Czechoslovak Mathematical Journal

Charles S. Kahane

A gradient estimate for solutions of the heat equation

Czechoslovak Mathematical Journal, Vol. 48 (1998), No. 4, 711–725

Persistent URL: http://dml.cz/dmlcz/127449

Czechoslovak Mathematical Journal, 48 (123) (1998), 711–725

A GRADIENT ESTIMATE FOR SOLUTIONS OF THE HEAT EQUATION

Charles S. Kahane, Nashville

(Received November 20, 1995)

1. Introduction

Consider a solution $u(x,t)$ of the following boundary-initial value problem for the heat equation:

$$\tag{1.1} \begin{cases} u_t(x,t) = u_{xx}(x,t) & \text{in } (a,b)\times(0,\infty), \\ u(a,t) = u(b,t) = 0 & \text{for } t > 0, \\ u(x,0) = f(x) & \text{for } a < x < b, \end{cases}$$

with $f(x)$ assumed, to begin with, to be in $C[a,b]$. Suppose further that $f(x) \in C^1[a,b]$ and vanishes at the endpoints, then by noting that $v = u_x$ is a solution of the problem

$$\tag{1.2} \begin{cases} v_t(x,t) = v_{xx}(x,t) & \text{in } (a,b)\times(0,\infty), \\ v_x(a,t) = v_x(b,t) = 0 & \text{for } t > 0, \\ v(x,0) = f'(x) & \text{for } a < x < b, \end{cases}$$

we obtain the estimate

$$\tag{1.3} |u_x(x,t)| \leqslant \max_{[a,b]} |f'(x)| \text{ for } (x,t) \in (a,b)\times(0,\infty)$$

as a consequence of the maximum principle.

The goal of this paper is to derive estimates of the same type as (1.3) for gradients of solutions u of the higher dimensional version of (1.1); that is, for solutions $u(x,t)$ with $(x,t) \in \Omega\times(0,\infty)$, Ω being a bounded domain in $\mathbb{R}^n$, $n \geqslant 2$, of the problem

$$\tag{1.4} \begin{cases} u_t = \Delta u & \text{in } \Omega\times(0,\infty), \\ u = 0 & \text{on } \partial\Omega\times(0,\infty), \\ u(x,0) = f(x) & \text{for } x \in \Omega. \end{cases}$$

Assuming that $f(x) \in C^1(\overline{\Omega})$ and vanishes on $\partial\Omega$, we shall be able to obtain estimates of the form (1.3), provided that $\partial\Omega$ is sufficiently smooth and most importantly satisfies the following mean curvature condition:

Let p be a typical point on $\partial\Omega$ and suppose that after suitable rotation and translation of our coordinate system placing p at the origin of the system, the portion of $\partial\Omega$ lying in a neighborhood of p is the surface described by the function

$$x_n = g(x_1, \dots, x_{n-1}) \tag{1.5}$$

where $(x_1, \dots, x_{n-1})$ varies over a neighborhood of $(x_1 = 0, \dots, x_{n-1} = 0)$ with $g(0, \dots, 0) = 0$ and with the positive x_n direction corresponding to the outward normal direction from $\partial\Omega$ at p. Then the mean curvature condition that we shall assume $\partial\Omega$ to satisfy is that

$$\sum_{j=1}^{n-1} \frac{\partial^2 g}{\partial x_j^2}\Big|_{x_j=0, j=1,\dots,n-1} \leqslant 0. \tag{1.6}$$

The precise statement of our result is

Theorem 1.1. *Assume $u(x,t)$ to be a solution of* (1.4) *with $f(x) \in C^1(\overline{\Omega})$ and vanishing on $\partial\Omega$. Suppose further that $\partial\Omega$ is C^3 and satisfies the mean curvature condition* (1.6). *Then for $(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n})$, the spatial gradient of u, hereinafter denoted by* $\operatorname{grad} u(x,t)$, *we have the estimate*

$$|\operatorname{grad} u(x,t)| \leqslant \max_{\overline{\Omega}} |\operatorname{grad} f(x)|, \ (x,t) \in \Omega \times (0, \infty). \tag{1.7}$$

The estimate depends crucially on $\partial\Omega$ satisfying the curvature condition (1.6). And we shall give an example showing that it fails when this condition does not hold.

We should point out that without this condition, but assuming only that $\partial\Omega$ is sufficiently smooth, Ladyzenskaja *et al.* in [2] have derived bounds for $\operatorname{grad} u(x,t)$ depending on $\max_{\overline{\Omega}} |\operatorname{grad} f(x))|$, $\max_{\overline{\Omega}} |f(x))|$, as well as $\partial\Omega$ (see [2] Theorem 4.1, p. 443 and Lemma 6.1, p. 589).

The plan of the paper is as follows: The proof of Theorem 1.1 will be described in Sections 2, 3, and 4. In Section 5 we will explain the construction of an example showing that (1.7) may fail when (1.6) does not hold.

In this section we will begin the proof of Theorem 1.1. We shall do this by endeavoring to apply a suitable version of the maximum principle to $|\operatorname{grad} u(x,t)|^2$ from which an estimate of the form (1.7) will then follow.

