Notes on Hölder Estimates for Parabolic PDE

Sébastien Picard

June 17, 2019

Abstract

These are lecture notes on parabolic differential equations, with a focus on estimates in Hölder spaces. The two main goals of our discussion are to obtain the parabolic Schauder estimate and the Krylov-Safonov estimate.

Contents

1 Maximum Principles

We start by establishing notation. Let $\Omega \subset \mathbf{R}^n$ be a bounded, connected open set with smooth boundary. Let $T > 0$. Let

$$a^{ij} : \overline{\Omega} \times [0,T] \to \mathbf{R}, \tag{1.1}$$

and

$$b^i : \overline{\Omega} \times [0,T] \to \mathbf{R} \tag{1.2}$$

be smooth functions satisfying, for some positive constant Λ,

$$a^{ij} = a^{ji}, \quad \Lambda^{-1}|\xi|^2 \le a^{ij}(x,t)\xi_i\xi_j \le \Lambda|\xi|^2, \quad |b^i| \le \Lambda, \tag{1.3}$$

for all $\xi \in \mathbf{R}^n$ and points $(x,t) \in \Omega \times [0,T]$.

We will use D for spacial derivatives: $D_i = \frac{\partial}{\partial x^i}$. Derivatives in time will be denoted ∂_t. We will study the parabolic operator

$$\partial_t - a^{ij} D_i D_j - b^i D_i, \tag{1.4}$$

where we use the summation convention where the summation is implied by repeated indices. For example,

$$a^{ij} D_i D_j u = \sum_{i,j} a^{ij} D_i D_j u. \tag{1.5}$$

We define $C^{2,1}(\overline{\Omega} \times [0,T])$ to be the class of functions

$$u : \overline{\Omega} \times [0,T] \to \mathbf{R} \tag{1.6}$$

with continuous spacial partial derivatives up to order 2, and which are once continuously differentiable in time.

1.1 Weak maximum principle

Theorem 1 *Suppose a function $u \in C^{2,1}(\overline{\Omega} \times [0,T])$ satisfies the parabolic inequality*

$$\left(\partial_t - a^{ij} D_i D_j - b^i D_i\right) u(x,t) \ge 0. \tag{1.7}$$

Then

$$u(x,t) \ge \min\left\{ \inf_{\Omega} u(x,0), \inf_{\partial\Omega \times [0,T]} u(x,t) \right\}. \tag{1.8}$$

Proof: Let $\varepsilon > 0$ and $v = u + \varepsilon t$. Suppose that on the compact set $\overline{\Omega} \times [0, T]$, the function $v(x,t)$ attains a minimum at (x_0, t_0). Suppose $t_0 > 0$ and $x_0 \in \Omega$. Then we must have

$$\partial_t v(x_0, t_0) \leq 0, \quad Dv(x_0, t_0) = 0, \quad D^2 v(x_0, t_0) \geq 0. \tag{1.9}$$

Thus

$$\partial_t v(x_0, t_0) \leq 0, \quad a^{ij} D_i D_j v(x_0, t_0) + b^k D_k v(x_0, t_0) \geq 0. \tag{1.10}$$

This implies

$$\partial_t u(x_0, t_0) \leq -\varepsilon, \quad (a^{ij} D_i D_j + b^i D_i) u(x_0, t_0) \geq 0, \tag{1.11}$$

hence

$$\left(\partial_t - a^{ij}(x,t) D_i D_j - b^i D_i\right) u(x_0, t_0) < 0, \tag{1.12}$$

a contradiction. Therefore $t_0 = 0$ or $x_0 \in \partial\Omega$.

$$v(x,t) \geq \min\left\{ \inf_{\Omega} v(x,0), \inf_{\partial\Omega \times [0,T]} v(x,t) \right\}. \tag{1.13}$$

In terms of u, this means

$$u(x,t) + \varepsilon T \geq \min\left\{ \inf_{\Omega} u(x,0), \inf_{\partial\Omega \times [0,T]} u(x,t) \right\}. \tag{1.14}$$

Letting $\varepsilon \to 0$, we have the desired inequality. □

Similarly, we have the following maximum principle.

Theorem 2 *Suppose a function* $u \in C^{2,1}(\overline{\Omega} \times [0,T])$ *satisfies the parabolic inequality*

$$\left(\partial_t - a^{ij} D_i D_j - b^i D_i\right) u(x,t) \leq 0. \tag{1.15}$$

Then

$$u(x,t) \leq \max\left\{ \sup_{\Omega} u(x,0), \sup_{\partial\Omega \times [0,T]} u(x,t) \right\}. \tag{1.16}$$

Proof: Let $v = -u$ and apply the previous theorem. □

As an application of the maximum principle, we obtain the comparison principle.

Theorem 3 *Suppose functions* $u, v \in C^{2,1}(\overline{\Omega}\times[0,T])$ *satisfy the parabolic inequality*

$$(\partial_t - a^{ij}D_iD_j - b^iD_i)\,u(x,t) \leq (\partial_t - a^{ij}D_iD_j - b^iD_i)\,v(x,t). \quad (1.17)$$

Further suppose

$$u(x,0) \leq v(x,0), \quad u|_{\partial\Omega} \leq v|_{\partial\Omega}. \quad (1.18)$$

Then

$$u(x,t) \leq v(x,t) \quad (1.19)$$

for all $(x,t) \in \overline{\Omega} \times [0,T]$.

Proof: Let $w = u - v$. Then

$$(\partial_t - a^{ij}D_iD_j - b^iD_i)w \leq 0. \quad (1.20)$$

By the maximum principle,

$$w(x,t) \leq 0, \quad (1.21)$$

since $w(x,0) \leq 0$ and $w|_{\partial\Omega} \leq 0$. □

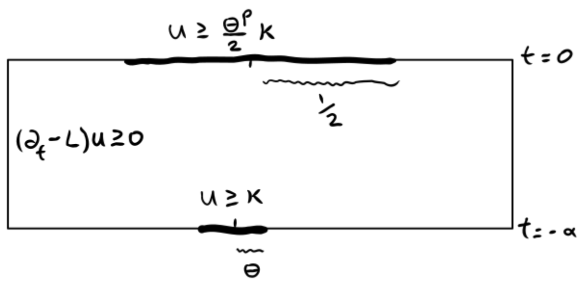
1.2 Strong maximum principle

The reference for this section is [8]. We let

$$L = a^{ij}D_iD_j + b^iD_i, \quad (1.22)$$

to simplify notation.

In this section, we prove a lemma which props up positive supersolutions (u satisfying $(\partial_t - L)u \geq 0$). The case with $R = 1$, $\alpha = 1$ and $\theta = 1/2$ is already useful, but for future reference we state a general version with flexible parameters. This lemma is due to Krylov and Safonov [7].



Lemma 1 *Let $R, \alpha > 0$. Let $u \in C^{2,1}(\overline{Q})$, where*

$$Q = \{(x,t) : |x| < R, \quad -\alpha R^2 < t < 0\}. \tag{1.23}$$

Suppose the coefficients of L, defined by (1.22) in Q, satisfy (1.3) for some $\Lambda > 0$. Suppose

$$(\partial_t - L)u \geq 0 \tag{1.24}$$

and $u \geq 0$ in Q. Let $\kappa > 0$ and $0 < \theta < 1$. Suppose

$$u(x, -\alpha R^2) \geq \kappa, \quad |x| < \theta R. \tag{1.25}$$

Then

$$u(x,0) \geq \theta^p \frac{\kappa}{2}, \quad |x| < \frac{R}{2}, \tag{1.26}$$

for p depending on α, Λ, R, θ.

Proof: Consider

$$\rho(t) = \frac{(1-\theta^2)}{\alpha}(t + \alpha R^2) + \theta^2 R^2. \tag{1.27}$$

This function is a linear interpolation, which from time $t = -\alpha R^2$ to $t = 0$ ranges from radius $\theta^2 R^2$ to R^2. Next, define

$$\chi(x,t) = \max\{\rho(t) - |x|^2, 0\}. \tag{1.28}$$

This function is a cutoff in space centered at $x = 0$ of radius ρ. The space-time barrier cutoff is

$$\Psi = \chi^2 \rho^{-q}, \quad q \geq 2. \tag{1.29}$$

We view Ψ as a function on

$$\tilde{Q} = \{(x,t) : |x|^2 < \rho(t), \quad -\alpha R^2 < t < 0\}, \tag{1.30}$$

where it can be freely differentiated. Differentiating in time gives

$$\partial_t \Psi = 2\chi \frac{1-\theta^2}{\alpha} \rho^{-q} - q\chi^2 \frac{1-\theta^2}{\alpha} \rho^{-q-1}. \tag{1.31}$$

Differentiating once in space, we obtain

$$D\Psi = 2\chi D\chi \rho^{-q}. \tag{1.32}$$

Differentiating twice in space, we obtain

$$\begin{aligned} D_i D_j \Psi &= 2\rho^{-q} D_i\chi D_j\chi + 2\chi\rho^{-q} D_i D_j\chi \\ &= 8\rho^{-q} x_i x_j - 4\chi\rho^{-q}\delta_{ij}. \end{aligned} \tag{1.33}$$

Thus

$$\begin{aligned} (\partial_t - L)\Psi &= 2\chi\frac{(1-\theta^2)}{\alpha}\rho^{-q} - q\frac{(1-\theta^2)}{\alpha}\chi^2\rho^{-q-1} \\ &\quad -\rho^{-q}\Big(8a^{ij}x_i x_j - 4\chi a^{ij}\delta_{ij} - 4\chi b^k x_k\Big). \end{aligned} \tag{1.34}$$

Let $\xi = \chi\rho^{-1}$. Regrouping

$$\begin{aligned} (\partial_t - L)\Psi &\le \rho^{1-q}\Big\{ - q\frac{(1-\theta^2)}{\alpha}\xi^2 - 8\Lambda^{-1}\rho^{-1}|x|^2 \\ &\quad + \Big(2\frac{(1-\theta^2)}{\alpha} + 4a^{ij}\delta_{ij} + 4b^k x_k\Big)\xi\Big\}. \end{aligned} \tag{1.35}$$

In $\tilde{Q}$, we have $|x|^2 = \rho - \chi$ and

$$-8\Lambda^{-1}\rho^{-1}|x|^2 = 8\Lambda^{-1}\xi - 8\Lambda^{-1}. \tag{1.36}$$

Therefore

$$(\partial_t - L)\Psi \le \rho^{1-q}\Big(- q\frac{(1-\theta^2)}{\alpha}\xi^2 - 8\Lambda^{-1} + C\xi\Big). \tag{1.37}$$

By taking $q \gg 1$ large and using $ab \le \frac{a^2}{2} + \frac{b^2}{2}$ on $C\xi$, we obtain

$$(\partial_t - L)\Psi \le 0. \tag{1.38}$$

We now take the following barrier function in $\tilde{Q}$

$$v = \kappa(\theta R)^{2q-4}\Psi. \tag{1.39}$$

We have $(\partial_t - L)v \le 0$, and at $t = -\alpha R^2$,

$$\begin{aligned} v(x, -\alpha R^2) &= \kappa(\theta R)^{2q-4}(\theta^2 R^2 - |x|^2)^2(\theta R)^{-2q} \\ &\le \kappa(\theta R)^{2q-4}(\theta^2 R^2)^2(\theta R)^{-2q} \\ &= \kappa \\ &\le u(x, -\alpha R^2). \end{aligned} \tag{1.40}$$

When $|x|^2 = \rho$, $v = 0 \leq u$. By the maximum principle,

$$v \leq u, \quad (x,t) \in \tilde{Q}. \tag{1.41}$$

(Actually, we used a version of the maximum principle where the parabolic domain is of the form $\Omega_t \times (-r, 0)$ for evolving domains Ω_t. We should verify that the proof of the weak maximum principle for $\Omega \times (0,T)$ goes through in this case.)

At $t = 0$ this means

$$\kappa(\theta R)^{2q-4}(R^2 - |x|^2)^2 (R)^{-2q} \leq u(x,0). \tag{1.42}$$

If we restrict x to $|x| < \frac{R}{2}$, we obtain the estimate. □

This prop-up lemma will be used again to derive the Krylov-Safonov estimate. For now, we note that the strong maximum principle follows as a consequence.

Theorem 4 *Let $T > 0$ and $\Omega \subset \mathbf{R}^n$ be a bounded, connected open set with smooth boundary. Suppose the coefficients of L, defined by (1.22) in $\Omega \times [0,T]$, satisfy (1.3) for some $\Lambda > 0$. Let $u : \overline{\Omega} \times [0,T] \to \mathbf{R}$ be a smooth function.*

- *If $(\partial_t - L)u \leq 0$ in $\Omega \times [0,T]$, and u attains a maximum at $(x_0, t_0) \in \Omega \times (0,T)$, then u is constant on $\Omega \times (0, t_0)$.*
- *If $(\partial_t - L)u \geq 0$ in $\Omega \times [0,T]$, and u attains a minimum at $(x_0, t_0) \in \Omega \times (0,T)$, then u is constant on $\Omega \times (0, t_0)$.*

Proof: We prove the first statement. Let $M = \sup_{\overline{\Omega} \times [0,T]} u$ and suppose $u(x_0, t_0) = M$ with $t_0 > 0$ and $x_0 \in \Omega$. Let $(\hat{x}, \hat{t})$ be a point in $\Omega \times (0, t_0)$. We need to show that $u(\hat{x}, \hat{t}) = M$.

Suppose on the contrary that $u(\hat{x}, \hat{t}) < M$. Consider the space-time path $\gamma : [0,1] \to \Omega \times [0,T]$ connecting $(\hat{x}, \hat{t})$ to (x_0, t_0) given by

$$\gamma(s) = (c(s), \hat{t} + s t_0), \tag{1.43}$$

where $c : [0,1] \to \Omega$ is a path in space connecting $\hat{x}$ to x_0. We claim the function $f : [0,1] \to \mathbf{R}$ given by

$$f(s) = M - u(\gamma(s)), \tag{1.44}$$

which satisfies $f(0) > 0$, stays greater than zero for all $s \in [0,1]$. This would imply $f(1) > 0$ which implies $u(x_0, t_0) < M$ which is a contradiction.

Indeed, suppose $f(s_0) = M - u(\gamma(s_0)) = 0$ at a first time s_0. Then $f(s) > 0$ for all $s < s_0$. By Lemma 1, $M - u > 0$ propagates forward in time, so we cannot have $M - u = 0$ at $\gamma(s_0)$. Therefore $f(s) > 0$ for all $s \in [0, 1]$, which completes the proof. $\square$

2 Parabolic Schauder Estimates

2.1 Parabolic Hölder spaces

The reference for this section is Krylov [6].

