Chapter 6
The Heat Equation

In physics, the term *heat* is used to describe the transfer of internal energy within a system of particles. When this transfer results from collective motion of particles in a gas or fluid, the process is called *convection*. The continuity equation developed in Sect. 3.1 describes convection by fluid flow, which is the special case called *advection*. Another form of convection is *conduction*, where the heat transfer caused by random collisions of individual particles.

The basic mathematical model for heat conduction is a PDE called the *heat equation*, developed by Joseph Fourier in the early 19th century. In this chapter we will discuss the derivation and develop some basic properties of this equation, our first example of a PDE of parabolic type.

6.1 Model Problem: Heat Flow in a Metal Rod

A metal rod that is sufficiently thin can be treated as one-dimensional system. Let $u(t,x)$ denote the temperature of the rod at time t and position x, with $x \in \mathbb{R}$ for now.

There are two physical principles that govern the flow of heat in the rod. The first is the relationship between thermal (internal) energy and temperature. Thermal energy is proportional to a product of density and temperature, by a constant c called the *specific heat* of the material. Thus, the total thermal energy in a segment $[a,b]$ is given by

$$\mathcal{U} = c\int_a^b \rho u \, dx. \tag{6.1}$$

We will assume that the density ρ is constant, although it could be variable in some applications.

The original version of the book was revised: Belated corrections from author have been incorporated. The erratum to the book is available at https://doi.org/10.1007/978-3-319-48936-0_14

© Springer International Publishing AG 2016

D. Borthwick, *Introduction to Partial Differential Equations*,
Universitext, DOI 10.1007/978-3-319-48936-0_6

The second principle is Fourier's law of heat conduction, which describes how heat flows from hotter regions to colder regions. In its one-dimensional form, Fourier's law says that the flux of thermal energy across a given point is given by

$$q = -k\frac{\partial u}{\partial x}, \tag{6.2}$$

where the constant $k > 0$ is the *thermal conductivity* of the material.

Assuming that the rod is thermally isolated, conservation of energy dictates the rate of change of the thermal energy within the segment is equal to the flux across its boundaries, i.e.,

$$\frac{d\mathcal{U}}{dt}(t) = q(t,a) - q(t,b). \tag{6.3}$$

As in our derivation of the local equation for conservation of mass, the combination of (6.1) and (6.3) yields an integral equation

$$\int_a^b \left(c\rho\frac{\partial u}{\partial t} + \frac{\partial q}{\partial x}\right) dx = 0.$$

Since a and b were arbitrary, this implies a local conservation law,

$$c\rho\frac{\partial u}{\partial t} + \frac{\partial q}{\partial x} = 0.$$

Using the formula for q from Fourier's law (6.2), we obtain the one-dimensional *heat equation*:

$$\frac{\partial u}{\partial t} - \frac{k}{c\rho}\frac{\partial^2 u}{\partial x^2} = 0. \tag{6.4}$$

For a rod of finite length, the solution u will satisfy boundary conditions that depend on how the rod interacts with its environment. If the rod is parametrized by $x \in [0, \ell]$ and we assume that each end is held at a fixed temperature, then this fixes the values at the endpoints,

$$u(t,0) = T_0, \quad u(t,\ell) = T_1 \tag{6.5}$$

for all t. These are inhomogeneous Dirichlet boundary conditions. In one dimension the inhomogeneous problem can be reduced very simply to the homogeneous case by noting that

$$u_0(x) := T_0\left(1 - \frac{x}{\ell}\right) + T_1\frac{x}{\ell}$$

gives an equilibrium solution to the heat equation satisfying the boundary conditions (6.5). By the superposition principle, $u - u_0$ satisfies the heat equation with homogeneous Dirichlet conditions.

