2 Heat Equation

2.1 Derivation

Ref: Strauss, Section 1.3.

Below we provide two derivations of the **heat equation**,

$$u_t - ku_{xx} = 0 \qquad k > 0. \tag{2.1}$$

This equation is also known as the **diffusion equation**.

2.1.1 Diffusion

Consider a liquid in which a dye is being diffused through the liquid. The dye will move from higher concentration to lower concentration. Let $u(x,t)$ be the concentration (mass per unit length) of the dye at position x in the pipe at time t. The total mass of dye in the pipe from x_0 to x_1 at time t is given by

$$M(t) = \int_{x_0}^{x_1} u(x,t)\,dx.$$

Therefore,

$$\frac{dM}{dt} = \int_{x_0}^{x_1} u_t(x,t)\,dx.$$

By Fick's Law,

$$\frac{dM}{dt} = \text{flow in} - \text{flow out} = ku_x(x_1,t) - ku_x(x_0,t),$$

where $k > 0$ is a proportionality constant. That is, the flow rate is proportional to the concentration gradient. Therefore,

$$\int_{x_0}^{x_1} u_t(x,t)\,dx = ku_x(x_1,t) - ku_x(x_0,t).$$

Now differentiating with respect to x_1, we have

$$u_t(x_1,t) = ku_{xx}(x_1,t).$$

Or,

$$u_t = ku_{xx}.$$

This is known as the *diffusion equation*.

2.1.2 Heat Flow

We now give an alternate derivation of (2.1) from the study of heat flow. Let D be a region in $\mathbb{R}^n$. Let $x = [x_1, \ldots, x_n]^T$ be a vector in $\mathbb{R}^n$. Let $u(x,t)$ be the temperature at point x, time t, and let $H(t)$ be the total amount of heat (in calories) contained in D. Let c be the specific heat of the material and ρ its density (mass per unit volume). Then

$$H(t) = \int_D c\rho u(x,t)\,dx.$$

Therefore, the change in heat is given by

$$\frac{dH}{dt} = \int_D c\rho u_t(x,t)\,dx.$$

Fourier's Law says that heat flows from hot to cold regions at a rate $\kappa > 0$ proportional to the temperature gradient. The only way heat will leave D is through the boundary. That is,

$$\frac{dH}{dt} = \int_{\partial D} \kappa \nabla u \cdot n\,dS.$$

where ∂D is the boundary of D, n is the outward unit normal vector to ∂D and dS is the surface measure over ∂D. Therefore, we have

$$\int_D c\rho u_t(x,t)\,dx = \int_{\partial D} \kappa \nabla u \cdot n\,dS.$$

Recall that for a vector field F, the Divergence Theorem says

$$\int_{\partial D} F \cdot n\,dS = \int_D \nabla \cdot F\,dx.$$

(*Ref: See Strauss, Appendix A.3.*) Therefore, we have

$$\int_D c\rho u_t(x,t)\,dx = \int_D \nabla \cdot (\kappa \nabla u)\,dx.$$

This leads us to the partial differential equation

$$c\rho u_t = \nabla \cdot (\kappa \nabla u).$$

If c, ρ and κ are constants, we are led to the heat equation

$$u_t = k\Delta u,$$

where $k = \kappa/c\rho > 0$ and $\Delta u = \sum_{i=1}^n u_{x_i x_i}$.

2.2 Heat Equation on an Interval in $\mathbb{R}$

2.2.1 Separation of Variables

Consider the initial/boundary value problem on an interval $\mathcal{I}$ in $\mathbb{R}$,

$$\begin{cases} u_t = k u_{xx} & x \in \mathcal{I}, t > 0 \\ u(x,0) = \phi(x) & x \in \mathcal{I} \\ u \text{ satisfies certain BCs.} & \end{cases} \tag{2.2}$$

In practice, the most common boundary conditions are the following:

1. Dirichlet ($\mathcal{I} = (0, l)$) : $u(0, t) = 0 = u(l, t)$.

2. Neumann ($\mathcal{I} = (0, l)$) : $u_x(0, t) = 0 = u_x(l, t)$.

3. Robin ($\mathcal{I} = (0, l)$) : $u_x(0, t) - a_0 u(0, t) = 0$ and $u_x(l, t) + a_l u(l, t) = 0$.

4. Periodic ($\mathcal{I} = (-l, l)$): $u(-l, t) = u(l, t)$ and $u_x(-l, t) = u_x(l, t)$.

We will give specific examples below where we consider some of these boundary conditions. First, however, we present the technique of *separation of variables*. This technique involves looking for a solution of a particular form. In particular, we look for a solution of the form

$$u(x, t) = X(x)T(t)$$

for functions X, T to be determined. Suppose we can find a solution of (2.2) of this form. Plugging a function $u = XT$ into the heat equation, we arrive at the equation

$$XT' - kX''T = 0.$$

Dividing this equation by kXT, we have

$$\frac{T'}{kT} = \frac{X''}{X} = -\lambda.$$

for some constant λ. Therefore, if there exists a solution $u(x, t) = X(x)T(t)$ of the heat equation, then T and X must satisfy the equations

$$\frac{T'}{kT} = -\lambda$$
$$\frac{X''}{X} = -\lambda$$

for some constant λ. In addition, in order for u to satisfy our boundary conditions, we need our function X to satisfy our boundary conditions. That is, we need to find functions X and scalars λ such that

$$\begin{cases} -X''(x) = \lambda X(x) & x \in \mathcal{I} \\ X \text{ satisfies our BCs.} \end{cases} \tag{2.3}$$

This problem is known as an **eigenvalue problem**. In particular, a constant λ which satisfies (2.3) for some function X, not identically zero, is called an **eigenvalue** of $-\partial_x^2$ for the given boundary conditions. The function X is called an **eigenfunction** with associated eigenvalue λ.

