HÖLDER ESTIMATES FOR SOLUTIONS OF DEGENERATE NONDIVergence ELLIPTIC AND PARABOLIC EQUATIONS

A. I. NAZAROV

To my teacher Nina Nikolaevna Uraltseva on the occasion of her birthday.

Abstract. We deal with a class of nondivergence type elliptic and parabolic equations degenerating at the coordinate hyperplanes. Assuming that the degeneration is coordinatewise and varies regularly, we prove the Hölder continuity of solutions. Also, the approximative solutions are considered.

§1. Introduction

Let \( n \geq 2 \), and let \( \Omega \) be a bounded domain in \( \mathbb{R}^n \) containing the origin. We consider an elliptic equation of nondivergence form

\[
Lu \equiv - \sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j} = f(x) \quad \text{in} \quad \Omega,
\]

with measurable coefficients. We suppose that a symmetric matrix \( A = (a_{ij}) \) has diagonal degeneration at coordinate hyperplanes \( x_k = 0, \ k = 1, \ldots, n \). Namely, we assume that

\[
A = \sqrt{\Lambda} \hat{A} \sqrt{\Lambda},
\]

where \( \Lambda \) is a diagonal matrix of specific form responsible for degeneration:

\[
\Lambda(x) = \text{diag}\{\lambda_1(|x_1|), \lambda_2(|x_2|), \ldots, \lambda_n(|x_n|)\}.
\]

We obtain local a priori estimates of the Hölder norm for solutions of (1.1) and for solutions of the corresponding parabolic equation

\[
Mu \equiv \partial_t u - \sum_{i,j=1}^{n} a_{ij}(x,t)u_{x_i x_j} = f(x,t) \quad \text{in} \quad Q = \Omega \times [0,T[.
\]

For uniformly elliptic and uniformly parabolic equations such estimates were established in the classical paper [1]; see also [2]. In [3], this result was generalized to the equations with unbounded lower-order coefficients. Another method for obtaining Hölder estimates was proposed in [4]. The Hölder continuity of solutions of some nonuniformly elliptic (parabolic) equations was proved in [5].

2010 Mathematics Subject Classification. Primary 35B45; Secondary 35J70, 35K65.
Key words and phrases. Nondivergence equations, Hölder estimates, degeneration, regularly varying functions.

The paper is supported by the grant NSh.227.2008.1 and by RFBR grant 08-01-00748.
Equations with diagonal degeneration (1.1)–(1.4) have been studied in some particular cases since the 1960s, see, e.g., [3], but only for a smooth coefficient matrix $\mathcal{A}$. Recently, these equations with only measurable coefficients have become of growing interest. One of the reasons is the active development of the theory of free boundary problems related to Gaussian curvature flows (see [7] and also [8, 9, 10]).

In [11], the Harnack inequality and Hölder estimates were obtained for solutions of an equation that can be reduced to (1.1)–(1.4) with

$$n = 2, \quad \lambda_1(t) = 1, \quad \lambda_2(t) = t^\alpha, \quad \alpha < 1.$$ 

It is important that in [11] equation (1.1) was considered in a domain located in the half-space $x_2 > 0$. Moreover, in essence, the Neumann condition was imposed on $\partial \Omega \cap \{x_2 = 0\}$. Thus, the even reflection with respect to $x_2 = 0$ reduces this problem to that of ours. On the other hand, a barrier function constructed in [11] is not available when a degeneration plane intersects the domain. It should be noted that the validity of the corresponding results of [11] for parabolic equations is doubtful. The point is that there is a gap in the proof of Theorem 3.2 in [11] (a version of the parabolic Aleksandrov-type maximum principle). Namely, it should be $\rho^\frac{1}{2}$ in place of $\rho^\frac{1}{4}$ on the right-hand side of the sup $u_+$ estimate. Since the proofs of subsequent statements in the parabolic case are omitted, it is not clear whether they remain valid after this correction.

In the paper [12] (see also [13]) the Hölder continuity of approximative solutions of the homogeneous equation (1.1)–(1.4) was proved under the condition

$$\lambda_k(t) = t^{\alpha_k}, \quad -\frac{1}{n-1} < \alpha_k < 1, \quad k = 1, \ldots, n.$$ 

The method of [12] [13] develops the techniques of [3] and is based on the Green function estimates for the uniformly elliptic equations with smooth coefficients approximating the equation conjugate to (1.1). Substantial analytic difficulties occur along this way, and the result was obtained only under hard restrictions on $\alpha_k$.

We establish the Hölder continuity of solutions of (1.1) and (1.5) under the condition of a regular behavior of the functions $\lambda_k$ at a neighborhood of the origin; see §2. In particular, the functions $\lambda_k(t) = t^{\alpha_k}$ are admissible for arbitrary $\alpha_k < 1$.

Our method is based on the classical barrier techniques. To construct the barriers, we uniformize the equation and reduce it to a uniformly elliptic (uniformly parabolic) equation with lower-order terms having strong singularities at the coordinate hyperplanes. For the reduced equations, barrier functions are constructed. The point is that our estimates for these barriers are uniform with respect to the distance from the coordinate hyperplanes. As we show, this approach can easily be modified to cover approximative solutions.

The paper is organized as follows. In §2, we collect auxiliary lemmas about regularly varying functions. §3 is devoted to uniformization of the equation and to barrier functions. In §4, Hölder continuity is proved for solutions of elliptic equations, and in §5 this is done for solutions of parabolic equations. Finally, in §6 we consider approximative solutions.

We introduce some notation. $K_{\rho}(x)$ stands for the $n$-dimensional cube centered at $x$, with $2\rho$-long edges parallel to coordinate axes; $K_{\rho} = K_{\rho}(0)$. If $x = (x_1, \bar{x})$ (i.e., $\bar{x}$ is the projection of $x$ to the hyperplane $x_1 = 0$), then $K_{\rho}^\prime(\bar{x})$ stands for the $(n-1)$-dimensional projection of $K_{\rho}(x)$.

For a set $E$, we denote by $|E|$ its $n$-dimensional $(n+1)$-dimensional in §5) Lebesgue measure. We also set $E^+ = E \cap \mathbb{R}_+^n$, where $\mathbb{R}_+^n$ is the positive orthant in $\mathbb{R}^n$.

We put $f_{\pm} = \max\{|f|, 0\}$. If $u$ is a continuous function, then $A_u^s = \{x : u(x) > s\}$ stands for its level set.
§2. About regularly varying functions

Recall that a positive function \( \varphi(\tau), \tau > 0 \), is \textit{regularly varying of order} \( \alpha \) near the origin \( \{0\} \) (we shall write \( \varphi \in \mathcal{R}_\alpha \)) if for any \( c > 0 \) we have
\[
\varphi(c\tau)/\varphi(\tau) \to c^\alpha \quad \text{as} \quad \tau \to +0.
\]

In a similar way we define regularly varying functions near infinity.

The properties of regularly varying functions are well known. We list some of them that we shall use. The proofs can be found, for example, in the classical book [14].

Proposition 2.1. 1. If \( \varphi \in \mathcal{R}_\alpha \), then for all \( 0 < c_1 < c_2 < +\infty \) the limit in (2.1) is uniform with respect to \( c \in [c_1, c_2] \).

2. If \( \varphi_1 \in \mathcal{R}_{\alpha_1}, \varphi_2 \in \mathcal{R}_{\alpha_2} \), then \( \varphi_1 \varphi_2 \in \mathcal{R}_{\alpha_1 + \alpha_2} \). If \( \varphi \in \mathcal{R}_\alpha \), then \( 1/\varphi \in \mathcal{R}_{-\alpha} \).

3. If \( \varphi_1 \in \mathcal{R}_{\alpha_1}, \varphi_2 \in \mathcal{R}_{\alpha_2}, \) and \( \alpha_1 > 0 \), then \( \varphi_2 \circ \varphi_1 \in \mathcal{R}_{\alpha_1 \alpha_2} \). In particular, \( \varphi_1^\alpha \in \mathcal{R}_{\alpha_1 \alpha_2} \).

4. If \( \varphi \in \mathcal{R}_\alpha, \alpha > 0 \), then \( \varphi \) increases monotonically in some neighborhood of the origin, and \( \varphi^{-1} \in \mathcal{R}_{1/\alpha} \).

