Parabolic equations: existence and properties of solutions
Lecture Notes

Swarnendu Sil*

Department of Mathematics
Indian Institute of Science
Bangalore, India

NCM AIS
Dec 4-23 2023, IISc

Contents

*swarnendusil@iisc.ac.in

1 Heat equation

1.1 Representation formulas for IVP

1.1.1 Heat Kernel

Consider the following IVP for the heat equation in n dimensions

$$\begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{1}$$

Taking the Fourier transform with respect to the space variable $x \in \mathbb{R}^n$, for a.e. $\xi \in \mathbb{R}^n$, we obtain the following IVP for an ODE

$$\begin{cases} \dfrac{d}{dt}\hat{u} + |\xi|^2 \hat{u} = 0 & \text{for } t \in (0,\infty), \\ \hat{u}(0,\xi) = \hat{g}(\xi) & \text{for } t = 0. \end{cases}$$

The ODE of course is easily solved. We have

$$\frac{\mathrm{d}\hat{u}}{\hat{u}} = -|\xi|^2 \,\mathrm{d}t.$$

Integrating from $t=0$ to $t=t$, we obtain

$$\int_{\hat{u}(0)}^{\hat{u}(t)} \frac{\mathrm{d}\hat{u}}{\hat{u}} = -|\xi|^2 \int_0^t \mathrm{d}t,$$

This implies

$$\log \frac{\hat{u}(t)}{\hat{u}(0)} = -|\xi|^2 t$$

Thus, we deduce

$$\hat{u} = \hat{g} e^{-t|\xi|^2}.$$

Hence, at least formally, the solution to (1) is given by

$$u = (\hat{u})^{\check{}} = \left(\hat{g} e^{-t|\xi|^2}\right)^{\check{}}.$$

Now suppose we can find a function E such that

$$\hat{E}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-t|\xi|^2}.$$

Then this means we can rewrite our solution as

$$u = \left(\hat{g} e^{-t|\xi|^2}\right)^{\check{}} = \left((2\pi)^{\frac{n}{2}} \hat{E}\hat{g}\right)^{\check{}} = \left([E * g]^{\hat{}}\right)^{\check{}} = E * g.$$

Can we find such a function E? Indeed we can. We just use the Fourier inversion formula and compute

$$E = \left(\hat{E}\right)^{\check{}} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle - t|\xi|^2} \, \mathrm{d}\xi = \frac{1}{(2\pi)^n} \left(\frac{\pi}{t}\right)^{\frac{n}{2}} e^{-\frac{|x|^2}{4t}} = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}.$$

Thus, we deduce,

$$u(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y) \, \mathrm{d}y \qquad \text{for } x \in \mathbb{R}^n, t > 0. \tag{2}$$

So far, what we have done is that we obtained a candidate function u given by (2), for the IVP for the Heat equation. But now we show that we have actually obtained a strong solution.

Theorem 1 (Representation formula for IVP for the homogeneous heat equation)**.** *Let $g \in L^\infty(\mathbb{R}^n) \cap C(\mathbb{R}^n)$ and let u be defined by (2). Then*

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

(ii) u satisfies

$$u_t - \Delta u = 0 \qquad \text{in } \mathbb{R}^n \times (0,\infty),$$

(iii) and we have

$$\lim_{(x,t)\to(x_0,t)} u(x,t) = g(x_0) \qquad \text{for every } x_0 \in \mathbb{R}^n.$$

In other words, u given by (2) is a strong solution of the IVP (1), which is C^∞ for any $t > 0$ and any $x \in \mathbb{R}^n$.

Proof. Since E is easily seen to be C^∞ for any $t > 0$ and any $x \in \mathbb{R}^n$, (i) is immediate. (ii) can be verified easily by noting that away from $t = 0$, E itself solves the Heat equation. For (iii), fix $x_0 \in \mathbb{R}^n$ and $\varepsilon > 0$. Now by continuity of g, choose $\delta > 0$ such that

$$|g(y) - g(x_0)| < \frac{\varepsilon}{2} \qquad \text{for all } y \in B_\delta(x_0).$$

Recall that we have seen how to calculate the integral of a Gaussian before. Using the same technique, it is easy to verify that for any $t > 0$, we have

$$\int_{\mathbb{R}^n} E(x-y,t)\ \mathrm{d}y = 1.$$

Now for any $x \in B_{\delta/2}(x_0)$ and any $t > 0$, we have

$$\begin{aligned}
|u(x,t) - g(x_0)| &= \left| \int_{\mathbb{R}^n} E(x-y,t)\left[g(y) - g(x_0)\right]\ \mathrm{d}y \right| \\
&\leq \int_{B_\delta(x_0)} E(x-y,t)\,|g(y) - g(x_0)|\ \mathrm{d}y \\
&\quad + \int_{\mathbb{R}^n \setminus B_\delta(x_0)} E(x-y,t)\,|g(y) - g(x_0)|\ \mathrm{d}y \\
&:= I_1 + I_2.
\end{aligned}$$

Clearly,

$$I_1 \leq \frac{\varepsilon}{2}\int_{B_\delta(x_0)} E(x-y,t)\ \mathrm{d}y \leq \frac{\varepsilon}{2}\int_{\mathbb{R}^n} E(x-y,t)\ \mathrm{d}y \leq \frac{\varepsilon}{2}.$$

Also, we have

$$I_2 \leq 2\,\|g\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n \setminus B_\delta(x_0)} E(x-y,t)\,\mathrm{d}y \leq \frac{C}{t^{\frac{n}{2}}}\int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x-y|^2}{4t}}\,\mathrm{d}y.$$

We need to show that the last expression converges to 0 as $t \to 0$, independently of $x \in B_{\delta/2}(x_0)$. The trick is to note that we can actually replace $|x - y|$ by $\frac{1}{2}|x_0 - y|$. Indeed, since $x \in B_{\delta/2}(x_0)$ and $y \notin B_\delta(x_0)$, we have

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

This implies

$$\frac{1}{2}|y - x_0| \leq |y - x|.$$

Thus, we deduce

$$\begin{aligned}
I_2 &\leq \frac{C}{t^{\frac{n}{2}}}\int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x-y|^2}{4t}}\,\mathrm{d}y \\
&\leq \frac{C}{t^{\frac{n}{2}}}\int_{\mathbb{R}^n \setminus B_\delta(x_0)} e^{-\frac{|x_0-y|^2}{16t}}\,\mathrm{d}y \to 0 \qquad \text{as } t \to 0+.
\end{aligned}$$

The result now follows easily. $\square$

Because of the importance of the function E, we give it a name.

Definition 2 (Heat Kernel)**.** *The function*

$$\Phi(x,t) := \begin{cases} \dfrac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} & \text{if } x \in \mathbb{R}^n, t > 0, \\ 0 & \text{if } x \in \mathbb{R}^n, t < 0, \end{cases}$$

*is called the **Heat kernel** or the **Gaussian kernel** of dimension n. This is also the fundamental solution for the Heat equation.*

1.1.2 Duhamel's principle

In this section, we study Duhamel's principle. We do not need a proof of Duhamel's principle as we can verify by direct computation that the solution of the inhomogeneous problem obtained by Duhamel's principle is a solution. However, it is still useful to understand why it works abstractly.

Theorem 3 (Duhamel's principle for ODEs)**.** *Consider the linear ODE*

$$\begin{cases} \dfrac{d}{dt} U = aU + f(t) & t \in (0,\infty), \\ U(0) = 0, \end{cases} \tag{3}$$

where $U : [0,\infty) \to \mathbb{R}$ is the unknown function, $a \in \mathbb{R}$ and $f : [0,\infty) \to \mathbb{R}$ is a given continuous function. Then the solution of (3) *is given by*

$$U(t) = \int_0^t U^s(t)\,\mathrm{d}s,$$

where $U^s : (s,\infty) \to \mathbb{R}$ solves

$$\begin{cases} \dfrac{d}{dt} U^s = aU^s & t \in (s,\infty), \\ U^s(s) = f(s). \end{cases} \tag{4}$$

Proof. The proof is simple, since we can write down the solution operator for the homogeneous linear equation explicitly. More precisely, we know that

$$V(t) = \zeta e^{at}$$

solves the equation

$$\begin{cases} \dfrac{d}{dt} V = aV & t \in (0,\infty), \\ V(0) = \zeta. \end{cases} \tag{5}$$

Thus, for any $0 \le s \le t$, we have

$$V(t) = \zeta e^{at} = \zeta e^{a(t-s)} e^{as} = e^{a(t-s)} V(s). \tag{6}$$

Using this for (4), we derive

$$U^s(t) = e^{a(t-s)} f(s).$$

Now we define

$$U(t) = \int_0^t U^s(t)\,\mathrm{d}s = \int_0^t e^{a(t-s)} f(s)\,\mathrm{d}s.$$

Then we have

$$\begin{aligned}\frac{d}{dt}U &= \frac{d}{dt}\left[\int_0^t e^{a(t-s)} f(s)\,\mathrm{d}s\right] \\ &= e^{a(t-s)} f(s)\big|_{s=t} + \int_0^t \frac{d}{dt}\left[e^{a(t-s)} f(s)\right]\mathrm{d}s \\ &= f(t) + a\left[\int_0^t e^{a(t-s)} f(s)\,\mathrm{d}s\right] \\ &= f(t) + aU.\end{aligned}$$

This proves the result. □

Now, using Fourier transform in $x \in \mathbb{R}^n$ to reduce the heat equation to an ODE, we now have the following, whose proof is left as an exercise.

Theorem 4 (Duhamel's principle for the heat equation)**.** *Consider the homogeneous initial value problem*

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = f(x,t) & \text{in } \mathbb{R}^n \times (0,\infty), \\ u = 0 & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{7}$$

where $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the unknown and $f : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is a given (smooth enough) function. Then the solution of (7) *is given by*

$$u(x,t) = \int_0^t u^s(x,t)\,\mathrm{d}s,$$

where $u^s : \mathbb{R}^n \times (s,\infty) \to \mathbb{R}$ solves

$$\begin{cases} u_t^s - \Delta u = 0 & \text{in } \mathbb{R}^n \times (s,\infty), \\ u^s = f & \text{on } \mathbb{R}^n \times \{t=s\}. \end{cases} \tag{8}$$

1.1.3 Representation formula for inhomogeneous IVP

Notation 5. *Let $f : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ be a real valued function, which we write as $f = f(x,t)$. We say $f \in C_1^2(\mathbb{R}^n \times (0,\infty))$ if*

$$f, \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x_i}, \frac{\partial^2 f}{\partial x_i \partial x_j} \in C(\mathbb{R}^n \times (0,\infty)) \qquad \text{for all } 1 \le i,j \le n.$$

Using the Duhamel's principle, we can now easily deduce that the following.

Theorem 6. *Let $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and $f \in C_1^2(\mathbb{R}^n \times (0,\infty))$ and f has compact support in $\mathbb{R}^n \times [0,\infty)$, i.e. $f \in C_c(\mathbb{R}^n \times [0,\infty))$. Define*

$$u(x,t) := \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\ \mathrm{d}y + \int_0^t \frac{1}{(4\pi(t-s))^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} f(y,s)\ \mathrm{d}y\mathrm{d}s,$$

Then we have

(i) $u \in C_1^2(\mathbb{R}^n \times (0,\infty))$,

(ii) *u satisfies*

$$u_t(x,t) - \Delta u(x,t) = f(x,t) \qquad \text{for all } x \in \mathbb{R}^n, t > 0.$$

(iii) $\lim_{(x,t)\to(x_0,0^+)} u(x,t) = g(x_0)$ *for every* $x_0 \in \mathbb{R}^n$.

The conclusion of the theorem of course means that u given by the formula furnishes a strong solution to the initial value problem

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

This can also be verified directly.

1.2 Mean value property

1.2.1 Heat Balls

We want to now show a mean value property for the heat equation. However, the formula is much more complicated as the relevant sets are no longer a sphere or balls. To discover what the sets should be, we recall the mean value formula for the harmonic functions and note that the reason balls appear in the mean value formula is because balls are the superlevel sets of the fundamental solutions. So along the same lines, we define **the heat ball**.

Definition 7 (Heat ball)**.** *Let $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$ and $r > 0$. We define the heat ball of radius r 'centered' at (x,t) as*

$$E(x,t;r) := \left\{ (y,s) \in \mathbb{R}^{n+1} : s \leq t, \Phi(x-y,t-s) \geq \frac{1}{r^n} \right\}.$$

This is a bounded region of the space-time whose boundary is the level set of the fundamental solution given by

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

Now a natural question that might arise is that superlevel sets could have been defined by

$$\left\{(y,s) \in \mathbb{R}^{n+1} : s \leq t, \Phi\left(x-y, t-s\right) \geq c\right\}$$

for any constant c. Why do we take $c = 1/r^n$ to define heat balls? The reason, once again, is the parabolic scaling. Note that for heat ball of radius r, the constant should of course, depend on r. But what should the exact nature of the dependence be? To figure this out, we notice that

$$\Phi\left(r\left(x-y\right), r^2\left(t-s\right)\right) = \frac{1}{r^n}\Phi\left(x-y, t-s\right) \qquad \text{for all } r > 0.$$

Thus, $c = 1/r^n$ is the correct dependence which ensures our heat balls respects the parabolic scaling. To figure out how the sets look like, we note that for any fixed $s < t$, we have

$$\left|x-y\right|^2 = 4n\left(t-s\right)\log\left(\frac{r}{\left[4\pi\left(t-s\right)\right]^{\frac{1}{2}}}\right),$$

which is a constant. Thus, each fixed 'time slice' is a ball in $\mathbb{R}^n$, but the radius varies in a rather complicated manner. However, it is not difficult to show that

$$\lim_{s\to t^-} 4n\left(t-s\right)\log\left(\frac{r}{\left[4\pi\left(t-s\right)\right]^{\frac{1}{2}}}\right) = 0.$$

Thus the so-called 'center' (x,t) is actually at the top, in fact at the center of the top of what looks like an ellipsoid and not somewhere in the middle at all. The heat ball is actually contained inside a rescaled parabolic cylinder. We leave it as an exercise to show

$$E\left(x,t;r\right) \subset \left\{(y,s) \in \mathbb{R}^{n+1} : |y-x| \leq \sqrt{\frac{n}{2\pi e}}r, t - \frac{r^2}{4\pi} \leq s \leq t\right\}.$$

1.2.2 Adjoint heat equation and backward heat kernel

Now we come to the question of how the mean value formula should look like. To have some idea, we consider any bounded open region $\mathcal{D} \subset \mathbb{R}^{n+1}$. Let $u, v \in C_c^\infty\left(\mathcal{D}\right)$. Integrating by parts, we deduce

$$\int_{\mathcal{D}} \left(u_t - \Delta u\right) v = \int_{\mathcal{D}} u\left(-v_t - \Delta v\right).$$

The operator

$$H^* := -\partial_t - \Delta$$

is called the **adjoint heat operator** or sometimes also called the **backward heat operator**. The reason for this last name is that formally, the operator is the heat equation in the variable $(x,-t)$. This analogy is actually an useful guideline. Let us define

Definition 8 (Backward Heat Kernel)**.** *The function*

$$\Phi_{back}(x,t) := \begin{cases} \dfrac{1}{(4\pi|t|)^{\frac{n}{2}}} e^{\frac{|x|^2}{4t}} & \text{if } x \in \mathbb{R}^n, t<0, \\ 0 & \text{if } x \in \mathbb{R}^n, t>0, \end{cases}$$

is called the ***backward heat kernel****.*

Now arguing exactly as in Theorem 1, we can establish

Theorem 9. *Let $g \in L^\infty(\mathbb{R}^n) \cap C(\mathbb{R}^n)$ and let u be defined by*

$$u(x,t) := \int_{\mathbb{R}^n} \Phi_{back}(x-y,t)\, g(y)\ \mathrm{d}y.$$

Then

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

(ii) *u satisfies*

$$-u_t - \Delta u = 0 \qquad \text{in } \mathbb{R}^n \times (-\infty,0),$$

(iii) *and we have*

$$\lim_{(x,t)\to(x_0,0^-)} u(x,t) = g(x_0) \qquad \text{for every } x_0 \in \mathbb{R}^n.$$

In other words, u is a strong solution to the following problem

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

Now note that the heat ball can also be written in terms of the backward heat kernel. Indeed, we have

$$\begin{aligned} \Phi_{\text{back}}(x,t) &= \frac{1}{(4\pi|t|)^{\frac{n}{2}}} e^{\frac{|x|^2}{4t}} \\ &= \frac{1}{(4\pi(-t))^{\frac{n}{2}}} e^{-\frac{|x|^2}{4(-t)}} \\ &= \Phi(x,-t) \ \text{ for any } x \in \mathbb{R}^n \text{ and any } t<0. \end{aligned}$$

Thus, we can write

$$E(x,t;r) := \left\{ (y,s) \in \mathbb{R}^{n+1} : s \le t, \Phi_{\text{back}}(x-y,s-t) \ge \frac{1}{r^n} \right\}.$$

1.2.3 Parabolic Greens identity

Now, for the rest we assume $(x,t) = (0,0)$. We can always get back to (x,t) by translation. Now, our plan is to derive a 'mean value inequality' of the form

$$u(0,0) \leq \int_{E(0,0;r)} w(y,s,r)\, u(y,s)\ \mathrm{d}y\mathrm{d}s,$$

for some suitable weight function $w(y,s,r)$. Once again, let $\mathcal{D} \subset \mathbb{R}^n$ be an ope, bounded region of space-time with smooth boundary and let $u, v \in C^{\infty}\left(\overline{\mathcal{D}}\right)$. Integrating by parts, we deduce the **Green's identity**

$$\int_{\mathcal{D}} (u_t - \Delta u)\, v = \int_{\mathcal{D}} u\,(-v_t - \Delta v) + \int_{\partial\mathcal{D}} \left[\tau uv - \frac{\partial u}{\partial \nu} v + u \frac{\partial v}{\partial \nu} v\right] \mathrm{d}\Sigma_{\partial\mathcal{D}}, \qquad (9)$$

where (ν, τ) denotes the exterior unit normal to $\partial\mathcal{D}$, i.e. τ denotes the t-component and ν denotes the x-component of the normal and $\mathrm{d}\Sigma_{\partial\mathcal{D}}$ is the 'surface measure' on $\partial\mathcal{D}$.

