CONTINUITY OF SOLUTIONS TO PARABOLIC EQUATIONS: NASH'S PROOF AND FABES-STROOCK'S MODIFICATION

Nash's Original Approach

Nash's theorem pertains to operators with nonsmooth coefficients in divergence form:

$$L_t = \partial_i(a^{ij}(x,t)\partial_j),$$

where only symmetry $a^{ij} = a^{ji}$ and a uniform ellipticity condition on the coefficients are assumed:

$$\lambda I \le (a^{ij}(x,t))_{i,j=1}^n \le \lambda^{-1} I$$

for all $(x,t) \in \mathbb{R}^n \times [0,\infty)$.

Theorem 1. *(Nash [3]) Suppose $u \in C^\infty(\mathbb{R}^n \times [0,\infty)) \cap L^\infty(\mathbb{R}^n \times [0,\infty))$ solves $\partial_t u = L_t u$. Then there exists $C = C(n,\alpha,\lambda) < \infty$ such that for all $(x_1,t_1),(x_2,t_2) \in \mathbb{R}^n \times (0,\infty)$ with $t_1 < t_2$, we have*

$$|u(x_1,t_1) - u(x_2,t_2)| \le C||u||_{L^\infty(\mathbb{R}^n\times[0,\infty))}\left(\frac{|x_1-x_2|^\alpha}{t_1^{\alpha/2}} + \left(\frac{t_2-t_1}{t_1}\right)^{\frac{\alpha}{2(1+\alpha)}}\right).$$

Remark 2. The assumption that u is smooth is unnecessary, which and an approximation argument shows that in fact the same result is true for any distributional solution. By an approximation argument, this has the important consequence that distributional solutions of $\partial_t u = L_t u$ are Holder continuous.

The starting point of Nash's proof is obtaining an on-diagonal upper bound for the fundamental solution

$$\Gamma : \{(x,t;y,s) \in (\mathbb{R}^n \times [0,\infty))^2; t > s\} \to (0,\infty)$$

of $\partial_t u = L_t u$, which is defined for each $(y,s) \in \mathbb{R}^n \times [0,\infty)$ by letting $\Gamma(\cdot,\cdot;y,s) \in C^\infty(\mathbb{R}^n \times (s,\infty))$ be the unique solution of $\partial_t u = L_t u$ which is bounded on compact time intervals, and which satisfies $\Gamma(\cdot,t;y,s) \to \delta_y$ in the sense of distributions as $t \searrow s$. The proof of the on-diagonal upper bound uses an inequality conceived by Nash and proved by Elias Stein (in the Euclidean setting), which is now called the Nash inequality:

$$||u||_{L^2(\mathbb{R}^n)}^{1+\frac{2}{n}} \le C(n)||\nabla u||_{L^2(\mathbb{R}^n)}||u||_{L^1(\mathbb{R}^n)}^{\frac{2}{n}} \qquad \forall u \in \mathcal{S}(\mathbb{R}^n).$$

Applying this inequality to $u_t(x) := \Gamma(x,t;0,0)$, and using $\int_{\mathbb{R}^n} u_t dx = 1$ gives

$$\frac{d}{dt}||u_t||_{L^2(\mathbb{R}^n)}^2 = 2\int_{\mathbb{R}^n} u_t\partial_i(a^{ij}(x,t)\partial_j u_t)dx \le -2\lambda||\nabla u_t||_{L^2(\mathbb{R}^n)}^2 \le -2c(n)||u_t||_{L^2(\mathbb{R}^n)}^{2+\frac{4}{n}},$$

which can be integrated to give $||u_t||_{L^2(\mathbb{R}^n)}^2 \le C(n,\lambda)t^{-\frac{n}{2}}$ for all $t > 0$, from which the semigroup property and similar estimates for the adjoint equation imply

$$\Gamma(x,t;y,s) \le \frac{C(n,\lambda)}{(t-s)^{\frac{n}{2}}}.$$

This derivation of the on-diagonal upper bound from the Nash inequality works in extremely general situations, and it is furthermore now known that Nash's inequality is equivalent to an L^2-Sobolev inequality.