To this end we will first show that

$$|\operatorname{grad} u(x,t)|^2 = \sum_{j=1}^{n} \Big(\frac{\partial u}{\partial x_j}\Big)^2$$

is sub-parabolic in $\Omega \times (0,\infty)$ meaning that

$$\Big[\Delta - \frac{\partial}{\partial t}\Big] |\operatorname{grad} u(x,t)|^2 \geqslant 0 \text{ in } \Omega \times (0,\infty). \tag{2.1}$$

Since sums of sub-parabolic functions are sub-parabolic, to establish (2.1) it suffices to prove that $(\frac{\partial u}{\partial x_j})^2$ is sub-parabolic for $j = 1,\dots,n$. But as each derivative $\frac{\partial u}{\partial x_j}$ is a solution of the heat equation, the sub-parabolicity of $(\frac{\partial u}{\partial x_j})^2$ is an immediate consequence of the following general result.

Proposition 2.1. *Suppose that $w(x,t)$ is a solution of the heat equation in $\Omega\times(0,\infty)$, and assume that $h(s)$ is a C^2 function on the real axis satisfying $h''(s) \geqslant 0$ for $s \in (-\infty,+\infty)$, then the function $h(w(x,t))$ is sub-parabolic in $\Omega \times (0,\infty)$.*

Proof. A straightforward calculation yields

$$\begin{aligned}\Big[\Delta - \frac{\partial}{\partial t}\Big] h(w(x,t)) &= h''(w(x,t)) \sum_{j=1}^{n} \Big(\frac{\partial w}{\partial x_j}\Big)^2 \\ &\geqslant 0 \text{ in } \Omega \times (0,\infty),\end{aligned}$$

due to the assumption $h'' \geqslant 0$.

Applying this proposition with $h(s) = s^2$ it follows that $|\operatorname{grad} u|^2$ is sub-parabolic. We can then obtain the estimate (1.7) on the basis of the maximum principle, by showing that the exterior normal derivative of $|\operatorname{grad} u|^2$ at any point on the lateral boundary $\partial\Omega \times (0,\infty)$ in non-positive:

$$\frac{\partial}{\partial n} |\operatorname{grad} u|^2 \Big|_{\partial\Omega\times(0,\infty)} \leqslant 0; \tag{2.2}$$

and we will devote the next two sections to establishing this as a result of the assumption (1.6).

For the sake of completeness we include a proof of the relevant version of the maximum principle that we are using. □

Theorem 2.2. *Let $v(x,t)$ be a non-negative sub-parabolic function in $\Omega \times (0,T]$ which is continuous in $\overline{\Omega} \times [0,T]$, with x and t derivatives continuous in $\overline{\Omega} \times (0,T]$ and with $\partial\Omega$ assumed to be C^1. Then if*

$$\frac{\partial v}{\partial n} \leqslant 0 \quad \text{along } \partial\Omega \times (0,T], \tag{2.3}$$

we have

$$v(x,t) \leqslant \max_{\overline{\Omega}} v(x,0) \quad \text{in } \Omega \times (0,T]. \tag{2.4}$$

Proof. Multiplying both sides of the inequality

$$\frac{\partial v}{\partial t} \leqslant \Delta v \quad \text{in } \Omega \times (0,T]$$

by v^p, $p \geqslant 1$, integrating over Ω and then integrating by parts, we obtain

$$\begin{aligned} \int_\Omega v^p \frac{\partial v}{\partial t} \, \mathrm{d}x &\leqslant \int_\Omega v^p \Delta v \, \mathrm{d}x \\ &= -\int_\Omega p v^{p-1} \left|\operatorname{grad} v\right|^2 \, \mathrm{d}x + \int_{\partial\Omega} v^p \frac{\partial v}{\partial n} \, \mathrm{d}\sigma \leqslant 0 \end{aligned}$$

for $t \in (0,T]$, in view of (2.3). Since the integral on the left is identical with

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_\Omega \frac{v^{p+1}}{p+1} \, \mathrm{d}x\Big),$$

we conclude that $\int_\Omega v^{p+1} \, \mathrm{d}x$ is a decreasing function of t:

$$\int_\Omega v^{p+1}(x,t) \, \mathrm{d}x \leqslant \int_\Omega v^{p+1}(x,s) \, \mathrm{d}x, \quad 0 < s < t \leqslant T.$$

Taking $p+1$ roots and passing to the limit as $p \to \infty$, this leads to

$$\max_{\overline{\Omega}} v(x,t) \leqslant \max_{\overline{\Omega}} v(x,s), \quad 0 < s < t \leqslant T.$$

Finally, by letting $s \downarrow 0$, we arrive at the desired result (2.4). $\square$

Section 3

In this section we will introduce a coordinate transformation on which we will base the proof of (2.2), the non-positivity of the exterior normal derivative of $|\mathrm{grad}\, u|^2$ on $\partial\Omega$.