1.2.4 Heat mean value property

Now, since in the heuristics of distributions,

$$H^{*}\Phi_{\text{back}} = \delta_{(0,0)} \qquad \text{in } \mathbb{R}^{n+1},$$

if we formally substitute $v = \Phi_{\text{back}}$ in (9), we would have

$$\int_{\mathcal{D}} u\,(-v_t - \Delta v) \approx \delta_{(0,0)}[u] = u(0,0)$$

for any domain $\mathcal{D}$ with $(0,0) \in \mathcal{D}$. In particular, this would hold, at least morally, if $\mathcal{D} = E(0,0;r)$. However, the boundary integral will complicate issues. Also, at any rate, we want to get rid of the boundary integral altogether. So we start fine-tuning our choice of v. Observe that we have

$$\Phi_{\text{back}}(y,s) = \frac{1}{r^n} \qquad \text{for any } (y,s) \in \partial E(0,0;r).$$

Thus, if we choose

$$v = \Phi_{\text{back}} - \frac{1}{r^n},$$

then we still have

$$H^{*}v = \delta_{(0,0)} \qquad \text{in } \mathbb{R}^{n+1},$$

but we would also have

$$\int_{\partial E(0,0;r)} \left[\tau uv - \frac{\partial u}{\partial \nu} v\right] \mathrm{d}\Sigma_{\partial E(0,0;r)} = 0,$$

as v vanishes on $\partial E(0,0;r)$. But the term

$$\int_{\partial E(0,0;r)} u \frac{\partial v}{\partial \nu} \mathrm{d}\Sigma_{\partial E(0,0;r)} = \int_{\partial E(0,0;r)} u \langle \nu, \nabla_x v \rangle \, \mathrm{d}\Sigma_{\partial E(0,0;r)}$$

still remains. We now want to further modify our choice of v to get rid of this term, without undoing the good work we have done so far. To this end, we select

$$v = \Phi_{\text{back}} - \frac{1}{r^n} + c \log\left(r^n \Phi_{\text{back}}\right) \tag{10}$$

for some $c \in \mathbb{R}$. Observe carefully that since $\Phi_{\text{back}} \equiv 1/r^n$ on $\partial E(0,0;r)$, the last term also vanishes on $\partial E(0,0;r)$ for any $c \in \mathbb{R}$. Now, we compute the spatial gradient. We have,

$$\nabla v = \nabla \Phi_{\text{back}} + c \left(\frac{r^n \nabla \Phi_{\text{back}}}{r^n \Phi_{\text{back}}} \right) = \left(1 + \frac{c}{\Phi_{\text{back}}} \right) \nabla \Phi_{\text{back}}.$$

Thus, we deduce

$$\nabla v = (1 + c r^n) \nabla \Phi_{\text{back}} \qquad \text{on } \partial E(0,0;r).$$

Hence, we can choose $c = -1/r^n$ to get ∇v to vanish on $\partial E(0,0;r)$. Thus, our final choice for v is

$$v = \Phi_{\text{back}} - \frac{1}{r^n} - \frac{1}{r^n} \log\left(r^n \Phi_{\text{back}}\right). \tag{11}$$

However, though we managed to get rid of all the boundary terms, the price to pay is that we no longer have $H^* v = \delta_{(0,0)}$. Instead, we now obtain

$$\begin{aligned} H^* v &= H^* \Phi_{\text{back}} - \frac{1}{r^n} H^* \left[\log\left(-r^n \Phi_{\text{back}}\right) \right] \\ &= H^* \Phi_{\text{back}} + \frac{1}{r^n} \partial_t \left[\log\left(-r^n \Phi_{\text{back}}\right) \right] + \frac{1}{r^n} \Delta \left[\log\left(-r^n \Phi_{\text{back}}\right) \right]. \end{aligned} \tag{12}$$

We compute

$$\begin{aligned} \partial_t \left[\log\left(-r^n \Phi_{\text{back}}\right) \right] &= \partial_t \left[\log r^n - \frac{n}{2} \log\left(4\pi(-t)\right) + \frac{|x|^2}{4t} \right] \\ &= -\frac{n}{2t} - \frac{|x|^2}{4t^2}. \end{aligned}$$

Similarly, we deduce

$$\begin{aligned} \Delta \left[\log\left(-r^n \Phi_{\text{back}}\right) \right] &= \Delta \left[\log r^n - \frac{n}{2} \log\left(4\pi(-t)\right) + \frac{|x|^2}{4t} \right] \\ &= \Delta \left[\frac{|x|^2}{4t} \right] = \operatorname{div}\left(\frac{x}{2t} \right) = \frac{n}{2t}. \end{aligned}$$

Plugging these last two expressions back in (12), we arrive at

$$H^*v = H^*\Phi_{\text{back}} - \frac{1}{r^n}\frac{|x|^2}{4t^2}.$$

Substituting formally into (9), we have

$$\int_{E(0,0;r)} (u_t - \Delta u)\, v = \int_{E(0,0;r)} uH^*v = u\,(0,0) - \frac{1}{4r^n}\int_{E(0,0;r)} \frac{|x|^2}{t^2} u\,(x,t)\ \mathrm{d}x\mathrm{d}t$$

Now note that $r^n\Phi_{\text{back}} \geq 1$ in $E\,(0,0;r)$. Putting $z = \log\left(r^n\Phi_{\text{back}}\right)$, we can write

$$r^n v = e^z - 1 - z.$$

Since $f\,(z) = e^z - 1 - z \geq 0$ for all $z \geq 0$, we see immediately that $v \geq 0$ in $E\,(0,0;r)$. Thus, if $u_t - \Delta u \leq 0$, we obtain

$$u\,(0,0) - \frac{1}{4r^n}\int_{E(0,0;r)} \frac{|x|^2}{t^2} u\,(x,t)\ \mathrm{d}x\mathrm{d}t \leq 0.$$

This yields the mean value formula, at least formally. But the heuristic with Dirac delta distribution can be made rigorous easily. One applies the Green's identity with the same choice of u and v, but with the heat ball truncated at the top to remove the singularity. More precisely, we take the domain to be

$$\mathcal{D}^{\varepsilon} = \left\{(y,s) \in \mathbb{R}^{n+1} : s \leq -\varepsilon, \Phi\,(-y,-s) \geq \frac{1}{r^n}\right\}.$$

for $\varepsilon > 0$ small. Since the singularity is now outside the domain, the integral

$$\int_{\mathcal{D}^{\varepsilon}} uH^*\Phi_{\text{back}} = 0 \qquad \text{for every } \varepsilon > 0.$$

But there would be non-zero boundary terms corresponding to the top boundary

$$\left\{(y,s) \in \mathbb{R}^{n+1} : s = -\varepsilon, \Phi\,(-y,-s) \geq \frac{1}{r^n}\right\}.$$

It is easy to check that these boundary integrals converge to $u\,(0,0)$ as $\varepsilon \to 0^+$. Details are left to the reader. These arguments establish the following.

Theorem 10 (The mean value inequality)**.** *Let $u \in C_1^2\,(\Omega_T)$ satisfy $u_t - \Delta u \leq 0$ in Ω_T. Then*

$$u\,(x,t) \leq \frac{1}{4r^n}\int_{E(x,t;r)} u\,(y,s)\,\frac{|x-y|^2}{(t-s)^2}\ \mathrm{d}y\mathrm{d}s$$

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

1.3 Maximum principle

1.3.1 Strong and weak maximum principle

Now we can show the strong maximum principle for the heat equation.

Theorem 11 (Strong maximum principle)**.** *Let* $u \in C_1^2(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ *satisfy*

$$u_t - \Delta u \leq 0 \qquad \text{in } \Omega_T.$$

Then 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* $\overline{\Omega_{t_0}}$.

Remark 12. *Note that the solution is constant only in all earlier times, but not necessarily after.*

Proof. Let

$$M = u(x_0, t_0) = \max_{\overline{\Omega_T}} u.$$

Then we choose $r > 0$ sufficiently small such that $E(x_0, t_0; r) \subset \Omega_T$. Then by the mean value inequalities,

$$\begin{aligned} M = u(x_0, t_0) &\leq \frac{1}{4r^n} \int_{E(x_0,t_0;r)} u(y, s) \frac{|x_0 - y|^2}{(t_0 - s)^2} \, \mathrm{d}y\mathrm{d}s \\ &\leq \frac{M}{4r^n} \int_{E(x_0,t_0;r)} \frac{|x_0 - y|^2}{(t_0 - s)^2} \, \mathrm{d}y\mathrm{d}s = M. \end{aligned}$$

Thus, $u \equiv M$ on $E(x_0, t_0; r)$. Now for any other point $(z, t) \in \overline{\Omega_{t_0}}$, by connectedness, we can join (z, t) with (x_0, t_0) by piecewise continuous line segment paths where the time is decreasing and covering those paths by heat balls, we have the result. □

The strong maximum principle obviously implies the weak maximum principle.

Theorem 13 (Parabolic weak maximum principle)**.** *Let* $u \in C_1^2(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ *satisfy*

$$u_t - \Delta u \leq 0 \qquad \text{in } \Omega_T.$$

Then

$$\max_{\overline{\Omega_T}} u = \max_{\Gamma_{\Omega_T}} u.$$

1.3.2 Consequences of the maximum principle

1.3.3 Infinite propagation speed

As a consequence of the maximum principle, we can show the infinite propagation speed property of the heat equations without using the fundamental solution.

Theorem 14 (Infinite propagation speed)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, smooth and* ***connected****. Let $T > 0$, $g : \overline{\Omega} \to \mathbb{R}$ be a continuous* ***nonnegative*** *function and $u \in C_1^2(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ satisfy*

$$\begin{cases} u_t - \Delta u = 0 & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g & \text{on } \Omega \times \{t = 0\}. \end{cases}$$

If $g > 0$ somewhere *in Ω, then $u > 0$* everywhere *in Ω_T.*

Proof. Since u satisfy the heat equation, it is also a supersolution of the heat equation and thus satisfies the parabolic strong and the weak **minimum** principle. But since $g \geq 0$, we have

$$\min_{\Gamma_{\Omega_T}} u = 0.$$

Thus the parabolic weak minimum principle implies $u \geq 0$ in Ω_T. On the other hand, if $u(x,t) = 0$ for some $(x,t) \in \Omega_T$, the parabolic strong minimum principle would imply that

$$u \equiv 0 \qquad \text{in } \overline{\Omega_t}.$$

But this would imply

$$u \equiv 0 \qquad \text{on } \Omega \times \{t = 0\}.$$

But this contradicts the fact that $g > 0$ somewhere in Ω. □

Remark 15. *This result of course also holds for the pure IVP and is much easier to establish. This follows directly from the representation formula. Check this.*

1.3.4 Uniqueness via the maximum principle

The maximum principle also allows us to prove uniqueness results for the initial-Dirichlet boundary value problems.

Theorem 16 (Uniqueness for the initial-Dirichlet boundary value problem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, smooth and* ***connected****. Let $T > 0$ and $f : \overline{\Omega_T} \to \mathbb{R}$, $g : \overline{\Gamma_{\Omega_T}} \to \mathbb{R}$ be continuous functions. Then there exists* ***at most one***

solution $u \in C_1^2(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ *of the following initial-Dirichlet boundary value problem*

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

Proof. If $u, v \in C_1^2(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ are two solutions of the problem, then setting $w = u - v$ and applying the maximum principle to w and $-w$, we deduce that we must have $w = 0$. □

We can also show uniqueness for the Cauchy problem. However, the Cauchy problem

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

actually admits infinitely many solutions and thus uniqueness does not hold without additional conditions.

Example 17 (Tychonoff's counterexample)**.** *Let* $\alpha > 1$ *be a real number and define*

$$g(t) = \begin{cases} e^{-\frac{1}{t^\alpha}} & \text{if } t > 0, \\ 0 & \text{if } t \leq 0. \end{cases}$$

Now we define the function

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

We can check that the series in question converge uniformly for any bounded x *and real* $t > 0$*. Also, one can check that* u *satisfies*

$$u_t - u_{xx} = 0 \qquad \text{in } \mathbb{R} \times (0, \infty)$$

and

$$\lim_{t \to 0^+} u(x, t) = 0 \qquad \text{uniformly in } x \text{ for any bounded } x.$$

Thus, for $n = 1$*,* u *is a nontrivial solution of the IVP*

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

Hence to ensure uniqueness, we need to impose additional conditions. One can check that the Tychonoff's solutions grow quite fast as $|x| \to \infty$, so one way to force uniqueness might be to impose growth conditions at spatial infinity.

Theorem 18 (Uniqueness for the initial value problem)**.** *Let $T > 0$, $g \in C(\mathbb{R}^n)$ and $f \in C(\mathbb{R}^n \times [0,T])$. Then there exists* ***at most one*** *solution $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ of the following initial value problem*

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

that satisfies the growth estimate

$$|u(x,t)| \leq Ce^{\alpha|x|^2}$$

for all $x \in \mathbb{R}^n$, $0 \leq t \leq T$, for some constants $C, \alpha > 0$.

The theorem is an immediate consequence of the following weak maximum principle.