For local estimates, the basic set is the parabolic cylinder

$$Q_r = B_r \times \{-r^2 < t \leq 0\}. \tag{2.1}$$

We will also sometimes write Q for any set the form $Q = \Omega \times I$, where $\Omega \subseteq \mathbf{R}^n$ is a domain and I is an interval in time. We denote points in Q by $p = (x, t)$, and introduce the metric

$$|p| = |x| + |t|^{1/2}. \tag{2.2}$$

This metric is defined so that we have the scaling property $|(\lambda x, \lambda^2 t)| = \lambda|(x, t)|$.

We will use D for spacial derivatives, and ∂_t for time derivatives. For example,

$$D_i u = \frac{\partial u}{\partial x^i}. \tag{2.3}$$

We will also sometimes use the notation $\dot{u} = \partial_t u$.

Let $u : Q \to \mathbf{R}$ be a function. The parabolic Hölder semi-norm is defined as

$$[u]_{\delta,\delta/2;Q} = \sup_{p \neq q \in Q} \frac{|u(p) - u(q)|}{|p - q|^\delta}. \tag{2.4}$$

The parabolic Hölder norm is defined as

$$\|u\|_{\delta,\delta/2;Q} = \|u\|_{L^\infty(Q)} + [u]_{\delta,\delta/2;Q}. \tag{2.5}$$

We will also use the second order semi-norm

$$[u]_{2+\delta,1+\delta/2;Q} = [\partial_t u]_{\delta,\delta/2;Q} + [D^2 u]_{\delta,\delta/2;Q}, \tag{2.6}$$

where

$$[D^2 u]_{\delta,\delta/2;Q} = \sum_{i,j} [D_i D_j u]_{\delta,\delta/2;Q}. \tag{2.7}$$

The second order Hölder norm is then

$$\begin{aligned} \|u\|_{2+\delta,1+\delta/2;Q} &= \|u\|_{L^\infty(Q)} + \|\partial_t u\|_{L^\infty(Q)} + \|Du\|_{L^\infty(Q)} \\ &\quad + \|D^2 u\|_{L^\infty(Q)} + [u]_{2+\delta,1+\delta/2;Q}. \end{aligned} \tag{2.8}$$

Since these norms are used to study parabolic equations, the idea is that one derivative in time is worth two in space. So for example, u_{xt} does not appear in the norm $\|u\|_{2+\delta,1+\delta/2;Q}$ because in terms of derivatives it is worth $1+2$.

We denote by $C^{\delta,\delta/2}(Q)$, $C^{2+\delta,1+\delta/2}(Q)$ the space of functions on Q which are bounded in the $\|\cdot\|_{\delta,\delta/2;Q}$, $\|\cdot\|_{2+\delta,1+\delta/2;Q}$ norm.

Proposition 1 *The spaces $C^{\delta,\delta/2}(Q)$ and $C^{2+\delta,1+\delta/2}(Q)$ are Banach spaces.*

Proof: See [6].

Proposition 2 *For two functions $f, g \in C^{\delta,\delta/2}(Q)$, we can estimate*

$$\|fg\|_{\delta,\delta/2;Q} \leq \|f\|_\infty \, [g]_{\delta,\delta/2;Q} + \|g\|_\infty \, [f]_{\delta,\delta/2;Q}. \tag{2.9}$$

Consequently,

$$\|fg\|_{\delta,\delta/2;Q} \leq \|f\|_{\delta,\delta/2;Q} \, \|g\|_{\delta,\delta/2;Q}. \tag{2.10}$$

Proof: This follows from

$$\begin{aligned} & |f(p)g(p) - f(q)g(q)| \\ = \; & |f(p)g(p) - f(p)g(q) + f(p)g(q) - f(q)g(q)| \\ \leq \; & |f(p)||g(p) - g(q)| + |g(q)||f(p) - f(q)|, \end{aligned} \tag{2.11}$$

for any $p, q \in Q$. □

Next, we note that to check parabolic Hölder continuity, we can check continuity in space and continuity in time seperately.

Lemma 2 *Define*

$$[u]'_{\delta,\delta/2;Q} = \sup_{(x,t),(y,t)\in Q} \frac{|u(x,t) - u(y,t)|}{|x-y|^\delta} + \sup_{(x,t),(x,s)\in Q} \frac{|u(x,t) - u(x,s)|}{|t-s|^{\delta/2}}. \tag{2.12}$$

Then $[u]_{\delta,\delta/2;Q}$ and $[u]'_{\delta,\delta/2;Q}$ are equivalent, and

$$2^{-1}[u]'_{\delta,\delta/2;Q} \leq [u]_{\delta,\delta/2;Q} \leq 2[u]'_{\delta,\delta/2;Q}. \tag{2.13}$$

Proof: Clearly, we have

$$\frac{|u(x,t)-u(y,t)|}{|x-y|^{\delta}}+\frac{|u(x,t)-u(x,s)|}{|t-s|^{\delta/2}}\leq 2\sup_{p\neq q\in Q}\frac{|u(p)-u(q)|}{|p-q|^{\delta}}, \quad (2.14)$$

which shows one side of the inequality. For the other, we recall that $0<\delta<1$ and use concavity of $s\to s^{\delta}$,

$$(|x|+|t|^{1/2})^{\delta}\geq\frac{2^{\delta}}{2}(|x|^{\delta}+|t|^{\delta/2}). \quad (2.15)$$

Therefore

$$\begin{aligned}\frac{|u(p)-u(q)|}{|p-q|^{\delta}} &\leq 2\frac{|u(x,t)-u(y,s)|}{|x-y|^{\delta}+|t-s|^{\delta/2}}\\ &\leq 2\frac{|u(x,t)-u(y,t)|}{|x-y|^{\delta}+|t-s|^{\delta/2}}+2\frac{|u(y,t)-u(y,s)|}{|x-y|^{\delta}+|t-s|^{\delta/2}}\\ &\leq 2\sup_{(x,t),(y,t)\in Q}\frac{|u(x,t)-u(y,t)|}{|x-y|^{\delta}}\\ &\quad+2\sup_{(x,t),(x,s)\in Q}\frac{|u(x,t)-u(x,s)|}{|t-s|^{\delta/2}}. \end{aligned} \quad (2.16)$$

This completes the proof. □

Next, we state a useful compactness property of Hölder spaces.

Lemma 3 *Let $Q\subset\mathbf{R}^{n+1}$ be a bounded domain. Suppose $u_n\in C^{\delta,\delta/2}(Q)$ is a sequence of functions such that*

$$\|u_n\|_{\delta,\delta/2;Q}\leq C, \quad (2.17)$$

for some uniform constant $C>0$. Let $0<\eta<\delta$. There exists a subsequence u_{n_k} and a function $u\in C^{\delta,\delta/2}(Q)$ such that (u_{n_k}) converges in $C^{\eta,\eta/2}(Q)$ to u.

Proof: By the Arzela-Ascoli theorem, there exists a continuous function u such that $u_{n_k}\to u$ uniformly. It follows then that

$$|u(p)-u(q)|=\lim|u_{n_k}(p)-u_{n_k}(q)|\leq C|p-q|^{\delta},\quad p,q\in Q. \quad (2.18)$$

Therefore $u\in C^{\delta,\delta/2}$. Let $v_k=u_{n_k}-u$. For $p,q\in Q$, we may expand

$$\frac{|v_k(p)-v_k(q)|}{|p-q|^{\eta}}=\left(\frac{|v_k(p)-v_k(q)|}{|p-q|^{\delta}}\,|v_k(p)-v_k(q)|^{\frac{\delta}{\eta}-1}\right)^{\frac{\eta}{\delta}}. \quad (2.19)$$

It follows that

$$[v_k]_{\eta,\eta/2;Q} \leq 2[v_k]_{\delta,\delta/2;Q}^{\frac{\eta}{\delta}} \|v_k\|_{L^\infty(Q)}^{1-\frac{\eta}{\delta}}. \tag{2.20}$$

Since $v_k \to 0$ uniformly and $[v_k]_{\delta,\delta/2;Q}$ is bounded, we see that

$$\|u_{n_k} - u\|_{\eta,\eta/2;Q} \to 0 \tag{2.21}$$

as $k \to \infty$. □

Lemma 4 *Let $Q \subset \mathbf{R}^{n+1}$ be a bounded domain. Suppose $u_n \in C^{2+\delta,1+\delta/2}(Q)$ is a sequence of functions such that*

$$\|u_n\|_{2+\delta,1+\delta/2;Q} \leq C, \tag{2.22}$$

for some uniform constant $C > 0$. Let $0 < \eta < \delta$. There exists a subsequence u_{n_k} and a function $u \in C^{2+\delta,1+\delta/2}(Q)$ such that (u_{n_k}), $\partial_t(u_{n_k})$, $D(u_{n_k})$, $D^2(u_{n_k})$ converge in $C^{\eta,\eta/2}(Q)$ to u, $\partial_t u$, Du, $D^2 u$.

Proof: This follows from repeated applications of the Arzela-Ascoli theorem and the previous argument. □

We note the following scaling properties. Let $u : Q_1 \to \mathbf{R}$ be rescaled to $u_\lambda = u(\lambda x, \lambda^2 t)$, which is defined on $Q_{\lambda^{-1}}$. Then

$$[u_\lambda]_{\delta,\delta/2;Q_{\lambda^{-1}}} = \lambda^\delta [u]_{\delta,\delta/2;Q_1}, \tag{2.23}$$

$$[u_\lambda]_{2+\delta,1+\delta/2;Q_{\lambda^{-1}}} = \lambda^{2+\delta} [u]_{2+\delta,1+\delta/2;Q_1}. \tag{2.24}$$

For the last part of this section, we study how to interpolate through weaker norms. As before, the general principle here is that one derivative in time is worth two in space.

Proposition 3 *Let $Q = \mathbf{R}^n \times (-\infty, 0]$ and $0 < \delta < 1$. There exists a constant $C(n)$ such that for any $u \in C^{2+\delta,1+\delta/2}(Q)$ and $\varepsilon > 0$, then*

$$\begin{aligned}
[u]_{\delta,\delta/2;Q} &\leq \varepsilon [u]_{2+\delta,1+\delta/2;Q} + C\varepsilon^{-\delta/2} \|u\|_{L^\infty(Q)}, & (2.25)\\
\|Du\|_{L^\infty(Q)} &\leq \varepsilon [u]_{2+\delta,1+\delta/2;Q} + C\varepsilon^{-1/(1+\delta)} \|u\|_{L^\infty(Q)}, & (2.26)\\
\|D^2 u\|_{L^\infty(Q)} &\leq \varepsilon [u]_{2+\delta,1+\delta/2;Q} + C\varepsilon^{-2/\delta} \|u\|_{L^\infty(Q)}, & (2.27)\\
\|\partial_t u\|_{L^\infty(Q)} &\leq \varepsilon [u]_{2+\delta,1+\delta/2;Q} + C\varepsilon^{-2/\delta} \|u\|_{L^\infty(Q)}, & (2.28)\\
[Du]_{\delta,\delta/2;Q} &\leq \varepsilon [u]_{2+\delta,1+\delta/2;Q} + C\varepsilon^{-(1+\delta)} \|u\|_{L^\infty(Q)}. & (2.29)
\end{aligned}$$

Proof: The first observation is to note that it suffices to prove these inequalities with $\varepsilon = 1$ and a constant depending on n in front of $[u]_{2+\delta,1+\delta/2;Q}$. The estimate with $\varepsilon > 0$ then follows from considering $u_\lambda = u(\lambda x, \lambda^2 t)$ and scaling the norms.

We introduce the notation

$$\delta_j u(x,t) = u(x + e_j, t) - u(x,t), \tag{2.30}$$

where e_j is the j-th canonical unit vector.

We start by proving (2.27). Let $(x,t) \in Q$. Repeated applications of the mean value theorem give the existence of points $y_0 \in B_1(x)$ and $y_1 \in B_2(x)$ such that

$$\begin{aligned} \delta_i \delta_j u(x,t) &= D_i(\delta_j u)(y_0,t) \\ &= D_i u(y_0 + e_j, t) - D_i u(y_0, t) \\ &= D_j D_i u(y_1, t). \end{aligned} \tag{2.31}$$

In particular, there exists $y_1 \in B_2(x)$ such that

$$|D_i D_j u(y_1,t)| \leq 4\|u\|_{L^\infty(Q)}. \tag{2.32}$$

Then

$$\begin{aligned} |D_i D_j u(x,t)| &\leq |D_i D_j u(y_1,t)| + |D_i D_j u(y_1,t) - D_i D_j u(x,t)| \\ &\leq 4\|u\|_{L^\infty(Q)} + 2[u]_{2+\delta,1+\delta/2;Q}, \end{aligned} \tag{2.33}$$

which proves (2.27) by the scaling argument discussed above. To be explicit, rescaling u gives the estimate

$$\lambda^2 \|D_i D_j u\|_{L^\infty} \leq 2\lambda^{2+\delta}[u]_{2+\delta,1+\delta/2;Q} + 4\|u\|_{L^\infty}, \tag{2.34}$$

for all $\lambda > 0$.

Next, we show (2.26). Let $(x,t) \in Q$. By the mean value theorem, there exists $y \in B_1(x)$ such that

$$u(x + e_j, t) - u(x,t) = D_j u(y,t). \tag{2.35}$$

Thus

$$|D_j u(y,t)| \leq 2\|u\|_{L^\infty(Q)}. \tag{2.36}$$

For (x,t), we may estimate

$$\begin{aligned} |D_j u(x,t)| &\leq |D_j u(y,t)| + |D_j u(y,t) - D_j u(x,t)| \\ &\leq 2\|u\|_{L^\infty(Q)} + \|D^2 u\|_{L^\infty(Q)}. \end{aligned} \tag{2.37}$$

Applying (2.27), we obtain (2.26).

Next, we show (2.28). Let $(x,t) \in Q$. Then

$$\begin{aligned} |\partial_t u(x,t)| &\leq |\partial_t u(x,t) - (u(x,t) - u(x,t-1))| \\ &\quad + |u(x,t) - u(x,t-1)|. \end{aligned} \tag{2.38}$$

By the mean value theorem, there exists $\theta \in (0,1)$ such that

$$|\partial_t u(x,t)| \leq |\partial_t u(x,t) - \partial_t u(x,t-\theta)| + 2\|u\|_{L^\infty(Q)}. \tag{2.39}$$

Therefore

$$|\partial_t u| \leq [\partial_t u]_{\delta,\delta/2;Q} + 2\|u\|_{L^\infty(Q)}, \tag{2.40}$$

which proves (2.28).