Another possible boundary assumption is that the ends are insulated, so that no thermal energy flows in or out. This means that q vanishes at the boundary, yielding the Neumann boundary conditions

$$\frac{\partial u}{\partial x}(t, 0) = \frac{\partial u}{\partial x}(t, \ell) = 0.$$

Example 6.1 On the bounded interval $[0, \pi]$, we can find product solutions to the heat equation using Lemma 5.1. For the Dirichlet boundary conditions $u(0) = u(\pi) = 0$, Theorem 5.2 gives the set of sine eigenfunctions (5.5). The corresponding heat equation solutions are

$$u(t, x) = e^{-n^2 t} \sin(nx)$$

for $n \in \mathbb{N}$. Note that all of these solutions decay exponentially to 0 as $t \to \infty$.

For insulated ends we switch to Neumann conditions and obtain the cosine modes. The resulting set of solutions is

$$u(t, x) = e^{-n^2 t} \cos(nx)$$

for $n \in \mathbb{N}_0$. In this case the $n = 0$ mode yields a constant solution.

In Chap. 8 we will discuss the construction of series solutions from these trigonometric families. $\lozenge$

The higher dimensional form of the heat equation can be derived by an argument similar to that given above. In $\mathbb{R}^n$, the thermal flux $\boldsymbol{q}$ is vector valued, and Fourier's law becomes the gradient formula

$$\boldsymbol{q} = -k\nabla u.$$

Local conservation of energy is expressed by the continuity equation (3.20),

$$c\rho\frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{q} = 0.$$

In combination, these yield the n-dimensional heat equation,

$$\frac{\partial u}{\partial t} - \frac{k}{c\rho}\Delta u = 0. \tag{6.6}$$

The importance of the heat equation as a model extends well beyond its original thermodynamic context. One of the most prominent examples of this is Albert Einstein's probabilistic derivation of the heat equation as a model for Brownian motion in 1905, in one of the set of papers for which he was later awarded the Nobel prize. Brownian motion is named for the botanist Robert Brown, who observed in 1827 that minute particles ejected by pollen grains drifted erratically when suspended in water, with a jittery motion for which no explanation was available at the time. Einstein theorized that the motion was caused by collisions with a large number of molecules whose velocities were distributed randomly. The existence of atoms and molecules was still unconfirmed in 1905, and Einstein's model provided crucial supporting evidence.

To summarize Einstein's argument, suppose that a total of n particles are distributed on the real line. In an interval of time τ, the position of each particle is assumed to change by a random amount according to a distribution function ϕ. To be more precise, the number of particles experiencing a displacement between σ and $\sigma + d\sigma$ is

$$dn = n\phi(\sigma)\, d\sigma.$$

The total number of particles is conserved, which imposes the condition

$$\int_{-\infty}^{\infty} \phi(\sigma)\, d\sigma = 1. \tag{6.7}$$

The distribution of displacements is assumed to be symmetric, $\phi(\sigma) = \phi(-\sigma)$, meaning that particles are equally likely to move left or right.

Suppose the distribution of particles at time t is given by a density function $\rho(t, x)$. Under the displacement hypothesis, the values of this density at times t and $t + \tau$ are related by

$$\rho(t + \tau, x) = \int_{-\infty}^{\infty} \rho(t, x - \sigma)\phi(\sigma)\, d\sigma. \tag{6.8}$$

To find an equation for ρ, Einstein takes the Taylor expansions of the density on both sides of (6.8), obtaining

$$\rho(t + \tau, x) = \rho(t, x) + \frac{\partial \rho}{\partial t}(t, x)\tau + \dots \tag{6.9}$$

on the left, and

$$\rho(t, x - \sigma) = \rho(t, x) - \frac{\partial \rho}{\partial x}(t, x)\sigma + \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2}(t, x)\sigma^2 + \dots$$

inside the integral. Integrating the latter expansion against ϕ gives, by (6.7) and the assumption that ϕ is even (which knocks out the linear term),

$$\int_{-\infty}^{\infty} \rho(t, x - \sigma)\phi(\sigma)\, d\sigma = \rho(t, x) + \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2}(t, x) \int_{-\infty}^{\infty} \sigma^2 \phi(\sigma)\, d\sigma + \dots.$$

Substituting this formula together with (6.9) into (6.8) and keeping the leading terms gives

$$\frac{\partial \rho}{\partial t}(t, x)\tau = \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2}(t, x) \int_{-\infty}^{\infty} \sigma^2 \phi(\sigma)\, d\sigma.$$