Our starting point for defining this transformation is the function

$$x_n = g(x_1, \ldots, x_{n-1}) \tag{3.1}$$

introduced in Section 1 and which describes the surface constituting that portion of $\partial\Omega$ lying in a sufficiently small neighborhood of the point $p \in \partial\Omega$, with p placed at the origin of our coordinate system. Our assumptions regarding g, were that it was a C^3 function for $(x_1, \ldots, x_{n-1})$ in a neighborhood of $(x_1 = 0, \ldots, x_{n-1} = 0)$ with

$$g(0, \ldots, 0) = 0. \tag{3.2}$$

We further assumed the positive x_n direction to correspond to the outward normal direction on $\partial\Omega$ at p. This means that the plane $x_n = 0$ is tangent to the surface described by (3.1) at the origin; and so

$$\Big(\frac{\partial}{\partial x_j}\Big) g(x_1, \ldots, x_{n-1})\Big|_{x_1=0,\ldots,x_{n-1}=0} = 0,\ j = 1, \ldots, n-1. \tag{3.3}$$

We now define a coordinate change from $\xi = (\xi_1, \ldots, \xi_n)$ to $x = (x_1, \ldots, x_n)$ in accordance with the following scheme: Starting from a point $(\xi_1, \ldots, \xi_{n-1}, g(\xi_1, \ldots, \xi_{n-1}))$ on the surface describing $\partial\Omega$, we proceed ξ_n units in the direction of the outward normal to the surface thereby arriving at the point with the coordinates $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$. The connection between the original point $(\xi_1, \ldots, \xi_{n-1}, g(\xi_1, \ldots, \xi_{n-1}))$ and $(x_1, \ldots, x_n)$ is given by

(3.4)

$$\begin{cases} x_j = \xi_j - g_{\xi_j}(\xi_1, \ldots, \xi_{n-1})\Big(1 + \sum\limits_{k=1}^{n-1} g^2_{\xi_k}(\xi_1, \ldots, \xi_{n-1})\Big)^{-\frac{1}{2}} \xi_n,\ j = 1, \ldots, n-1, \\ \text{and} \\ x_n = g(\xi_1, \ldots, \xi_{n-1}) + \Big(1 + \sum\limits_{k=1}^{n-1} g^2_{\xi_k}(\xi_1, \ldots, \xi_{n-1})\Big)^{-\frac{1}{2}} \xi_n. \end{cases}$$

We may view these equations as defining either a coordinate change from $\xi = (\xi_1, \ldots, \xi_n)$ to $x = (x_1, \ldots, x_n)$ or the other way around. For short we will write equations (3.4) as $x = x(\xi)$ and componentwise as $x_j = x_j(\xi) = x_j(\xi_1, \ldots, \xi_n)$, $j = 1, \ldots, n$. Similarly, the inverse transformation, which we will show in a moment, exists, will be denoted by $\xi = \xi(x)$ and componentwise by $\xi_j = \xi_j(x) = \xi_j(x_1, \ldots, x_n)$, $j = 1, \ldots, n$.

For our purposes, the essential point about this transformation is that differentiation in the outward normal direction on $\partial\Omega$ corresponds to differentiation with respect to ξ_n when $\xi_n = 0$. More precisely, if $\varphi(x)$ represents a function in the x coordinates and $\psi(\xi)$ represents the corresponding function in the ξ coordinates, i.e., $\psi(\xi) = \varphi(x(\xi))$, then

$$\frac{\partial\varphi(x)}{\partial n}\bigg|_{\partial\Omega} = \frac{\partial\psi(\xi)}{\partial\xi_n}\bigg|_{\xi_n=0}. \tag{3.5}$$

In the two propositions which follow we describe the main analytic properties of this transformation. For the first of these, which concerns the existence of the inverse transformation, we need only to assume that g is C^2.

Proposition 3.1. *Assume that $g(\xi_1, \dots, \xi_{n-1})$ is a C^2 function of $(\xi_1, \dots, \xi_{n-1})$ in some neighborhood of $(\xi_1 = 0, \dots, \xi_{n-1} = 0)$, satisfying the conditions* (3.2) *and* (3.3). *Then the equations* (3.4) *define a non-singular C^1 transformation $x = x(\xi)$ in a neighborhood of $\xi = 0$, which sends $\xi = 0$ into $x = 0$ and whose Jacobian at the origin is the identity matrix:*

$$\frac{\partial x}{\partial\xi}\bigg|_{\xi=0} = I. \tag{3.6}$$

Consequently, the inverse transformation $\xi = \xi(x)$ exists in a neighborhood of $x = 0$, is C^1 there, sends $x = 0$ into $\xi = 0$, and its Jacobian at the origin is also the identity matrix:

$$\frac{\partial\xi}{\partial x}\bigg|_{x=0} = I. \tag{3.7}$$

Proof. In view of (3.2), $x = x(\xi)$ sends $\xi = 0$ into $x = 0$. By the inverse function theorem all the other assertions made in the proposition will follow the moment (3.6) is established. A straightforward calculation using (3.3), show that

$$\frac{\partial x_j}{\partial\xi_k}\bigg|_{\xi=0} = \delta_{jk}\ j,\ k = 1, \dots, n, \tag{3.8}$$

δ_{jk} denoting the Kronecker delta; which proves (3.6).

