POINTEWISE DIFFERENTIABILITY OF WEAK SOLUTIONS OF PARABOLIC EQUATIONS WITH MEASURABLE COEFFICIENTS

Pawel Strzelecki
Uniwersytet Warszawski, Instytut Matematyki
ul. Banacha 2, PL-02-097 Warszawa, Polska (Poland); pawelst@plearn.bitnet

Abstract. We prove that weak solutions of a parabolic equation with measurable coefficients $u_t - \text{div}[a(x,t) \cdot u_x] = (b_x, u_x)$ are totally differentiable (in the classical sense) almost everywhere with respect to the Lebesgue measure if $u_t \in L_{\text{loc}}^{\infty,1}$.

1. Introduction

In this note, we consider weak solutions of a parabolic equation

$$
\frac{\partial u}{\partial t} - \sum_{k,l=1}^{n} \frac{\partial}{\partial x_k} \left( a_{kl}(x,t) \cdot \frac{\partial u}{\partial x_l} \right) = \sum_{l=1}^{n} b_l(x,t) \cdot \frac{\partial u}{\partial x_l}.
$$

We assume that the equation (1) is defined in an open domain $Q = G \times (0,T) \subset \mathbb{R}^{n+1}$, where $G$ is an open domain in $\mathbb{R}^n$ and $T > 0$, and that the coefficients $a_{kl} = a_{lk}$, $b_l$ are bounded measurable functions of $(x,t)$ fulfilling the following conditions with some $K \geq 1$:

$$
K^{-1} \leq \sum_{k,l} a_{kl}(x,t) V_k V_l \leq K
$$

for all $(x,t) \in Q$ and all unit vectors $V \in \mathbb{R}^n$, and

$$
\sum_{k} (b_k(x,t))^2 \leq K^2
$$

for all $(x,t) \in Q$. Let $W^{1,2}(Q)$ denote the Sobolev space of square integrable functions on $Q$ with first order distributional partial derivatives in $L^2(Q)$. Then $u \in W^{1,2}(Q)$ is called a weak solution of (1) if and only if the integral identity

$$
\iint_{Q} \left( \varphi \frac{\partial u}{\partial t} + \sum_{k,l} \frac{\partial \varphi}{\partial x_k} a_{kl} \frac{\partial u}{\partial x_l} - \sum_{l} \varphi b_l \frac{\partial u}{\partial x_l} \right) dx \, dt = 0
$$

1991 Mathematics Subject Classification: Primary 35K10; Secondary 58G11.

The author is partially supported by a grant from KBN.
holds for every infinitely differentiable function \( \varphi = \varphi(x,t) \) which, for every fixed \( t \), has compact support as a function of \( x_1, \ldots, x_n \) and vanishes on the parabolic boundary \( \Gamma \) of the cylinder \( Q \), \( \Gamma = (\partial G \times [0,T]) \cup (G \times \{0\}) \). If \( u \) is a weak solution to (1), then (4) holds also for every function \( \varphi \in W^{1,2}(Q) \) having zero trace on \( \Gamma \). In the sequel, we shall abbreviate (4) to

\[
(5) \quad \iint_{\mathcal{Q}} (\varphi u_t + (\varphi_x, au_x) - \varphi(b, u_x)) \, dx \, dt = 0.
\]

Known theorems of Moser [10], Ivanov [7], Aronson and Serrin [1], Trudinger [14], DiBenedetto [5] and others ensure Hölder continuity of different classes of solutions of linear and nonlinear parabolic equations with measurable coefficients. For elliptic equations with measurable coefficients, apart from the fundamental de Giorgi type theorems ascertaining Hölder continuity of weak solutions a number of results are known about almost everywhere (abbreviated to a.e.) differentiability ofweak solutions, giving additional geometric information about the regularity of such solutions. We shall mention here three examples: a theorem of Mori [8] stating the a.e. differentiability of quasiconformal mappings (in fact, the same theorem has been also proved by Bojarski in his paper [2] even without the assumption of bijectivity; a concise proof of this stronger result may be found in the paper of Gehring and Lehto [6])—the connection with elliptic equations is via the Beltrami equation; a theorem of Bojarski [3] about the a.e. differentiability of weak solutions of \( \text{div} (a(x)u_x) = 0 \), and a result of Reshetnyak [11] concerning the more general case of the quasilinear equation \( \text{div} A(x, u, u_x) = B(x, u, u_x) \).