Theorem 19 (Parabolic weak maximum principle in $\mathbb{R}^n$)**.** *Let $T > 0$ and let $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ satisfy*

$$u_t - \Delta u \leq 0 \qquad \text{in } \mathbb{R}^n \times (0,T].$$

Then if there exist constants $C, \alpha > 0$ such that u satisfies the growth estimate

$$|u(x,t)| \leq Ce^{\alpha|x|^2}$$

for all $x \in \mathbb{R}^n$, $0 \leq t \leq T$, then

$$\sup_{\mathbb{R}^n \times [0,T]} u = \sup_{\mathbb{R}^n} u(\cdot, 0).$$

Proof. The idea is to use the parabolic weak maximum principle for bounded domains. But since we want to estimate the supremum of the function by the supremum on the bottom part of the parabolic cylinder only, we need to apply it to a function for which the values on the lateral boundary can be made small. To do trhis, first we assume that $T > 0$ is small enough to satisfy

$$4\alpha T < 1.$$

Then there exists $\beta > 0$ such that

$$4\alpha(T+\beta) < 1.$$

Now pick $\varepsilon > 0$ and define

$$v(x,t) := u(x,t) - \frac{\varepsilon}{(T+\beta-t)^{\frac{n}{2}}} e^{\frac{|x|^2}{4(T+\beta-t)}} \qquad \text{for any } x \in \mathbb{R}^n, 0 \leq t \leq T.$$

It is easy to check that

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

Thus, for any radius $r > 0$, by applying the weak maximum principle to v in the parabolic cylinder Ω_T, where $\Omega = B_r(0)$, we deduce

$$\max_{\overline{\Omega_T}} v = \max_{\Gamma_{\Omega_T}} v.$$

At the bottom part of the boundary, we have

$$v(x,0) = u(x,0) - \frac{\varepsilon}{(T+\beta)^{\frac{n}{2}}} e^{\frac{|x|^2}{4(T+\beta)}} \leq \sup_{B_r(0)} u(x,0).$$

But on the lateral part of the parabolic boundary Γ_{Ω_T}, we have, by the growth estimate

$$\begin{aligned} v(x,t) &\leq Ce^{\alpha|x|^2} - \frac{\varepsilon}{(T+\beta-t)^{\frac{n}{2}}} e^{\frac{|x|^2}{4(T+\beta-t)}} \\ &= Ce^{\alpha r^2} - \frac{\varepsilon}{(T+\beta-t)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\beta-t)}}. \end{aligned}$$

But since $4\alpha(T+\beta) < 1$, the second term with the negative sign has a higher exponent. Hence the RHS tends to $-\infty$ as $r \to \infty$. thus, letting $r \to \infty$, we deduce

$$\sup_{\mathbb{R}^n \times [0,T]} v \leq \sup_{\mathbb{R}^n} u(\cdot,0).$$

But since $\varepsilon > 0$ is arbitrary, this implies the result if T is sufficiently small. If T is not small enough, we subdivide the time interval into smaller ones, each of which satisfy the smallness assumption and apply the result successively to each one. This completes the proof. □

1.4 Interior regularity of strong solutions

We now show that any strong solution of the heat equation is automatically smooth in the interior of its domain. The circle of ideas are the same as in the case of the Laplace equation. We first derive apriori estimates for smooth solutions and then approximate any strong solution by smooth ones to derive the regularity conclusion. As before, there is no circularity here. Since our estimates are local, we need to use localization. To this end, we need to introduce a notation.

Notation 20. *For any $(x,t) \in \mathbb{R}^n \times \mathbb{R}$ and for any $r > 0$, the notation $C(x,t;r)$ stands for the closed circular cylinder radius r, height r^2 and the top center point (x,t), i.e.*

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

1.4.1 Local representation formula

We first begin by showing that we can derive apriori estimates for the derivatives for any **smooth** solution of the heat equation. First we derive a local representation formula.

Theorem 21. *Let $T > 0$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded. Let $u \in C^{\infty}(\Omega_T)$ satisfy*

$$u_t - \Delta u = 0 \qquad\qquad \text{in } \Omega \times (0, T).$$

Let $(x_0, t_0) \in \Omega \times (0, T]$ and $r > 0$ such that $C(x_0, t_0; r) \subset\subset \Omega \times (0, T]$. Let $\zeta \in C^{\infty}(\Omega_T)$ be such that $0 \le \zeta \le 1$ and

$$\zeta \equiv 1 \quad \text{in } C(x_0, t_0; 3r/4), \qquad \zeta \equiv 0 \quad \text{in } \mathbb{R}^n \times [0, t_0] \setminus C(x_0, t_0; r)$$

and $\zeta = 0$ near the curved and the bottom boundary of the cylinder $C(x_0, t_0; r)$.

Then we have

$$u(x, t) = \int_{t_0 - r^2}^{t} \int_{|y - x_0| < r} K(x, t, y, s)\, u(y, s)\ \mathrm{d}y\mathrm{d}s,$$

for any $(x, t) \in C(x_0, t_0; r/2)$, where

$$K(x, t, y, s) := \Phi(x - y, t - s)\left[\frac{\partial \zeta}{\partial s}(y, s) + \Delta\zeta(y, s)\right] + 2\left\langle \nabla_y \Phi(x - y, t - s), \nabla\zeta(y, s)\right\rangle,$$

where $\Phi(\cdot, \cdot)$ denotes the fundamental solution of the heat equation.

Proof. Set

$$v(x, t) := \zeta(x, t)\, u(x, t).$$

Then $v = 0$ on $\mathbb{R}^n \times \{t = 0\}$ and

$$v_t = \zeta_t u + \zeta u_t \qquad \text{and} \qquad \Delta v = \zeta \Delta u + 2\langle \nabla\zeta, \nabla u\rangle + u\Delta\zeta.$$

Hence, we have

$$\begin{aligned} v_t - \Delta v &= \zeta(u_t - \Delta u) + \zeta_t u - 2\langle \nabla\zeta, \nabla u\rangle - u\Delta\zeta \\ &= \zeta_t u - 2\langle \nabla\zeta, \nabla u\rangle - u\Delta\zeta, \end{aligned}$$

in $\mathbb{R}^n \times (0, t_0)$. Notice that the function

$$f := \zeta_t u - 2\langle \nabla\zeta, \nabla u\rangle - u\Delta\zeta \tag{13}$$

is actually defined everywhere in Ω_T and $f \in C_c^\infty(\Omega_T)$. By a slight abuse of notation, we would denote the extension by zero outside Ω_T of this function to $\mathbb{R}^n \times (0,\infty)$ by f as well. Clearly, v satisfies

$$\begin{cases} v_t - \Delta v = f & \text{in } \mathbb{R}^n \times (0, t_0)\,, \\ v = 0 & \text{on } \mathbb{R}^n \times \{t=0\}\,. \end{cases}$$

Define

$$\tilde{v}(x,t) := \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\, f(y,s)\, \mathrm{d}y\mathrm{d}s \text{ for all } (x,t) \in \mathbb{R}^n \times (0,\infty)\,.$$

By Theorem 6, $\tilde{v}$ solves initial value-problem

$$\begin{cases} \tilde{v}_t - \Delta \tilde{v} = f & \text{in } \mathbb{R}^n \times (0, \infty)\,, \\ \tilde{v} = 0 & \text{on } \mathbb{R}^n \times \{t=0\}\,. \end{cases}$$

and thus v, $\tilde{v}$ are both strong solutions of

$$\begin{cases} w_t - \Delta w = f & \text{in } \mathbb{R}^n \times (0, t_0)\,, \\ w = 0 & \text{on } \mathbb{R}^n \times \{t=0\}\,. \end{cases}$$

It is easy to see that both functions are bounded (by a constant, depending on u and ζ) and thus, by the uniqueness conclusion of Theorem 16, we have

$$v(x,t) = \tilde{v}(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\, f(y,s)\, \mathrm{d}y\mathrm{d}s$$

in $\mathbb{R}^n \times (0, t_0)$. Since $\zeta \equiv 1$ in $C(x_0, t_0; 3r/4)$ and $C(x_0, t_0; r/2) \subset C(x_0, t_0; 3r/4)$, we deduce

$$u(x,t) = v(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\, f(y,s)\, \mathrm{d}y\mathrm{d}s$$

for all $(x,t) \in C(x_0, t_0; r/2)$. Looking at (13), we arrive at

$$\begin{aligned} u(x,t) = & \int_{t_0 - r^2}^{t} \int_{|y-x_0|<r} \Phi(x-y, t-s) \left[\frac{\partial \zeta}{\partial s}(y,s) - \Delta \zeta(y,s) \right] u(y,s)\, \mathrm{d}y\mathrm{d}s \\ & -2 \int_{t_0 - r^2}^{t} \int_{|y-x_0|<r} \Phi(x-y, t-s) \langle \nabla \zeta(y,s), \nabla u(y,s) \rangle)\, \mathrm{d}y\mathrm{d}s, \end{aligned} \tag{14}$$

for all $(x,t) \in C(x_0, t_0; r/2)$. Now, since ζ vanishes near the curved boundary of $C(x_0, t_0; r)$, for every $t_0 - r^2 \leq s \leq t_0$, the map

$$y \mapsto \zeta(y,s) \qquad \text{is compactly supported inside the ball } B_r(x_0)\,.$$

Thus, integrating by parts in the y variable, we have

$$-2\int_{t_0-r^2}^{t}\int_{|y-x_0|<r}\Phi(x-y,t-s)\langle\nabla\zeta(y,s),\nabla u(y,s)\rangle)\,\mathrm{d}y\mathrm{d}s$$

$$=-2\int_{t_0-r^2}^{t}\int_{|y-x_0|<r}\langle\Phi(x-y,t-s)\nabla\zeta(y,s),\nabla u(y,s)\rangle)\,\mathrm{d}y\mathrm{d}s$$

$$=2\int_{t_0-r^2}^{t}\int_{|y-x_0|<r}\mathrm{div}_y\left[\Phi(x-y,t-s)\nabla\zeta(y,s)\right]u(y,s)\;\mathrm{d}y\mathrm{d}s$$

$$=2\int_{t_0-r^2}^{t}\int_{|y-x_0|<r}\langle\nabla_y\Phi(x-y,t-s),\nabla\zeta(y,s)\rangle u(y,s)\;\mathrm{d}y\mathrm{d}s$$

$$+2\int_{t_0-r^2}^{t}\int_{|y-x_0|<r}\Phi(x-y,t-s)\Delta\zeta(y,s)u(y,s)\;\mathrm{d}y\mathrm{d}s.$$

Substituting this in (14), we obtain the desired formula. □

Remark 22. *Notice that although we have assumed $u\in C^\infty(\Omega_T)$, the proof only needed u to be smooth in a neighborhood of $C(x_0,t_0;r)$ in Ω_T.*

1.4.2 Interior derivative estimates

Now we are ready to obtain our derivative estimates.

Theorem 23 (Interior apriori estimate for derivatives)**.** *Let $T>0$ and let $\Omega\subset\mathbb{R}^n$ be open and bounded. Let $u\in C^\infty(\Omega_T)$ satisfy*

$$u_t-\Delta u=0 \qquad\qquad \text{in } \Omega\times(0,T).$$

Then for any pair of nonnegative integers k,l, there exists a constant $C_{kl}>0$, independent of u, such that

$$\max_{C(x_0,t_0;r/2)}\left|D_x^\alpha D_t^l u\right|\le\frac{C_{kl}}{r^{k+2l+n+2}}\|u\|_{L^1(C(x_0,t_0;r))},$$

for every multiindex α with $|\alpha|=k$ for every cylinder

$$C(x_0,t_0;r/2)\subset C(x_0,t_0;r)\subset\subset\Omega\times(0,T].$$

Proof. First we begin by noting that we can assume, without loss of generality that $(x_0,t_0)=(0,0)$ and $r=1$. Indeed, fix a point $(x_0,t_0)\in\Omega\times(0,T]$ and a radius $r>0$ such that $C(x_0,t_0;r)\subset\subset\Omega\times(0,T]$. Set

$$v(x,t):=u\left(x_0+rx,t_0+r^2t\right).$$

Then v is smooth in a neighborhood of $C(0,0;1)$ and

$$v_t - \Delta v = 0$$

in a neighborhood of $C(0,0;1)$. Moreover, we have

$$\|v\|_{L^1(C(0,0;1))} = \frac{1}{r^{n+2}} \|u\|_{L^1(C(x_0,t_0;r))}$$

and

$$D_x^\alpha D_t^l v(x,t) = r^{2l+k} D_x^\alpha D_t^l u\left(x_0 + rx, t_0 + r^2 t\right)$$

for any (x,t) in a neighborhood of $C(0,0;1)$, for any pair of nonnegative integers k, l, and any multiindex α with $|\alpha| = k$. In view of the last two inequalities, proving the result for v would give us the result for u.

Now we return to the proof, where we have assumed $(x_0, t_0) = (0,0)$ and $r = 1$. Applying Theorem 21, we have

$$u(x,t) = \int_{-1}^{t} \int_{|y|<1} K(x,t,y,s)\, u(y,s)\ \mathrm{d}y\mathrm{d}s, \qquad \text{for any } (x,t) \in C(0,0;1/2),$$

where

$$K(x,t,y,s) := \Phi(x-y,t-s)\left[\frac{\partial \zeta}{\partial s}(y,s) + \Delta\zeta(y,s)\right] + 2\left\langle \nabla_y \Phi(x-y,t-s), \nabla\zeta(y,s)\right\rangle,$$

where $\Phi(\cdot,\cdot)$ denotes the fundamental solution of the heat equation and ζ is a cut-off function as in Theorem 21. Thus, for any pair of nonnegative integers k, l, and any multiindex α with $|\alpha| = k$, we deduce

$$\begin{aligned}
\left|D_x^\alpha D_t^l u(x,t)\right| &\le \int_{-1}^{t} \int_{|y|<1} \left|D_x^\alpha D_t^l K(x,t,y,s)\right| |u(y,s)|\ \mathrm{d}y\mathrm{d}s \\
&\le \int_{-1}^{0} \int_{|y|<1} \left|D_x^\alpha D_t^l K(x,t,y,s)\right| |u(y,s)|\ \mathrm{d}y\mathrm{d}s \\
&= \int_{C(0,0;1)} \left|D_x^\alpha D_t^l K(x,t,y,s)\right| |u(y,s)|\ \mathrm{d}y\mathrm{d}s \\
&\le C_{kl} \|u\|_{L^1(C(0,0;1))},
\end{aligned}$$

for some constant $C_{kl} > 0$ for any $(x,t) \in C(0,0;1/2)$. This completes the proof. $\square$

Remark 24. *Notice that once again, we only need u to be smooth in a neighborhood of $C(x_0,t_0;r)$ in Ω_T.*

1.4.3 Smoothness of solutions

Now we are in a position to use the apriori estimate for smooth solutions to prove that any strong solution of heat equation is smooth in the interior and satisfies the interior derivative estimtes.

Theorem 25 (Interior estimate for derivatives)**.** *Let $T > 0$ and let $\Omega \subset \mathbb{R}^n$ be open and bounded. Let $u \in C_1^2(\Omega_T)$ satisfy*

$$u_t - \Delta u = 0 \qquad\qquad \text{in } \Omega \times (0,T).$$

Then $u \in C^\infty(\Omega \times (0,T))$ and for any pair of nonnegative integers k,l, there exists a constant $C_{kl} > 0$, independent of u, such that

$$\max_{C(x_0,t_0;r/2)} \left| D_x^\alpha D_t^l u \right| \leq \frac{C_{kl}}{r^{k+2l+n+2}} \|u\|_{L^1(C(x_0,t_0;r))},$$

for every multiindex α with $|\alpha| = k$ for every cylinder

$$C(x_0,t_0;r/2) \subset C(x_0,t_0;r) \subset\subset \Omega \times (0,T).$$

Proof. Fix a point $(x_0,t_0) \in \Omega \times (0,T)$ and a radius $r > 0$ such that $C(x_0,t_0;r) \subset\subset \Omega \times (0,T)$. Let ϕ be a mollifying kernel in x and t. Let U is an open neighborhood of $C(x_0,t_0;r)$ such that $U \subset\subset \Omega \times (0,T)$. Then there exists $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$, the functions

$$u^\varepsilon := u * \phi_\varepsilon$$

is smooth in U. It is easy to check that these functions all satisfy

$$u_t^\varepsilon - \Delta u^\varepsilon = 0 \qquad \text{in } U.$$

By standard properties of convolutions, we also have

$$u^\varepsilon \to u \qquad \text{in } L^1(U).$$

Note that the linearity of the heat equation implies that for any two $0 < \varepsilon_1, \varepsilon_2 < \varepsilon_0$, the function $u^{\varepsilon_1} - u^{\varepsilon_2}$ also satisfies the heat equation in U. Thus, applying Theorem 23, for any pair of nonnegative integer k,l and for every multiindex α with $|\alpha| = k$, we deduce,

$$\begin{aligned}\max_{C(x_0,t_0;r/2)} \left| D_x^\alpha D_t^l \left(u^{\varepsilon_1} - u^{\varepsilon_2}\right) \right| &\leq \frac{C_{kl}}{r^{k+2l+n+2}} \|u^{\varepsilon_1} - u^{\varepsilon_2}\|_{L^1(C(x_0,t_0;r))} \\ &\leq \frac{C_{kl}}{r^{k+2l+n+2}} \|u^{\varepsilon_1} - u^{\varepsilon_2}\|_{L^1(U)} \to 0\end{aligned}$$

as $\varepsilon_1, \varepsilon_2 \to 0$. This shows that for any pair of nonnegative integer k,l and for every multiindex α with $|\alpha| = k$, the sequence of smooth functions $\left\{D_x^\alpha D_t^l u^\varepsilon\right\}_{\varepsilon > 0}$ are uniformly Cauchy and thus converges uniformly to a continuous function $v^{\alpha,l}$ in $C(x_0,t_0;r/2)$. In particular, $u^\varepsilon \to u$ in $C(x_0,t_0;r/2)$. Now, standard arguments imply that $u \in C^\infty\left(C\left(x_0, t_0; r/2\right)\right)$ and for any pair of nonnegative integer k, l and for every multiindex α with $|\alpha| = k$, we have

$$D_x^\alpha D_t^l u^\varepsilon \to v^{\alpha,l} = D_x^\alpha D_t^l u \qquad \text{uniformly in } C\left(x_0, t_0; r/2\right).$$

Since $(x_0, t_0) \in \Omega \times (0, T)$ is arbitrary, this implies $u \in C^\infty\left(\Omega \times (0, T)\right)$ and concludes the proof. □

Remark 26. *The result is actually valid for the top i.e. $t = T$ as well, as the estimates hold for $t = T$. Only the approximation by mollification argument needs to change a bit.*

1.5 Energy methods

Energy methods are a useful tool for the case of heat equation. It is the mathematical expression of the physical fact that evolution via the heat equation can only lose energy, but can not create energy.

