Chapter 3

The heat equation

The Fourier transform was originally introduced by Joseph Fourier in an 1807 paper in order to construct a solution of the heat equation on an interval $0 < x < 2\pi$, and we will also use it to do something similar for the equation

$$\begin{aligned} &\partial_t u = \tfrac{1}{2}\partial_x^2 u \ , \qquad t \in \mathbb{R}^1_+ \ , \quad x \in \mathbb{R}^1 \\ &u(0,x) = f(x) \ , \end{aligned} \tag{3.1}$$

The first thing to notice is that the equation (3.1) is invariant under *Brownian scaling*, which is the change of variables

$$t' = \varepsilon t \ , \quad x' = \sqrt{\varepsilon} x \ .$$

This rescaling transformation is closely related to certain principles of probability and the properties of Brownian motion. To see how the equation behaves, the change of variables gives that

$$\partial_t = \varepsilon \partial_{t'} \ , \quad \partial_x = \sqrt{\varepsilon}\partial_{x'} \ ,$$

and therefore if $u(t,x)$ solves the heat equation (3.1) then so does $u' := u'(t',x') = u(\varepsilon t, \sqrt{\varepsilon} x)$;

$$\varepsilon\big(\partial_{t'} u' - \tfrac{1}{2}\partial_{x'}^2\big)u'(t',x') = \big(\partial_t - \tfrac{1}{2}\partial_x^2\big)u(\varepsilon t, \sqrt{\varepsilon}x) = 0 \ .$$

We will see that the ratio $x/\sqrt{t}$, which is invariant under Brownian scaling, appears in a central way in expressions for the solution, and indeed it plays an related role in probability.

3.1. The heat kernel

A derivation of the solution of (3.1) by Fourier synthesis starts with the assumption that the solution $u(t,x)$ is sufficiently well behaved that is satisfies the hypotheses of the Fourier inversaion formula. This will be verified *a postiori.* Writing

$$u(t,x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\xi x} \hat{u}(t,\xi)\, d\xi \ ,$$

then solutions of (3.1) must satisfy

$$0 = \big(\partial_t - \tfrac{1}{2}\partial_x^2\big) u(t,x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\xi x} \big(\partial_t - \tfrac{1}{2}(i\xi)^2\big) \hat{u}(t,\xi)\, d\xi \ .$$

We have used Proposition 2.3 to express the spatial derivatives in terms of the Fourier transform. Reasoning as before, the Fourier transform of the solution must satisfy the family of ODEs

$$\frac{d}{dt}\hat{u} + \tfrac{1}{2}\xi^2 \hat{u} = 0 \ ,$$

parametrized by ξ. The solution is

$$\hat{u}(t,\xi) = e^{-\frac{1}{2}\xi^2 t} \hat{u}(0,\xi) = e^{-\frac{1}{2}\xi^2 t} \hat{f}(\xi) \ ,$$

where $\hat{f}(\xi)$ is the Fourier transform of the initial data. With this we can express the solution in integral operator form;

$$u(t,x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\xi x} \big(e^{-\frac{1}{2}\xi^2 t} \hat{f}(\xi)\big)\, d\xi \tag{3.2}$$

$$= \frac{1}{2\pi} \iint e^{i\xi(x-y)} e^{-\frac{1}{2}\xi^2 t} f(y)\, dy d\xi \tag{3.3}$$

$$\int \Big(\frac{1}{2\pi} \int e^{i\xi(x-y)} e^{-\frac{1}{2}\xi^2 t}\, d\xi\Big) f(y)\, dy := \int H(t, x-y) f(y)\, dy \ . \tag{3.4}$$

Justification of the exchanges of the order of integration comes from the Fubini theorem, as long as the initial data satisfies $f \in L^1(\mathbb{R}^1)$. The function $H(t, x-y)$ is the *heat kernel*, the integral kernel for the solution operator $\mathbf{H}(t)$ for the heat equation with the initial data $f(x)$;

$$u(t,x) = \big(\mathbf{H}(t) f\big)(x) \ .$$

Figure 1. Superposition of the graph of the heat kernel at several times.