The reason we assumed g to be C^3 rather than just C^2 is because we will need to take second derivatives of $x(\xi)$ and $\xi(x)$. The relevant facts concerning these second derivatives are contained in □

Proposition 3.2. *Under the assumption that $g(\xi_1,\dots,\xi_{n-1})$ is C^3 in a neighborhood of $(\xi_1 = 0,\dots,\xi_{n-1} = 0)$, the transformation $x = x(\xi)$ and its inverse $\xi = \xi(x)$ referred to in Proposition* 3.1 *are C^2 in neighborhoods of $x = 0$ and $\xi = 0$, respectively.*

Furthermore, the second derivatives of these transformations at the origin are related in the following way:

$$\frac{\partial}{\partial \xi_m}\Big(\frac{\partial \xi_j}{\partial x_l}\Big)\Big|_0 = -\frac{\partial}{\partial \xi_m}\Big(\frac{\partial x_j}{\partial \xi_l}\Big)\Big|_0, \tag{3.9}$$

and

$$\frac{\partial}{\partial x_m}\Big(\frac{\partial \xi_j}{\partial x_l}\Big)\Big|_0 = -\frac{\partial}{\partial \xi_m}\Big(\frac{\partial x_j}{\partial \xi_l}\Big)\Big|_0, \tag{3.10}$$

$j, l, m = 1,\dots,n$, and where the evaluation notation here means that we evaluate these derivatives at $\xi = 0$ or, equivalently, at $x = 0$.

In particular, we have

$$\frac{\partial^2 \xi_n}{\partial x_i^2}\Big|_0 = -\frac{\partial^2 x_n}{\partial \xi_i^2}\Big|_0 = -g_{\xi_i\xi_i}(0,\dots,0), \quad i = i,\dots,n-1, \tag{3.11}$$

and

$$\frac{\partial^2 \xi_n}{\partial x_n^2}\Big|_0 = -\frac{\partial^2 x_n}{\partial \xi_n^2}\Big|_0 = 0, \tag{3.12}$$

as well as

$$\frac{\partial}{\partial \xi_n}\Big(\frac{\partial \xi_n}{\partial x_n}\Big)\Big|_0 = -\frac{\partial^2 x_n}{\partial \xi_n^2}\Big|_0 = 0. \tag{3.13}$$

P r o o f. To establish (3.9) and (3.10) we begin by observing that as $x = x(\xi)$ and $\xi = \xi(x)$ are inverse to each other, so also are their Jacobian matrices: $\frac{\partial x}{\partial \xi}\frac{\partial \xi}{\partial x} = I$. In terms of the Jacobian entries this means that

$$\sum_{k=1}^{n} \frac{\partial x_j}{\partial \xi_k}\frac{\partial \xi_k}{\partial x_l} = \delta_{jl}, \quad j, l = 1,\dots,n.$$

Differentiating with respect to ξ_m, $m = 1,\dots,n$, yields

$$\sum_{k=1}^{n} \Big[\frac{\partial}{\partial \xi_m}\Big(\frac{\partial x_j}{\partial \xi_k}\Big)\frac{\partial \xi_k}{\partial \chi_l} + \frac{\partial x_j}{\partial \xi_k}\frac{\partial}{\partial \xi_m}\Big(\frac{\partial \xi_k}{\partial x_l}\Big)\Big] = 0.$$

Evaluating at $x = \xi = 0$, making use of (3.7) and (3.8) according to which $\frac{\partial \xi_k}{\partial x_l}\big|_0 = \delta_{kl}$ and $\frac{\partial x_j}{\partial \xi_k}\big|_0 = \delta_{jk}$, we arrive at

$$\sum_{k=1}^{n} \Big[\frac{\partial}{\partial \xi_m} \Big(\frac{\partial x_j}{\partial \xi_k} \Big)\Big|_0 \delta_{kl} + \delta_{jk} \frac{\partial}{\partial \xi_m} \Big(\frac{\partial \xi_k}{\partial x_l} \Big)\Big|_0 \Big] = 0;$$

and hence

$$\frac{\partial}{\partial \xi_m} \Big(\frac{\partial x_j}{\partial \xi_l} \Big)\Big|_0 + \frac{\partial}{\partial \xi_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0 = 0$$

for $j, l, m = 1, \ldots, n$, which is (3.9).

To prove (3.10) we note that as the right sides of (3.9) and (3.10) are identical, (3.10) will follow the moment we can show that the left sides of (3.9) and (3.10) are the same:

$$\frac{\partial}{\partial \xi_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0 = \frac{\partial}{\partial x_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0. \tag{3.14}$$

To do so we apply the chain rule to carry out the indicated differentiation on the left side of (3.14):

$$\frac{\partial}{\partial \xi_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} \Big(\frac{\partial \xi_j}{\partial x_l} \Big) \frac{\partial x_k}{\partial \xi_m}.$$

Evaluating this at $x = \xi = 0$, using $\frac{\partial x_k}{\partial \xi_m}\big|_0 = \delta_{km}$ results in

$$\frac{\partial}{\partial \xi_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0 = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0 \delta_{km} = \frac{\partial}{\partial x_m} \Big(\frac{\partial \xi_j}{\partial x_l} \Big)\Big|_0,$$

which is (3.14).

Finally, we turn to the verification of (3.11)–(3.13). To accomplish this for (3.11) and (3.12) we apply (3.10) with $j = n$ and $m = l = i$:

$$\frac{\partial^2 \xi_n}{\partial x_i^2}\Big|_0 = -\frac{\partial^2 x_n}{\partial \xi_i^2}\Big|_0, \quad i = 1, \ldots n.$$

Using (3.4) we can explicitly calculate the derivative on the right side of this equation. Carrying this out we find that

$$\frac{\partial^2 x_n}{\partial \xi_i^2}\Big|_0 = \begin{cases} g_{\xi_i \xi_i}(0, \ldots, 0) & \text{for } i = 1, \ldots n-1, \\ 0 & \text{for } i = n, \end{cases} \tag{3.15}$$

which establishes (3.11) and (3.12).