Nash then considers two additional functionals defined using Γ: the moment

$$M(t) := \int_{\mathbb{R}^n} \Gamma(x,t;0,0)|x|dx,$$

and the entropy

$$Q(t) := -\int_{\mathbb{R}^n} \Gamma(x,t;0,0)\log\Gamma(x,t;0,0)dx.$$

In the Euclidean case $a^{ij} \equiv \delta^{ij}$, a direct computation gives $M(t) = \sqrt{2nt}$ and $Q(t) = \frac{n}{2}\log(4\pi t) + \frac{n}{2}$. Nash shows that these inequalities holds in great generality:

$$Q(t) \geq \frac{n}{2}\log(4\pi t) - C(n,\lambda),$$

$$C(n,\lambda)^{-1}\sqrt{t} \leq M(t) \leq C(n,\lambda)\sqrt{t},$$

where the lower bounds for Q, M are both elementary consequences of the on-diagonal heat kernel upper bound, while the upper bound for M requires proving an estimate of the form

$$Q'(t) \geq c(n,\lambda)M'(t)^2.$$

Next, Nash proves a a weighted lower bound for the logarithm of the normaized fundamental solution:

$$G(t) := \int_{\mathbb{R}^n} e^{-|\xi|^2}\log(t^{\frac{n}{2}}\Gamma(t^{\frac{1}{2}}x,t;0,0)+\delta)dx \geq -C(n,\lambda)\sqrt{-\log\delta},$$

which he terms 'The G Bound'. The proof relies on a somewhat arduous but elementary computation, which at some point uses in a crucial way the on-diagonal heat kernel upper bound, and the upper and lower bounds for $M(t)$, to show that

$$G'(t) \geq c(n,\lambda)G(t)^2 + C(n,\lambda)\log\delta$$

when $G(t)$ is sufficiently negative, from which the G Bound easily follows. An elementary (but not obvious) consequence of the G Bound is that

$$A(t) := \frac{1}{2}\int_{\mathbb{R}^n} |\Gamma(x,t;x_1,0) - \Gamma(x,t;x_2,0)|dy \leq \psi\left(\frac{|x_1-x_2|}{\sqrt{t}}\right)$$

for some positive increasing function $\psi : (0,\infty) \to (0,1)$. In other words, any fundamental solutions based at nearby points overlap at least some controlled amount.

To show spatial continuity, Nash uses an inductive scheme based on the above inequality for $A(t)$, simultaneously estimating $A(t)$ and the moments

$$M_a(t) := \int_{\mathbb{R}^n} |x-x_0|T_a(x,t)dx, \quad M_b(t) := \int_{\mathbb{R}^n} |x-x_0|T_b(x,t)dx,$$

where $x_0 := \frac{x_1+x_2}{2}$ and

$$T_a(x,t) := (\Gamma(x,t;x_1,0) - \Gamma(x,t;x_2,0))_+\,,$$
$$T_b(x,t) := (\Gamma(x,t;x_2,0) - \Gamma(x,t;x_1,0))_+\,.$$

Here, T_a can roughly be thought of as the fundamental solution based at $(x_1,0)$, with a chunk of the solution near x_2 subtracted off. Nash fixes $\sigma := 1-(1-\psi(1))/4$, and lets t_k be the first time $t > 0$ where the (nonincreasing) quantity A satisfies $A(t_k) = \sigma^k$. It follows from the estimate for $A(t)$ and the semigroup property that

$$A(t) \leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \psi\left(\frac{|x-y|}{\sqrt{t-s}}\right)\chi(x,y,s)dxdy, \tag{0.1}$$

where χ is a density function on $\mathbb{R}^n \times \mathbb{R}^n$ which is mostly concentrated near (x_1,x_2):

$$\frac{A(t)}{2} \leq \int_{B(x_1,2M_a(t)/A(t))}\int_{B(x_2,2M_b(t)/A(t))} \chi(x,y,s)dydx \leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi(x,y,s)dydx \leq A(t) \leq 1.$$

If we know $\operatorname{supp}(\chi(\cdot,\cdot,s)) \subseteq B\left((x_1,x_2), \frac{\tau^{\frac{1}{2}}|x_1-x_2|}{\sqrt{s}}\right)$, we could estimate $A(t+\tau) \leq \psi(2)A(t)$, and so easily estimate t_k inductively. $\chi(\cdot,\cdot,t)$ will never be compactly supported, however, so we need to estimate the "nearby part" of the right hand side of (0.1) using the fact that then $\psi < 1-\epsilon$ there, and then estimate the "far-away part" in terms of the moment bound $M_k := \{M_a(t_k), M_b(t_k)\}$:

$$A(t') \leq \sigma^k \left(\frac{3}{4} + \frac{1}{4}\psi\left(\frac{4M_k}{\sigma^k\sqrt{t'-t_k}}\right)\right),$$

which implies $t_{k+1} \leq t_k + 16\sigma^{-2k}M_k^2$. Although we have a rough estimate for the moments of the form $M_k(t) \leq C(n,\lambda)\sqrt{t}$, we need an upper bound that approaches zero as $|x_1 - x_2| \to 0$. To find such a bound, we can use that T_a satisfies $\partial_t T_a \leq L_t T_a$ in the weak sense along with the moment estimate to get

