The Heat Equation

The **heat equation**, also known as **diffusion equation**, describes in typical physical applications the evolution in time of the density u of some quantity such as heat, chemical concentration, population, etc. Let V be any smooth subdomain, in which there is no source or sink. Then the rate of change of the total quantity within V equals the negative of the net flux $\mathbf{F}$ through ∂V:

$$\frac{d}{dt}\int_V u dx = -\int_{\partial V} \mathbf{F}\cdot \nu dS.$$

The divergence theorem tells us

$$\frac{d}{dt}\int_V u dx = -\int_V \operatorname{div}\mathbf{F} dx.$$

Since V is arbitrary, we should have

$$u_t = -\operatorname{div}\mathbf{F}.$$

For many applications $\mathbf{F}$ is proportional to the (spatial) gradient $Du = (u_{x_1}, u_{x_2}, \cdots, u_{x_n})$ of u, but points in opposite direction (flux is from regions of higher to lower concentration):

$$\mathbf{F} = -aDu \quad (a > 0).$$

Therefore we obtain the equation

$$u_t = a\operatorname{div}(Du) = a\Delta u,$$

which is called the heat equation when $a = 1$.

If there is a source in Ω, we should obtain the following nonhomogeneous equation

$$u_t - \Delta u = f(x,t) \quad x\in\Omega, \quad t\in(0,\infty).$$

4.1. Fundamental solution of heat equation

As in Laplace's equation case, we would like to find some special solutions to the heat equation. The textbook gives one way to find such a solution, and a problem in the book gives another way. Here we discuss yet another way of finding a special solution to the heat equation.

4.1.1. Fundamental solution and the heat kernel. We first make the following observations.

(1) If $u_j(x,t)$ are solution to the *one-dimensional* heat equation $u_t = u_{xx}$ for $x \in \mathbb{R}$ and $t > 0$, $j = 1, \dots, n$, then

$$u(x_1, \dots, x_n, t) = u_1(x_1, t) \cdots u_n(x_n, t)$$

is a solution to the heat equation $u_t = \Delta u$ for $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and $t > 0$. This simple fact is left as an exercise.

(2) If $u(x,t)$ is a solution to $u_t = u_{xx}$ with $x \in \mathbb{R}$, then so is $w(\lambda)u(\lambda x, \lambda^2 t)$ for any real λ. Especially, we expect to have a solution of the form $u(x,t) = w(t)v(\frac{x^2}{t})$. A direct computation yields

$$u_t(x,t) = w'(t)v(\frac{x^2}{t}) - w(t)v'(\frac{x^2}{t})\frac{x^2}{t^2};$$
$$u_x(x,t) = w(t)v'(\frac{x^2}{t})\frac{2x}{t};$$
$$u_{xx}(x,t) = w(t)v''(\frac{x^2}{t})\frac{4x^2}{t^2} + w(t)v'(\frac{x^2}{t})\frac{2}{t}.$$

In order to satisfy the heat equation $u_t = u_{xx}$, we need

$$w'(t)v(\frac{x^2}{t}) - \frac{w(t)}{t}\left[v''(\frac{x^2}{t})\frac{4x^2}{t} + 2v'(\frac{x^2}{t}) + v'(\frac{x^2}{t})\frac{x^2}{t}\right] = 0.$$

Let $s = \frac{x^2}{t}$. Then separation of variables yields that

$$\frac{w'(t)t}{w(t)} = \frac{4sv''(s) + 2v'(s) + sv'(s)}{v(s)}.$$

Therefore, both sides of this equality must be constant, say, λ; so

$$s(4v'' + v') + \frac{1}{2}(4v' - 2\lambda v) = 0.$$

This equation is satisfied if we choose $\lambda = -1/2$ and $4v' + v = 0$; hence $v(s) = e^{-\frac{s}{4}}$. In this case,

$$\frac{w'(t)t}{w(t)} = -\frac{1}{2},$$

from which we have $w(t) = t^{-\frac{1}{2}}$. Therefore

$$u = u_1(x,t) = \frac{1}{\sqrt{t}}\, e^{-\frac{x^2}{4t}}$$

is a solution of $u_t = u_{xx}$ for $x \in \mathbb{R}$. By the observation (1) above, function

$$u(x_1, x_2, \cdots, x_n, t) = \frac{1}{t^{n/2}}\, e^{-\frac{|x|^2}{4t}}$$

is a solution of the heat equation $u_t = \Delta u$ for $t > 0$ and $x \in \mathbb{R}^n$.

Definition 4.1. The function

$$\Phi(x_1, \cdots, x_n, t) = \begin{cases} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} & (t > 0), \\ 0 & (t \le 0) \end{cases}$$

is called the **fundamental solution of heat equation** $u_t = \Delta u$.

The constant $\frac{1}{(4\pi)^{n/2}}$ in the fundamental solution $\Phi(x,t)$ is due to the following

Lemma 4.1. *For each $t > 0$,*

$$\int_{\mathbb{R}^n} \Phi(x_1, \dots, x_n, t) dx_1 \cdots dx_n = 1.$$

Proof. This is a straight forward computation using the fact

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

□

Remark 4.1. (i) The fundamental solution $\Phi(x,t)$ of the heat equation is singular at $(0,0)$ and C^∞ in $(\mathbb{R}^n \times \mathbb{R}) \setminus \{(0,0)\}$. Furthermore, all *spatial* derivatives $D^\alpha \Phi$ are integrable on $\mathbb{R}^n$ for each $t > 0$. Moreover, as $\Phi(x,t) = t^{-n/2}\Phi(\frac{x}{\sqrt{t}}, 1)$,

$$\int_{\mathbb{R}^n} |D^\alpha \Phi(x,t)|\, dx = t^{-\frac{|\alpha|}{2}} \int_{\mathbb{R}^n} |D^\alpha \Phi(y,1)|\, dy.$$

(ii) As $t \to 0^+$, $\Phi(x,t) \to 0$ $(x \neq 0)$ and $\Phi(0,t) \to \infty$. In fact, below we show that $\Phi(\cdot, t) \to \delta_0$ in distribution on $\mathbb{R}^n$ as $t \to 0^+$.

Definition 4.2. We call the function

$$K(x,y,t) = \Phi(x-y,t) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x-y|^2}{4t}} \quad (x,\, y \in \mathbb{R}^n,\ t > 0)$$

the **heat kernel** in $\mathbb{R}^n$.

4.1.2. Initial-value problem. Consider the **initial-value problem** or **Cauchy problem** of the heat equation

$$\begin{cases} u_t = \Delta u, & x \in \mathbb{R}^n,\ t \in (0,\infty), \\ u(x,0) = g(x), & x \in \mathbb{R}^n. \end{cases} \tag{4.1}$$

Define

$$\begin{aligned} u(x,t) &= \int_{\mathbb{R}^n} \Phi(x-y,t) g(y) dy = \int_{\mathbb{R}^n} K(x,y,t) g(y)\, dy \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\, dy \quad (x \in \mathbb{R}^n,\ t > 0). \end{aligned} \tag{4.2}$$

Theorem 4.2. *Let $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and u be defined by (4.2). Then*

(i) $u \in C^\infty(\mathbb{R}^n \times (0,\infty)) \cap L^\infty(\mathbb{R}^n \times (0,\infty))$,

(ii) $u_t = \Delta u$ *on* $\mathbb{R}^n \times (0,\infty)$,

(iii) for each $x_0 \in \mathbb{R}^n$,

$$\lim_{x \to x_0,\, t \to 0^+} \int_{\mathbb{R}^n} \Phi(x-y,t) g(y) dy = g(x_0).$$

Proof. 1. Clearly, $K(x,y,t) > 0$ and $\int_{\mathbb{R}^n} K(x,y,t)\, dy = 1$ for all $x \in \mathbb{R}^n$, $t > 0$. Hence

$$|u(x,t)| \le \|g\|_{L^\infty} \int_{\mathbb{R}^n} K(x,y,t)\, dy = \|g\|_{L^\infty} \quad (x \in \mathbb{R}^n,\ t > 0).$$

Furthermore, since $K(x,y,t)$ and along all derivatives are uniformly bounded on $\mathbb{R}^n \times \mathbb{R}^n \times [\delta, \infty)$, we see that $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$ and

$$u_t(x,t) - \Delta u(x,t) = \int_{\mathbb{R}^n} [(\Phi_t - \Delta_x \Phi)(x-y,t)] g(y)\, dy = 0 \quad (x \in \mathbb{R}^n,\ t > 0).$$

2. Fix $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$. Choose $\delta > 0$ such that

$$|g(y) - g(x_0)| < \varepsilon \quad \forall\, |y - x_0| < \delta,\ y \in \mathbb{R}^n.$$

Then if $|x - x_0| < \delta/2$, we have

$$|u(x,t) - g(x_0)| = \left| \int_{\mathbb{R}^n} \Phi(x-y,t)[g(y) - g(x_0)]\, dy \right|$$

$$\le \int_{B(x_0,\delta)} \Phi(x-y,t)|g(y) - g(x_0)|\, dy + \int_{\mathbb{R}^n \setminus B(x_0,\delta)} \Phi(x-y,t)|g(y) - g(x)| dy$$

$$:= I + J.$$

Now $I \le \varepsilon \int_{\mathbb{R}^n} \Phi(x-y,t)\, dy = \varepsilon$. Furthermore, if $|x - x_0| < \delta/2$ and $|y - x_0| \ge \delta$, then

$$|y - x_0| \le |y - x| + |x - x_0| < |y - x| + \frac{\delta}{2} \le |x - y| + \frac{1}{2}|y - x_0|.$$

Thus $|y - x| \ge \frac{1}{2}|y - x_0|$. Consequently,

$$J \le 2\|g\|_{L^\infty} \int_{\mathbb{R}^n \setminus B(x_0,\delta)} \Phi(x-y,t)\, dy = \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B(x_0,\delta)} e^{-\frac{|x-y|^2}{4t}}\, dy$$

$$\le \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B(x_0,\delta)} e^{-\frac{|y-x_0|^2}{16t}}\, dy = C \int_{\mathbb{R}^n \setminus B(0,\delta/\sqrt{t})} e^{-\frac{|z|^2}{16}}\, dz \to 0,$$

as $t \to 0^+$. Hence if $|x - x_0| < \delta/2$ and $t > 0$ is sufficiently small, $|u(x,t) - g(x_0)| < 2\varepsilon$. □

Remark 4.2. (a) Solution $u(x,t)$ defined by (4.2) depends on the values of $g(y)$ at all points $y \in \mathbb{R}^n$ no matter how far y and x are away. Even the initial datum g is compactly supported, its **domain of influence** on the solution $u(x,t)$ is still all of $x \in \mathbb{R}^n$ as long as $t > 0$. This phenomenon is known as the **infinite speed of propagation** of disturbances of the heat equation.