Complete the square to evaluate this Fourier integral definition of the heat kernel;

$$\begin{aligned} H(t,x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\xi x} e^{-\frac{1}{2}\xi^2 t}\, d\xi \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}x^2/t} \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{\frac{1}{2}x^2/t + i\xi x - \frac{1}{2}\xi^2 t}\, d\xi \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}x^2/t} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\xi\sqrt{t} - ix/\sqrt{t})^2}\, d\xi \ . \end{aligned}$$

Substitute $\xi' = \xi\sqrt{t} - ix/\sqrt{t}$ (which is actually an application of Cauchy's theorem of complex analysis), we find the expression

$$H(t,x) = \frac{1}{2\pi} e^{-\frac{1}{2}x^2/t} \int e^{-\frac{1}{2}(\xi')^2} \frac{d\xi'}{\sqrt{t}} d\xi' = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}x^2/t} \ . \tag{3.5}$$

Theorem 3.1. *The solution of the heat equation for initial data $f \in L^1(\mathbb{R}^1)$ is given by the convolution of the initial data with the heat kernel;*

$$u(t,x) = \int_{-\infty}^{+\infty} H(t, x-y) f(y)\, dy = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}(x-y)^2/t} f(y)\, dy \ .$$

Take note of the invariance under Brownian scaling of the quantity

$$\frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}x^2/t}\, dx \ ,$$

which is the Gaussian measure that is relevant to the heat equation.

This procedure via the heat kernel gives a unique solution among the class of bounded solutions, or among solutions which are in $L^1(\mathbb{R}^1_x)$, or simply among solutions which do not grow too rapidly as $|x| \to +\infty$. It is a surprising fact, which will come up in a later discussion, that solutions of the heat equation are not unique if we admit ones that are badly unbounded [A. N. Tykonov (1977)].

The heat kernel $H(t, x-y)$ for $t > 0$ has its maximum at $x = y$, where the maximim value is $\frac{1}{\sqrt{2\pi t}}$, and width $\sqrt{t}$ (defined as the distance between its center and inflection point). A graph of the heat kernel, taken at several snapshots in time, is given in Figure 1.

It is reasonable to expect that we will obtain the same solution $u(t,x)$ at time $t>0$ starting with data $f(x)$ at time $t=0$, if we alternatively solved the heat equations up to time $0<s<t$, and then starting again at s with data $u(s,x)$, continued the solution for time $t-s$. That, is, in operator notation

$$\mathbf{H}(t)f = \mathbf{H}(t-s)\mathbf{H}(s) \ , \qquad 0 < s < t \ . \tag{3.6}$$

This is the *semigroup property*, and it is related to the question of uniqueness.

3.2. Convolution operators

The evolution operator for the heat equation is an example of a convolution operator, with convolution kernel the heat kernel $H(t,x)$.

Definition 3.2. Let $h(x) \in L^1(\mathbb{R}^1)$, and define the *convolution product* with $f(x) \in L^1(\mathbb{R}^1)$ to be

$$\big(h * f\big)(x) = \int_{-\infty}^{+\infty} h(x-x')f(x')\,dx' \ .$$

In these terms the evolution operator for the heat equation can be written as

$$u(t,x) = \mathbf{H}(t)f(x) = \big(H(t,\cdot) * f\big)(x) = \int_{-\infty}^{+\infty} H(t,x-x')f(x')\,dx' \ .$$

Convolution operators with kernels $h \in L^1(\mathbb{R}^1)$ have a number of convenient features, the most elementary ones are covered in the following proposition.

