Chapter 6

The heat equation

In this chapter, we will present a short and far from exhaustive theoretical study of the heat equation, then describe and analyze a few approximation methods. We will mostly work in one dimension of space, some of the results having an immediate counterpart in higher dimensions, others not.

Let Ω be an open subset of $\mathbb{R}^d$, $T \in R_+$. We note $Q = \Omega \times]0,T[$. Recall that the heat equation is

$$\frac{\partial u}{\partial t} - \Delta u = f \text{ in } Q,$$

together with an initial condition

$$u(x,0) = u_0(x) \text{ in } \Omega,$$

and boundary values, for instance Dirichlet boundary values

$$u(x,t) = g(x,t) \text{ on } \partial\Omega \times]0,T[,$$

where f, u_0 and g are given functions.

When $d = 1$, we take $\Omega =]0,1[$ without loss of generality for a bounded Ω and the problem reads

$$\begin{cases} \dfrac{\partial u}{\partial t} - \dfrac{\partial^2 u}{\partial x^2} = f \text{ in } Q, \\ u(x,0) = u_0(x) \text{ in } \Omega, \\ u(0,t) = g(0,t), u(1,t) = g(1,t) \text{ in }]0,T[, \end{cases} \tag{6.1}$$

where $g(0,\cdot)$ and $g(1,\cdot)$ are two given scalar valued functions defined on $]0,T[$.

6.1 The maximum principle for the heat equation

We have seen a version of the maximum principle for a second order elliptic equation, in one dimension of space. Parabolic equations also satisfy their own version of the maximum principle.

Proposition 6.1.1 *We assume that u is a solution of problem* (6.1) *that belongs to $C^0(\bar{Q}) \cap C^2(Q \cup (\Omega \times \{T\}))$. If $f \geq 0$ in Q, then u attains its minimum on $(\Omega \times \{0\}) \cup (\partial\Omega \times [0,T])$.*

Proof. Let us first assume that $f > 0$ on $Q \cup (\Omega \times \{T\})$. The set $\bar{Q}$ is compact and u is continuous on $\bar{Q}$, thus it attains its minimum somewhere in $\bar{Q}$, say at point (x_0,t_0).

If $(x_0,t_0) \in Q$, then $\frac{\partial u}{\partial t}(x_0,t_0) = 0$ and $\frac{\partial^2 u}{\partial x^2}(x_0,t_0) \geq 0$ since u is C^2 in a neighborhood of (x_0,t_0), so that $\big(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\big)(x_0,t_0) \leq 0$, which contradicts $f(x_0,t_0) > 0$.

Therefore $(x_0,t_0) \in \partial Q = \big((\Omega \times \{0\}) \cup (\partial\Omega \times [0,T])\big) \cup (\Omega \times \{T\})$. Assume that $(x_0,t_0) \in \Omega \times \{T\}$, *i.e.*, that $x_0 \in \Omega$ and $t_0 = T$. It follows again, since as a function in the variable x for $t = T$, u is also C^2, that $\frac{\partial^2 u}{\partial x^2}(x_0,T) \geq 0$ so that $\frac{\partial u}{\partial t}(x_0,T) = \frac{\partial^2 u}{\partial x^2}(x_0,T) + f(x_0,T) > 0$. Thus there exists $t < T$ such that $u(x_0,t) < u(x_0,T)$, which is consequently not a minimum for u.

The only possibility left is that $(x_0,t_0) \in K = (\Omega \times \{0\}) \cup (\partial\Omega \times [0,T])$.

Consider now the case $f \geq 0$. Let $\varepsilon > 0$ and $u_\varepsilon(x,t) = u(x,t) + \varepsilon x(1-x)$. In particular $u(x,t) \leq u_\varepsilon(x,t)$ in $\bar{Q}$. We have

$$\frac{\partial u_\varepsilon}{\partial t} - \frac{\partial^2 u_\varepsilon}{\partial x^2} = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} + 2\varepsilon = f + 2\varepsilon > 0.$$

By the previous argument, u_ε attains its minimum at a point $(x_\varepsilon,t_\varepsilon)$ of K. We have

$$u(x_0,t_0) \leq u(x_\varepsilon,t_\varepsilon) \leq u_\varepsilon(x_\varepsilon,t_\varepsilon) \leq u_\varepsilon(x_0,t_0) = u(x_0,t_0) + \varepsilon x_0(1-x_0).$$

In particular,

$$u(x_0,t_0) \leq u(x_\varepsilon,t_\varepsilon) \leq u(x_0,t_0) + \varepsilon x_0(1-x_0).$$

We now let $\varepsilon \to 0$. Since K is compact, we may extract a subsequence such that $(x_\varepsilon,t_\varepsilon) \to (\bar{x},\bar{t}) \in K$. Passing to the limit in the above inequalities and using the continuity of u, we obtain

$$u(x_0,t_0) = u(\bar{x},\bar{t}),$$

with $(\bar{x},\bar{t}) \in K$ and the minimum is attained on K. □

Remark 6.1.1 The meaning of the maximum principle is that the minimum temperature is either attained at $t = 0$ or on the boundary of Ω at some other time $t \in]0,T]$, but not in $\Omega \times]0,T]$. It is also valid in any dimension d, using the same proof and the maximum principle for the Laplacian, that we have not proved here. The result cannot be refined further since $u = 0$ attains its minimum at any point in K. □

The maximum principle has many consequences, some of which we now list.

Corollary 6.1.1 *If $f \geq 0$, $g \geq 0$ and $u_0 \geq 0$, then $u \geq 0$ in $\bar{Q}$.*

Proof. Clear since the minimum of u is either of the form $u_0(x_0)$ or $g(x_0,t_0)$. □

Remark 6.1.2 This form of the maximum principle is again a monotonicity result. The interesting physical interpretation is that if you heat up a room, the walls are kept at a positive temperature and the initial temperature is positive, then the temperature in the room stays positive everywhere and at any time. □

We also have a stability result in the C^0 norm.

Corollary 6.1.2 *If $f = 0$ and $g = 0$, then $\|u\|_{C^0(\bar{Q})} = \|u_0\|_{C^0(\bar{\Omega})}$.*

Proof. Let $v_+ = u + \|u_0\|_{C^0(\bar{\Omega})}$. We have

$$\frac{\partial v_+}{\partial t} - \frac{\partial^2 v_+}{\partial x^2} = 0,$$

since $f = 0$,

$$v_+(0,t) = v_+(1,t) = \|u_0\|_{C^0(\bar{\Omega})} \geq 0,$$

since $g = 0$ and

$$v_+(x,0) = u_0 + \|u_0\|_{C^0(\bar{\Omega})} \geq 0.$$

By Corollary 6.1.1, $v_+ \geq 0$ in $\bar{Q}$, or in other words $u(x,t) \geq -\|u_0\|_{C^0(\bar{\Omega})}$ in $\bar{Q}$. Changing u in $-u$, we also have $u(x,t) \leq \|u_0\|_{C^0(\bar{\Omega})}$ in $\bar{Q}$, hence the result. □

Such a stability result immediately entails a uniqueness result.

Proposition 6.1.2 *Problem* (6.1) *has at most one solution in $C^0(\bar{Q}) \cap C^2(Q)$.*

Proof. Indeed, if u_1 and u_2 are two solutions, then $v = u_1 - u_2$ satisfies the hypotheses of Corollary 6.1.2 with $u_0 = 0$ on $\Omega \times [0,T-\eta])$ for all $\eta > 0$. □

6.2 Construction of a regular solution

We will see several different ways of constructing solutions to the heat equation. Let us start with an elementary construction using Fourier series. It should be recalled that Joseph Fourier invented what became Fourier series in the 1800s, exactly for the purpose of solving the heat equation.

We consider the case when $f = 0$, no heat source, and $g = 0$, homogeneous Dirichlet boundary condition, the only nonzero data being the initial condition u_0.

Proposition 6.2.1 *Let $u_0 \in C^0([0,1])$ be piecewise C^1 and such that $u_0(0) = u_0(1) = 0$. Then we have*

$$u_0(x) = \sum_{k=1}^{+\infty} b_k \sin(k\pi x)$$

for all $x \in [0,1]$, with $\sum_{k=1}^{+\infty} |b_k| < +\infty$.

Proof. We first extend u_0 by imparity by setting $\widetilde{u}_0(x) = u_0(x)$ for $x \in [0,1]$ and $\widetilde{u}_0(x) = -u_0(-x)$ for $x \in [-1,0[$. The resulting function is odd and continuous on $[-1,1]$ by construction since $u_0(0) = 0$ and still piecewise C^1.

Secondly, we extend $\widetilde{u}_0$ to $\mathbb{R}$ by 2-periodicity by setting $\widetilde{\widetilde{u}}_0(x) = \widetilde{u}_0(x - 2\lfloor \frac{x+1}{2} \rfloor)$. This function is continuous since $\widetilde{u}_0(-1) = \widetilde{u}_0(1) = 0$, piecewise C^1 and 2-periodic by construction. Therefore, by Dirichlet's theorem, it can be expanded in Fourier series

$$\widetilde{\widetilde{u}}_0(x) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} a_k \cos(k\pi x) + \sum_{k=1}^{+\infty} b_k \sin(k\pi x),$$

with $\sum_{k=1}^{+\infty}(|a_k| + |b_k|) < +\infty$, hence the series is normally convergent. Now $\widetilde{\widetilde{u}}_0$ is also odd by construction, so that all a_k Fourier coefficients vanish. Restricting the above expansion to $x \in [0,1]$, we obtain the result. □

Theorem 6.2.1 *Let u_0 be as above. Then the function defined by*

$$u(x,t) = \sum_{k=1}^{+\infty} b_k \sin(k\pi x) e^{-k^2\pi^2 t}$$

belongs to $C^0(\mathbb{R} \times [0,+\infty[) \cap C^\infty(\mathbb{R} \times]0,+\infty[)$. Its restriction to $\bar{Q}$ solves problem (6.1) with data $f = 0$, $g = 0$.

Proof. We first need to show that the series above is convergent in some sense and that its sum belongs to the function spaces indicated in the theorem. Normal convergence on $\mathbb{R} \times [0,+\infty[$ is obvious since $|b_k \sin(k\pi x)e^{-k^2\pi^2 t}| \leq |b_k|$, thus u exists and is continuous $\mathbb{R} \times [0,+\infty[$.