(b) The solution u defined by (4.2) is in $C^\infty(\mathbb{R}^n \times (0,\infty))$ even if g is not continuous; actually u is real analytic in $\mathbb{R}^n \times (0,\infty)$. This phenomenon is known as the **smoothing effect** of the heat kernel.

4.1.3. Nonhomogeneous initial value problems. We consider the initial-value problem for heat equation with source:

$$\begin{cases} u_t - \Delta u = f(x,t) & (x \in \mathbb{R}^n,\ t > 0), \\ u(x,0) = 0 & (x \in \mathbb{R}^n). \end{cases}$$

A general method for solving nonhomogeneous problems of general linear evolution equations using the solutions of homogeneous problem with variable initial data is known as **Duhamel's principle**. We use the idea of this method to solve the above nonhomogeneous heat equation.

Given $s > 0$, we solve the following homogeneous problem

$$\begin{cases} \tilde{u}_t - \Delta \tilde{u} = 0 & \text{in } \mathbb{R}^n \times (s,\infty), \\ \tilde{u}(x,s) = f(x,s) & \text{for } x \in \mathbb{R}^n \end{cases} \tag{4.3}$$

to obtain a solution $\tilde{u}(x,t) = U(x,t;s)$, indicating the dependence on $s > 0$. By letting $\tilde{u}(x,t) = v(x,t-s)$ for a function v on $\mathbb{R}^n \times (0,\infty)$; then v solves the heat equation with initial data $v(x,0) = f(x,s)$. We use the heat kernel to have such a solution v as

$$v(x,t) = \int_{\mathbb{R}^n} K(x,y,t) f(y,s)\, dy \quad (x \in \mathbb{R}^n,\ t > 0).$$

In this way we obtain a solution $\tilde{u}$ to (4.3) as

$$\tilde{u}(x,t) = U(x,t;s) = \int_{\mathbb{R}^n} K(x,y,t-s)f(y,s)\,dy \quad (x \in \mathbb{R}^n,\ t > s > 0).$$

Then **Duhamel's principle** asserts that the function

$$u(x,t) = \int_0^t U(x,t;s)\,ds \quad (x \in \mathbb{R}^n,\ t > 0)$$

would be a solution to the original nonhomogeneous problem. Formally, $u(x,0) = 0$ and

$$u_t(x,t) = U(x,t;t) + \int_0^t U_t(x,t;s)\,ds = f(x,t) + \int_0^t \Delta U(x,t;s)\,ds = f(x,t) + \Delta u(x,t).$$

However, we have to justify the differentiation under the integral.

Rewriting,

$$\begin{aligned} u(x,t) &= \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s)f(y,s)\,dyds \\ &= \int_0^t \frac{1}{(4\pi(t-s))^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} f(y,s)\,dyds \quad (x \in \mathbb{R}^n,\ t > 0). \end{aligned} \tag{4.4}$$

In the following, we use $C_1^2(\Omega \times I)$ to denote the space of functions $u(x,t)$ on $\Omega \times I$ such that $u, u_t, u_{x_i}, u_{x_i x_j}$ are in $C(\Omega \times I)$.

Theorem 4.3. *Assume $f \in C_1^2(\mathbb{R}^n \times [0,\infty))$ is such that f, f_t, f_{x_i} and $f_{x_i x_j}$ are bounded on $\mathbb{R}^n \times [0,\infty)$. Define $u(x,t)$ by (4.4). Then $u \in C_1^2(\mathbb{R}^n \times (0,\infty))$ and satisfies*

(i) $u_t(x,t) - \Delta u(x,t) = f(x,t)$ $(x \in \mathbb{R}^n,\ t > 0)$,

(ii) for each $x_0 \in \mathbb{R}^n$,

$$\lim_{x \to x_0,\, t \to 0^+} u(x,t) = 0.$$

Proof. By the change of variables, we have that

$$u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)f(x-y,t-s)\,dyds \quad (x \in \mathbb{R}^n,\ t > 0).$$

As f, f_t, f_{x_i} and $f_{x_i x_j}$ are all bounded on $\mathbb{R}^n \times [0,\infty)$, it follows that, for all $x \in \mathbb{R}^n$ and $t > 0$,

$$u_t(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)f_t(x-y,t-s)\,dyds + \int_{\mathbb{R}^n} \Phi(y,t)f(x-y,0)\,dy$$

and

$$u_{x_i x_j} = \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)f_{x_i x_j}(x-y,t-s)\,dyds \quad (i,j = 1,2,\cdots,n).$$

This proves that $u \in C^2_1(\mathbb{R}^n \times (0,\infty))$. So

$$\begin{aligned}
u_t(x,t) - \Delta u(x,t) =& \int_0^t \int_{\mathbb{R}^n} \Phi(y,s)[(\partial_t - \Delta_x) f(x-y,t-s)]\, dyds \\
&+ \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy \\
=& \int_\varepsilon^t \int_{\mathbb{R}^n} \Phi(y,s)[(-\partial_s - \Delta_y) f(x-y,t-s)]\, dyds \\
&+ \int_0^\varepsilon \int_{\mathbb{R}^n} \Phi(y,s)[(-\partial_s - \Delta_y) f(x-y,t-s)]\, dyds \\
&+ \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy \\
:=& I_\varepsilon + J_\varepsilon + N.
\end{aligned}$$

Now $|J_\varepsilon| \le C \int_0^\varepsilon \int_{\mathbb{R}^n} \Phi(y,s) dyds \le C\varepsilon$. By integration by parts, we have

$$\begin{aligned}
I_\varepsilon =& \int_\varepsilon^t \int_{\mathbb{R}^n} [(\partial_s - \Delta_y)\Phi(y,s)] f(x-y,t-s)\, dyds \\
&+ \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y,t-\varepsilon)\, dy - \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy \\
=& \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y,t)\, dy + \int_{\mathbb{R}^n} \Phi(y,\varepsilon)[f(x-y,t-\varepsilon) - f(x-y,t)]\, dy \\
&- \int_{\mathbb{R}^n} \Phi(y,t) f(x-y,0)\, dy,
\end{aligned}$$

using $(\partial_s - \Delta_y)\Phi(y,s) = 0$. Therefore,

$$u_t(x,t) - \Delta u(x,t) = \lim_{\varepsilon\to 0^+} (I_\varepsilon + J_\varepsilon + N)$$

$$= \lim_{\varepsilon\to 0^+} \left[\int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y,t)\, dy + \int_{\mathbb{R}^n} \Phi(y,\varepsilon)[f(x-y,t-\varepsilon) - f(x-y,t)]\, dy \right]$$

$$= \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^n} \Phi(y,\varepsilon) f(x-y,t)\, dy = \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^n} \Phi(x-y,\varepsilon) f(y,t)\, dy = f(x,t),$$

by (iii) of Theorem 4.2. Finally, we also have $\|u(\cdot,t)\|_{L^\infty} \le t\|f\|_{L^\infty} \to 0$ as $t \to 0^+$. This completes the proof. □

4.1.4. Nonuniqueness of the heat equation. The Cauchy problem (4.1) of the heat equation does not have unique solution. In fact we have the following result.

Theorem 4.4 (Tychonoff's solution)**.** *There are infinitely many solutions to Problem* (4.1).

Proof. We only need to construct infinitely many nonzero solutions to the one-dimensional heat equation with 0 initial condition.

We first solve the following Cauchy problem

$$\begin{cases} u_t = u_{xx}, & x \in \mathbb{R},\ t \in (-\infty,\infty), \\ u(0,t) = g(t), & u_x(0,t) = 0, \quad t \in \mathbb{R}, \end{cases} \tag{4.5}$$

by formally expanding u as a Taylor series of x,

$$u(x,t) = \sum_{j=0}^\infty g_j(t) x^j.$$

A formal computation from $u_t = u_{xx}$ requires

$$g_0(t) = g(t),\ g_1(t) = 0,\ g_j'(t) = (j+2)(j+1)g_{j+2},\ j = 0, 1, 2, \ldots;$$

therefore

$$g_{2k}(t) = \frac{1}{(2k)!} g^{(k)}(t), \quad g_{2k+1}(t) = 0, \quad k = 0, 1, \ldots,$$

which leads to

$$u(x,t) = \sum_{k=0}^{\infty} \frac{g^{(k)}(t)}{(2k)!} x^{2k}. \tag{4.6}$$

We then would like to choose a function g so that this series defines a true solution of (4.5). For this purpose, let $g(t)$ be defined by

$$g(t) = \begin{cases} e^{-t^{-\alpha}}, & t > 0, \\ 0, & t \le 0, \end{cases}$$

where $\alpha > 1$ is a constant. Then it is a good exercise to show that there exists a number $\theta = \theta(\alpha) > 0$ with

$$|g^k(t)| \le \frac{k!}{(\theta t)^k} e^{-\frac{1}{2}t^{-\alpha}} \quad (k = 0, 1, 2, \ldots,\ \ t > 0),$$

and hence the function u defined above is a honest solution of (4.5) that also satisfies $u(x,0) = 0$; such a solution is called a **Tychonoff solution** to the heat equation with zero initial datum, and it is not identically zero as $u(0,t) = g(t)$ for all $t > 0$.

Actually, the Tychonoff solution $u(x,t)$ is an entire function of x for any real t, but is not analytic in t. Of course, for different $\alpha > 1$ we have the different Tychonoff solutions. ☐

4.2. Maximum principles for second-order linear parabolic equations

We now jump to the study of an important property called the **weak maximum principle** satisfied by the solutions of general second-order linear parabolic equations, including the heat equation.

4.2.1. Parabolic cylinder and second-order linear parabolic equations. We assume Ω is a bounded open set in $\mathbb{R}^n$, $T > 0$. Consider the **parabolic cylinder**

$$\Omega_T = \Omega \times (0,T] = \{(x,t) \mid x \in \Omega,\ t \in (0,T]\}.$$

We define the **parabolic boundary** of Ω_T to be

$$\Gamma_T = \partial'\Omega_T := \overline{\Omega_T} \setminus \Omega_T = (\partial\Omega \times [0,T]) \cup (\Omega \times \{t = 0\}). \tag{4.7}$$

We consider the second-order linear differential operator of the form

$$(\partial_t + L)u = u_t + Lu = u_t - \sum_{ij=1}^{n} a^{ij}(x,t) D_{ij}u + \sum_{i=1}^{n} b^i(x,t) D_i u + c(x,t)u,$$

where a^{ij}, b^i and c are given functions in Ω_T. Without loss of generality, we assume $a^{ij} = a^{ji}$.