In this paper, we prove a similar result for solutions of the parabolic equation (1), namely the following

**Theorem.** Each weak solution \( u \in W^{1,2}(Q) \) of the equation (1), such that \( u_t \in L_{\text{loc}}^{\infty,1}(Q) \), is differentiable almost everywhere with respect to the Lebesgue measure in \( Q \).

We recall two definitions. A measurable function \( \psi \) is said to be of class \( L^{p,q}(Q) \) if

\[
\int_0^T \left( \int_G |\psi(x,t)|^p \, dx \right)^{q/p} \, dt < +\infty,
\]

with obvious modifications in the case when \( p \) or \( q \) are equal to \( +\infty \). We say that a function \( u \) is totally differentiable at a point \( X_0 = (x_0, t_0) \) if there exists a linear map \( L \) from \( \mathbb{R}^{n+1} \) to \( \mathbb{R} \) such that

\[
u(X_0 + h) - u(X_0) - Lh = o(|h|) \quad \text{for } h \in \mathbb{R}^{n+1} \text{ as } h \to 0.
\]

The assumption that \( u_t \in L_{\text{loc}}^{\infty,1} \) (which at first glance may seem somewhat artificial) is used only in the last part of the proof, when we apply the classical
Rademacher–Stepanoff differentiability criterion (see [13]). The idea of the proof comes from Bojarski’s paper [3]. The main point of the reasoning is provided by the Main Lemma (whose exact formulation is given in Section 2), which roughly speaking states that all the weak solutions of (1) are in fact at almost every point Lipschitz continuous in the space directions and Hölder continuous with exponent equal to 1/2 in the time direction. The essential difference between the parabolic case and the elliptic one is that in the latter case the pointwise differentiability a.e. is easily derived (and there is no need of any additional assumption) from the fact that the solution is at almost every point not only Hölder continuous, but also Lipschitz continuous in all directions.

**Notation.** $X = (x, t), Y = (y, s)$ denote points of $\mathbb{R}^{n+1}$. By $R(a) = R(X_0, a)$ we denote a rectangle with center at $X_0$ and edges parallel to the coordinate axes in $\mathbb{R}^{n+1}$ such that those parallel to $x_1, \ldots, x_n$ are of length $a$, and the edge parallel to that of length $a^2$,

$$R(X_0, a) = \{(x, t) : \max_{1 \leq i \leq n} |x_0, x_i| < a/2, |t_0 - t| < a^2/2\}.$$ 

Rectangles of this type will sometimes be called *parabolic*. The barred integral $\overline{\int_A f} dx dt$ denotes the average value $f_A$ of a function $f$ on a measurable set $A$, $f_A := |A|^{-1} \int_A f$ dx dt. We also use the following notation: $u_x = (u_{x_1}, \ldots, u_{x_n})$ denotes the gradient of $u$ with respect to space variables only, whereas $\nabla u = (u_x, u_t)$ stands for the full gradient.

A function $u \in W^{1,2}(Q)$ is called a subsolution of (1) if

$$\overline{\iint_Q (\varphi u_t + (\varphi_x, a u_x) - \varphi(b, u_x))} dx dt \leq 0$$

for every $\varphi \in C^\infty(\overline{Q})$, $\varphi \geq 0$, $\varphi = 0$ on $\Gamma$.

**2. The main lemma**

*If* $u$ *is a weak solution of (1) in* $Q$ *and* $R(X_0, 2a) = R(2a) \subset Q$, $2a < 1$, *then*

$$\text{ess sup}_{X, Y \in R(a)} |u(X) - u(Y)|^2 \leq C \cdot a^2 \overline{\iint_{R(2a)}} |\nabla u(x, t)|^2 dx dt,$$

*where* $C$ *is a constant depending only on* $n$ *and* $K$. 

Proof. Our proof uses the iteration technique of Moser and with some minor modifications is based on his proof of local boundedness of weak solutions of the equation \( u_t - \text{div} [a(x,t)u_x] = 0 \) (cf. [10, Theorem 3]).

