$D\Psi = 0, \qquad \Psi_t \geq 0, \qquad \Delta\Psi \leq 0.$$

Substituting these relations into the preceding inequality gives

$$\frac{1}{2} H(x_0, Du^\varepsilon)^2 + D_x H(x_0, Du^\varepsilon) \cdot Du^\varepsilon \leq C_0^2.$$

Thus $|Du^\varepsilon(x_0, t_0)| < R_0$. Combining the initial-time and positive-time cases gives

$$\sup_{\mathbb{R}^n \times [0,T]} \Psi \leq \frac{1}{2} \max\{\|Du_0\|^2_\infty, R_0^2\}.$$

At a maximum with $t_0 = T$, the inequality $\Psi_t \geq 0$ is understood using the derivative from earlier times. This is why a maximum on the closed time interval can be used.

If the maximum is attained, we have therefore proved the desired spatial gradient estimate. Together with Step 1, this gives the two bounds in the statement of §7. The remaining point is to justify the same estimate when a maximum is not attained.

Remark: maxima on the whole space (Tran Remark 1.27)

On the unbounded domain $\mathbb{R}^n$, the maximum of Ψ need not be attained. Following Tran Remark 1.27, introduce

$$\Psi_\delta(x,t) = \Psi(x,t) - \delta\sqrt{1+|x|^2}, \qquad \delta > 0.$$

Fix ε and T first. The preliminary classical gradient bound is finite on this strip, so $\Psi_\delta \to -\infty$ as $|x| \to \infty$, uniformly in $t \in [0,T]$. Hence Ψ_δ attains a maximum.

Let $\chi(x) = \sqrt{1+|x|^2}$. If the maximizing time is positive, then at the maximizing point

$$D\Psi = \delta D\chi, \qquad \Psi_t \geq 0, \qquad \Delta\Psi \leq \delta\Delta\chi \leq n\delta.$$

For this fixed ε and T, the number

$$A_{\varepsilon,T} = \sup_{\mathbb{R}^n\times[0,T]} |D_p H(x, Du^\varepsilon(x,t))|$$

is finite. Substitution into the differential inequality gives

$$Q(x, Du^\varepsilon) \leq C_0^2 + \delta(A_{\varepsilon,T} + n\varepsilon).$$

Choose δ small enough that the last error is less than one. The definition of R_0 then bounds the gradient at this maximum by R_0. If the maximizing time is zero, the bound follows from Du_0. In both cases,

$$\Psi(x,t) - \delta\chi(x) \leq \max\Psi_\delta \leq \frac{1}{2}\max\{\|Du_0\|_\infty^2, R_0^2\}.$$

Let $\delta \downarrow 0$ with ε and T fixed. Since the right-hand side is independent of both, and T is arbitrary,

$$\boxed{\|Du^\varepsilon\|_{L^\infty(\mathbb{R}^n\times[0,\infty))} \leq \max\{\|Du_0\|_\infty, R_0\}.}$$

Together with Step 1, this proves the target estimate in §7, with a possible choice

$$C = C_0 + \max\{\|Du_0\|_\infty, R_0\}.$$

Neither C_0 nor R_0 depends on ε or T. The order of the argument matters: first fix ε and T, then let $\delta \downarrow 0$. No uniform bound on $A_{\varepsilon,T}$ is assumed in advance. The parameter δ is only a spatial penalization used to attain a maximum; it is distinct from the viscosity parameter ε.

Interpretation

The Bernstein method estimates the gradient not by direct integration but by treating the quantity $\frac{1}{2}|Du^\varepsilon|^2$ as a new unknown and applying the maximum principle to it. The argument works because the Hamiltonian has enough coercive structure to prevent arbitrarily large gradient maxima.

9. Examples: Applying the Bernstein Method

Example 8.1: a quadratic Hamiltonian

Consider

$$H(p) = \frac{1}{2}|p|^2.$$

Then H is independent of x, so $D_x H = 0$, and

$$\frac{1}{2}H(p)^2 + D_x H(x,p)\cdot p = \frac{1}{2}\left(\frac{1}{2}|p|^2\right)^2 = \frac{1}{8}|p|^4.$$

Hence

$$\frac{1}{8}|p|^4 \to \infty \qquad \text{as } |p| \to \infty.$$

Therefore the coercivity assumption in (1.27) is satisfied.

If

$$u_0 \in C^2(\mathbb{R}^n) \cap \mathrm{BUC}(\mathbb{R}^n), \qquad \|u_0\|_{C^2(\mathbb{R}^n)} < \infty,$$

then all the assumptions of Theorem 1.25 hold. Consequently, the estimate proved in §§7–8 gives a constant $C > 0$, independent of all sufficiently small $\varepsilon > 0$, such that

$$\|u_t^\varepsilon\|_{L^\infty} + \|Du^\varepsilon\|_{L^\infty} \le C.$$

By Theorem 1.13, we may extract a subsequence $u^{\varepsilon_j} \to u$ locally uniformly, and the limit u is a viscosity solution of

$$u_t + \frac{1}{2}|Du|^2 = 0, \qquad u(x,0) = u_0(x).$$

This is the model example of the theory: the viscous approximations are smooth, the Bernstein method yields viscosity-independent estimates, and the limit is a viscosity solution of the first-order Hamilton–Jacobi equation.