1.5.1 Energy dissipation

Let $T > 0$ and let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. For any function $u \in C\left(\overline{\Omega_T}\right)$, then for any $0 \leq t \leq T$, we define the **energy** of u at time t as

$$E_u(t) := \int_\Omega \left|u(x,t)\right|^2 \, \mathrm{d}x.$$

Theorem 27 (Energy dissipation)**.** *Let $T > 0$ and let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. Let $u \in C_1^2\left(\overline{\Omega_T}\right)$ satisfy*

$$\begin{cases} u_t - \Delta u = 0 & \text{in } \Omega \times (0,T), \\ u = 0 & \text{on } \partial\Omega \times (0,T). \end{cases}$$

Then we have

$$\frac{d}{dt} E_u(t) \leq 0 \qquad \text{for all } 0 < t < T.$$

In particular, we have

$$E_u(t) \leq E_u(0) \qquad \text{for all } 0 \leq t \leq T.$$

Proof. We have

$$\begin{aligned} \frac{d}{dt} E_u(t) &= \frac{d}{dt}\left(\int_\Omega \left|u(x,t)\right|^2 \, \mathrm{d}x\right) \\ &= 2\int_\Omega u(x,t)\, u_t(x,t) \, \mathrm{d}x \\ &= 2\int_\Omega u(x,t)\, \Delta u(x,t) \, \mathrm{d}x \\ &= -2\int_\Omega \left|\nabla u(x,t)\right|^2 \, \mathrm{d}x \leq 0, \end{aligned}$$

for any $0 < t < T$. Now, integrating this with respect to t, we deduce

$$E_u(t) - E_u(0) = \int_0^t \frac{d}{dt} E_u(t) \ \mathrm{d}t \leq 0,$$

for any $0 < t \leq T$. This completes the proof. □

1.5.2 Uniqueness via energy methods

As a consequence of energy dissipation, we can now provide another proof of the uniqueness for the initial-boundary value problem for heat equation without using the maximum principle.

Theorem 28 (Uniqueness for the initial-Dirichlet boundary value problem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, smooth. Let $T > 0$ and $f : \overline{\Omega_T} \to \mathbb{R}$, $g : \overline{\Gamma_{\Omega_T}} \to \mathbb{R}$ be continuous functions. Then there exists* ***at most one*** *solution $u \in C_1^2\left(\overline{\Omega_T}\right)$ of the following initial-Dirichlet boundary value problem*

$$\begin{cases} u_t - \Delta u = f & \text{in } \Omega \times (0,T), \\ u = g & \text{on } \Gamma_{\Omega_T}. \end{cases}$$

Proof. If $u, v \in C_1^2\left(\overline{\Omega_T}\right)$ are two solutions of the problem, then setting $w = u - v$ we deduce that $w \in C_1^2\left(\overline{\Omega_T}\right)$ satisfy

$$\begin{cases} w_t - \Delta w = 0 & \text{in } \Omega \times (0,T), \\ w = 0 & \text{on } \Gamma_{\Omega_T}. \end{cases}$$

Thus, bu energy dissipation, for any $0 \leq t \leq T$, we have

$$\int_\Omega |w(x,t)|^2 \ \mathrm{d}x = E_w(t) \leq E_w(0) = 0.$$

This proves $w \equiv 0$ in Ω_T and completes the proof. □

1.6 Harnack inequality

1.6.1 Harnack inequality for the IVP

Theorem 29. *Let $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ be a solution of the following initial value problem*

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

Assume $u \geq 0$ in $\mathbb{R}^n \times (0,T)$ and $u \in L^\infty(\mathbb{R}^n \times [0,T])$. Then for any compact $K \subset \mathbb{R}^n$ and any $0 < t_1 < t_2 < T$ there exists a constant $C = C(K, t_1, t_2) > 0$ such that

$$\sup_{x \in K} u(x, t_1) \leq C \inf_{y \in K} u(y, t_2).$$

Proof. By the representation formula and uniqueness of the Cauchy problem, we have

$$u(x_2, t_2) = \int_{\mathbb{R}^n} \frac{1}{(4\pi t_2)^{\frac{n}{2}}} e^{-\frac{|x_2-y|^2}{4t_2}} g(y)dy.$$

Now, for $t_1 < t_2$ whenever $|x_1|, |x_2| \leq \Lambda < \infty$, there exists a constant $C = C(|t_1 - t_2|, \Lambda)$ so that

$$-\frac{|x_2-y|^2}{4t_2} \geq -\frac{|x_1-y|^2}{4t_1} - C. \quad \forall y \in \mathbb{R}^n$$

Consequently,

$$u(x_2, t_2) \geq \left(\frac{t_1}{t_2}\right)^{\frac{n}{2}} e^{-C} \int_{\mathbb{R}^n} \frac{1}{(t_1)^{\frac{n}{2}}} e^{-\frac{|x_1-y|^2}{4t_1}} g(y)dy = \left(\frac{t_1}{t_2}\right)^{\frac{n}{2}} e^{-C} u(x_1, t_1).$$

This completes the proof. □

In the above, we have used the following elementary estimate, whose proof is left as an exercise.

Lemma 30. *If $K \subset \mathbb{R}^n$ is compact and $0 < t_1 < t_2 < \infty$, then there exists a constant $C > 0$ depending on K and $t_2, t_1 > 0$, such that*

$$\frac{|x_1-y|^2}{t_2} \leq \frac{|x_2-y|^2}{t_1} + C \quad \forall x_1, x_2 \in K, y \in \mathbb{R}^n.$$

Exercise 31. *Let $g : \mathbb{R}^n \to [0, \infty)$ a smooth function with compact support such that $g(0) = 1$. Set*

$$u(x,t) := \int_{\mathbb{R}^n} \Phi(x-y,t)g(y) \quad t > 0$$

Show that

$$\inf_{x \in \mathbb{R}^n} u(x,t) = 0 \quad \textit{for all } t > 0$$

and

$$\sup_{x \in \mathbb{R}^n} u(x,t) > 0 \textit{ for all } t > 0.$$

Why does this not contradict Harnack's principle in Theorem 29?

Exercise 32. *Let us consider one space-dimension. Let $\xi \in \mathbb{R}$ be given and u defined as*

$$u_\xi(x,t) := (t+1)^{-\frac{1}{2}} e^{-\frac{x+\xi)^2}{4(t+1)}}.$$

Show that u is a solution of $(\partial_t - \Delta) u = 0$ in $\mathbb{R} \times (0, \infty)$. Moroever show for each fixed $t > 0$ there is no constant $C = C(t) > 0$ such that

$$\sup_{x \in [-1,1]} u_\xi(x,t) \leq C \inf_{y \in [-1,1]} u_\xi(y,t) \quad \forall \xi \in \mathbb{R}^n.$$

Why does this not contradict Harnack's principle in Theorem 29?

1.6.2 Harnack for the IBVP

We now want to prove the Harnack inequality for the Dirichlet IVP for the heat equation. For this, we need a slightly more general version of the weak maximum principle.

Theorem 33 (Weak maximum principle with first order terms)**.** *Let $B \in C(\Omega_T; \mathbb{R}^n)$ and let $u \in C^2_1(\Omega_T) \cap C\left(\overline{\Omega_T}\right)$ satisfy*

$$u_t - \Delta u - \langle B(x,t), \nabla u\rangle \leq 0 \qquad \textit{in } \Omega_T.$$

Then

$$\max_{\overline{\Omega_T}} u = \max_{\Gamma_{\Omega_T}} u.$$

Proof. Without loss of generality, we can assume

$$u_t - \Delta u - \langle B(x,t), \nabla u\rangle < 0 \qquad \text{in } \Omega_T.$$

Indeed, otherwise we consider $v(x,t) = u(x,t) - \mu t$ and let $\mu \to 0^+$ at the end. Now suppose, if possible, that $(x_0, t_0) \in \Omega_T$ is a point such that

$$u(x_0, t_0) = \max_{\overline{\Omega_T}} u.$$

Thus, we have $u_t(x_0,t_0) \geq 0$ (in fact $u_t(x_0,t_0) = 0$ unless $t_0 = T$), $\nabla u(x_0,t_0) = 0$ and $D^2 u(x_0,t_0)$ is nonpositive definite. Thus, we have

$$\Delta u(x_0,t_0) = \operatorname{trace} D^2 u(x_0,t_0) \leq 0.$$

But these inequalities imply

$$u_t(x_0,t_0) - \Delta u(x_0,t_0) - \langle B(x_0,t_0), \nabla u(x_0,t_0)\rangle \geq 0,$$

contradicting our assumption. This completes the proof. □

Theorem 34 (Harnack Inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Let $T > 0$. Let $u \in C^2_1(\Omega_T)$ be a* ***nonnegative*** *solution of*

$$u_t - \Delta u = 0 \qquad\qquad \textit{in } \Omega_T.$$

Suppose $\tilde{\Omega} \subset\subset \Omega$ is connected. For each $0 < t_1 < t_2 \leq T$, there exists a constant $C = C\left(\tilde{\Omega}, t_1, t_2\right) > 0$ such that

$$\sup_{\tilde{\Omega}} u(\cdot, t_1) \leq C \inf_{\tilde{\Omega}} u(\cdot, t_2).$$

Proof. By the interior regularity results proved earlier, we can assume u is smooth. Without loss of generality we may also assume that $\inf u > 0$ in Ω_T, as otherwise we would consider $u + \varepsilon$ and let $\varepsilon \to 0^+$ at the end. By a covering argument, it is also enough to prove the result when $\tilde{\Omega}$ is an open ball $B_R \subset\subset \Omega$. We set

$$v := \log u \quad \text{ in } \Omega_T.$$

We want to establish the differential inequality

$$v_t \geq \alpha|\nabla v|^2 - \beta \qquad \text{in } B_R \times [t_1, t_2]\,, \tag{15}$$

with constants $\alpha, \beta > 0$, which depends only on the ball B_R, the times t_1, t_2 and possibly also on the dimension n, but is independent of v.

First let us show that (15) is enough to prove the result. Fix $x_1, x_2 \in B_R$. Then we have

$$\begin{aligned} v(x_2, t_2) - v(x_1, t_1) &= \int_0^1 \frac{d}{ds} v\left(sx_2 + (1-s)x_1, st_2 + (1-s)t_1\right)\, \mathrm{d}s \\ &= \int_0^1 \left[\langle \nabla v, x_2 - x_1\rangle + v_t (t_2 - t_1)\right]\, \mathrm{d}s \\ &\overset{(15)}{\geq} \int_0^1 \left[-|\nabla v|\,|x_2 - x_1| + (t_2 - t_1)\left[\alpha|\nabla v|^2 - \beta\right]\right]\, \mathrm{d}s. \end{aligned}$$

Now the last integrand is a quadratic polynomial in $|\nabla v|$ with $\alpha(t_2 - t_1) > 0$. Thus, elementary manipulations by completing the square yields

$$\begin{aligned} &\alpha(t_2 - t_1)|\nabla v|^2 - |x_2 - x_1|\,|\nabla v| - \beta(t_2 - t_1) \\ &\qquad = \alpha(t_2 - t_1)\left(|\nabla v| + \frac{|x_2 - x_1|}{2\alpha(t_2 - t_1)}\right)^2 - \beta(t_2 - t_1) - \frac{|x_2 - x_1|^2}{4\alpha(t_2 - t_1)} \\ &\qquad \geq -\beta(t_2 - t_1) - \frac{|x_2 - x_1|^2}{4\alpha(t_2 - t_1)} \\ &\qquad \geq -\beta(t_2 - t_1) - \frac{R^2}{\alpha(t_2 - t_1)} \\ &\qquad := -\gamma. \end{aligned}$$

Thus, we deduce

$$\begin{aligned} v(x_2, t_2) - v(x_1, t_1) &\geq \int_0^1 \left[-|\nabla v|\,|x_2 - x_1| + (t_2 - t_1)\left[\alpha|\nabla v|^2 - \beta\right]\right]\, \mathrm{d}s \\ &\geq -\int_0^1 \gamma\, \mathrm{d}s = -\gamma. \end{aligned}$$

But by definition of v, this implies

$$u(x_1, t_1) \leq e^{\gamma} u(x_2, t_2)\,.$$

Since $x_1, x_2 \in B_R$ is arbitrary, this implies the desired Harnack inequality.

Thus it only remains to establish (15). We easily compute

$$u_t = e^v v_t, \quad \nabla u = e^v \nabla v \quad \text{and} \quad \Delta u = e^v \left(\Delta v + |\nabla v|^2\right).$$

This yields

$$0 = u_t - \Delta u = e^v \left(v_t - \Delta v - |\nabla v|^2\right).$$

Set $w := \Delta v$ and $\tilde{w} := |\nabla v|^2$ and note that the above equation implies

$$v_t = \Delta v + |\nabla v|^2 = w + \tilde{w} \qquad \text{in } \Omega_T. \tag{16}$$

Note that (15) is a lower bound for v_t. The term $|\nabla v|^2$ is nonnegative, but we do not know anything about the sign of Δv. However, if we can find a constant $0 < \alpha < 1$ such that

$$\Delta v + (1-\alpha)|\nabla v|^2$$

is bounded below, i.e.

$$\Delta v + (1-\alpha)|\nabla v|^2 \geq -\beta,$$

for some constant $\beta > 0$, this would establish (15). The crucial identity that helps us enormously in the computations is the following one.

$$\Delta |\nabla v|^2 = 2\left|D^2 v\right|^2 + 2\left\langle \nabla v, \nabla \Delta v\right\rangle. \tag{17}$$

This can be established by direct computation. Now applying the Laplacian to both sides of (16) and using (17), we immediately obtain

$$w_t - \Delta w - 2\nabla v \cdot \nabla w = 2\left|D^2 v\right|^2. \tag{18}$$

On the other hand, applying the gradient to both sides of (16) and using (17), we deduce

$$\begin{aligned}
\tilde{w}_t &= 2\left\langle \nabla v, \nabla v_t\right\rangle \\
&= 2\left\langle \nabla v, \nabla\left(\Delta v + |\nabla v|^2\right)\right\rangle \\
&= 2\left\langle \nabla v, \nabla \Delta v\right\rangle + 2\left\langle \nabla v, \nabla |\nabla v|^2\right\rangle \\
&= \Delta |\nabla v|^2 - 2\left|D^2 v\right|^2 + 2\left\langle \nabla v, \nabla |\nabla v|^2\right\rangle \\
&= \Delta \tilde{w} - 2\left|D^2 v\right|^2 + 2\left\langle \nabla v, \nabla \tilde{w}\right\rangle.
\end{aligned}$$

Thus, we arrive at

$$\tilde{w}_t - \Delta \tilde{w} - 2\left\langle \nabla v, \nabla \tilde{w}\right\rangle = -2\left|D^2 v\right|^2. \tag{19}$$

Now, noting (18) and (19), we see that the uniformly parabolic operator

$$\mathcal{P}U := \partial_t U - \Delta U - 2\langle \nabla v, \nabla U \rangle,$$

applied to w and $\tilde{w}$ have **opposite signs**. Thus, we can take $\alpha = 1/2$ and set

$$\hat{w} := w + \frac{1}{2}\tilde{w}$$

to discover

$$\hat{w}_t - \Delta \hat{w}_t - 2\langle \nabla v, \nabla \hat{w} \rangle = \left| D^2 v \right|^2 \geq 0.$$