Lastly, the remaining relation (3.13) follows by applying (3.9) with $j = m = l = n$ which yields

$$\frac{\partial}{\partial \xi_n}\Big(\frac{\partial \xi_n}{\partial x_n}\Big)\Big|_0 = -\frac{\partial^2 x_n}{\partial \xi_n^2}\Big|_0;$$

and since, as we just observed in (3.15), the derivative on the right vanishes, (3.13) is proved. □

Section 4

In this section we will prove the non-positivity of the normal derivative of $|\operatorname{grad} u|^2$ on $\partial\Omega$, i.e. for p an arbitrary point on $\partial\Omega$ we will show that

$$\frac{\partial}{\partial n}|\operatorname{grad} u|^2\Big|_p \leqslant 0. \tag{4.1}$$

(We suppress any mention of the time variable t because it plays no role in our computations; it is to be understood as being fixed at an arbitrary positive value.)

In order to establish (4.1) we introduce the transformation $x = x(\xi)$ defined in the previous section, in which $\xi_n = 0$ corresponds to $\partial\Omega$ and p corresponds to $\xi = 0$. Our first step will be to compute $|\operatorname{grad} u|^2$ in terms of the ξ coordinates. For this purpose let $v(\xi, t)$ denote the function $u(x, t)$ referred to ξ coordinates, i.e. $v(\xi, t) = u(x(\xi), t)$. We then find that

$$|\operatorname{grad} u(x,t)|^2 = \sum_{i=1}^{n}\Big(\frac{\partial u}{\partial x_i}\Big)^2 = \sum_{1\leqslant j,k\leqslant n} b_{jk}\frac{\partial v}{\partial \xi_j}\frac{\partial v}{\partial \xi_k} \tag{4.2}$$

where

$$b_{jk} = \sum_{i=1}^{n}\frac{\partial \xi_j}{\partial x_i}\frac{\partial \xi_k}{\partial x_i}, \quad j,\ k = 1,\dots,n. \tag{4.3}$$

Next we wish to calculate the normal derivative $\frac{\partial}{\partial n}|\operatorname{grad} u|^2$ on $\partial\Omega$ in the ξ coordinates. Since differentiation in the outward normal direction on $\partial\Omega$ corresponds to differentiation with respect to ξ_n when $\xi_n = 0$ (see (3.5)), it follows that

$$\begin{aligned}\frac{\partial}{\partial n}|\operatorname{grad} u|^2\Big|_{\partial\Omega} &= \frac{\partial}{\partial \xi_n}\Big(\sum_{1\leqslant j,k\leqslant n} b_{jk}\frac{\partial v}{\partial \xi_j}\frac{\partial v}{\partial \xi_k}\Big)\Big|_{\xi_n=0}\\ &= \sum_{1\leqslant j,k\leqslant n}\frac{\partial}{\partial \xi_n}(b_{jk})\frac{\partial v}{\partial \xi_j}\frac{\partial v}{\partial \xi_k} + \sum_{1\leqslant j,k\leqslant n} 2b_{jk}\frac{\partial^2 v}{\partial \xi_n\partial \xi_j}\frac{\partial v}{\partial \xi_k}\Big|_{\xi_n=0}.\end{aligned}$$

But as u vanishes on $\partial\Omega$, v vanishes when $\xi_n = 0$ and consequently so do all derivatives $\frac{\partial v}{\partial \xi_k}$ with $k \neq n$ vanish when $\xi_n = 0$. Hence

$$\frac{\partial}{\partial n}\left|\operatorname{grad} u\right|^2\Big|_{\partial\Omega} = \frac{\partial}{\partial \xi_n}(b_{nn})\Big(\frac{\partial v}{\partial \xi_n}\Big)^2 + \sum_{j=1}^{n} 2b_{jn}\frac{\partial^2 v}{\partial \xi_n \partial \xi_j}\frac{\partial v}{\partial \xi_n}\Big|_{\xi_n=0}.$$

Finally, evaluating at the point p which corresponds to $\xi = 0$, we arrive at

$$\frac{\partial}{\partial n}\left|\operatorname{grad} u\right|^2\Big|_{p} = 2b_{nn}\frac{\partial^2 v}{\partial \xi_n^2}\frac{\partial v}{\partial \xi_n}\Big|_{\xi=0}, \tag{4.4}$$

because, as we shall in a moment,

$$\frac{\partial}{\partial \xi_n}(b_{nn})\Big|_{\xi=0} = 0 \tag{4.5}$$

and

$$b_{jn}\big|_{\xi=0} = 0 \quad \text{for } j = 1, \dots, n-1. \tag{4.6}$$

To verify (4.6), we set $\xi = 0$ in the definition (4.3) for b_{nj} and then use the evaluation $\frac{\partial \xi_j}{\partial x_i}\big|_{x=0} = \frac{\partial \xi_j}{\partial x_i}\big|_{\xi=0} = \delta_{ji}$ (see (3.7)) as follows:

$$b_{jn}\big|_{\xi=0} = \sum_{i=1}^{n}\frac{\partial \xi_j}{\partial x_i}\frac{\partial \xi_n}{\partial x_i}\Big|_{\xi=0} = \sum_{i=1}^{n}\delta_{ji}\delta_{ni} = \delta_{jj}\delta_{nj} = 0$$

for $j \neq n$.