We start with two inequalities, which play the same role as the Cacciopoli estimates in the elliptic case. They are valid for arbitrary solutions and nonnegative subsolutions of (1):

\[
\text{(8)} \quad \text{ess sup}_{s \in I'} \int_{R' \cap \{s = t\}} v^2 \, dx \leq C \cdot \left( \frac{1}{(\theta - \theta')^2} + \frac{1}{(\tau - \tau')^2} \right) \cdot \iint_R v^2 \, dx \, dt,
\]

\[
\text{(9)} \quad \iint_{R'} |v_x|^2 \, dx \, dt \leq C \cdot \left( \frac{1}{(\theta - \theta')^2} + \frac{1}{(\tau - \tau')^2} \right) \cdot \iint_R v^2 \, dx \, dt,
\]

where

\[
R = P \times I = \{ x : \max_i |x_{0,i} - x_i| < \theta/2 \} \times (t_2 - \tau, t_2),
\]

\[
R' = P' \times I' = \{ x : \max_i |x_{0,i} - x_i| < \theta'/2 \} \times (t_2 - \tau', t_2),
\]

\(0 < \theta' < \theta, \, 0 < \tau' < \tau\), and the integration on the left hand side of (8) is performed over a cross section \( t = s \) and then the essential maximum is taken (of course, \( R' \) is a subrectangle of \( R \)). To prove (8) and (9), we take \( \varphi = v \cdot \psi^2 \) as a testing function (assume that \( \psi \in C^\infty(Q), \, \psi \geq 0, \, \psi = 0 \) on \( \Gamma \)), so that

\[
(\varphi_x, av_x) = 2\psi v \cdot (\psi_x, av_x) + \psi^2 \cdot (v_x, av_x),
\]

and from (4) (or (6), if \( v \) is merely a nonnegative subsolution) we obtain:

\[
\iint (v v_t \psi^2 + \psi^2 \cdot (v_x, av_x)) \, dx \, dt
\]

\[
\leq 2 \iint |v \psi(\psi_x, av_x)| \, dx \, dt + \iint \psi^2 |v(b, v_x)| \, dx \, dt.
\]

We estimate the integrands on the right hand side using the symmetry of the matrix \( (a_{kl}) \), Schwarz inequality and the assumptions (2), (3):

\[
|\psi v(\psi_x, av_x)| \leq (v^2 \cdot (\psi_x, av_x) \cdot \psi^2 (v_x, av_x))^{1/2}
\]

\[
\leq 2K^2 v^2 \cdot (\psi_x, av_x) + 8^{-1}K^{-2} \psi^2 \cdot (v_x, av_x)
\]

\[
\leq 2K^3 v^2 |\psi_x|^2 + (8K)^{-1} \psi^2 |v_x|^2;
\]

\[
\psi^2 |v(b, v_x)| \leq 2K^3 \psi^2 v^2 + (8K)^{-1} \psi^2 |v_x|^2.
\]
On the left hand side, by (2), $\psi^2(v_x, av_x) \geq K^{-1} \psi^2 |v_x|^2$; suitable arrangement of all the terms leads to
\[
\frac{1}{2} \iint (v^2)_{t} \psi^2 \, dx \, dt + \frac{1}{2K} \iint \psi^2 |v_x|^2 \, dx \, dt \leq 4K^3 \iint v^2 (\psi^2 + |\psi_x|^2) \, dx \, dt
\]
and hence
\[
\iint (v^2 \psi^2)_{t} \, dx \, dt + \frac{1}{K} \iint \psi^2 |v_x|^2 \, dx \, dt
\]
(10)
\[
\leq 8K^3 \iint v^2 (\psi^2 + |\psi_x|^2 + |\psi_x|) \, dx \, dt.
\]
In this inequality we integrate over the rectangle $R_s = P \times (t_2 - \tau, s)$. The above argument may be applied to $R_s$ since we allow the testing function $\varphi$ to take nonzero values on the upper part of the boundary $t = s$. If we choose $\psi \equiv 1$ in $R'$, $\psi = 0$ for $t = t_2 - \tau$ and $s$ such that the integral
\[
\int_{P' \times \{t=s\}} v^2(x, s) \, dx
\]
exceeds one half of the essential supremum of this quantity taken over $t \in I'$, then, disregarding the second integral on the left hand side of (10), we have
\[
\text{ess sup}_{t \in I'} \int_{x \in P'} v^2(x, t) \, dx \leq 2 \int_{P' \times \{s\}} v^2(x, s) \, dx \leq 2 \iint (v^2 \psi^2)_{t} \, dx \, dt
\]
\[
\leq 16K^3 \iint v^2 (\psi^2 + |\psi_x|^2 + |\psi_x|) \, dx \, dt.
\]
On the other hand, if we set $s = t_2$ in (10), drop the first integral on the left and replace the domain of integration in the second one by $R'$, we obtain:
\[
\iint_{R'} |v_x|^2 \, dx \, dt \leq 8K^4 \iint_{R} v^2 (\psi^2 + |\psi_x|^2 + |\psi_x|) \, dx \, dt.
\]
The last two inequalities are just (8) and (9): if only $\psi$ is chosen to be identically 1 in $R'$, $\psi = 0$ on the parabolic boundary of $R$ and piecewise linear in the remaining parts of $Q$, then one finds:
\[
\psi^2 + |\psi_x|^2 + |\psi_x| \leq C_n \cdot \left( \frac{1}{(\varphi - \varphi')^2} + \frac{1}{\tau - \tau'} \right)
\]
for \( q, q', \tau, \tau' < 1 \).