This implies $\hat{w}$ is a supersolution to the uniformly parabolic operator $\mathcal{P}$ and thus would satisfy the weak minimum principle. Hence we can establish a lower bound for $\hat{w}$ if we knew a lower bound for $\hat{w}$ on Γ_{Ω_T}. Unfortunately, we do not have any information about the values of $\hat{w}$ on Γ_{Ω_T}. So instead we plan to modify $\hat{w}$ to make it vanish on Γ_{Ω_T}, but hopefully still keeping it a supersolution of $\mathcal{P}$, which would establish a lower bound. To this end, we choose a cutoff function $\zeta \in C^\infty(\Omega_T)$ such that $0 \leq \zeta \leq 1$ in Ω_T and

$$\zeta \equiv 1 \text{ in } B_R \times [t_1, t_2] \qquad \text{and} \qquad \zeta \equiv 0 \text{ near } \Gamma_T.$$

Now we set

$$W := \zeta^4 \hat{w} + \mu t \qquad \text{in } \Omega_T,$$

for some $\mu > 0$ that we are going to choose suitably later. Now direct computation yields

$$\begin{aligned} W_t - \Delta W - 2\langle \nabla v, \nabla W \rangle &= \zeta^4 \left(\hat{w}_t - \Delta \hat{w}_t - 2\nabla v \cdot \nabla \hat{w} \right) + \mu + R(x,t), \\ &= \zeta^4 \left| D^2 v \right|^2 + \mu + R(x,t), \end{aligned} \tag{20}$$

where the reminder term is given by

$$\begin{aligned} R(x,t) := 4\zeta^2 \zeta_t \hat{w} - 8\zeta^3 \langle \nabla v, \nabla \zeta \rangle \hat{w} - 8\zeta^3 \langle \nabla \hat{w}, \nabla \zeta \rangle \\ - 12\zeta^2 \left| \nabla \zeta \right|^2 \hat{w} - 4\zeta^2 \hat{w} \Delta \zeta. \end{aligned}$$

Now we claim that (20) implies W can not have a negative minima in Ω_T if $\mu > 0$ is chosen large enough. Suppose, for the sake of contradiction, that W attains a negative minima at $(x_0, t_0) \in \Omega_T$. Then, we have $W_t(x_0, t_0) \leq 0$ ($W_t(x_0, t_0) = 0$ if $t_0 < T$), $\nabla W(x_0, t_0) = 0$ and $D^2 W(x_0, t_0)$ is nonnegative definite. Thus $\Delta W(x_0, t_0) = \operatorname{trace} D^2 W(x_0, t_0) \geq 0$. Thus, the LHS of (20) is nonpositive at the point (x_0, t_0). But we are going to show now that we can choose $\mu > 0$ large enough (independently of (x_0, t_0)) such that

$$\zeta^4(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^2 + \mu + R(x_0, t_0) > 0.$$

This would yield the contradiction and rule out a negative minima for W.

Now note that since $W = \zeta^4 \hat{w} + \mu t$ is assumed to attain a negative minima at (x_0, t_0), we must have

$$\zeta(x_0, t_0) > 0 \qquad \text{and} \qquad \Delta v(x_0, t_0) + \frac{1}{2} |\nabla v(x_0, t_0)|^2 = \hat{w}(x_0, t_0) < 0.$$

This last inequality implies the inequalities

$$|\hat{w}(x_0, t_0)| \leq |\Delta v(x_0, t_0)| \leq c \left| D^2 v(x_0, t_0) \right|, \tag{21}$$

$$|\nabla v(x_0, t_0)|^2 \leq 2 |\Delta v(x_0, t_0)| \leq c \left| D^2 v(x_0, t_0) \right|. \tag{22}$$

A brief glance at the expression for R would make it clear that now it only remains to find a way to estimate $\nabla \hat{w}$. For this, we use the following fact

$$0 = \nabla W(x_0, t_0) = 4\zeta^3(x_0, t_0) \nabla \zeta(x_0, t_0) \hat{w}(x_0, t_0) + \zeta^4(x_0, t_0) \nabla \hat{w}(x_0, t_0).$$

Using the fact that $\zeta(x_0, t_0) \neq 0$, this yields

$$|\zeta(x_0, t_0) \nabla \hat{w}(x_0, t_0)| = |-4\hat{w}(x_0, t_0) \nabla \zeta(x_0, t_0)| \leq 4 |\hat{w}(x_0, t_0) \nabla \zeta(x_0, t_0)|. \tag{23}$$

Now, using (21), (22) and (23) and finally using Young's inequality with $\varepsilon > 0$, we deduce the estimate

$$\begin{aligned} |R(x_0, t_0)| &\leq c\zeta^2(x_0, t_0) \left| D^2 v(x_0, t_0) \right| + c\zeta^3(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^{\frac{3}{2}} \\ &\leq \varepsilon \zeta^4(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^2 + C_\varepsilon, \end{aligned} \tag{24}$$

for any $\varepsilon > 0$. Thus, choosing $0 < \varepsilon < 1$, we deduce the estimate

$$\begin{aligned} &\zeta^4(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^2 + \mu + R(x_0, t_0) \\ &\qquad \geq \zeta^4(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^2 + \mu - |R(x_0, t_0)| \\ &\qquad \geq (1 - \varepsilon) \zeta^4(x_0, t_0) \left| D^2 v(x_0, t_0) \right|^2 + \mu + C_\varepsilon \\ &\qquad \geq \mu + C_\varepsilon > 0, \end{aligned}$$

if $\mu > 0$ is chosen to satisfy $\mu > C_\varepsilon$. This implies $W = \zeta^4 \hat{w} + \mu t \geq 0$ in Ω_T. But since $\zeta \equiv 1$ in $B_R \times [t_1, t_2]$, we deduce

$$\hat{w} + \mu t \geq 0 \qquad \text{in } B_R \times [t_1, t_2].$$

This finally implies the estimate

$$\Delta v + \frac{1}{2} |\nabla v|^2 = \hat{w} \geq -\mu t_2 \qquad \text{in } B_R \times [t_1, t_2].$$

Returning to (16), this implies

$$v_t = \Delta v + |\nabla v|^2 \geq \frac{1}{2} |\nabla v|^2 - \mu t_2 \qquad \text{in } B_R \times [t_1, t_2].$$

This establishes (15) and completes the proof. □

2 L^2 estimates

2.1 First estimates and existence

2.1.1 First L^2 estimate

Theorem 35 (First apriori L^2 estimate)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. Let $T > 0$. Let*

$$\begin{aligned} A = A(x,t) &:= (a_{ij}(x,t))_{1\leq i,j\leq n} \in L^\infty\left(\Omega_T; \mathbb{R}^{n\times n}\right), \\ B = B(x,t) &:= (b_i(x,t))_{1\leq i\leq n} \in L^\infty\left(\Omega_T; \mathbb{R}^n\right), \\ c &= c(x,t) \in L^\infty(\Omega_T). \end{aligned}$$

Let A be uniformly elliptic in Ω_T with constant $\lambda > 0$, i.e. there exists some constant $\lambda > 0$ such that we have

$$\langle A(x,t)\xi, \xi\rangle \geq \lambda |\xi|^2$$

for a.e. $x \in \Omega$ and a.e. $t \in (0,T)$.

Let

$$\begin{aligned} g \in L^2(\Omega), \qquad & f \in L^2\left((0,T); L^2(\Omega)\right) \\ F := (F_i)_{1\leq i\leq n} &\in L^2\left((0,T); L^2(\Omega; \mathbb{R}^n)\right). \end{aligned}$$

Let $u \in C^\infty\left(\overline{\Omega_T}\right)$ be a smooth solution to

$$\begin{cases} u_t - \operatorname{div}(A\nabla u) + \langle B, \nabla u\rangle + cu = f(x,t) - \operatorname{div}_x F(x,t) & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g & \text{on } \Omega \times \{t = 0\}. \end{cases}$$

Then there exists a constant $C = C\left(\lambda, \|A\|_{L^\infty}, \|B\|_{L^\infty}, \|c\|_{L^\infty}, \Omega, T\right) > 0$ such that

$$\begin{aligned} \max_{0\leq t\leq T} \|u(t)\|_{L^2(\Omega)} + \|u\|_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|u_t\|_{L^2((0,T);H^{-1}(\Omega))} \\ \leq C\left(\|f\|_{L^2((0,T);L^2(\Omega))} + \|F\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} + \|g\|_{L^2(\Omega)}\right). \quad (25) \end{aligned}$$

Proof. Multiply the equation by u and for a.e. $0 < s < T$, integrate over Ω and integrate by parts. Noticing that u vanishes on $\partial\Omega$, we obtain

$$\begin{aligned} 0 &= \int_\Omega u_t u + \int_\Omega \langle A\nabla u, \nabla u\rangle + \int_\Omega \langle B, \nabla u\rangle u + \int_\Omega c|u|^2 - \int_\Omega fu - \int_\Omega \langle F, \nabla u\rangle \\ &\geq \frac{1}{2}\frac{d}{dt}\left(\int_\Omega |u|^2\right) + \lambda\int_\Omega |\nabla u|^2 - \left(|I_1| + |I_2| + |I_3| + |I_4|\right). \end{aligned}$$

Now, we have

$$|I_1| = \left|\int_\Omega \langle B, \nabla u\rangle\, u\right| \leq \int_\Omega |\langle B, \nabla u\rangle\, u|$$
$$\leq \|B\|_{L^\infty} \int_\Omega |\nabla u|\, |u| \leq \varepsilon \int_\Omega |\nabla u|^2 + C\int_\Omega |u|^2 .$$

$$|I_2| = \left|\int_\Omega c\, |u|^2\right| \leq \|c\|_{L^\infty} \int_\Omega |u|^2 .$$

$$|I_3| = \left|\int_\Omega fu\right| \leq \int_\Omega |f|\, |u| \leq C\int_\Omega |f|^2 + C\int_\Omega |u|^2$$

$$|I_4| = \left|\int_\Omega \langle F, \nabla u\rangle\right| \leq \int_\Omega |F|\, |\nabla u| \leq \varepsilon \int_\Omega |\nabla u|^2 + C\int_\Omega |F|^2 .$$

Choosing $\varepsilon > 0$ small enough, we deduce

$$\frac{d}{dt}\left(\int_\Omega |u|^2\right) + \lambda \int_\Omega |\nabla u|^2 \leq C_1 \left(\int_\Omega |u|^2 + \int_\Omega |f|^2 + \int_\Omega |F|^2\right)$$

for some constant $C_1 > 0$. But this implies

$$\frac{d}{dt}\left(e^{-C_1 t}\int_\Omega |u|^2\right) = e^{-C_1 t}\left[\frac{d}{dt}\left(\int_\Omega |u|^2\right) - C_1\int_\Omega |u|^2\right]$$
$$\leq e^{-C_1 t}\left[C_1\left(\int_\Omega |f|^2 + \int_\Omega |F|^2\right) - \lambda\int_\Omega |\nabla u|^2\right]$$

So we arrive at

$$\frac{d}{dt}\left(e^{-C_1 t}\int_\Omega |u|^2\right) + \lambda e^{-C_1 t}\int_\Omega |\nabla u|^2 \leq C_1 e^{-C_1 t}\left(\int_\Omega |f|^2 + \int_\Omega |F|^2\right)$$
$$\leq C_1\left(\int_\Omega |f|^2 + \int_\Omega |F|^2\right)$$

as $C_1 t > 0$ and thus $e^{-C_1 t} < 1$. Integrating with respect to t from 0 to s, where $0 < s < T$, we deduce

$$e^{-C_1 s}\int_\Omega |u(s)|^2 - \int_\Omega |u(0)|^2 + \lambda\int_0^s e^{-C_1 t}\int_\Omega |\nabla u(t)|^2 \, \mathrm{d}t$$
$$\leq C\left(\int_0^s \int_\Omega |f|^2 \, \mathrm{d}t + \int_0^s \int_\Omega |F|^2 \, \mathrm{d}t\right)$$
$$\leq C\left(\int_0^T \int_\Omega |f|^2 \, \mathrm{d}t + \int_0^T \int_\Omega |F|^2 \, \mathrm{d}t\right).$$
$$= C\left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|F\|^2_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))}\right).$$

Thus, using the obvious estimate $e^{-C_1 T} < e^{-C_1 t}$ for all $0 \leq t \leq T$, we have

$$
\begin{aligned}
e^{-C_1 s} \int_\Omega |u(s)|^2 + \lambda e^{-C_1 T} \int_0^s \int_\Omega |\nabla u(t)|^2 \ \mathrm{d}t & \\
\leq C \left(\int_0^T \int_\Omega |f|^2 \ \mathrm{d}t + \int_0^T \int_\Omega |F|^2 \ \mathrm{d}t \right) + \int_\Omega |u(0)|^2 .
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\int_\Omega |u(s)|^2 &+ \lambda e^{-C_1 T} \int_0^s \int_\Omega |\nabla u(t)|^2 \ \mathrm{d}t \\
&\leq \int_\Omega |u(s)|^2 + \lambda e^{-C_1 (T-s)} \int_0^s \int_\Omega |\nabla u(t)|^2 \ \mathrm{d}t \\
&\leq C e^{C_1 s} \left(\int_0^T \int_\Omega |f|^2 \ \mathrm{d}t + \int_0^T \int_\Omega |F|^2 \ \mathrm{d}t + \int_\Omega |u(0)|^2 \right) \\
&\leq C e^{C_1 T} \left(\int_0^T \int_\Omega |f|^2 \ \mathrm{d}t + \int_0^T \int_\Omega |F|^2 \ \mathrm{d}t + \int_\Omega |u(0)|^2 \right) \\
&= C e^{C_1 T} \left(\int_0^T \int_\Omega |f|^2 \ \mathrm{d}t + \int_0^T \int_\Omega |F|^2 \ \mathrm{d}t + \int_\Omega |g|^2 \right).
\end{aligned}
$$

Taking supremum over $0 \leq s \leq T$, we deduce

$$
\begin{aligned}
\sup_{0 \leq t \leq T} \|u(t)\|^2_{L^2(\Omega)} &+ \|u\|^2_{L^2\left((0,T); H_0^1(\Omega)\right)} \\
&\leq C \left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|F\|^2_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} + \|g\|^2_{L^2(\Omega)} \right). \qquad (26)
\end{aligned}
$$

Now multiplying the equation by $\phi \in L^2\left((0,T); H_0^1(\Omega)\right)$ and integrating over Ω and integrating by parts, we have

$$
0 = \int_\Omega u_t \phi + \int_\Omega \langle A \nabla u, \nabla \phi \rangle + \int_\Omega \langle B, \nabla u \rangle \phi + \int_\Omega c u \phi - \int_\Omega f \phi - \int_\Omega \langle F, \nabla \phi \rangle .
$$

Thus, we have

$$
\int_\Omega u_t \phi = - \int_\Omega \langle A \nabla u, \nabla \phi \rangle - \int_\Omega \langle B, \nabla u \rangle \phi - \int_\Omega c u \phi + \int_\Omega f \phi + \int_\Omega \langle F, \nabla \phi \rangle .
$$

Hence, we deduce

$$
\left| \int_\Omega u_t \phi \right| \leq |I_1| + |I_2| + |I_3| + |I_4| + |I_5| ,
$$

where we have

$$
\begin{aligned}
|I_1| &= \left|\int_\Omega \langle A\nabla u, \nabla\phi\rangle\right| \leq \|A\|_{L^\infty} \int_\Omega |\nabla u|\,|\nabla\phi| \leq \|A\|_{L^\infty} \|\nabla u\|_{L^2(\Omega)} \|\nabla\phi\|_{L^2(\Omega)}\,, \\
|I_2| &= \left|\int_\Omega \langle B, \nabla u\rangle\,\phi\right| \leq \|B\|_{L^\infty} \int_\Omega |\nabla u|\,|\phi| \leq \|B\|_{L^\infty} \|\nabla u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}\,, \\
|I_3| &= \left|\int_\Omega cu\phi\right| \leq \|c\|_{L^\infty} \int_\Omega |u|\,|\phi| \leq \|c\|_{L^\infty} \|u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}\,, \\
|I_4| &= \left|\int_\Omega f\phi\right| \leq \int_\Omega |f|\,|\phi| \leq \|f\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}\,, \\
|I_5| &= \left|\int_\Omega \langle F, \nabla\phi\rangle\right| \leq \int_\Omega |F|\,|\nabla\phi| \leq \|F\|_{L^2(\Omega)} \|\nabla\phi\|_{L^2(\Omega)}\,.
\end{aligned}
$$