5. If \( \varphi \in \mathcal{R}_\alpha, \alpha > -1 \), then the integral \( \psi(t) = \int_0^t \varphi(\tau) \, d\tau \) converges, and \( \psi \in \mathcal{R}_{\alpha + 1} \).

6. If \( \varphi \in \mathcal{R}_{\alpha_1} \), and \( \tau \varphi'(\tau)/\varphi(\tau) \to \alpha \) as \( \tau \to +0 \), then \( \varphi \in \mathcal{R}_\alpha \). Moreover, if \( \varphi \in \mathcal{R}_\alpha \), then there exists a continuously differentiable function \( \varphi_1 \in \mathcal{R}_\alpha \) such that
\[
\frac{\varphi_1(\tau)}{\varphi(\tau)} \to 1, \quad \frac{\tau \varphi_1'(\tau)}{\varphi_1(\tau)} \to \alpha \quad \text{as} \quad \tau \to +0.
\]

In what follows we need some statements about measures generated by regularly varying densities. The first statement resembles the well-known Chebyshev inequality.

Lemma 2.2. Suppose \( \varphi_k \in \mathcal{R}_{\alpha_k}, k = 1, 2, \) with \( \alpha_1 > -1, \alpha_2 > -1, \alpha_1 + \alpha_2 > -1 \). For \( -1 \leq t_1 < t_2 \leq 1 \) we have
\[
(2.3) \quad \frac{(t_2 - t_1) \int_{t_1}^{t_2} \varphi_1(|\tau|) \varphi_2(|\tau|) \, d\tau}{\int_{t_1}^{t_2} \varphi_1(|\tau|) \, d\tau} \geq 1.
\]

Prove. Without loss of generality, \( t_2 \geq |t_1| \). Note that, replacing \( \varphi_1 \) and \( \varphi_2 \) with functions equivalent to them at the origin, we keep relation (2.3). Thus, part 6 of Proposition 2.1 allows us to assume that the functions \( \varphi_1 \) and \( \varphi_2 \) are smooth, and
\[
(2.4) \quad \lim_{t \to +0} \frac{t \varphi_1'(t)}{\varphi_1(t)} = \alpha_1, \quad \lim_{t \to +0} \frac{t \varphi_2'(t)}{\varphi_2(t)} = \alpha_2.
\]

By Proposition 2.1, parts 2 and 5, the functions
\[
\psi_1(t) = \int_0^t \varphi_1(\tau) \, d\tau, \quad \psi_2(t) = \int_0^t \varphi_2(\tau) \, d\tau, \quad \psi(t) = \int_0^t \varphi_1(\tau) \varphi_2(\tau) \, d\tau
\]
are regularly varying with exponents \( \alpha_1 + 1, \alpha_2 + 1, \) and \( \alpha_1 + \alpha_2 + 1, \) respectively. We take a number \( N \geq 3 \) such that
\[
(2.5) \quad N^{\alpha_1 + 1} \geq 4, \quad N^{\alpha_2 + 1} \geq 4, \quad N^{\alpha_1 + \alpha_2 + 1} \geq 4.
\]

Upon equivalent replacement, we may assume that
\[
(2.6) \quad \frac{\psi_1(Nt)}{\psi_1(t)} \geq \frac{N^{\alpha_1 + 1}}{2}, \quad \frac{\psi_2(Nt)}{\psi_2(t)} \geq \frac{N^{\alpha_2 + 1}}{2}, \quad \frac{\psi(Nt)}{\psi(t)} \geq \frac{N^{\alpha_1 + \alpha_2 + 1}}{2}
\]
\[1\]For \( \alpha = 0 \) such a function is also called a \textit{slowly varying function} near the origin.
Lemma 2.3. It is also "thin" with respect to a measure with regularly varying density.

(2.9) \( \psi \) (the last relation follows from part 1 in Proposition 2.1).

□

Theorem 1.14, we conclude that if the measure \( \rho \) is located "near" the coordinate hyperplanes \( x_k = 0 \) with \( k \leq m \) and "far" from the coordinate hyperplanes \( x_k = 0 \) with \( k > m \) (here \( 0 \leq m \leq n \)). Namely,

(2.7) \( x_k^0 < \frac{N + 1}{N - 1} \cdot \rho, \quad k = 1, \ldots, m; \quad x_k^0 > \frac{N + 1}{N - 1} \cdot \rho, \quad k = m + 1, \ldots, n. \)

Note that replacing \( \varphi_k \) with equivalent functions leads to multiplication of the measure \( |E| \) of any set \( E \) by a quantity bounded by positive constants from above and from below. So, this does not influence the claim. Therefore, we may assume that the functions \( \psi_k(t) = \int_0^t \varphi_k(\tau) \, d\tau \) satisfy

(2.8) \( \frac{\psi_k(Nt)}{\psi_k(t)} \geq \frac{N \alpha_k + 1}{2}, \quad 0 < t \leq 1/N. \)

Inequalities (2.7) and (2.8) imply that

(2.9) \( \int_{x_k^0 - \rho}^{x_k^0 + \rho} \varphi_k(\tau) \, d\tau \approx \begin{cases} \psi_k(\rho) \approx \rho \varphi_k(\rho) & \text{if } k = 1, \ldots, m, \\ \rho \varphi_k(x_k^0) & \text{if } k = m + 1, \ldots, n, \end{cases} \)

whence

(2.10) \( |K_\rho(x^0)| \approx \rho^n \varphi_1(\rho) \cdots \varphi_m(\rho) \varphi_{m+1}(\rho) \cdots \varphi_n(\rho). \)

Now we estimate the measure \( |E \cap K_\rho(x^0)| \). By the “bathtub principle” (see, e.g., [15, Theorem 1.14]), we conclude that if the measure \( |E \cap K_\rho(x^0)| \) is fixed, then the maximal value of \( |E \cap K_\rho(x^0)| \) is attained if \( E \) is the level set

\( E_s = \{ x : \varphi_1(|x_1|) \cdots \varphi_n(|x_n|) > s \} \).
with a suitably chosen \( s \) (replacing the \( \varphi_k \) with equivalent functions, we may assume without loss of generality that all sets \( \{ \varphi_1(x_1) \ldots \varphi_n(x_n) = s \} \) have zero measure).

Obviously, \( K_\rho(x) \) lies in the parallelepiped
\[
\mathcal{P}(\rho) = \{ x : |x_k| < 3\rho, \ k \leq m; \ x_k^0 - \rho < x_k < x_k^0 + \rho, \ k > m \}.
\]

On the other hand, from (2.10) we see that \( |K_\rho(x)|_* \sim |\mathcal{P}(\rho)|_*. \) Therefore, it suffices to estimate \( |E_s \cap \mathcal{P}(\rho)|_* \), and even \( |E_s \cap \mathcal{P}(\rho)|_* \), by symmetry.

We introduce the sets
\[
\tilde{E}_1(\delta, \rho) = \mathcal{P}^+(\rho) \cap \{ x : x_k > \delta \rho, \ k \leq m \}; \quad \tilde{E}_2(\delta, \rho) = \mathcal{P}^-(\rho) \setminus \tilde{E}_1(\delta, \rho).
\]

Then, obviously, \( |E_s \cap \mathcal{P}(\rho)|_* \lesssim |E_s \cap \tilde{E}_1(\delta, \rho)|_* + |\tilde{E}_2(\delta, \rho)|_* \).