Einstein then assumes that the value

$$D = \frac{1}{2\tau}\int_{-\infty}^{\infty} \sigma^2 \phi(\sigma)\, d\sigma \tag{6.10}$$

is a fixed constant, so that the equation for ρ becomes

$$\frac{\partial \rho}{\partial t} - D\frac{\partial^2 \rho}{\partial x^2} = 0,$$

i.e., the heat equation. Remarkably, the function ϕ representing the random distribution of displacements plays no role in the final equation, except in the value of the constant D. This fact is related to a fundamental result in probability called the central limit theorem.

Diffusion models involving random motions of particles are prevalent in physics, biology, and chemistry. The same statistical principles appear in other applications as well, for example in models of the spread of infection in medicine, or in the study of fluctuating financial markets. The heat equation plays a fundamental role in all of these applications.

6.2 Scale-Invariant Solution

Let us consider the heat equation on $\mathbb{R}$, with physical constants normalized to 1,

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0. \tag{6.11}$$

Note that the equation is invariant under the rescaling $(t, x) \mapsto (\lambda^2 y, \lambda x)$, with λ a nonzero constant. This suggests a change of variables to the scale-invariant ratio $y := x/\sqrt{t}$ might simplify the equation.

Let us try to find a solution of the form $u(t, x) = q(y)$ for $t > 0$. By the chain rule,

$$\frac{\partial u}{\partial t} = -\frac{y}{2t}q', \qquad \frac{\partial^2 u}{\partial x^2} = \frac{1}{t}q''.$$

Thus, as an equation for q, (6.11) reduces to an ODE,

$$q'' = -\frac{y}{2}q'.$$

This can be solved for q' by separation of variables for ODE, as described in Sect. 2.5,

$$q'(y) = q'(0)e^{-y^2/4}.$$

A further integration yields

$$q(y) = q'(0) \int_0^y e^{-y'^2/4}\, dy' + q(0).$$

In the original coordinates this translates to

$$u(t, x) = C_1 \int_0^{\frac{x}{\sqrt{t}}} e^{-y^2/4}\, dy + C_2.$$

It is easy to confirm that this solves (6.11) for $t > 0$. To see what happens as $t \to 0$, note that

$$\int_0^\infty e^{-y^2/4}\, dy = \sqrt{\pi},$$

by the computations from Exercise 2.5. Thus

$$\lim_{t\to 0} u(t, x) = \begin{cases} C_1\sqrt{\pi} + C_2, & x > 0, \\ 0, & x = 0, \\ -C_1\sqrt{\pi} + C_2, & x < 0. \end{cases} \tag{6.12}$$

In view of this limiting behavior, let us consider the particular solution U defined by setting $C_1 = \frac{1}{\sqrt{4\pi}}$, $C_2 = \frac{1}{2}$,

$$U(t, x) := \frac{1}{\sqrt{4\pi}} \int_0^{\frac{x}{\sqrt{t}}} e^{-y^2/4}\, dy + \frac{1}{2}. \tag{6.13}$$

This solution is plotted for some small values of t in Fig. 6.1. By (6.12), $\lim_{t\to 0} U(t, x) = \Theta(x)$, the *Heaviside step function* defined by

$$\Theta(x) := \begin{cases} 1, & x > 0, \\ \frac{1}{2}, & x = 0, \\ 0, & x < 0. \end{cases}$$