Next, we show (2.25). By Lemma 2, we can consider variations in time and space separately. Let $(x,t),(y,t) \in Q$. If $|x-y| \geq 1$, then

$$\frac{|u(x,t) - u(y,t)|}{|x-y|^\delta} \leq 2\|u\|_{L^\infty(Q)}, \tag{2.41}$$

and if $|x-y| \leq 1$, then

$$\frac{|u(x,t) - u(y,t)|}{|x-y|^\delta} \leq \|Du\|_{L^\infty(Q)}, \tag{2.42}$$

by the mean value theorem. A similar argument holds for variations in time. Therefore

$$[u]_{\delta,\delta/2;Q} \leq C(\|Du\|_{L^\infty(Q)} + \|\partial_t u\|_{L^\infty(Q)} + \|u\|_{L^\infty(Q)}), \tag{2.43}$$

where C is an absolute constant. This proves (2.25) after applying (2.28) and (2.26).

Lastly, we show (2.29). Variations in space can be handled as before:

$$\frac{|D_i u(x,t) - D_i u(y,t)|}{|x-y|^\delta} \leq \|D^2 u\|_{L^\infty(Q)} + 2\|Du\|_{L^\infty(Q)}. \tag{2.44}$$

Next, we deal with variations in time. Let $(x,t),(x,s) \in Q$, and denote $d = |t-s|^{1/2}$. If $d \geq 1$, we directly have

$$\frac{|D_i u(x,t) - D_i u(x,s)|}{|t-s|^{\delta/2}} \leq 2\|Du\|_{L^\infty}. \tag{2.45}$$

If $0 < d < 1$, we instead start with the estimate

$$\begin{aligned}
& |D_i u(x,t) - D_i u(x,s)| \\
\leq\ & |D_i u(x,t) - d^{-1}\{u(x,t) - u(x+de_i,t)\}| \\
& + d^{-1}|\{u(x,t) - u(x+de_i,t)\} - \{u(x,s) - u(x+de_i,s)\}| \\
& + |d^{-1}\{u(x,s) - u(x+de_i,s)\} - D_i u(x,s)|. \qquad (2.46)
\end{aligned}$$

There exists $0 \leq \theta_1, \theta_2 \leq 1$ such that

$$\begin{aligned}
& |D_i u(x,t) - D_i u(x,s)| \\
\leq\ & |D_i u(x,t) - D_i u(x+\theta_1 de_i,t)| + d^{-1}|u(x,t) - u(x,s)| \\
& + d^{-1}|u(x+de_i,t) - u(x+de_i,s)| \\
& + |D_i u(x,s) - D_i u(x+\theta_2 de_i,s)| \\
\leq\ & 2d\|D^2 u\|_{L^\infty} + 2d\|\partial_t u\|_{L^\infty}. \qquad (2.47)
\end{aligned}$$

Since $0 < d < 1$, we get

$$\frac{|D_i u(x,t) - D_i u(x,s)|}{|t-s|^{\delta/2}} \leq 2(\|D^2 u\|_{L^\infty} + \|\partial_t u\|_{L^\infty}). \qquad (2.48)$$

Collecting everything, we have obtained the estimate

$$\|Du\|_{\delta,\delta/2;Q} \leq C(\|Du\|_{L^\infty(Q)} + \|D^2 u\|_{L^\infty(Q)} + \|\partial_t u\|_{L^\infty(Q)}), \quad (2.49)$$

As before, a scaling argument completes the proof. □

2.2 Interior derivative estimates for the heat equation

Recall the notation $Q_R = B_R(0) \times (-R^2, 0]$.

Proposition 4 *Let $R > 0$ and suppose $(\partial_t - \Delta)u = 0$ in Q_R. Then*

$$\sup_{Q_{R/2}} |Du| \leq \frac{C(n)}{R}\|u\|_{L^\infty(Q_R)}. \qquad (2.50)$$

Proof: Consider $v(x,t) = u(Rx, R^2 t)$, which is a function defined on Q_1. This function satisfies $(\partial_t - \Delta)v = 0$. Proving

$$\sup_{Q_{1/2}} |Dv| \leq C(n)\|v\|_{L^\infty(Q_1)}, \qquad (2.51)$$

would imply the desired estimate on u. Therefore, we can assume that $R = 1$ in the statement of the proposition.

Let η be a space-time bump function compactly supported in $B_1 \times (-1, 1)$ and identically 1 in a neighbourhood of $Q_{1/2}$. Consider the test function

$$G = \eta^2|Du|^2 + Au^2, \tag{2.52}$$

where $A > 0$ is a constant to be determined. We compute

$$\partial_t G = 2(\partial_t\eta)\eta|Du|^2 + 2\eta^2\langle D\partial_t u, Du\rangle + 2Au\partial_t u. \tag{2.53}$$

Differentiating twice in space gives

$$\begin{aligned} \Delta G &= 2(\Delta\eta)\eta|Du|^2 + 2|D\eta|^2|Du|^2 + 4\eta\langle D\eta, D|Du|^2\rangle \\ &\quad + 2\eta^2|DDu|^2 + 2\eta^2\langle D\Delta u, u\rangle \\ &\quad + 2A|Du|^2 + 2Au\Delta u. \end{aligned} \tag{2.54}$$

The evolution of G is then

$$\begin{aligned} (\partial_t - \Delta)G &= 2\eta|Du|^2(\partial_t - \Delta)\eta + 2\eta^2\langle D(\partial_t - \Delta)u, Du\rangle \\ &\quad + 2Au(\partial_t - \Delta)u - 2|D\eta|^2|Du|^2 \\ &\quad - 4\eta\langle D\eta, D|Du|^2\rangle - 2\eta^2|DDu|^2 \\ &\quad - 2A|Du|^2. \end{aligned} \tag{2.55}$$

Therefore

$$\begin{aligned} (\partial_t - \Delta)G &\le \Big[(2\eta)(\partial_t - \Delta)\eta - 2|D\eta|^2 - 2A\Big]|Du|^2 \\ &\quad + 8\eta|D\eta||DDu||Du| - 2\eta^2|DDu|^2. \end{aligned} \tag{2.56}$$

Using $2ab \le a^2 + b^2$, we estimate

$$\begin{aligned} 8\eta|D\eta||DDu||Du| &= 2(\sqrt{2}\eta|DDu|)(2\sqrt{2}|D\eta||Du|) \\ &\le 2\eta^2|DDu|^2 + 8|D\eta|^2|Du|^2. \end{aligned} \tag{2.57}$$

Hence

$$(\partial_t - \Delta)G \le \Big[(2\eta)(\partial_t - \Delta)\eta + 6|D\eta|^2 - 2A\Big]|Du|^2 \le 0, \tag{2.58}$$

for $A \gg 1$ depending on η. By the maximum principle, G attains its maximum either at the initial time or on the spacial boundary. In both cases, $\eta = 0$ at the maximum of G. Hence

$$G(x,t) \le A\|u\|^2_{L^\infty(Q)}. \tag{2.59}$$

For any $(\hat{x}, \hat{t}) \in Q_{1/2}$, we have $G(\hat{x}, \hat{t}) = |Du|^2(\hat{x}, \hat{t}) + Au^2(\hat{x}, \hat{t})$ since the cutoff function η is equal to 1 there. Therefore

$$|Du|^2(\hat{x}, \hat{t}) \leq G(\hat{x}, \hat{t}) \leq A\|u\|_{L^\infty(Q)}^2, \tag{2.60}$$

and the estimate follows. □

Proposition 5 *Suppose $u : Q_R \to \mathbf{R}$ is a smooth function satisfying $(\partial_t - \Delta)u = 0$. Then*

$$\|\partial_t^k D^\ell u\|_{L^\infty(Q_{R/2})} \leq \frac{C(n, k, \ell)}{R^{\ell+2k}} \|u\|_{L^\infty(Q_R)}. \tag{2.61}$$

Proof: As before, by scaling we may assume $R = 1$. We already proved the statement for $k = 0$, $\ell = 1$. Since $(\partial_t - \Delta)Du = 0$, applying this estimate leads to

$$\sup_{Q_{1/4}} |D^2 u| \leq C\|Du\|_{L^\infty(Q_{1/2})} \leq C\|u\|_{L^\infty(Q_1)}, \tag{2.62}$$

where C here is a generic constant depending on n, k, ℓ which may change line-by-line. By covering $Q_{1/2}$ with cylinders $Q_{1/4}(x, t)$, we obtain

$$\sup_{Q_{1/2}} |D^2 u| \leq C\|u\|_{L^\infty(Q_1)}, \tag{2.63}$$

which proves the statement for $k = 0$ and $\ell = 2$. For $k = 0$ and arbitrary ℓ, we use that $(\partial_t - \Delta)D^\ell u = 0$ and repeat the argument. For $k \geq 1$, we note that since

$$\partial_t^k D^\ell u = (\Delta)^k D^\ell u, \tag{2.64}$$

time derivatives can be converted to spacial derivatives. □

Proposition 6 *Suppose $u : \mathbf{R}^n \times (-\infty, 0) \to \mathbf{R}$ is a bounded smooth function satisfying $(\partial_t - \Delta)u = 0$. Then u is a constant function.*

Proof: Let $(x, t) \in \mathbf{R}^n \times (-\infty, 0)$. Let $R_0 > 0$ be such that $(x, t) \in Q_{R_0}$. Then for all $R \geq 2R_0$,

$$|Du|(x, t) \leq \frac{C(n)}{R}\|u\|_{L^\infty}. \tag{2.65}$$

Letting $R \to \infty$ shows that $Du \equiv 0$. Thus u is constant in space, and since $\partial_t u = \Delta u$, it follows that u is also constant in time. □

2.3 Schauder estimates for the heat equation

We start by proving Schauder estimates for the heat equation, following the blow-up argument of L. Simon [11]. See also e.g. [4, 12].

Proposition 7 *Let* $0 < \delta < 1$. *Suppose* $u \in C^{2+\delta,1+\delta/2}(\mathbf{R}^n \times (-\infty, 0])$ *satisfies*

$$(\partial_t - \Delta)u = f, \tag{2.66}$$

where $f \in C^{\delta,\delta/2}(\mathbf{R}^n \times (-\infty, 0])$. *Then*

$$[D^2 u]_{\delta,\delta/2} \leq C\|f\|_{\delta,\delta/2}, \tag{2.67}$$

where C *depends on* n *and* δ.

Proof: We prove this by contradiction. Suppose the statement is false, and there exists a sequence of function $u_k \in C^{2+\delta,1+\delta/2}$ satisfying $(\partial_t - \Delta)u_k = f_k$ with $f_k \in C^{\delta,\delta/2}$ such that

$$[D^2 u_k]_{\delta,\delta/2} \geq k\|f_k\|_{\delta,\delta/2}. \tag{2.68}$$

Replacing u_k with

$$\frac{u_k}{\|f_k\|_{\delta,\delta/2}}, \tag{2.69}$$

we obtain a sequence satisfying

$$(\partial_t - \Delta)u_k = f_k, \tag{2.70}$$

$$\|f_k\|_{\delta,\delta/2} \leq 1, \tag{2.71}$$

and

$$[D^2 u_k]_{\delta,\delta/2} = M_k \to \infty. \tag{2.72}$$

After taking a subsequence of u_k and relabeling, there is a direction $D_i D_j$ such that a sequence of points $p_k, q_k \in \mathbf{R}^n \times (-\infty, 0]$ satisfies

$$\frac{|D_i D_j u_k(p_k) - D_i D_j u_k(q_k)|}{|p_k - q_k|^\delta} \geq c_n M_k. \tag{2.73}$$

Here c_n is a small constant depending on n. We will also sometimes use the notation $p_k = (x_k, t_k)$ and $q_k = (y_k, s_k)$. We let

$$\lambda_k = |p_k - q_k| = |x_k - y_k| + |t_k - s_k|^{1/2}. \tag{2.74}$$

Consider the rescaled functions $\hat{u}_k : \mathbf{R}^n \times (-\infty, 0] \to \mathbf{R}$ defined by

$$\hat{u}_k(x,t) = M_k^{-1} \lambda_k^{-2-\delta} u_k(x_k + \lambda_k x, t_k + \lambda_k^2 t). \tag{2.75}$$

By the scaling property of Hölder norms (2.24), these functions satisfy

$$[D^2\hat{u}_k]_{\delta,\delta/2} \le M_k^{-1}[D^2 u_k]_{\delta,\delta/2}. \tag{2.76}$$

We shift $\hat{u}_k$ to define

$$\begin{aligned} &\hat{v}_k(x,t) \qquad (2.77)\\ = \;& \hat{u}_k(x,t) - \hat{u}_k(0) - D_i\hat{u}_k(0)x^i - \partial_t\hat{u}_k(0)t - \frac{1}{2}D_pD_q\hat{u}_k(0)x^px^q. \end{aligned}$$

This sequence $\hat{v}_k : \mathbf{R}^n \times (-\infty, 0] \to \mathbf{R}$ satisfies

$$\hat{v}_k(0) = 0,\ D\hat{v}_k(0) = 0,\ \partial_t\hat{v}_k(0) = 0,\ D^2\hat{v}_k(0) = 0. \tag{2.78}$$

Furthermore, $\hat{v}_k$ satisfies

$$[D^2\hat{v}_k]_{\delta,\delta/2} \le 1, \tag{2.79}$$

and also satisfies the evolution equation

$$(\partial_t - \Delta)\hat{v}_k = M_k^{-1}\lambda_k^{-\delta}\Big[f_k(x_k + \lambda_k x, t_k + \lambda_k^2 t) - f_k(x_k, t_k)\Big]. \tag{2.80}$$

Let $R > 0$ be fixed. Since $D^2\hat{v}_k$ vanishes at the origin and has bounded Hölder norm, we may estimate

$$\|D^2\hat{v}_k\|_{Q_R} \le C(R). \tag{2.81}$$

Similarly, we may estimate the gradient and uniform norms of $\hat{v}_k$, and conclude

$$\|\hat{v}_k\|_{2+\delta,1+\delta/2;Q_R} \le C(R). \tag{2.82}$$

The bound on $\partial_t\hat{v}_k$ can be deduced by using the equation (2.80). It follows that there exists a limiting function $v_\infty \in C^{2+\delta,1+\delta/2}$ such that a subsequence $\hat{v}_k$ converges to v_∞ in $C^{2,1}(Q_R)$. Since R was arbitrary, the function is defined on the whole space

$$v_\infty : \mathbf{R}^n \times (-\infty, 0] \to \mathbf{R}. \tag{2.83}$$

Furthermore, the $C^{2,1}$ convergences gives

$$(\partial_t - \Delta)v_\infty = 0, \tag{2.84}$$

$$v_\infty(0) = 0,\ Dv_\infty(0) = 0,\ \partial_t v_\infty(0) = 0,\ D^2v_\infty(0) = 0. \tag{2.85}$$

Furthermore, we have

$$[D^2 v_\infty]_{\delta,\delta/2} \leq 1. \tag{2.86}$$

This implies

$$|D^2 v_\infty(p)| \leq |p|^\delta. \tag{2.87}$$

Since $(\partial_t - \Delta)v_\infty = 0$, we have that v_∞ is smooth by regularity of the heat equation. Furthermore, $(\partial_t - \Delta)D^2 v_\infty = 0$ and the interior estimates give

$$\sup_{Q_{R/2}} |DD^2 v_\infty| \leq \frac{C}{R}\|D^2 v_\infty\|_{L^\infty(Q_R)} \leq \frac{CR^\delta}{R}. \tag{2.88}$$

Letting $R \to \infty$, we see that $D^2 v_\infty$ is a constant. This is a contradiction, since a subsequence of

$$o_k = (\lambda_k^{-1}(y_k - x_k), \lambda_k^{-2}(s_k - t_k)) \tag{2.89}$$

converges to o_∞ with $|o_\infty| = 1$, and taking the limit of

$$|D_i D_j \hat{v}_k(0) - D_i D_j \hat{v}_k(o_k)| \tag{2.90}$$

while using (2.73) gives

$$|D_i D_j v_\infty(0) - D_i D_j v_\infty(o_\infty)| \geq c_n. \tag{2.91}$$

Here we used that $D^2 \hat{v}_k(o_k)$ converges to $D^2 v_\infty(o_\infty)$ by uniform convergence of $D^2 \hat{v}_k$ on compact sets (2.81). $\square$

Note: it follows that $[\partial_t u]_{\delta,\delta/2}$ also satisfies the same estimate (2.92), since $\partial_t u = \Delta u + f$. Thus

$$[u]_{2+\delta,1+\delta/2} \leq C\|f\|_{\delta,\delta/2}, \tag{2.92}$$

where as before C depends only on n and δ.

