

## Weak maximum principle for the heat equation

Harald Hanche-Olsen

In this note, we consider the standard heat equation

$$u_t - \Delta u = 0 \quad \text{in } \Omega_T$$

where  $\Omega \subset \mathbb{R}^n$  is a *bounded* region,  $\Omega_T = (0, T) \times \Omega$  with  $T > 0$ , and

$$u \in C(\overline{\Omega_T}) \cap C^2(\Omega_T).$$

We think of  $\Omega_T$  as an open *cylinder* with base  $\Omega$  and height  $T$ . Its closure is a closed cylinder:  $\overline{\Omega_T} = [0, T] \times \overline{\Omega}$ .

**Definition.** The *parabolic boundary* of  $\Omega_T$  is the set

$$\Gamma = (\{0\} \times \overline{\Omega}) \cup ([0, T] \times \partial\Omega).$$

Clearly,  $\Gamma$  is contained in the normal boundary  $\partial\Omega_T$ ; the difference is

$$\partial\Omega_T \setminus \Gamma = \{T\} \times \Omega.$$

We call  $\{T\} \times \Omega$  the *final boundary* of  $\Omega_T$  (nonstandard nomenclature).

**Observation.** If a  $C^2$  function  $v$  has a maximum at some point in  $\Omega_T$ , then  $v_t = 0$  and  $\Delta v \leq 0$  at that point, so we get  $v_t - \Delta v \geq 0$  there. Moreover, this holds at the final boundary as well, the only difference being that there, we can only conclude  $v_t \geq 0$  and  $\Delta v \leq 0$ . In other words,

$$v_t - \Delta v \geq 0 \quad \text{at any maximum in } \overline{\Omega_T} \setminus \Gamma.$$

We must face a minor technical glitch: The above statement requires that  $v$  is  $C^2$  up to and including the final boundary of  $\Omega_T$ . This complicates the proof of the following theorem, but only a little.

**Theorem 1** (The weak maximum principle). *Assume that  $u \in C(\overline{\Omega_T}) \cap C^2(\Omega_T)$  satisfies*

$$u_t - \Delta u \leq 0.$$

*Then  $u(t, \mathbf{x}) \leq \max_{\Gamma} u$  for all  $(t, \mathbf{x}) \in \overline{\Omega_T}$ . In other words,  $u$  achieves its maximum on the parabolic boundary.*

*Proof.* First, to deal with the “minor technical glitch” mentioned above, we shall strengthen the assumptions somewhat, and assume that  $u \in C^2((0, T] \times \Omega)$ . We will remove this extra assumption at the end.

Now let  $\varepsilon > 0$ , and put  $v(t, \mathbf{x}) = u(t, \mathbf{x}) - \varepsilon t$ . Then  $v_t - \Delta v \leq -\varepsilon < 0$ , and so it follows *immediately* from the Observation above that  $v$  cannot achieve its maximum anywhere other than at  $\Gamma$ . On the other hand, since  $v$  is continuous and  $\overline{\Omega_T}$  is compact,  $v$  does have a maximum in  $\overline{\Omega_T}$ , and so we must conclude that  $v(t, \mathbf{x}) \leq \max_{\Gamma} v$  for any  $(t, \mathbf{x}) \in \overline{\Omega_T}$ . But then  $u(t, \mathbf{x}) = v(t, \mathbf{x}) + \varepsilon t \leq \max_{\Gamma} v + \varepsilon T \leq \max_{\Gamma} u + \varepsilon T$ . Since this holds for any  $\varepsilon > 0$ , it finally follows that  $u(t, \mathbf{x}) \leq \max_{\Gamma} u$ , and the proof is complete, with the strengthened assumptions.

We now drop the requirement that  $u \in C^2((0, T] \times \Omega)$ . Given any point  $(t, \mathbf{x}) \in \Omega_T$ , pick some  $T'$  with  $t < T' < T$ . Then  $u \in C^2((0, T'] \times \Omega)$ , so the first part shows that  $u(t, \mathbf{x}) \leq \max_{\Gamma_{T'}} u$ . Here  $\Gamma_{T'}$  is the parabolic boundary of  $\Omega_{T'}$ . But  $\Gamma_{T'} \subset \Gamma$ , so we also have  $u(t, \mathbf{x}) \leq \max_{\Gamma} u$ . Finally, this also holds for  $t = T$ , since  $u$  is continuous on  $\overline{\Omega_T}$ . This, at last, completes the proof. ■

It should come as no surprise that there is also a *minimum* principle. It is proved by replacing  $u$  by  $-u$  in Theorem 1.

**Corollary 2** (The weak minimum principle). *Assume that  $u \in C(\overline{\Omega_T}) \cap C^2(\Omega_T)$  satisfies*

$$u_t - \Delta u \geq 0.$$

*Then  $u(t, \mathbf{x}) \geq \min_{\Gamma} u$  for all  $(t, \mathbf{x}) \in \overline{\Omega_T}$ . In other words,  $u$  achieves its minimum on the parabolic boundary.*

We will mostly be concerned with solutions of the heat equation  $u_t - \Delta u = 0$ , and for these, both the maximum principle and the minimum principle can be used. But we may also wish to study inhomogeneous equations  $u_t - \Delta u = f$ , and if  $f$  has a definite sign, one or the other principle will apply.