The fact that $U(t, x)$ has such a simple limit as $t \to 0$ can be used to derive a more general integral formula. Suppose we want to solve (6.11) for the initial condition

$$u(t, x) = \varphi(x)$$

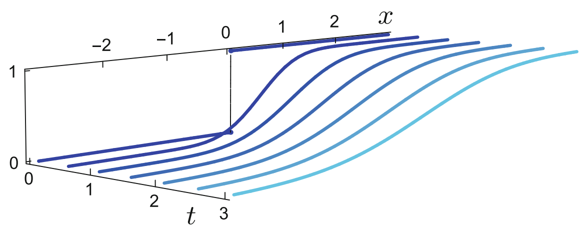

Fig. 6.1 The heat solution $U(t, x)$ for times from $t = 0$ to 3

with $\varphi \in C^{\infty}_{\mathrm{cpt}}(\mathbb{R})$. The key observation is that φ can be reproduced by integrating its derivative against the Heaviside function,

$$\begin{aligned}\int_{-\infty}^{\infty} \varphi'(z)\Theta(x-z)\,dz &= \int_{-\infty}^{x} \varphi(z)\,dx' \\ &= \varphi(x).\end{aligned}$$

This suggests that we could solve the heat equation with initial condition φ by setting

$$u(t,x) := \int_{-\infty}^{\infty} \varphi'(z)U(t,x-z)\,dz.$$

For $t > 0$ the function $U(t, x-z)$ is continuously differentiable, so we can integrate by parts to rewrite this as

$$\begin{aligned}u(t,x) &= -\int_{-\infty}^{\infty} \varphi(z)\frac{\partial U}{\partial z}(t,x-z)\,dz \\ &= \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} \varphi(z)e^{-(x-z)^2/4t}\,dz.\end{aligned} \tag{6.14}$$

In terms of the function

$$H_t(x) := \frac{1}{\sqrt{4\pi t}}e^{-x^2/4t},$$

the solution is

$$u(t,x) = \int_{-\infty}^{\infty} H_t(x-z)\varphi(z)\,dz.$$

This integral is called the *convolution* of H_t with φ and denoted $u = H_t * \varphi$.

In the next section, we will generalize this convolution formula to higher dimension, and check that it does yield a solution of the heat equation satisfying the desired initial condition.

6.3 Integral Solution Formula

Consider the heat equation on $\mathbb{R}^n$,

$$\frac{\partial u}{\partial t} - \Delta u = 0 \tag{6.15}$$

for $t > 0$, with initial condition

$$u(0,\boldsymbol{x}) = g(\boldsymbol{x})$$

for $\boldsymbol{x} \in \mathbb{R}^n$.

Inspired by the calculations in Sect. 6.2, we define

$$H_t(\boldsymbol{x}) := (4\pi t)^{-\frac{n}{2}} e^{|\boldsymbol{x}|^2/4t} \tag{6.16}$$

for $t > 0$. The normalizing factor $(4\pi t)^{-\frac{n}{2}}$ is chosen so that

$$\int_{\mathbb{R}^n} H_t(\boldsymbol{x})\, d^n\boldsymbol{x} = 1. \tag{6.17}$$

(See Exercise 2.5 for the computation of integrals of this type.)

Direct differentiation shows that for $t > 0$,

$$\frac{\partial}{\partial t} H_t - \Delta H_t = 0, \tag{6.18}$$

so that $H_t(\boldsymbol{x})$ is a solution of the heat equation. However, the limit of $H_t(\boldsymbol{x})$ as $t \to 0$ is 0 for $\boldsymbol{x} \neq 0$ and ∞ for $\boldsymbol{x} = 0$, which does not seem to make sense as a distribution of temperatures. (We will return to discuss the interpretation of this initial condition in Chap. 12.)

With (6.14) as motivation, our goal in this section is to show that the convolution $u = H_t * g$ satisfies the heat equation on $\mathbb{R}^n$ for a continuous and bounded initial condition g. A function that acts on other functions by convolution is an *integral kernel*, and H_t is specifically called the *heat kernel* on $\mathbb{R}^n$. It is also called the *fundamental solution* of the heat equation, for reasons we will explain in Chap. 12.