Definition 4.3. The operator $\partial_t + L$ is called **parabolic** in Ω_T if there exists $\lambda(x,t) > 0$ in Ω_T such that

$$\sum_{i,j=1}^{n} a^{ij}(x,t)\xi_i\xi_j \ge \lambda(x,t)|\xi|^2 \quad \forall\,(x,t) \in \Omega_T,\ \xi \in \mathbb{R}^n.$$

If $\lambda(x,t) \geq \lambda_0 > 0$ in Ω_T for some constant $\lambda_0 > 0$, we say $\partial_t + L$ is **uniformly parabolic** in Ω_T.

4.2.2. Weak maximum principle.

Theorem 4.5 (Weak maximum principle)**.** *Let Ω be bounded open in $\mathbb{R}^n$. Assume $\partial_t + L$ is parabolic in Ω_T. Let $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$ satisfy $u_t + Lu \leq 0$ in Ω_T* (that is, u is a **subsolution of** $\partial_t + L$)*. Then*

$$(a) \qquad \max_{\overline{\Omega_T}} u = \max_{\partial'\Omega_T} u \quad \text{if } c(x,t) = 0 \text{ in } \Omega_T;$$

$$(b) \qquad \max_{\overline{\Omega_T}} u \leq \max_{\partial'\Omega_T} u^+ \quad \text{if } c(x,t) \geq 0 \text{ in } \Omega_T.$$

Proof. Fix $\varepsilon > 0$ and let $v = u - \varepsilon t$. Then

$$v_t + Lv = u_t + Lu - [\varepsilon + c(x,t)\varepsilon t] \leq -\varepsilon < 0 \quad \forall\,(x,t) \in \Omega_T$$

in both cases of (a) and (b) above. We claim

$$(a) \ \max_{\overline{\Omega_T}} v = \max_{\partial'\Omega_T} v \ \text{ if } c = 0 \text{ in } \Omega_T; \quad (b) \ \max_{\overline{\Omega_T}} v \leq \max_{\partial'\Omega_T} v^+ \ \text{ if } c \geq 0 \text{ in } \Omega_T. \tag{4.8}$$

Let $v(x_0,t_0) = \max_{\overline{\Omega_T}} v$ for some $(x_0,t_0) \in \overline{\Omega_T}$. If $(x_0,t_0) \in \Omega_T$, then $Dv(x_0,t_0) = 0$, $v_t(x_0,t_0) \geq 0$ and $(D_{ij}v(x_0,t_0)) \leq 0$ (nonnegative definite in the matrix sense). Hence

$$v_t + Lv = v_t - \sum_{i,j=1}^n a^{ij} D_{ij} v + cv \quad \text{at } (x_0,t_0). \tag{4.9}$$

In case (b) we assume $v(x_0,t_0) = \max_{\overline{\Omega_T}} v > 0$ since otherwise (4.8)(b) is obvious. Therefore, by (4.9), in both cases of (a) and (b) of (4.8), we have $v_t + Lv \geq 0$ at (x_0,t_0), which is contradiction to $v_t + Lv \leq -\varepsilon < 0$. Consequently, we must have $(x_0,t_0) \in \partial'\Omega_T$, which proves (4.8). Finally, by taking $\varepsilon \to 0^+$, the results for u follow. $\square$

Theorem 4.6 (Estimate for general $c(x,t)$)**.** *Let Ω be bounded open in $\mathbb{R}^n$. Assume $\partial_t + L$ is parabolic in Ω_T with the coefficeint $c(x,t)$ bounded in Ω_T. Let $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$ satisfy $u_t + Lu \leq 0$ in Ω_T. Then*

$$\max_{\overline{\Omega_T}} u \leq e^{CT} \max_{\partial'\Omega_T} u^+,$$

where $C \geq 0$ is any constant satisfying $C + c(x,t) \geq 0$ in Ω_T.

Proof. The inequality is obvious if $\max_{\overline{\Omega_T}} u \leq 0$. So we assume $\max_{\overline{\Omega_T}} u > 0$. Consider $w(x,t) = e^{-Ct} u(x,t)$. Then

$$w_t + Lw = e^{-Ct}(u_t + Lu) - Ce^{-Ct}u = e^{-Ct}(u_t + Lu) - Cw$$

and hence $w_t + (Lw + Cw) = e^{-Ct}(u_t + Lu) \leq 0$ in Ω_T. Observe that the operator $\partial_t + \tilde{L}$, where $\tilde{L}w = Lw + Cw$, is parabolic and has the zeroth order coefficient $\tilde{c}(x,t) = c(x,t) + C \geq 0$ in Ω_T. Hence, by the weak maximum principle,

$$\max_{\overline{\Omega_T}} w \leq \max_{\partial'\Omega_T} w^+,$$

which implies

$$\max_{\overline{\Omega_T}} u = \max_{\overline{\Omega_T}} (e^{Ct} w) \leq e^{CT} \max_{\overline{\Omega_T}} w \leq e^{CT} \max_{\partial'\Omega_T} w^+ \leq e^{CT} \max_{\partial'\Omega_T} u^+,$$

completing the proof. $\square$

Similarly, the **weak minimum principle** with $c(x,t) \ge 0$ and the estimate of minimum with general $c(x,t)$ also hold for a **supersolution** u; for example, we have the following result

Theorem 4.7. *Let Ω be bounded open in $\mathbb{R}^n$. Assume $\partial_t + L$ is parabolic in Ω_T with the coefficeint $c(x,t)$ bounded in Ω_T. Let $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$ satisfy $u_t + Lu \ge 0$ in Ω_T. Then*

$$\min_{\overline{\Omega_T}} u \ge e^{CT} \min_{\partial'\Omega_T} u^-,$$

where $u^- = -(-u)^+$ is the negative part of u and $C \ge 0$ is any constant satisfying $C + c(x,t) \ge 0$ in Ω_T.

4.2.3. Uniqueness of mixed-value problems. The estimate for general $c(x,t)$ implies the uniqueness of **mixed-value problem** for parabolic equations *regardless the sign of* $c(x,t)$. This is different from the elliptic equations.

Theorem 4.8. *Let Ω be bounded open in $\mathbb{R}^n$. Assume $\partial_t + L$ is parabolic in Ω_T with bounded coefficients in Ω_T. Then, given any f, h and g, the* mixed-value problem *or* initial boundary value problem

$$\begin{cases} u_t + Lu = f(x,t) & (x,t) \in \Omega_T, \\ u(x,t) = h(x,t) & x \in \partial\Omega,\ t \in [0,T], \\ u(x,0) = g(x) & x \in \Omega \end{cases}$$

can have at most one solution u in $C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$.

4.2.4. Uniqueness for the heat equation by energy method. Let us go back to the heat equation. We study another method for uniqueness of the heat equation for more smooth solutions on smooth domains.

Assume Ω is a bounded smooth domain in $\mathbb{R}^n$. Let $u \in C_1^2(\overline{\Omega_T})$ be a solution to the homogeneous mixed-value problem

$$u_t = \Delta u \ \text{ in } \Omega_T, \quad u = 0 \ \text{ on } \partial'\Omega_T.$$

Of course, by the previous uniqueness theorem, we know that $u \equiv 0$. However, we give another proof based on integration by parts; this is known as the **energy method**. Let $e(t)$ be the "energy" defined by

$$e(t) = \int_\Omega u(x,t)^2\, dx \quad (0 \le t \le T).$$

Then $e(0) = 0$, $e(t) \ge 0$ and is differentiable in $(0,T)$ and, by Green's identity,

$$e'(t) = 2\int_\Omega u u_t\, dx = 2\int_\Omega u\Delta u\, dx = -2\int_\Omega |Du|^2\, dx \le 0,$$

which implies that $e(t)$ is non-increasing in $(0,T)$. Hence $e(t) \le e(0) = 0$; so $e(t) \equiv 0$. This proves that $u \equiv 0$ in Ω_T.

The energy method can also be used to obtain other types of uniqueness result.

Theorem 4.9 (Uniqueness of mixed-Neumann value problem)**.** *Given any functions f, h and g, the* **mixed-Neumann value problem**

$$\begin{cases} u_t - \Delta u = f(x,t) & (x,t) \in \Omega_T, \\ \frac{\partial u}{\partial \nu}(x,t) = h(x,t), & x \in \partial\Omega,\ t \in [0,T], \\ u(x,0) = g(x), & x \in \Omega \end{cases}$$

can have at most one solution u in $C_1^2(\bar{\Omega} \times [0,T])$.

Proof. The proof is standard by showing the difference of any two solutions must be zero, which can be proved by a similar energy method as above. □

Theorem 4.10 (Backward uniqueness of the heat equation)**.** *Let $u \in C^2(\overline{\Omega_T})$ solve*

$$u_t = \Delta u \ \text{ in } \Omega_T, \quad u(x,t) = 0 \ \text{ on } \partial\Omega \times [0,T].$$

If $u(x,T) = 0$, then $u \equiv 0$ in Ω_T.

Proof. 1. Let $e(t) = \int_\Omega u(x,t)^2\,dx$ for $0 \le t \le T$. As above $e'(t) = -2\int_\Omega |Du|^2\,dx$ and hence

$$e''(t) = -4\int_\Omega Du \cdot Du_t = 4\int_\Omega (\Delta u)^2\,dx.$$

Now, by Hölder's inequality (or the Cauchy-Schwartz inequality)

$$\int_\Omega |Du|^2\,dx = -\int_\Omega u\Delta u\,dx \le \left(\int_\Omega u^2\,dx\right)^{1/2}\left(\int_\Omega (\Delta u)^2\,dx\right)^{1/2}$$

and so

$$(e'(t))^2 = 4\left(\int_\Omega |Du|^2\,dx\right)^2 \le e(t)e''(t).$$

2. If $e(t) = 0$ for all $0 \le t \le T$, then we are done. Otherwise, suppose $e(t) \not\equiv 0$ on $[0,T]$. Since $e(T) = 0$, we can find $0 \le t_1 < t_2 \le T$ such that

$$e(t) > 0 \ \text{ on } t \in [t_1, t_2), \quad e(t_2) = 0.$$

Set $f(t) = \ln e(t)$ for $t \in [t_1, t_2)$. Then

$$f''(t) = \frac{e''(t)e(t) - e'(t)^2}{e(t)^2} \ge 0 \quad (t_1 \le t < t_2).$$

This shows that f is convex on $[t_1, t_2)$; hence $f((1-\lambda)t_1 + \lambda t) \le (1-\lambda)f(t_1) + \lambda f(t)$ and thus $e((1-\lambda)t_1 + \lambda t) \le e(t_1)^{1-\lambda}e(t)^\lambda$ for all $0 < \lambda < 1$ and $t_1 < t < t_2$. Letting $t \to t_2^-$ and in view of $e(t_2) = 0$ we have

$$0 \le e((1-\lambda)t_1 + \lambda t_2) \le e(t_1)^{1-\lambda}e(t_2)^\lambda = 0 \quad \forall\, 0 < \lambda < 1,$$

which is a contradiction to the assumption $e(t) > 0$ on $[t_1, t_2)$. □

Remark 4.3. The **backward uniqueness theorem** for the heat equation asserts that if two temperature distributions on Ω agree at some time $T > 0$ and have the same boundary values at all earlier times $0 \le t \le T$ then these temperature distributions must be identical within Ω at all earlier times.