2.4 Interior Schauder estimates for linear parabolic equations

The reference for this section is Krylov [6].

We will now consider a linear operator of the form

$$L = a^{ij}(x,t)D_i D_j + b^i(x,t)D_i + c(x,t), \tag{2.93}$$

where a^{ij}, b^i, c are smooth functions and

$$a^{ij}(x,t) = a^{ji}(x,t). \tag{2.94}$$

First, we state a lemma.

Lemma 5 *Let $\delta \in (0,1)$. Suppose $u(x,t) \in C^{2+\delta,1+\delta/2}(\mathbf{R}^n \times (-\infty,0])$, and a^{ij} is a constant symmetric positive definite matrix such that*

$$\Lambda^{-1}|\xi|^2 \le a^{ij}\xi_i\xi_j \le \Lambda|\xi|^2, \quad \xi \in \mathbf{R}^n, \tag{2.95}$$

for some $\Lambda > 0$. Then

$$[u]_{2+\delta,1+\delta/2} \le C\|(\partial_t - a^{ij}D_iD_j)u\|_{\delta,\delta/2}, \tag{2.96}$$

where C depends on n, Λ, δ.

Proof: Let B be an orthogonal matrix such that $B^p{}_j a^{jk} B^q{}_k = \delta^{pq}$. Let $v(Bx,t) = u(x,t)$. Then

$$u_{ij} = B^p{}_i v_{pq} B^q{}_j, \tag{2.97}$$

hence $a^{ij}u_{ji} = \delta^{pq}v_{qp}$. Thus $(\partial_t - \Delta)v = (\partial_t - a^{ij}D_iD_j)u$ and we may apply (2.92) to v. $\square$

Next, we prove a lemma which allows varying coefficients in the operator L by using a technique sometimes called "freezing coefficients". This lemma has a strong assumption on the support of u, and is not yet very practical, but it is used as a intermediate step towards the interior Schauder estimates.

Lemma 6 *Let $\delta \in (0,1)$. Suppose L, as defined in (2.93), satisfies*

$$\begin{aligned} &\|a,b,c\|_{\delta,\delta/2} \le \Lambda, \\ &\Lambda^{-1}|\xi|^2 \le a^{ij}(x,t)\xi_i\xi_j \le \Lambda|\xi|^2, \quad \xi \in \mathbf{R}^n, \end{aligned} \tag{2.98}$$

for some $\Lambda > 0$, and where the coefficients a^{ij}, b^i, c are defined on Q_1. There exists $0 < \kappa < (1/2)$ depending on n, Λ, and δ, with the following property.

Let $u \in C^{2+\delta,1+\delta/2}(Q_1)$ be a function such that

$$\operatorname{supp} u(\cdot,t) \subset B_\kappa, \quad \text{for all } t \in (-1,0]. \tag{2.99}$$

Further suppose $u \equiv 0$ *for* $t \le -\kappa^2$. *If*

$$(\partial_t - L)u = f, \tag{2.100}$$

for $f \in C^{\delta,\delta/2}(Q_1)$, *then*

$$\|u\|_{2+\delta,1+\delta/2;Q_1} \le C(\|f\|_{\delta,\delta/2;Q_1} + \|u\|_{L^\infty(Q_1)}), \tag{2.101}$$

where C *depends on* n, Λ, δ, *and* κ.

Proof: Let u be as in the statement. We may identify u with a function defined on $\mathbf{R}^n \times (-\infty, 0]$ by extension by zero. We apply the previous lemma to obtain

$$[u]_{2+\delta,1+\delta/2} \le C_0\|(\partial_t - a^{ij}(0)D_iD_j)u\|_{\delta,\delta/2;Q_\kappa}. \tag{2.102}$$

Next, we write

$$\partial_t - a^{ij}(0)D_iD_j = (\partial_t - L) + (a^{ij} - a^{ij}(0))D_iD_j + b^kD_k + c. \tag{2.103}$$

By a standard estimate in Hölder spaces (Proposition 2),

$$\begin{aligned} & \|(\partial_t - a^{ij}(0)D_iD_j)u\|_{\delta,\delta/2;Q_\kappa} \\ \le\ & \|f\|_{\delta,\delta/2} + \|a^{ij} - a^{ij}(0)\|_{L^\infty(Q_\kappa)}\|D^2u\|_{\delta,\delta/2} \\ & +\|a^{ij} - a^{ij}(0)\|_{\delta,\delta/2;\,Q_\kappa}\|D^2u\|_\infty \\ & +\|b^k\|_{\delta,\delta/2}\|Du\|_{\delta,\delta/2} + \|c\|_{\delta,\delta/2}\|u\|_{\delta,\delta/2}. \end{aligned} \tag{2.104}$$

We may estimate

$$\|(a^{ij} - a^{ij}(0))\|_{L^\infty(Q_\kappa)}\|D^2u\|_{\delta,\delta/2;\,Q_\kappa} \le \Lambda(2\kappa)^\delta\|u\|_{2+\delta,1+\delta/2}. \tag{2.105}$$

Therefore

$$\begin{aligned} [u]_{2+\delta,1+\delta/2} \ \le\ & C_0\Lambda(2\kappa)^\delta\|u\|_{2+\delta,1+\delta/2} + C\|f\|_{\delta,\delta/2} \\ & +C(\|u\|_{C^{2,1}} + [Du]_{\delta,\delta/2} + [u]_{\delta,\delta/2}), \end{aligned} \tag{2.106}$$

where C denotes a generic constant depending on $n, \Lambda, \delta, \kappa$ which may change line by line. Choose κ such that $C_0\Lambda(2\kappa)^\delta < (1/4)$. Then

$$\begin{aligned} \|u\|_{2+\delta,1+\delta/2} \ \le\ & \frac{1}{4}\|u\|_{2+\delta,1+\delta/2} + C\|f\|_{\delta,\delta/2} \\ & +C(\|u\|_{C^{2,1}} + [Du]_{\delta,\delta/2} + [u]_{\delta,\delta/2}). \end{aligned} \tag{2.107}$$

Recall that we view u as defined on $\mathbf{R}^n \times (-\infty, 0]$. By the interpolation inequalities (Proposition 3), we can estimate

$$\begin{aligned} & C(\|u\|_{C^{2,1}} + [Du]_{\delta,\delta/2} + [u]_{\delta,\delta/2}) \\ \leq \ & \frac{1}{4}\|u\|_{2+\delta,1+\delta/2} + C\|u\|_{L^\infty}. \end{aligned} \tag{2.108}$$

Thus

$$\|u\|_{2+\delta,1+\delta/2} \leq \frac{1}{2}\|u\|_{2+\delta,1+\delta/2} + C\|f\|_{\delta,\delta/2} + C\|u\|_{L^\infty}, \tag{2.109}$$

and the estimate follows. $\square$

We now prove the interior Schauder estimate, which is the main result of this section.

Theorem 5 *Let $\delta \in (0,1)$ and $R > 0$. Suppose $u(x,t) \in C^{2+\delta,1+\delta/2}(Q_{2R})$ satisfies*

$$(\partial_t - L)u = f, \tag{2.110}$$

with $f \in C^{\delta,\delta/2}(Q_{2R})$, L as in (2.93), and

$$\|a,b,c\|_{\delta,\delta/2} \leq \Lambda, \quad \Lambda^{-1}|\xi|^2 \leq a^{ij}(x,t)\xi_i\xi_j \leq \Lambda|\xi|^2, \quad \xi \in \mathbf{R}^n, \tag{2.111}$$

for some $\Lambda > 0$. Then

$$\|u\|_{2+\delta,1+\delta/2;\,Q_R} \leq C(\|f\|_{\delta,\delta/2;\,Q_{2R}} + \|u\|_{L^\infty(Q_{2R})}), \tag{2.112}$$

where C depends on n, Λ, δ, R.

Proof: We prove the theorem when $R = 1$; the general case follows by rescaling. As before, C is a constant which may change line by line, but only depends on Λ, n, δ. Given the previous lemma, the proof of this estimate follows an iteration argument using cutoff functions whose support grows slightly at each step.

Let $(x_0, t_0) \in Q_1$. Let $0 < \kappa \leq (1/2)$ be as in the previous lemma. Define

$$P_\ell = B_{R_\ell}(x_0) \times \{-R_\ell^2 < t - t_0 \leq 0\}, \quad R_\ell = \kappa \sum_{i=1}^{\ell} 2^{-i}. \tag{2.113}$$

Then P_ℓ expands from $Q_{\kappa/2}(x_0,t_0)$ towards $Q_\kappa(x_0,t_0)$. Let ζ_ℓ be a cutoff function such that $\zeta_\ell \equiv 1$ in P_ℓ and

$$\operatorname{supp}\zeta_\ell \subset B_{R_{\ell+1}} \times (t_0 - R_{\ell+1}^2, t_0 + R_{\ell+1}^2). \tag{2.114}$$

Since

$$R_{\ell+1} - R_\ell = \kappa 2^{-(\ell+1)}, \tag{2.115}$$

the slope of ζ_ℓ goes like $\kappa^{-1}2^{\ell+1}$ in the spacial direction and $\kappa^{-2}2^{\ell+1}$ in the time direction. Let C, depending only on κ and n, be such that

$$\|\zeta_\ell\|_{2+\delta,1+\delta/2} \leq C2^{3\ell} = C\rho^{-\ell}, \quad \rho = 2^{-3}. \tag{2.116}$$

With this setup, $u\zeta_\ell$ is a function defined on $Q_1(x_0,t_0)$ with support contained in $B_\kappa(x_0)$ at each time and identically zero for $t \leq t_0 - \kappa^2$. We may also view $u\zeta_\ell$ as a function defined on $Q = \mathbf{R}^n \times (-\infty, t_0]$.

We are in a position to apply Lemma 6 (after translating the origin $(0,0)$ to (x_0,t_0)), hence

$$\|u\zeta_\ell\|_{2+\delta,1+\delta/2;Q} \leq C(\|(\partial_t - L)(u\zeta_\ell)\|_{\delta,\delta/2;Q} + \|u\zeta_\ell\|_{L^\infty(Q)}). \tag{2.117}$$

The evolution of $u\zeta_\ell$ is

$$\begin{aligned} & (\partial_t - L)(u\zeta_\ell) \\ = \ & f\zeta_\ell + u\partial_t\zeta_\ell - ub^iD_i\zeta_\ell - 2a^{ij}D_juD_i\zeta_\ell - ua^{ij}D_iD_j\zeta_\ell. \end{aligned} \tag{2.118}$$

Therefore, estimating the Hölder norms (Proposition 2) gives

$$\begin{aligned} \|u\zeta_\ell\|_{2+\delta,1+\delta/2;Q} \ \leq \ & C\rho^{-\ell}(\|u\|_{\delta,\delta/2;P_{\ell+1}} + \|Du\|_{\delta,\delta/2;P_{\ell+1}}) \\ & +C\rho^{-\ell}(\|f\|_{\delta,\delta/2;Q_2} + \|u\|_{L^\infty(Q_2)}). \end{aligned} \tag{2.119}$$

We now use the next cutoff $\zeta_{\ell+1}$ in the sequence.

$$\begin{aligned} \|u\zeta_\ell\|_{2+\delta,1+\delta/2;Q} \ \leq \ & C\rho^{-\ell}(\|u\zeta_{\ell+1}\|_{\delta,\delta/2;Q} + \|D(u\zeta_{\ell+1})\|_{\delta,\delta/2;Q}) \\ & +C\rho^{-\ell}(\|f\|_{\delta,\delta/2;Q_2} + \|u\|_{L^\infty(Q_2)}). \end{aligned} \tag{2.120}$$

By interpolation inequalities (Proposition 3), for any $0 < \varepsilon < 1$, there holds

$$\begin{aligned} \|u\zeta_\ell\|_{2+\delta,1+\delta/2;Q} \ \leq \ & C_1\rho^{-\ell}\varepsilon\|u\zeta_{\ell+1}\|_{2+\delta,1+\delta/2;Q} \\ & +C_2\varepsilon^{-2}\rho^{-\ell}(\|f\|_{\delta,\delta/2;Q_2} + \|u\|_{L^\infty(Q_2)}). \end{aligned} \tag{2.121}$$

Let $\varepsilon = \frac{1}{2C_1}\rho^3\rho^\ell$. We have

$$\begin{aligned} & \|u\zeta_\ell\|_{2+\delta,1+\delta/2;Q} \\ \leq \ & \frac{\rho^3}{2}\|u\zeta_{\ell+1}\|_{2+\delta,1+\delta/2;Q} + C\rho^{-3\ell}(\|f\|_{\delta,\delta/2} + \|u\|_{L^\infty}), \end{aligned} \tag{2.122}$$

for $C = C_2(2C_1\rho^{-3})^2$.