This implies,

$$
\left|\int_\Omega u_t\phi\right| \leq C\left(\|u(t)\|_{H_0^1(\Omega)} + \|f(t)\|_{L^2(\Omega)} + \|F(t)\|_{L^2(\Omega)}\right)\|\phi(t)\|_{H_0^1(\Omega)}
$$

By the dual characterization of the $H^{-1}(\Omega)$ norm, this means

$$
\|u_t(t)\|_{H^{-1}(\Omega)} \leq C\left(\|u(t)\|_{H_0^1(\Omega)} + \|f(t)\|_{L^2(\Omega)} + \|F(t)\|_{L^2(\Omega)}\right).
$$

Squaring both sides and integrating with respect to t from 0 to T, we derive

$$
\begin{aligned}
&\|u_t\|^2_{L^2((0,T);H^{-1}(\Omega))} \\
&\qquad \leq C\left(\|u\|^2_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|f\|^2_{L^2((0,T);L^2(\Omega))} + \|F\|^2_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))}\right).
\end{aligned}
$$

Combined with (26), this implies the estimate

$$
\begin{aligned}
&\|u_t\|^2_{L^2((0,T);H^{-1}(\Omega))} \\
&\qquad\quad \leq C\left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|F\|^2_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} + \|g\|^2_{L^2(\Omega)}\right). \quad (27)
\end{aligned}
$$

Now by the theory of time-dependent Sobolev spaces, we know that

$$
\begin{cases} u \in L^2\left((0,T); H_0^1(\Omega)\right) \\ u_t \in L^2\left((0,T); H^{-1}(\Omega)\right) \end{cases} \Rightarrow u \in C\left([0,T]; L^2(\Omega)\right)
$$

along with the estimate

$$
\max_{0\leq t\leq T} \|u(t)\|_{L^2(\Omega)} \leq C\left(\|u\|_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|u_t\|_{L^2((0,T);H^{-1}(\Omega))}\right). \qquad (28)
$$

This allows the supremum in (26) to improve to a maximum and this combined with (27) implies (25) and completes the proof. □

2.1.2 Existence and uniqueness of weak solutions

The following assumptions would always be in force for this subsection. Let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. Let $T > 0$. Let

$$\begin{aligned} A = A(x,t) &:= (a_{ij}(x,t))_{1\leq i,j\leq n} \in L^{\infty}\left(\Omega_T; \mathbb{R}^{n\times n}\right), \\ B = B(x,t) &:= (b_i(x,t))_{1\leq i\leq n} \in L^{\infty}\left(\Omega_T; \mathbb{R}^{n}\right), \\ c &= c(x,t) \in L^{\infty}\left(\Omega_T; \mathbb{R}^{n}\right). \end{aligned}$$

Let A be uniformly elliptic in Ω_T with constant $\lambda > 0$, i.e. there exists some constant $\lambda > 0$ such that we have

$$\langle A(x,t)\xi, \xi\rangle \geq \lambda |\xi|^2$$

for a.e. $x \in \Omega$ and a.e. $t \in (0,T)$.

In view of the apriori estimates in the last section, it makes sense to define the concept of weak solutions as follows.

Definition 36 (Weak solutions)**.** *Given any $g \in L^2(\Omega)$, $f \in L^2\left((0,T); L^2(\Omega)\right)$ and $F \in L^2\left((0,T); L^2(\Omega;\mathbb{R}^n)\right)$, a function $u \in L^2\left((0,T); H^1_0(\Omega)\right)$ is called a* weak solution *to*

$$\begin{cases} u_t - \operatorname{div}(A\nabla u) + \langle B, \nabla u\rangle + cu = f - \operatorname{div} F & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g & \text{on } \Omega \times \{t = 0\}, \end{cases} \tag{29}$$

if the following are satisfied.

(i) $u \in L^2\left((0,T); H^1_0(\Omega)\right) \cap C\left([0,T]; L^2(\Omega)\right)$ *and* $u_t \in L^2\left((0,T); H^{-1}(\Omega)\right)$,

(ii) $u(0) = g$ *in* $L^2(\Omega)$ *and*

(iii) u *satisfies*

$$\int_0^T \int_\Omega \left[u_t\phi + \langle A\nabla u, \nabla\phi\rangle + \langle B, \nabla u\rangle \phi + cu\phi\right]\,\mathrm{d}x\mathrm{d}t = \int_0^T \int_\Omega \left[f\phi + \langle F, \nabla \phi\rangle\right]\,\mathrm{d}x\mathrm{d}t, \tag{30}$$

for every $\phi \in L^2\left((0,T); H^1_0(\Omega)\right)$.

Note that this definition, combined with estimate (25) immediately imply the uniqueness of weak solutions.

Theorem 37. *Given any $f \in L^2\left((0,T); L^2(\Omega)\right)$, $F \in L^2\left((0,T); L^2(\Omega;\mathbb{R}^n)\right)$ and any $g \in L^2(\Omega)$, there exists **at most one** weak solution to* (29).

Proof. If u, v are two weak solutions of (29), then $w = u - v$ is a weak solution of (29) with $f = 0$, $F = 0$ and $g = 0$. But then estimate (25) implies,

$$\max_{0 \leq t \leq T} \|w(t)\|_{L^2(\Omega)} = 0.$$

This implies $w \equiv 0$ and finishes the proof. □

Now using standard approximation and the **Galerkin** approximations, Theorem 35 implies the existence of weak solutions.

Theorem 38 (Existence of weak solutions)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. Let $T > 0$. Let*

$$\begin{aligned} A = A(x,t) &:= (a_{ij}(x,t))_{1 \leq i,j \leq n} \in L^\infty\left(\Omega_T; \mathbb{R}^{n \times n}\right), \\ B = B(x,t) &:= (b_i(x,t))_{1 \leq i \leq n} \in L^\infty(\Omega_T; \mathbb{R}^n), \\ c = c(x,t) &\in L^\infty(\Omega_T; \mathbb{R}^n). \end{aligned}$$

Let A be uniformly elliptic in Ω_T with constant $\lambda > 0$, i.e. there exists some constant $\lambda > 0$ such that we have

$$\langle A(x,t)\xi, \xi \rangle \geq \lambda |\xi|^2$$

for a.e. $x \in \Omega$ and a.e. $t \in (0,T)$.
Then given any $g \in L^2(\Omega)$, $f \in L^2\left((0,T); L^2(\Omega)\right)$ and $F \in L^2\left((0,T); L^2(\Omega; \mathbb{R}^n)\right)$, there exists a ***unique*** weak solution *u to* (29)*. Moreover, there exists a constant $C = C\left(\lambda, \|A\|_{L^\infty}, \|B\|_{L^\infty}, \|c\|_{L^\infty}, \Omega, T\right) > 0$ such that*

$$\begin{aligned} &\max_{0 \leq t \leq T} \|u(t)\|_{L^2(\Omega)} + \|u\|_{L^2\left((0,T); H_0^1(\Omega)\right)} + \|u_t\|_{L^2((0,T); H^{-1}(\Omega))} \\ &\qquad \leq C\left(\|f\|_{L^2((0,T); L^2(\Omega))} + \|F\|_{L^2((0,T); L^2(\Omega; \mathbb{R}^n))} + \|g\|_{L^2(\Omega)}\right). \end{aligned}$$

Proof. **Step 1: Reduction to smooth data** We begin by claiming that it is enough to show the theorem under the additional assumption that f, F, g, A, B, c are smooth up to the boundary i.e. we can assume without loss of generality, that $f \in C^\infty\left(\overline{\Omega_T}\right)$, $F \in C^\infty\left(\overline{\Omega_T}; \mathbb{R}^n\right)$, $g \in C^\infty\left(\overline{\Omega}\right)$, $A \in C^\infty\left(\overline{\Omega_T}; \mathbb{R}^{n \times n}\right)$, $B \in C^\infty\left(\overline{\Omega_T}; \mathbb{R}^n\right)$ and $c \in C^\infty\left(\overline{\Omega_T}\right)$, Indeed, by standard approximation results in Sobolev spaces, we can find sequences

$$\{f^\mu\}_{\mu \in \mathbb{N}} \subset C^\infty\left(\overline{\Omega_T}\right), \{F^\mu\}_{\mu \in \mathbb{N}} \subset C^\infty\left(\overline{\Omega_T}; \mathbb{R}^n\right), \{g^\mu\}_{\mu \in \mathbb{N}} C^\infty\left(\overline{\Omega}\right)$$

such that

$$\begin{aligned} f^\mu &\to f && \text{in } L^2\left((0,T); L^2(\Omega)\right), \\ F^\mu &\to F && \text{in } L^2\left((0,T); L^2(\Omega; \mathbb{R}^n)\right), \\ g^\mu &\to g && \text{in } L^2(\Omega). \end{aligned}$$

Extending by zero outside Ω_T if necessary and mollifying, we can also find sequences

$$\{A^\mu\}_{\mu\in\mathbb{N}} \subset C^\infty\left(\overline{\Omega_T} : \mathbb{R}^{n\times n}\right), \{B^\mu\}_{\mu\in\mathbb{N}} \subset C^\infty\left(\overline{\Omega_T}; \mathbb{R}^n\right), \{c^\mu\}_{\mu\in\mathbb{N}} C^\infty\left(\overline{\Omega_T}\right)$$

such that

$$\langle A^\mu(x,t)\xi, \xi\rangle \geq \frac{\lambda}{2}|\xi|^2$$

for every $(x,t) \in \Omega_T$ for all $\mu \in \mathbb{N}$ and

$$\begin{aligned} &A^\mu \to A, \quad B^\mu \to B \text{ and } c^\mu \to c \qquad \text{strongly in } L^p(\Omega_T) \text{ for any } 1 \leq p < \infty, \\ &\|A^\mu\|_{L^\infty} \leq \|A\|_{L^\infty}, \|B^\mu\|_{L^\infty} \leq \|B\|_{L^\infty} \text{ and } \|c^\mu\|_{L^\infty} \leq \|c\|_{L^\infty} \qquad \text{for all } \mu \in \mathbb{N}. \end{aligned}$$

Now if we can prove the theorem when the data is smooth, we can find a sequence of weak solutions $\{u^\mu\}_{\mu\in\mathbb{N}}$, such that for every $\mu \in \mathbb{N}$, u^μ is the weak solution to (29) with data f^μ, F^μ, g^μ, A^μ, B^μ and c^μ and satisfies the estimates

$$\begin{aligned} &\max_{0\leq t\leq T}\|u^\mu(t)\|_{L^2(\Omega)} + \|u^\mu\|_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|u_t^\mu\|_{L^2((0,T);H^{-1}(\Omega))} \\ &\qquad \leq C\left(\|f^\mu\|_{L^2((0,T);L^2(\Omega))} + \|F^\mu\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} + \|g^\mu\|_{L^2(\Omega)}\right). \end{aligned}$$

Note that the crucial point here is that the constant C in estimate (25) depends only on the ellipticity constant λ and the bounds on the L^∞ norms of A, B, c and thus can be chosen to be independent of $\mu \in \mathbb{N}$. Since the RHS of the above estimates can be bounded by

$$C\left(\|f\|_{L^2((0,T);L^2(\Omega))} + \|F\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} + \|g\|_{L^2(\Omega)}\right)$$

for all $\mu \in \mathbb{N}$, using standard functional analysis arguments, we deduce that up to the extraction of a subsequence that we do not relabel, we have

$$\begin{aligned} u^\mu &\rightharpoonup u && \text{in } L^2\left((0,T); H_0^1(\Omega)\right), \\ u_t^\mu &\rightharpoonup u && \text{in } L^2\left((0,T); H^{-1}(\Omega)\right), \\ u^\mu &\to u && \text{in } C\left([0,T]; L^2(\Omega)\right), \end{aligned}$$

for some $u \in L^2\left((0,T); H_0^1(\Omega)\right) \cap C\left([0,T]; L^2(\Omega)\right)$ with $u_t \in L^2\left((0,T); H^{-1}(\Omega)\right)$ and $u(0) = g$.

We can also verify that u is a weak solution of (29) with data f, F, g, A, B, c. Indeed, we have

$$\begin{aligned} \|A^\mu \nabla u^\mu\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} &\leq \|A^\mu\|_{L^\infty(\Omega_T;\mathbb{R}^{n\times n})} \|\nabla u^\mu\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} \\ &\leq \|A\|_{L^\infty(\Omega_T;\mathbb{R}^{n\times n})} \|\nabla u^\mu\|_{L^2((0,T);L^2(\Omega;\mathbb{R}^n))} \leq C, \end{aligned}$$

as $\{\nabla u^{\mu}\}_{\mu\in\mathbb{N}}$ is uniformly bounded in $L^2\left((0,T);L^2(\Omega;\mathbb{R}^n)\right)$. This means, up to the extraction of a subsequence which we do not relabel, we have

$$A^{\mu}\nabla u^{\mu} \rightharpoonup v$$

for some $v \in L^2\left((0,T);L^2(\Omega;\mathbb{R}^n)\right)$. We claim that $v = A\nabla u$. To see this, consider $\zeta \in C_c^{\infty}(\Omega_T;\mathbb{R}^n)$ and note that

$$\begin{aligned}\int_0^T\int_{\Omega}\langle A^{\mu}\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t \\ = \int_0^T\int_{\Omega}\langle (A^{\mu}-A)\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t + \int_0^T\int_{\Omega}\langle A\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t\end{aligned}$$

Now, for any $p > 2$, we have

$$\begin{aligned}\left|\int_{\Omega_T}\langle (A^{\mu}-A)\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t\right| &\leq \|A^{\mu}-A\|_{L^p(\Omega_T)}\|\nabla u^{\mu}\|_{L^2(\Omega_T)}\|\zeta\|_{L^{\frac{2p}{p-2}}(\Omega_T)} \\ &\leq C\|\zeta\|_{L^{\frac{2p}{p-2}}(\Omega_T)}\|A^{\mu}-A\|_{L^p(\Omega_T)} \to 0,\end{aligned}$$

by the strong convergence of A^{μ} to A in L^p. But by the weak convergence u^{μ} in $L^2\left((0,T);H_0^1(\Omega)\right)$, we have

$$\int_0^T\int_{\Omega}\langle A\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t \to \int_0^T\int_{\Omega}\langle A\nabla u,\zeta\rangle\ \mathrm{d}x\mathrm{d}t.$$

Thus, the last two estimate implies

$$\int_0^T\int_{\Omega}\langle A^{\mu}\nabla u^{\mu},\zeta\rangle\ \mathrm{d}x\mathrm{d}t \to \int_0^T\int_{\Omega}\langle A\nabla u,\zeta\rangle\ \mathrm{d}x\mathrm{d}t,$$

for any $\zeta \in C_c^{\infty}(\Omega_T;\mathbb{R}^n)$. Thus, by uniqueness of weak limits, we have, $v = A\nabla u$. Hence, we have

$$A^{\mu}\nabla u^{\mu} \rightharpoonup A\nabla u \qquad \text{weakly in } L^2\left((0,T);L^2(\Omega;\mathbb{R}^n)\right).$$

This implies

$$\int_0^T\int_{\Omega}\langle A^{\mu}\nabla u^{\mu},\nabla\phi\rangle\,\mathrm{d}x\mathrm{d}t \to \int_0^T\int_{\Omega}\langle A^{\mu}\nabla u^{\mu},\nabla\phi\rangle\,\mathrm{d}x\mathrm{d}t$$

for any $\phi \in L^2\left((0,T);H_0^1(\Omega)\right)$. Easier arguments would show that

$$\int_0^T\int_{\Omega}\langle B^{\mu},u^{\mu}\rangle\,\phi\mathrm{d}x\mathrm{d}t \to \int_0^T\int_{\Omega}\langle B,u\rangle\,\phi\mathrm{d}x\mathrm{d}t$$

and

$$\int_0^T\int_{\Omega}c^{\mu}u^{\mu}\phi\mathrm{d}x\mathrm{d}t \to \int_0^T\int_{\Omega}cu\phi\mathrm{d}x\mathrm{d}t$$

for any $\phi \in L^2\left((0,T); H_0^1(\Omega)\right)$.