4.3. Uniqueness for Cauchy problems of the heat equation

We now study the maximum principle and uniqueness theorem for the heat equation on $\mathbb{R}^n \times (0,T]$.

4.3.1. Maximum principle for Cauchy problems.

Theorem 4.11 (Weak maximum principle for Cauchy problem). *Let* $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ *satisfy*

$$\begin{cases} u_t - \Delta u \le 0, & \text{in } \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & \text{for } x \in \mathbb{R}^n, \\ u(x,t) \le M e^{a|x|^2} & \text{on } \mathbb{R}^n \times [0,T] \end{cases}$$

for some constants M, $a > 0$. *Then*

$$u(x,t) \le \sup_{z \in \mathbb{R}^n} g(z) \quad \forall\, x \in \mathbb{R}^n,\ 0 \le t \le T.$$

Proof. 1. Without loss of generality, assume that $\sup_{\mathbb{R}^n} g < \infty$. We first assume that $4aT < 1$ and $\epsilon > 0$ and $\gamma > 0$ are two numbers such that $4(a+\gamma)(T+\epsilon) = 1$. Let $\mu > 0$ and consider the function

$$v(x,t) = u(x,t) - \mu w(x,t), \quad (x \in \mathbb{R}^n,\ 0 \le t \le T),$$

where $w(x,t)$ is the function defined by

$$w(x,t) = \frac{1}{(T+\epsilon-t)^{n/2}} e^{\frac{|x|^2}{4(T+\epsilon-t)}}$$

A direct calculation shows that $w(x,t)$ satisfies the heat equation $w_t = \Delta w$ on $\mathbb{R}^n \times [0,T]$, and hence

$$v_t - \Delta v \le 0 \quad \text{in } \mathbb{R}^n \times (0,T].$$

2. Clearly,

$$v(x,0) < u(x,0) = g(x) \le \sup_{\mathbb{R}^n} g \quad \forall\, x \in \mathbb{R}^n.$$

For $r > 0$, if $|x| = r$ and $0 \le t \le T$, then

$$v(x,t) = u(x,t) - \mu w(x,t) = u(x,t) - \frac{\mu}{(T+\epsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\epsilon-t)}}$$

$$\le M e^{ar^2} - \frac{\mu}{(T+\epsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\epsilon-t)}} \le M e^{ar^2} - \frac{\mu}{(T+\epsilon)^{n/2}} e^{\frac{r^2}{4(T+\epsilon)}}$$

$$:= M e^{ar^2} - \mu(4(a+\gamma))^{n/2} e^{(a+\gamma)r^2}.$$

Hence we can choose a number $R > 0$ sufficiently large (depending on μ) so that

$$M e^{ar^2} - \mu(4(a+\gamma))^{n/2} e^{(a+\gamma)r^2} \le \sup_{\mathbb{R}^n} g \quad \forall\, r \ge R.$$

Therefore, we have that

$$v(x,t) \le \sup_{\mathbb{R}^n} g \quad \forall\, |x| \ge R,\ 0 \le t \le T, \tag{4.10}$$

and thus that $v(x,t) \le \sup_{\mathbb{R}^n} g$ for all $(x,t) \in \partial'\Omega_T^R$, where $\Omega^R = \{x \in \mathbb{R}^n : |x| < R\}$. So, applying the weak maximum principle to the circular cylinder $\bar{\Omega}_T^R$, we have

$$v(x,t) \le \sup_{\mathbb{R}^n} g \quad \forall\, |x| \le R,\ 0 \le t \le T. \tag{4.11}$$

Consequently, combining (4.10) and (4.11), we have

$$v(x,t) \le \sup_{\mathbb{R}^n} g \quad \forall\, x \in \mathbb{R}^n,\ \ 0 \le t \le T;$$

this implies that

$$u(x,t) \le \sup_{\mathbb{R}^n} g + \mu w(x,t) \quad \forall\, x \in \mathbb{R}^n,\ 0 \le t \le T.$$

Letting $\mu \to 0^+$, we obtain that $u(x,t) \le \sup_{\mathbb{R}^n} g$ for all $x \in \mathbb{R}^n$ and $0 \le t \le T$.

3. Finally, if $4aT \ge 1$, we repeatedly apply the result on intervals $[0,T_1]$, $[T_1, 2T_1], \cdots$, where $T_1 = \frac{1}{8a}$, until we reach to time T. □

4.3.2. Uniqueness under growth condition. The following uniqueness result is an immediate consequence of the previous theorem.

Theorem 4.12 (Uniqueness under growth condition)**.** *Given any f, g, there exists at most one solution $u \in C^2_1(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ to the Cauchy problem*

$$\begin{cases} u_t - \Delta u = f(x,t) & \text{in } \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & \text{on } \mathbb{R}^n \end{cases}$$

that satisfies the growth condition $|u(x,t)| \le Me^{a|x|^2}$ in $\mathbb{R}^n \times (0,T)$ for some constants M, $a > 0$.

4.3.3. Nonnegative solutions. From the example of Tychonoff's solutions above, we know that the initial data can not determine the solution uniquely and some additional information is needed for uniqueness; for example, the suitable growth condition at infinity or the boundary conditions when the domain Ω is bounded.

We now discuss another important uniqueness result due to D.V. Widder concerning the *nonnegative solutions* of the heat equation.

Theorem 4.13 (Widder's theorem). *Let u be continuous for $x \in \mathbb{R}$, $0 \le t < T$, and let u_t, u_x and u_{xx} exist and be continuous for $x \in \mathbb{R}$, $0 < t < T$. Assume that*

$$u_t = u_{xx}, \quad u(x,0) = g(x), \quad u(x,t) \ge 0.$$

Then u is determined uniquely and represented by

$$u(x,t) = \int_{\mathbb{R}} K(x,y,t) g(y)\,dy.$$

Proof. The idea is to show that (i) $u(x,t) \ge \int_{\mathbb{R}} K(x,y,t)g(y)dy$ (the representation formula gives the smallest nonnegative solution), and (ii) $w(x,t) = u(x,t) - \int_{\mathbb{R}} K(x,y,t)g(y)dy$ must be identically zero.

1. For $a > 1$ define function $\zeta^a(x)$ by $\zeta^a(x) = 1$ for $|x| \le a-1$; $\zeta^a(x) = 0$ for $|x| \ge a$; $\zeta^a(x) = a - |x|$ for $a - 1 < |x| < a$. Consider the expression

$$v^a(x,t) = \int_{\mathbb{R}} K(x,y,t)\zeta^a(y)g(y)dy.$$

Since $\zeta^a(y)g(y)$ is bounded and continuous on $\mathbb{R}$, we know that

$$v^a_t - v^a_{xx} = 0 \ \text{ for } x \in \mathbb{R},\ 0 < t < T, \quad v^a(x,0) = \zeta^a(x)g(x).$$

Let M_a be the maximum of $g(x)$ on $|x| \le a$. Using the inequality (**Prove it!**)

$$K(x,y,t) \le \frac{1}{\sqrt{2\pi e}|x-y|} \quad \forall\, x,\, y \in \mathbb{R},\ x \ne y,\ \ t > 0,$$

we have that, for $|x| > a$,

$$0 \le v^a(x,t) \le M_a \int_{-a}^{a} K(x,y,t)dy \le \frac{2M_a a}{\sqrt{2\pi e}} \frac{1}{|x| - a}.$$

Let $\epsilon > 0$ and let $\rho > a + \frac{2M_a a}{\epsilon\sqrt{2\pi e}}$. Then

$$\begin{cases} v^a(x,t) \le \epsilon \le \epsilon + u(x,t), & |x| = \rho,\ 0 < t < T, \\ v^a(x,0) \le g(x) \le \epsilon + u(x,0), & |x| \le \rho. \end{cases}$$

By the weak maximum principle,

$$v^a(x,t) \le \epsilon + u(x,t) \quad \text{for } |x| \le \rho,\ 0 \le t < T.$$

Let $\rho \to \infty$ and we find the same inequality for all $x \in \mathbb{R}$, $0 \le t < T$. Letting $\epsilon \to 0^+$, it follows that

$$v^a(x,t) \le u(x,t), \quad \forall\, x \in \mathbb{R},\ 0 \le t < T.$$

Since ζ^a is a non-decreasing bounded functions of a we find that

$$v(x,t) := \lim_{a\to\infty} v^a(x,t) = \int_{\mathbb{R}} K(x,y,t)g(y)dy$$

exists for $x \in \mathbb{R}$, $0 \le t < T$ and $0 \le v(x,t) \le u(x,t)$. Furthermore, as $v^a(x,t) \le v(x,t) \le u(x,t)$ for $x \in \mathbb{R}$, $0 \le t < T$ it also follows that $v(x,t)$ is continuous on $\mathbb{R} \times [0,T)$ with $v(x,0) = g(x)$.