$$M_a(t') \leq M_a(t) + C(n,\lambda)A(t)\sqrt{t'-t}.$$

Because $M_a(0) = |x_1 - x_0| = \frac{|x_1-x_2|}{2}$, we get the following by induction:

$$M_{k+1} \leq M_k + C(n,\lambda)\sigma^k \cdot 4\sigma^{-k}M_k \leq \cdots \leq C(n,\lambda)^k|x_1-x_2|,$$

so combining estimates and using $t_0 = 0$, we have $t_{k+1} \leq B^k|x_1-x_2|^2$, where $B = B(n,\lambda) < \infty$. This easily implies

$$A(t) \leq C(n,\lambda,\alpha)\left(\frac{|x_1-x_2|}{\sqrt{t}}\right)^\alpha,$$

where $\alpha := -2(\log B)^{-1}\log\sigma \in (0,1)$, which in turn implies Theorem 1 when $t_1 = t_2$, using the semigroup property. For the estimate in time, we note that any solution $\partial_t u = L_t u$ satisfies

$$|u(x,t)-u(x,s)| \leq \int_{B(x,\rho)} |u(y,s)-u(x,s)| \cdot \Gamma(x,t;y,s)dy + 2||u||_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n\setminus B(x,\rho)} \Gamma(x,t;y,s)dy$$

by the semigroup property, where $\rho > 0$ is to be determined. The first integral can be estimated using the spatial Holder estimate for u, while the second integral can be estimated using the moment bound for Γ. Combining estimates and optimizing ρ then gives Theorem 1.

Gaussian Bounds on the Fundamental Solution

The strategy of Fabes-Stroock [2] is to derive Theorem (1) as a consequence of the following two-sided Gaussian bounds for the fundamental solution, originally proved by Aronson.

Theorem 3. *(Aronson [1]) There exists $C = C(n,\lambda) < \infty$ such that for all $(x_1,t_1),(x_2,t_2) \in \mathbb{R}^n \times [0,\infty)$ with $t_1 < t_2$, we have*

$$\frac{1}{C(t_2-t_1)^{\frac{n}{2}}}\exp\left(-\frac{C|x_1-x_2|^2}{t_2-t_1}\right) \leq \Gamma(x_2,t_2;x_1,t_1) \leq \frac{C}{(t_2-t_1)^{\frac{n}{2}}}\exp\left(-\frac{|x_1-x_2|^2}{C(t_2-t_1)}\right).$$

The proof of Theorem 3 in [1] relies on the parabolic Harnack inequality, but Fabes-Stroock are able to give an independent proof of this fact, and in fact use this to prove the parabolic Harnack inequality. The proof of the on-diagonal upper bound is similar to Nash's, where $\frac{d}{dt}||u||_{L^p(\mathbb{R}^n)}$ is estimated using Nash's inequality for a positive solution of $\partial_t u = L_t u$, generalizing Nash's estimate for the L^2 norm. Moreover, we actually estimate the kernel corresponding to the "twisted" operators $L_t^\psi := e^{-\psi} \circ L_t \circ e^\psi$, where $\psi(x) := \langle \alpha, x\rangle$ for a fixed direction $\alpha \in \mathbb{R}^n$, which gives the slightly more complicated evolution estimate

$$\frac{d}{dt}||u_t||_{L^{2p}(\mathbb{R}^n)} \leq -\frac{c(n,\lambda)}{2p}||u_t||_{L^{2p}(\mathbb{R}^n)}^{1+\frac{4p}{n}}||u_t||_{L^p(\mathbb{R}^n)}^{-\frac{4p}{n}} + \frac{|\alpha|^2 p}{\lambda}||u_t||_{L^{2p}(\mathbb{R}^n)}$$

for positive solutions of $\partial_t u = L_t^\psi u$. Then an entirely elementary ODE comparison argument lets us estimate $L^{2^{k+1}}$-norms of u in terms of L^{2^k}-norms. Iterating these estimate leads to

$$\sup_{\mathbb{R}^n} u_t \leq C(n,\lambda) t^{-\frac{n}{4}} e^{2\lambda^{-1}|\alpha|^2 t} ||u_0||_{L^2(\mathbb{R}^n)}.$$

By duality, we can easily obtain

$$||u_t||_{L^2(\mathbb{R}^n)} \leq C(n,\lambda) t^{-\frac{n}{4}} e^{2\lambda^{-1}|\alpha|^2 t} ||u_0||_{L^2(\mathbb{R}^n)}.$$

Combining estimates using the semigroup property, we get the following on-diagonal upper bound for the fundamental solution Γ^ψ of the twisted operator:

$$\Gamma^\psi(x,t;y,s) \leq \frac{C(n,\lambda)}{t^{\frac{n}{2}}} e^{4\lambda^{-1}|\alpha|^2 t}.$$

Now we observe that the on-diagonal upper bounds for the twisted operators imply a Gaussian upper bound for the original operator, by noting the following relation:

$$\Gamma^\psi(x,t;y,0) = e^{-\langle x-y,\alpha\rangle}\Gamma(x,t;y,0),$$

and then taking $\alpha := \frac{\lambda}{8t} y - x$.