Let us now consider differentiability. Now if u is supposed to coincide with u_0 at $t = 0$, and u_0 is only piecewise C^1, we cannot expect u to be C^∞ up to $t = 0$, hence the exclusion of $t = 0$. In order to use theorems on the differentiation of series, we actually need to stay away from $t = 0$ as will become clear in the proof. Let us thus chose $\varepsilon > 0$ and work for $t \geq \varepsilon$. It is convenient to notice that

$$\sin(k\pi x)e^{-k^2\pi^2 t} = \Im(e^{ik\pi x - k^2\pi^2 t}).$$

Therefore, for any natural integers p and q, we have

$$\frac{\partial^{p+q}}{\partial t^p \partial x^q}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big) = (k\pi)^p(-k^2\pi^2)^q \Im(i^p e^{ik\pi x - k^2\pi^2 t}).$$

Thus

$$\begin{aligned}\Big|b_k \frac{\partial^{p+q}}{\partial t^p \partial x^q}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big)\Big| &\leq |b_k|\pi^{p+2q}k^{p+2q}e^{-k^2\pi^2 t} \\ &\leq |b_k|\pi^{p+2q}k^{p+2q}e^{-k^2\pi^2 \varepsilon}\end{aligned}$$

for $t \geq \varepsilon$. Since $b_k = \frac{1}{2}\int_{-1}^{1} \sin(k\pi x)\widetilde{u}_0(x)\,dx$, we have $|b_k| \leq \|u_0\|_{C^0}$, thus

$$\Big|b_k \frac{\partial^{p+q}}{\partial t^p \partial x^q}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big)\Big| \leq C_{p,q}k^{p+2q}e^{-k^2\pi^2 \varepsilon},$$

because $t \geq \varepsilon$. The right-hand side is the general term of a convergent series due to the $e^{-k^2\pi^2\varepsilon}$ term with $\varepsilon > 0$, thus the left-hand side is the general term of a normally, thus uniformly convergent series, for any p and q. Therefore, u is of class C^∞ on $\mathbb{R} \times]\varepsilon,+\infty[$, for all $\varepsilon > 0$, thus is in $C^\infty(\mathbb{R} \times]0,+\infty[)$. Moreover, we have

$$\frac{\partial^{p+q}u}{\partial t^p \partial x^q}(x,t) = \sum_{k=1}^{+\infty} b_k \frac{\partial^{p+q}}{\partial t^p \partial x^q}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big)$$

for all $(x,t) \in \mathbb{R} \times]0,+\infty[$. In particular, we have

$$\frac{\partial u}{\partial t}(x,t) = \sum_{k=1}^{+\infty} b_k \frac{\partial}{\partial t}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big) = -\sum_{k=1}^{+\infty} b_k k^2\pi^2 \sin(k\pi x)e^{-k^2\pi^2 t}$$

and

$$\frac{\partial^2 u}{\partial x^2}(x,t) = \sum_{k=1}^{+\infty} b_k \frac{\partial^2}{\partial x^2}\big(\sin(k\pi x)e^{-k^2\pi^2 t}\big) = -\sum_{k=1}^{+\infty} b_k k^2\pi^2 \sin(k\pi x)e^{-k^2\pi^2 t}$$

so that
$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$$
and u solves the heat equation.

For the boundary conditions, we note that for all integers $k \geq 1$, $\sin(k\pi \times 0) = \sin(k\pi \times 1) = 0$, so that
$$u(0,t) = u(1,t) = 0$$
for all $t \in \mathbb{R}_+$. Finally,
$$u(x,0) = \sum_{k=1}^{+\infty} b_k \sin(k\pi x) e^{-k^2\pi^2 \times 0} = \sum_{k=1}^{+\infty} b_k \sin(k\pi x) = u_0(x),$$
and the initial condition is satisfied. □

Remark 6.2.1 It is worth noticing that both boundary conditions and initial condition make sense because u is continuous. Moreover, the regularity of u is such that the previous uniqueness result applies, thus we have found the one and only one solution in that class.

An important feature of the heat equation, and more generally of parabolic equations, is that whatever regularity u_0 may have, if $f = 0$, then the solution u becomes C^∞ instantly for $t > 0$. This is a *smoothing* effect.

For $t \geq 0$ fixed, the series that gives $x \mapsto u(x,t)$ is also the Fourier series of the odd and 2-periodic $\mathbb{R}$-extension of this function. The exponential term $e^{-k^2\pi^2 t}$ makes these Fourier coefficients decrease rapidly, which indicates that the sum is smooth (with respect to x), but we knew that already, in x and t.

The smoothing effect also shows why the backward heat equation is ill-posed. Indeed, there can be no solution to the backward heat equation with an initial condition that is not C^∞, since an initial condition for the backward heat equation is a final condition for the forward heat equation. It is not even clear that all C^∞ functions can be reached by the evolution of the heat equation. Therefore, time is irreversible in the heat equation.

We can see the same effect in the series, since for $t < 0$, $-k^2\pi^2 t > 0$ and the exponential term instead of ensuring extremely fast convergence of the series, thus smoothing it at the same time, becomes explosive. The only way the series can converge for $t < 0$ is for the Fourier coefficients b_k of the initial condition to be rapidly decreasing, so as to compensate for the exponential term. Again, a function with rapidly decreasing Fourier coefficients is very smooth. □

The above solution of the heat equation exhibits rapid uniform decay in time.

Proposition 6.2.2 *There exists a constant C such that*

$$|u(x,t)| \leq Ce^{-\pi^2 t}.$$

In particular, $u(x,t) \to 0$ *when* $t \to +\infty$, *uniformly with respect to x.*

Proof. Indeed, $e^{-k^2\pi^2 t} \leq e^{-\pi^2 t}$ for all k and all $t \geq 0$, so that

$$|u(x,t)| \leq \sum_{k=1}^{+\infty} |b_k| e^{-k^2\pi^2 t} \leq e^{-\pi^2 t} \sum_{k=1}^{+\infty} |b_k|,$$

hence the result since $\sum_{k=1}^{+\infty} |b_k| < +\infty$. □

Remark 6.2.2 If we remember the physical interpretation of the heat equation, keeping the walls of a room at 0 degree is tantamount to having paper-thin walls and a huge ice cube surrounding the room. If there is no heat source inside the room, it is not contrary to physical intuition that the temperature inside should drop to 0 degree pretty quickly, if it was positive at $t = 0$. All the heat inside flows outside (the heat flux is proportional to the opposite of the temperature gradient).□

Apart from proving the existence of a solution in a particular case, the Fourier series expansion is also a very precise numerical method, provided the Fourier coefficients of the initial condition are known with good accuracy. In effect, we have a coarse error estimate

$$\Big| u(x,t) - \sum_{k=1}^{N} b_k \sin(k\pi x) e^{-k^2\pi^2 t} \Big| \leq \Big(\sum_{k=N+1}^{+\infty} |b_k| \Big) e^{-(N+1)^2\pi^2 t},$$

so that truncating the series with only a few terms can be expected to give excellent precision as soon as $t > 0$ is noticeably nonzero, and depending on the Fourier coefficients of u_0 in the neighborhood of $t = 0$. Of course, the sine and exponential functions are already implemented in all computer languages.

Let us take a simple continuous, piecewise C^1 initial condition such as this:

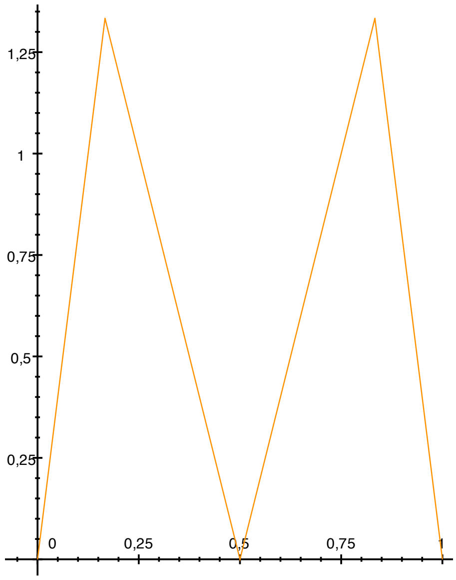
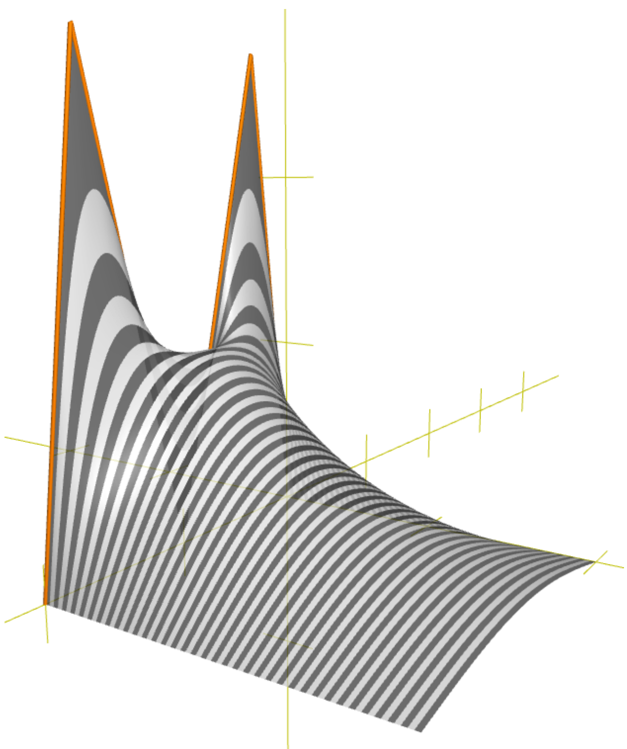
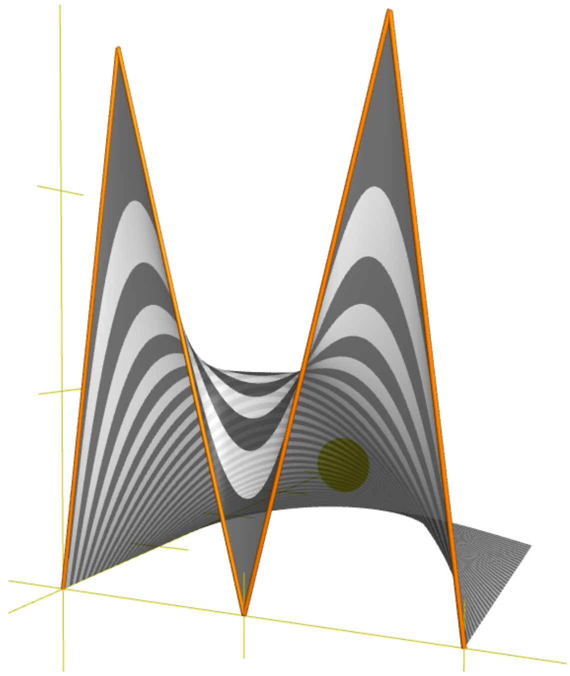
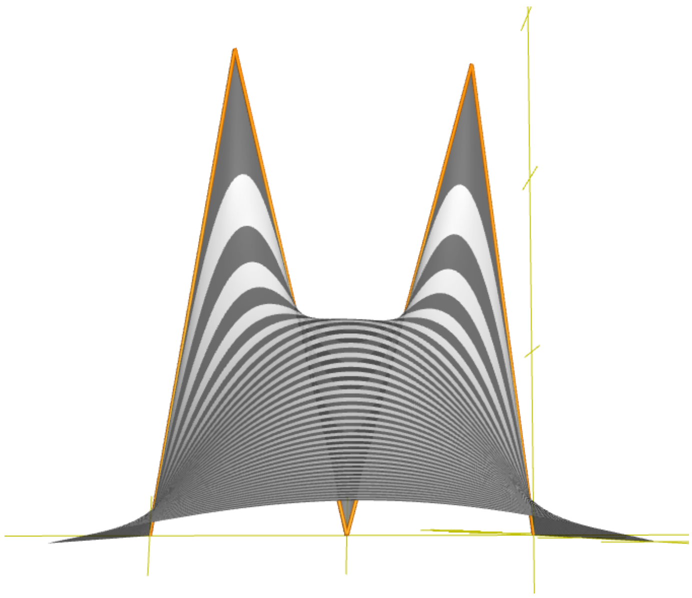


Figure 1. An admissible initial value u_0.

Then six terms in the Fourier series already give a very good approximation. Following are several views of the graph of u plotted in (x,t) space. The grey stripes show the graphs of $x \mapsto u(x,t)$ for a discrete sample of values of t.