If \( u \) is a solution of (1), and \( g \) a nonnegative convex function of class \( C^1 \) (or piecewise \( C^1 \)), then \( v = g(u) \) is a nonnegative subsolution of (1). The argument is essentially the same as in Moser’s paper [10, p. 117]. This allows us to write the inequalities (8) and (9) for functions \( v = |u - u_0|^{\tau} \), where \( \tau \geq 1 \) and \( u_0 = u_{R(a)} = \iint_{R(a)} u(x,t) \, dx \, dt \).

The inequalities (8) and (9) will now be combined with a general inequality (which is a consequence the Sobolev embedding theorem) in order to obtain the so-called weak inverse Hölder inequality. For a rectangle \( R = P \times I \) (\( x \in P, t \in I \)) with edges of length \( q \) in the space direction and of length \( \tau \) in the time direction we define:

\[
H_{q, \tau}(w) \overset{\text{def}}{=} \iint_{R} w^2 \, dx \, dt,
\]

\[
D_{q, \tau}(w) \overset{\text{def}}{=} q^2 \iint_{R} |w_x|^2 \, dx \, dt,
\]

\[
M_{q, \tau}(w) \overset{\text{def}}{=} \operatorname{ess sup}_{t \in I} \int_{P} w^2(x,t) \, dx.
\]

As a consequence of the Sobolev inequality, written for the \( x \) variables and then integrated with respect to the time \( t \), one obtains the following

**Lemma.** (Moser [10]). For every function \( w \) for which the integrals \( H, D, \) and \( M \) exist we have

\[
H(w^{\kappa}) \leq C(n) \cdot M(w)^{\kappa-1} \cdot (H(w) + D(w)),
\]

where \( \kappa = 1 + 2/n \) for \( n = 3, 4, \ldots \) and \( \kappa = 5/3 \) for \( n = 1, 2 \).

Expressing (8) and (9) in terms of \( H, D \) and \( M \) we obtain for \( \frac{1}{4} q < q' < q \) and \( \frac{1}{4} \tau < \tau' < \tau \)

\[
M_{q', \tau'}(v) \leq C(n, K) \cdot \left( \frac{\tau'}{(q - q')^2} + \frac{\tau'}{\tau - \tau'} \right) \cdot H_{q, \tau}(v),
\]

\[
D_{q', \tau'}(v) \leq C(n, K) \cdot \left( \frac{q'^2}{(q - q')^2} + \frac{q'^2}{\tau - \tau'} \right) \cdot H_{q, \tau}(v).
\]

These inequalities together with the above lemma imply:

\[
H_{q', \tau'}(v^{\kappa}) \leq A \cdot (H_{q, \tau}(v))^{\kappa},
\]
where
\[ A = C_1(n, K) \cdot \left( \frac{\tau'}{(\varrho - \varrho')^2} + \frac{\varrho'}{\tau - \tau'} \right)^{\kappa - 1} \cdot \left( 1 + \frac{\varrho'^2}{(\varrho - \varrho')^2} + \frac{\varrho'^2}{\tau - \tau'} \right). \]