For the lower bound, we consider a quantity similar to that estimated by Nash:

$$G := \int_{\mathbb{R}^n} e^{-|y|^2} \log \Gamma(x,1;y,0) dy,$$

which is Nash's quantity $G(1)$, but with the δ in the term $\log(\Gamma(x,1;y,0)+\delta)$ removed and the integral is over the initial time slice. In fact, Fabes-Stroock show that $G \geq -C(n,\lambda)$ for all $x \in B(0,1)$, whereas the lower bound Nash proved for G degenerates as $\delta \searrow 0$. Moreover, the proof of this G-estimate is considerably simpler than Nash's, without any deep new ideas. Once this bound is established, we get a near-diagonal lower bound for Γ using the semigroup property and Jensen's inequality:

$$\begin{aligned}\log\Gamma(x,2;y,0) &\geq \log\left(\int_{\mathbb{R}^n} \Gamma(x,2;z,1)\Gamma(z,1;y,0)e^{-\pi|z|^2}dz\right)\\ &\geq \int_{\mathbb{R}^n} \left(\log\Gamma(x,2;z,1)+\log\Gamma(z,1;y,0)\right)e^{-\pi|z|^2}dz \geq -2C(n,\lambda)\end{aligned}$$

whenever $x,y \in B(0,1)$. By parabolic rescaling, and iterating this inequality (again using the semigroup property), we get the desired Gaussian lower bound.

Holder Continuity via Gaussian Bounds

To give a proof of Theorem 1 using Theorem 3, we first prove a lower bound for the fundamental solution $\widehat{\Gamma}(x,t;y,s)$ with Dirichlet boundary conditions on $B(0,1)$. The strong Markov property for Γ gives the formula

$$\widehat{\Gamma}(x,t;y,0) = \Gamma(x,t;y,0) - \int_{\partial B(0,1)\times[0,t)} \Gamma(x,t;\xi,\tau)d\mu_{0,y}(\xi,\tau)$$

for $x,y \in B(0,1)$, where $\mu_{0,y}$ is a positive measure with total mass ≤ 1. Now use the near-diagonal lower bound to estimate $\Gamma(x,t;y,0)$ from below, then the Gaussian upper bound to estimate the integral, obtaining

$$\widehat{\Gamma}(x,t;y,0) \geq c(n,\lambda,\delta,\gamma)$$

for all $x,y \in B(0,\delta)$, $s \geq 0$, and $t \in (s+\gamma, s+1)$. By a standard iteration procedure, Theorem 1 will follow from proving the oscillation estimate for some $\alpha = \alpha(\lambda,n) \in (0,1)$:

$$\operatorname{osc}_{B(x,\frac{1}{2})\times[\frac{3}{4},1]} u \leq \alpha \operatorname{osc}_{B(x,1)\times[0,1]} u$$

for all solutions $u \in C^\infty(\mathbb{R}^n \times [0,1[)$ of $\partial_t u = L_t u$. Set $M(r) := \sup_{B(x,r)\times[1-r^2,1]} u$, $m(r) := \inf_{B(x,r)\times[1-r^2,1]}$. By replacing u with $-u$, we can assume

$$S := \{x \in B(0,1); u(x,0) \geq \frac{1}{2}(M(1)+m(1))\}$$

satisfies $|S| \geq \frac{\omega_n}{2}$. Then, for all $(x,t) \in B(0,\frac{1}{2}) \times [\frac{3}{4},1]$,

$$\begin{aligned} u(x,t) - m(1) \geq & \int_{B(x,1)} (u(y,0)-m(1))\,\widehat{\Gamma}(x,t;y,0)dy \geq \frac{1}{2}(M(1)-m(1)) \int_S \widehat{\Gamma}(x,t;y,0)dy \\ \geq & c(n,\lambda)(M(1)-m(1)). \end{aligned}$$

Taking the infimum over (x,t) and rearranging gives the claim.

References

[1] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.

[2] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338, 1986.

[3] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.