Step 2: Galerkin approximations Hence for the rest of the proof, we would assume

$$f \in C^\infty\left(\overline{\Omega_T}\right), F \in C^\infty\left(\overline{\Omega_T}; \mathbb{R}^n\right) \text{ and } g \in C^\infty\left(\overline{\Omega}\right).$$

Now we use the **Galerkin approximations**. Let $\{\psi_k\}_{k\in\mathbb{N}}$ be the eigenfunctions of the Dirichlet Laplacian on Ω, i.e. for each $k \in \mathbb{N}$, ψ_k is the weak solution to

$$\begin{cases} -\Delta \psi_k = \lambda_k \psi_k & \text{in } \Omega, \\ \psi_k = 0 & \text{on } \partial\Omega, \end{cases}$$

where $0 < \lambda_1 < \lambda_2 \leq \ldots$ be the sequence of eigenvalues (repeated with multiplicity). Assume that we have orthogonalized the sequence in $H_0^1(\Omega)$ and orthonormalized in $L^2(\Omega)$. Define

$$X_m := \operatorname{Span}\{\psi_1, \psi_2, \ldots, \psi_m\} \subset H_0^1(\Omega) \qquad \text{for any } m \in \mathbb{N}.$$

We plan to project the problem (29) to the subspace X_m. More precisely, for each $m \in \mathbb{N}$, let

$$\begin{aligned} f^m(x,t) &:= \sum_{k=1}^m \left(\int_\Omega f(x,t)\,\psi_k(x)\ \mathrm{d}x \right) \psi_k(x), \\ F^m(x,t) &:= \sum_{k=1}^m \left(\int_\Omega F(x,t)\,\psi_k(x)\ \mathrm{d}x \right) \psi_k(x), \\ g^m(x) &:= \sum_{k=1}^m \left(\int_\Omega g(x)\,\psi_k(x)\ \mathrm{d}x \right) \psi_k(x). \end{aligned}$$

We want to find functions $\alpha^{k,m} : [0,T] \to \mathbb{R}$ such that

$$u^m(x,t) = \sum_{k=1}^m \alpha^{k,m}(t)\,\psi_k(x) \tag{31}$$

is a weak solution to

$$\begin{cases} u_t^m - \operatorname{div}(A\nabla u^m) + \langle B, \nabla u^m \rangle + cu^m = f^m - \operatorname{div} F^m & \text{in } \Omega_T, \\ u^m = 0 & \text{on } \partial\Omega \times [0,T], \\ u^m = g^m & \text{on } \Omega \times \{t=0\}. \end{cases} \tag{32}$$

Using $u^m(x,0) = g^m$ and the fact that $\{\psi_k\}_{k\in\mathbb{N}}$ is an orthonormal Schauder basis in $L^2(\Omega)$, we see that we must have

$$\alpha^{k,m}(0) = \int_\Omega g(x)\,\psi_k(x)\ \mathrm{d}x \qquad \text{for all } 1 \leq k \leq m. \tag{33}$$

Now fix a $1 \leq i \leq m$ and multiplying the equation by $\psi_i(x)$ and integrating by parts, we have

$$\int_\Omega \left[u_t^m \psi_i + \langle A\nabla u^m, \nabla\psi_i\rangle + \langle B, \nabla u^m\rangle \psi_i + cu^m\psi_i\right] \mathrm{d}x$$
$$= \int_\Omega \left[f^m\psi_i + \langle F^m, \nabla\psi_i\rangle\right] \mathrm{d}x \quad (34)$$

Now, from (31), we calculate

$$u_t^m = \sum_{k=1}^m \frac{d}{dt}\left(\alpha^{k,m}\right)(t)\,\psi_k(x),$$
$$\nabla u^m = \sum_{k=1}^m \alpha^{k,m}(t)\,\nabla\psi_k(x)$$

Since $\{\psi_k\}_{k\in\mathbb{N}}$ is orthonormal in $L^2(\Omega)$, we have

$$\int_\Omega u_t^m\psi_i(x) = \frac{d}{dt}\left(\alpha^{i,m}\right)(t),$$
$$\int_\Omega cu^m\psi_i = c(x,t)\,\alpha^{i,m}(t),$$
$$\int_\Omega f^m\psi_i = \int_\Omega f(x,t)\,\psi_i.$$

Substituting in (34), we deduce

$$\frac{d}{dt}\left(\alpha^{i,m}\right)(t)$$
$$= -\sum_{k=1}^m \alpha^{k,m}(t)\int_\Omega \left[\langle A(x,t)\nabla\psi_k, \nabla\psi_i\rangle + \langle B(x,t), \nabla\psi_k\rangle\psi_i\right]$$
$$- c(x,t)\,\alpha^{i,m}(t) + \int_\Omega f(x,t)\,\psi_i$$
$$+ \sum_{k=1}^m \int_\Omega \left\langle \int_\Omega F(x,t)\,\psi_k(x)\; \mathrm{d}x, \nabla\psi_i\right\rangle \psi_k.$$

Since the argument can be repeated for every $1 \leq i \leq m$, we get the following initial value problem for the system of m linear ODEs

$$\begin{cases} \dfrac{d}{dt}\alpha^m(t) = P^m(t)\,\alpha(t) + q^m(t), \\ \alpha^m(0) = \alpha_0^m, \end{cases} \quad (35)$$

where $\alpha^m(t) = \left\{\alpha^{i,m}(t)\right\}_{1\leq i\leq m}$, $q^m(t) = \left\{q^{i,m}(t)\right\}_{1\leq i\leq m}$, are m-vector fields and $P^m(t) = \left\{P_{ij}^m(t)\right\}_{1\leq i,j\leq m}$, is an $m\times m$ matrix field, $\alpha_0^m = \left\{\alpha_0^{i,m}\right\}_{1\leq i\leq m}$

is a fixed vector in $\mathbb{R}^m$ and

$$\alpha_0^{i,m} = \int_\Omega g\psi_i,$$

$$q^{i,m}(t) = \int_\Omega f(x,t)\psi_i + \sum_{k=1}^m \int_\Omega \left\langle \int_\Omega F(x,t)\psi_k(x)\ \mathrm{d}x, \nabla\psi_i \right\rangle \psi_k,$$

$$P_{ij}^m(t) = -\int_\Omega \left[\langle A(x,t)\nabla\psi_j, \nabla\psi_i\rangle + \langle B(x,t), \nabla\psi_j\rangle \psi_i\right] - c(x,t)\delta_{ij},$$

for all $1 \leq i,j \leq m$, where δ_{ij} is the Kronecker delta. Now since this is an IVP for a system of linear ODEs with smooth coefficients, there exists a unique smooth solution. Thus, we can construct a smooth u^m which is a weak solution to (32). Now since by Bessel's inequality, we have

$$\begin{aligned} f^m &\to f && \text{in } L^2\left((0,T); L^2(\Omega)\right), \\ F^m &\to F && \text{in } L^2\left((0,T); L^2(\Omega;\mathbb{R}^n)\right), \\ g^m &\to g && \text{in } L^2(\Omega). \end{aligned}$$

Thus, using the apriori estimate (25) and arguing as in Step 1, we conclude that up to the extraction of a subsequence that we do not relabel, we have

$$\begin{aligned} u^m &\rightharpoonup u && \text{in } L^2\left((0,T); H_0^1(\Omega)\right), \\ u_t^m &\rightharpoonup u && \text{in } L^2\left((0,T); H^{-1}(\Omega)\right), \\ u^m &\to u && \text{in } C\left([0,T]; L^2(\Omega)\right), \end{aligned}$$

for some $u \in L^2\left((0,T); H_0^1(\Omega)\right) \cap C\left([0,T]; L^2(\Omega)\right)$ with $u_t \in L^2\left((0,T); H^{-1}(\Omega)\right)$ and $u(0) = g$. It is easy to verify that u is a weak solution of (29) with data f, F, g, A, B, c and satisfies the estimate (25). This completes the proof. □

2.2 Higher L^2 estimates

Now we want to derive L^2 estimates for higher derivatives. However, since the equation intermingles spatial derivatives and time derivative, to derive higher estimates, we need to have some compatibility condition between the initial data g and time derivatives of source term. We begin with L^2 estimates for the Hessian, which requires a rather mild compatibility, i.e. g to have zero trace on the boundary.

Theorem 39 (Apriori L^2 estimate for D^2u)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, bounded and smooth. Let $T > 0$. Let*

$$\begin{aligned} A = A(x,t) &:= \left(a_{ij}(x,t)\right)_{1\leq i,j\leq n} \in W^{1,\infty}\left(\Omega_T; \mathbb{R}^{n\times n}\right), \\ B = B(x,t) &:= \left(b_i(x,t)\right)_{1\leq i\leq n} \in W^{1,\infty}\left(\Omega_T; \mathbb{R}^n\right), \\ c = c(x,t) &\in W^{1,\infty}\left(\Omega_T\right). \end{aligned}$$

Let A be uniformly elliptic in Ω_T with constant $\lambda > 0$, i.e. there exists some constant $\lambda > 0$ such that we have

$$\langle A(x,t)\xi, \xi\rangle \geq \lambda |\xi|^2$$

for a.e. $x \in \Omega$ and a.e. $t \in (0,T)$. Let $u \in C^\infty\left(\overline{\Omega_T}\right)$ be a smooth solution to

$$\begin{cases} u_t - \operatorname{div}(A\nabla u) + \langle B, \nabla u\rangle + cu = f - \operatorname{div}_x F & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g & \text{on } \Omega \times \{t = 0\}. \end{cases}$$

(i) Suppose

$$g \in H_0^1(\Omega), \qquad f \in L^2\left((0,T); L^2(\Omega)\right), \qquad F \in L^2\left((0,T); H^1(\Omega; \mathbb{R}^n)\right).$$

Then there exists a constant

$$C = C\left(\lambda, \|A\|_{W^{1,\infty}}, \|B\|_{W^{1,\infty}}, \|c\|_{W^{1,\infty}}, \Omega, T\right) > 0$$

such that

$$\sup_{0\leq t\leq T} \|u(t)\|_{H_0^1(\Omega)} + \|u\|_{L^2((0,T);H^2(\Omega))} + \|u_t\|_{L^2((0,T);L^2(\Omega))} \leq C\left(\|f\|_{L^2((0,T);L^2(\Omega))} + \|F\|_{L^2((0,T);H^1(\Omega;\mathbb{R}^n))} + \|g\|_{H_0^1(\Omega)}\right). \quad (36)$$

(ii) Suppose

$$g \in H_0^1(\Omega) \cap H^2(\Omega), \quad f \in H^1\left((0,T); L^2(\Omega)\right), \quad F \in H^1\left((0,T); H^1(\Omega; \mathbb{R}^n)\right).$$

Then there exists a constant

$$C = C\left(\lambda, \|A\|_{W^{1,\infty}}, \|B\|_{W^{1,\infty}}, \|c\|_{W^{1,\infty}}, \Omega, T\right) > 0$$

such that

$$\sup_{0\leq t\leq T}\left(\|u(t)\|_{H^2(\Omega)} + \|u_t(t)\|_{L^2(\Omega)}\right) + \|u_t\|_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|u_{tt}\|_{L^2((0,T);H^{-1}(\Omega))} \leq C\left(\|f\|_{L^2((0,T);L^2(\Omega))} + \|F\|_{L^2((0,T);H^1(\Omega;\mathbb{R}^n))} + \|g\|_{H_0^1(\Omega)}\right). \quad (37)$$

Proof. By a similar approximation argument, we can assume if necessary that A, B, c are also smooth. So we would not care about smoothness at all while deriving the apriori estimates. Also note that for both parts, we can just assume $F \equiv 0$, as the hypothesis allows us to absorb the term $\operatorname{div}_x F$ into f. We begin by proving Part (i).

Part (i): Multiply the equation by u_t and for a.e. $0 < s < T$, integrate over Ω and integrate by parts. Note that u_t also vanishes on $\partial\Omega$, as $u \equiv 0$ on $\partial\Omega$ for all time $0 \leq t \leq T$, (for $t = 0$, note carefully that this is implied by the explicit assumption $g \in H_0^1$), its time derivative must also vanish. Hence we obtain

$$\begin{aligned} 0 &= \int_\Omega |u_t|^2 + \int_\Omega \langle A\nabla u, \nabla u_t\rangle + \int_\Omega \langle B, \nabla u\rangle\, u_t + \int_\Omega cuu_t - \int_\Omega fu_t \\ &\geq \int_\Omega |u_t|^2 + \frac{1}{2}\frac{d}{dt}\left(\int_\Omega \langle A\nabla u, \nabla u\rangle\right) - \frac{1}{2}\int_\Omega \langle A_t\nabla u, \nabla u\rangle - \left(|I_1| + |I_2| + |I_3|\right). \end{aligned}$$

Now, we have

$$\begin{aligned} |I_1| = \left|\int_\Omega \langle B, \nabla u\rangle\, u_t\right| &\leq \int_\Omega |\langle B, \nabla u\rangle\, u_t| \\ &\leq \|B\|_{L^\infty} \int_\Omega |\nabla u|\, |u_t| \leq \varepsilon \int_\Omega |u_t|^2 + C\int_\Omega |\nabla u|^2 . \end{aligned}$$

$$|I_2| = \left|\int_\Omega cuu_t\right| \leq \|c\|_{L^\infty} \int_\Omega |\nabla u|\, |u_t| \leq \varepsilon \int_\Omega |u_t|^2 + C\int_\Omega |\nabla u|^2 .$$

$$|I_3| = \left|\int_\Omega fu_t\right| \leq \int_\Omega |f|\, |u_t| \leq \varepsilon \int_\Omega |u_t|^2 + C\int_\Omega |f|^2 .$$

We also have

$$\left|\frac{1}{2}\int_\Omega \langle A_t\nabla u, \nabla u\rangle\right| \leq \frac{1}{2}\|A_t\|_{L^\infty}\int_\Omega |\nabla u|^2 .$$

Choosing $\varepsilon > 0$ small enough, we deduce

$$\int_\Omega |u_t|^2 + \frac{d}{dt}\left(\int_\Omega \langle A\nabla u, \nabla u\rangle\right) \leq C_1\left(\int_\Omega |\nabla u|^2 + \int_\Omega |f|^2\right).$$

for some constant $C_1 > 0$. Integrating with respect to t from 0 to s, where $0 < s < T$, we deduce

$$\begin{aligned} \int_0^s \int_\Omega |u_t|^2 &+ \int_\Omega \langle A(s)\nabla u(s), \nabla u(s)\rangle - \int_\Omega \langle A(0)\nabla u(0), \nabla u(0)\rangle \\ &\leq C\left(\int_0^s \int_\Omega |f|^2 \,\mathrm{d}t + \int_0^s\int_\Omega |\nabla u|^2 \,\mathrm{d}t\right) \\ &\leq C\left(\int_0^T \int_\Omega |f|^2 \,\mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \,\mathrm{d}t\right). \\ &= C\left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H_0^1(\Omega)\right)}\right). \end{aligned}$$

Using ellipticity of $A(s)$, we deduce

$$
\begin{aligned}
&\int_0^s \int_\Omega |u_t|^2 + \lambda \int_\Omega |\nabla u(s)|^2 \\
&\leq \int_0^s \int_\Omega |u_t|^2 + \int_\Omega \langle A(s) \nabla u(s), \nabla u(s) \rangle \\
&\leq \int_\Omega \langle A(0) \nabla u(0), \nabla u(0) \rangle + C \left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \right) \\
&\leq \int_\Omega \langle A(0) \nabla g, \nabla g \rangle + C \left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \right) \\
&\leq \|A\|_{L^\infty} \int_\Omega |\nabla g|^2 + C \left(\|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \right) \\
&\leq C \left(\|g\|^2_{H^1_0(\Omega)} + \|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \right).
\end{aligned}
$$

Hence, taking supremum over $0 \leq s \leq T$, we deduce

$$
\begin{aligned}
&\sup_{0 \leq t \leq T} \|u(t)\|^2_{H^1_0(\Omega)} + \|u_t\|^2_{L^2((0,T);L^2(\Omega))} \\
&\qquad \leq C \left(\|g\|^2_{H^1_0(\Omega)} + \|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \right). \qquad (38)
\end{aligned}
$$

On the other hand, for each $0 < s < T$, we have

$$
-\operatorname{div}(A(s) \nabla u(s)) + \langle B(s), \nabla u(s) \rangle + c(s) u(s) = f(s) - u_t(s) \quad \text{in } \Omega,
$$

Since the LHS of the above equation is an uniformly elliptic operator, using elliptic regularity, more precisely, up to the boundary $W^{2,2}$ estimates imply the estimate

$$
\int_\Omega \left| D^2 u(s) \right|^2 \leq C \left(\int_\Omega |\nabla u(s)|^2 + \int_\Omega |f(s)|^2 + \int_\Omega |u_t(s)|^2 \right). \qquad (39)
$$

Integrating with respect to s from 0 to T, we arrive at

$$
\begin{aligned}
&\|u\|^2_{L^2((0,T);H^2(\Omega))} \\
&\qquad \leq C \left(\|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} + \|f\|^2_{L^2((0,T);L^2(\Omega))} + \|u_t\|^2_{L^2((0,T);L^2(\Omega))} \right). \qquad (40)
\end{aligned}
$$

Now, combining (38), (40) with (25), we obtain (36). This completes the proof of (i).