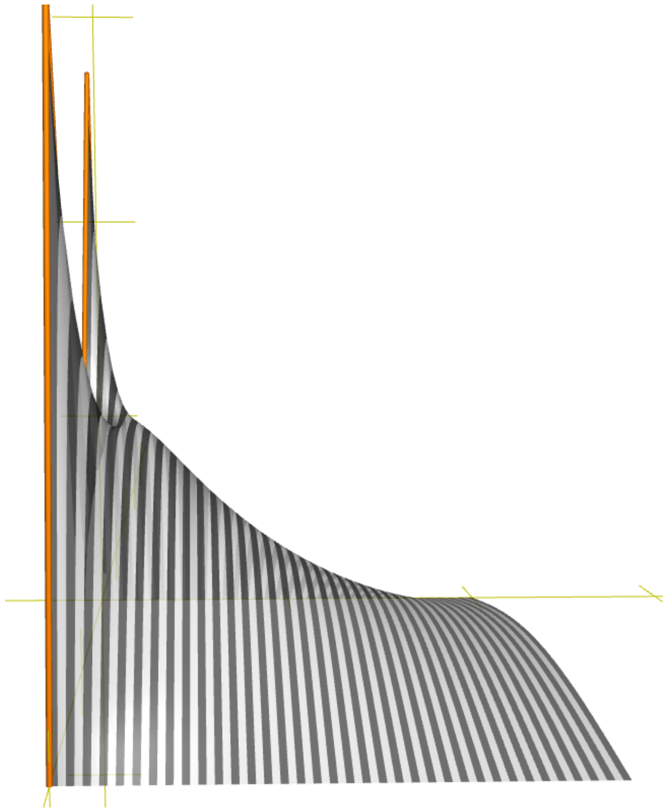


Figure 2. Various views of the corresponding solution u, using Fourier series.

We see the exponential decay in time, the smoothing effect, and also the fact that the first nonzero term in the series rapidly becomes dominant as t increases, as can be expected from the exponential terms. Note also the continuity as $t \to 0^+$, and the fact that the time derivative goes to $\pm\infty$ when (x,t) tends to a point $(x_0,0)$ where the second space derivative of the initial condition is in a sense infinite[1], *i.e.* the first space derivative is discontinuous. We also see that the minimum is attained where the maximum principle says it must be attained[2].

The Fourier expansion even gives quite good results for cases that are not covered by the preceding analysis, for instance for an initial condition that does not satisfy the Dirichlet boundary condition, such as $u_0(x) = 1$! Here with 20 terms in the series, and of course a Gibbs phenomenon around the discontinuities.

[1]Or more accurately a Dirac mass.
[2]Which is reassuring.

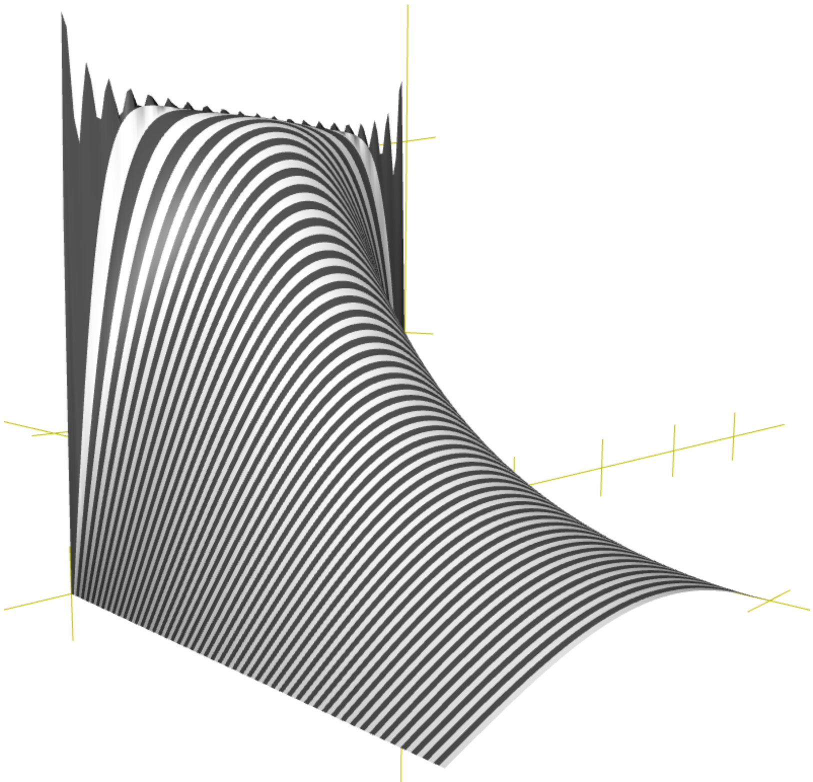


Figure 3. Fourier series and discontinuous solutions, 20 terms.

The same with 100 terms, the Gibbs phenomenon is still there, see Figures 5 and 6, but does not show on Figure 4 for sampling reasons: it occurs on a length scale that is too small to be captured by the graphics program, recall that high frequency oscillations in space are damped extremely rapidly in time by the exponential term.

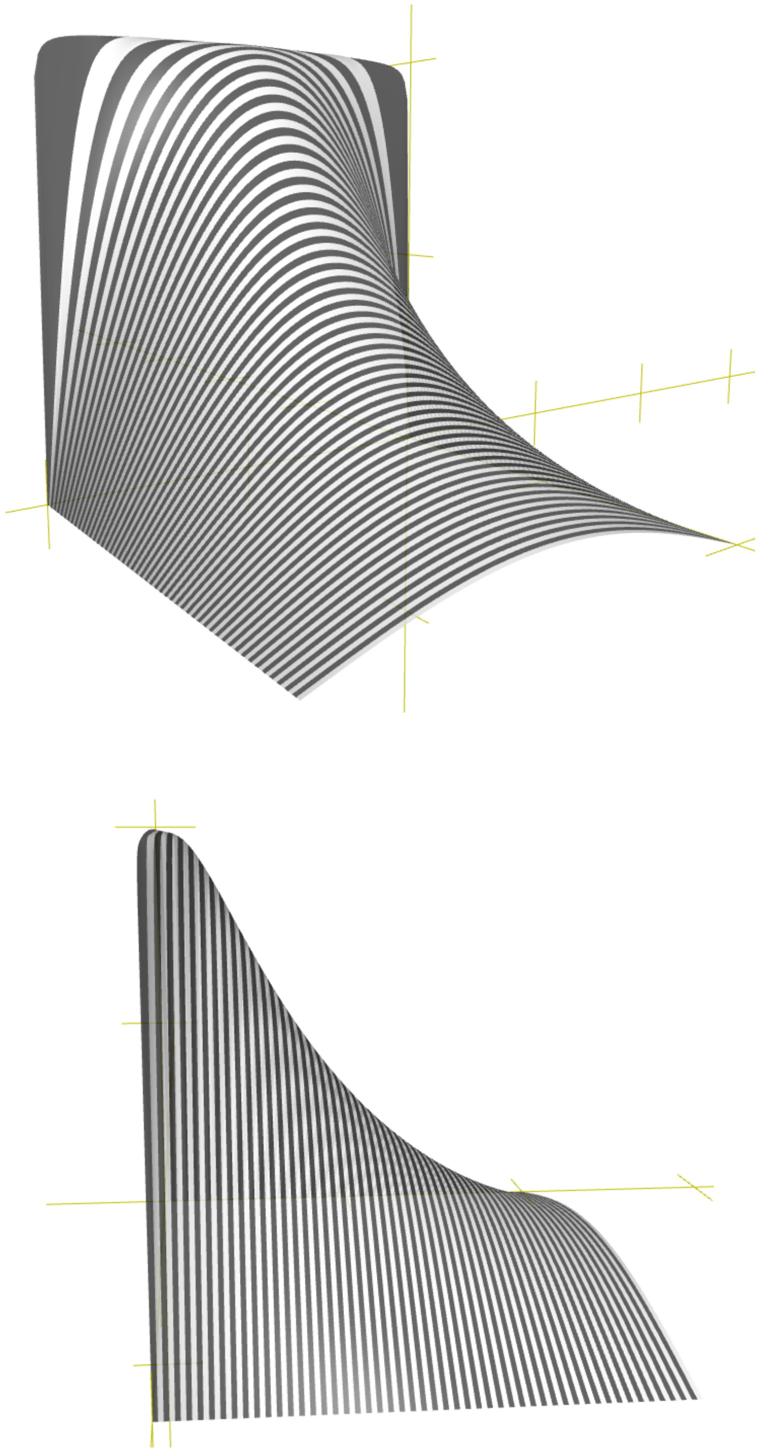


Figure 4. Fourier series and discontinuous solutions, 100 terms.

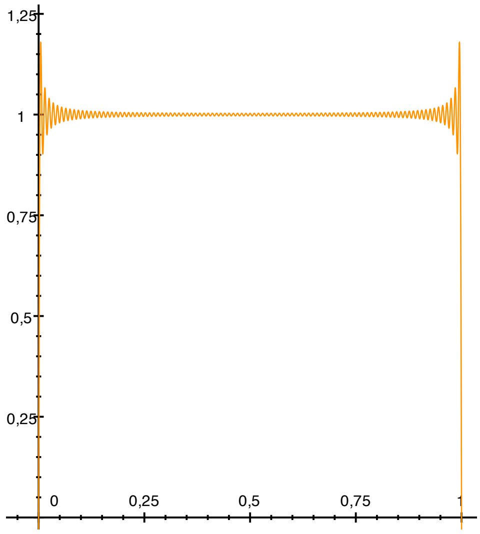


Figure 5. One hundred terms in the Fourier series of $\widetilde{u}_0$, with Gibbs phenomenon around 0 and 1.

Here are a few close-ups of the 100 term Fourier series expansion of u near $(x,t) = (0,0)$.

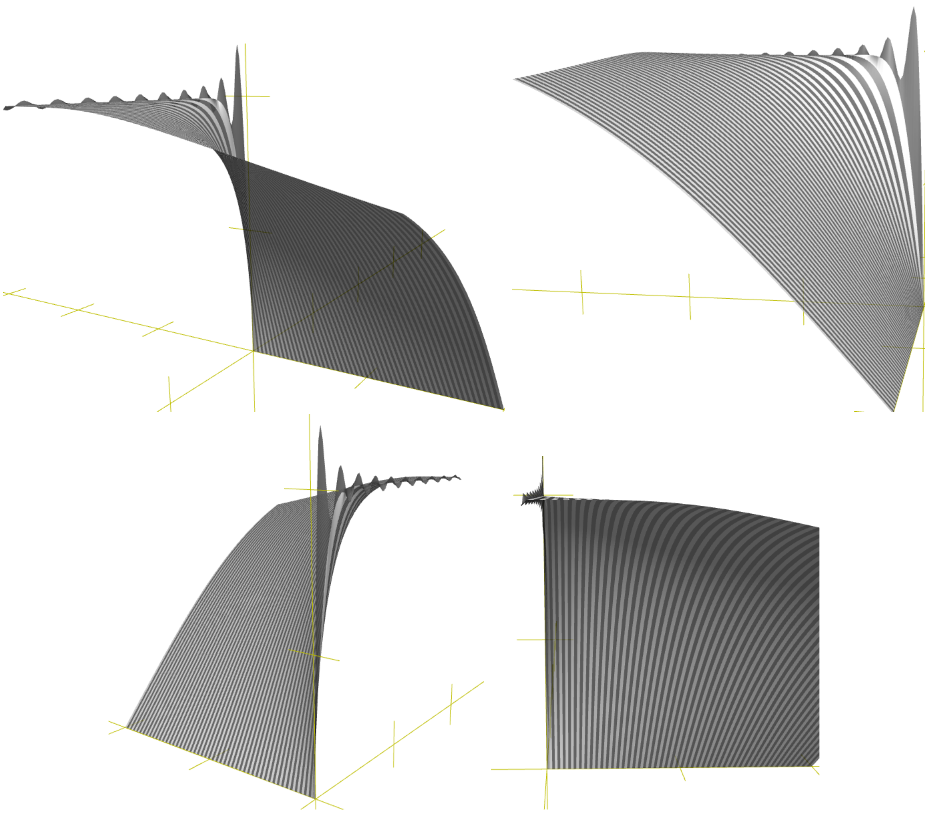


Figure 6. Gibbs phenomenon and bad approximation of discontinuity, up close.