Proposition 3.3. Let $h(x), f(x), g(x) \in L^1(\mathbb{R}^1)$. Then
(i) The convolution product is commutative and associative;

$$\begin{aligned} \big(h * f\big)(x) &= \big(f * h\big)(x) \\ h * \big(g * f\big) &= \big(h * g\big) * f(x) \ . \end{aligned}$$

(ii) If in addition $\partial_x f \in L^1(\mathbb{R}^1)$ then differentiation commutes with the operation of convolution;

$$\partial_x\big(h * f\big)(x) = h * \big(\partial_x f\big)(x) \ .$$

(iii) Under Fourier transform, convolution with $h(x)$ becomes a multiplication operator

$$\widehat{h * f}(\xi) = \sqrt{2\pi}\hat{h}(\xi)\hat{f}(\xi) \ .$$

Proof. Address the statement (iii) first of all;

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-i\xi x}\Big(\int_{-\infty}^{+\infty} h(x-x')f(x')\,dx'\Big)\,dx$$
$$= \frac{1}{\sqrt{2\pi}}\iint e^{-i\xi(x-x')}h(x-x')e^{-i\xi x'}f(x')\,dx'dx\ ,$$

by using Fubini's Theorem. Then this quantity can be rewritten as

$$= \sqrt{2\pi}\Big(\frac{1}{\sqrt{2\pi}}\int e^{-i\xi x'}f(x')\Big(\frac{1}{\sqrt{2\pi}}\int e^{-i\xi(x-x')}h(x-x')\,dx\Big)\,dx'\Big)$$
$$= \sqrt{2\pi}\hat{f}(\xi)\hat{h}(\xi)\ ,$$

which is the statement of (iii). This can now be used to prove (i) and (ii). Indeed, since

$$\widehat{h*f}(\xi) = \sqrt{2\pi}\hat{h}(\xi)\hat{f}(\xi)$$

the order of multiplication does not play a role, and therefore the properties of commutativity and associativity clearly hold. Furthermore,

$$\widehat{\partial_x(h*f)}(\xi) = i\xi\widehat{h*f}(\xi) = \sqrt{2\pi}(i\xi)\hat{h}(\xi)\hat{f}(\xi)\ ,$$

and the derivative can be seen either to be acting on the convolution product or on f alone. The fact is that the Fourier transform has simultaneously diagonalized the operations of convolution with $h(x)$ and differentiation, from which the results of the proposition are correlaries. $\square$

3.3. The maximum principle

This fundamental principle is a feature of solutions of parabolic equations such as the heat equation; we will encounter it as well when we take up the topic of elliptic equations such as Lapace's equation.

Theorem 3.4 (maximum principle)**.** *Suppose that the initial data satisfies $f(x) \geq 0$ for all $x \in \mathbb{R}^1$. Then either $u(t,x) > 0$ for all $t > 0$, or else both $u(t,x) \equiv 0$ and $f(x) \equiv 0$.*

The conclusion is powerful because of the fact that the existence of one zero of the solution implies that both the whole solution and the initial data mush vanish identically.

Proof. Observe that the heat kernel $H(t,x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{1}{2}x^2/t} > 0$ for all $t > 0$. Hence if $f(x) \geq 0$ then clearly

$$u(t,x) = \big(H(t,\cdot)*f\big)(x) = \int_{-\infty}^{+\infty} H(t,x-x')f(x')\,dx' \geq 0\ .$$

This non-strict inequality is called the *weak maximum principle*. By a further argument we can prove that $u(t,x) > 0$ unless $f \equiv 0$. Suppose that

$f(x) \geq 0$ and define sets $B_\delta = \{x : f(x) > \delta\}$. If f is not identically zero (excluding sets of zero Lebesgue measure) then B_δ is nonempty for some δ and there is a bounded set $A \subseteq B_\delta$ with positive measure. Then

$$\begin{aligned} u(t,x) &= \int_{-\infty}^{+\infty} H(t, x-x') f(x')\, dx' \\ &\geq \int_A H(t, x-x') f(x')\, dx' \\ &\geq \int_A H(t, x-x') \delta\, dx' \ . \end{aligned}$$

This last expression is surely positive, by the positive character of the heat kernel $H(t,x)$. There is more precise information available on a lower bound for $u(t,x)$ if we consider the Gaussian nature of the heat kernel;

$$u(t,x) \geq \delta \operatorname{meas}(A) \inf_{x' \in A} \Big(\frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-x')^2}{2t}} \Big) \ .$$