Relations (2.9) and (2.10) imply that, for sufficiently small \( \delta \), we have
\[
|\tilde{E}_2(\delta, \rho)|_* \lesssim |K_p(x)|_* \cdot \sum_{k=1}^m \frac{\psi_k(\delta \rho)}{\psi_k(\rho)}.
\]

Therefore, for any \( \varepsilon > 0 \) there exists \( \delta > 0 \), depending only on \( \varepsilon \) and \( \varphi_k \), such that
\[
|\tilde{E}_2(\delta, \rho)|_* \lesssim \frac{\varepsilon}{2} |K_p(x)|_*.
\]

Next, in the segment \([\hat{\delta}, 1]\) we have \( \varphi_k(\rho z) \approx z^{\alpha_k} \varphi_k(\rho) \). Hence, using the second line in (2.9), we obtain the following relation on the set \( E_1(\hat{\delta}, \rho) \) (here \( z_k = \frac{z}{\rho} \), \( \hat{N} \) depends only on \( \hat{\delta} \) and \( \varphi_k \), and \( \mu \) stands for an arbitrary positive number):
\[
\begin{align*}
\varphi_1(x_1) \ldots \varphi_n(x_n) &> \mu \varphi_1(\rho) \ldots \varphi_m(\rho) \varphi_{m+1}(x_{m+1}) \ldots \varphi_n(x_n) \\
\Rightarrow \varphi_1(x_1) \ldots \varphi_n(x_n) &> \mu \frac{\mu}{\hat{N}}.
\end{align*}
\]

We set \( s = \mu \varphi_1(\rho) \ldots \varphi_m(\rho) \varphi_{m+1}(x_{m+1}) \ldots \varphi_n(x_n) \). Then
\[
|E_s \cap \tilde{E}_1(\hat{\delta}, \rho)|_* \lesssim \mu \int \varphi_1(x_1) \ldots \varphi_n(x_n) \, dx_1 \ldots dx_n
\]
\[
\lesssim \hat{N} \mu^\alpha \varphi_1(\rho) \ldots \varphi_m(\rho) \varphi_{m+1}(x_{m+1}) \ldots \varphi_n(x_n) \cdot \int \varphi_1(x_1) \ldots \varphi_m(\rho) \, dx_1 \ldots dx_m.
\]

By (2.10), there exists \( \mu \), depending only on \( \varepsilon \), \( \hat{N} \), and \( \alpha_k \), \( k = 1, \ldots, m \), such that \( |E_s \cap \tilde{E}_1(\hat{\delta}, \rho)|_* \lesssim \frac{\varepsilon}{2} |K_\rho(x)|_* \). We take the smallest admissible \( \mu \) and \( \delta \) depending on \( \mu \), \( \hat{N} \), and the collection \( \alpha_k \) such that \( |K_\rho^+ \cap \{ z : z_1^{\alpha_1} \ldots z_m^{\alpha_m} > \mu/\hat{N} \}| \geq \delta \) for any \( m \leq n \) and for any choice of \( m \) exponents \( \alpha_k \) out of \( n \). This completes the proof. \( \Box \)

### §3. Coordinate Transformation and Barriers

We suppose that the functions in (1.4) satisfy
\[
(3.1) \quad \lambda_k \in \mathcal{R}_{\alpha_k}, \quad \alpha_k < 1, \quad k = 1, \ldots, n.
\]

Note that, after replacing \( \lambda_k \) with functions equivalent at zero, equation (1.1) keeps its structure, and the constant \( \nu \) in (1.3) can be controlled. Thus, by part 6 of Proposition 2.1, we may assume without loss of generality that \( \lambda_k \in \mathcal{C}^1([0, 1]) \) and \( t \lambda_k'(t)/\lambda_k(t) \to \alpha_k \) as \( t \to 0 \). We extend the \( \lambda_k \) to the negative half-axis as even functions.
Next, since Hölder continuity is a local property, it suffices to establish this property in a cube $K_{R_0}(x^0) \Subset \Omega$ with a small (given) edge of length $2R_0$ and arbitrary center $x^0$. At the same time, if $x^0$ is situated “far” from the $k$th coordinate hyperplane (i.e., $|x^0_k| \geq 2R_0$), we may assume (shifting the origin and changing $\nu$ if needed) that $\lambda_k \equiv 1$ and $x^0_k = 0$. Thus, it suffices to consider the case where $x^0$ is “near” the origin (i.e., $|x^0_k| < 2R_0$, $k = 1, \ldots, n$).

**Remark 1.** Some regularly varying functions arise in this section by transformation of $\lambda_k$. Also we consider some terms with finite limits at zero. It is easily seen that there exists $R_0 \leq 1$, depending only on $(\lambda_i)$, such that if $|x_k| < 3R_0$, then all such terms are “closely approximated” by their limit values (surely, this statement needs to be specified in each specific case).

We introduce a new coordinate system:

\[
y_k = \int_0^{x_k} \frac{d\tau}{\sqrt{\lambda_k(\tau)}}, \quad k = 1, \ldots, n.
\]

Obviously, the transformation (3.2) (as well as its inverse) satisfies the Hölder condition in $\bar{\Omega}$. Thus, the Hölder exponent and Hölder constant of an arbitrary function in the original coordinates depend only on its Hölder exponent (respectively, constant) in the new coordinates and on the collection $(\lambda_i)$.

In the $y$-coordinates, equation (1.1) is written as follows (we keep the original notation for the “transplanted” functions):

\[
\tilde{L} u = -\sum_{i,j=1}^n \tilde{a}_{ij}(y) u_{y_i y_j} + \sum_{i=1}^n \tilde{a}_{ii}(y) \beta_i(y) u_{y_i} = \tilde{f}(y),
\]

where

\[
\beta_k(y_k) = \frac{\lambda_k'(x_k)}{2\sqrt{\lambda_k(x_k)}}.
\]

Note that

\[
\sigma_k = \lim_{y_k \to 0} y_k \beta_k(y_k) = \lim_{x_k \to 0} \frac{x_k \lambda'(x_k)}{\lambda_k(x_k)} \cdot \lim_{x_k \to 0} \frac{\int_0^{x_k} \frac{d\tau}{\sqrt{\lambda_k(\tau)}}}{2x_k/\sqrt{\lambda_k(x_k)}} = \frac{\alpha_k}{2} - \alpha_k \in ]-1,1[.
\]

By Remark 1, we may assume that if $|x_k| < 3R_0$, then

\[
\sigma_k - \frac{1 - |\sigma_k|}{2} \leq y_k \beta_k(y_k) \leq \sigma_k + \frac{1 - |\sigma_k|}{2}.
\]

The image of the cube $K_{3R_0}$ under the transformation (3.2) is a rectangular parallelepiped to be denoted by $\widehat{\Omega}$.

Now we introduce a set of “one-dimensional” functions that will be the base for constructing barriers.

**Lemma 3.1.** Let $K_\rho(y^0) \subset \widehat{\Omega}$. Then there exist $B_k$, depending only on $y^0_k$ and $\lambda_k$ ($k = 1, \ldots, n$), such that the functions $w_k(y_k) = -y_k^2 + B_k x_k$ (here $x_k(y_k)$ stands for the function inverse to (3.2)) satisfy the following relations:

1) $0 < \nu(1 - (\sigma_k)_+) \leq \tilde{L} w_k \leq 4\nu^{-1}$ in $K_\rho(y^0)$;

2) $w_k$ attains its maximum at $y_k^0$, moreover,

\[
C_1 \rho^2 \leq w_k(y_k^0) - w_k(y_k^0 + \rho) \leq C_2 \rho^2;
\]

\[
C_1 \rho^2 \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \leq C_2 \rho^2,
\]

where $C_1$, $C_2$ are positive constants depending only on the collection $(\lambda_i)$. 


3) there exists \( \vartheta > 0 \), depending only on \((\lambda_i)\), such that if \(|1 - \frac{\sigma}{y_k}| \leq \vartheta\), then

\[
\hat{C}_1 \rho^2 \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \leq \hat{C}_2 \rho^2,
\]

where \( \hat{C}_1, \hat{C}_2 \) are constants depending only on \((\lambda_i)\), and \( \hat{C}_2 \leq \frac{5}{4} \hat{C}_1 \).

**Proof.** A direct calculation gives \( \hat{C}(y_k^0) = 2\tilde{a}_{kk}(1 - y_k\beta_k(y_k)) \). Since \( Lx_k = 0 \), we see that, by (3.4), property 1 is satisfied for any choice of \( B_k \).

Next, we note that for \( y_k^0 = 0 \) we can put \( B_k = 0 \), and property 2 becomes trivial. Otherwise, if \( y_k^0 \neq 0 \), then by symmetry we may assume that \( y_k^0 > 0 \). Since \( dx_k/dy_k = \sqrt{\lambda_k(x_k)} \), the condition \( w_k'(y_k^0) = 0 \) will be satisfied if we put \( B_k = \frac{2y_k^0}{\sqrt{\lambda_k(x_k(y_k^0))}} \). It should be noted that, by parts 3–5 of Proposition 2.1,

\[
\frac{y_k}{\sqrt{\lambda_k(x_k(y_k))}} \in \mathcal{R}_{1-\sigma_k} \text{ is a monotone increasing function,}
\]

i.e., \( B_k \) depends on \( y_k^0 \) monotonically; thus, for \( B_k \) fixed, the function \( w_k \) increases on \([-\infty, y_k^0]\) and decreases on \([y_k^0, +\infty)\].