We shall apply (13) to the successive powers of solution and to the sequence of nested rectangles with common upper boundary. For \( m = 0, 1, 2, \ldots \) let \( v_m = |u - u_{R(a)}|^{\kappa^m} \) (so that \( v_{m+1} = v_m^2 \)) and define rectangles \( R_m \)
\[ R_m = \{ x : \max_i |x_{0,i} - x_i| < \varrho_m/2 \} \times (t_0 - \tau_m + 2a^2, t_0 + 2a^2), \]
where
\[ \varrho_m = (1 + 2^{-m}) \cdot a, \]
\[ \tau_m = \left( 3^{1/2} + (2 - 3^{1/2}) \cdot 2^{-m} \right)^2 \cdot a^2. \]
It is clear that \( R(2a) = R_0 \supset R_1 \supset \ldots \supset \bigcap_{m=0}^{\infty} R_m \supset R(a) \). Then, letting \( v = v_m, \varrho = \varrho_m, \tau = \tau_m, \varrho' = \varrho_{m+1}, \) and \( \tau' = \tau_{m+1} \) in (13), we find after a simple calculation the following inequality for \( H_m := H_{\varrho_m, \tau_m}(v_m) \):
\[ H_{m+1} \leq C_1(n, K) \cdot 20^{\kappa - 1} \cdot 6 \cdot (9^\kappa)^{m+1} \cdot (H_m)^\kappa \leq C_2^{m+1} \cdot (H_m)^\kappa. \]
Hence, by induction,
\[ (H_{m+1})^{1/\kappa^{m+1}} \leq C_3 \sum_{j=1}^{m+1} j/j^{\kappa^j} \cdot H_0 \leq C_3 \cdot H_0 \]
\[ = C_3 \cdot \iint_{R(2a)} |u(x,t) - u_{R(a)}|^2 \, dx \, dt. \]
If we now define, for \( m \in \mathbb{N} \),
\[ \Phi_m \overset{\text{def}}{=} \left( \iint_{R(a)} \iint_{R(a)} |u(x,t) - u(y,s)|^{2 \cdot \kappa^m} \, dx \, dt \, dy \, ds \right)^{1/\kappa^m}, \]
then we can use the Minkowski inequality to find that:
\[ \Phi_m \leq 4 \left( \iint_{R(a)} |u(x,t) - u_{R(a)}|^{2 \cdot \kappa^m} \, dx \, dt \right)^{1/\kappa^m} \]
\[ \leq 2^{n+4} \left( \iint_{R_m} |u(x,t) - u_{R(a)}|^{2 \cdot \kappa^m} \, dx \, dt \right)^{1/\kappa^m} = 2^{n+4} (H_m)^{1/\kappa^m}, \]
and thanks to (14):

\[ (15) \quad \Phi_m \leq C_4 \cdot \iint_{R(2a)} |u(x, t) - u_{R(a)}|^2 \, dx \, dt. \]

Finally, we shall apply Poincaré's inequality to estimate the right hand side of (15). Assume with no loss of generality that \( X_0 = 0 \), then change integration variables from \((x, t) \in R(2a)\) to \((y, s) = (x/a, t/2a^2)\), estimate the integral (over cube!) with the help of Poincaré's inequality in its standard form, and change variables back to \((x, t)\) to obtain:

\[ \Phi_m \leq C_5 \cdot a^2 \cdot \iint_{R(2a)} |\nabla u(x, t)|^2 \, dx \, dt. \]

This inequality obviously implies (7) upon letting \( m \to +\infty \). The proof of the Main Lemma is complete. \( \Box \)

3. Proof of the theorem

It is possible to show that for each function \( f \in L^1 \)

\[ (16) \quad \lim_{a \to 0} \int_{R(X_0, a)} f(X) \, dX = f(X_0) \quad \text{for a.e. } X_0 \]

(see [12, Chapter 1, Section 5.3 (d)], or [4, Chapter 1, Section 3]). Let \( X_0 = (x_0, t_0) \in Q = G \times (0, T) \) be a point for which (16) holds (with \( f \) replaced by \(|\nabla u|^2\)). In order to apply Stepanoff's criterion and finish the proof it is sufficient to show that, for almost all such points \( X_0 \),

\[ \limsup_{h \to 0} \frac{|u(X_0 + h) - u(X_0)|}{|h|} < +\infty. \]