The verification of (4.5) is similarly straightforward: We differentiate the defining equation (4.3) for b_{nn} with respect to ξ_n and then set $\xi = 0$ thereby obtaining

$$\begin{aligned}\frac{\partial}{\partial \xi_n}(b_{nn})\Big|_{\xi=0} &= \frac{\partial}{\partial \xi_n}\sum_{i=1}^{n}\Big(\frac{\partial \xi_n}{\partial x_i}\Big)^2\Big|_{\xi=0} = \sum_{i=1}^{n} 2\frac{\partial \xi_n}{\partial x_i}\Big|_{\xi=0}\frac{\partial}{\partial \xi_n}\Big(\frac{\partial \xi_n}{\partial x_i}\Big)\Big|_{\xi=0} \\ &= \sum_{i=1}^{n} 2\delta_{ni}\frac{\partial}{\partial \xi_n}\Big(\frac{\partial \xi_n}{\partial x_i}\Big)\Big|_{\xi=0} = 2\frac{\partial}{\partial \xi_n}\Big(\frac{\partial \xi_n}{\partial x_n}\Big)\Big|_{\xi=0} = 0\end{aligned}$$

in view of (3.13); and this proves (4.5).

The next step in proving (4.1) involves the Laplacian of u; we shall need the expression for Δu in terms of the ξ coordinates. This is given by

$$\Delta u = \sum_{1\leqslant j,k\leqslant n} b_{jk}\frac{\partial^2 v}{\partial \xi_j \partial \xi_k} + \sum_{j=1}^{n} c_j\frac{\partial v}{\partial \xi_j}, \tag{4.7}$$

where the b_{jk} are as defined in (4.3) and

$$c_j = \sum_{i=1}^{n} \frac{\partial^2 \xi_j}{\partial x_i^2}, \ j = 1, \ldots, n. \tag{4.8}$$

Now by standard regularity theory [1], the derivatives $u_{x_j x_k}$ and u_t have continuous extensions to the lateral boundary $\partial\Omega \times (0, \infty)$; and so the equation $u_t - \Delta u = 0$ is satisfied on $\partial\Omega \times (0, \infty)$. But as u vanishes on $\partial\Omega \times (0, \infty)$, u_t also vanishes on $\partial\Omega \times (0, \infty)$. Consequently we must have $\Delta u = 0$ on $\partial\Omega$ for $t > 0$. Therefore, evaluating the left side of (4.7) on $\partial\Omega$ and correspondingly, the right side on $\xi_n = 0$, we find that

$$0 = \Delta u\big|_{\partial\Omega} = \sum_{1 \leqslant j,k \leqslant n} b_{jk} \frac{\partial^2 v}{\partial \xi_j \partial \xi_k} + \sum_{j=1}^{n} c_j \frac{\partial v}{\partial \xi_j}\Big|_{\xi_n = 0}.$$

Taking into account the vanishing of v when $\xi_n = 0$, which implies that all the derivatives of v which do not involve ξ_n also vanish when $\xi_n = 0$, the preceeding becomes

$$0 = b_{nn} \frac{\partial^2 v}{\partial \xi_n^2} + \sum_{j=1}^{n-1} 2b_{jn} \frac{\partial^2 v}{\partial \xi_j \partial \xi_n} + c_n \frac{\partial v}{\partial \xi_n}\Big|_{\xi_n = 0}.$$

Evaluating at $\xi = 0$ then yields

$$b_{nn} \frac{\partial^2 v}{\partial \xi_n^2}\Big|_{\xi=0} = -c_n \frac{\partial v}{\partial \xi_n}\Big|_{\xi=0}, \tag{4.9}$$

in view of (4.6).

We now multiply both sides of (4.9) by $2\frac{\partial v}{\partial \xi_n}\big|_{\xi=0}$ and insert the resulting expression for $2b_{nn}\frac{\partial^2 v}{\partial \xi_n^2}\frac{\partial v}{\partial \xi_n}\big|_{\xi=0}$ into (4.4) thereby obtaining

$$\frac{\partial}{\partial n} |\text{grad}\, u|^2 \Big|_p = -2c_n \Big(\frac{\partial v}{\partial \xi_n}\Big)^2\Big|_{\xi=0}.$$

Finally, to evaluate c_n as defined by (4.8) at $\xi = 0$, we use (3.11) and (3.12):

$$c_n\Big|_{\xi=0} = \sum_{i=1}^{n} \frac{\partial^2 \xi_n}{\partial x_i^2}\Big|_{\xi=0} = -\sum_{i=1}^{n-1} g_{\xi_i \xi_i}(0, \ldots, 0).$$

Therefore

$$\frac{\partial}{\partial n} |\text{grad}\, u|^2 \Big|_p = 2\Big(\sum_{i=1}^{n-1} g_{\xi_i \xi_i}(0, \ldots, 0)\Big)\Big(\frac{\partial v}{\partial \xi_n}\Big)^2\Big|_{\xi=0} \leqslant 0$$

because of the mean curvature hypothesis (1.6). This proves the desired result (4.1) and completes the proof of Theorem 1.1.