Therefore, in order to find a solution of (2.2) of the form $u(x, t) = X(x)T(t)$ our first goal is to find all solutions of our eigenvalue problem (2.3). Let's look at some examples below.

Example 1. (Dirichlet Boundary Conditions) Find all solutions to the eigenvalue problem

$$\begin{cases} -X'' = \lambda X & 0 < x < l \\ X(0) = 0 = X(l). \end{cases} \tag{2.4}$$

Any positive eigenvalues? First, we check if we have any positive eigenvalues. That is, we check if there exists any $\lambda = \beta^2 > 0$. Our eigenvalue problem (2.4) becomes

$$\begin{cases} X'' + \beta^2 X = 0 & 0 < x < l \\ X(0) = 0 = X(l). \end{cases}$$

The solutions of this ODE are given by

$$X(x) = C\cos(\beta x) + D\sin(\beta x).$$

The boundary condition

$$X(0) = 0 \implies C = 0.$$

The boundary condition

$$X(l) = 0 \implies \sin(\beta l) = 0 \implies \beta = \frac{n\pi}{l} \qquad n = 1, 2, \ldots.$$

Therefore, we have a sequence of positive eigenvalues

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2$$

with corresponding eigenfunctions

$$X_n(x) = D_n \sin\left(\frac{n\pi}{l}x\right).$$

Is zero an eigenvalue? Next, we look to see if zero is an eigenvalue. If zero is an eigenvalue, our eigenvalue problem (2.4) becomes

$$\begin{cases} X'' = 0 & 0 < x < l \\ X(0) = 0 = X(l). \end{cases}$$

The general solution of the ODE is given by

$$X(x) = C + Dx.$$

The boundary condition

$$X(0) = 0 \implies C = 0.$$

The boundary condition

$$X(l) = 0 \implies D = 0.$$

Therefore, the only solution of the eigenvalue problem for $\lambda = 0$ is $X(x) = 0$. By definition, the zero function is not an eigenfunction. Therefore, $\lambda = 0$ is not an eigenvalue.

Any negative eigenvalues? Last, we check for negative eigenvalues. That is, we look for an eigenvalue $\lambda = -\gamma^2$. In this case, our eigenvalue problem (2.4) becomes

$$\begin{cases} X'' - \gamma^2 X = 0 & 0 < x < l \\ X(0) = 0 = X(l). \end{cases}$$

The solutions of this ODE are given by

$$X(x) = C\cosh(\gamma x) + D\sinh(\gamma x).$$

The boundary condition

$$X(0) = 0 \implies C = 0.$$

The boundary condition

$$X(l) = 0 \implies D = 0.$$

Therefore, there are no negative eigenvalues.

Consequently, *all* the solutions of (2.4) are given by

$$\boxed{\lambda_n = \left(\frac{n\pi}{l}\right)^2 \qquad X_n(x) = D_n \sin\left(\frac{n\pi}{l}x\right) \qquad n = 1, 2, \ldots.}$$

◇

Example 2. (Periodic Boundary Conditions) Find all solutions to the eigenvalue problem

$$\begin{cases} -X'' = \lambda X & -l < x < l \\ X(-l) = X(l),\ X'(-l) = X'(l). \end{cases} \tag{2.5}$$

Any positive eigenvalues? First, we check if we have any positive eigenvalues. That is, we check if there exists any $\lambda = \beta^2 > 0$. Our eigenvalue problem (2.5) becomes

$$\begin{cases} X'' + \beta^2 X = 0 & -l < x < l \\ X(-l) = X(l),\ X'(-l) = X'(l). \end{cases}$$

The solutions of this ODE are given by

$$X(x) = C\cos(\beta x) + D\sin(\beta x).$$

The boundary condition

$$X(-l) = X(l) \implies D\sin(\beta l) = 0 \implies D = 0 \text{ or } \beta = \frac{n\pi}{l}.$$

The boundary condition

$$X'(-l) = X'(l) \implies C\beta\sin(\beta l) = 0 \implies C = 0 \text{ or } \beta = \frac{n\pi}{l}.$$

Therefore, we have a sequence of positive eigenvalues

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2$$

with corresponding eigenfunctions

$$X_n(x) = C_n \cos\left(\frac{n\pi}{l}x\right) + D_n \sin\left(\frac{n\pi}{l}x\right).$$

Is zero an eigenvalue? Next, we look to see if zero is an eigenvalue. If zero is an eigenvalue, our eigenvalue problem (2.5) becomes

$$\begin{cases} X'' = 0 & -l < x < l \\ X(-l) = X(l), \; X'(-l) = X'(l). \end{cases}$$