$\square$

3.4. Conservation laws and the evolution of moments

The heat equation is often called the *diffusion equation*, and indeed the physical interpretation of a solution is of a heat distribution or a particle density distribution that is evolving in time according to equation (3.1). That is, in probabilistic terms, the quantity

$$P_t[a,b) = \int_a^b u(t,x)\, dx$$

represents the probability of the outcome of a random event taking its value in the interval $[a,b)$ at time t. For instance $P_t[a,b)$ might represent the probability that a random particle lies in the interval at that time. For this interpretation we should normalize the initial data $f(x)$ so that $\int_{\mathbb{R}^1} f(x)\, dx = 1$, and in particular we should require that $f \in L^1(\mathbb{R}^1)$. We can derive the heat equation itself from this interpretation, given one more piece of information, namely Fourier's law of heat conservation and flux. Take $[a,b)$ to be an arbitrary interval, the change in heat over this interval in time is given by

$$\partial_t P_t[a,b) = \partial_t \int_a^b u(t,x)\, dx = F(a) - F(b) \ ,$$

where Fourier's law of heat conservation posits that this change is given by the flux of heat across the boundary. Namely, $F(a)$ is interpreted to be the flux of heat across the boundary point a into the interval, and $F(b)$ is interpreted to be the flux of hear out of the interval across the boundary point b. Fourier's law for the heat flux gives the form as

$$F(a) := -\tfrac{1}{2}\partial_x u(t,a) \ ,$$

and similarly for b, so that

$$\begin{aligned}\partial_t P_t[a,b) &= \int_a^b \partial_t u(t,x)\,dx = F(a) - F(b) \\ &= \tfrac{1}{2}\partial_x u(t,b) - \tfrac{1}{2}\partial_x u(t,a) = \int_a^b \tfrac{1}{2}\partial_x^2 u(t,x)\,dx \ .\end{aligned}$$

Since the interval $[a,b)$ is arbitrary, the heat density $u(t,x)$ must satisfy the heat equation $\partial_t u = \frac{1}{2}\partial_x^2 u$.

This interpretation must satisfy two conditions in order to be consistent; (i) the total amount of heat must be conserved, and (ii) if the initial heat distribution $f(x) \geq 0$ then for all $t \geq 0$ we should have $u(t,x) \geq 0$. The second condition is already satisfied for us by the maximum principle. Regarding the first, there is the following result.

Theorem 3.5. *If $u(t,x) \in L^1(\mathbb{R}^1)$ is a solution of* (3.1) *then for all $t \geq 0$*

$$\int_{-\infty}^{+\infty} u(t,x)\,dx = \int_{-\infty}^{+\infty} f(x)\,dx \ .$$

Proof.

$$\begin{aligned}\partial_t \int_{-\infty}^{+\infty} u(t,x)\,dx &= \int_{-\infty}^{+\infty} \partial_t u(t,x)\,dx \\ &= \int_{-\infty}^{+\infty} \tfrac{1}{2}\partial_x^2 u(t,x)\,dx = \tfrac{1}{2}\lim_{R\to+\infty}\big(\partial_x u(t,R) - \partial_x u(t,-R)\big) = 0 \ .\end{aligned}$$

□

In fact the proof is incomplete for we should also have shown that when $f(x) \in L^1(\mathbb{R}^1)$ then for $t > 0$ we have $\partial_x^2 u(t,x) \in L^1(\mathbb{R}^1)$, and then justify the exchanges of differentiation and integration and the integrations by parts. These facts will be addressed in the later Chapter on the properties of the Fourier transform.