Theorem 6.2 *For a bounded function $g \in C^0(\mathbb{R}^n)$, the heat equation*

$$\left(\frac{\partial}{\partial t} - \Delta\right) u = 0, \qquad u|_{t=0} = g, \tag{6.19}$$

admits a classical solution given by

$$u(t, \boldsymbol{x}) = H_t * g(\boldsymbol{x}). \tag{6.20}$$

Proof Explicitly, the formula (6.20) says that

$$u(t, \boldsymbol{x}) = (4\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-|\boldsymbol{x}-\boldsymbol{y}|^2/4t} g(\boldsymbol{y})\, d^n\boldsymbol{y}. \tag{6.21}$$

Because the domain is infinite here, we should treat differentiation under the integral with some care. To justify this, the key point is that the partial derivatives of H_t can be estimated by expressions of the form $c_1(t, \boldsymbol{x})e^{-c_2(t,\boldsymbol{x})|\boldsymbol{y}|^2}$, where the dependence of the constants c_1 and c_2 on t and $\boldsymbol{x}$ is continuous. We will not go into the technical details, but this makes it relatively straightforward to check that differentiating under the integral works in this case. In particular, the fact that $H_t(\boldsymbol{x})$ solves the heat equation implies (6.19).

To check the initial condition, fix some $\boldsymbol{x} \in \mathbb{R}^n$. A change of variables to $\boldsymbol{w} = (\boldsymbol{y} - \boldsymbol{x})/\sqrt{t}$ in (6.21) gives

$$\begin{aligned} u(t, \boldsymbol{x}) &= (4\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-|\boldsymbol{w}|^2} g\left(\boldsymbol{x} + t^{\frac{1}{2}} \boldsymbol{w}\right) d^n \boldsymbol{w} \\ &= \int_{\mathbb{R}^n} H_1(\boldsymbol{w}) g\left(\boldsymbol{x} + t^{\frac{1}{2}} \boldsymbol{w}\right) d^n \boldsymbol{w} \end{aligned}$$

By (6.17) we can also write

$$g(\boldsymbol{x}) = \int_{\mathbb{R}^n} H_1(\boldsymbol{w}) g(\boldsymbol{x})\, d^n \boldsymbol{w}.$$

Thus the difference we are trying to estimate is

$$u(t, \boldsymbol{x}) - g(\boldsymbol{x}) = \int_{\mathbb{R}^n} H_1(\boldsymbol{w}) \left[g\left(\boldsymbol{x} + t^{\frac{1}{2}} \boldsymbol{w}\right) - g(\boldsymbol{x})\right] d^n \boldsymbol{w}. \tag{6.22}$$

Given $\varepsilon > 0$, we can use the exponential decay of $H_1(\boldsymbol{w})$ as $|\boldsymbol{w}| \to \infty$ to choose R so that

$$\int_{|\boldsymbol{w}| \geq R} H_1(\boldsymbol{w})\, d^n \boldsymbol{w} < \varepsilon.$$

Since g is bounded by assumption, there exists a constant M so that $|g| \leq M$. The "large-$\boldsymbol{w}$" piece of (6.22) can thus be estimated by

$$\int_{|\boldsymbol{w}| \geq R} H_1(\boldsymbol{w}) \left|g\left(\boldsymbol{x} + t^{\frac{1}{2}} \boldsymbol{w}\right) - g(\boldsymbol{x})\right| d^n \boldsymbol{w} \leq 2M\varepsilon. \tag{6.23}$$

On the other hand, by the continuity of g we can choose $\delta > 0$ so that for $\boldsymbol{y}$ such that for $|\boldsymbol{x} - \boldsymbol{y}| < \delta$, we have

$$|g(\boldsymbol{x} - \boldsymbol{y}) - g(\boldsymbol{x})| \leq \varepsilon.$$

Thus for $|\boldsymbol{w}| < R$ and $t < \delta^2/R^2$,

$$\left|g\left(\boldsymbol{x} + t^{\frac{1}{2}} \boldsymbol{w}\right) - g(\boldsymbol{x})\right| \leq \varepsilon.$$

It follows that for $t < \delta^2/R^2$,

$$\begin{aligned}\int_{|\boldsymbol{w}|<R} H_1(\boldsymbol{w}) \left| g\left(\boldsymbol{x} + t^{\frac{1}{2}}\boldsymbol{w}\right) - g(\boldsymbol{x})\right| \, d^n\boldsymbol{w} &\le \varepsilon \int_{|\boldsymbol{w}|<R} H_1(\boldsymbol{w})\, d^n\boldsymbol{w} \\ &\le \varepsilon \int_{\mathbb{R}^n} H_1(\boldsymbol{w})\, d^n\boldsymbol{w} \\ &= \varepsilon.\end{aligned} \tag{6.24}$$