Section 5

In this section we will sketch the construction of an example showing that the mean curvature condition (1.6) is required to establish the estimate (1.7).

Our example will be constructed with Ω in $\mathbb{R}^2$, in which case (1.6) just amounts to a convexity condition. Accordingly, we seek our example so that Ω is some simple non-convex set in $\mathbb{R}^2$; and such Ω's are furnished by the circular sectors S_α whose polar coordinate description is

$$S_\alpha = \{(r,\theta)\colon\ 0 < r < 1,\ 0 < \theta < \alpha\}$$

provided that the central angle $\alpha > \pi$. (These S_α's will not quite do for our example because their boundaries are not C^3; however, the example that we will ultimately devise will be based on a "smoothed out" version of $S_{\frac{3\pi}{2}}$.)

Next, we construct solutions of the heat equation in $S_\alpha \times (0,\infty)$ with the aid of the Bessel functions of index p:

$$J_p(s) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(p+m+1)} \left(\frac{s}{2}\right)^{2m+p}. \tag{5.1}$$

Because of the differential equation

$$s^2 J_p''(s) + sJ_p'(s) + (s^2 - p^2)J_p(s) = 0$$

satisfied by $J_p(s)$, it is easily verified that the functions

$$g(r,\theta) = J_{\frac{n\pi}{\alpha}}(\lambda r)\sin\left(\frac{n\pi\theta}{\alpha}\right),\ \ n = 1,2,\dots \tag{5.2}$$

are solutions of

$$\Delta g = -\lambda^2 g \tag{5.3}$$

for all $r > 0$ and all $\theta \in (-\infty,\infty)$. In particular $g(r,\theta)$ satisfies this equation inside S_α and vanishes on the boundary of S_α if λ is a zero of $J_{\frac{n\pi}{\alpha}}(s)$; i.e. $J_{\frac{n\pi}{\alpha}}(\lambda) = 0$.

Now consider the circular sector $S_{\frac{3\pi}{2}}$ with central angle $\frac{3\pi}{2}$, and let μ and ν denote any pair of distinct zeros for the function $J_{\frac{2}{3}}(s)$. It follows from the above, that for any choice of a and b, the function

$$u = \left[aJ_{\frac{2}{3}}(\mu r)e^{-\mu^2 t} + bJ_{\frac{2}{3}}(\nu r)e^{-\nu^2 t}\right]\sin\left(\frac{2}{3}\theta\right) \tag{5.4}$$

will be a solution of the heat equation in $S_{\frac{3\pi}{2}} \times (0,\infty)$, which vanishes on $\partial S_{\frac{3\pi}{2}} \times (0,\infty)$ and which takes on the initial values

$$f = \left[aJ_{\frac{2}{3}}(\mu r) + bJ_{\frac{2}{3}}(\nu r)\right] \sin\left(\frac{2}{3}\theta\right) \tag{5.5}$$

in $S_{\frac{3\pi}{2}}$.

The function f is clearly continuous in the closure of $S_{\frac{3\pi}{2}}$ and vanishes on its boundary. In regard to its differentiabilty properties, an examination of the series representing f:

$$f = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(\frac{2}{3}+m+1)} \left[a\left(\frac{\mu}{2}\right)^{2m+\frac{2}{3}} + b\left(\frac{\nu}{2}\right)^{2m+\frac{2}{3}}\right] r^{2m+\frac{2}{3}} \sin\left(\frac{2}{3}\theta\right)$$

reveals that, although every term with $m > 0$ is C^1 in the closure of $S_{\frac{3\pi}{2}}$, in general, this is not so for the term corresponding to $m = 0$ because

$$\left|\text{grad}\left[r^{\frac{2}{3}} \sin\left(\frac{2}{3}\theta\right)\right]\right|^2 = \left(\frac{2}{3}\right)^2 r^{-\frac{2}{3}} \to \infty \quad \text{as } r \downarrow 0. \tag{5.6}$$

Consequently, f will not be in $C^1(\overline{S}_{\frac{3\pi}{2}})$ unless the term corresponding to $m = 0$ does not appear; and this will be the case if the bracketed factor in that term vanishes:

$$a\left(\frac{\mu}{2}\right)^{\frac{2}{3}} + b\left(\frac{\nu}{2}\right)^{\frac{2}{3}} = 0. \tag{5.7}$$

We now choose a and b in accordance with this condition, thus assuring that the resulting function f is in $C^1(\overline{S}_{\frac{3\pi}{2}})$ and vanishes on $\partial S_{\frac{3\pi}{2}}$. Nevertheless, the solution u given by (5.4) of the initial boundary value problem (1.4) which is generated by this initial data f, does not have a bounded gradient in $S_{\frac{3\pi}{2}}$ for fixed $t > 0$. Again, this follows from a series representation, namely

$$u = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(\frac{2}{3}+m+1)} \left[a\left(\frac{\mu}{2}\right)^{2m+\frac{2}{3}} \mathrm{e}^{-\mu^2 t} + b\left(\frac{\nu}{2}\right)^{2m+\frac{2}{3}} \mathrm{e}^{-\nu^2 t}\right] r^{2m+\frac{2}{3}} \sin\left(\frac{2}{3}\theta\right).$$