The interpretation of a solution $u(t,x)$ of (3.1) as the density of a probability measure motivates a discussion of the time evolution of its moments. For a given distribution $f(x)\,dx$ define the k^{th} moment to be the quantity

$$m_k(f) = \int_{-\infty}^{+\infty} x^k f(x)\,dx \ . \tag{3.7}$$

For an arbitrary $f(x) \in L^1(\mathbb{R}^1)$ this may be infinite or not well defined, but for suficiently well localized functions it makes perfectly good sense. To give examples, take the case of the heat kernel itself, then

$$m_0(H(t,\cdot)) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx = 1$$
$$m_1(H(t,\cdot)) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} x e^{-\frac{x^2}{2t}}\, dx = 0$$
$$m_2(H(t,\cdot)) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} x^2 e^{-\frac{x^2}{2t}}\, dx = t \ .$$

Now consider the initial vaue problem for the heat equation, our solution procedure yields $u(t,x) = H(t,\cdot) * f(x)$, and the resulting heat distribution $u(t,x)\,dx$ has moments which are given by

$$m_k(u(t,\cdot)) = \int_{-\infty}^{+\infty} x^k u(t,x)\, dx = \iint x^k \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-x')^2}{2t}} f(x')\, dx' dx \ .$$

Proposition 3.6. The first two moments of the solution $u(t,x)$ are conserved;

$$m_0(u(t,\cdot)) = m_0(f) \ , \qquad m_1(u(t,\cdot)) = m_1(f) \ .$$

To deduce the evolution properties of the higher moments of $u(t,x)\,dx$ we will derive further elementary properties of the Fourier transform.

Proposition 3.7. For sufficiently well behaved functions $g(x)$, the Fourier transform has the property that

$$\mathcal{F}(x^k g(x)) = (i\xi)^k \hat{g}(\xi) \ ,$$
$$\int_{-\infty}^{+\infty} g(x)\, dx = \sqrt{2\pi}\hat{g}(0) \ .$$

Proof. This first statement is essentially dual to that of Proposition 2.3. Observe by integrations by parts that

$$\int e^{-i\xi x}\big(x^k g(x)\big)\, dx = \int \big((i\partial_\xi)^k e^{-i\xi x}\big) g(x)\, dx = (i\partial_\xi)^k \int e^{-i\xi x} g(x)\, dx \ ,$$

which is the first statement of the proposition. To show the second statement,

$$\int g(x)\, dx = \int e^{-i\xi x} g(x)\, dx\Big|_{\xi=0} = \sqrt{2\pi}\hat{g}(0) \ .$$

Hypotheses on the function $g(x)$ are necessary in order to justify the integrations by parts and the limits. □

We use this information in order to compute moments; since

$$\widehat{x^k u}(t,\xi) = (i\partial_\xi)^k \hat{u}(t,\xi) = \sqrt{2\pi}(i\partial_\xi)^k \big(\hat{H}(t,\xi)\hat{f}(\xi)\big) \ ,$$

therefore

$$m_k(u(t,\cdot)) = \int x^k u(t,x)\,dx = \sqrt{2\pi}\big((i\partial_\xi)^k \hat{u}(t,\xi)\big)\big|_{\xi=0}$$
$$= 2\pi\big((i\partial_\xi)^k \hat{H}(t,\xi)\hat{f}(\xi)\big)\big|_{\xi=0} .$$

In particular for $k = 0$,

$$m_0(u(t,\cdot)) = 2\pi\hat{H}(t,0)\hat{f}(0) = \sqrt{2\pi}\hat{f}(0) = m_0(f) ,$$

which gives the first statement of Proposition 3.6. For $k = 1$,

$$\begin{aligned} m_1(u(t,\cdot)) &= 2\pi i\partial_\xi\big(\hat{H}(t,\xi)\hat{f}(\xi)\big)\big|_{\xi=0} \\ &= 2\pi i\big(\partial_\xi\hat{H}(t,\xi)\hat{f}(\xi) + \hat{H}(t,\xi)\partial_\xi\hat{f}(\xi)\big)\big|_{\xi=0} \\ &= 2\pi\hat{H}(t,0)\partial_\xi\hat{f}(\xi)\big|_{\xi=0} = m_1(f) . \end{aligned}$$