Combining the estimates (6.23) and (6.24) gives

$$|u(t, \boldsymbol{x}) - g(\boldsymbol{x})| \le (2M + 1)\varepsilon$$

for $0 < t < \delta^2/R^2$. Since ε was arbitrary, this shows that

$$\lim_{t\to 0} u(t, \boldsymbol{x}) = g(\boldsymbol{x}). \qquad \square$$

Without extra restrictions on u, the solution of (6.19) is not necessarily unique. However, if we start from a bounded initial condition g, then it is physically reasonable to assume that u is bounded over finite time intervals.

Theorem 6.3 *Under the assumption that $u(t, \cdot)$ is bounded on $[0, T] \times \mathbb{R}^n$ for each $T > 0$, the solution of the heat equation (6.19) is unique.*

We will develop tools to prove this result (maximum principles) in Chap. 9. The statement can be improved by weakening the boundedness hypothesis to an assumption of exponential growth. The counterexamples to uniqueness exhibit superexponential growth and are not considered valid as physical solutions.

In combination, Theorems 6.2 and 6.3, show that a bounded solution of the heat equation on $\mathbb{R}^n$ with continuous initial data satisfies (6.21). The function $H_t(\boldsymbol{x})$ is C^∞ in both variables for $t > 0$. As we noted in the proof of Theorem 6.2, differentiation under the integral is justified in (6.21), so this regularity can be extended to general solutions.

Theorem 6.4 *Suppose that u is a bounded solution of the heat equation (6.19) for a bounded initial condition $g \in C^0(\mathbb{R}^n)$. Then $u \in C^\infty((0, \infty) \times \mathbb{R}^n)$.*

Similar regularity results hold for the heat equation in other contexts, for example on a bounded domain. We will discuss some of these cases later in Sect. 8.6. This behavior, i.e., smoothness of solutions that does not depend on the regularity of the initial data, is characteristic of parabolic equations.

Another interesting feature of the heat kernel is the fact that $H_t(\boldsymbol{x})$ is strictly positive for all $t > 0$ and $\boldsymbol{x} \in \mathbb{R}^n$. This means that if g is nonnegative and not identically zero, then u is nonzero at all points $\boldsymbol{x} \in \mathbb{R}^n$ for $t > 0$. Compare this to the Huygens principle that we observed in Chap. 4, which says that for solutions of the wave equation the range of influence of a point is limited by the (finite) propagation speed. The heat equation exhibits *infinite propagation speed*.

We can see the origin of the infinite propagation speed in Einstein's diffusion model from Sect. 6.1. In (6.10) the value D, which is assumed to be constant, is the average squared displacement per unit time. The fact that D is fixed implies that in the continuum limit $\tau \to 0$, the average absolute displacement per unit time diverges. Hence the infinite propagation speed is built in to the construction of the model. It reflects the fact that models of diffusion are inherently statistical, and not expected to be accurate on a microscopic scale.

6.4 Inhomogeneous Problem

Duhamel's method, which was used to construct solutions of the inhomogeneous wave equation in Sect. 4.4, was originally developed in the context of the heat equation. There are slight differences in the setup, but the basic idea of translating a forcing term into an initial condition applies in both settings.

Consider the equation on $\mathbb{R}^n$,

$$\frac{\partial u}{\partial t} - \Delta u = f \tag{6.25}$$

for $t > 0$, with initial condition $u(0, \boldsymbol{x}) = 0$. For $s \geq 0$, let $\eta_s(t, \boldsymbol{x})$ be the solution of the homogeneous heat equation (6.19) for $t \geq s$, subject to the initial condition

$$\eta_s(t, \boldsymbol{x})\big|_{t=s} = f(s, \boldsymbol{x}). \tag{6.26}$$

We claim that the solution is then given by the integral

$$u(t, \boldsymbol{x}) := \int_0^t \eta_s(t, \boldsymbol{x})\, ds.$$

Using the formula for η_s provided by Theorem 6.2, the proposed solution can be written

$$u(t, \boldsymbol{x}) = \int_0^t \int_{\mathbb{R}^n} H_{t-s}(\boldsymbol{x} - \boldsymbol{y}) f(s, \boldsymbol{y})\, d^n\boldsymbol{y}\, ds. \tag{6.27}$$

To justify this formula, we must investigate carefully what happens near the point $t = s$.