Just as in the series for f, all the terms with $m > 0$ are in $C^1(\overline{S}_{\frac{3\pi}{2}})$ (for fixed $t > 0$) and the series converges so rapidly that the sum of all these terms is also in $C^1(\overline{S}_{\frac{3\pi}{2}})$. Only the term corresponding to $m = 0$ fails to be in $C^1(\overline{S}_{\frac{3\pi}{2}})$; in fact its gradient in $S_{\frac{3\pi}{2}}$ is unbounded on account of (5.6). Moreover, unlike the situation for f, this term actually appears in the series because the corresponding bracketed factor here does not vanish:

$$a\left(\frac{\mu}{2}\right)^{\frac{2}{3}} \mathrm{e}^{-\mu^2 t} + b\left(\frac{\nu}{2}\right)^{\frac{2}{3}} \mathrm{e}^{-\nu^2 t} \neq 0 \quad \text{for } t > 0. \tag{5.8}$$

It does not vanish because functions of t of the form on the left of (5.8) vanish at most for only one value of $t \in (-\infty, +\infty)$, and in view of (5.7) it already vanishes at $t = 0$.

In summary then we have constructed a solution u of (1.4) in $S_{\frac{3\pi}{2}} \times (0, \infty)$ whose gradient is unbounded for any $t > 0$, even though the initial function f meets all the requirements of Theorem 1.1. This does not yet provide us with the desired example showing the necessity of the mean curvature condition (1.6), because the underlying domain $S_{\frac{3\pi}{2}}$ does not have a C^3 boundary as assumed in the theorem. However, we can produce such an example based on the considerations above, by means of an appropriate approximation procedure which we describe without proof.

First, we approximate $S_{\frac{3\pi}{2}}$ by a sequence of expanding domains $\Omega_n \subset S_{\frac{3\pi}{2}}$, with C^∞ boundaries and which "converge" to $S_{\frac{3\pi}{2}}$ in the set theoretic sense:

$$\bigcup_{n=1}^{\infty} \Omega_n = S_{\frac{3\pi}{2}}. \tag{5.9}$$

At the same time, by multiplying f by suitable cut-off functions which vanish near $\partial\Omega$, we can construct a sequence of functions $\{f_n\}$ with $f_n \in C^1(\overline{\Omega}_n)$, vanishing on $\partial\Omega_n$, and converging to f in the sense that

$$\sup_{\overline{\Omega}_n} |f_n - f| \to 0 \text{ as } n \to \infty; \tag{5.10}$$

while

$$\operatorname{grad} f_n \to \operatorname{grad} f \quad \text{as } n \to \infty, \text{ pointwise in } \Omega \tag{5.11}$$

and boundedly, meaning that

$$\sup_{\overline{\Omega}_n} |\operatorname{grad} f_n| \leqslant B \quad \text{for all } n \tag{5.12}$$

and some number B (for B we can take a suitable constant multiple of $\sup_{\overline{\Omega}} |\operatorname{grad} f|$).

Next, consider the solutions u_n in $\Omega_n \times (0, \infty)$ of (1.4) generated by the initial data f_n. Because of the convergence (5.9) of Ω_n to Ω and the convergence (5.10) of f_n to f, the solutions u_n converge to the solution u in $S_{\frac{3\pi}{2}} \times (0, \infty)$ constructed above:

$$u_n \to u \quad \text{as } n \to \infty,$$

pointwise and boundedly in $S_{\frac{3\pi}{2}} \times (0, \infty)$. In turn this implies that for the gradients of the u_n's we have

$$\operatorname{grad} u_n \to \operatorname{grad} u \quad \text{as } n \to \infty, \tag{5.13}$$

pointwise in $S_{\frac{3\pi}{2}} \times (0,\infty)$.

Suppose now that we had an estimate of the form (1.7) holding without assuming the curvature condition (1.6) but only assuming that $\partial\Omega$ is C^3 and that $f \in C^1(\overline{\Omega})$ with f vanishing on $\partial\Omega$. Then that estimate would hold for the u_n's generated by the f_n's:

$$|\operatorname{grad} u_n| \leqslant \sup_{\overline{\Omega}_n} |\operatorname{grad} f_n| \tag{5.14}$$

in $\Omega_n \times (0,\infty)$. In view of (5.12) this would imply that

$$|\operatorname{grad} u_n| \leqslant B \quad \text{in } \Omega_n \times (0,\infty), \text{ for all } n.$$

Sending $n \to \infty$, we would then obtain, because of (5.13),

$$|\operatorname{grad} u| \leqslant B \quad \text{in } S_{\frac{3\pi}{2}} \times (0,\infty).$$

But this is a contradiction, since we known that the gradient of the function u constructed above is not bounded in $S_{\frac{3\pi}{2}} \times (0,\infty)$. It follows that the estimate (5.14) cannot hold for all the functions u_n; and so we will have arrived at our desired example.

References

[1] *A. Friedman*: Partial Differential Equations of Parabolic Type. Prentice Hall Inc., Englewood Cliffs, 1964.

[2] *O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural'ceva*: Linear and Quasilinear Equations of Parabolic Type. Amer. Math. Soc., Providence, 1968.

[3] *M. H. Protter and H. F. Weinberger*: Maximum Principles in Differential Equations. Prentice Hall Inc., Englewood Cliffs, 1967.

Author's address: Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37240, U.S.A.