6.3 Spaces of Hilbert space valued functions

In order to work with more general solutions and less smooth data, we need to introduce a few new function spaces. Let V denote a separable Hilbert space. Let $T > 0$ be given. The space $C^0([0,T];V)$ of continuous functions from $[0,T]$ with values in V is a Banach space for its natural norm

$$\|f\|_{C^0([0,T];V)} = \max_{t\in[0,T]} \|f(t)\|_V.$$

A V-valued function on $]0,T[$ is differentiable at point $t \in]0,T[$ if there exists a vector $f'(t) \in V$ such that

$$\left\| \frac{f(t+h)-f(t)}{h} - f'(t) \right\|_V \to 0 \text{ when } h \to 0.$$

Of course, $f'(t)$ is called the derivative of f at point t. A function is clearly continuous at all its points of differentiability. If f is differentiable at all points t, then its derivative becomes a V-valued function. We can define

$$C^1([0,T];V) = \{f \in C^0([0,T];V); f' \in C^0([0,T];V)\}$$

in the sense that f' has a continuous extension at 0 and T. When equipped with its natural norm

$$\|f\|_{C^1([0,T];V)} = \max\big(\|f\|_{C^0([0,T];V)}, \|f'\|_{C^0([0,T];V)}\big),$$

it is a Banach space. More generally, we can define $C^k([0,T];V)$ for all positive integers k. All of these notions are perfectly classical and work the same as in the real-valued case.

Measurability (and integrability) issues are a little trickier in the infinite dimensional valued case than in the finite dimensional valued case. There are different types of measurability and integrals when V is a Banach space or a more general topological vector space. We stick to the simplest notions. Besides we will not use V-valued integrals here. We equip $[0,T]$ with the Lebesgue σ-algebra.

Definition 6.3.1 *A function $f\colon [0,T] \to V$ is called a* simple function *if there exists a finite measurable partition of $[0,T]$, $(E_i)_{i=1,\dots,k}$, and a finite set of vectors $v_i \in V$ such that*

$$f(t) = \sum_{i=1}^{k} \mathbf{1}_{E_i}(t) v_i,$$

for all $t \in [0,T]$.

In other words, f only takes a finite number of values in V and is equal to v_i exactly on the Lebesgue measurable set E_i. It should be noted that for each t, there is one and only one nonzero term $\mathbf{1}_{E_i}(t)$ in the sum, due to the fact that the sets E_i form a partition of $[0,T]$.

Definition 6.3.2 *A function $f\colon [0,T] \to V$ is said to be* measurable *if there exists a negligible set $N \subset [0,T]$ and a sequence of simple functions f_n such that $\|f_n(t) - f(t)\|_V \to 0$ when $n \to +\infty$ for all $t \notin N$.*

We also say that f is an almost everywhere limit of simple functions. When $V = \mathbb{R}$, this notion of measurability coincides with the usual one. It is easy to see that a continuous function is measurable.

Proposition 6.3.1 *Let $f\colon [0,T] \to V$ be a measurable function. Then the function $N_V f\colon [0,T] \to \mathbb{R}_+$, $N_V f(t) = \|f(t)\|_V$, is a measurable function in the usual sense.*

Proof. Let $f_n(t) = \sum_{i=1}^{k_n} \mathbf{1}_{E_{n,i}}(t) v_{n,i}$ be a sequence of simple functions that converges a.e. to f. Since the norm of V is continuous from V into $\mathbb{R}$, we have $N_V f_n \to N_V f$ a.e. Now due to the fact that for all t, there is at most one nonzero term in the sum, we also have $\|f_n(t)\|_V = \sum_{i=1}^{k_n} \mathbf{1}_{E_{n,i}}(t) \|v_{n,i}\|_V$, hence $N_V f_n$ is a real-valued simple function. Therefore $N_V f$ is measurable. $\square$

Definition 6.3.3 *We say that two measurable functions $f_1, f_2\colon [0,T] \to V$ are* equal almost everywhere *if there exists a negligible set $N \subset [0,T]$ such that $f_1(t) = f_2(t)$ for all $t \notin N$.*

Almost everywhere equality is an equivalence relation and from now on, we will not distinguish between a function and its equivalence class. The V-valued Lebesgue spaces are defined as would be expected

$$L^p(0,T;V) = \big\{ f\colon [0,T] \to V, \text{ measurable and such that } N_V f \in L^p(0,T) \big\}$$

equipped with the norm

$$\|f\|_{L^p(0,T;V)} = \|N_V f\|_{L^p(0,T)}$$

are Banach spaces. For $p = 2$, $L^2(0,T;V)$ is a Hilbert space for the scalar product

$$(f|g)_{L^2(0,T;V)} = \int_0^T \big(f(t)|g(t)\big)_V \, dt.$$

The Hilbert norm reads explicitly

$$\|f\|_{L^2(0,T;V)} = \Big(\int_0^T \|f(t)\|_V^2 \, dt \Big)^{1/2}.$$

Obviously,
$$C^0([0,T];V) \hookrightarrow L^p(0,T;V),$$
for all $p \in [1,+\infty]$.

It is also possible to define V-valued Sobolev spaces and V-valued distributions, but we will not use these spaces here.

In the case when V is itself a function space on an open set Ω of $\mathbb{R}^d$, there is a natural connection between V-valued functions on $[0,T]$ and real valued functions on $\bar{Q} = \bar{\Omega} \times [0,T]$ in $d+1$ variables. Let us give an example of this.

Proposition 6.3.2 *The spaces $L^2(0,T;L^2(\Omega))$ and $L^2(Q)$ are canonically isometric.*

Proof. We leave aside the measurability questions, which are delicate. First of all, let us take $f \in L^2(Q)$. We thus have $\int_Q f(x,t)^2\,dxdt < +\infty$. By Fubini's theorem applied to f^2, we thus have that
$$\int_\Omega f(x,t)^2\,dx < +\infty \text{ for almost all } t \in [0,T]$$
and
$$\int_Q f(x,t)^2\,dxdt = \int_0^T \Big(\int_\Omega f(x,t)^2\,dx\Big)\,dt.$$
Therefore, if we set $\widetilde{f}(t) = f(\cdot,t)$, then we see that $\widetilde{f}(t) \in L^2(\Omega)$ for almost all t. Thus we can let $\widetilde{f}(t) = 0$ for those t for which the initial $\widetilde{f}$ is not in $L^2(\Omega)$ and $\widetilde{f}$ is then an $L^2(\Omega)$-valued function. Moreover, the second relation then reads
$$\|f\|^2_{L^2(\Omega)} = \|\widetilde{f}\|^2_{L^2(0,T;L^2(\Omega))},$$
hence the isometry.

Conversely, taking $\widetilde{f} \in L^2(0,T;L^2(\Omega))$, then for all t, $\widetilde{f}$ is a function in the variable $x \in \Omega$ that belongs to $L^2(\Omega)$. If we thus set $f(x,t) = \widetilde{f}(t)(x)$, we define a function on Q which is such that $\int_0^T \big(\int_\Omega f(x,t)^2\,dx\big)\,dt < +\infty$. By Fubini's theorem again, it follows that $f \in L^2(Q)$ and we have the isometry. □

It is thus possible to switch between the two points of view: function in one variable with values in a function space on a d dimensional domain and function in $d+1$ variables. If $\widetilde{f} \in C^1([0,T];L^2(\Omega))$, then the associated f is in $L^2(Q)$ and it can be shown that its distributional derivative $\frac{\partial f}{\partial t}$ is in $C^0([0,T];L^2(\Omega))$ and $\widetilde{\frac{\partial f}{\partial t}} = (\widetilde{f})'$.

We will also encounter such situations as $f \in C^0([0,T];H) \cap L^2(0,T;V)$ with two (or more) different spaces $V \subset H$, meaning that $f(t)$ is unambiguously defined as an element of H for all t, and continuous with values in H, and the same $f(t)$ is in V for almost all t and square integrable with values in V. It is allowed to exit V on a negligible subset of $[0,T]$. Moreover, if T is a continuous linear operator from V to H, then we can define Tf by $(Tf)(t) = T(f(t))$. This definition commutes with all previous notions.

Proposition 6.3.3 *If $f \in C^k([0,T];V)$ then $Tf \in C^k([0,T];H)$ with $(Tf)^{(j)} = T(f^{(j)})$ for $j \le k$, and if $f \in L^2(0,T;V)$ then $Tf \in L^2(0,T;H)$.*

Proof. We start with the continuity. We have

$$\|Tf(t+h) - Tf(t)\|_H = \|T(f(t+h) - f(t))\|_H \le \|T\| \|f(t+h) - f(t)\|_V \underset{h\to 0}{\longrightarrow} 0.$$

Therefore $Tf \in C^0([0,T];H)$. Moreover, we have $\|Tf(t)\|_H \le \|T\| \|f(t)\|_V$ so that taking the max for $t \in [0,T]$ on both sides

$$\|Tf\|_{C^0([0,T];H)} \le \|T\| \|f\|_{C^0([0,T];V)}.$$

Similarly

$$\left\| \frac{Tf(t+h) - Tf(t)}{h} - Tf'(t) \right\|_H \le \|T\| \left\| \frac{f(t+h) - f(t)}{h} - f'(t) \right\|_V \underset{h\to 0}{\longrightarrow} 0$$

and so on for the successive derivatives and their norms. Finally,

$$\int_0^T \|Tf(t)\|_H^2 \, dt \le \|T\|^2 \int_0^T \|f(t)\|_V^2 \, dt < +\infty,$$

leaving aside measurability issues, which are not difficult here. Of course, the above inequality is nothing but

$$\|Tf\|_{L^2(0,T;H)} \le \|T\| \|f\|_{L^2(0,T;V)},$$

as with the C^k spaces. □

To get an idea of how this can be used, just take $V = H^2(\Omega)$, $H = L^2(\Omega)$ and $T = -\Delta$.

6.4 Energy estimates, stability, uniqueness

In this section, we consider solutions of problem (6.1) with data that is considerably less smooth than in the previous section. We assume that the solutions considered are regular enough so that all computations are justified. As the proof in arbitrary dimension of space d work the same as in one dimension, we will let $\Omega \in \mathbb{R}^d$ bounded and $Q = \Omega \times]0,T[$.

We start with a lemma.