2. Let $w = u - v$, which is continuous for $x \in \mathbb{R}$, $0 \le t < T$, $w_t = w_{xx}$, $w(x,t) \ge 0$ and $w(x,0) = 0$. It remains to show that $w \equiv 0$. (That is, we reduced the problem to the case $g = 0$.) We introduce a new function

$$W(x,t) = \int_0^t w(x,s)ds.$$

Clearly $W_t = w$ and W are nonnegative continuous on $\mathbb{R} \times [0,T)$ with $W(x,0) = 0$. The existence of W_x, W_{xx} is not obvious since $w_x(x,s)$ may not exist for $s = 0$. To prove the existence of W_x, W_{xx}, we use the difference quotient operator

$$\delta_h f(x,t) = \frac{f(x+h,t) - f(x,t)}{h} \quad (h \ne 0).$$

Note that if f_x exists then $\delta_h f(x,t) = \int_0^1 f_x(x+zh,t)\,dz$. Hence

$$\begin{aligned} \delta_H w(x,t) - \delta_h w(x,t) &= \int_0^1 (w_x(x+zH,t) - w_x(x+zh,t))\,dz \\ &= \int_0^1 \int_h^H w_{xx}(x+zp,t)z\,dpdz = \int_0^1 \int_h^H w_t(x+zp,t)z\,dpdz, \end{aligned}$$

and so by integration with respect to t it follows that for any $\varepsilon \in (0,t)$

$$\begin{aligned} \delta_H W(x,t) - \delta_h W(x,t) &= \delta_H W(x,\varepsilon) - \delta_h W(x,\varepsilon) \\ &\quad + \int_0^1 \int_h^H (w(x+zp,t) - w(x+zp,\varepsilon))z\,dpdz. \end{aligned}$$

Letting $\varepsilon \to 0^+$, since W and w are continuous at $t = 0$, gives

$$\delta_H W(x,t) - \delta_h W(x,t) = \int_0^1 \int_h^H w(x+zp,t)z\,dpdz. \tag{4.12}$$

This proves that $W_x(x,t) = \lim_{h\to 0} \delta_h W(x,t)$ exists. Letting $h \to 0^+$ in (4.12) and rewriting, we obtain

$$W(x+H,t) = W(x,t) + HW_x(x,t) + \int_0^1 \int_0^H Hw(x+zp,t)z\,dpdz,$$

and thus $W(x+H,t)$ is twice differentiable with respect to H and hence is also twice differentiable with respect to x; moreover,

$$W_{xx}(x+H,t) = W_{HH}(x+H,t) = \int_0^1 (2zw(x+zH,t) + Hz^2 w_x(x+zH,t))\,dz.$$

For $H \to 0^+$ we find that $W_{xx}(x,t) = w(x,t) = W_t(x,t) \geq 0$. So $W(x,t)$ is convex in x; hence

$$2W(x,t) \leq W(x+H,t) + W(x-H,t)$$

for any $H > 0$. Integrating the inequality with respect to H from 0 to $x > 0$ we find

$$2xW(x,t) \leq \int_0^x W(x+H,t)dH + \int_0^x W(x-H,t)dH = \int_0^{2x} W(y,t)dy.$$

From Step 1, for all $t > s,\ x > 0$,

$$W(0,t) \geq \int_{\mathbb{R}} K(0,y,t-s)W(y,s)dy \geq \int_0^{2x} K(0,y,t-s)W(y,s)dy$$
$$\geq \frac{e^{-x^2/(t-s)}}{\sqrt{4\pi(t-s)}} \int_0^{2x} W(y,s)dy.$$

The similar argument can be carried out to the case $x < 0$ (or applied to $W(-x,t)$). Therefore we deduce that

$$W(x,s) \leq \sqrt{\frac{\pi(t-s)}{x^2}}\, e^{x^2/(t-s)} W(0,t) \quad (x \in \mathbb{R},\ t > s \geq 0). \tag{4.13}$$

3. Now, for each $\epsilon \in (0, T/2)$, consider $W(x,s)$ for $x \in \mathbb{R}$ and $0 \leq s \leq T - 2\epsilon$. Then $W(x,s)$ is bounded for $|x| \leq \sqrt{\pi T}$ and $0 \leq s \leq T - 2\epsilon$. Using (4.13) with $t = T - \epsilon$, we obtain that, for $|x| \geq \sqrt{\pi T}$ and $0 \leq s \leq T - 2\epsilon$,

$$W(x,s) \leq e^{x^2/\epsilon} W(0, T-\epsilon).$$

Hence W satisfies the assumption of Theorem (4.11). It follows that $W(x,s) = 0$ for $x \in \mathbb{R}^n$ and $s \in [0, T-2\epsilon]$. Since $\epsilon > 0$ is arbitrary, $W(x,s) = 0$ for $s \in [0,T)$; that is $u(x,t) = v(x,t)$. ☐

4.4. Regularity for the heat equation

We now establish the regularity of solution for the heat equation.

Theorem 4.14 (Smoothness)**.** *Let $u \in C_1^2(\Omega_T)$ solve the heat equation in Ω_T. Then $u \in C^\infty(\Omega_T)$.*

Proof. 1. Consider the closed circular cylinder (called a **parabolic cylinder**)

$$C(x,t;r) = \{(y,s) \mid |y-x| \leq r,\ t - r^2 \leq s \leq t\}.$$

Fix $(x_0, t_0) \in \Omega_T$. Choose $r > 0$ small enough so that $C = C(x_0, t_0; r) \subset \Omega_T$. Consider two smaller cylinders

$$C' = C(x_0, t_0; \frac{3}{4}r), \quad C'' = C(x_0, t_0; \frac{1}{2}r).$$

Let $\phi \in C^\infty(\mathbb{R}^n)$ and $\psi \in C^\infty([0,t_0])$ be such that $0 \leq \phi \leq 1$, $\phi = 0$ on $\mathbb{R}^n \setminus B(x_0, r)$, $\phi = 1$ on $B(x_0, \frac{3}{4}r)$, $0 \leq \psi \leq 1$, $\psi(t) = 1$ for $t_0 - (\frac{3}{4}r)^2 \leq t \leq t_0$ and $\psi(t) = 0$ for $0 \leq t \leq t_0 - r^2$. Define $\zeta(x,t) = \phi(x)\psi(t)$; then $\zeta \in C^\infty(\mathbb{R}^n \times [0,t_0])$ with $\zeta = 0$ on $(\mathbb{R}^n \times [0,t_0]) \setminus C$, $\zeta = 1$ on C' and $0 \leq \zeta \leq 1$.

2. We temporarily assume that $u \in C^\infty(\Omega_T)$. This seems a cyclic argument to what is to be proved, but the following arguments aim to establishing an identity that is valid for all C^2_1-solutions. Let $v = \zeta u$ on $\mathbb{R}^n \times [0, t_0]$. Then $v \in C^\infty(\mathbb{R}^n \times [0, t_0])$ and $v = 0$ on $\mathbb{R}^n \times \{t = 0\}$. Furthermore,

$$v_t - \Delta v = \zeta_t u - 2D\zeta \cdot Du - u\Delta\zeta := \tilde{f} \quad \text{in } \mathbb{R}^n \times [0, t_0].$$

Note that $\tilde{f}$ is C^∞ and equals zero on $(\mathbb{R}^n \times [0, t_0]) \setminus C$; moreover, v is bounded on $\mathbb{R}^n \times [0, t_0]$. Hence, by the uniqueness of bounded solution, we have

$$v(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)\, dyds \quad (x \in \mathbb{R}^n,\ t \in (0, t_0]).$$

Given any $(x,t) \in C''$, since $u(x,t) = v(x,t)$, it follows that

$$\begin{aligned} u(x,t) &= \iint_C \Phi(x-y, t-s)\tilde{f}(y,s)\, dyds \\ &= \iint_C \Phi(x-y, t-s)[(\zeta_s - \Delta\zeta)u(y,s) - 2D\zeta \cdot Du]\, dyds \\ &= \iint_C [\Phi(x-y, t-s)(\zeta_s + \Delta\zeta) + 2D_y\Phi(x-y, t-s) \cdot D\zeta]u(y,s)\, dyds \\ &= \iint_{C \setminus C'} [\Phi(x-y, t-s)(\zeta_s + \Delta\zeta) + 2D_y\Phi(x-y, t-s) \cdot D\zeta]u(y,s)\, dyds. \end{aligned}$$

Let $\Gamma(x,y,t,s) = \Phi(x-y, t-s)(\zeta_s + \Delta\zeta)(y,s) + 2D_y\Phi(x-y, t-s) \cdot D\zeta(y,s)$; then

$$u(x,t) = \iint_{C \setminus C'} \Gamma(x,y,t,s)u(y,s)\, dyds \quad \forall\, (x,t) \in C''. \tag{4.14}$$

3. We have derived the identity (4.14) under the assumption that u is a C^∞-solution of the heat equation in Ω_T. Using the standard mollifying technique, this formula is valid for all $C^2_1(\Omega_T)$-solution of the heat equation. From this formula, since $\Gamma(x,y,t,s)$ is C^∞ in $(x,t) \in C''$ for all $(y,s) \in C \setminus C'$, we deduce that $u \in C^\infty(C'')$; in particular, u is C^∞ at $(x_0, t_0) \in \Omega_T$. □

Theorem 4.15 (Estimates on derivatives)**.** *There exist constants $C_{k,l}$ for $k, l = 0, 1, 2, \cdots$ such that*

$$\max_{C(x,t;\frac{r}{2})} |D^\alpha_x D^l_t u| \le \frac{C_{k,l}}{r^{k+2l+n+2}} \|u\|_{L^1(C(x,t;r))} \quad (|\alpha| = k)$$

for all cylinders $C(x,t;r) \subset \Omega_T$ and all $C^2_1(\Omega_T)$-solutions u of the heat equation in Ω_T.

Proof. 1. Fix some point $(x_0, t_0) \in \Omega_T$. Upon shifting the coordinates we may assume the point is $(0,0)$. Suppose first that the cylinder $C(1) = C(0,0;1) \subset \Omega_T$; then by (4.14) above,

$$u(x,t) = \iint_{C(1) \setminus C(\frac{3}{4})} \Gamma(x,y,t,s)u(y,s)\, dyds \quad \forall\, (x,t) \in C(\frac{1}{2}).$$

Consequently, for all $k, l = 0, 1, 2, \cdots$ and all α with $|\alpha| = k$,

$$|D^\alpha_x D^l_t u(x,t)| \le \iint_{C(1) \setminus C(\frac{3}{4})} |D^\alpha_x D^l_t \Gamma(x,y,t,s)||u(y,s)|\, dyds \le C_{k,l}\|u\|_{L^1(C(1))}, \tag{4.15}$$

for all $(x,t) \in C(\frac{1}{2})$, where $C_{k,l}$ are some constants.