Iterating this estimate, we start with

$$\begin{aligned} &\|u\zeta_1\|_{2+\delta,1+\delta/2;Q} \\ \leq\ & \frac{\rho^3}{2}\|u\zeta_2\|_{2+\delta,1+\delta/2} + C\rho^{-3}(\|f\|_{\delta,\delta/2} + \|u\|_{L^\infty}), \end{aligned} \tag{2.123}$$

followed by

$$\begin{aligned} \|u\zeta_1\|_{2+\delta,1+\delta/2;Q} \leq\ & \left(\frac{\rho^3}{2}\right)^2 \|u\zeta_3\|_{2+\delta,1+\delta/2;Q} \\ & + (1+\frac{1}{2})C\rho^{-3}(\|f\|_{\delta,\delta/2} + \|u\|_{L^\infty}), \end{aligned} \tag{2.124}$$

etc, which gives

$$\begin{aligned} \|u\zeta_1\|_{2+\delta,1+\delta/2;Q} \leq\ & \left(\frac{\rho^3}{2}\right)^k \|u\zeta_{k+1}\|_{2+\delta,1+\delta/2;Q} \\ & + (1+\frac{1}{2}+\cdots+\frac{1}{2^{k-1}})C\rho^{-3}(\|f\|_{\delta,\delta/2} + \|u\|_{L^\infty}), \end{aligned} \tag{2.125}$$

for each integer k. We note that

$$\|u\zeta_k\|_{2+\delta,1+\delta/2;Q} \leq C\rho^{-k}(1 + \|u\|_{2+\delta,1+\delta/2;Q_2}), \tag{2.126}$$

hence upon taking the limit as $k \to \infty$, the first term in (2.125) goes to zero. Since $\zeta_1 \equiv 1$ on P_1, this gives the estimate

$$\|u\|_{2+\delta,1+\delta;Q_{\kappa/2}(x_0,t_0)} \leq C(\|f\|_{\delta,\delta/2} + \|u\|_{L^\infty}). \tag{2.127}$$

We can now cover Q_1 with cylinders $Q_{\kappa/2}(x_0,t_0)$ and obtain the interior Schauder estimate. □

Theorem 6 *Let $\delta \in (0,1)$, $R > 0$ and k be a non-negative integer. Suppose $u(x,t) \in C^{2+\delta,1+\delta/2}(Q_{2R})$ satisfies*

$$(\partial_t - L)u = f, \tag{2.128}$$

where L is as in (2.93) and

$$\begin{aligned} &\|D^\alpha a^{ij}\|_{\delta,\delta/2}, \|D^\alpha b^i\|_{\delta,\delta/2}, \|D^\alpha c\|_{\delta,\delta/2} \leq \Lambda, \\ &\Lambda^{-1}|\xi|^2 \leq a^{ij}\xi_i\xi_j \leq \Lambda|\xi|^2, \quad \xi \in \mathbf{R}^n, \\ &D^\alpha f \in C^{\delta,\delta/2}(Q_{2R}), \end{aligned} \tag{2.129}$$

for all multi-index $|\alpha| \le k$.

Then $D^\alpha u \in C^{2+\delta,1+\delta/2}(Q_R)$ and

$$\begin{aligned} &\|D^k u\|_{2+\delta,1+\delta/2;\,Q_R} \\ \le\ & C\left(\sum_{|\alpha|\le k}\|D^\alpha f\|_{\delta,\delta/2;\,Q_{2R}} + \|u\|_{L^\infty(Q_{2R})}\right), \end{aligned} \tag{2.130}$$

where C depends on n, δ, Λ, R.

Proof: We use the notation

$$\delta_{h,j}u(x,t) = \frac{1}{h}(u(x+he_j,t) - u(x,t)). \tag{2.131}$$

Let $v = \delta_{h,j}u$. Then v satisfies

$$\begin{aligned} (\partial_t - L)v(x,t) &= \delta_{h,j}f(x,t) - \delta_{h,j}a^{pq}(x,t)D_pD_qu(x+he_j,t) \\ &\quad -\delta_{h,j}b^p(x,t)D_pu(x+he_j,t) \\ &\quad -\delta_{h,j}c(x,t)u(x+he_j,t). \end{aligned} \tag{2.132}$$

Applying Schauder estimates and taking the limit as $h \to 0$ gives the estimate when $k = 1$, and we repeat this for higher order estimates. We omit the details. $\square$

3 Krylov-Safonov Estimates

3.1 Krylov-Tso ABP estimate

The reference for this section is [8].

Let $Q_1 = B_1(0) \times (-1, 0]$. For a function $u \in C^{2,1}(\bar{Q}_1)$, we denote the upper contact set by

$$\begin{aligned} \Gamma^+(u) &= \Big\{(x,t) : u(x,t) \geq 0, && (3.1) \\ &\quad u(y,s) \leq u(x,t) + Du(x,t)\cdot(y-x), \quad \text{for all } y \in B_1(0),\ s \leq t\Big\}. \end{aligned}$$

The first observation is

Lemma 7 *For $(x,t) \in \Gamma^+(u)$, we have $\partial_t u(x,t) \geq 0$ and $D^2 u(x,t) \leq 0$.*

Proof: Setting $y = x$ in the definition, we obtain $u(x,t) - u(x,s) \geq 0$ for $s \leq t$. Dividing by $(t-s)$ and taking the limit gives $\partial_t u(x,t) \geq 0$. Setting $s = t$ in the definition, we see that

$$h(y) = u(y,t) - u(x,t) - Du(x,t)\cdot(y-x), \tag{3.2}$$

is defined on $B_1(0)$, $h \leq 0$, and h attains a maximum at x. Therefore $D^2 h|_{y=x} \leq 0$, which translates to $D^2 u(x,t) \leq 0$. □

Next, we state a lemma which we will use in the proof of the local maximum principle.

Lemma 8 *Suppose $u : Q_1 \to \mathbf{R}$ satisfies $u|_{\partial B_1} = 0$. For $(x,t) \in \Gamma^+(u)$,*

$$|Du(x,t)| \leq \frac{|u(x,t)|}{d(x, \partial B)}. \tag{3.3}$$

Proof: Let $(x,t) \in \Gamma^+(u)$. Consider $y \in B_1(0)$ along the ray in $B_1(0)$ starting at x in the direction $-Du(x)$. Then

$$u(y,t) \leq u(x,t) - |Du(x,t)||y-x|. \tag{3.4}$$

Taking the limit as y tends to $y_\infty \in \partial B_1(0)$, we have that $u(y,t) \to 0$, and

$$0 \leq u(x,t) - |Du(x,t)||y_\infty - x|. \tag{3.5}$$

Since $d(x, \partial B) \leq |x - y_\infty|$, we obtain (3.3). $\square$

The key estimate of this section was first proved by Krylov [5] and Tso [10], building on prior work in the elliptic case by Aleksandrov-Bakelman-Pucci (ABP).

Proposition 8 *Suppose $u \in C^{2,1}(\bar{Q}_1)$ with $u|_{\partial B_1} \leq 0$ and $u|_{t=-1} \leq 0$. Then*

$$\sup_{Q_1} u \leq C(n) \left(\int_{\Gamma^+(u)} |\partial_t u \, \mathrm{det} D^2 u| \right)^{1/(n+1)}. \tag{3.6}$$

Proof: Suppose $u(x_0, t_0) = M$ attains a positive maximum. Define $\Phi : Q_1 \to \mathbf{R}^{n+1}$ by

$$\Phi(x,t) = \langle Du(x,t), u(x,t) - Du(x,t) \cdot (x - x_0) \rangle. \tag{3.7}$$

Direct computation gives

$$D\Phi = \begin{bmatrix} u_{ij} & u_{it} \\ -\sum_k u_{kj}(x-x_0)_k & \partial_t u - \sum_k u_{kt}(x-x_0)_k \end{bmatrix}. \tag{3.8}$$

To take the determinant, we may multiply the ith row by $(x - x_0)_i$ and add it to the last row. Then

$$\mathrm{det} D\Phi = \det \begin{bmatrix} u_{ij} & u_{it} \\ 0 & \partial_t u \end{bmatrix} \tag{3.9}$$

and so

$$|\mathrm{det} D\Phi| = |\partial_t u \, \mathrm{det} D^2 u|. \tag{3.10}$$

By the change of variables formula,

$$\int_{\Gamma^+(u)} |\partial_t u \, \mathrm{det} D^2 u| \geq \int_{\Phi(\Gamma^+(u))} 1. \tag{3.11}$$

(Since Φ is not a diffeomorphism, we have an inequality instead of equality. For more details on this step, see Lemma 1.4 in [1].)

We claim $P \subset \Phi(\Gamma^+(u))$, where

$$P = \left\{ (\xi, h) \in \mathbf{R}^{n+1} : |\xi| < \frac{h}{2},\ 0 < h < \frac{M}{2} \right\}. \tag{3.12}$$

Indeed, let $(\xi, h) \in P$. Consider the hyperplane $\ell : B_1 \to \mathbf{R}$ defined by

$$\ell(x) = h + \xi \cdot (x - x_0). \tag{3.13}$$

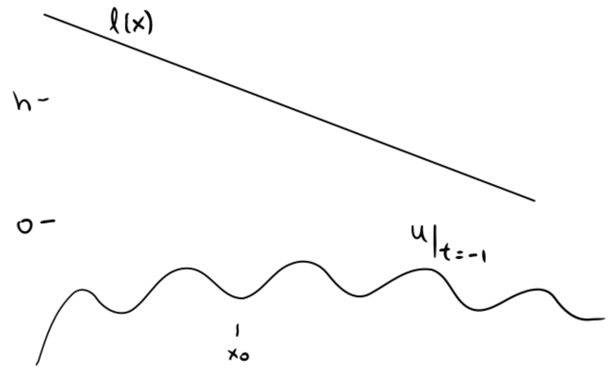
Then

$$\ell \geq -2|\xi| + h > 0, \tag{3.14}$$

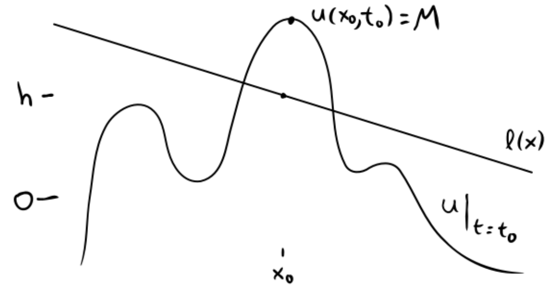
and furthermore $\ell(x_0) = h < u(x_0, t_0)$. It follows that

$$g(x,t) = u(x,t) - \ell(x) \tag{3.15}$$

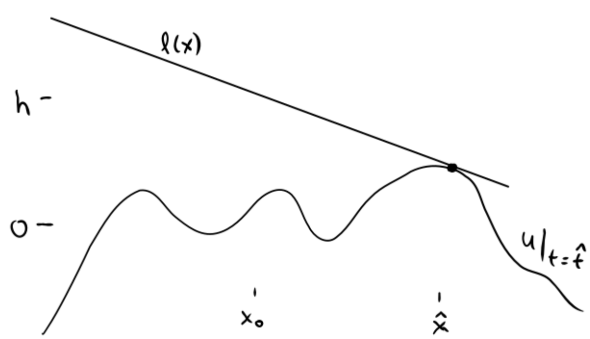
is such that $g|_{t=-1} < 0$ and $g|_{\partial B} < 0$,



but $g(x_0, t_0) > 0$.



Let $(\hat{x}, \hat{t})$ be the first time when $g(\hat{x}, \hat{t}) = 0$.



It follows that $\hat{x} \notin \partial B$, hence we have the critical equation $Dg(\hat{x}, \hat{t}) = 0$. This implies

$$\xi = Du(\hat{x}, \hat{t}), \quad h = u(\hat{x}, \hat{t}) - Du(\hat{x}, \hat{t}) \cdot (\hat{x} - x_0). \tag{3.16}$$

Therefore $\Phi(\hat{x}, \hat{t}) = (\xi, h)$. We see that $(\hat{x}, \hat{t}) \in \Gamma^+(u)$ since $u(\hat{x}, \hat{t}) \geq 0$ and $u \leq \ell(x)$ for all times prior to $\hat{t}$. This proves the claim, and hence

$$|P| \leq \int_{\Gamma^+(u)} |\partial_t u \, \det D^2 u|. \tag{3.17}$$

The mass of P is proportional to M^{n+1}, giving the estimate. □

For parabolic equations with term $b^i D_i$, we will need the following variant. As before, let $M = \sup_Q u^+$, and suppose $M > 0$ with $u|_{\partial B} \leq 0$ and $u|_{t=-1} \leq 0$. For $0 < \theta < 1$, define

$$\Gamma^+_\theta(u) = \Gamma^+ \cap \{|Du| < \theta \frac{M}{2}\}. \tag{3.18}$$

A similar argument gives $P_\theta \subset \Phi(\Gamma^+_\theta(u))$, where

$$P_\theta = \{(\xi, h) \in \mathbf{R}^{n+1} : |\xi| \leq \theta h, \ 0 \leq h \leq \frac{M}{2}\}. \tag{3.19}$$

By (3.17), since the mass of $|P_\theta|$ is proportional to M^{n+1} and θ^n, we obtain

$$\sup_Q u \leq \frac{C(n)}{\theta^{n/(n+1)}} \left(\int_{\Gamma^+_\theta(u)} |\partial_t u \, \det D^2 u| \right)^{1/(n+1)}. \tag{3.20}$$

We can now state the maximum principle. For bounded coefficients $a^{ij}(x,t)$ and $b^i(x,t)$ satisfying

$$a^{ij} = a^{ji}, \quad \Lambda^{-1}|\xi|^2 \leq a^{ij}(x,t)\xi_i \xi_j \leq \Lambda |\xi|^2, \text{ for all } \xi \in \mathbf{R}^n, \tag{3.21}$$

$$|b^i(x,t)| \leq \Lambda, \tag{3.22}$$

on Q_1 for some $\Lambda > 0$, we consider the operator

$$L = a^{ij} D_i D_j + b^i D_i. \tag{3.23}$$

The ABP maximum principle states:

Theorem 7 *Suppose $u \in C^{2,1}(\bar{Q}_1)$ satisfies*

$$(\partial_t - a^{ij} D_i D_j - b^i D_i) u \leq f, \tag{3.24}$$

$$u|_{\partial B} \leq 0, \quad u|_{t=-1} \leq 0, \tag{3.25}$$

where $f \in L^{n+1}(Q_1)$ and the coefficients a^{ij}, b^i, satisfy (3.21). Then

$$\sup_{Q_1} u \leq C \|f\|_{L^{n+1}(\Gamma^+(u))}, \tag{3.26}$$

where C depends on the dimension n, ellipticity constant Λ, and $\|b^i\|_{L^{n+1}}$.