The general solution of the ODE is given by

$$X(x) = C + Dx.$$

The boundary condition

$$X(-l) = X(l) \implies D = 0.$$

The boundary condition

$$X'(-l) = X'(l) \text{ is automatically satisfied if } D = 0.$$

Therefore, $\lambda = 0$ is an eigenvalue with corresponding eigenfunction

$$X_0(x) = C_0.$$

Any negative eigenvalues? Last, we check for negative eigenvalues. That is, we look for an eigenvalue $\lambda = -\gamma^2$. In this case, our eigenvalue problem (2.5) becomes

$$\begin{cases} X'' - \gamma^2 X = 0 & -l < x < l \\ X(-l) = X(l), \; X'(-l) = X'(l). \end{cases}$$

The solutions of this ODE are given by

$$X(x) = C\cosh(\gamma x) + D\sinh(\gamma x).$$

The boundary condition

$$X(-l) = X(l) \implies D\sinh(\gamma l) = 0 \implies D = 0.$$

The boundary condition

$$X'(-l) = X'(l) \implies C\gamma\sinh(\gamma l) = 0 \implies C = 0.$$

Therefore, there are no negative eigenvalues.

Consequently, *all* the solutions of (2.5) are given by

$$\boxed{\begin{array}{lll} \lambda_n = \left(\dfrac{n\pi}{l}\right)^2 & X_n(x) = C_n \cos\left(\dfrac{n\pi}{l}x\right) + D_n \sin\left(\dfrac{n\pi}{l}x\right) & n = 1, 2, \ldots \\ \lambda_0 = 0 & X_0(x) = C_0. & \end{array}}$$

$\diamond$

Now that we have done a couple of examples of solving eigenvalue problems, we return to using the method of separation of variables to solve (2.2). Recall that in order for a function of the form $u(x,t) = X(x)T(t)$ to be a solution of the heat equation on an interval $\mathcal{I} \subset \mathbb{R}$ which satisfies given boundary conditions, we need X to be a solution of the eigenvalue problem,

$$\begin{cases} X'' = -\lambda X & x \in \mathcal{I} \\ X \text{ satisfies certain BCs} \end{cases}$$

for some scalar λ and T to be a solution of the ODE

$$-T' = k\lambda T.$$

We have given some examples above of how to solve the eigenvalue problem. Once we have solved the eigenvalue problem, we need to solve our equation for T. In particular, for any scalar λ, the solution of the ODE for T is given by

$$T(t) = Ae^{-k\lambda t}$$

for an arbitrary constant A. Therefore, for each eigenfunction X_n with corresponding eigenvalue λ_n, we have a solution T_n such that the function

$$u_n(x,t) = T_n(t)X_n(x)$$

is a solution of the heat equation on the interval $\mathcal{I}$ which satisfies our boundary conditions. Note that we have not yet accounted for our initial condition $u(x,0) = \phi(x)$. We will look at that next. First, we remark that if $\{u_n\}$ is a sequence of solutions of the heat equation on $\mathcal{I}$ which satisfy our boundary conditions, than any *finite* linear combination of these solutions will also give us a solution. That is,

$$u(x,t) \equiv \sum_{n=1}^{N} u_n(x,t)$$

will be a solution of the heat equation on $\mathcal{I}$ which satisfies our boundary conditions, assuming each u_n is such a solution. In fact, one can show that an *infinite* series of the form

$$u(x,t) \equiv \sum_{n=1}^{\infty} u_n(x,t)$$

will also be a solution of the heat equation, under proper convergence assumptions of this series. We will omit discussion of this issue here.

2.2.2 Satisfying our Initial Conditions

We return to trying to satisfy our initial conditions. Assume we have found all solutions of our eigenvalue problem. We let $\{X_n\}$ denote our sequence of eigenfunctions and $\{\lambda_n\}$ denote our sequence of eigenvalues. Then for each λ_n, we have a solution T_n of our equation for T. Let

$$u(x,t) = \sum_n X_n(x)T_n(t) = \sum_n A_n X_n(x)e^{-k\lambda_n t}.$$

Our goal is to choose A_n appropriately such that our initial condition is satisfied. In particular, we need to choose A_n such that

$$u(x,0) = \sum_n A_n X_n(x) = \phi(x).$$

In order to find A_n satisfying this condition, we use the following *orthogonality* property of eigenfunctions.

First, we make some definitions. For two real-valued functions f and g defined on Ω,

$$\langle f, g \rangle = \int_\Omega f(x) g(x)\, dx$$

is defined as the $\mathbf{L^2}$ **inner product** of f and g on Ω. The $\mathbf{L^2}$ **norm** of f on Ω is defined as

$$||f||^2_{L^2(\Omega)} = \langle f, f \rangle = \int_\Omega |f(x)|^2\, dx.$$

We say functions f and g are **orthogonal** on $\Omega \subset \mathbb{R}^n$ if

$$\langle f, g \rangle = \int_\Omega f(x) g(x)\, dx = 0.$$

We say boundary conditions are **symmetric** if

$$[f'(x)g(x) - f(x)g'(x)]|_{x=a}^{x=b} = 0$$

for all functions f and g satisfying the boundary conditions.