2. Now suppose $C(r) = C(0,0;r) \subset \Omega_T$. We rescale u by defining

$$v(x,t) = u(rx, r^2t) \quad \forall\, (x,t) \in C(1).$$

Then $v_t - \Delta v = 0$ in $C(1)$. Note that for all $|\alpha| = k$,

$$D_x^\alpha D_t^l v(x,t) = r^{2l+k} D_y^\alpha D_s^l u(rx, r^2 t)$$

and $\|v\|_{L^1(C(1))} = \frac{1}{r^{n+2}} \|u\|_{L^1(C(r))}$. Then, with (4.15) applied to v, we deduce

$$|D_y^\alpha D_s^l u(y,s)| \le \frac{C_{k,l}}{r^{k+2l+n+2}} \|u\|_{L^1(C(r))}$$

for all $(y,s) \in C(\frac{r}{2})$. This completes the proof. □

4.5. Mean value property and strong maximum principle for the heat equation

In this section we derive for the heat equation some kind of analogue of the mean value property of harmonic functions.

Definition 4.4. Fix $x \in \mathbb{R}^n,\ t \in \mathbb{R},\ r > 0$. We define the **heat ball** to be

$$E(x,t;r) = \{(y,s) \mid \Phi(x-y, t-s) \ge \frac{1}{r^n}\}.$$

Remark 4.4. The heat ball $E(x,t;r)$ is a bounded region in space-time whose boundary is the level set $\Phi(x-y,t-s) = r^{-n}$. By elementary calculations, it is easy to see that a heat ball is contained in a rescaled parabolic cylinder:

$$E(x,t;r) \subset \left\{ (y,s) \;\middle|\; |y-x| \le \sqrt{\frac{n}{2\pi e}}\, r,\ t - \frac{r^2}{4\pi} \le s \le t \right\}.$$

4.5.1. Mean-value property for the heat equation.

4.5.1. Mean-value property for the heat equation. We have the following mean-value property for *subsolutions* of the heat equation, from which we easily obtain the **mean-value property for the heat equation**.

Theorem 4.16. *Let $u \in C_1^2(\Omega_T)$ satisfy $u_t - \Delta u \le 0$ in Ω_T. Then*

$$u(x,t) \le \frac{1}{4r^n} \iint_{E(x,t;r)} u(y,s) \frac{|x-y|^2}{(t-s)^2} dyds$$

for all heat balls $E(x,t;r) \subset \Omega_T$.

Proof. 1. We may assume that $x = 0, t = 0$ and write $E(r) = E(0,0;r)$. Set

$$\phi(r) = \frac{1}{r^n} \iint_{E(r)} u(y,s) \frac{|y|^2}{s^2} dyds = \iint_{E(1)} u(ry, r^2 s) \frac{|y|^2}{s^2} dyds.$$

Let $v(y,s) = u(ry, r^2 s)$; then $v_s(y,s) \le \Delta v(y,s)$ in $E(1)$. Note also that $D_x u(ry, r^2 s) = \frac{1}{r} Dv(y,s)$ and $u_t(ry, r^2 s) = \frac{1}{r^2} v_s(y,s)$. Let

$$H(y,s) = \Phi(y,-s) = \frac{1}{(-4\pi s)^{n/2}} e^{\frac{|y|^2}{4s}} \quad (y \in \mathbb{R}^n,\ s < 0).$$

Then $H(y,s) = 1$ for $(y,s) \in \partial E(1)$; moreover, $\ln H(y,s) = -\frac{n}{2}\ln(-4\pi s) + \frac{|y|^2}{4s} \ge 0$ for all $(y,s) \in E(1)$, $4(\ln H)_s + \frac{|y|^2}{s^2} = -\frac{2n}{s}$ and $D(\ln H) = D_y(\ln H) = \frac{y}{2s}$.

2. We calculate $\phi'(r)$ as follows.

$$
\begin{aligned}
\phi'(r) =& \iint_{E(1)} \left[D_x u(ry, r^2 s) \cdot y + 2rs u_t(ry, r^2 s)\right] \frac{|y|^2}{s^2}\, dyds \\
=& \frac{1}{r} \iint_{E(1)} \left[y \cdot Dv(y,s) + 2s v_s(y,s)\right] \frac{|y|^2}{s^2}\, dyds \\
=& \frac{1}{r} \iint_{E(1)} \left[y \cdot Dv \,\frac{|y|^2}{s^2} + 4 v_s\, y \cdot D(\ln H)\right] dyds \\
=& \frac{1}{r} \iint_{E(1)} y \cdot Dv \,\frac{|y|^2}{s^2} dyds - \frac{4}{r} \iint_{E(1)} (n v_s + y \cdot (Dv)_s) \ln H\, dyds \\
=& \frac{1}{r} \iint_{E(1)} y \cdot Dv \,\frac{|y|^2}{s^2} dyds - \frac{4}{r} \iint_{E(1)} \left[n v_s \,\ln H - y \cdot Dv\, (\ln H)_s\right] dyds \\
=& -\frac{1}{r}\left[\iint_{E(1)} y \cdot Dv \,\frac{2n}{s} dyds + 4n \iint_{E(1)} v_s\, \ln H dyds\right] \\
\geq& -\frac{1}{r}\left[\int_{E(1)} y \cdot Dv \,\frac{2n}{s} dyds + 4n \iint_{E(1)} \Delta v\, \ln H dyds\right] \\
=& -\frac{2n}{r} \iint_{E(1)} \left[\frac{y}{s} \cdot Dv - 2Dv \cdot D(\ln H)\right] dyds = 0.
\end{aligned}
$$

Consequently, $\phi(r)$ is nondecreasing for $r > 0$; hence

$$
\phi(r) \geq \phi(0^+) = u(0,0) \iint_{E(1)} \frac{|y|^2}{s^2} dyds.
$$

3. It remains to show that $\iint_{E(1)} \frac{|y|^2}{s^2} dyds = 4$. Note that

$$
E(1) = \{(y,s) \mid -\frac{1}{4\pi} \leq s < 0,\ |y|^2 \leq 2ns \ln(-4\pi s)\}.
$$

Using the change of variables $(y,s) \mapsto (y,-\tau)$, where $\tau = -\ln(-4\pi s)$ (and hence $s = -\frac{e^{-\tau}}{4\pi}$), the set $E(1)$ is mapped one-to-one and onto the set $\tilde{E}(1) = \{(y,\tau) \mid \tau \in (0,\infty),\ |y|^2 \leq \frac{n}{2\pi}\tau e^{-\tau}\}$. Therefore, using $ds = \frac{1}{4\pi} e^{-\tau} d\tau$,

$$
\begin{aligned}
\iint_{E(1)} \frac{|y|^2}{s^2} dyds =& \int_0^\infty \int_{|y|^2 \leq \frac{n}{2\pi}\tau e^{-\tau}} 4\pi |y|^2 e^{\tau} dy d\tau \\
=& 4\pi n \alpha_n \int_0^\infty \int_0^{\sqrt{\frac{n}{2\pi}\tau e^{-\tau}}} r^{n+1} e^{\tau} dr d\tau \\
=& \frac{4\pi n\alpha_n}{n+2} \left(\frac{n}{2\pi}\right)^{\frac{n+2}{2}} \int_0^\infty \tau^{\frac{n+2}{2}} e^{-\frac{n}{2}\tau} d\tau \\
=& \frac{4\pi n\alpha_n}{n+2} \left(\frac{n}{2\pi}\right)^{\frac{n+2}{2}} \left(\frac{n}{2}\right)^{-\frac{n}{2}-2} \int_0^\infty t^{\frac{n+2}{2}} e^{-t} dt \quad (\text{using } \tau = \tfrac{2}{n} t) \\
=& \frac{4n\alpha_n}{n+2} \frac{2}{n} \left(\frac{1}{\pi}\right)^{n/2} \Gamma\left(\frac{n}{2}+2\right) \\
=& 2n\alpha_n \left(\frac{1}{\pi}\right)^{n/2} \Gamma\left(\frac{n}{2}\right) = 4,
\end{aligned}
$$

noting that $\alpha_n = \frac{2}{n\Gamma(\frac{n}{2})} \pi^{n/2}$. □

4.5.2. Strong maximum principle.

Theorem 4.17 (Strong Maximum Principle). *Assume that* $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$ *satisfies* $u_t - \Delta u \le 0$ *in* Ω_T.

(i) *We have that* $\max_{\overline{\Omega_T}} u = \max_{\partial' \Omega_T} u$.

(ii) *If* Ω *is connected and there exists a point* $(x_0, t_0) \in \Omega_T$ *such that* $u(x_0, t_0) = \max_{\overline{\Omega_T}} u$, *then* u *is constant in* $\overline{\Omega_{t_0}}$.

Proof. 1. The part (i) is just the weak maximum principle for the heat equation; it also follows from part (ii). (**Explain why?**) So we only prove (ii).

2. Let $M = u(x_0, t_0) = \max_{\overline{\Omega_T}} u$. Then for all sufficiently small $r > 0$, $E(x_0, t_0; r) \subset \Omega_T$, and we employ the mean value theorem to get

$$M = u(x_0, t_0) \le \frac{1}{4r^n} \int_{E(x_0,t_0;r)} u(y,s) \frac{|x_0 - y|^2}{(t_0 - s)^2} dyds \le \frac{M}{4r^n} \int_{E(x_0,t_0;r)} \frac{|x_0 - y|^2}{(t_0 - s)^2} dyds = M.$$

Therefore, $u(y, s) \equiv M$ on $E(x_0, t_0; r)$. Draw any line segment L in Ω_T connecting (x_0, t_0) to some point $(y_0, s_0) \in \Omega_T$ with $s_0 < t_0$. Consider

$$r_0 = \inf\{s \in [s_0, t_0] \,|\, u(x, t) = M \quad \forall\, (x, t) \in L,\ s \le t \le t_0\}.$$

Since u is continuous, the infimum is attained. We claim $r_0 = s_0$ and hence $u \equiv M$ on L. Suppose, for the contrary, that $r_0 > s_0$. Then $u(z_0, r_0) = M$ for some point $(z_0, r_0) \in L$. From the previous argument, $u(x, t) = M$ on $E(z_0, r_0; r)$ for all sufficiently small r. Note that $E(z_0, r_0; r)$ contains $L \cap \{r_0 - \sigma < t \le r_0\}$ for some $\sigma > 0$. We obtain a contradiction.

3. Now fix any $x \in \Omega$ and $t \in (0, t_0)$. We show $u(x, t) = M$. Since Ω is connected, so is Ω_T; hence, there exists a piece-wise continuous line segments $\{L_i\}$ connecting points (x_0, t_0) to (x, t) in such a way that the t-coordinates of the endpoints of L_i are decreasing. By Step 2, we know $u \equiv M$ on each L_i and hence $u(x, t) = M$. Consequently, $u \equiv M$ in Ω_{t_0}. □

Remark 4.5. (a) If a solution u to the heat equation attains its maximum (or minimum) at an interior point (x_0, t_0) then u is constant at all *earlier times* $t \le t_0$. However, the solution may change values at the later times $t > t_0$ if, for example, the boundary conditions alter after t_0.