Proof: We will apply the arithmetic-geometric mean inequality

$$(\det AB)^{1/(n+1)} \leq \frac{1}{n+1} \mathrm{Tr}\, AB \tag{3.27}$$

for positive-definite matrices A and B. Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & a^{ij} \end{pmatrix}, \quad B = \begin{pmatrix} \partial_t u & 0 \\ 0 & -u_{ij} \end{pmatrix}. \tag{3.28}$$

By Lemma 7, B is positive-definite on $\Gamma^+(u)$. Thus we have

$$\begin{aligned} (\partial_t u)(-\det D^2 u) &\leq C(\partial_t u - a^{ij} D_i D_j u)^{n+1} \\ &\leq C|f|^{n+1} + C|Du|^{n+1} \sum_i |b^i|^{n+1}. \end{aligned} \tag{3.29}$$

Here we used $(a+b)^{n+1} \leq C(n)(a^{n+1} + b^{n+1})$ for $a, b \geq 0$ and the convention that the constant C may change line by line.

By (3.20), for any $0 < \theta < 1$ we have

$$\begin{aligned} \sup_Q u &\leq \frac{C}{\theta^{n/(n+1)}} \|f\|_{L^{n+1}} \\ &\quad + \frac{C}{\theta^{n/(n+1)}} \sum_i \left(\int_{\Gamma^+_\theta(u)} |Du|^{n+1} |b^i|^{n+1} \right)^{1/(n+1)}. \end{aligned} \tag{3.30}$$

On $\Gamma^+_\theta(u)$ we have $|Du| \leq \theta(\sup_{Q_1} u^+)$, hence

$$\sup_{Q_1} u \leq \frac{C}{\theta^{n/(n+1)}} \|f\|_{L^{n+1}} + \theta^{1/(n+1)} C \|b\|_{L^{n+1}} \sup_{Q_1} u^+. \tag{3.31}$$

Choosing θ small enough gives the estimate. □

3.2 Local maximum principle

The reference for this section is [8]. In this section, we obtain a local estimate for the supremum of u. By local, we mean that we do not assume anything about $u : Q_1 \to \mathbf{R}$ at the boundary $u|_{\partial B}$ and $u|_{t=-1}$, and only obtain an estimate inside $Q_{1/2}$.

Theorem 8 *If $(\partial_t - a^{ij} D_i D_j)u \leq f$ in Q_1 with $\Lambda^{-1}\delta^{ij} \leq a^{ij}(x,t) \leq \Lambda\delta^{ij}$ and $a^{ij} = a^{ji}$, then for all $p > 0$ there exists a constant C depending on n, p, Λ such that*

$$\sup_{Q_{1/2}} u \leq C(\|u\|_{L^p(Q_1)} + \|f\|_{L^{n+1}(Q_1)}). \tag{3.32}$$

Proof: A useful cutoff function on Q_1 that will be used in this proof is

$$\eta(x,t) = (1-|x|^2)^\beta (1+t)^\beta, \quad \beta > 2. \tag{3.33}$$

We compute its evolution.

$$\partial_t \eta = \beta(1+t)^{-1}\eta, \tag{3.34}$$

$$D_i \eta = -2\beta x_i (1-|x|^2)^{\beta-1}(1+t)^\beta, \tag{3.35}$$

$$\begin{aligned}(\partial_t - a^{ij}D_iD_j)\eta &= \beta(1-|x|^2)^\beta(1+t)^{\beta-1} \\ &\quad +2\beta(\textstyle\sum a^{ii})(1-|x|^2)^{\beta-1}(1+t)^\beta \\ &\quad -4\beta(\beta-1)a^{ij}x_ix_j(1-|x|^2)^{\beta-2}(1+t)^\beta.\end{aligned} \tag{3.36}$$

This is equal to

$$\begin{aligned}(\partial_t - a^{ij}D_iD_j)\eta &= \beta(1+t)^{-1}\eta + 2\beta(\textstyle\sum a^{ii})(1-|x|^2)^{-1}\eta \\ &\quad -4\beta(\beta-1)a^{ij}x_ix_j(1-|x|^2)^{-2}\eta.\end{aligned} \tag{3.37}$$

We can estimate, using $(1-|x|^2) < 1$ and $(1+t) < 1$,

$$(\partial_t - a^{ij}D_iD_j)\eta \leq C\frac{\eta}{(1-|x|^2)^2(1+t)^2} = C\eta^{1-(2/\beta)}, \tag{3.38}$$

for $C(\Lambda, \beta)$. We will also use the following estimate obtained from (3.35),

$$|D\eta| \leq C\eta\, \eta^{-(1/\beta)}. \tag{3.39}$$

We will use the cutoff function η to localize u. Let $v = \eta u$. Compute the evolution

$$\begin{aligned}
(\partial_t - a^{ij}D_iD_j)v &= u(\partial_t - a^{ij}D_iD_j)\eta + \eta(\partial_t - a^{ij}D_iD_j)u - 2a^{ij}D_i\eta D_j u \\
&\leq u(\partial_t - a^{ij}D_iD_j)\eta + f\eta + 2\Lambda|D\eta||Du| \\
&\leq C|v|\eta^{-(2/\beta)} + f\eta + C\eta\eta^{-(1/\beta)}|Du|. \qquad (3.40)
\end{aligned}$$

We use the contact set to estimate $|Du|$. By (3.3), on $\Gamma^+(v)$ we have the estimate

$$|Dv(x,t)| \leq \frac{|v(x,t)|}{d(x,\partial B)} \leq \frac{|v(x,t)|}{1-|x|}. \qquad (3.41)$$

Since $\eta Du + uD\eta = Dv$, we can derive

$$\begin{aligned}
|Du| &\leq \eta^{-1}|Dv| + \eta^{-1}|u||D\eta| \\
&\leq |u|(1-|x|)^{-1} + C|u|\eta^{-(1/\beta)} \\
&\leq C|u|\eta^{-(1/\beta)}. \qquad (3.42)
\end{aligned}$$

In the last line we used $(1-|x|)^{-1} = (1+|x|)(1-|x|^2)^{-1}$. Putting everything together,

$$(\partial_t - a^{ij}D_iD_j)v \leq \varphi(x,t), \qquad (3.43)$$

for some function $\varphi(x,t)$ which satisfies

$$\varphi(x,t) \leq C|v|\eta^{-(2/\beta)} + f\eta, \qquad (3.44)$$

on $\Gamma^+(v)$. By the ABP estimate,

$$\sup_Q v \leq C(\|v\eta^{-(2/\beta)}\|_{n+1} + \|f\|_{n+1}). \qquad (3.45)$$

Using $v = \eta u$, we have

$$\sup_Q v \leq C((\sup v)^{1-(2/\beta)}\|u^{(2/\beta)}\|_{n+1} + \|f\|_{n+1}). \qquad (3.46)$$

For any $a, b, \varepsilon > 0$, by Young's inequality we have

$$ab \leq \varepsilon\frac{a^p}{p} + \frac{1}{\varepsilon}\frac{b^q}{q}, \qquad p = \frac{\beta}{\beta-2}, \quad q = \frac{\beta}{2}. \qquad (3.47)$$

Applying this to $(\sup v)^{1-(2/\beta)}$ and $\|u^{2/\beta}\|_{n+1}$, we obtain

$$\sup_Q v \leq C\left(\left(\int_{Q_1} |u|^{(2(n+1)/\beta)}\right)^{\beta/(2(n+1))} + \|f\|_{n+1}\right). \qquad (3.48)$$

Note

$$\sup_{Q_{1/2}} u \leq (1-(1/4)^2)^{-\beta}(1-(1/4)^2)^{-\beta} \sup_{Q_{1/2}} v \leq C \sup_{Q} v, \tag{3.49}$$

hence

$$\sup_{Q_{1/2}} u \leq C \left(\left(\int_{Q_1} |u|^{(2(n+1)/\beta)} \right)^{\beta/(2(n+1))} + \|f\|_{n+1} \right), \tag{3.50}$$

and this gives the desired estimate for $p = 2(n+1)/\beta$. The only restriction on β in our argument was the condition $\beta > 2$. This proves the result for all $0 < p < n+1$.

Since we have established the estimate for small $p > 0$, we also have it for large p since can always increase the L^p norm to make the estimate worse. □

3.3 Weak Harnack inequality

The reference for this section are the notes by C. Mooney [9]. I learned these estimates together with B. Choi, who also wrote his own set of notes [2].

We start by recall the prop-up lemma (Lemma 1), which was proved earlier to establish the strong maximum principle. We previously allowed terms $b^i D_i$, which we now set to zero. We note that $\delta > 0$ in the following lemma is independent of r by parabolic rescaling $u(rx, r^2 t)$.

Lemma 9 *Let $r, \alpha > 0$. Suppose $(\partial_t - a^{ij} D_i D_j)u \geq 0$ with $u \geq 0$ in $B_r(0) \times (-\alpha^2 r^2, 0]$, and*

$$a^{ij} = a^{ji}, \quad \Lambda^{-1}\delta^{ij} \leq a^{ij} \leq \Lambda \delta^{ij}. \tag{3.51}$$

Let $M > 0$ and $0 < \theta < 1$. Suppose

$$u(x, -\alpha^2 r^2) \geq M, \quad |x| < \theta r. \tag{3.52}$$

Then

$$u(x, 0) \geq \delta M, \quad |x| < \frac{r}{2}, \tag{3.53}$$

for $\delta > 0$ depending on α, Λ, θ, n.

The basic measure estimate for supersolutions is the following.

Proposition 9 *Suppose* $(\partial_t - a^{ij}D_iD_j)u \geq 0$ *in* Q_1 *with* $a^{ij} = a^{ji}$ *and* $\Lambda^{-1}\delta^{ij} \leq a^{ij}(x,t) \leq \Lambda\delta^{ij}$*. Further assume that* $u \geq 0$ *and* $u(0,0) \leq 1$*. Then there exists* $0 < \mu < 1$ *depending only on* n *and* Λ *such that*

$$|\{u \leq 2\} \cap Q_1| \geq \mu |Q_1|. \tag{3.54}$$

Proof: Consider

$$v = -u + 2(1+t)(1-|x|^2). \tag{3.55}$$

Then $v(0,0) \geq 1$, $v|_{t=-1} \leq 0$, $v|_{\partial B_1} \leq 0$, and $(\partial_t - a^{ij}D_iD_j)v \leq C$. Applying the ABP estimate (Theorem 7),

$$1 \leq C|\Gamma^+(v)|. \tag{3.56}$$

Note that $\Gamma^+(v)$ is contained in $\{v \geq 0\}$ by definition, hence $\Gamma^+(v) \subseteq \{u \leq 2\}$. $\square$

Next, we adapt this measure estimate to space-time paraboloids, which is the most technical part of the proof. We denote the downwards space-time parabola by

$$E^-_{1/2} = \{|x|^2 \leq -t, \quad -(1/2^2) \leq t \leq 0\}. \tag{3.57}$$

In addition to the downwards space-time parabola $E^-_{1/2}$, we will also use

$$E^+_r(x_0,t_0) = \{(x,t) : |x-x_0|^2 \leq t - t_0, \quad t_0 \leq t \leq t_0 + r^2 \leq 0\}, \tag{3.58}$$

which is a forward space-time parabola. We will also often use the parabolic cylinder

$$Q_r(x_0,t_0) = B_r(x_0) \times (-r^2 + t_0, t_0]. \tag{3.59}$$

Proposition 10 *Suppose* $(\partial_t - a^{ij}D_iD_j)u \geq 0$ *in* Q_3 *with* $a^{ij} = a^{ji}$ *and* $\Lambda^{-1}\delta^{ij} \leq a^{ij} \leq \Lambda\delta^{ij}$ *and* $u \geq 0$*.*

There exists $M > 1$ *and* $0 < \mu_0 < 1$ *depending only on* n *and* Λ *with the following property. Let* $(x_0,t_0) \in E^-_{1/2}$ *and* $0 < r \leq 1/2$*. If*

$$F_r = E^+_r(x_0,t_0) \cap E^-_{1/2} \tag{3.60}$$

intersects $\{u \leq 1\}$ *at height* $t = t_0 + r^2$*, then*

$$|\{u \leq M\} \cap F_r| \geq \mu_0\, |F_r|. \tag{3.61}$$

Proof: We claim there exists a uniform $\theta = \theta(n) > 0$ such that there exists a cylinder $Q_{2\theta r}(\bar{x}, \bar{t})$ contained inside $E_r^+(x_0, t_0) \cap E_{1/2}^-$. The proof of this can be found in [13], Lemma 2.2.

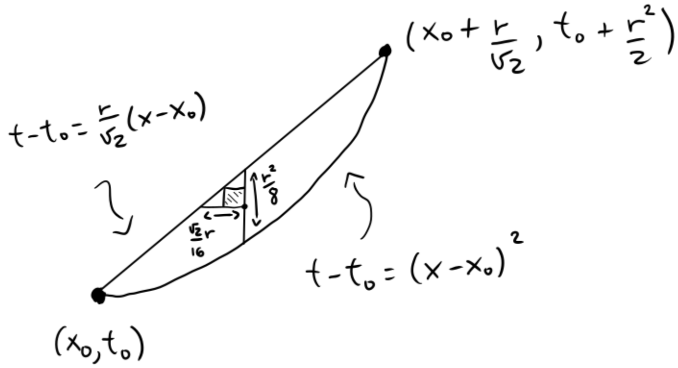
As an illustration, we provide full details of the claim if space has dimension $n = 1$. By symmetry, we may assume $x_0 \leq 0$. First observe that $(x_0 + \frac{r}{\sqrt{2}}, t_0 + \frac{r^2}{2})$ is inside $E_r^+(x_0, t_0) \cap E_{1/2}^-$. This point is clearly on $E_r^+(x_0, t_0)$, and it is in $E_{1/2}^-$, since the condition $(x_0 + \frac{r}{\sqrt{2}})^2 \leq -t_0 - \frac{r^2}{2}$ becomes

$$|x_0|^2 + r^2 - \sqrt{2} r\, |x_0| \leq |t_0|, \tag{3.62}$$

subject to

$$0 \leq r^2 \leq |t_0|, \quad |x_0| \leq |t_0|^{1/2}. \tag{3.63}$$

This can be checked by assuming $|t_0| = 1$ without loss of generality and optimizing the left-hand side of (3.62). By convexity, the line segment from (x_0, y_0) to $(x_0 + \frac{r}{\sqrt{2}}, t_0 + \frac{r^2}{2})$ is inside $E_r^+(x_0, t_0) \cap E_{1/2}^-$, and we can fit a cylinder of scale r in the area between this line segment and the paraboloid $E_r^+(x_0, t_0)$.