For this particular Hamiltonian, a sharper gradient estimate follows directly from the equation for Ψ. Since $D_x H = 0$,

$$\Psi_t + Du^\varepsilon \cdot D\Psi - \varepsilon\Delta\Psi = -\varepsilon|D^2 u^\varepsilon|^2 \le 0.$$

The classical maximum principle gives

$$\|Du^\varepsilon(\cdot,t)\|_\infty \le \|Du_0\|_\infty.$$

The more general structural argument in §8 is useful when $D_x H \neq 0$.

Example 8.2: a bounded spatial potential

Let

$$H(x,p) = \frac{1}{2}|p|^2 + P(x), \qquad P \in C^2(\mathbb{R}^n), \quad \|P\|_{C^2} < \infty.$$

Now $D_x H = DP(x)$ and $D_p H = p$. If $A = \|P\|_\infty$ and $B = \|DP\|_\infty$, the same square inequality used in §8 gives

$$\begin{aligned} Q(x,p) &= \frac{1}{2}\left(\frac{1}{2}|p|^2 + P(x)\right)^2 + DP(x)\cdot p \\ &\ge \frac{1}{16}|p|^4 - \frac{1}{2}A^2 - B|p|. \end{aligned}$$

The lower bound tends to $+\infty$ as $|p| \to \infty$, uniformly in x. The other assumptions in (1.27) follow from the bounded derivatives of P. Thus the Bernstein estimate and vanishing viscosity construction also apply to

$$u_t + \frac{1}{2}|Du|^2 + P(x) = 0$$

with initial data satisfying (1.26). This example shows how the quartic growth in H^2 controls the possibly negative term $D_x H \cdot p$.

10. Summary

The existence theory in this week is built from three ingredients, once classical solutions of the regularized problems are available.

1. **Bernstein estimates** provide bounds independent of sufficiently small ε.
2. **Compactness** gives a subsequence converging locally uniformly, including at the initial time.
3. **Stability of viscosity inequalities** shows that its limit solves the first-order equation.

The consistency theorem shows that a viscosity solution satisfies the classical PDE at every point of differentiability. Its proof uses the touching-function theorem from Week 7. Stability, in contrast, uses convergence of the functions and Hamiltonians, without differentiating the solutions.

Week 7 established the viscosity characterization of a control-defined value function through the DPP. This week established a PDE construction by approximation. Identifying solutions from different constructions, and upgrading subsequential convergence to convergence of the whole family, requires uniqueness for the same equation and data.

11. Exercises

Exercise 8.1. Adapt the proof of Lemma 1.12 to a strict local minimum. Then consider $u_k(x) = x/k$, $u(x) = 0$, and $\varphi(x) = 0$ on $(-1, 1)$. Explain why ordinary, non-strict contact at 0 does not by itself guarantee nearby local maxima of $u_k - \varphi$.

Exercise 8.2. In the proof of Theorem 1.13, explain precisely where the uniform bound

$$\|u_t^\varepsilon\|_{L^\infty} + \|Du^\varepsilon\|_{L^\infty} \le C$$

is used. Separate the role of the estimates in compactness from the role of local uniform convergence in the viscosity test. Explain how the initial condition passes to the limit.

Exercise 8.3. Prove Theorem 1.16 using Lemma 1.15 and the definitions of viscosity subsolution and viscosity supersolution.

Exercise 8.4. In Step 1 of the Bernstein proof, identify exactly where the assumption

$$\|u_0\|_{C^2(\mathbb{R}^n)} < \infty$$

is used. State the classical parabolic comparison result used in that step and explain how it differs from the first-order viscosity comparison theorems to be studied in Weeks 9–10.

Exercise 8.5. Verify carefully that for $H(p) = \frac{1}{2}|p|^2$,

$$\frac{1}{2}H(p)^2 + D_x H(x,p)\cdot p = \frac{1}{8}|p|^4 \to \infty \qquad \text{as } |p| \to \infty.$$

Also derive $\|Du^\varepsilon(\cdot,t)\|_\infty \le \|Du_0\|_\infty$ from the differential inequality for Ψ.

Exercise 8.6. Let

$$\Psi_\delta(x,t) = \frac{1}{2}|Du^\varepsilon(x,t)|^2 - \delta\sqrt{1+|x|^2}.$$

For fixed ε and T, use the preliminary classical gradient bound to show that Ψ_δ attains a maximum on $\mathbb{R}^n \times [0,T]$. Compute $D\chi$ and $\Delta\chi$ for $\chi(x) = \sqrt{1+|x|^2}$, and verify the derivative equalities and inequalities at a maximizing point used in §8. Explain why δ is sent to zero with ε fixed.

Exercise 8.7. For the exact stationary profiles in §4, verify

$$0 \le -2\varepsilon \log\cosh\left(\frac{x}{2\varepsilon}\right) + |x| \le 2\varepsilon\log 2.$$

Use $\frac{1}{2}e^{|s|} \le \cosh s \le e^{|s|}$. Compute $\varepsilon(u^\varepsilon)''(0)$ and explain why its nonzero limit does not contradict the viscosity limiting argument.

Exercise 8.8. Assume, in addition to the hypotheses of §4, that a comparison principle gives uniqueness for the limiting Cauchy problem in the class containing the subsequential limits. Prove that the whole family u^ε converges locally uniformly. If convergence fails on a compact set, extract a sequence remaining a fixed distance from the proposed limit, then apply compactness and uniqueness to that sequence.

12. References

- H. V. Tran, *Hamilton–Jacobi Equations: Theory and Applications*, AMS Graduate Studies in Mathematics 213, 2021.