We have

\[
w_k(y_k^0) - w_k(y_k^0 + \rho) = \rho^2 + 2y_k^0 \left( \frac{y_k^0}{\sqrt{\lambda_k(x_k(y_k^0))}} \cdot w_k'(y_k^0 + \theta \rho) \right),
\]

Expanding \( x_k(y_k^0 + \rho) \) by the Taylor formula, we get

\[
w_k(y_k^0) - w_k(y_k^0 + \rho) = \rho^2 \left( 1 - \frac{y_k^0}{2\sqrt{\lambda_k(x_k(y_k^0))}} \cdot \lambda_k'(x_k(y_k^0 + \theta \rho)) \right),
\]

where \( \theta \in [0, 1] \).

Formula (3.7) yields

\[
\left| \frac{y_k^0 \cdot \lambda_k'(x_k(y_k^0 + \theta \rho))}{2\sqrt{\lambda_k(x_k(y_k^0))}} \right| \leq \left| (y_k^0 + \theta \rho) \cdot \frac{\lambda_k'(x_k(y_k^0 + \theta \rho))}{2\sqrt{\lambda_k(x_k(y_k^0 + \theta \rho))}} \right| = |(y_k^0 + \theta \rho)\beta_k(y_k^0 + \theta \rho)|,
\]

whence \( \frac{1 - \sigma_k}{2} \rho^2 \leq w_k(y_k^0) - w_k(y_k^0 + \rho) \leq \frac{3 - \sigma_k}{2} \rho^2 \).

The second inequality in (3.5) is somewhat more complicated. As in (3.8), we have

\[
w_k(y_k^0) - w_k(y_k^0 - \rho) = \rho^2 \left( 1 - \frac{y_k^0}{2\sqrt{\lambda_k(x_k(y_k^0))}} \cdot \lambda_k'(x_k(y_k^0 - \theta \rho)) \right).
\]

First, suppose that \( \rho \leq \delta y_k^0 \), where \( \delta \in [0, 1] \) is a quantity to be chosen later. By part 1 of Proposition 2.1, Remark 1, and relation (3.7), we may assume that

\[
\frac{y_k^0 \sqrt{\lambda_k(x_k(y_k^0 - \theta \rho))}}{(y_k^0 - \theta \rho)\sqrt{\lambda_k(x_k(y_k^0))}} \leq \left( \frac{y_k^0}{y_k^0 - \theta \rho} \right)^{1 - \sigma_k} \frac{4 - |\sigma_k|}{3} \leq \frac{4 - |\sigma_k|}{3(1 - \delta)^{1 - \sigma_k}},
\]

and therefore

\[
\left| \frac{y_k^0}{2\sqrt{\lambda_k(x_k(y_k^0))}} \cdot \lambda_k'(x_k(y_k^0 - \theta \rho)) \right| \leq \frac{4 - |\sigma_k|}{3(1 - \delta)^{1 - \sigma_k}} \cdot |(y_k^0 - \theta \rho)\beta_k(y_k^0 - \theta \rho)|.
\]

Choosing \( \delta = 1 - (1 - |\sigma_k|)^2/9 \), we arrive at

\[
\frac{1 - |\sigma_k|}{2} \frac{\rho^2}{|\sigma_k|} \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \leq \frac{7 + 5|\sigma_k|}{2} \frac{\rho^2}{|\sigma_k|}.
\]

Now, let \( \delta y_k^0 \leq \rho \leq 2y_k^0 \). Then the above estimate yields

\[
w_k(y_k^0) - w_k(y_k^0 - \rho) \geq w_k(y_k^0) - w_k(y_k^0 - \delta y_k^0) \geq \frac{1 - |\sigma_k|}{2} \frac{(\delta y_k^0)^2}{|\sigma_k|} \geq \frac{1 - |\sigma_k|}{2} \frac{\delta^2}{8} \rho^2.
\]
On the other hand, \( \frac{\sigma_k}{\sqrt{n_k(x_k)}} = x_k \cdot dy_k/dx_k \), and \( y_k(x_k) \in \mathcal{R}_{2-\alpha_k} \). By part 6 of Proposition 2.1 and Remark 1, we may assume that \( B_kx_k(y_k^0) \leq 2(2 - \alpha_k)(y_k^0)^2 \), whence

\[
w_k(y_k^0) - w_k(y_k^0 - \rho) \leq w_k(y_k^0) - w_k(-y_k^0) = 2B_kx_k(y_k^0) \leq 4(2 - \alpha_k)(y_k^0)^2 \leq 4\frac{2 - \alpha_k}{\delta^2}\rho^2.
\]

Finally, let \( \rho \geq 2y_k^0 \). Then \(-\rho \leq y_k^0 - \rho \leq -\frac{\rho}{2}\), so that

\[
\frac{\rho^2}{4} \leq -w_k(-\frac{\rho}{2}) \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \\
\leq w_k(y_k^0 - \rho) \leq 2B_kx_k(\rho) + \rho^2 \leq (9 - 4\alpha_k)\rho^2.
\]

This completes the proof of part 2.

Under the conditions of part 3, we have \( y_k^0 > 0 \) and \( |y_k^0 - \rho| \leq \vartheta y_k^0 \). Therefore,

\[
B_k(x_k(y_k^0) - x_k(\vartheta y_k^0)) = (1 - \vartheta^2)(y_k^0)^2 \\
= w_k(y_k^0) - w_k(\vartheta y_k^0) \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \leq w_k(y_k^0) - w_k(-\vartheta y_k^0) \\
= B_k(x_k(y_k^0) + x_k(\vartheta y_k^0)) - (1 - \vartheta^2)(y_k^0)^2.
\]

Since \( x_k(y_k) \in \mathcal{R}_{1+\sigma_k} \), Remark 1 allows us to assume that \( x_k(\vartheta y_k^0) \leq 2\vartheta^{1+\alpha_k}x_k(y_k^0) \).

Next, by part 6 of Proposition 2.1 and Remark 1, we may assume that

\[
\frac{(1 - \vartheta)(2 - \alpha_k)}{(1 + \vartheta)^2}\rho^2 \leq (1 - \vartheta)(2 - \alpha_k)(y_k^0)^2 \leq B_kx_k(y_k^0) \\
\leq (1 + \vartheta)(2 - \alpha_k)(y_k^0)^2 \leq \frac{(1 + \vartheta)(2 - \alpha_k)}{(1 - \vartheta)^2}\rho^2.
\]

Consequently,

\[
(1 - \alpha_k - c(\vartheta, \sigma_k))\rho^2 \leq w_k(y_k^0) - w_k(y_k^0 - \rho) \leq (1 - \alpha_k + c(\vartheta, \sigma_k))\rho^2,
\]

where \( c \to 0 \) as \( \vartheta \to 0 \).

It remains to choose \( \vartheta \) so that \( c(\vartheta, \sigma_k) \leq \frac{1 - \alpha_k}{\vartheta} \) for all \( k = 1, \ldots, n \).

\section*{4. Elliptic case}

In all lemmas of this section we assume that \( v \in W^2_\text{loc}(K_{3R_0}) \) is a function that, after passage to the \( y \)-coordinates, is nonnegative in the cube \( K_\rho(y^0) \subset \tilde{H} \) and satisfies the inequality \( \tilde{L}v \geq -f(y) \) a.e. We introduce the notation \( \lambda = \text{det}(\Lambda) = \prod_{k=1}^n \lambda_k \).

We follow the classical pattern of [3]. The first lemma shows that if in some cube the level set \( A^v_s \) has sufficiently large measure, then some smaller cube lies in the level set \( A^v_{s/2} \) (up to correction for the right-hand side of the equation). A folklore title “the thin set lemma” is explained by the fact that the set under the level \( s \) is “thin” (i.e., has small measure).