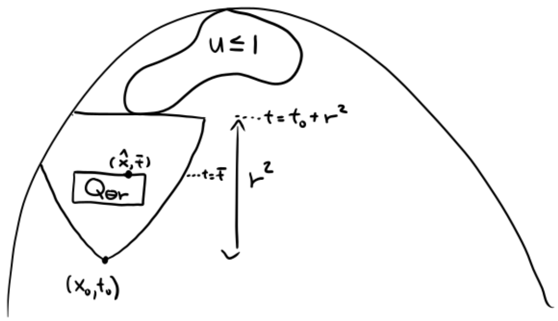
Next, we claim there is a constant $M > 1$ depending on n, Λ, and a point $(\hat{x}, \bar{t})$ on the top of $Q_{\theta r}(\bar{x}, \bar{t})$ such that $u(\hat{x}, \bar{t}) \leq M/2$. Indeed, suppose that

$$u(x, \bar{t}) \geq M/2, \quad x \in B_{\theta r}(\bar{x}). \tag{3.64}$$

Note that u is defined on $B_{4r}(\bar{x}) \times [\bar{t}, t_0 + r^2] \subset Q_3$, and if we let

$$t_0 + r^2 - \bar{t} = \alpha^2 r^2, \tag{3.65}$$

then $\alpha \in [1, 2]$.



We apply the prop-up lemma (Lemma 9) with α in a compact range to obtain $\delta(n, \Lambda, \theta) > 0$ such that

$$u(x, t_0 + r^2) \geq \delta \frac{M}{2}, \quad x \in B_{2r}(\bar{x}). \tag{3.66}$$

Thus $u \geq 1$ on the top of $E_r^+(x_0, t_0)$, which is the set $B_r(x_0) \times \{t_0 + r^2\}$. For $M \gg 1$, this contradicts the assumption that this upper boundary should intersect $\{u \leq 1\}$.

Hence $Q_{\theta r}(\hat{x}, \bar{t}) \subset Q_{2\theta r} \subset E_r^+(x_0, t_0) \cap E_{1/2}^-$ has the property $u(\hat{x}, \bar{t}) \leq M/2$. By the measure estimate (Proposition 9),

$$|\{u \leq M\} \cap Q_{\theta r}(\hat{x}, \bar{t})| \geq \mu_0 |Q_{\theta r}(\hat{x}, \bar{t})|. \tag{3.67}$$

Finally, we have

$$|\{u \leq M\} \cap E_r^+(x_0, t_0) \cap E_{1/2}^-| \geq \mu_1 |E_r^+(x_0, t_0) \cap E_{1/2}^-|, \tag{3.68}$$

as we may convert back to the intersection of paraboloids since the r scale is the same, as discussed in the claim at the start of the proof. □

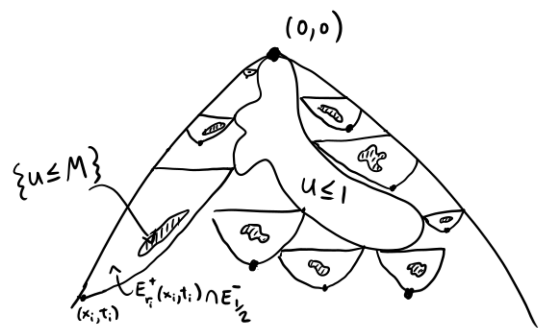
Next, we combine the measure estimate with a covering argument to obtain measure decay.

Proposition 11 *Suppose* $(\partial_t - a^{ij} D_i D_j) u \geq 0$ *in* Q_3 *with* $a^{ij} = a^{ji}$ *and* $\Lambda^{-1} \delta^{ij} \leq a^{ij}(x, t) \leq \Lambda \delta^{ij}$. *Further assume that* $u \geq 0$ *and* $u(0, 0) \leq 1$. *Then there exists* $M > 1$ *and* $0 < \mu < 1$ *depending only on* n *and* Λ *such that*

$$|\{u > M^{k+1}\} \cap E_{1/2}^-| \leq (1 - \mu) |\{u > M^k\} \cap E_{1/2}^-|. \tag{3.69}$$

Proof: First, we prove the estimate with $k = 0$. Choose $(x_i, t_i) \in \{u > 1\} \cap E^-_{1/2}$ (if this set is empty, the estimate holds trivially). There exists $r^2 = -t_i$ such that $(0,0) \in E^+_r(x_i, t_i)$, and we know $u(0,0) \leq 1$, hence we may take the minimal $r_i = r(x_i, t_i) > 0$ such that $E^+_{r_i}(x_i, t_i) \cap E^-_{1/2}$ intersects $\{u \leq 1\}$.

The sets $E^+_{r_i}(x_i, t_i) \cap E^-_{1/2}$ cover $\{u > 1\} \cap E^-_{1/2}$ as we vary (x_i, t_i). By the previous proposition, each of these sets contains a region where $u \leq M$ and we have a lower bound on the measure of this region.



By a version of the Vitali covering lemma due to Y. Wang, we can choose countable points (x_i, t_i) such that $E^+_{r_i}(x_i, t_i) \cap E^-_{1/2}$ are disjoint and their dilation still covers $\{u > 1\} \cap E^-_{1/2}$. To be more precise, define the dilation of

$$E^+_i = E^+_{r_i}(x_i, t_i) = \{(x,t) : |x - x_i|^2 \leq t - t_i, \quad t_i \leq t \leq t_i + r_i^2\} \quad (3.70)$$

to be

$$\tilde{E}^+_i = \{(\sqrt{2}+1)^{-2}|x - x_i|^2 \leq t - t_i + 3r_i^2, \;\; t_i - 3r^2 \leq t \leq t_i + r_i^2\}. \quad (3.71)$$

A calculation of the volumes gives the ratio

$$\frac{|E_i^+|}{|\tilde{E}_i^+|} = \kappa(n) > 0, \tag{3.72}$$

where $\kappa(n)$ is independent of r_i, x_i, t_i.

Lemma 10 *(Lemma 2.4 in [13]) Let D be a bounded subset of $\mathbf{R}^{n+1}$. Let $r(x,t)$ be a positive and bounded function on D. Cover D with the sets $E^+_{r(x,t)}(x,t)$. Then there exists a countable, disjoint subcollection*

$$E_i^+ = E_{r_i}^+(x_i, t_i), \tag{3.73}$$

such that

$$D \subseteq \bigcup_i \tilde{E}_i^+. \tag{3.74}$$

We will use this lemma and refer to [13] for the proof. Take such a disjoint subcollection E_i^+ such that $\bigcup \tilde{E}_i^+$ covers $\{u > 1\} \cap E^-_{1/2}$. Then by Proposition 10,

$$\begin{aligned} |\{u > M\} \cap E^-_{1/2}| &= |\{u > 1\} \cap E^-_{1/2}| - |\{u \le M\} \cap \{u > 1\} \cap E^-_{1/2}| \\ &\le |\{u > 1\} \cap E^-_{1/2}| - \sum_i |\{u \le M\} \cap E_i^+ \cap E^-_{1/2}| \\ &\le |\{u > 1\} \cap E^-_{1/2}| - \mu_0 \sum_i |E_i^+ \cap E^-_{1/2}|. \end{aligned} \tag{3.75}$$

The measure of $E_i^+ \cap E^-_{1/2}$ is comparable to the measure of E_i^+, by the remark in the first paragraph of the proof of Proposition 10. Thus

$$\begin{aligned} |\{u > M\} \cap E^-_{1/2}| &\le |\{u > 1\} \cap E^-_{1/2}| - \mu_1 \sum_i |E_i^+| \\ &= |\{u > 1\} \cap E^-_{1/2}| - \mu_1 \kappa \sum_i |\tilde{E}_i^+| \\ &\le |\{u > 1\} \cap E^-_{1/2}| - \mu_1 \kappa |\{u > 1\} \cap E^-_{1/2}|. \end{aligned} \tag{3.76}$$

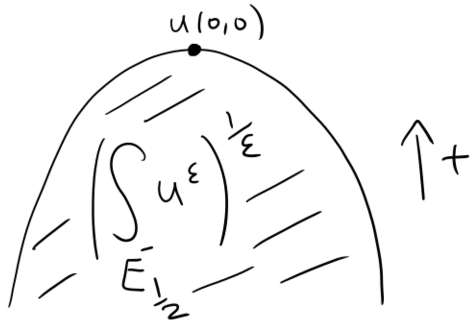
This is the desired result when $k = 0$. For arbitrary k, consider u/M^k to complete the proof. □

Theorem 9 *Suppose $(\partial_t - a^{ij} D_i D_j)u \ge 0$ in Q_3 where the coefficients satisfy*

$$a^{ij}(x,t) = a^{ji}(x,t), \quad \Lambda^{-1}\delta^{ij} \le a^{ij}(x,t) \le \Lambda \delta^{ij}. \tag{3.77}$$

Further assume that $u \geq 0$. Then there exists $\varepsilon > 0$ and $C > 1$ depending on Λ and n such that

$$\left(\int_{E^-_{1/2}} |u|^\varepsilon \right)^{1/\varepsilon} \leq C u(0,0). \tag{3.78}$$



Proof: First, suppose $u(0,0) = 1$. Define

$$P_k = E^-_{1/2} \cap \{M^k < u \leq M^{k+1}\}, \tag{3.79}$$

and split up the integral as follows:

$$\begin{aligned} \int_{E^-_{1/2}} |u|^\varepsilon &= \int_{\{u \leq 1\} \cap E^-_{1/2}} |u|^\varepsilon + \sum_{k=0}^{\infty} \int_{P_k} |u|^\varepsilon \\ &\leq |\{u \leq 1\} \cap E^-_{1/2}| + \sum_{k=0}^{\infty} \int_{P_k} M^{(k+1)\varepsilon}. \end{aligned} \tag{3.80}$$

By the iterated measure estimate (Proposition 11),

$$|P_k| \leq (1-\mu)^k |\{u > 1\} \cap E^-_{1/2}|. \tag{3.81}$$

Therefore

$$\int_{E^-_{1/2}} |u|^\varepsilon \leq |E^-_{1/2}| \left\{ 1 + M^\varepsilon \sum_{k=0}^{\infty} (1-\mu)^k M^{k\varepsilon} \right\}. \tag{3.82}$$

Choose $\varepsilon > 0$ small enough such that the geometric series converges. We then have

$$\int_{E^-_{1/2}} |u|^\varepsilon \leq C(n, \Lambda, \varepsilon). \tag{3.83}$$

Next, in the case when $u(0,0) \neq 1$ and is nonzero, we can consider $v(x,t) = u(x,t)/u(0,0)$ and apply (3.83) to v. When $u(0,0) = 0$, we apply the strong maximum principle and conclude $u \equiv 0$. $\square$

Theorem 10 *(Weak Harnack inequality) Suppose* $(\partial_t - a^{ij}D_iD_j)u \geq 0$ *in* Q_3 *with* $a^{ij} = a^{ji}$ *and* $\Lambda^{-1}\delta^{ij} \leq a^{ij} \leq \Lambda\delta^{ij}$*. Further assume that* $u \geq 0$*. Then there exists* $\varepsilon > 0$ *and* $C > 1$ *depending on* Λ *and* n *such that*

$$\left(\int_{B_{1/8}\times[-3/16,-2/16]} |u|^{\varepsilon}\right)^{1/\varepsilon} \leq C \inf_{B_{1/8}\times[-1/16,0]} u. \tag{3.84}$$

Proof: Let $(\hat{x},\hat{t})$ be the infimum of u on $\bar{B}_{1/8}\times[-1/16,0]$. By the previous result, the value of u at $(\hat{x},\hat{t})$ is bounded below by the L^{ε} norm of u on a space-time parabola

$$E^{-}_{1/2}(\hat{x},\hat{t}) = \{|x-\hat{x}|^2 \leq -t+\hat{t}, \;\; -\frac{1}{4}+\hat{t} \leq t \leq \hat{t}\}. \tag{3.85}$$

The set $B_{1/8}\times[-3/16,-2/16]$ is contained in this space-time parabola. □

There is a version of this estimate with a source term $f \in L^{n+1}$ on the right-hand side. To obtain the weak Harnack inequality with the L^{n+1} norm of f, we would need to redo the estimate with a few extra steps. Instead, we take a shortcut and get the following weaker inequality, which is still good enough for many purposes.

Corollary 1 *Suppose* $(\partial_t - a^{ij}D_iD_j)u \geq f$ *in* Q_3 *with* $a^{ij} = a^{ji}$ *and* $\Lambda^{-1}\delta^{ij} \leq a^{ij}(x,t) \leq \Lambda\delta^{ij}$ *and* $f \in L^{\infty}(Q_1)$*. Further assume that* $u \geq 0$*. Then there exists* $\varepsilon > 0$ *and* $C > 1$ *depending on* Λ *and* n *such that*

$$\left(\int_{B_{1/8}\times[-3/16,-2/16]} |u|^{\varepsilon}\right)^{1/\varepsilon} \leq C\left\{\inf_{B_{1/8}\times[-1/16,0]} u + \|f\|_{L^{\infty}(Q_1)}\right\}.$$

Proof: Let

$$v = u + (t+3^2)\|f\|_{L^{\infty}}. \tag{3.86}$$

Then $v \geq 0$ and $(\partial_t - a^{ij}D_iD_j)v \geq 0$ on Q_3. Thus by the previous theorem,

$$\left(\int_{Q^{-}_{1/8}} \left|u+(t+3^2)\|f\|_{L^{\infty}}\right|^{\varepsilon}\right)^{1/\varepsilon} \leq C \inf_{Q_{1/8}}\left\{u+(t+3^2)\|f\|_{L^{\infty}}\right\}, \tag{3.87}$$

where $Q^{-}_{1/8} = B_{1/8}\times[-3/16,-2/16]$. Therefore

$$\left(\int_{Q^{-}_{1/8}} |u|^{\varepsilon}\right)^{1/\varepsilon} \leq C \inf_{Q_{1/8}} u + C\sup_{Q_{1/8}}\left\{(t+3^2)\|f\|_{L^{\infty}}\right\}, \tag{3.88}$$

as required. □

3.4 Hölder estimates

We can now obtain the Harnack inequality of Krylov-Safonov [7] by combining our results.