Lemma 3. *Consider the eigenvalue problem* (2.3) *with symmetric boundary conditions. If X_n, X_m are two eigenfunctions of* (2.3) *with distinct eigenvalues, then X_n and X_m are orthogonal.*

Proof. Let $\mathcal{I} = [a,b]$.

$$\begin{aligned}
\lambda_n \int_a^b X_n(x) X_m(x)\, dx &= -\int_a^b X_n''(x) X_m(x)\, dx \\
&= \int_a^b X_n'(x) X_m'(x)\, dx - X_n'(x) X_m(x)|_{x=a}^{x=b} \\
&= -\int_a^b X_n(x) X_m''(x)\, dx + [X_n(x) X_m'(x) - X_n'(x) X_m(x)]|_{x=a}^{x=b} \\
&= -\lambda_m \int_a^b X_n(x) X_m(x)\, dx,
\end{aligned}$$

using the fact that the boundary conditions are symmetric. Therefore,

$$(\lambda_n - \lambda_m) \int_a^b X_n(x) X_m(x)\, dx = 0,$$

but $\lambda_n \neq \lambda_m$ because the eigenvalues are assumed to be distinct. Therefore,

$$\int_a^b X_n(x)X_m(x)\,dx = 0,$$

as claimed.

$\square$

We can use this lemma to find coefficients A_n such that

$$\sum_n A_n X_n(x) = \phi(x).$$

In particular, multiplying both sides of this equation by X_m for a fixed m and integrating over $\mathcal{I}$, we have

$$A_m \langle X_m, X_m \rangle = \langle X_m, \phi \rangle,$$

which implies

$$\boxed{A_m = \frac{\langle X_m, \phi \rangle}{\langle X_m, X_m \rangle}.}$$

Example 4. (Dirichlet Boundary Conditions) In the case of Dirichlet boundary conditions on the interval $[0, l]$, we showed earlier that our eigenvalues and eigenfunctions are given by

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2, \quad X_n(x) = \sin\left(\frac{n\pi}{l}x\right) \qquad n = 1, 2, \ldots.$$

Our solutions for T_n are given by

$$T_n(t) = A_n e^{-k\lambda_n t} = A_n e^{-k(n\pi/l)^2 t}.$$

Now let

$$u(x,t) = \sum_{n=1}^{\infty} X_n(x)T_n(t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right) e^{-k(n\pi/l)^2 t}.$$

Using the fact that Dirichlet boundary conditions are symmetric (check this!), our coefficients A_m are given by

$$A_m = \frac{\langle X_m, \phi \rangle}{\langle X_m, X_m \rangle} = \frac{\int_0^l \sin(m\pi x/l)\,\phi(x)\,dx}{\int_0^l \sin^2(m\pi x/l)\,dx} = \frac{2}{l}\int_0^l \sin\left(\frac{m\pi}{l}x\right)\phi(x)\,dx.$$

Therefore, the solution of (2.2) on the interval $\mathcal{I} = [0, l]$ with Dirichlet boundary conditions is given by

$$\boxed{u(x,t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right) e^{-k(n\pi/l)^2 t}}$$

where

$$\boxed{A_n = \frac{2}{l}\int_0^l \sin\left(\frac{n\pi}{l}x\right)\phi(x)\,dx.}$$

$\diamond$

Example 5. (Periodic Boundary Conditions) In the case of periodic boundary conditions on the interval $[-l, l]$, we showed earlier that our eigenvalues and eigenfunctions are given by

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2, \quad X_n(x) = \begin{cases} \cos\left(\frac{n\pi}{l}x\right) \\ \sin\left(\frac{n\pi}{l}x\right) \end{cases} \qquad n = 1, 2, \ldots$$

$$\lambda_0 = 0, \quad X_0(x) = C_0.$$

Therefore, our solutions for T_n are given by

$$T_n(t) = A_n e^{-k\lambda_n t} = \begin{cases} A_n e^{-k(n\pi/l)^2 t} & n = 1, 2, \ldots \\ A_0 & n = 0. \end{cases}$$

Now using the fact that for any integer $n \geq 0$, $u_n(x,t) = X_n(x)T_n(t)$ is a solution of the heat equation which satisfies our periodic boundary conditions, we define

$$u(x,t) = \sum_n X_n(x)T_n(t) = A_0 + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi}{l}x\right) + B_n \sin\left(\frac{n\pi}{l}x\right)\right] e^{-k(n\pi/l)^2 t}.$$

Now, taking into account our initial condition, we want

$$u(x,0) = A_0 + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi}{l}x\right) + B_n \sin\left(\frac{n\pi}{l}x\right)\right] = \phi(x).$$