Theorem 6.5 *Assuming that $f \in C^2([0, \infty) \times \mathbb{R}^n)$ and is compactly supported, the formula (6.27) yields a classical solution to the inhomogeneous heat equation (6.25).*

Proof We can see that u is at least C^2 by changing variables in the integral formula to obtain

$$u(t, \boldsymbol{x}) = \int_0^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) f(t - s, \boldsymbol{x} - \boldsymbol{y})\, d^n\boldsymbol{y}\, ds.$$

Since $H_s(\boldsymbol{y})$ is smooth near $s = t$ and f is compactly supported, differentiation under the integral is justified. This gives

$$\frac{\partial u}{\partial t}(t,\boldsymbol{x}) = \int_0^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \frac{\partial f}{\partial t}(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds + \int_{\mathbb{R}^n} H_t(\boldsymbol{y}) \frac{\partial f}{\partial t}(0,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}, \tag{6.28}$$

and

$$\Delta u(t,\boldsymbol{x}) = \int_0^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \Delta_{\boldsymbol{x}} f(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds. \tag{6.29}$$

Our goal is to integrate by parts in these formulas, to exploit the fact that H_s solves the heat equation. Here we must be careful, because of the singular behavior of H_s at $s = 0$.

To deal with this singularity, we split the integral at $s = \varepsilon$. For the first integral in (6.28), switching the t derivative to an s derivative and integrating by parts gives

$$\begin{aligned}
&\int_\varepsilon^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \frac{\partial f}{\partial t}(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \\
&\quad = -\int_\varepsilon^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \frac{\partial f}{\partial s}(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \\
&\quad = \int_\varepsilon^t \int_{\mathbb{R}^n} \frac{\partial H_s}{\partial s}(\boldsymbol{y}) f(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \\
&\qquad - \int_{\mathbb{R}^n} H_t(\boldsymbol{y}) f(0,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y} + \int_{\mathbb{R}^n} H_\varepsilon(\boldsymbol{y}) f(t-\varepsilon,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}
\end{aligned}$$

The corresponding result for (6.29) has no boundary terms because of the compact support of f,

$$\begin{aligned}
&\int_\varepsilon^t \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \Delta_{\boldsymbol{x}} f(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \\
&\quad = \int_\varepsilon^t \int_{\mathbb{R}^n} \Delta_{\boldsymbol{y}} H_s(\boldsymbol{y}) f(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds
\end{aligned}$$

Applying these integrations by parts to (6.28) and (6.29), and using the fact that $(\frac{\partial}{\partial s} - \Delta)H_s = 0$, we obtain

$$\left(\frac{\partial}{\partial t} - \Delta\right) u(t,\boldsymbol{x}) = \int_{\mathbb{R}^n} H_\varepsilon(\boldsymbol{y}) f(t-\varepsilon,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y} + \int_0^\varepsilon \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \left(\frac{\partial}{\partial t} - \Delta_{\boldsymbol{x}}\right) f(t-s,\boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds. \tag{6.30}$$

Since $H_s > 0$, the second term can be estimated by

$$\left| \int_0^{\varepsilon} \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \left(\frac{\partial}{\partial t} - \Delta_{\boldsymbol{x}} \right) f(t-s, \boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \right| \\ \leq C \int_0^{\varepsilon} \int_{\mathbb{R}^n} H_s(\boldsymbol{y})\, d^n\boldsymbol{y}\, ds,$$

where C is the maximum value of $\left|(\frac{\partial}{\partial t} - \Delta) f\right|$ (which exists by the assumption of compact support). By (6.17), the integral of $H_s(\boldsymbol{y})$ over $\boldsymbol{y} \in \mathbb{R}^n$ evaluates to 1, yielding the estimate

$$\left| \int_0^{\varepsilon} \int_{\mathbb{R}^n} H_s(\boldsymbol{y}) \left(\frac{\partial}{\partial t} - \Delta_{\boldsymbol{x}} \right) f(t-s, \boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}\, ds \right| \leq C\varepsilon.$$