Lemma 6.4.1 *Let $u \in C^1([0,T];L^2(\Omega))$. Then the function $t \mapsto \frac{1}{2}\int_\Omega u(x,t)^2\,dx$ is of class $C^1([0,T])$ and its derivative is given by $t \mapsto \int_\Omega u(x,t)u'(x,t)\,dx$.*

Proof. Let $E(t) = \frac{1}{2}\int_\Omega u(x,t)^2\,dx$. We write

$$\frac{E(t+h)-E(t)}{h} = \frac{1}{2}\int_\Omega \big(u(x,t+h)+u(x,t)\big)\Big(\frac{u(x,t+h)-u(x,t)}{h}\Big)\,dx.$$

Now, by L^2-valued continuity, $u(\cdot,t+h) \to u(\cdot,t)$ in $L^2(\Omega)$ when $h \to 0$. By L^2-valued differentiability, $\frac{u(\cdot,t+h)-u(\cdot,t)}{h} \to u'(\cdot,t)$ in $L^2(\Omega)$ when $h \to 0$. Therefore,

$$\frac{E(t+h)-E(t)}{h} \to \int_\Omega u(x,t)u'(x,t)\,dx$$

when $h \to 0$ by the Cauchy-Schwarz inequality. By the same inequality, the right-hand side is a continuous function of t. □

Remark 6.4.1 This result can be construed as a kind of differentiation under the integral sign, since

$$\frac{d}{dt}\Big(\int_\Omega u(x,t)^2\,dx\Big) = 2\int_\Omega \frac{\partial u}{\partial t}(x,t)u(x,t)\,dx = \int_\Omega \frac{\partial(u^2)}{\partial t}(x,t)\,dx,$$

with the identification $\frac{\partial u}{\partial t} = u'$. □

Proposition 6.4.1 *Assume that $g = 0$ (homogeneous Dirichlet conditions), $u_0 \in L^2(\Omega)$ and $f \in L^2(Q)$. Then, if $u \in C^1([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$, then*

$$\|u\|_{C^0([0,T];L^2(\Omega))} \le \|u_0\|_{L^2(\Omega)} + C\|f\|_{L^2(Q)}, \tag{6.2}$$

where C is the Poincaré inequality constant.

Proof. Since $u \in C^1([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$, we have that $u(\cdot,t)$ and $\frac{\partial u}{\partial t}(\cdot,t)$ belong to $L^2(\Omega)$ for all t and that $u(\cdot,t)$ belongs to $H_0^1(\Omega)$ for almost all t. The meaning of the partial differential equation in this context is thus that $u' - \Delta u = f$ where $u' \in C^0([0,T];L^2(\Omega))$, $f \in L^2(0,T;L^2(\Omega))$ and $\Delta u \in L^2(0,T;H^{-1}(\Omega))$, see Corollary 2.6.4, so that the equation is well defined in this sense. Of course, it also coincides with the distributional equation on Q.

For almost all $s \in [0,T]$, both sides of the equation are in $H^{-1}(\Omega)$. We thus take the duality bracket by u and obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{ds}\Big(\int_\Omega u(x,s)^2\,dx\Big) + \int_\Omega \|\nabla u(x,s)\|^2\,dx &= \int_\Omega f(x,s)u(x,s)\,dx \\ &\le \Big(\int_\Omega f(x,s)^2\,dx\Big)^{\frac{1}{2}}\Big(\int_\Omega u(x,s)^2\,dx\Big)^{\frac{1}{2}}\end{aligned}$$

by Lemma 6.4.1 and Cauchy-Schwarz. Because of the homogeneous Dirichlet condition, we have Poincaré's inequality

$$\Big(\int_\Omega u(x,s)^2\,dx\Big)^{\frac{1}{2}} \le C\Big(\int_\Omega \|\nabla u(x,s)\|^2\,dx\Big)^{\frac{1}{2}},$$

and using Young's inequality $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$, we obtain

$$\frac{1}{2}\frac{d}{ds}\Big(\int_\Omega u(x,s)^2\,dx\Big) + \int_\Omega \|\nabla u(x,s)\|^2\,dx \\ \le \frac{C^2}{2}\int_\Omega f(x,s)^2\,dx + \frac{1}{2}\int_\Omega \|\nabla u(x,s)\|^2\,dx,$$

so that

$$\frac{1}{2}\frac{d}{ds}\Big(\int_\Omega u(x,s)^2\,dx\Big) \le \frac{1}{2}\frac{d}{ds}\Big(\int_\Omega u(x,s)^2\,dx\Big) + \frac{1}{2}\int_\Omega \|\nabla u(x,s)\|^2\,dx \\ \le \frac{C^2}{2}\int_\Omega f(x,s)^2\,dx.$$

We integrate the above inequality between 0 and t with respect to s and obtain

$$\int_\Omega u(x,t)^2\,dx - \int_\Omega u(x,0)^2\,dx \le C^2\int_0^t\int_\Omega f(x,s)^2\,dx\,ds \le C^2\int_0^T\int_\Omega f(x,s)^2\,dx\,ds,$$

for all $t \in [0,T]$ due to Lemma 6.4.1, and since $u(x,0) = u_0(x)$, it follows that

$$\|u(\cdot,t)\|_{L^2(\Omega)} \le \Big(\|u_0\|^2_{L^2(\Omega)} + C^2\|f\|^2_{L^2(Q)}\Big)^{1/2} \tag{6.3}$$

hence the result, since $\sqrt{a^2+b^2} \le a+b$ for all a,b positive. □

Remark 6.4.2 The quantity $E(t) = \frac{1}{2}\int_\Omega u(x,t)^2\,dx$ is the *energy* (up to physical constants), hence the term "energy estimate". It follows from the proof that the energy is decreasing when $f = 0$. In addition, it is quite clear also from the proof that if $f \in L^2(\Omega \times \mathbb{R}_+)$, then the energy estimates remains valid for all times, *i.e.*,

$$\sup_{t\in\mathbb{R}_+} \|u(\cdot,t)\|_{L^2(\Omega)} \le \|u_0\|_{L^2(\Omega)} + C\|f\|_{L^2(\Omega\times\mathbb{R}_+)},$$

provided such a solution exists.

Let us note that the energy estimate can be proved under lower regularity hypotheses, namely that $u \in C^0([0,T];L^2(\Omega) \cap L^2(0,T;H^1_0(\Omega))$. The first space in the intersection gives a precise meaning to the initial condition in L^2. □

As in the case of the maximum principle, the energy estimate has consequences in terms of uniqueness and stability.

Corollary 6.4.1 *There is at most one solution u belonging to $C^1([0,T];L^2(\Omega))\cap L^2(0,T;H^1(\Omega))$ to the heat equation with initial data $u_0\in L^2(\Omega)$, right-hand side $f\in L^2(Q)$ and Dirichlet boundary condition $g\in L^2(0,T;H^{1/2}(\partial\Omega))$.*

Proof. Let u_1 and u_2 be two such solutions. Then, their difference u_1-u_2 belongs to $C^1([0,T];L^2(\Omega))\cap L^2(0,T;H^1_0(\Omega))$ and is a solution of the heat equation with zero right-hand side and initial condition. By estimate (6.2), it follows that we have $u_1-u_2=0$. □

Again this also holds in $C^0([0,T];L^2(\Omega)\cap L^2(0,T;H^1_0(\Omega))$. Stability or continuous dependence on the data is straightforward.

Corollary 6.4.2 *Let $u_i\in C^1([0,T];L^2(\Omega))\cap L^2(0,T;H^1_0(\Omega))$ be solutions corresponding to initial conditions $u_{0,i}\in L^2(\Omega)$ and right-hand sides $f_i\in L^2(Q)$. Then*

$$\|u_1-u_2\|_{C^0([0,T];L^2(\Omega))}\leq\|u_{0,1}-u_{0,2}\|_{L^2(\Omega)}+C\|f_1-f_2\|_{L^2(Q)}. \tag{6.4}$$

Proof. Clear. □

When there is no heat source, $f=0$, we can expect some kind of exponential decay as in the regular case. Here, the energy is the relevant quantity.

Proposition 6.4.2 *If $f=0$, then we have*

$$E(t)\leq e^{-\frac{2t}{C^2}}E(0)=\frac{e^{-\frac{2t}{C^2}}}{2}\|u_0\|^2_{L^2(\Omega)}, \tag{6.5}$$

where C is the Poincaré inequality constant.

Proof. As before, we have

$$\frac{d}{dt}\Big(\frac{1}{2}\int_\Omega u(x,t)^2\,dx\Big)+\int_\Omega\|\nabla u(x,t)\|^2\,dx=0.$$

Thus

$$\frac{dE}{dt}(t)=-\int_\Omega\|\nabla u(x,t)\|^2\,dx\leq-\frac{1}{C^2}\int_\Omega u(x,t)^2\,dx=-\frac{2}{C^2}E(t),$$

by Poincaré's inequality. Solving this differential inequality, we obtain the result. □

Remark 6.4.3 A function in L^2 is not bounded in general, thus we cannot expect uniform decay of the temperature as in the regular case. However, it can be shown that u is of class C^∞ as soon as $t>0$, which is the same smoothing effect as before. Thus, there is also a uniform exponential decay but starting away from $t=0$. In fact, it can be shown that u is C^∞ on any open subset where f is C^∞, in particular when it is equal to 0. This property of the heat operator is called *hypoellilticity*. □

6.5 Variational formulation and existence of weak solutions

So far, we still have no existence result for $f \neq 0$ or for $f = 0$ and $u_0 \in L^2(\Omega)$. For this, we need to recast the problem in variational form. We only consider the homogeneous Dirichlet boundary condition, since a non homogeneous Dirichlet condition can be transformed into a homogeneous one via an appropriate lift of the boundary data. We start with regularity hypotheses that are a little to strong, but not by much.

Proposition 6.5.1 *Let $u_0 \in L^2(\Omega)$, $f \in L^2(Q)$. Consider u in $C^1([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ such that $u' - \Delta u = f$ for almost all t and $u(0) = u_0$. Then we have, for all $v \in H_0^1(\Omega)$,*

$$\frac{d}{dt}\big((u(t)|v)_{L^2(\Omega)}\big) + a(u(t),v) = (f(t)|v)_{L^2(\Omega)}$$

almost everywhere in $[0,T]$, where $a(u,v) = \int_\Omega \nabla u \cdot \nabla v\,dx$, and

$$(u(0)|v)_{L^2(\Omega)} = (u_0|v)_{L^2(\Omega)}.$$

Conversely, a solution in $C^1([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ of the above two variational equations is a solution of the initial-boundary value problem for the heat equation with initial data u_0 and right-hand side f.

Proof. We have already seen that each term in the equation $u' - \Delta u = f$ is at worst in $L^2(0,T;H^{-1}(\Omega))$. It is therefore meaningful to take the duality bracket of each one of them with $v \in H_0^1(\Omega)$, so that we have

$$\langle u'(t), v\rangle - \langle \Delta u(t), v\rangle = \langle f(t), v\rangle,$$

for almost all t.

Arguing as in the proof of Lemma 6.4.1, we see that the real-valued function $t \mapsto (u(t)|v)_{L^2(\Omega)}$ is of class C^1 and that $\frac{d}{dt}\big((u(t)|v)_{L^2(\Omega)}\big) = \int_\Omega u'(x,t)v(x)\,dx = \langle u'(t), v\rangle$ since $u'(t) \in L^2(\Omega)$. Similarly, $\langle f(t), v\rangle = \int_\Omega f(x,t)v(x)\,dx$. Finally, by Corollary 2.6.4, we have $-\langle \Delta u(t), v\rangle = a(u(t),v)$ so that the first equation is established. The second equation is trivial.