It remains only to find coefficients satisfying this equation. From our earlier discussion, using the fact that periodic boundary conditions are symmetric (check this!), we know that eigenfunctions corresponding to distinct eigenvalues will be orthogonal. For periodic boundary conditions, however, we have two eigenfunctions for each positive eigenvalue. Specifically, $\cos\left(\frac{n\pi}{l}x\right)$ and $\sin\left(\frac{n\pi}{l}x\right)$ are both eigenfunctions corresponding to the eigenvalue $\lambda_n = (n\pi/l)^2$. Could we be so lucky that these eigenfunctions would also be orthogonal? By a straightforward calculation, one can show that

$$\int_{-l}^{l} \cos\left(\frac{n\pi}{l}x\right) \sin\left(\frac{n\pi}{l}x\right) \, dx = 0.$$

They are orthogonal! This is not merely coincidental. In fact, for any eigenvalue λ of (2.3) with multiplicty m (meaning it has m linearly independent eigenfunctions), the eigenfunctions may always be chosen to be orthogonal. This process is known as the *Gram-Schmidt orthogonalization method.*

The fact that all our eigenfunctions are mutually orthogonal will allow us to calculate coefficients A_n, B_n so that our initial condition is satisfied. Using the technique described above, and letting $\langle f, g\rangle$ denote the L^2 inner product on $[-l, l]$, we see that

$$\begin{aligned}
A_0 &= \frac{\langle 1, \phi\rangle}{\langle 1, 1\rangle} = \frac{1}{2l}\int_{-l}^{l} \phi(x)\, dx \\
A_n &= \frac{\langle \cos(n\pi x/l), \phi\rangle}{\langle \cos(n\pi x/l), \cos(n\pi x/l)\rangle} = \frac{1}{l}\int_{-l}^{l} \cos\left(\frac{n\pi}{l}x\right)\phi(x)\, dx \\
B_n &= \frac{\langle \sin(n\pi x/l), \phi\rangle}{\langle \sin(n\pi x/l), \sin(n\pi x/l)\rangle} = \frac{1}{l}\int_{-l}^{l} \sin\left(\frac{n\pi}{l}x\right)\phi(x)\, dx.
\end{aligned}$$

Therefore, the solution of (2.2) on the interval $\mathcal{I} = [-l, l]$ with periodic boundary conditions is given by

$$\boxed{u(x,t) = A_0 + \sum_{n=1}^{\infty}\left[A_n \cos\left(\frac{n\pi}{l}x\right) + B_n \sin\left(\frac{n\pi}{l}x\right)\right] e^{-k(n\pi/l)^2 t}}$$

where

$$\boxed{\begin{aligned}
A_0 &= \frac{1}{2l}\int_{-l}^{l} \phi(x)\, dx \\
A_n &= \frac{1}{l}\int_{-l}^{l} \cos\left(\frac{n\pi}{l}x\right)\phi(x)\, dx \\
B_n &= \frac{1}{l}\int_{-l}^{l} \sin\left(\frac{n\pi}{l}x\right)\phi(x)\, dx.
\end{aligned}}$$

$\diamond$

2.2.3 Fourier Series

In the case of Dirichlet boundary conditions above, we looked for coefficients so that

$$\phi(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right).$$

We showed that *if* we could write our function ϕ in terms of this infinite series, our coefficients would be given by the formula

$$A_n = \frac{2}{l}\int_{0}^{l} \sin\left(\frac{n\pi}{l}x\right)\phi(x)\, dx.$$

For a given function ϕ defined on $(0, l)$ the infinite series

$$\phi \sim \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right) \text{ where } A_n \equiv \frac{2}{l}\int_{0}^{l} \sin\left(\frac{n\pi}{l}x\right)\phi(x)\, dx$$

is called the **Fourier sine series** of ϕ. *Note: The notation '$\sim$' just means the series associated with ϕ. It doesn't imply that the series necessarily converges to ϕ.*

In the case of periodic boundary conditions above, we looked for coefficients so that

$$\phi(x) = A_0 + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi}{l}x\right) + B_n \sin\left(\frac{n\pi}{l}x\right)\right]. \tag{2.6}$$

We showed that in this case our coefficients must be given by

$$\begin{aligned} A_0 &= \frac{1}{2l}\int_{-l}^{l} \phi(x)\,dx \\ A_n &= \frac{1}{l}\int_{-l}^{l} \cos\left(\frac{n\pi}{l}x\right)\phi(x)\,dx \\ B_n &= \frac{1}{l}\int_{-l}^{l} \sin\left(\frac{n\pi}{l}x\right)\phi(x)\,dx. \end{aligned} \tag{2.7}$$

For a given function ϕ defined on $(-l, l)$ the series

$$\phi \sim A_0 + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi}{l}x\right) + B_n \sin\left(\frac{n\pi}{l}x\right)\right]$$

where A_n, B_n are defined in (2.7) is called the **full Fourier series** of ϕ.