**Lemma 4.1.** There exist constants \( \zeta \in ]0, 1[ \), \( \eta \in ]0, 1[ \), \( \gamma > 0 \), and \( C_3 > 0 \), depending only on \( n, \nu \), and the collection \( (\lambda_k) \), such that if the inequality \( |K_\rho(y^0) \cap A^v_s| \geq (1 - \zeta)|K_\rho(y^0)| \) is fulfilled for some \( s > 0 \), then

\[
v \geq \frac{s}{2} - C_3\rho^2 \| f \|_{n,K_\rho(y^0)} \quad \text{in} \quad K_{\eta\rho}(y^0).
\]

**Proof.** We construct the barrier function

\[
W(y) = \frac{1}{C_1\rho^2} \sum_{k=1}^n w_k(y_k) + \tilde{C},
\]
where the $u_k$ are the functions on the cube $K_{\rho}(y^0)$ that were defined in Lemma 3.1, $C_1$ is the constant occurring in (3.5), and $\tilde{C}$ is determined by the condition $W(y^0) = 1$. From (3.5) it follows that $W|_{\partial K_{\rho}(y^0)} \leq 0$.

We pass to $x$-coordinates. Denote by $x^0$ the image of $y^0$, and by $\Pi_{\rho}(x^0)$ the image of $K_{\rho}(y^0)$ (this is a cuboid). We keep the original notation for “transplanted” functions.

Obviously, the function $sW(x) - v(x)$ is nonpositive on $\partial \Pi_{\rho}(x^0)$. Applying the Aleksandrov maximum principle (see [10]), we obtain
\[
\left( \max_{\Pi_{\rho}(x^0)} (sW - v) \right)^n \leq \frac{N_1(n) |\Pi_{\rho}(x^0)| \cdot \int_{A_{0}^{W-v} \cap \Pi_{\rho}(x^0)} (\mathcal{L}(sW(x) - v(x)))^n}{\det(A)} dx 
\leq \frac{N_1(n) |\Pi_{\rho}(x^0)| \cdot \int_{A_{0}^{W-v} \cap K_{\rho}(y^0)} (\tilde{\mathcal{L}}(sW(y) - v(y)))^n}{\lambda(x(y))} dy.
\]

By part 1 of Lemma 3.1, $\tilde{\mathcal{L}}(sW - v) \leq \frac{4m\nu^{-1}}{C_1 \rho^2} s + f_+(x)$. Therefore,
\[
(sW - v)^n \leq \frac{N_2(n)}{\nu} |\Pi_{\rho}(x^0)| \cdot \left( s \frac{|A_{0}^{W-v} \cap K_{\rho}(y^0)|}{\nu C_1 \rho^2} + \left\| f_+ \right\|_{n,K_{\rho}(y^0)} \right),
\]
where, as in Lemma 2.3,
\[
|E|_* = \int_E dy_1 \ldots dy_n.
\]

Since $W \leq 1$, it is easily seen that $(A_{0}^{W-v} \cap K_{\rho}(y^0)) \subset K_{\rho}(y^0) \setminus A_s$. Denoting $\varepsilon = \frac{|\Pi_{\rho}(x^0)\cap A_s^\nu|}{|K_{\rho}(y^0)|}$, we obtain
\[
(4.3) \quad v(y) \geq sW(y) - \frac{N_2(n)}{\nu} |\Pi_{\rho}(x^0)| \cdot \left( s \frac{|A_{0}^{W-v} \cap K_{\rho}(y^0)|}{\nu C_1 \rho^2} + \left\| f_+ \right\|_{n,K_{\rho}(y^0)} \right).
\]

Note that
\[
\frac{|\Pi_{\rho}(x^0)| \cdot |K_{\rho}(y^0)|}{\rho^{2n}} = \prod_{k=1}^{n} \frac{\int_{y_k^{0} - \rho}^{y_k^{0} + \rho} \sqrt{\lambda_k(x_k(y_k))} dy_k \cdot \int_{y_k^{0} - \rho}^{y_k^{0} + \rho} dy_k}{\rho^2}.
\]

By Lemma 2.2, the last expression is bounded from above by a constant depending only on the collection $(\lambda_i)$. Therefore, inequalities (4.3) and (3.5) show that, for $y \in K_{\rho}(y^0)$,
\[
(4.4) \quad v(y) \geq s \left( 1 - \frac{n C_2}{C_1} \eta^2 - N_3 \varepsilon \right) - N_4 |\Pi_{\rho}(x^0)| \cdot \left\| f_+ \right\|_{n,K_{\rho}(y^0)} (\lambda_{K_{\rho}(y^0)})\]

(here $C_2$ is the constant from (3.5), while $N_3$ and $N_4$ depend only on $n$, $\nu$, and $(\lambda_i)$).

Since $\sqrt{\lambda_k(x_k(y_k))} \in \mathcal{R}_{\sigma_k}$, formula (2.10) yields
\[
(4.5) \quad |\Pi_{\rho}(x^0)| \simeq \rho^n \prod_{k=1}^{m} \sqrt{\lambda_k(x_k(y_k))} \cdot \prod_{k=m+1}^{n} \sqrt{\lambda_k(x_k(y_k))} \leq N_5 \rho^{m\gamma},
\]
where $\gamma = \frac{1}{\nu} - \frac{1}{\nu} \sum_{k=1}^{n} (\sigma_k) - > 0$, and $N_5$ depends only on the collection $(\lambda_i)$.

We substitute (4.5) in (4.4) and put $\eta = \sqrt{C_1/(4n C_2)}$, $C_3 = N_3 N_5^{1/\nu}$. It remains to note that, by Lemma 2.3, the estimate $|K_{\rho}(y^0) \setminus A_s^\nu| \leq \zeta \cdot |K_{\rho}(y^0)|$ for sufficiently small $\zeta$ ensures that $\varepsilon \leq (4N_3)^{-n}$, and we arrive at (4.1). \hfill $\square$

The next lemma shows that a positive lower bound for $v$ on a face of a sufficiently large subcube situated close to a coordinate hyperplane can be transferred across this hyperplane by a thin “bridge”.
Lemma 4.2. There exist constants \( \hat{\theta} \in [0, \frac{1}{2}] \) and \( C_4 > 0 \), depending only on \( n, \nu \), and \( (\lambda_i) \), such that if \( 0 \leq \psi_m \leq (1 - \hat{\theta}) \rho \) for some \( m \) (for definiteness, let \( m = 1 \)), and if for some \( s > 0 \) the inequality \( v \geq s \) is fulfilled on the set \( \{ \hat{\theta}^2 \rho \} \times K'(1-\hat{\theta})\rho(y^0) \), then

\[
(4.6) \quad v \geq \frac{s}{2} - C_4\rho^\gamma \left\| \frac{f_+}{(1+\hat{\theta}^2)\rho^\gamma} \right\|_{n, K'(y^0)} \quad \text{in the cuboid} \quad K'(1-\hat{\theta})\rho(y^0),
\]

where \( \gamma \) is the same constant as in Lemma 4.1.

Proof. Without loss of generality, we may assume that \( \hat{\theta} \leq \theta_1 \), where \( \theta_1 \) is the constant occurring in part 3 of Lemma 3.1. We denote \( y^0 = (-\hat{\theta}\rho, y^0) \) and define the barrier function

\[
(4.7) \quad \tilde{W}(y) = \frac{1}{C_1((1-\hat{\theta})\rho)^2} \sum_{k=2}^n w_k(y_k) - \frac{w_1(y_1)}{C_1((1-\hat{\theta})\rho)^2} + \tilde{C},
\]

where the \( w_k \) are functions on the cube \( K_\rho(y^0) \) that were defined in Lemma 3.1, \( C_1 \) and \( \tilde{C} \) are constants from (3.5) and (3.6), respectively, and \( \tilde{C} \) is determined by the condition \( \tilde{W}(\hat{\theta}^2 \rho, y^0) = 1 \).