Part (ii): Differentiating the equation with respect to time, we have

$$
\begin{aligned}
&u_{tt} - \operatorname{div}(A \nabla u_t) + \langle B, \nabla u_t \rangle + c u_t \\
&\qquad = f_t + \operatorname{div}(A_t \nabla u) - \langle B_t, \nabla u \rangle - c_t u. \qquad (41)
\end{aligned}
$$

Multiplying this equation by u_t and for a.e. $0 < s < T$, integrating over Ω and integrating by parts, we deduce

$$
\begin{aligned}
0 = & \int_\Omega \langle u_{tt}, u_t \rangle + \int_\Omega \langle A \nabla u_t, \nabla u_t \rangle + \int_\Omega \langle B, \nabla u_t \rangle u_t + \int_\Omega c \left| u_t \right|^2 \\
& - \int_\Omega f_t u_t - \int_\Omega \langle A_t \nabla u, \nabla u_t \rangle - \int_\Omega \langle B_t \nabla u \rangle u_t - \int_\Omega c_t u u_t \\
\geq & \frac{1}{2} \frac{d}{dt} \left(\int_\Omega \left| u_t \right|^2 \right) + \lambda \int_\Omega \left| \nabla u_t \right|^2 - \left(|I_1| + |I_2| + |I_3| + |I_4| + |I_5| + |I_6| \right).
\end{aligned}
$$

Now, we have

$$
\begin{aligned}
|I_1| = \left| \int_\Omega \langle B, \nabla u_t \rangle u_t \right| & \leq \int_\Omega \left| \langle B, \nabla u_t \rangle u_t \right| \\
& \leq \|B\|_{L^\infty} \int_\Omega \left| \nabla u_t \right| \left| u_t \right| \leq \varepsilon \int_\Omega \left| \nabla u_t \right|^2 + C \int_\Omega \left| u_t \right|^2.
\end{aligned}
$$

$$
|I_2| = \left| \int_\Omega c \left| u_t \right|^2 \right| \leq \|c\|_{L^\infty} \int_\Omega \left| u_t \right|^2.
$$

$$
|I_3| = \left| \int_\Omega f_t u_t \right| \leq \int_\Omega |f_t| \left| u_t \right| \leq C \int_\Omega \left| u_t \right|^2 + C \int_\Omega |f_t|^2.
$$

$$
|I_4| = \left| \int_\Omega \langle A_t \nabla u, \nabla u_t \rangle \right| \leq \|A_t\|_{L^\infty} \int_\Omega |\nabla u| \left| \nabla u_t \right| \leq \varepsilon \int_\Omega \left| \nabla u_t \right|^2 + C \int_\Omega |\nabla u|^2.
$$

$$
\begin{aligned}
|I_5| = \left| \int_\Omega \langle B_t, \nabla u \rangle u_t \right| & \leq \int_\Omega \left| \langle B_t, \nabla u \rangle u_t \right| \\
& \leq \|B_t\|_{L^\infty} \int_\Omega |\nabla u| \left| u_t \right| \leq \varepsilon \int_\Omega \left| \nabla u_t \right|^2 + C \int_\Omega \left| u_t \right|^2.
\end{aligned}
$$

$$
\begin{aligned}
|I_6| = \left| \int_\Omega c_t u u_t \right| & \leq \|c_t\|_{L^\infty} \int_\Omega \left| u_t \right| |u| \\
& \leq C \int_\Omega |u|^2 + C \int_\Omega \left| u_t \right|^2 \leq C \int_\Omega |\nabla u|^2 + C \int_\Omega \left| u_t \right|^2,
\end{aligned}
$$

where we have used the Poincaré inequality in the last line. Choosing $\varepsilon > 0$ small enough, we deduce

$$
\frac{d}{dt} \left(\int_\Omega \left| u_t \right|^2 \right) + \lambda \int_\Omega \left| \nabla u_t \right|^2 \leq C_1 \left(\int_\Omega |\nabla u|^2 + \int_\Omega \left| u_t \right|^2 + \int_\Omega |f_t|^2 \right).
$$

for some constant $C_1 > 0$. But this implies

$$\begin{aligned}\frac{d}{dt}\left(e^{-C_1 t}\int_\Omega |u_t|^2\right) &= e^{-C_1 t}\left[\frac{d}{dt}\left(\int_\Omega |u_t|^2\right) - C_1\int_\Omega |u_t|^2\right]\\ &\le e^{-C_1 t}\left[C_1\left(\int_\Omega |f_t|^2 + \int_\Omega |\nabla u|^2\right) - \lambda\int_\Omega |\nabla u_t|^2\right]\end{aligned}$$

So we arrive at

$$\begin{aligned}\frac{d}{dt}\left(e^{-C_1 t}\int_\Omega |u_t|^2\right) + \lambda e^{-C_1 t}\int_\Omega |\nabla u_t|^2 &\le C_1 e^{-C_1 t}\left(\int_\Omega |f_t|^2 + \int_\Omega |\nabla u|^2\right)\\ &\le C_1\left(\int_\Omega |f_t|^2 + \int_\Omega |\nabla u|^2\right)\end{aligned}$$

as $C_1 t > 0$ and thus $e^{-C_1 t} < 1$. Integrating with respect to t from 0 to s, where $0 < s < T$, we deduce

$$\begin{aligned}&e^{-C_1 s}\int_\Omega |u_t(s)|^2 - \int_\Omega |u_t(0)|^2 + \lambda\int_0^s e^{-C_1 t}\int_\Omega |\nabla u_t(t)|^2 \ \mathrm{d}t\\ &\qquad\le C\left(\int_0^s\int_\Omega |f_t|^2 \ \mathrm{d}t + \int_0^s\int_\Omega |\nabla u|^2 \ \mathrm{d}t\right)\\ &\qquad\le C\left(\int_0^T\int_\Omega |f_t|^2 \ \mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \ \mathrm{d}t\right).\\ &\qquad= C\left(\|f_t\|^2_{L^2((0,T);L^2(\Omega))} + \|\nabla u\|^2_{L^2\left((0,T);H_0^1(\Omega)\right)}\right).\end{aligned}$$

Thus, using the obvious estimate $e^{-C_1 T} < e^{-C_1 t}$ for all $0 \le t \le T$, we have

$$\begin{aligned}&e^{-C_1 s}\int_\Omega |u_t(s)|^2 + \lambda e^{-C_1 T}\int_0^s\int_\Omega |\nabla u_t(t)|^2 \ \mathrm{d}t\\ &\qquad\le C\left(\int_0^T\int_\Omega |f_t|^2 \ \mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \ \mathrm{d}t\right) + \int_\Omega |u_t(0)|^2.\end{aligned}$$

Thus,

$$\begin{aligned}&\int_\Omega |u_t(s)|^2 + \lambda e^{-C_1 T}\int_0^s\int_\Omega |\nabla u_t(t)|^2 \ \mathrm{d}t\\ &\qquad\le \int_\Omega |u_t(s)|^2 + \lambda e^{-C_1(T-s)}\int_0^s\int_\Omega |\nabla u_t(t)|^2 \ \mathrm{d}t\\ &\qquad\le Ce^{C_1 s}\left(\int_0^T\int_\Omega |f_t|^2 \ \mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \ \mathrm{d}t + \int_\Omega |u_t(0)|^2\right)\\ &\qquad\le Ce^{C_1 T}\left(\int_0^T\int_\Omega |f_t|^2 \ \mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \ \mathrm{d}t + \int_\Omega |u_t(0)|^2\right).\end{aligned}$$

Now, we want to estimate the last integral on the right. From the equation, passing to the limits as $t \to 0$, we have

$$u_t(0) = f(0) - \mathcal{L}_0 u(0) = f(0) - \mathcal{L}_0 g,$$

where the uniformly elliptic spatial differential operator $\mathcal{L}_0$ is defined by

$$\mathcal{L}_0 w := -\operatorname{div}(A(0)\nabla w) + \langle B(0), \nabla w\rangle + c(0)\,w.$$

Since this is a second order operator involving only the spatial variables, it is easy to see that we have the estimate

$$\int_\Omega |\mathcal{L}_0 g|^2 \leq C\, \|g\|^2_{H^2(\Omega)}.$$

Thus, we have

$$\int_\Omega |u_t(0)|^2 \leq \int_\Omega |f(0)|^2 + \int_\Omega |\mathcal{L}_0 g(0)|^2 \leq \int_\Omega |f(0)|^2 + C\,\|g\|^2_{H^2(\Omega)}.$$

Plugging this estimate back into the last one, we deduce

$$\begin{aligned}
&\int_\Omega |u_t(s)|^2 + \lambda e^{-C_1 T}\int_0^s \int_\Omega |\nabla u_t(t)|^2 \,\mathrm{d}t \\
&\quad\leq C e^{C_1 T}\left(\int_0^T\int_\Omega |f_t|^2 \,\mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \,\mathrm{d}t + \int_\Omega |u_t(0)|^2\right) \\
&\quad\leq C e^{C_1 T}\left(\int_0^T\int_\Omega |f_t|^2 \,\mathrm{d}t + \int_0^T\int_\Omega |\nabla u|^2 \,\mathrm{d}t + \int_\Omega |f(0)|^2 + \|g\|^2_{H^2(\Omega)}\right) \\
&\quad\leq C e^{C_1 T}\left(\|f\|^2_{H^1((0,T);L^2(\Omega))} + \int_0^T\int_\Omega |\nabla u|^2 \,\mathrm{d}t + \|g\|^2_{H^2(\Omega)}\right).
\end{aligned}$$

Taking supremum over $0 \leq s \leq T$, we deduce

$$\begin{aligned}
&\sup_{0\leq t\leq T} \|u_t(t)\|^2_{L^2(\Omega)} + \|u_t\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} \\
&\qquad\qquad \leq C\left(\|f\|^2_{H^1((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H^1_0(\Omega)\right)} + \|g\|^2_{H^2(\Omega)}\right). \qquad (42)
\end{aligned}$$

Now multiplying the equation (41) by $\phi \in L^2\left((0,T);H^1_0(\Omega)\right)$ and integrating over Ω and integrating by parts, we have

$$\begin{aligned}
0 = &\int_\Omega u_{tt}\phi + \int_\Omega \langle A\nabla u_t, \nabla\phi\rangle + \int_\Omega \langle B, \nabla u_t\rangle\,\phi + \int_\Omega c u_t\phi - \int_\Omega f_t\phi \\
&\qquad + \int_\Omega \langle A_t\nabla u, \nabla\phi\rangle + \int_\Omega \langle B_t, \nabla u\rangle\,\phi + \int_\Omega c_t u\phi.
\end{aligned}$$

Hence, we deduce

$$\left| \int_\Omega u_{tt} \phi \right| \le |I_1| + |I_2| + |I_3| + |I_4| + |I_5| + |I_6| + |I_7| \,,$$

where we have

$$\begin{aligned}
|I_1| &= \left| \int_\Omega \langle A \nabla u_t, \nabla \phi \rangle \right| \le \|A\|_{L^\infty} \|\nabla u_t\|_{L^2(\Omega)} \|\nabla \phi\|_{L^2(\Omega)} \,, \\
|I_2| &= \left| \int_\Omega \langle B, \nabla u_t \rangle \phi \right| \le \|B\|_{L^\infty} \int_\Omega |\nabla u_t| \, |\phi| \le \|B\|_{L^\infty} \|\nabla u_t\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} \,, \\
|I_3| &= \left| \int_\Omega c u_t \phi \right| \le \|c\|_{L^\infty} \int_\Omega |u_t| \, |\phi| \le \|c\|_{L^\infty} \|u_t\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} \,, \\
|I_4| &= \left| \int_\Omega f_t \phi \right| \le \int_\Omega |f_t| \, |\phi| \le \|f_t\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} \,, \\
|I_5| &= \left| \int_\Omega \langle A_t \nabla u, \nabla \phi \rangle \right| \le \|A_t\|_{L^\infty} \|\nabla u\|_{L^2(\Omega)} \|\nabla \phi\|_{L^2(\Omega)} \,, \\
|I_6| &= \left| \int_\Omega \langle B_t, \nabla u \rangle \phi \right| \le \|B_t\|_{L^\infty} \|\nabla u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} \,, \\
|I_7| &= \left| \int_\Omega c_t u \phi \right| \le \|c_t\|_{L^\infty} \|u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} \,.
\end{aligned}$$

This implies,

$$\left| \int_\Omega u_{tt} \phi \right| \le C \left(\|u_t(t)\|_{H_0^1(\Omega)} + \|f_t(t)\|_{L^2(\Omega)} \right) \|\phi(t)\|_{H_0^1(\Omega)}$$

By the dual characterization of the $H^{-1}(\Omega)$ norm, this means

$$\|u_{tt}(t)\|_{H^{-1}(\Omega)} \le C \left(\|u_t(t)\|_{H_0^1(\Omega)} + \|f_t(t)\|_{L^2(\Omega)} \right).$$

Squaring both sides and integrating with respect to t from 0 to T, we derive

$$\|u_{tt}\|^2_{L^2((0,T);H^{-1}(\Omega))} \le C \left(\|u_t\|^2_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|f\|^2_{H^1((0,T);L^2(\Omega))} \right).$$

Combined with (42), this implies the estimate

$$\|u_{tt}\|^2_{L^2((0,T);H^{-1}(\Omega))} \le C \left(\|f\|^2_{H^1((0,T);L^2(\Omega))} + \|u\|^2_{L^2\left((0,T);H_0^1(\Omega)\right)} + \|g\|^2_{H^2(\Omega)} \right). \quad (43)$$

Now Sobolev embedding for Sobolev spaces involving time implies (38)

Now (42), (43) and the theory of time-dependent Sobolev spaces again implies that

$$\begin{cases} u_t \in L^2\left((0,T); H_0^1(\Omega)\right) \\ u_{tt} \in L^2\left((0,T); H^{-1}(\Omega)\right) \end{cases} \Rightarrow u_t \in C\left([0,T]; L^2(\Omega)\right)$$

along with the estimate

$$\begin{aligned}
&\max_{0\leq t\leq T}\left\|u_t\left(t\right)\right\|_{L^2(\Omega)}\\
&\qquad\leq C\left(\left\|u_t\right\|_{L^2\left((0,T);H_0^1(\Omega)\right)}+\left\|u_{tt}\right\|_{L^2((0,T);H^{-1}(\Omega))}\right)\\
&\qquad\leq C\left(\left\|f\right\|_{H^1((0,T);L^2(\Omega))}+\left\|u\right\|_{L^2\left((0,T);H_0^1(\Omega)\right)}+\left\|g\right\|_{H^2(\Omega)}\right). \qquad (44)
\end{aligned}$$

But now returning to (39) and taking supremum over $0\leq s\leq T$, we get

$$\sup_{0\leq t\leq T}\left\|u\left(t\right)\right\|_{H^2(\Omega)}\leq C\sup_{0\leq t\leq T}\left(\left\|u\left(t\right)\right\|_{H_0^1(\Omega)}+\left\|f\left(t\right)\right\|_{L^2(\Omega)}+\left\|u_t\left(t\right)\right\|_{L^2(\Omega)}\right). \qquad (45)$$

Now the estimates (25), (38), (26), (42), (43) and (45) together implies (37). This completes the proof. □

These estimates can be iterated to gain higher and higher regularity if A, B, c, f, g are sufficiently regular and f and g satisfy the compatibility conditions. See Theorem 6, Chapter 7 in [1].

References

[1] Evans, L. C. *Partial differential equations*, vol. 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1998.