More generally, for a sequence of eigenfunctions $\{X_n\}$ of (2.3) which satisfy certain boundary conditions, we define the **general Fourier series** of a function ϕ as

$$\phi \sim \sum_n A_n X_n(x) \text{ where } A_n \equiv \frac{\langle X_n, \phi\rangle}{\langle X_n, X_n\rangle}.$$

Remark. Consider the example above where we looked to solve the heat equation on an interval with Dirichlet boundary conditions. (A similar remark holds for the case of periodic or other boundary conditions.) In order that our initial condition be satisfied, we needed to find coefficients A_n such that

$$\phi(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right).$$

We showed that *if* ϕ could be represented in terms of this infinite series, then our coefficients must be given by

$$A_n = \frac{2}{l}\int_0^l \sin\left(\frac{n\pi}{l}x\right)\phi(x)\,dx.$$

While this is the necessary form of our coefficients and defining

$$u(x,t) \equiv \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right) e^{-k(n\pi/l)^2 t}$$

for A_n defined above is the appropriate *formal* definition of our solution, in order to verify that this construction actually satisfies our initial/boundary value problem (2.2) with Dirichlet boundary conditions, we would need to verify the following.

1. The Fourier sine series of ϕ converges to ϕ (in some sense).

2. The *infinite* series actually satisfies the heat equation.

We will not discuss these issues here, but refer the reader to convergence results in Strauss as well as the notes from 220a. Later, in the course, we will prove an L^2 convergence result of eigenfunctions.

Complex Form of Full Fourier Series. It is sometimes useful to write the full Fourier series in complex form. We do so as follows. The eigenfunctions associated with the full Fourier series are given by

$$\left\{\cos\left(\frac{n\pi}{l}x\right), \sin\left(\frac{n\pi}{l}x\right)\right\}$$

for $n = 0, 1, 2, \ldots$. Using deMoivre's formula,

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

we can write

$$\begin{aligned}
\cos\left(\frac{n\pi}{l}x\right) &= \frac{e^{in\pi x/l} + e^{-in\pi x/l}}{2} \\
\sin\left(\frac{n\pi}{l}x\right) &= \frac{e^{in\pi x/l} - e^{-in\pi x/l}}{2i}.
\end{aligned}$$

Now, of course, any linear combination of eigenfunctions is also an eigenfunction. Therefore, we see that

$$\begin{aligned}
\cos\left(\frac{n\pi}{l}x\right) + i\sin\left(\frac{n\pi}{l}x\right) &= e^{in\pi x/l} \\
\cos\left(\frac{n\pi}{l}x\right) - i\sin\left(\frac{n\pi}{l}x\right) &= e^{-in\pi x/l}
\end{aligned}$$

are also eigenfunctions. Therefore, the eigenfunctions associated with the full Fourier series can be written as

$$\left\{e^{in\pi x/l}\right\} \qquad n = \ldots, -2, -1, 0, 1, 2, \ldots.$$

Now, let's suppose we can represent a given function ϕ as an infinite series expansion in terms of these eigenfunctions. That is, we want to find coefficients C_n such that

$$\phi(x) = \sum_{n=-\infty}^{\infty} C_n e^{in\pi x/l}.$$

As described earlier, eigenfunctions corresponding to distinct eigenvalues will be orthogonal, as periodic boundary conditions are symmetric. Therefore, we have

$$\int_{-l}^{l} e^{in\pi x/l} e^{im\pi x/l}\, dx = 0 \qquad \text{for } m \neq n.$$

For eigenfunctions corresponding to the same eigenvalue, we need to check the L^2 inner product. In particular, for the eigenvalue $\lambda_n = (n\pi/l)^2$, we have two eigenfunctions: $e^{in\pi x/l}$ and $e^{-in\pi x/l}$. By a straightforward calculation, we see that

$$\int_{-l}^{l} e^{in\pi x/l} e^{in\pi x/l}\, dx = 0 \qquad \text{for } n \neq 0$$
$$\int_{-l}^{l} e^{in\pi x/l} e^{-in\pi x/l}\, dx = 2l.$$

Therefore, our coefficients C_n would need to be given by

$$C_n = \frac{1}{2l}\int_{-l}^{l} \phi(x) e^{-in\pi x/l}\, dx.$$

Consequently, the complex form of the full Fourier series for a function ϕ defined on $(-l,l)$ is given by

$$\boxed{\phi \sim \sum_{n=-\infty}^{\infty} C_n e^{in\pi x/l} \text{ where } C_n = \frac{1}{2l}\int_{-l}^{l} \phi(x) e^{-in\pi x/l}\, dx.}$$

2.3 Fourier Transforms

2.3.1 Motivation

Ref: Strauss, Section 12.3

We would now like to turn to studying the heat equation on the whole real line. Consider the initial-value problem,

$$\begin{cases} u_t = k u_{xx}, & -\infty < x < \infty \\ u(x,0) = \phi(x). \end{cases} \tag{2.8}$$

In the case of the heat equation on an interval, we found a solution u using Fourier series. For the case of the heat equation on the whole real line, the Fourier series will be replaced by the Fourier transform.