**Corollary 3** (Uniqueness for the heat equation). *There exists at most one solution  $u \in C(\overline{\Omega_T}) \cap C^2(\Omega_T)$  to the problem*

$$\begin{aligned} u_t - \Delta u &= f & \text{in } \Omega_T, \\ u &= g & \text{on } \Gamma. \end{aligned}$$

*Here,  $f$  and  $g$  are given functions on  $\Omega_T$  and  $\Gamma$ , respectively. (Thus  $g$  combines initial values and boundary values in one function.)*

*Proof.* Let  $u$  be the difference between two solutions to this problem: Then  $u$  solves the same problem, but with  $f = 0$  and  $g = 0$ . Thus  $u$  achieves both its minimum and maximum on  $\Gamma$ , but  $u = 0$  there, so  $u = 0$  everywhere. ■

The following corollary is proved in essentially the same way, by applying the minimum and maximum principles to  $u_1 - u_2$ . Note that it immediately implies the preceding corollary by taking  $g_1 = g_2$ .

**Corollary 4** (Continuous dependence on data). *Let  $u_1$  and  $u_2$  satisfy*

$$\left. \begin{array}{l} u_{it} - \Delta u_i = f \quad \text{in } \Omega_T, \\ u_i = g_i \quad \text{on } \Gamma, \end{array} \right\} \text{ for } i = 1, 2.$$

Then  $|u_1 - u_2| \leq \max_{\Gamma} |g_1 - g_2|$ .

**Exercise 1** (Continuous dependence on data, improved). Assume that  $u_1$  and  $u_2$  satisfy

$$\left. \begin{array}{l} u_{it} - \Delta u_i = f_i \quad \text{in } \Omega_T, \\ u_i = g_i \quad \text{on } \Gamma, \end{array} \right\} \text{ for } i = 1, 2.$$

Let  $\varphi = \sup_{\Omega_T} |f_1 - f_2|$  and  $\gamma = \max_{\Gamma} |g_1 - g_2|$ , and show that  $|u_1 - u_2| \leq \gamma + \varphi T$ .

Note that for any  $t$ , we can pick  $T = t$ , so we really get  $|u_1 - u_2| \leq \gamma + \varphi t$ .

*Hint:* Apply the maximum principle to  $u_1 - u_2 - \varphi t$  and  $u_2 - u_1 - \varphi t$ .

**Exercise 2.** Show that the maximum (and minimum) principle continues to hold if  $u_t - \Delta u$  is replaced by the more general

$$u_t - \nabla \cdot (A \nabla u)$$

where the (constant) real  $n \times n$  matrix  $A$  is symmetric and positive definite.

Here are some ingredients for a proof:

- The *Hessian* of  $u$  is defined to be the (symmetric!)  $n \times n$  matrix  $Hu$  with entries  $u_{x_i x_j}$ . At an interior maximum point,  $Hu$  is negative semidefinite, i.e.,  $\mathbf{y}^T H \mathbf{y} \leq 0$  for all  $\mathbf{y} \in \mathbb{R}^n$ . (Short proof: Take the second derivative of  $u(\mathbf{x} + s\mathbf{y})$  with respect to  $s$  where  $\mathbf{x}$  is a maximum point, and put  $s = 0$ .)
- The *Frobenius inner product* of two real matrices  $A$  and  $B$  is

$$\langle A, B \rangle_{\text{F}} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij} = \text{tr}(A^T B).$$

It turns out that

$$\nabla \cdot (A \nabla u) = \langle A, Hu \rangle_{\text{F}}.$$

- It is known that if  $A$  and  $B$  are positive semidefinite, then  $\langle A, B \rangle_{\text{F}} \geq 0$ . (Short proof: Since  $A$  is symmetric, we can write  $\langle A, B \rangle_{\text{F}} = \text{tr}(AB)$ .  $A$  will have a positive semidefinite square root  $A^{1/2}$ . A standard result on the trace gives  $\text{tr}(AB) = \text{tr}(A^{1/2} A^{1/2} B) = \text{tr}(A^{1/2} B A^{1/2})$ , but  $A^{1/2} B A^{1/2}$  is positive semidefinite, and such matrices have nonnegative trace.)

**Exercise 3.** Show that the maximum (and minimum) principle continues to hold if  $u_t - \Delta u$  is replaced by the even more general

$$u_t - \nabla \cdot (A \nabla u) + b(\nabla u),$$

where the real matrix  $A$  is symmetric and positive definite, provided the continuous function  $b$  satisfies  $b(\mathbf{0}) = 0$ . (For a simple and common example, let  $b(\nabla u) = \mathbf{b} \cdot \nabla u$ .)

**Remark.** In many PDE texts, the term  $\nabla \cdot (A \nabla u)$  is written out in detail as

$$\nabla \cdot (A \nabla u) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} u_{x_i x_j}.$$

Pedantically speaking, considering the order in which derivatives are taken, that should be

$$\nabla \cdot (A \nabla u) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} u_{x_j x_i},$$

but this makes no difference, due to symmetry.