We choose an additional point \( Y \) with the time coordinate equal to that of \( X_0 \) and with spatial coordinates equal to those of \( X = X_0 + h = (x, t) \). The triangle inequality gives:

\[
|u(X_0) - u(X)| \leq |u(X_0) - u(Y)| + |u(Y) - \int_P u(y, t_0) \, dy| \\
+ \left| \int_P [u(y, t_0) - u(y, t)] \, dy \right| + \left| \int_P u(y, t) \, dy - u(X) \right| \\
= S_1 + S_2 + S_3 + S_4.
\]
In view of the Main Lemma, $S_1$ on the right hand side of the above inequality is bounded from above by a constant times $|h|$. In order to estimate $S_2$ and $S_4$ we are going to make use of the local Hölder continuity of weak solutions of (1). Select $P$ to be an $n$-dimensional cube in the space $\mathbb{R}^n$ of space variables, with center at $x$ and edge $|h|^{\gamma}$, where $\gamma = 1/\alpha$ is chosen to be equal to the reciprocal of the Hölder exponent of the function $u$. Then, for $y \in P$, one has $|u(Y) - u(y, t_0)| \leq C \cdot (|h|^{\gamma})^{\alpha} = C \cdot |h|$, and in a similar manner $|u(X) - u(y, t)| \leq C \cdot |h|$. Finally, to estimate $S_3$ we apply the assumption $u_t \in L_{loc}^{\infty, 1}$. Let $I$ denote the interval $(t_0, t)$. Then $|I| \leq |h|$ and we see that

$$\left| \int_P [u(y, t_0) - u(y, t)] \, dy \right| \leq |h| \cdot \int_P \int_I |u_t(y, s)| \, ds \, dy \leq |h| \cdot \int_I M(s) \, ds,$$

where

$$M(s) = \text{ess sup}_y |u_t(y, s)|$$

is a locally integrable function of one real variable $s$. The classical differentiation theorem of Lebesgue implies now that for $\mathbb{R}^{n+1} \ni h \to 0$, i.e. for $t \to t_0$ we have

$$\int_I M(s) \, ds \longrightarrow M(t_0) \quad \text{for a.e. } t_0.$$

Putting all these estimates together, we conclude that for sufficiently small $|h|$: $|u(X_0 + h) - u(X_0)| \leq C \cdot |h|$, a.e. $X_0$, where the constant $C$ depends on $n$, $K$, $|\nabla u(X_0)|^2$, $M(t_0)$ and the local Hölder norm of $u$. We may now apply Stepanoff’s criterion and obtain the a.e. differentiability of $u$ in the classical sense. The proof is complete. \Box

4. Remarks

1) The Main Lemma is obtained before we even know that weak solutions of (1) are Hölder continuous. In fact, the proof is simpler than Moser’s original proof of Hölder continuity of weak solutions of $u_t - \text{div} [a(x, t)u_x] = 0$ (we do not need Moser’s remarkable version of John and Nirenberg’s lemma on BMO functions).

2) The method of proof used in [11] evidently fails in the parabolic case, because the linear transformations $X \mapsto hX + X_0$ do not conserve the structure of (1)—this is due to the asymmetry of derivatives with respect to time and space variables.

3) There is one general problem connected with the last part of our proof: Is it true that for $f \in L^2$ (or generally for $f \in L^p$, with $p > 1$)

$$\limsup_{h \to 0} \int_P \int_{t-h}^{t+h} |f(y, s)| \, ds \, dy < +\infty \quad \text{for a.e. } (x, t),$$
where $P$ may be an arbitrary $n$-dimensional cube with center $x'$ not too far from $x$, i.e. $\text{dist}(x', x) < h$, and edge of length $h^{\gamma'}$ for some fixed $\gamma > 1$? A positive answer to the above question could give the possibility to state our Theorem without the assumption about the time derivative of solution.

4) The Main Lemma and the Theorem can be generalized to the case of quasilinear parabolic equations. Further development of the topics considered here, together with their possible applications to the regularity theory of parabolic systems, are object of a forthcoming paper.

Acknowledgement. I would like to express my sincere gratitude to Professor Bogdan Bojarski and Dr. Jan Herczyński for their helpful comments concerning the first draft of this paper. I am also extremely grateful to Dr. Wojciech Chojnacki for his encouragement and advice.

References


Received 19 March 1991