Preliminary Reflections on the Hamilton-Jacobi Equation

Introduction to Viscosity Solutions

As Lions pointed out, the concept of viscosity solutions allows us to work with HJ equations even when solutions are nowhere differentiable. It's somewhat similar to the idea of weak solutions in distribution theory, but instead of integration by parts, we're essentially using "differentiation by parts," and everything happens inside the nonlinearity. This approach is very useful when passing to the limit in approximations.

For linear PDEs, we often define weak solutions by integrating against test functions, since the adjoint operator L^* is explicit. But for nonlinear equations like Hamilton-Jacobi, this isn't practical—so it makes sense to switch to a more local approach. That's where viscosity solutions come in: we just compare the solution with smooth functions that touch it from above or below.

There are two main types of HJ equations: time-dependent (on $[0,T]\times\mathbb{R}^n$) and time-independent. When dealing with the time-dependent case, we need to be extra careful at $t=0$ and $t=T$. For example, in the textbook, Lemma 1.22 shows how to use a function like $\epsilon/(T-t)$ to prove the subsolution test still holds at $t=T$. But this kind of trick doesn't always work at both time boundaries.

This becomes especially important when we use the doubling variables method to prove uniqueness. In the time-independent case, things are simpler. But in the time-dependent case, we have to prove that the maximum point of our function doesn't lie at $t=0$.

Comments on the Adjoint Method and Convergence Results

One clear advantage of the adjoint method over the doubling variables method is that it linearizes the Hamilton-Jacobi (HJ) equation. This makes it possible to apply classical techniques such as integration by parts. In the textbook, there are two important convergence results for the time-independent case:

- **Lemma 1.44**: In general condition, for the regularized solution u^ε and the corresponding adjoint density σ^ε, we have the second derivative estimate:
$$\varepsilon \int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 \sigma^\varepsilon \, dx \leq C.$$
- **Theorem 1.47**: In convex condition, let u^ε be the smooth approximation and u be the Lipschitz viscosity solution. Then for any test function $r \in C_c^\infty(\mathbb{R}^n)$, the averaged error satisfies:
$$\left| \int_{\mathbb{R}^n} (u^\varepsilon(x) - u(x))\, r(x+y)\, dx \right| \leq C \left(1 + \|Dr\|_{L^1(\mathbb{R}^n)}\right)^{1/2} \varepsilon.$$

From these, we see that to study $\partial_\varepsilon u^\varepsilon$, it is enough to understand the properties of the adjoint solution σ^ε and Δu^ε. What I find surprising is that, in this context, the time-dependent case is actually simpler than the time-independent one.

In Professor Tran's 2011 paper, the time-dependent results of Evans (2010) were extended to the time-independent setting. As the paper mentions, the adjoint method does not apply directly to the static case due to issues with existence, uniqueness, and the nonnegativity of the adjoint solution.

To resolve this, the paper transforms the time-independent problem into a degenerate time-dependent one and then applies the adjoint method. I believe a similar idea might be useful for extending the results of Cirant–Goffi from the time-dependent to the time-independent case.

One thing I found confusing is that the proof for the time-independent case in the textbook appears different from that in the paper. I'm not sure whether this is due to later developments on the fundamental solution, or simply reflects a different strategic choice.

Remarks on the Cirant–Goffi Paper

To better understand the Cirant–Goffi paper, I first compared it with the convergence results like Lemma 1.44 and Theorem 1.47 discussed earlier.

One difficulty with Theorem 1.47 is that the convergence estimate depends on the derivative $\|Dr\|_{L^1}$, which blows up as the test function r weakly converges to a Dirac mass δ_x. This makes it challenging to directly obtain pointwise convergence estimates from the time-independent theory.

In contrast, the time-dependent case might actually handle this situation better. Consider the structure of the fundamental solution ρ to the adjoint equation, which in the time-dependent setting satisfies:

$$\begin{cases} \partial_t \rho - \varepsilon \Delta \rho - \operatorname{div}(D_p H(Du_\varepsilon)\rho) = 0 & \text{in } Q_{\tau,T} := \mathbb{T}^n \times (\tau, T), \\ \rho(\tau) = \delta_{\bar{x}} & \text{in } \mathbb{T}^n. \end{cases}$$

This setup is quite helpful: the singularity $\delta_{\bar{x}}$ only appears at the initial time $t = \tau$. For $t > \tau$, the solution becomes smoother and decays away from the singularity. This reminds me of how kernel functions behave in Fourier analysis—they may blow up at a point but decay rapidly elsewhere.

This analogy suggests a useful techniques that we often use in Fourier analysis: we might try to split the analysis into two time regimes—when t is close to τ, and when it is not—and use different tools for each. This could lead to improved convergence rate estimates in the time-dependent setting.

One of the key technical contributions of the Cirant–Goffi paper is how it handles the singularity at $t = \tau$ using time-dependent weights. Specifically, they prove estimates of the form:

$$\int_{\tau+\delta}^{T} \int_{\mathbb{T}^n} |D^2 u_\varepsilon|^2 \rho \, dx dt \lesssim \frac{1}{\delta^\alpha},$$

for some $\delta > 0$ and $\alpha > 1$, as shown in equation (8) of the paper.

This, combined with previous estimates, allows the authors to split the time interval (τ, T) into $(\tau, \tau + \delta) \cup (\tau + \delta, T)$, and estimate integrals like:

$$\left| \iint_{\mathbb{T}^n \times (\tau,T)} \Delta u_\varepsilon \rho \right| \lesssim \frac{1}{\delta^\alpha} + \sqrt{\frac{\delta}{\varepsilon}},$$

which leads to convergence rates of order $\mathcal{O}(\varepsilon^\beta)$ for $\beta < 1$, by choosing $\delta = \delta(\varepsilon)$ appropriately.

What I found elegant in the Cirant–Goffi paper is how different time weights are used to cancel the singularity at $t = \tau$, leading to different convergence rates.

By multiplying the equation with $(t - \tau)^{\alpha}$ and using the uniform convexity of H along with the classical Bernstein method which is shown in Theorem 1.47, they obtain an $\mathcal{O}(\varepsilon^{\beta})$ estimate. Similarly, in Section 4, by multiplying with $X(t) = (t - \tau)(T - t)$ and using the known decay properties of $|D\rho|$, they derive an $\mathcal{O}(\varepsilon \log \varepsilon)$ bound.