This proves the second statement of Proposition 3.6, for initial data with finite zeroth and first moments, at least. The general case is now clear;

$$\begin{aligned} m_k(u(t,\cdot)) &= 2\pi(i\partial_\xi)^k\big(\hat{H}(t,\xi)\hat{f}(\xi)\big)|_{\xi=0} \\ &= 2\pi i^k \sum_{j=0}^{k} \binom{j}{k} \big(\partial_\xi^j\hat{H}(t,\xi)\partial_\xi^{k-j}\hat{f}(\xi)\big)|_{\xi=0} \\ &= \sum_{\substack{j=0 \\ j \text{ even}}}^{k} m_j(H(t,\cdot))m_{k-j}(f) . \end{aligned}$$

We used the Leibnitz product rule for for differentiation in the second line, and in the third line we used that a odd moments of the heat kernel are zero, as the kernel itself is an even function of x.

A property of solutions of the heat equation, related to that of conservation of heat as in Theorem 3.5, is that the evolution has a certain contraction property on a number of common spaces of functions.

Proposition 3.8. Let $u(t,x)$ solve the heat equation (3.1) with initial data $f \in L^2(\mathbb{R}^1)$. Then the L^2-norm of $u(t,\cdot)$ is a decreasing function of t;

$$\|u(t,\cdot)\|_{L^2} \le \|f\|_{L^2} .$$

Proof. Compute the time derivative of the norm;

$$\begin{aligned} \partial_t\|u(t,\cdot)\|_{L^2}^2 &= \partial_t \int u^2(t,x)\,dx \\ &= \int 2u(t,x)\partial_t u(t,x)\,dx = \int 2u(t,x)\tfrac{1}{2}\partial_x^2 u(t,x)\,dx \\ &= -\int |\nabla u(t,x)|^2\,dx \le 0 . \end{aligned}$$

Thus we find that

$$\|u(t,\cdot)\|_{L^2}^2 - \|f\|_{L^2}^2 = \int_0^t \partial_t \|u(s,\cdot)\|_{L^2}^2 \, ds \le 0 \ .$$

One needs to revisit this calculation in order to justify the integration by parts. □

We have already seen that the setting of $L^1(\mathbb{R}^1)$ functions is sometimes more natural for the heat equation than that of $L^1(\mathbb{R}^1)$. It turns out that time evolution by heat flow is also a contraction in $L^1(\mathbb{R}^1)$.

Proposition 3.9. Let $f \in L^1(\mathbb{R}^1)$ and let $u(t,x)$ solve the heat equation with initial data f. Then $\|u(t,\cdot)\|_{L^1}$ is a nonincreasing function of t;

$$\|u(t,\cdot)\|_{L^1} \le \|f\|_{L^1} \ .$$

Proof. The easy case is for $f \in L^1$ to be of one sign; say for convenience that $f(x) \ge 0$. The the maximum principle implies that $u(t,x) \ge 0$ as well, so that

$$\begin{aligned} \partial_t \int |u(t,x)| \, dx &= \partial_t \int u(t,x) \, dx \\ &= \int \partial_t \, u(t,x) \, dx = \int \tfrac{1}{2} \partial_x^2 u(t,x) \, dx = 0 \ , \end{aligned}$$

and the L^1-norm is unchanged by the flow. However when the initial data $f(x)$ changes sign, the L^1-norm will in fact be decreasing in time. To show this, decompose the initial data $f(x) = f_+(x) - f_-(x)$, respectively its positive and negative parts. Both $f_\pm(x) \ge 0$, and furthermore they have essentially disjoint support; $\{f_+ \neq 0\} \cap \{f_- \neq 0\} = \emptyset$. The L^1-norm of f is then the sum of the L^1-norms; indeed

$$\int |f(x)| \, dx = \int (f_+(x) + f_-(x)) \, dx = \int_{\mathrm{supp}(f_+)} f_+ \, dx + \int_{\mathrm{supp}(f_-)} f_- \, dx \ .$$