(b) Suppose that u solves the heat equation in Ω_T and equals zero on $\partial\Omega \times [0, T]$. If the initial data $u(x, 0) = g(x)$ is nonnegative and is positive somewhere, then u is positive everywhere within Ω_T. This is another illustration of **infinite speed of propagation** of disturbances for the heat equation.

4.6. Harnack's inequality for second-order linear parabolic equations

Assume Ω is a bounded open set in $\mathbb{R}^n$. Let

$$(\partial_t + L)u = u_t + Lu = u_t - \sum_{i,j=1}^n a^{ij}(x,t) D_{ij} u + \sum_{i=1}^n b^i(x,t) D_i u + c(x,t) u, \tag{4.16}$$

where a^{ij}, b^i and c are smooth functions on $\overline{\Omega_T}$, with $a^{ij} = a^{ji}$ satisfying, for a constant $\theta > 0$,

$$\sum_{i,j=1}^n a^{ij}(x,t)\xi_i \xi_j \ge \theta |\xi|^2 \quad \forall\, (x,t) \in \Omega_T,\ \xi \in \mathbb{R}^n.$$

4.6.1. Harnack's inequality.

Theorem 4.18. *Let $V \subset\subset \Omega$ be connected and $0 < t_1 < t_2 \le T$. Then, there is a constant C depending only on V, t_1, t_2 and the coefficients of L such that*

$$\sup_V u(.,t_1) \le C \inf_V u(.,t_2)$$

for all nonnegative smooth solutions u of $u_t + Lu = 0$ in Ω_T.

Proof. Without loss of generality, assume V is a ball since in general V can be covered by finitely many balls. Let $\varepsilon > 0$; the idea is to show that there is some $C > 0$, depending only on V, t_1, t_2 and the coefficients of L but independent of ε, such that

$$v(x_2,t_2) - v(x_1,t_1) \ge -C \quad (\forall\, x_1,\, x_2 \in V), \tag{4.17}$$

where $v = \ln(u+\varepsilon)$ and u is any nonnegative smooth solution of $u_t + Lu = 0$ in Ω_T. Once this is proved, we have $u(x_1,t_1) + \varepsilon \le e^C(u(x_2,t_2)+\varepsilon)$ and so letting $\varepsilon \to 0^+$ gives the desired inequality

$$\sup_V u(.,t_1) \le e^C \inf_V u(.,t_2).$$

Now that

$$\begin{aligned} v(x_2,t_2) - v(x_1,t_1) &= \int_0^1 \frac{d}{ds} v(sx_2 + (1-s)x_1, st_2 + (1-s)t_1)ds \\ &= \int_0^1 [(x_2-x_1)\cdot Dv + (t_2-t_1)v_t]ds, \end{aligned}$$

and so, to show (4.17), it is enough to show that

$$v_t \ge \nu |Dv|^2 - C \quad \text{ in } V \times [t_1,t_2], \tag{4.18}$$

for some constants $\nu > 0$ and $C > 0$. A direct computation shows that

$$D_j v = \frac{D_j u}{u+\varepsilon}, \quad D_{ij} v = \frac{D_{ij} u}{u+\varepsilon} - D_i v D_j v,$$

and so, using $u_t = \sum_{ij} a^{ij} D_{ij} u - \sum_i b^i D_i u - cu$, we have

$$\begin{aligned} v_t = \frac{u_t}{u+\varepsilon} &= \sum_{ij} a^{ij} D_{ij} v + \sum_{ij} a^{ij} D_i v D_j v - \sum_i b^i D_i v - c\frac{u}{u+\varepsilon} \\ &:= \alpha + \beta + \gamma + f, \end{aligned}$$

where

$$\alpha = \sum_{ij} a^{ij} D_{ij} v, \;\; \beta = \sum_{ij} a^{ij} D_i v D_j v, \;\; \gamma = -\sum_i b^i D_i v, \;\; f = -\frac{cu}{u+\varepsilon}.$$

Note that $\beta \ge \theta |Dv|^2$ and

$$|f| \le C, \quad |Df| \le C(1+|Dv|), \quad |D^2 f| \le C(1 + |D^2 v| + |Dv|^2),$$

where constant $C > 0$ depends only on $c(x,t)$. Therefore, to establish (4.18), it suffices to show that

$$w := \alpha + \kappa\beta + \gamma \ge -C \quad \text{on } V \times [t_1,t_2] \tag{4.19}$$

for some constants $C > 0$ and $\kappa \in (0, 1/2)$ depending only on V, t_1, t_2 and the coefficients of L. To this end, we calculate

$$\begin{aligned} D_k\alpha &= \sum_{ij} a^{ij} D_{ijk}v + \sum_{ij} D_k a^{ij} D_{ij}v := \sum_{ij} a^{ij} D_{ijk}v + R_1, \\ D_k\beta &= 2\sum_{ij} a^{ij} D_{ik}vD_jv + \sum_{ij} D_k a^{ij} D_iv D_jv := 2\sum_{ij} a^{ij} D_{ik}vD_jv + R_2, \\ D_k\gamma &= -\sum_i b^i D_{ik}v - \sum_i D_k b^i D_iv := -\sum_i b^i D_{ik}v + R_3, \end{aligned}$$

where $|R_1| \le C|D^2v|,\ |R_2| \le C|Dv|^2,\ |R_3| \le C|Dv|$ with some constant $C > 0$ depending only on a^{ij} and b^i. In the following, if not specifically stated, R_k or $\tilde{R}_k$ will always denote a term satisfying

$$|R_k|,\ |\tilde{R}_k| \le C(|D^2v||Dv| + |D^2v| + |Dv|^2 + |Dv| + 1)$$

with a constant $C > 0$ depending only on the coefficients of L. Below we will also use C or C_k to denote a constant that depends only on the coefficients of L and $\theta > 0$ but could be different in different estimates. Note that

$$D_{kl}\beta = 2\sum_{ij} a^{ij} D_{ikl}vD_jv + 2\sum_{ij} a^{ij} D_{ik}vD_{jl}v + R_4,$$

and hence

$$\begin{aligned} \sum_{kl} a^{kl} D_{kl}\beta &= 2\sum_{ijkl} a^{ij}a^{kl} D_{ikl}vD_jv + 2\sum_{ijkl} a^{ij}a^{kl} D_{ik}vD_{jl}v + \tilde{R}_4 \\ &= 2\sum_{ij} a^{ij} D_jv(D_i\alpha - R_1) + 2\sum_{ijkl} a^{ij}a^{kl} D_{ik}vD_{jl}v + \tilde{R}_4 \\ &= \sum_i \tilde{b}^i D_i\alpha + 2\sum_{ijkl} a^{ij}a^{kl} D_{ik}vD_{jl}v + \tilde{R}_5, \end{aligned}$$

where $\tilde{b}^i = 2\sum_j a^{ij}D_jv$. Therefore, from the definition of α, using $v_t = \alpha + \beta + \gamma + f$, we have that

$$\begin{aligned} \alpha_t &= \sum_{kl} a^{kl} D_{kl}v_t + \sum_{kl} a^{kl}_t D_{kl}v \\ &= \sum_{kl} a^{kl} D_{kl}(\alpha + \gamma) + \sum_{kl} a^{kl} D_{kl}\beta + \sum_{kl} a^{kl} D_{kl}f + \sum_{kl} a^{kl}_t D_{kl}v \\ &= \sum_{kl} a^{kl} D_{kl}(\alpha + \gamma) + \sum_i \tilde{b}^i D_i\alpha + 2\sum_{ijkl} a^{ij}a^{kl} D_{ik}vD_{jl}v + R_6 \\ &\ge \sum_{kl} a^{kl} D_{kl}(\alpha + \gamma) + \sum_i \tilde{b}^i D_i\alpha + 2\theta^2|D^2v|^2 + R_6 \\ &\ge \sum_{kl} a^{kl} D_{kl}(\alpha + \gamma) + \sum_i \tilde{b}^i D_i\alpha + \theta^2|D^2v|^2 - C|Dv|^2 - C, \end{aligned} \tag{4.20}$$

where we have used the fact that, for all $n \times n$-symmetric matrices A, B,

$$\operatorname{tr}(ABAB) \ge \theta^2|B|^2 \quad \text{if } A \ge \theta I_n.$$

Similarly, we compute

$$
\begin{aligned}
\beta_t &= 2\sum_{kl} a^{kl} D_k v D_l v_t + \sum_{kl} a_t^{kl} D_k v D_l v \\
&= 2\sum_{kl} a^{kl} D_k v D_l(\alpha+\beta+\gamma+f) + \sum_{kl} a_t^{kl} D_k v D_l v \\
&= \sum_l \tilde{b}^l D_l\alpha + \sum_l \tilde{b}^l D_l\beta + 2\sum_{kl} a^{kl} D_k v D_l(\gamma+f) + \sum_{kl} a_t^{kl} D_k v D_l v \\
&= \sum_{kl} a^{kl} D_{kl}\beta - 2\sum_{ijkl} a^{ij}a^{kl} D_{ik}v D_{jl}v - \tilde{R}_5 + \sum_l \tilde{b}^l D_l\beta + R_7 \\
&\geq \sum_{kl} a^{kl} D_{kl}\beta + \sum_l \tilde{b}^l D_l\beta - C(|D^2v|^2 + |Dv|^2 + 1),
\end{aligned}
\tag{4.21}
$$

and

$$
\begin{aligned}
\gamma_t &= -\sum_k b^k D_k v_t - \sum_k b_t^k D_k v = -\sum_k b^k D_k(\alpha+\beta+\gamma+f) - \sum_k b_t^k D_k v \\
&= \sum_k \tilde{b}^k D_k\gamma + R_8.
\end{aligned}
$$

Hence, for $w = \alpha + \kappa\beta + \gamma$, with $\kappa \in (0, 1/2)$ sufficiently small, we have

$$
w_t - \sum_{kl} a^{kl} D_{kl} w - \sum_k \tilde{b}^k D_k w \geq \frac{\theta^2}{2}|D^2v|^2 - C|Dv|^2 - C. \tag{4.22}
$$

Lemma 4.19. *Let $\zeta \in C^\infty(\Omega_T)$ be a cutoff function such that $0 \leq \zeta \leq 1$ and $\zeta \equiv 1$ on $V \times [t_1, t_2]$ and $\zeta = 0$ on $\partial'\Omega_T$. Then there is a number $\mu > 0$ depending only on V, t_1, t_2 and the coefficients of L such that $H(x,t) = \zeta^4 w + \mu t \geq 0$ in Ω_T.*