Above, we discussed the complex form of the full Fourier series for a given function ϕ. In particular, for a function ϕ defined on the interval $[-l,l]$ we define its full Fourier series as

$$\phi \sim \sum_{n=-\infty}^{\infty} C_n e^{in\pi x/l} \text{ where } C_n = \frac{1}{2l}\int_{-l}^{l} \phi(x) e^{-in\pi x/l}\, dx.$$

Plugging the coefficients C_n into the infinite series, we see that

$$\phi \sim \sum_{n=-\infty}^{\infty} \left[\frac{1}{2l}\int_{-l}^{l} \phi(y) e^{-in\pi y/l}\, dy\right] e^{in\pi x/l}.$$

Now, letting $k = n\pi/l$, we can write this as

$$\phi \sim \sum_{n=-\infty}^{\infty} \left[\frac{1}{2\pi}\int_{-l}^{l} \phi(y) e^{-i(y-x)k}\, dy\right] \frac{\pi}{l}.$$

The distance between points k is given by $\Delta k = \pi/l$. As $l \to +\infty$, we can think of $\Delta k \to dk$ and the infinite sum becoming an integral. Roughly, we have

$$\phi \sim \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(y) e^{-i(y-x)k} \, dy \, dk.$$

Below we will use this motivation to define the Fourier transform of a function ϕ and then show how the Fourier transform can be used to solve the heat equation (among others) on the whole real line, and more generally in $\mathbb{R}^n$.

2.3.2 Definitions and Properties of the Fourier Transform

We say $f \in L^1(\mathbb{R}^n)$ if

$$\int_{\mathbb{R}^n} |f(x)| \, dx < +\infty.$$

For $f \in L^1(\mathbb{R}^n)$, we define its **Fourier transform** at a point $\xi \in \mathbb{R}^n$ as

$$\widehat{f}(\xi) \equiv \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \, dx.$$

We define its **inverse Fourier transform** at the point $\xi \in \mathbb{R}^n$ as

$$\check{f}(\xi) \equiv \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(x) \, dx.$$

Remark. Sometimes the constants in front are defined differently. I.e. - in some books, the Fourier transform is defined with a constant of $\frac{1}{(2\pi)^n}$ instead of $\frac{1}{(2\pi)^{n/2}}$, and in accordance the inverse Fourier transform is defined with a constant 1 replacing the constant $\frac{1}{(2\pi)^{n/2}}$ above.

Theorem 6. (Plancherel's Theorem) *If $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\widehat{u}, \check{u} \in L^2(\mathbb{R}^n)$ and*

$$||\widehat{u}||_{L^2(\mathbb{R}^n)} = ||\check{u}||_{L^2(\mathbb{R}^n)} = ||u||_{L^2(\mathbb{R}^n)}.$$

In order to prove this theorem, we need to prove some preliminary facts.

Claim 7. *Let*

$$f(x) = e^{-\epsilon|x|^2}.$$

Then

$$\widehat{f}(\xi) = \frac{1}{(2\epsilon)^{n/2}} e^{-|\xi|^2/4\epsilon}.$$

Proof.

$$\begin{aligned}
\widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} e^{-\epsilon|x|^2} \, dx \\
&= \frac{1}{(2\pi)^{n/2}} \left(\int_{-\infty}^{\infty} e^{-ix_1\xi_1} e^{-\epsilon x_1^2} \, dx_1 \right) \cdots \left(\int_{-\infty}^{\infty} e^{-ix_n\xi_n} e^{-\epsilon x_n^2} \, dx_n \right).
\end{aligned}$$

Therefore, we just need to look at

$$\int_{-\infty}^{\infty} e^{-ix\xi} e^{-\epsilon x^2}\, dx.$$

Completing the square, we have

$$\begin{aligned}
-\epsilon x^2 - ix\xi &= -\epsilon \left[x^2 + \frac{i\xi x}{\epsilon} + \left(\frac{i\xi}{2\epsilon}\right)^2 - \left(\frac{i\xi}{2\epsilon}\right)^2 \right] \\
&= -\epsilon \left[x + \left(\frac{i\xi}{2\epsilon}\right) \right]^2 + \epsilon \left(\frac{i\xi}{2\epsilon}\right)^2.
\end{aligned}$$

Therefore, we have

$$\int_{-\infty}^{\infty} e^{-ix\xi} e^{-\epsilon x^2}\, dx = \int_{-\infty}^{\infty} e^{-\epsilon[x+(i\xi/2\epsilon)]^2} e^{-\xi^2/4\epsilon}\, dx.$$

Now making the change of variables, $z = [x + (i\xi/2\epsilon)]$, we have

$$\int_{-\infty}^{\infty} e^{-\epsilon[x+(i\xi/2\epsilon)]^2} e^{-\xi^2/4\epsilon}\, dx = e^{-\xi^2/4\epsilon} \int_{\Gamma} e^{-\epsilon z^2}\, dz$$

where Γ is the line in the complex plane given by

$$\Gamma \equiv \left\{ z \in \mathbb{C} : y = x + \frac{i\xi}{2\epsilon}, x \in \mathbb{R} \right\}.$$

Without loss of generality, we assume $\xi > 0$. A similar analysis works if $\xi < 0$. Now

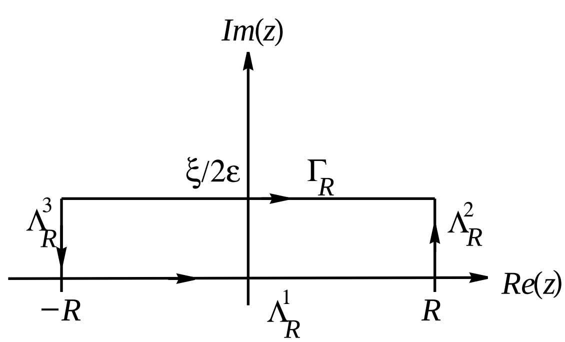
$$\int_{\Gamma} e^{-\epsilon z^2}\, dz = \lim_{R\to+\infty} \int_{\Gamma_R} e^{-\epsilon z^2}\, dz$$

where Γ_R is the line segment in the complex plane given by

$$\Gamma_R \equiv \left\{ z \in \mathbb{C} : z = x + \frac{i\xi}{2\epsilon}, |x| \le R \right\}.$$