Theorem 11 *Let $R > 0$. Suppose $u \in C^{2,1}(Q_R)$ solves*

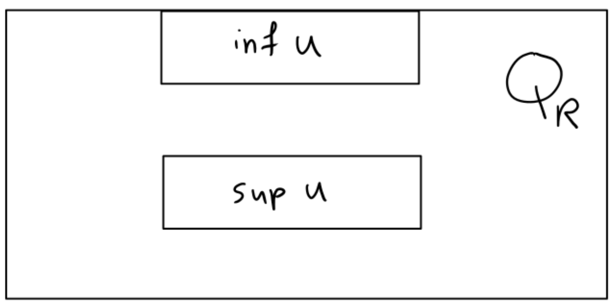
$$(\partial_t - a^{ij} D_i D_j) u = f \tag{3.89}$$

with

$$a^{ij}(x,t) = a^{ji}(x,t), \quad \Lambda^{-1}\delta^{ij} \le a^{ij}(x,t) \le \Lambda \delta^{ij} \tag{3.90}$$

for some $\Lambda > 0$, and $f \in L^\infty(Q_R)$. Further assume that $u \ge 0$. Then there exists $C > 1$ depending on Λ and n such that

$$\sup_{Q_{R/2}(0,-R^2/2)} u \le C \left\{ \inf_{Q_{R/2}(0,0)} u + R^2 \|f\|_{L^\infty(Q_R)} \right\}. \tag{3.91}$$



Proof: First, suppose u is defined in Q_3. Let $\varepsilon > 0$ be as in Corollary 1. By Theorem 8

$$\sup_{Q_{1/16}(0,-1/8)} u \le C \left\{ \|u\|_{L^\varepsilon(Q_{1/8}(0,-1/8))} + \|f\|_{L^\infty(Q_3)} \right\}. \tag{3.92}$$

Next, by Corollary 1,

$$\|u\|_{L^\varepsilon(Q_{1/8}(0,-1/8))} \le C \left\{ \inf_{Q_{1/8}} u + \|f\|_{L^\infty(Q_3)} \right\}. \tag{3.93}$$

Putting both together

$$\sup_{Q_{1/16}(0,-1/8)} u \le C \left\{ \inf_{Q_{1/8}} u + \|f\|_{L^\infty(Q_3)} \right\}. \tag{3.94}$$

More generally, if u is defined on Q_{3R}, we perform a rescaling and consider $v(x,t) = u(Rx, R^2t)$. Then v is defined on Q_3 and satisfies

$$(\partial_t - a^{ij} D_i D_j) v(x,t) = R^2 f(Rx, R^2 t), \tag{3.95}$$

and applying the Harnack inequality to v on Q_3 gives the Harnack inequality for u on Q_{3R}.

$$\sup_{Q_{R/16}(0,-R^2/8)} u \leq C \left\{ \inf_{Q_{R/16}} u + R^2 \|f\|_{L^\infty(Q_{3R})} \right\}. \tag{3.96}$$

Here we replaced $\inf_{Q_{R/8}}$ with $\inf_{Q_{R/16}}$ at the cost of making the inequality a bit worse, but now the right-hand side matches the left-hand side. Inequality (3.96) is the type of Harnack inequality that we are looking for, but for purely aesthetic reasons, we will try to get rid of the factors of $R/16$, $3R$, etc, to get the simpler estimate (3.91) stated in the theorem.

To get (3.91), we set $R = 1$ and use a covering argument. Fix $\rho = 1/6$. Suppose u is defined on Q_1. Suppose the supremum of u on the closure of $Q_{1/2}(0, -1/2)$ is attained at (x_0, t_0) and the infimum of u on the closure of $Q_{1/2}$ is attained at (x_∞, t_∞). Make a path of cylinders

$$C_i = Q_{\rho/16}(x_i, t_i), \quad x_i \in B_{1/2}(0), \ t_i \in [t_0, t_\infty], \tag{3.97}$$

such that $C_i \cap C_{i+1}$ is non-empty for i odd, and $C_i \mapsto C_{i+1}$ for i even is a shift forward in time by $\rho^2/8$. Choose these cylinders such that

$$C_0 = Q_{\rho/16}(x_0, t_0), \quad (x_\infty, t_\infty) \in C_N. \tag{3.98}$$

Note $Q_{3\rho}(p_i) \subseteq Q_1$. Thus we may apply (3.96)

$$\begin{aligned}
\sup_{Q_{\rho/16}(x_0,t_0)} u &\leq C \left\{ \inf_{C_1} u + \|f\|_{L^\infty(Q_1)} \right\} \\
&\leq C \left\{ \inf_{C_1 \cap C_2} u + \|f\|_{L^\infty(Q_1)} \right\} \\
&\leq C \left\{ \sup_{C_1 \cap C_2} u + \|f\|_{L^\infty(Q_1)} \right\} \\
&\leq C \left\{ \sup_{C_2} u + \|f\|_{L^\infty(Q_1)} \right\} \\
&\leq C \left\{ \inf_{C_3} u + \|f\|_{L^\infty(Q_1)} \right\}.
\end{aligned} \tag{3.99}$$

Within $(6)^2(8)$ steps shifting cylinders in time, and $(6)(16)$ steps moving cylinders in space, we can reach C_N which contains (x_∞, t_∞). Then

$$\sup_{Q_{\rho/16}(x_0,t_0)} u \le C\left\{ \inf_{C_N} u + \|f\|_{L^\infty(Q_1)} \right\}, \tag{3.100}$$

and by definition of (x_0, t_0) and (x_∞, t_∞), we conclude

$$\sup_{Q_{1/2}(0,-1/2)} u \le C\left\{ \inf_{Q_{1/2}} u + \|f\|_{L^\infty(Q_1)} \right\}. \tag{3.101}$$

The general version with u defined on Q_R now holds by rescaling. $\square$

We can now use the Harnack inequality to obtain an estimate on Hölder norms. As previously noted, the theory developed here also goes through when $f \in L^{n+1}$, but we only treat the case $f \in L^\infty$ for simplicity.

Theorem 12 *[7] Suppose $u(x,t)$ is a $C^{2,1}(Q_1)$ solution of*

$$\left(\partial_t - a^{ij}(x,t) D_i D_j\right) u(x,t) = f(x,t), \tag{3.102}$$

where $f \in L^\infty(Q_1)$ and

$$a^{ij}(x,t) = a^{ji}(x,t), \quad \Lambda^{-1}\delta^{ij} \le a^{ij}(x,t) \le \Lambda \delta^{ij}. \tag{3.103}$$

Then there exists $0 < \delta < 1$ and $C > 1$ depending on n and Λ such that

$$\|u\|_{\delta/2,\delta;\, Q_{1/2}} \le C(\|u\|_{L^\infty(Q_1)} + \|f\|_{L^\infty(Q_1)}). \tag{3.104}$$

Proof: Set $F_0 = \|f\|_{L^\infty(Q_1)}$. Let $p = (x_0, t_0) \in Q_{1/2}$ and $Q_r(p) \subseteq Q_{1/2}$ with

$$0 < r < \frac{1}{2}. \tag{3.105}$$

Define

$$M(r) = \sup_{Q_r(p)} u, \quad m(r) = \inf_{Q_r(p)} u, \tag{3.106}$$

and

$$\omega(r) = M(r) - m(r). \tag{3.107}$$

For simplicity, we write

$$Q_{r/2}(x_0, t_0 - r^2/2) = Q^-_{r/2}(p). \tag{3.108}$$

Applying the Harnack inequality to $M(r) - u \geq 0$ on Q_r gives

$$\sup_{Q^-_{r/2}(p)} (M(r) - u) \leq C\left\{ \inf_{Q_{r/2}(p)} (M(r) - u) + r^2 F_0 \right\}, \tag{3.109}$$

$$M(r) - \inf_{Q^-_{r/2}(p)} u \leq C\left\{ M(r) - \sup_{Q_{r/2}(p)} u + r^2 F_0 \right\}. \tag{3.110}$$

Applying the Harnack inequality to $u - m(r) \geq 0$ in Q_r gives

$$\sup_{Q^-_{r/2}(p)} (u - m(r)) \leq C\left\{ \inf_{Q_{r/2}(p)} (u - m(r)) + r^2 F_0 \right\}, \tag{3.111}$$

$$\sup_{Q^-_{r/2}(p)} u - m(r) \leq C\left\{ \inf_{Q_{r/2}(p)} u - m(r) + r^2 F_0 \right\}. \tag{3.112}$$

Adding (3.110) and (3.112), we obtain

$$\begin{aligned} & \left\{ \sup_{Q^-_{r/2}(p)} u - \inf_{Q^-_{r/2}(p)} u \right\} + C\omega(r/2) \\ \leq \quad & (C-1)\omega(r) + 2Cr^2 F_0, \end{aligned} \tag{3.113}$$

which implies

$$\omega(r/2) \leq \frac{C-1}{C}\omega(r) + 2F_0 r^2. \tag{3.114}$$

We now appeal to the following useful lemma (e.g. [3] Lemma 4.19), which is often used in estimates of Hölder norms.

Lemma 11 *Let $\omega : (0, R] \to [0, \infty)$ and $\sigma : (0, R] \to [0, \infty)$ be increasing functions. Suppose there exists $0 < \gamma < 1$ and $0 < \tau < 1$ such that for all $r \leq R$,*

$$\omega(\tau r) \leq \gamma\omega(r) + \sigma(r). \tag{3.115}$$

Then for any $0 < \mu < 1$ and $r \leq R$, there holds

$$\omega(r) \leq C\left\{ \left(\frac{r}{R}\right)^{\delta} \omega(R) + \sigma(r^{\mu} R^{1-\mu}) \right\} \tag{3.116}$$

where $C = C(\gamma, \tau)$ and $\delta = (1-\mu)\log\gamma/\log\tau$.

Applying this lemma with $R = 1/2$, $\gamma = (C-1)/C$, $\tau = 1/2$, $\sigma(r) = 2F_0 r^2$, gives

$$\omega(r) \le C\left\{\omega\left(\frac{1}{2}\right) r^{\delta} + F_0 r^{2\mu}\right\}. \tag{3.117}$$

Choosing μ such that $2\mu \ge \delta$, then

$$\sup_{q_1,q_2\in Q_r(p)} \left\{u(q_1) - u(q_2)\right\} \le C r^{\delta}\left\{\|u\|_{L^\infty(Q_1)} + \|f\|_{L^\infty(Q_1)}\right\}. \tag{3.118}$$

Translating $Q_r(p) \subseteq Q_{1/2}$ gives the desired estimate. □

Proof of Lemma 11: Fix $r_1 \le R$. Then for all $r \le r_1$,

$$\omega(\tau r) \le \gamma\omega(r) + \sigma(r_1), \tag{3.119}$$

since σ is increasing. Applying this inequality again, we have

$$\omega(\tau^2 r) \le \gamma^2\omega(r) + (1+\gamma)\sigma(r_1). \tag{3.120}$$

Iterating, we obtain

$$\omega(\tau^k r) \le \gamma^k\omega(r) + \sigma(r_1)\sum_{i=0}^{k-1}\gamma^i \le \gamma^k\omega(r) + \frac{\sigma(r_1)}{1-\gamma}, \tag{3.121}$$

for any integer k. We now choose k such that

$$\tau^k r_1 < r \le \tau^{k-1} r_1. \tag{3.122}$$

Then, since ω is increasing,

$$\omega(r) \le \omega(\tau^{k-1} r_1) \le \gamma^{k-1}\omega(R) + \frac{\sigma(r_1)}{1-\gamma}. \tag{3.123}$$

From

$$k\log\tau \le \log r - \log r_1, \tag{3.124}$$

we obtain

$$k\log\gamma \le \frac{\log\gamma}{\log\tau}(\log r - \log r_1), \tag{3.125}$$

and hence

$$\gamma^k \le \left(\frac{r}{r_1}\right)^{\log\gamma/\log\tau}. \tag{3.126}$$

Therefore

$$\omega(r) \le \gamma^{-1}\left(\frac{r}{r_1}\right)^{\log\gamma/\log\tau}\omega(R) + \frac{\sigma(r_1)}{1-\gamma}. \tag{3.127}$$

We now let $r_1 = r^\mu R^{1-\mu}$. Then

$$\omega(r) \le \gamma^{-1}\left(\frac{r^{1-\mu}}{R^{1-\mu}}\right)^{\log\gamma/\log\tau}\omega(R) + \frac{\sigma(r^\mu R^{1-\mu})}{1-\gamma}. \tag{3.128}$$

This completes the proof of Lemma 11. □

Lastly, we note the usual Harnack inequality for stationary solutions, which is a special case of (3.91).

Theorem 13 *Let $R > 0$. Suppose $u \in C^2(B_R)$ solves*

$$a^{ij}(x)D_iD_ju(x) = f(x) \tag{3.129}$$

with $a^{ij} = a^{ji}$, $\Lambda^{-1}\delta^{ij} \le a^{ij}(x) \le \Lambda\delta^{ij}$ and $f \in L^\infty(B_R)$. Assume $u \ge 0$. Then there exists $C > 1$ depending on n, Λ such that

$$\sup_{B_{R/2}} u \le C\left\{\inf_{B_{R/2}} u + R^2\|f\|_{L^\infty(B_R)}\right\}. \tag{3.130}$$

Acknowledgements: I learned this material through numerous discussions with Beomjun Choi, Tristan Collins, Teng Fei, Karsten Gimre, Pengfei Guan, Zhijie Huang, Pei-Ken Hung, Duong Phong, Connor Mooney, Keaton Naff, Freid Tong, and Xiangwen Zhang.

References

[1] Y.-Z. Chen and L.-C. Wu, *Second Order Elliptic Equations and Elliptic Systems*, Translations of Mathematical Monographs Volume 174, American Mathematical Society, 1998.

[2] B. Choi, *Parabolic Krylov-Safonov Harnack Inequality*, notes available on the author's website.

[3] Q. Han and F. Lin, *Elliptic Partial Differential Equations*, Courant lecture notes in mathematics, Vol 1. American Mathematical Society, 2011.

[4] R. Haslhofer, *PDE II - Schauder estimates*, notes available on the author's website.

[5] N.V. Krylov, *Sequences of convex functions and estimates of the maximum of the solution of a parabolic equation*, Siberian Mathematical Journal 17, no. 2 (1976), 226-236.

[6] N.V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics, Vol. 12. American Mathematical Society, Providence, RI, 1996.

[7] N.V. Krylov and M.V. Safonov, *A property of the solutions of parabolic equations with measurable coefficients*, Izv. Akad. Nauk SSSR Ser. Mat., 44(1), 1980.

[8] G.M. Lieberman, *Second Order Parabolic Differential Equations*, World Scientific, Singapore New Jersey London Hong Kong, 1996.

[9] C. Mooney, *Parabolic PDE*, notes available on the author's website.

[10] K. Tso, *On an Aleksandrov-Bakelman type maximum principle for second-order parabolic equations*, Communications in Partial Differential Equations 10, no. 5 (1985), 543-553.

[11] L. Simon, *Schauder estimates by scaling*, Calculus of Variations and Partial Differential Equations 5, no. 5 (1997), 391-407.

[12] G. Szekelyhidi, *An introduction to extremal Kahler metrics*, Vol. 152. American Mathematical Society, 2014.

[13] Y. Wang, *Small perturbation solutions for parabolic equations*, Indiana University Mathematics Journal 62, no. 2 (2013), 671-697.