Proof. Suppose that H is negative somewhere on Ω_T. Then, since $H = 0$ on $\partial'\Omega_T$, the negative minimum of H on $\overline{\Omega_T}$ must be attained at some point $(x_0, t_0) \in \Omega_T$. At this point (x_0, t_0), we must have that $\zeta > 0$ and $D_k H = \zeta^3(4wD_k\zeta + \zeta D_k w) = 0$, and so $\zeta D_k w = -4wD_k\zeta$; moreover, we have

$$
0 \geq H_t - \sum_{kl} a^{kl} D_{kl} H - \sum_k \tilde{b}^k D_k H
$$

at (x_0, t_0). Using $D_{kl}(\zeta^4 w) = \zeta^4 D_{kl} w + D_l(\zeta^4) D_k w + D_k(\zeta^4) D_l w + w D_{kl}(\zeta^4)$, we have

$$
\begin{aligned}
0 \geq& \mu + 4\zeta^3\zeta_t w + \zeta^4\Big(w_t - \sum_{kl} a^{kl} D_{kl} w - \sum_k \tilde{b}^k D_k w\Big) \\
&\quad - \sum_k \tilde{b}^k w D_k(\zeta^4) - 2\sum_{kl} a^{kl} D_k w D_l(\zeta^4) - \sum_{kl} a^{kl} w D_{kl}(\zeta^4) \\
=& \mu + \zeta^4\Big(w_t - \sum_{kl} a^{kl} D_{kl} w - \sum_k \tilde{b}^k D_k w\Big) + M \\
\geq& \mu + \zeta^4\Big(\frac{\theta^2}{2}|D^2v|^2 - C|Dv|^2 - C\Big) + M
\end{aligned}
\tag{4.23}
$$

at (x_0, t_0), in views of (4.22). Since $|\tilde{b}^k| \leq C|Dv|$ and $\zeta D_k w = -4wD_k\zeta$ and hence $D_l(\zeta^4) D_k w = -16w\zeta^2 D_k\zeta D_l\zeta$ at (x_0, t_0), we have that

$$
|M| \leq C(\zeta^2|w| + \zeta^3|w||Dv|).
$$

At (x_0, t_0), since $H = \zeta^4 w + \mu t < 0$, we have $w = \alpha + \kappa\beta + \gamma < 0$ and thus $\kappa\beta \leq -\alpha - \gamma$. Since $\beta \geq \theta|Dv|^2$, we have, at (x_0, t_0), that $|Dv|^2 \leq C(|D^2v| + |Dv|)$; so the inequalities

$$|Dv|^2 \leq C(1 + |D^2v|), \quad |w| \leq C(1 + |D^2v|), \\ |w||Dv| \leq C(1 + |D^2v|)^{3/2} \leq C[2^{3/2}|D^2v|^{3/2} + 2^{3/2}] \tag{4.24}$$

hold at (x_0, t_0). Consequently, at (x_0, t_0),

$$|M| \leq C(\zeta^2|D^2v| + \zeta^3|D^2v|^{3/2} + 1) \leq \delta\zeta^4|D^2v|^2 + C_\delta,$$

for each $\delta > 0$; here we have used **Young's inequality** with $\delta > 0$. Now let $\delta = \frac{\theta^2}{4}$; then by (4.23) and (4.24) we have

$$\begin{aligned} 0 &\geq \mu + \zeta^4\Big(\frac{\theta^2}{4}|D^2v|^2 - C|Dv|^2 - C\Big) - C_1 \\ &\geq \mu + \zeta^4\Big(\frac{\theta^2}{4}|D^2v|^2 - C_2|D^2v| - C_3\Big) - C_1 \\ &= \mu + \zeta^4\Big(\frac{\theta}{2}|D^2v| - \frac{C_2}{\theta}\Big)^2 - \zeta^4\Big(\frac{C_2^2}{\theta^2} + C_3\Big) - C_1 \end{aligned}$$

at (x_0, t_0), which implies

$$\mu \leq \zeta^4\Big(\frac{C_2^2}{\theta^2} + C_3\Big) + C_1 \leq \frac{C_2^2}{\theta^2} + C_3 + C_1.$$

Therefore, if $\mu = \frac{C_2^2}{\theta^2} + C_3 + C_1 + 1$, then $H(x,t) = \zeta^4 w + \mu t \geq 0$ in Ω_T. □

To complete the proof of our main theorem, let $\mu > 0$ be a constant determined in the lemma above. Since $H = w + \mu t \geq 0$ in $V \times [t_1, t_2]$, we have that $w \geq -\mu t_2 = -C$ in $V \times [t_1, t_2]$. This proves (4.19) and hence completes the proof. □

4.6.2. Elliptic Harnack's inequality. Using Harnack's inequality for the parabolic equation we can derive Harnack's inequality for elliptic equations since when the coefficients of L depend only on x, a solution to elliptic equation $Lu = 0$ can be thought as a steady-state solution of the parabolic equation $u_t + Lu = 0$.

Corollary 4.20 (**Harnack's inequality for elliptic equations**). *Given any connected open bounded subset $V \subset\subset \Omega$, there exists a constant $C(L, V) > 0$ such that*

$$\sup_{x \in V} u(x) \leq C(L, V) \inf_{x \in V} u(x)$$

for all nonnegative solutions $u \in C^2(\Omega)$ to a uniformly elliptic equation $Lu = 0$ in Ω.

4.6.3. Strong maximum principle.

Theorem 4.21 (**Strong maximum principle**). *Let $\partial_t + L$ be as given in* (4.16) *and Ω be open, bounded and connected. Assume $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega_T})$, $u_t + Lu \leq 0$ in Ω_T and $u(x_0, t_0) = M = \max_{\overline{\Omega_T}} u$ at some point $(x_0, t_0) \in \Omega_T$. Then*

(i) *$u \equiv M$ in Ω_{t_0} if $c \equiv 0$.*
(ii) *$u \equiv 0$ in Ω_{t_0} if $M = 0$.*
(iii) *$u \equiv M$ in Ω_{t_0} if $c \geq 0$ and $M \geq 0$.*

Proof. 1. Let $c \equiv 0$. Select any smooth open set $W \subset\subset \Omega$ with $x_0 \in W$. Let v be the solution of $v_t + Lv = 0$ in W_T and $v|_{\partial' W_T} = u|_{\partial' W_T}$. (Here and below we assume the existence of such a solution.) Then by the weak maximum principle, we have $u(x,t) \le v(x,t) \le M$ for all $(x,t) \in W_T$. From this we deduce that $v(x_0,t_0) = u(x_0,t_0) = M$ is maximum of v. Let $w = M - v(x,t)$. Then $w_t + Lw = 0$ and $w(x,t) \ge 0$ in W_T. Choose any connected set $V \subset\subset W$ with $x_0 \in V$. Let $0 < t < t_0$. Using Harnack's inequality we have a constant $C = C(L, V, t, t_0)$ such that $0 \le w(x,t) \le C \inf_{y \in V} w(y,t_0) = Cw(x_0,t_0) = 0$, which implies that $w \equiv 0$ on V_{t_0}. Since V is arbitrary, this implies $w \equiv 0$ in W_{t_0}; but then $v \equiv M$ in W_{t_0}. Since $v = u$ on $\partial' W_T$, we have $u \equiv M$ on $\partial' W_{t_0}$. Since $W \subset\subset \Omega$ is arbitrary, we have $u \equiv M$ in Ω_{t_0}.

2. Let $M = 0$. As in Step 1, select any smooth open set $W \subset\subset \Omega$ with $x_0 \in W$. Let v be the solution of $v_t + Lv = 0$ in W_T and $v|_{\partial' W_T} = u|_{\partial' W_T}$. Then, based on Theorem 4.6, we have $u(x,t) \le v(x,t) \le 0$ for all $(x,t) \in W_T$. So $v(x_0,t_0) = u(x_0,t_0) = 0$ is maximum of v. Let $w = -v(x,t)$. Then $w_t + Lw = 0$ and $w(x,t) \ge 0$ in W_T. As in Step 1, Harnack's inequality implies that $w = -v \equiv 0$ in W_{t_0}; but then $u \equiv 0$ on $\partial' W_{t_0}$. Since $W \subset\subset \Omega$ is arbitrary, we have $u \equiv 0$ in Ω_{t_0}.

3. Let $c \ge 0$ and $M > 0$. Let $\tilde{L}$ be the operator obtained from L by dropping the zeroth term. Let v be the solution of $v_t + \tilde{L}v = 0$ in W_T and $v|_{\partial' W_T} = u^+|_{\partial' W_T}$. Then $0 \le v(x,t) \le M$. Consider the set $U = \{(x,t) \in W_{t_0} \,|\, u(x,t) > 0\}$ (which may not be a parabolic cylinder). Then $\tilde{L}(u - v) \le 0$ in the interior of U and $u - v \le 0$ on $\partial' W_{t_0} \cap \partial U$. As in the proof of weak maximum principle for cylindrical domains, we can prove that $u - v \le 0$ in U. This implies $v(x_0,t_0) = M$. Let $w = M - v(x,t)$. Then $w_t + \tilde{L}w = 0$ and $w(x,t) \ge 0$ in W_{t_0}. As in Step 1, Harnack's inequality implies that $w \equiv 0$ in W_{t_0}; but then $v \equiv M$ in W_{t_0}. Since $v = u^+$ on $\partial' W_T$, we have $u^+ \equiv M$ on $\partial' W_{t_0}$. Since $W \subset\subset \Omega$ is arbitrary, we have $u^+ \equiv M$ in Ω_{t_0}; hence $u \equiv M$ in Ω_{t_0}. □

Remark 4.6. Assume $u \in C_1^2(\Omega_T) \cap C^1(\overline{\Omega_T})$ and $u(x,t) < u(x_0,t_0)$ for all $x \in \Omega,\ 0 < t < t_0$, where $x_0 \in \partial\Omega$ and $0 < t_0 < T$. Assume one of the following conditions holds:

$$\text{(a) } c(x,t) \equiv 0; \quad \text{(b) } u(x_0,t_0) = 0; \quad \text{(c) } c(x,t) \ge 0 \text{ and } u(x_0,t_0) > 0.$$

Then the **parabolic Hopf's inequality** $\dfrac{\partial u}{\partial \nu}(x_0,t_0) > 0$ holds provided the exterior normal derivative exists at (x_0,t_0) and $\partial\Omega$ satisfies the **interior ball condition** at x_0.