We can therefore take $\varepsilon \to 0$ in (6.30) to obtain

$$\left(\frac{\partial}{\partial t} - \Delta \right) u(t, \boldsymbol{x}) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} H_\varepsilon(\boldsymbol{y}) f(t-\varepsilon, \boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y}.$$

The remaining limit is very close to the limit computed in the proof of Theorem 6.2, except that t is replaced by $t - \varepsilon$. A simple modification of that argument shows that

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} H_\varepsilon(\boldsymbol{y}) f(t-\varepsilon, \boldsymbol{x}-\boldsymbol{y})\, d^n\boldsymbol{y} = f(t, \boldsymbol{x}).$$

This completes the proof that u satisfies (6.25). □

6.5 Exercises

6.1 Biological processes and chemical reactions are frequently described by *reaction-diffusion equations*, consisting of the heat equation modified by a reaction term. Consider the simplest such equation,

$$\frac{\partial u}{\partial t} + \gamma u - \Delta u = 0.$$

on $\mathbb{R}^n$ with initial condition $u(0, \boldsymbol{x}) = f(\boldsymbol{x})$. Assuming f is continuous and bounded, find a formula for the solution. Hint: use a substitution of the form $u \to e^{-at} u$ to reduce this to the ordinary heat equation.

6.2 For $t \geq 0$, $x \geq 0$, suppose that $u(t, x)$ satisfies the one-dimensional heat equation (6.11) with the initial condition $u(0, x) = 0$ for $x \geq 0$ and the boundary condition

$$u(t, 0) = A\cos(\omega t)$$

for $t \geq 0$. Under the additional requirement that $u(t, \cdot)$ is bounded, find a solution $u(t, x)$. Hint: use separation of variables and assume that the temporal components have the form $e^{\pm i\omega t}$.

6.3 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with piecewise C^1 boundary. Suppose that $u(t, \boldsymbol{x})$ satisfies the heat equation

$$\frac{\partial u}{\partial t} - \Delta u = 0,$$

on $(0, \infty) \times \Omega$. Following the discussion from Sect. 6.1, we define the total thermal energy at time t by

$$\mathcal{U}[t] = \int_\Omega u(t, \boldsymbol{x})\, d^n\boldsymbol{x}.$$

(a) Assume that u satisfies Neumann boundary conditions,

$$\left.\frac{\partial u}{\partial \nu}\right|_{\partial\Omega} = 0,$$

(the insulated case). Show that $\mathcal{U}$ is constant.

(b) Assume that u is positive in the interior of Ω and equals 0 on the boundary. Show that $\mathcal{U}(t)$ is decreasing in this case.

6.4 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with piecewise C^1 boundary. Suppose that $u(t, \boldsymbol{x})$ is real-valued and satisfies the heat equation

$$\frac{\partial u}{\partial t} - \Delta u = 0,$$

on $(0, \infty) \times \Omega$. Define

$$\eta(t) := \int_\Omega u(t, \boldsymbol{x})^2\, d^n\boldsymbol{x}. \tag{6.31}$$

(a) Assume that u satisfies the Dirichlet boundary conditions:

$$u(t, \boldsymbol{x})|_{\boldsymbol{x}\in\partial\Omega} = 0$$

for $t \geq 0$. Show that η decreases as a function of t.

(b) Use (a) to show that a solution u satisfying boundary and initial conditions

$$u|_{t=0} = g, \quad u|_{\boldsymbol{x}\in\partial\Omega} = h,$$

for some continuous functions g on Ω and h on $\partial\Omega$, is uniquely determined by g and h.