Conversely, let us be given a solution u of the variational problem. Since $H_0^1(\Omega)$ is dense in $L^2(\Omega)$, the second equation implies that $u(0) = u_0$. Moreover, the above calculations can be carried out backwards, so that

$$\langle u'(t) - \Delta u(t) - f(t), v\rangle = 0,$$

for almost all t and all $v \in H_0^1(\Omega)$. Consequently, for almost all t, $u'(t) - \Delta u(t) - f(t) = 0$ as an element of $H^{-1}(\Omega)$, hence the heat equation with right-hand side f is satisfied in this sense. $\square$

As we said above, the regularity in time assumed above is a bit too high. Indeed, the variational formulation makes sense in a slightly less regular context. This leads to the following definition.

Definition 6.5.1 *The variational formulation of the heat equation with homogeneous Dirichlet boundary condition, initial data $u_0 \in L^2(\Omega)$ and right-hand side $f \in L^2(Q)$ consists in looking for $u \in C^0([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ such that, for all $v \in H_0^1(\Omega)$,*

$$\begin{cases} \left((u|v)_{L^2(\Omega)}\right)' + a(u,v) = (f|v)_{L^2(\Omega)} \text{ in the sense of } \mathscr{D}'(]0,T[), \\ \qquad (u(0)|v)_{L^2(\Omega)} = (u_0|v)_{L^2(\Omega)}. \end{cases} \tag{6.6}$$

Remark 6.5.1 Let us check that this definition makes sense. First of all, since $u \in C^0([0,T];L^2(\Omega))$ and v does not depend on t, we see that the function $t \mapsto (u|v)_{L^2(\Omega)}$ is continuous on $[0,T]$, hence its derivative is a distribution on $]0,T[$. Likewise, since $u \in L^2(0,T;H_0^1(\Omega))$, the function $t \mapsto a(u,v)$ is in $L^1(0,T)$ by Cauchy-Schwarz, hence a distribution on $]0,T[$ and the same holds for $t \mapsto (f|v)_{L^2(\Omega)}$. Therefore, the first equation in (6.6) is well defined in the distributional sense.

We have already seen that the second equation is equivalent to $u(0) = u_0$, and the continuity of u with respect to t with values in $L^2(\Omega)$ makes this initial condition relevant. □

We use the variational formulation to prove existence and uniqueness of solutions. We will write the proof in the 1d case, $\Omega =]0,1[$. The general case is entirely similar. For all $k \in \mathbb{N}^*$, we let $\phi_k(x) = \sqrt{2}\sin(k\pi x)$ and $\lambda_k = k^2\pi^2$. It is well-known that the family $(\phi_k)_{k\in\mathbb{N}^*}$ is a Hilbert base of $L^2(0,1)$ and a total orthogonal family in $H_0^1(]0,1[)$ (recall that a total family is a family that spans a dense vector space). Moreover, for all $w \in H_0^1(\Omega)$, we have

$$a(w,\phi_k) = \int_0^1 w'\phi_k'\,dx = -\int_0^1 w\phi_k''\,dx = \lambda_k (w|\phi_k)_{L^2(\Omega)}.$$

Theorem 6.5.1 *Let $u_0 \in L^2(\Omega)$, $f \in L^2(Q)$. There exists a unique solution $u \in C^0([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ of problem* (6.6)*, which is given by*

$$u(t) = \sum_{k=1}^{+\infty} u_k(t)\phi_k, \tag{6.7}$$

where

$$u_k(t) = (u_0|\phi_k)_{L^2(\Omega)} e^{-\lambda_k t} + \int_0^t (f(s)|\phi_k)_{L^2(\Omega)} e^{-\lambda_k(t-s)}\,ds \tag{6.8}$$

and the series converges in $C^0([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$.

Proof. We start with the uniqueness. Let $u \in C^0([0,T];L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ be a solution of (6.6). For all $t \in [0,T]$, $u(t)$ is thus an element of $L^2(\Omega)$ and can therefore be expanded on the Hilbert basis $(\phi_k)_{k\in\mathbb{N}^*}$. Consequently, we have for all t

$$u(t) = \sum_{k=1}^{+\infty} u_k(t)\phi_k \tag{6.9}$$

with

$$u_k(t) = (u(t)|\phi_k)_{L^2(\Omega)}$$

for all $k \in \mathbb{N}^*$ and the series converges in $L^2(\Omega)$. Now $\phi_k \in H_0^1(\Omega)$ is a legitimate test-function in problem (6.6). In particular, since $u(t) \in H_0^1(\Omega)$ almost everywhere, we have

$$a(u(t),\phi_k) = \lambda_k u_k(t)$$

almost everywhere, hence everywhere since the right-hand side is continuous (the left-hand side is L^1). We thus have

$$\begin{cases} u_k'(t) + \lambda_k u_k(t) = (f(t)|\phi_k)_{L^2(\Omega)} \text{ in the sense of } \mathscr{D}'(]0,T[), \\ \qquad\qquad u_k(0) = (u_0|\phi_k)_{L^2(\Omega)}, \end{cases}$$

for each $k \in \mathbb{N}^*$. Now this is a Cauchy problem for a linear ordinary differential equation, and there are no other distributional solutions than the classical solution obtained by variation of the constant, or Duhamel's formula:

$$u_k(t) = (u_0|\phi_k)_{L^2(\Omega)} e^{-\lambda_k t} + \int_0^t (f(s)|\phi_k)_{L^2(\Omega)} e^{-\lambda_k (t-s)}\,ds,$$

which is exactly formula (6.8).[3] Hence the uniqueness.

We now use the above series to prove existence. First recall that $u_0 \in L^2(\Omega)$, therefore

$$\|u_0\|^2_{L^2(\Omega)} = \sum_{k=1}^{+\infty} (u_0|\phi_k)^2_{L^2(\Omega)}$$

by Plancherel's formula. Similarly, $f \in L^2(Q)$ and

$$\|f\|^2_{L^2(Q)} = \int_0^T \sum_{k=1}^{+\infty} (f(t)|\phi_k)^2_{L^2(\Omega)}\,dt.$$

Let us set $u_{0,k} = (u_0|\phi_k)_{L^2(\Omega)}$ and $f_k(t) = (f(t)|\phi_k)_{L^2(\Omega)}$. We are going to show that the series in formula (6.9) converges in both spaces $C^0(0,T;L^2(\Omega))$ and $L^2(0,T;H_0^1(\Omega))$ and that its sum u is a solution of the variational problem. To do

[3] Observe that the function u_k is continuous in t.

this, we will show that the partial sums $U_n(t) = \sum_{k=1}^n u_k(t)\phi_k$ are Cauchy sequences for both norms. Let $p < q$ be to given integers and let us estimate $U_p - U_q$.

First of all, due to the continuity of $t \mapsto u_k(t)$, the partial sums U_n are continuous with values in $L^2(\Omega)$. Moreover, for all $t \in [0,T]$, we have

$$\begin{aligned}
\|U_p(t) - U_q(t)\|^2_{L^2(\Omega)} &= \Big\| \sum_{k=p+1}^{q} u_k(t)\phi_k \Big\|^2_{L^2(\Omega)} \\
&= \sum_{k=p+1}^{q} u_k(t)^2 \\
&\le 2 \sum_{k=p+1}^{q} \Big[u_{0,k}^2 + \Big(\int_0^t |f_k(s)|\,ds \Big)^2 \Big] \\
&\le 2 \sum_{k=p+1}^{q} \Big[u_{0,k}^2 + t \int_0^t f_k(s)^2\,ds \Big] \\
&\le 2 \sum_{k=p+1}^{q} u_{0,k}^2 + 2T \sum_{k=p+1}^{q} \int_0^T f_k(s)^2\,ds
\end{aligned}$$

since all the exponential terms are less than 1 and by Cauchy-Schwarz. Therefore

$$\begin{aligned}
\|U_p - U_q\|^2_{C^0([0,T];L^2(\Omega))} &= \max_{t\in[0,T]} \|U_p(t) - U_q(t)\|^2_{L^2(\Omega)} \\
&\le 2 \sum_{k=p+1}^{q} u_{0,k}^2 + 2T \sum_{k=p+1}^{q} \int_0^T f_k(s)^2\,ds
\end{aligned}$$

can be made as small as we wish by taking p large enough, due to the hypotheses on u_0 and f, and the sequence is consequently Cauchy in $C^0(0,T;L^2(\Omega))$.

Similarly, the partial sums are obviously in $L^2(0,T;H_0^1(\Omega))$, in fact they even are continuous with values in $H_0^1(\Omega)$, although this continuity will not persist in the limit. We use the H_0^1 seminorm, so that $|v|^2_{H_0^1(\Omega)} = a(v,v)$ (for a more general parabolic equation, H_0^1-ellipticity of the bilinear form would here come into play).

We recall that the family $(\phi_k)_{k\in\mathbb{N}^*}$ is also orthogonal in $H_0^1(\Omega)$, thus

$$\begin{aligned}
|U_p(t) - U_q(t)|^2_{H_0^1(\Omega)} &= \sum_{k=p+1}^{q} u_k(t)^2 a(\phi_k,\phi_k) \\
&= \sum_{k=p+1}^{q} \lambda_k u_k(t)^2,
\end{aligned}$$

so that integrating between 0 and T, we obtain

$$\|U_p - U_q\|^2_{L^2(0,T;H_0^1(\Omega))} = \sum_{k=p+1}^{q} \int_0^T \lambda_k u_k(t)^2\,dt.$$

Let us estimate each term in the sum on the right. We have

$$\lambda_k u_k(t)^2 \leq 2\lambda_k \Big(u_{0,k}^2 e^{-2\lambda_k t} + T \int_0^t f_k(s)^2 e^{-2\lambda_k(t-s)}\, ds \Big),$$

so that

$$\begin{aligned}
\int_0^T \lambda_k u_k(t)^2\, dt &\leq 2\lambda_k \Big(u_{0,k}^2 \int_0^T e^{-2\lambda_k t}\, dt + T \int_0^T \int_0^t f_k(s)^2 e^{-2\lambda_k(t-s)}\, ds dt \Big) \\
&= (1 - e^{-2\lambda_k T}) u_{0,k}^2 + 2\lambda_k T \int_0^T f_k(s)^2 \int_s^T e^{-2\lambda_k(t-s)}\, dt ds \\
&= (1 - e^{-2\lambda_k T}) u_{0,k}^2 + T \int_0^T (1 - e^{-2\lambda_k(T-s)}) f_k(s)^2\, ds \\
&\leq u_{0,k}^2 + T \int_0^T f_k(s)^2\, ds.
\end{aligned}$$

Therefore

$$\|U_p - U_q\|^2_{L^2(0,T;H_0^1(\Omega))} \leq \sum_{k=p+1}^{q} u_{0,k}^2 + T \sum_{k=p+1}^{q} \int_0^T f_k(s)^2\, ds,$$

which can again be made as small as we wish by taking p large enough, and the sequence is Cauchy in $L^2(0,T;H_0^1(\Omega))$.