Consider the following \( \tilde{\theta}(\tilde{\theta} + 1)\rho \)-thick rectangular parallelepiped adjacent to the hyperplane \( y_1 = \tilde{\theta}^2 \rho \):

\[
\tilde{\Pi}_\rho = \{ -\tilde{\theta} \rho, \tilde{\theta}^2 \rho \times K'(1-\tilde{\theta})\rho(y^0) \cap (K_\rho(y^0) \cap K_\rho(y^0)).
\]

Relations (3.5) and (3.6) imply that \( \tilde{W} \leq 1 \) on \( \partial \tilde{\Pi}_\rho \), except for the face \( y_1 = \tilde{\theta}^2 \rho \). Next, by part 1 of Lemma 3.1,

\[
\tilde{\mathcal{L}} \tilde{W} \leq \frac{1}{\rho^2} \cdot \frac{(4(n-1)\nu^{-1})}{C_1(1-\hat{\theta})^2} - \frac{\nu(1-(\sigma_1)_+)}{C_1\hat{\theta}^2(\hat{\theta} + 1)^2}.
\]

Let \( \vartheta_1 \) denote the positive root of the expression in brackets. Then for \( \hat{\theta} \leq \vartheta_1 \) we have \( \tilde{\mathcal{L}} \tilde{W} \leq 0 \).

We pass to \( x \)-coordinates and denote by \( \hat{\Pi}_\rho \) the image of \( \tilde{\Pi}_\rho \) (it is a cuboid). We keep the original notation for “transplanted” functions.

Obviously, the function \( s\tilde{W}(x) - v(x) \) is nonpositive on \( \partial \hat{\Pi}_\rho \). Applying the Aleksandrov maximum principle, we obtain

\[
(\max_{\hat{\Pi}_\rho} (s\tilde{W} - v)_+) \leq N_1(n)|\hat{\Pi}_\rho| \int_{\hat{\Pi}_\rho} \frac{(\mathcal{L}(s\tilde{W}(x) - v(x)))^n}{\det(A)}\,dx
\]

\[
\leq \frac{N_1(n)}{\nu^n} |\hat{\Pi}_\rho| \cdot \left\| \frac{f_+}{\sqrt{\lambda}} \right\|_{n, \hat{\Pi}_\rho}^n.
\]

As in (4.5), we have \( |\hat{\Pi}_\rho| \leq N_6 \rho^{n\gamma} \), where \( N_6 \) depends only on the collection \( (\lambda_i) \).

Recalling (3.5) and (3.6), for \( y \in \{ -\hat{\theta}^2 \rho, \hat{\theta}^2 \rho \times K'(y^0) \} \) we obtain

\[
v(y) \geq s \left[ \tilde{W}(y^0) - \frac{(n-1)C_2\hat{\theta}^2}{C_1(1-\hat{\theta})^2} + \frac{(1-\hat{\theta})^2}{(1+\hat{\theta})^2} \right] - C_4\rho^\gamma \left\| \frac{f_+}{\sqrt{\lambda}} \right\|_{n, K'(y^0)}
\]

(here \( C_2 \) is the constant from (3.5)).

By (3.6), \( \tilde{W}(y^0) \geq 1 - \frac{C_2}{C_1} \geq -\frac{1}{2} \). Therefore, there exists \( \hat{\theta} \leq \min\{\vartheta, \vartheta_1\} \) such that the expression in square brackets is at least \( \frac{1}{2} \). This gives (4.6). \( \Box \)
Theorem 4.4. Suppose a function \( u \in W^{2,1}_{n,\text{loc}}(\Omega) \) satisfies equation (1.1), and \( |u| \leq M_0 \) in \( \Omega \). Suppose that the coefficients in (1.1) satisfy conditions (1.2)–(1.4) and (3.1), and that \( \frac{1}{\det(\Lambda)} \in L_{\infty}(\Omega) \). Then for any subdomain \( \Omega' \subseteq \Omega \) we have

\[
\|u\|_{C^{1,\gamma}(\Omega')} \leq M_{\gamma},
\]

where \( \gamma \) is completely determined by \( n, \nu \), and the collection \( \{\lambda_i\} \), while \( M_{\gamma} \) depends also on \( M_0, \frac{1}{\det(\Lambda)} \|_{n,\Omega}, \text{ and dist}(\Omega',\partial\Omega) \).

Lemma 4.3. There exist constants \( \varkappa \in [0,1] \) and \( C_5 > 0 \), depending only on \( n, \nu \), and \( (\lambda_i) \), such that if the inequality \( |K_3(y^\theta) \cap A_1^\nu| \geq (1 - \varkappa) |K_3(y^\theta)| \) is fulfilled for some \( s > 0 \), then

\[
v \geq \varkappa s - C_5 \rho \|f\|_{n,K_3(y^\theta)} \quad \text{in} \quad K_3(y^\theta)
\]

(4.8)

(\text{here} \( \varkappa \in [0,1] \) \text{ and } \gamma > 0 \text{ are constants from Lemma 4.1}).

Proof. Without loss of generality, we may assume that \( y^0 \in \mathbb{R}^n_+ \). First, let \( f = 0 \). Applying Lemma 4.1 to the cube \( K_{\frac{3}{2}}(y^\theta) \), we see that \( v \geq \frac{s}{2} \) on \( K_{\frac{3}{2}}(y^\theta) \).

Put \( \hat{\eta} = \min\{\eta, \hat{\eta}^3\} \) (here \( \hat{\eta} \) is the constant from Lemma 4.2). Note that the coefficients \( \beta_i \) are bounded on the set \( K_3(y^\theta) \cap \{y_k > \frac{3}{2} \rho, \ k = 1, \ldots, n\} \) and that, by (3.4), their \( L_n \)-norms over this set are bounded from above by some constant independent of \( \rho \) and \( y^\theta \). The “lemma on spreading of ink-stains” (see [3, Lemma 2.2]) allows us to obtain the estimate \( v \geq \varkappa s \) on the set \( K_{(1 - \frac{3}{2} \rho)}(y^\theta) \cap \{y_k \geq \frac{3}{2} \rho, \ k = 1, \ldots, n\} \) (here \( \varkappa \) depends only on \( n, \nu \), and \( (\lambda_i) \)).

Assume that the cube \( K_{(1 - \frac{3}{2} \rho)}(y^\theta) \) intersects some coordinate hyperplane (let it be the hyperplane \( y_1 = 0 \)). We consider all cubes \( K_{\frac{3}{2} \rho} \) that lie completely in the half-spaces \( y_k > 0, k = 2, \ldots, n \), and lean out of the half-space \( y_1 > 0 \) at a depth of \( \frac{3}{2} \rho \). Applying Lemma 4.2 to these cubes, we obtain the estimate \( v \geq \frac{s}{2} \) in the set \( K_{(1 - \frac{3}{2} \rho)}(y^\theta) \cap \{y_1 \leq \frac{3}{2} \rho; y_k > \frac{3}{2} \rho, \ k = 2, \ldots, n\} \). Then we again refer to the “lemma on spreading of ink-stains” (in the second orthant), obtaining the estimate \( v \geq \varkappa s \) in the set \( K_{(1 - \frac{3}{2} \rho)}(y^\theta) \cap \{y_k \geq \frac{3}{2} \rho, \ k = 2, \ldots, n\} \) except for a “tunnel” of depth \( \frac{3}{2} \rho \) in the neighborhood of the hyperplane \( y_1 = 0 \) (here \( \varkappa \) depends only on \( n, \nu \), and \( (\lambda_i) \)).

Proceeding as above, after at most \( n \) steps we obtain the estimate \( v \geq \varkappa s \) in the cube \( K_{(1 - \frac{3}{2} \rho)}(y^\theta) \) and for a cross-shaped “tunnel” of depth \( \frac{3}{2} \rho \) in the neighborhood of the hyperplanes \( y_k = 0, k = 1, \ldots, n \). Finally, we apply Lemma 4.2 to cubes \( K_{\frac{3}{2}} \) that lean out of the half-space \( y_k > 0 \) at a depth of \( \frac{3}{2} \rho \). This gives the estimate \( v \geq \varkappa s \) in the cube \( K_{\frac{3}{2}}(y^\theta) \).