The previous result can now be applied to the two initial data, respectively f_+ and f_-,

$$\|u_+(t,\cdot)\|_{L^1} = \|f_+\|_{L^1} \ , \qquad \|u_-(t,\cdot)\|_{L^1} = \|f_-\|_{L^1} \ .$$

The solution is obtained from the sum, $u(t,x) = u_+(t,x) - u_-(t,x)$, however the maximum principle applies, so that as long as neither of $f_\pm$ is zero almost everywhere, then $\mathrm{supp}(u_+) = \mathrm{supp}(u_-) = \mathbb{R}^1$. We therefore have the estimate

$$\begin{aligned} \|u(t,\cdot)\|_{L^1} &= \|u_+ - u_-\|_{L^1} \\ &\le \|u_+\|_{L^1} + \|u_-\|_{L^1} = \|f_+\|_{L^1} + \|f_-\|_{L^1} = \|f\|_{L^1} \ . \end{aligned}$$

Since we are assuming that both u_+ and u_- are not zero, then there must be some cancellation in the inequality of the second line, and hence the inequality is in fact strict. $\square$

Most of the results of this chapter hold in general for the heat equation posed in higher dimensional space $\mathbb{R}^n$, for any dimension n.

3.5. Gradient flow

The last section of the chater has to do with an interpretation of the heat equations as a gradient flow for the *energy functional*

$$E(u) = \frac{1}{4}\int |\nabla u|^2 \, dx \ . \tag{3.8}$$

To illustrate gradient flow with a finite dimensional example, let

$$e(v) \, : \, \mathbb{R}^d \to \mathbb{R}$$

be a C^1 function of $v \in \mathbb{R}^d$. Its gradient $\nabla e(v)$ (defined in terms of the Euclidian inner product on $\mathbb{R}^d$) is given by the formula

$$\frac{d}{d\sigma}\Big|_{\sigma=0} e(v+\sigma V) = \langle \nabla e(v), V\rangle \ .$$

The (negative) gradient flow of e is the solution map of the ordinary differential equation

$$\dot{v} = -\nabla e(v) \ .$$

The function $e(v)$ is decreasing along trajectories of this equation, indeed

$$\frac{d}{dt} e(v(t)) = \langle \nabla e(v), \dot{v}(t)\rangle = -\langle \nabla e(v), \nabla e(v)\rangle \leq 0 \ .$$

Such flows play a role in Morse theory, among many other things.

The heat equation defines a flow in space of functions. Using a similar calculation, we can identify this as a gradient flow for the energy functional $E(u)$.

$$\begin{aligned}
&\frac{d}{d\sigma}\Big|_{\sigma=0} E(u+\sigma w) \\
&= \frac{d}{d\sigma}\Big|_{\sigma=0} \frac{1}{4}\int |\nabla u + \sigma w|^2 \, dx \\
&= \frac{d}{d\sigma}\Big|_{\sigma=0} \frac{1}{4}\int |\nabla u|^2 + 2\sigma \nabla u \cdot \nabla w + \sigma^2 |\nabla w|^2 \, dx \\
&= \tfrac{1}{2}\int \sigma \nabla u \cdot \nabla w \, dx \ .
\end{aligned}$$

In order to describe this quantity in terms of the L^2 inner product one integrates by parts;

$$\frac{d}{d\sigma}\Big|_{\sigma=0} E(u+\sigma w) = -\int \tfrac{1}{2}\Delta u \, w \, dx \ ,$$

which identifies $\mathrm{grad} E(u) = -\Delta u$. The (negative) gradient flow is therefore

$$\partial_t u = -\mathrm{grad} E(u) = \tfrac{1}{2}\Delta u \ ,$$

which is precisely the heat equation (1.8).

Exercises: Chapter 3

Exercise 3.1. Show that the heat kernel satisfies the identity

$$H(t,x) = \int_{-\infty}^{+\infty} H(t-s, x-x')H(x',s)\,dx' \ , \quad 0 < s < t \ ,$$

which is a key step of a proof of the semigroup property (3.6) for the solution operator $\mathbf{H}(t)$ of the heat equation.