Finally, it remains to be seen that the function u defined by the series and which belongs to $C^0(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$ is a solution of the variational problem (6.6). The initial condition is obvious even in non variational form since

$$u(0) = \sum_{k=1}^{+\infty} \Big(u_{0,k} e^0 + \int_0^0 f_k(s) e^{\lambda_k s}\, ds \Big) \phi_k = \sum_{k=1}^{+\infty} u_{0,k} \phi_k = u_0.$$

Regarding the evolution equation, we obtain from the ordinary differential equations for u_k that for all $v \in \operatorname{span}((\phi_k)_{k \in \mathbb{N}^*})$

$$\big((u|v)_{L^2(\Omega)} \big)' + a(u,v) = (f(t)|v)_{L^2(\Omega)} \text{ in the sense of } \mathscr{D}'(]0,T[)$$

since any such v is a linear combination of the ϕ_k.

Let now $v \in H_0^1(\Omega)$ be arbitrary and $v_n \in \operatorname{span}((\phi_k)_{k \in \mathbb{N}^*})$ be such that $v_n \to v$ in $H_0^1(\Omega)$. For any $\varphi \in \mathscr{D}(]0,T[)$, we thus have

$$-\int_0^T (u(t)|v_n)_{L^2(\Omega)} \varphi'(t)\, dt + \int_0^T a(u(t),v_n) \varphi(t)\, dt = \int_0^T (f(t)|v_n)_{L^2(\Omega)} \varphi(t)\, dt.$$

It is then fairly obvious that each term in the above relation passes to the limit as $n \to +\infty$, thus establishing the evolution equation. $\square$

Remark 6.5.2 Formula (6.7)–(6.8) clearly generalizes the expansion obtained in Theorem 6.2.1. □

Remark 6.5.3 The recovery of a bona-fide solution of the heat equation from the above variational solution would require the use of Hilbert space valued distributions and integrals. Let us just say that it can be done. There are other approaches to the heat equation, for instance using semigroups. □

Remark 6.5.4 The d-dimensional heat equation can be solved along the exact same lines, replacing the functions ϕ_k and scalars λ_k by the eigenfunctions and eigenvalues of the minus Laplacian in $H_0^1(\Omega)$, *i.e.*, the solutions of $-\Delta\phi_k = \lambda_k\phi_k$, $\phi_k \in H_0^1(\Omega)$, $\phi_k \neq 0$. This eigenvalue problem for Ω bounded only has solutions for λ_k in a sequence $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots$ such that $\lambda_k \to +\infty$ when $k \to +\infty$. Of course, the eigenvalues and eigenfunctions depend on the shape of Ω, see Chapter 1, Section 1.6. □

We also have an energy decay and stability estimate in the present context.

Proposition 6.5.2 *The solution u of problem* (6.6) *satisfies*

$$\|u(t)\|_{L^2(\Omega)} \leq \|u_0\|_{L^2(\Omega)} e^{-\lambda_1 t} + \int_0^t \|f(s)\|_{L^2(\Omega)} e^{-\lambda_1(t-s)}\,ds, \tag{6.10}$$

for all $t \in [0,T]$.

Proof. This is a consequence of the series expansion. We first observe the following fact. Let g be a L^1-function from $[0,T]$ to a Euclidean space E (*i.e.*, a finite dimensional Hilbert space). Then the integral $\int_0^t g(s)\,ds$ is well defined as a vector of E by choosing a basis of E and integrating g componentwise. Moreover, since E is Euclidean, there exists a unit vector e such that

$$\Big\|\int_0^t g(s)\,ds\Big\|_E = \Big(\int_0^t g(s)\,ds\Big)\cdot e = \int_0^t g(s)\cdot e\,ds \leq \int_0^t \|g(s)\|_E\,ds,$$

by the Cauchy-Schwarz inequality.

We now turn to estimate (6.10). We have

$$u(t) = \sum_{k=1}^{+\infty}\Big(u_{0,k}e^{-\lambda_k t} + \int_0^t f_k(s)e^{-\lambda_k(t-s)}\,ds\Big)\phi_k$$

so that by the triangle inequality

$$\|u(t)\|_{L^2(\Omega)} \leq \Big\|\sum_{k=1}^{+\infty} u_{0,k}e^{-\lambda_k t}\phi_k\Big\|_{L^2(\Omega)} + \Big\|\sum_{k=1}^{+\infty}\int_0^t f_k(s)e^{-\lambda_k(t-s)}\,ds\phi_k\Big\|_{L^2(\Omega)}.$$

For the first term, we note that

$$\Big\|\sum_{k=1}^{+\infty} u_{0,k}e^{-\lambda_k t}\phi_k\Big\|_{L^2(\Omega)} = \Big(\sum_{k=1}^{+\infty} u_{0,k}^2 e^{-2\lambda_k t}\Big)^{\frac{1}{2}} \le \Big(\sum_{k=1}^{+\infty} u_{0,k}^2 e^{-2\lambda_1 t}\Big)^{\frac{1}{2}} = \|u_0\|_{L^2(\Omega)}e^{-\lambda_1 t}$$

since the sequence of eigenvalues λ_k is increasing. For the second term, we resort to the observation above with $E = \mathrm{span}(\phi_1, \dots, \phi_n)$ equipped with the L^2-norm, and deduce that

$$\begin{aligned}\Big\|\sum_{k=1}^{n}\int_0^t f_k(s)e^{-\lambda_k(t-s)}\,ds\phi_k\Big\|_{L^2(\Omega)} &= \Big\|\int_0^t \sum_{k=1}^{n} f_k(s)e^{-\lambda_k(t-s)}\phi_k\,ds\Big\|_{L^2(\Omega)}\\ &\le \int_0^t \Big\|\sum_{k=1}^{n} f_k(s)e^{-\lambda_k(t-s)}\phi_k\Big\|_{L^2(\Omega)}\,ds\\ &= \int_0^t \Big(\sum_{k=1}^{n} f_k(s)^2 e^{-2\lambda_k(t-s)}\Big)^{\frac{1}{2}}\,ds\\ &\le \int_0^t \Big(\sum_{k=1}^{n} f_k(s)^2\Big)^{\frac{1}{2}} e^{-\lambda_1(t-s)}\,ds.\end{aligned}$$

We now let $n \to +\infty$ and conclude by the convergence of the left-hand side series in $L^2(\Omega)$ and by the Lebesgue monotone convergence theorem for the right-hand side term. □

Remark 6.5.5 We recover the exponential decay of the energy when $f = 0$. □

6.6 The heat equation on $\mathbb{R}$

Even though it is unphysical, the heat equation on $\mathbb{R}^d$ is nonetheless interesting from the point of view of mathematics. We will consider the case $d = 1$, with the straightforward extension to a general d left to the reader. Let us thus consider the initial value problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) - \dfrac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) \text{ in } \mathbb{R}\times\,]0,T[,\\ u(x,0) = u_0(x) \text{ on } \mathbb{R}.\end{cases} \tag{6.11}$$

Note that there is no boundary data since $\mathbb{R}$ has no boundary. They are going to be replaced by some kind of asymptotic behavior at infinity.

Let us introduce an extremely important function.

Definition 6.6.1 *The function defined on* $\mathbb{R}^2$ *by*

$$E(x,t) = \begin{cases} \dfrac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \textit{ for } t > 0, \\ 0 \textit{ for } t \leq 0, \end{cases} \tag{6.12}$$

is called the (one-dimensional) heat kernel.

We note that for $t > 0$ fixed, the function $x \mapsto E(x,t)$ is a Gaussian. When $t \to 0^+$, the Gaussian becomes increasingly spiked. Indeed, we see that $E(x,t) = \frac{1}{\sqrt{4t}} E\left(\frac{x}{\sqrt{4t}}, \frac{1}{4}\right)$. In particular, $E(0,t) \to +\infty$, whereas $E(x,t) \to 0$ for all $x \neq 0$.

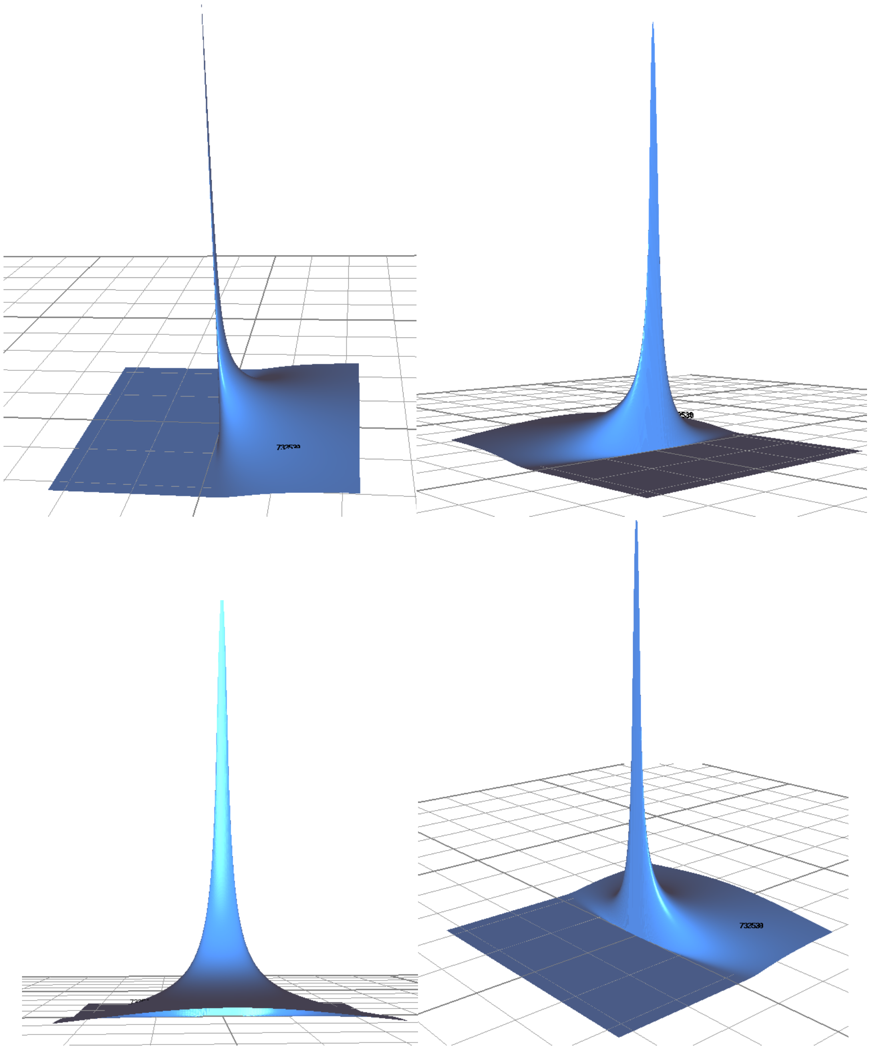


Figure 7. Various views of the graph of the heat kernel.