To treat the general case, it suffices to note that at every step we reduce the estimate by a quantity of the form \( c \rho \|f\|_{n,K_3(y^\theta)} \|\hat{\epsilon}\|_{n,K_3(y^\theta)} \), where a constant \( c \) depends only on \( n, \nu \), and \( (\lambda_i) \). This gives (4.8). \( \square \)

Lemma 4.3 is a direct analog of [3, Lemma 2.3], and now the Hölder estimate for solutions of (3.3) can be deduced in a standard way; see [3, §2]. Returning to \( x \)-coordinates, we obtain an estimate for solutions of equation (1.1). We only formulate the final result.

Theorem 4.4. Suppose a function \( u \in W^{2,1}_{n,\text{loc}}(\Omega) \) satisfies equation (1.1), and \( |u| \leq M_0 \) in \( \Omega \). Suppose that the coefficients in (1.1) satisfy conditions (1.2)–(1.4) and (3.1), and that \( \frac{1}{\det(\Lambda)} \in L_{\infty}(\Omega) \). Then for any subdomain \( \Omega' \subseteq \Omega \) we have

\[
\|u\|_{C^{1,\gamma}(\Omega')} \leq M_{\gamma},
\]

where \( \gamma \) is completely determined by \( n, \nu \), and the collection \( (\lambda_i) \), while \( M_{\gamma} \) depends also on \( M_0, \frac{1}{\det(\Lambda)} \|_{n,\Omega}, \text{ and dist}(\Omega',\partial\Omega) \).
§5. Parabolic Case

In this section we consider equation (1.5) under the assumptions (1.2)–(1.4) and (3.1). In essence, the general line to establish the Hölder estimates is the same as in the preceding section. Therefore, here we only formulate the corresponding lemmas and give some hints to the proofs. We denote $Q_\rho(x,t) = K_\rho(x) \times |t - \rho^2|$, $Q'_\rho(x,t) = K'_\rho(x) \times |t - \rho^2|$. The proof of the next lemma repeats that of Lemma 2.3 word for word.

**Lemma 5.1.** Suppose $\varphi_k \in \mathcal{R}_{\alpha_k}$, $\alpha_k > -1$, $k = 1, \ldots, n$. For measurable sets $E \subset \mathbb{R}^{n+1}$, define the measure

$$ |E|_* = \int_E \varphi_1(|x_1|) \ldots \varphi_n(|x_n|) \, dx_1 \ldots dx_n \, dt. $$

Then for any $\varepsilon > 0$ there exists $\delta > 0$, depending only on $\varepsilon$ and the collection $(\varphi_k)$, such that for all cylinders $Q_\rho(x^0, t^0) \subset K_1 \times |0,T|$ and an arbitrary measurable set $E$ we have

$$ |E \cap Q_\rho(x^0, t^0)| \leq \delta \cdot |Q_\rho(x^0, t^0)| \quad \Rightarrow \quad |E \cap Q_\rho(x^0, t^0)|_* \leq \varepsilon \cdot |Q_\rho(x^0, t^0)|_*.$$

After the coordinate transformation (3.2), equation (1.5) becomes

$$ (5.1) \quad \widehat{M} u = \partial_t u - \sum_{i,j=1}^n \tilde{a}_{ij}(y,t) u_{y_i y_j} + \sum_{i=1}^n \tilde{a}_{i1}(y,t) u_{y_i} = f(y,t) $$

(we keep the original notation for “transplanted” functions, as before).

In Lemmas 5.2–5.4 we suppose that the function $v \in W^{2,1}_{n+1}(K_3R_0 \times |0,T|)$ after passage to the $y$-coordinates is nonnegative in the cylinder $Q_\rho(y^0, t^0) \subset (\Pi \times |0,T|)$ and satisfies the inequality $\widehat{M} v \geq -f(y,t)$ a.e.

**Lemma 5.2.** There exist constants $\tilde{\zeta} \in ]0,1[\setminus \mathbb{Q}$, $\tilde{\eta} \in ]0,1[\setminus \mathbb{Q}$, $\tilde{\gamma} > 0$, and $C_6 > 0$, depending only on $n$, $\nu$, and $(\lambda_i)$, such that if the inequality $|Q_\rho(y^0, t^0) \cap A_s^*| \geq (1 - \tilde{\zeta})|Q_\rho(y^0, t^0)|$ is fulfilled for some $s > 0$, then

$$ v \geq \frac{s}{2} - C_6 \rho \tilde{\zeta} \left\| \frac{f_+}{\sqrt{\lambda}} \right\|_{n+1, Q_\rho(y^0, t^0)} \quad \text{in} \quad Q_{\tilde{\eta} \rho}(y^0, t^0). $$

**Proof.** The proof follows the lines of Lemma 4.1. Instead of the barrier function $W$ (see (4.2)), we introduce the function $W(y,t) = W(y) + \frac{t^0 - t}{\rho^2}$, and instead of the Aleksandrov maximum principle we use its parabolic version (see [17] and also [18, 19]). This gives an analog of inequality (4.3):

$$ v(y,t) \geq s W(y,t) - N_7 |\Pi_\rho(x^0)| \frac{1}{\sqrt{\rho^2}} \left( \frac{s |Q_\rho(y^0, t^0)|_*}{\rho^2} + \left\| \frac{f_+}{\sqrt{\lambda}} \right\|_{n+1, Q_\rho(y^0, t^0)} \right), $$

where $N_7$ depends only on $n$, $\nu$, and $(\lambda_i)$, and $\varepsilon = \frac{|Q_\rho(y^0, t^0)|_*}{\rho^2}$. The rest of the proof runs practically without changes. \hfill $\Box$

**Lemma 5.3.** There exist constants $\tilde{\vartheta} \in ]0, \frac{1}{4}[\setminus \mathbb{Q}$ and $C_7 > 0$, depending only on $n$, $\nu$, and $(\lambda_i)$, such that if $0 \leq y_m^0 \leq (1 - \tilde{\vartheta}) \rho$ for some $m$ (for definiteness, let $m = 1$), and if for some $s > 0$ the inequality $v \geq s$ is fulfilled on the set $\{\tilde{\vartheta} \rho \times Q'_\rho(y^0, t^0)\}$, then

$$ v \geq \frac{s}{2} - C_7 \rho \tilde{\vartheta} \left\| \frac{f_+}{\sqrt{\lambda}} \right\|_{n+1, Q_\rho(y^0, t^0)} \quad \text{in the cylinder} \quad |\tilde{\vartheta} \rho \times Q'_\rho(y^0, t^0)|,$$

where $\tilde{\vartheta}$ is the constant from Lemma 5.2.
Proof. The proof follows the lines of that of Lemma 4.2, but instead of the barrier function \( \tilde{W} \) (see (4.7)) we introduce the function \( \tilde{W}(y, t) = \tilde{W}(y) + \frac{t}{\sqrt{\lambda}} \), and use the parabolic version of the Aleksandrov maximum principle.

**Lemma 5.4.** There exist constants \( \tilde{c} \in [0, 1] \) and \( C_3 > 0 \), depending only on \( n, \nu \), and \( (\lambda_i) \), such that if the inequality \( |Q_{\tilde{z}}(y^0, t^0 - 3\tilde{c}^2) \cap A^c_{\tilde{z}}| \geq (1 - \tilde{c})|Q_{\tilde{z}}(y^0, t^0)| \) is fulfilled for some \( s > 0 \), then

\[
\forall \geq \tilde{c}s - C_3\rho \tilde{\gamma} \| f_{\lambda} \|_{\nu, L^1(Q_{\tilde{z}})} \quad \text{in} \quad Q_{\tilde{z}}(y^0, t^0)
\]

(here \( \tilde{c} \in [0, 1] \) and \( \tilde{\gamma} > 0 \) are constants from Lemma 5.2).

Proof. The proof follows the lines of that of Lemma 4.3. Instead of \[3\] Lemma 3.2, taking account of the fact that the coefficients \( \beta_i \) are bounded on the set \( Q_{\rho}(y^0, t^0) \cap \{ y_k > \frac{3}{n+1} \rho, \quad k = 1, \ldots, n \} \), and, by (3.4), their \( L_{n+2} \)-norms over this set are bounded from above by some constant independent of \( \rho \) and \( y^0 \).

Lemma 5.4 is a direct analog of \[3\] Lemma 3.3, and now the Hölder estimates for solutions of equation (5.1) are deduced in a standard way; see \[3\] §3. Returning to \( x \)-coordinates, we obtain an estimate for solutions of equation (1.5). We only formulate the final result.

**Theorem 5.5.** Suppose a function \( u \in W^{2,1}_{n+1, \text{loc}}(Q) \) satisfies equation (1.5), and \( |u| \leq M_0 \) in \( Q \). Suppose that the coefficients in (1.5) satisfy conditions (1.2)–(1.4) and (3.1), and that \( \frac{f}{\det(x)} \in L_{n+1}(Q) \). Then for any subcylinder \( Q' = \Omega' \times [\delta, T] \), \( \Omega' \subseteq \Omega \), we have

\[
\| u \|_{C^\gamma(Q')} \leq M_{\tilde{\gamma}},
\]

where \( \tilde{\gamma} \) is completely determined by \( n, \nu \), and the collection \( (\lambda_i) \), while \( M_{\tilde{\gamma}} \), depends also on \( M_0, \| \frac{f}{\det(x)} \|_{n+1, \text{loc}}, \delta, \) and \( \text{dist}(\Omega', \partial \Omega) \).

**Remark 2.** In (5.2), as usual, the Hölder norm is with respect to the parabolic metric

\[
d_{\text{par}}((x^1, t^1), (x^2, t^2)) = |x^1 - x^2| + |t^1 - t^2|^{1/2}.
\]

\section{6. About approximative solutions}

If the coefficients in (1.1) are only measurable, in the general situation one cannot ensure the existence of a solution \( u \in W^{2,1}_{n, \text{loc}}(\Omega) \) even in the uniformly elliptic case \( (\lambda_k \equiv 1, \quad k = 1, \ldots, n) \). Therefore, approximative solutions of equation (1.1) are often considered. Namely, for \( \varepsilon \in [0, \varepsilon_0] \) we set

\[
\lambda_i^{(\varepsilon)}(t) = \begin{cases} 
\lambda_i(t) & \text{if } |t| \geq \varepsilon; \\
\lambda_i(\varepsilon) & \text{if } |t| \leq \varepsilon.
\end{cases}
\]

Next, we approximate the matrix \( \tilde{A} \) in the sense of a.e. convergence with smooth matrix functions \( \tilde{A}^{(\varepsilon)} \) that satisfy the uniform ellipticity condition with \( \nu \) independent of \( \varepsilon \). Finally, we introduce the matrices

\[
\Lambda^{(\varepsilon)}(x) = \text{diag}\{\lambda_1^{(\varepsilon)}(|x_1|), \lambda_2^{(\varepsilon)}(|x_2|), \ldots, \lambda_n^{(\varepsilon)}(|x_n|)\},
\]

\[
\tilde{A}^{(\varepsilon)} = \sqrt{\Lambda^{(\varepsilon)} \tilde{A}^{(\varepsilon)} \Lambda^{(\varepsilon)}},
\]
Lemma 6.1. Let the cuboid that is the image of the coordinate transformation by formula (3.2) with \( \left( \begin{array}{c} \kappa \\ \nu \end{array} \right) \) in place of \( \lambda_k \) and denote by \( \Pi^{(\varepsilon)} \) the cuboid that is the image of \( K_{3R_0} \) under this transformation. Note that in the new coordinate system the operator \( L^{(\varepsilon)} \) reads as follows:

\[
L^{(\varepsilon)}u^{(\varepsilon)} = - \sum_{i,j=1}^{n} a^{(\varepsilon)}_{ij}(x)u^{(\varepsilon)}_{x_i x_j} = f^{(\varepsilon)}(x),
\]

where \( f^{(\varepsilon)} \) is a sequence of smooth functions such that \( \left\| \frac{f^{(\varepsilon)}-f}{\det(\Lambda)} \right\|_{L^{\infty}(\Omega)} \to 0 \) as \( \varepsilon \to 0 \).

The standard elliptic theory (see, e.g., [20, Chapter III]) shows that equations (6.1) have classical solutions \( u^{(\varepsilon)} \in C^\infty(\Omega) \) for all \( \varepsilon > 0 \). A function \( u \) is called an approximative solution of equation (1.1) if there exists a subsequence of equations (6.1) and their solutions \( u^{(\varepsilon_k)} \) convergent to \( u \) a.e. in \( \Omega \) as \( \varepsilon_k \to 0 \).

Remark 3. The result of [11] shows that the Dirichlet problem for a uniformly elliptic equation (1.1) has approximative solutions for any continuous boundary data. Moreover, solutions of the Dirichlet problems for equations (6.1) converge to the approximative solution (1.1) if there exists a subsequence of equations (6.1) and their solutions \( u^{(\varepsilon_k)} \) convergent to \( u \) a.e. in \( \Omega \) as \( \varepsilon_k \to 0 \).

In this section we establish the Hölder property for an arbitrary approximative solution of equation (1.1) under the assumptions (1.2)–(1.4), (3.1). To proceed, we introduce a coordinate transformation by formula (3.2) with \( \lambda_k^{(\varepsilon)} \) in place of \( \lambda_k \) and denote by \( \Pi^{(\varepsilon)} \) the cuboid that is the image of \( K_{3R_0} \) under this transformation. Note that in the new coordinate system the operator \( L^{(\varepsilon)} \) reads as follows:

\[
\tilde{L}^{(\varepsilon)}u = - \sum_{i,j=1}^{n} a^{(\varepsilon)}_{ij}(y)u_{y_i y_j} + \sum_{i=1}^{n} a^{(\varepsilon)}_{ii}(y)\beta^{(\varepsilon)}_i(y_k)u_{y_i},
\]

where \( \beta^{(\varepsilon)}_k(y_k) = \frac{\lambda_k^{(\varepsilon)}(x_k)}{2\lambda_k^{(\varepsilon)}(x_k)} \cdot \chi_{[\varepsilon, +\infty]}(|x_k|). \)

Lemma 6.1. Let \( \rho > 0 \). There exists \( \delta > 0 \), depending only on \( \rho, n, \nu, \) and the collection (\( \lambda_k \)), such that for \( 0 < \varepsilon \leq \delta \) the following is true.

Suppose that a function \( v \in C^2(\tilde{\Pi}^{(\varepsilon)}) \) is nonnegative in the cube \( K_{\rho}(y^0) \subset \tilde{\Pi}^{(\varepsilon)} \) and satisfies the inequality \( \tilde{L}^{(\varepsilon)}v \geq -f^{(\varepsilon)}(y) \). If the inequality \( |K_2(y^0)\cap A|^s \geq (1-\frac{s}{2})|K_2(y^0)| \) is fulfilled for some \( s > 0 \), then

\[
v \geq \frac{\zeta}{2} s - 2C_5\rho^\gamma \left\| \frac{f^{(\varepsilon)}}{\det(\Lambda)} \right\|_{L^{\infty}(\Omega, K_{\rho}(y^0))} \quad \text{in} \quad \tilde{K}_{2}(y^0)
\]

(here \( \zeta \) and \( \gamma \) are constants from Lemma 4.1, and \( \zeta, C_5 \) are constants from Lemma 4.3).

Proof. The proof repeats the lines of §4. It suffices to note that all the quantities in the proofs of Lemmas 4.1–4.3 tend to their limits uniformly as \( \varepsilon \to 0 \).

Lemma 6.1 implies that if a nonnegative function \( v \) in the cube \( K_{\rho}(y^0) \subset \tilde{\Pi} \) satisfies the inequality \( \tilde{L}v \geq -f(y) \) in the approximative sense, then the statement of Lemma 4.3 holds true for that function (possibly, after changing on a zero measure set), with the replacement \( \zeta \to \frac{\zeta}{2}, \gamma \to \frac{\gamma}{2}, C_5 \to 2C_5 \). Since all further arguments from [3, §2] do not require the smoothness of functions, Theorem 4.4 remains valid for approximative solutions.
Approximative solutions of equation (1.5) can be defined in a similar way. Theorem 5.5 remains valid for them.

I am thankful to Yu. A. Alkhutov who kindly provided me with the manuscript of the paper [12].

REFERENCES


Department of Mathematics and Mechanics, St. Petersburg State University, Universitetskii Prospekt 28, Petrodvorets, St. Petersburg 198504, Russia

E-mail address: al.il.nazarov@gmail.com

Received 8/SEP/2008

Translated by THE AUTHOR