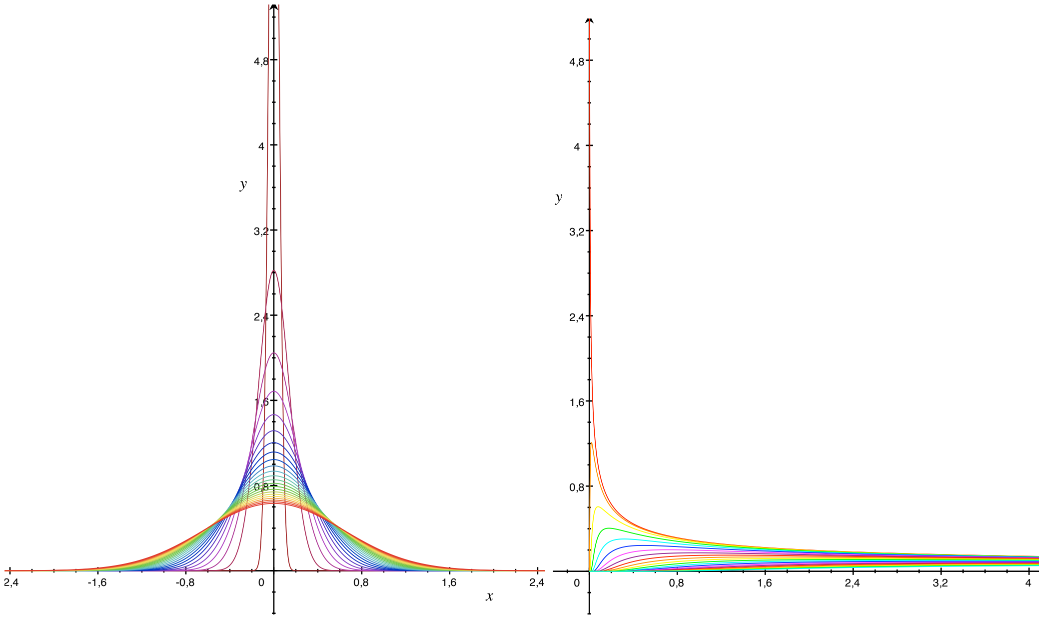


Figure 8. The heat kernel at t fixed (left) and x fixed (right).

Proposition 6.6.1 *We have $E \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$, hence $E \in \mathscr{D}'(\mathbb{R}^2)$.*

Proof. Clearly $E \in C^\infty(\mathbb{R}^2 \setminus \{(0,0)\})$, therefore the only potential local integrability problem is in a compact neighborhood of $(0,0)$. Since E takes positive values, it suffices to integrate it on the square $[-a,a]^2$ for some $a > 0$. Since E vanishes for $t \leq 0$, only the upper half square is left. We have

$$\begin{aligned}\int_{-a}^{a}\int_{0}^{a} E(x,t)\,dxdt &\leq \int_{-\infty}^{+\infty}\int_{0}^{a} E(x,t)\,dxdt\\ &= \frac{1}{2\sqrt{\pi}}\int_{-\infty}^{+\infty}\int_{0}^{a}\frac{1}{\sqrt{t}}e^{-\frac{x^2}{4t}}\,dxdt\\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\int_{0}^{a} e^{-y^2}\,dydt = a < +\infty\end{aligned}$$

where we have performed the change of variables $x = 2\sqrt{t}y$ and because the Gaussian integral value is well-known, $\int_{-\infty}^{+\infty} e^{-y^2}\,dy = \sqrt{\pi}$. □

The heat kernel is the *fundamental solution* or *elementary solution* of the heat equation.

Proposition 6.6.2 *We have*

$$\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2} = \delta_0, \tag{6.13}$$

where δ_0 is the Dirac distribution at 0.

Proof. Given $\varphi \in \mathscr{D}(\mathbb{R}^2)$, our goal is to compute the value of the duality bracket

$$\Big\langle \frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}, \varphi \Big\rangle = \Big\langle E, -\frac{\partial \varphi}{\partial t} - \frac{\partial^2 \varphi}{\partial x^2} \Big\rangle.$$

We have already noticed that E is of class C^∞ everywhere except at $(x,t) = (0,0)$. Its distributional derivatives thus coincide with its classical derivatives on $\mathbb{R}^2 \setminus \{(0,0)\}$. Let us first compute these derivatives using brute force for $t > 0$ (only mild force is needed for $t < 0$). We thus have

$$\begin{aligned}
\frac{\partial E}{\partial t} &= \frac{1}{2\sqrt{\pi}}\Big(-\frac{1}{2t^{3/2}} + \frac{x^2}{4t^{5/2}}\Big)e^{-\frac{x^2}{4t}},\\
\frac{\partial E}{\partial x} &= -\frac{1}{4\sqrt{\pi}}\frac{x}{t^{3/2}}e^{-\frac{x^2}{4t}},\\
\frac{\partial^2 E}{\partial x^2} &= -\frac{1}{4\sqrt{\pi}}\Big(\frac{1}{t^{3/2}} - \frac{x^2}{2t^{5/2}}\Big)e^{-\frac{x^2}{4t}},
\end{aligned}$$

so that $\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2} = 0$ on $\mathbb{R}^2 \setminus \{(0,0)\}$. Therefore the support of the distribution $\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}$ is included in $\{(0,0)\}$.

Let us now work in the distributional sense. We take a test-function $\varphi \in \mathscr{D}(\mathbb{R}^2)$. We have

$$\begin{aligned}
\Big\langle \frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}, \varphi \Big\rangle &= -\Big\langle E, \frac{\partial \varphi}{\partial t} + \frac{\partial^2 \varphi}{\partial x^2} \Big\rangle\\
&= -\int_{-\infty}^{+\infty}\int_0^{+\infty} E(x,t)\Big(\frac{\partial \varphi}{\partial t} + \frac{\partial^2 \varphi}{\partial x^2}\Big)(x,t)\,dt\,dx,
\end{aligned}$$

since E is L^1_{loc} and vanishes for $t \le 0$. The derivatives $\frac{\partial \varphi}{\partial t}$ and $\frac{\partial^2 \varphi}{\partial x^2}$ have compact support, hence $E\big(\frac{\partial \varphi}{\partial t} + \frac{\partial^2 \varphi}{\partial x^2}\big)$ is in $L^1(\mathbb{R}^2)$ and the Lebesgue dominated convergence theorem implies that

$$\Big\langle \frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}, \varphi \Big\rangle = -\lim_{n\to+\infty}\int_{-\infty}^{+\infty}\int_{\frac{1}{n}}^{+\infty} E(x,t)\Big(\frac{\partial \varphi}{\partial t} + \frac{\partial^2 \varphi}{\partial x^2}\Big)(x,t)\,dt\,dx.$$

Now on the set $\mathbb{R} \times [\frac{1}{n}, +\infty[$, all the functions are C^∞ and we can integrate by parts, so that

$$\begin{aligned}
\Big\langle \frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}, \varphi \Big\rangle = \lim_{n\to+\infty}\Big\{\int_{-\infty}^{+\infty}\int_{\frac{1}{n}}^{+\infty}\Big(\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}\Big)(x,t)\varphi(x,t)\,dt\,dx\\
+ \int_{-\infty}^{+\infty} E(x,n^{-1})\varphi(x,n^{-1})\,dx\Big\}.
\end{aligned}$$

We have already seen that $\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2} = 0$ on $\mathbb{R} \times [\frac{1}{n}, +\infty[$ so that the first integral vanishes. Let us study the second integral. We perform the change of variables $y = \frac{\sqrt{n}x}{2}$. This yields

$$\int_{-\infty}^{+\infty} E(x, n^{-1})\varphi(x, n^{-1})\,dx = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-y^2}\varphi(2n^{-1/2}y, n^{-1})\,dy \to \varphi(0,0)$$

by the dominated convergence theorem. Therefore, we have shown that

$$\Big\langle \frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2}, \varphi \Big\rangle = \varphi(0,0) = \langle \delta_0, \varphi \rangle,$$

and the proposition is proved. □

The heat kernel can be used to express the solution in various function spaces. Let us give an example.

Proposition 6.6.3 *Let $u_0 \in L^1(\mathbb{R})$ and $f \in L^1(\mathbb{R} \times \mathbb{R}_+)$. Then*

$$u(x,t) = \int_{-\infty}^{+\infty} E(x-y,t)u_0(y)\,dy + \int_{-\infty}^{+\infty}\int_0^t E(x-y,t-s)f(y,s)\,dsdy \quad (6.14)$$

is a solution of problem (6.11).

Proof. We write the proof for u_0 and f continuous and bounded, for simplicity. First of all, all the integrals make sense and define a function on $\mathbb{R} \times \mathbb{R}_+^*$. Moreover, it is easy to check that all partial derivatives of the heat kernel are integrable on $\mathbb{R} \times [a, +\infty[$ for all $a > 0$. Therefore, we can differentiate under the integral signs without any problems as soon as the second argument of E stays bounded away from 0. Let us set

$$v(x,t) = \int_{-\infty}^{+\infty} E(x-y,t)u_0(y)\,dy$$

and

$$w(x,t) = \int_{-\infty}^{+\infty}\int_0^t E(x-y,t-s)f(y,s)\,dsdy.$$

By the observation above, we have that

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = \int_{-\infty}^{+\infty}\Big(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\Big)E(x-y,t)u_0(y)\,dy = 0$$

for all $t > 0$. We need to exert a little more care to deal with w. Setting

$$w_n(x,t) = \int_{-\infty}^{+\infty}\int_0^{t-1/n} E(x-y,t-s)f(y,s)\,dsdy,$$

we see that $w_n \to w$ in the sense of $\mathscr{D}'(\mathbb{R}\times\mathbb{R}_+^*)$ by the dominated convergence theorem. Therefore

$$\frac{\partial w_n}{\partial t}-\frac{\partial^2 w_n}{\partial x^2}\to\frac{\partial w}{\partial t}-\frac{\partial^2 w}{\partial x^2} \text{ in } \mathscr{D}'(\mathbb{R}\times\mathbb{R}_+^*).$$

We have no problem computing the effect of the heat operator on w_n:

$$\begin{aligned}\frac{\partial w_n}{\partial t}-\frac{\partial^2 w_n}{\partial x^2}&=\int_{-\infty}^{+\infty}\int_0^{t-1/n}\Big(\frac{\partial}{\partial t}-\frac{\partial^2}{\partial x^2}\Big)E(x-y,t-s)f(y,s)\,dsdy\\&\qquad+\int_{-\infty}^{+\infty}E(x-y,n^{-1})f(y,t-n^{-1})\,dy\\&=\int_{-\infty}^{+\infty}E(x-y,n^{-1})f(y,t-n^{-1})\,dy\to f(x,t)\end{aligned}$$

simply, and even uniformly on compact sets, as we have essentially already seen before. Hence

$$\frac{\partial w}{\partial t}-\frac{\partial^2 w}{\partial x^2}=f.$$

Concerning the initial condition, we obviously have $w(x,0)=0$ and $w(x,t)\to 0$ when $t\to 0$. Indeed, $\big|\int_{-\infty}^{+\infty}E(x-y,t-s)f(y,s)\,dy\big|\le\sup|f|$. On the other hand, we have $v(x,t)\to u_0(x)$ when $t\to 0$ by the same change of variable as above. □

Remark 6.6.1 i) The analysis is a little more difficult when u_0 or f are not continuous.

ii) We have not made precise in which space the above solution is unique.

iii) The solution u is C^∞ on any open set of $\mathbb{R}\times\mathbb{R}_+^*$ where f is C^∞. This is again hypoellipticity.

iv) When u_0 is in L^∞, the result remains. In particular, when u_0 is periodic and $f=0$, then u is also periodic in x, and if u_0 is periodic and vanishes at some point, we obtain the same solution as the one obtained via Fourier series by restricting u to a space-time strip based on a period.

v) Notice an interesting phenomenon: when u_0 is positive with compact support and $f=0$, we have $u(x,t)>0$ for all $x\in\mathbb{R}$ and all $t>0$. In other words, a compactly supported initial distribution of temperature instantly spreads to the whole of $\mathbb{R}$. Thus, the heat equation propagates energy at infinite speed, which is strongly non physical. However, the validity of the heat equation as a model of temperature evolution is still extremely good for all classical physics and engineering applications. □