Now define Λ_R^1, Λ_R^2 and Λ_R^3 as shown in the picture below. That is,

$$\begin{aligned}
\Lambda_R^1 &\equiv \{x \in \mathbb{R} : |x| \le R\} \\
\Lambda_R^2 &\equiv \left\{ z \in \mathbb{C} : z = x + iy, x, y \in \mathbb{R}, x = R, 0 \le y \le \frac{\xi}{2\epsilon} \right\} \\
\Lambda_R^3 &\equiv \left\{ z \in \mathbb{C} : z = x + iy; x, y \in \mathbb{R}; x = -R, 0 \le y \le \frac{\xi}{2\epsilon} \right\}.
\end{aligned}$$



From complex analysis, we know that

$$\int_C e^{-\epsilon z^2}\, dz = 0$$

where C is the closed curve given by $C = \Gamma_R \cup \Lambda_R^1 \cup \Lambda_R^2 \cup \Lambda_R^3$ traversed in the counter-clockwise direction. Therefore, we have

$$\int_{\Gamma_R} e^{-\epsilon z^2}\, dz = \int_{\Lambda_R} e^{-\epsilon z^2}\, dz$$

where the integral on the right-hand side is the line integral given by $\Lambda_R = \Lambda_R^3 \cup \Lambda_R^1 \cup \Lambda_R^2$ traversed in the direction shown. Therefore,

$$\int_{\Gamma} e^{-\epsilon z^2}\, dz = \lim_{R \to +\infty} \int_{\Lambda_R} e^{-\epsilon z^2}\, dz.$$

But, as $R \to +\infty$,

$$\begin{aligned} &\int_{\Lambda_R^j} e^{-\epsilon z^2}\, dz \to 0 \qquad \text{for} j = 2, 3 \\ &\int_{\Lambda_R^1} e^{-\epsilon z^2}\, dz \to \int_{-\infty}^{\infty} e^{-\epsilon x^2}\, dx. \end{aligned}$$

Therefore,

$$\int_{\Gamma} e^{-\epsilon z^2}\, dz = \int_{-\infty}^{\infty} e^{-\epsilon x^2}\, dx.$$

Consequently, we have

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ix\xi} e^{-\epsilon x^2}\, dx &= e^{-\xi^2/4\epsilon} \int_{\Gamma} e^{-\epsilon z^2}\, dz \\ &= e^{-\xi^2/4\epsilon} \int_{-\infty}^{\infty} e^{-\epsilon x^2}\, dx \\ &= e^{-\xi^2/4\epsilon} \int_{-\infty}^{\infty} e^{-\widetilde{x}^2} \frac{d\widetilde{x}}{\sqrt{\epsilon}} \\ &= \frac{e^{-\xi^2/4\epsilon}}{\sqrt{\epsilon}} \sqrt{\pi}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}} \left(\frac{e^{-\xi_1^2/4\epsilon}}{\sqrt{\epsilon}} \sqrt{\pi} \right) \cdots \left(\frac{e^{-\xi_n^2/4\epsilon}}{\sqrt{\epsilon}} \sqrt{\pi} \right) \\ &= \frac{1}{(2\epsilon)^{n/2}} e^{-|\xi|^2/4\epsilon}, \end{aligned}$$

as claimed. $\square$

Claim 8. *Let*

$$w(x) = u * v(x) = \int_{\mathbb{R}^n} u(x-y)v(y)\,dy \in L^1(\mathbb{R}^n).$$

That is, let w be the convolution of u and v. Then

$$\widehat{w}(\xi) = \widehat{u * v}(\xi) = (2\pi)^{n/2}\widehat{u}(\xi)\widehat{v}(\xi).$$

Proof. By definition,

$$\begin{aligned}
\widehat{w}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} w(x)\,dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \int_{\mathbb{R}^n} u(x-y)v(y)\,dy\,dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} e^{-i(x-y)\cdot\xi} u(x-y)\,dx\right] e^{-iy\cdot\xi} v(y)\,dy \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} (2\pi)^{n/2}\widehat{u}(\xi) e^{-iy\cdot\xi} v(y)\,dy \\
&= \widehat{u}(\xi) \int_{\mathbb{R}^n} e^{-iy\cdot\xi} v(y)\,dy \\
&= \widehat{u}(\xi)(2\pi)^{n/2}\widehat{v}(\xi) \\
&= (2\pi)^{n/2}\widehat{u}(\xi)\widehat{v}(\xi).
\end{aligned}$$

□

Now we use these two claims to prove Plancherel's Theorem.

Proof of Theorem 6. By assumption, $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Let

$$\begin{aligned}
v(x) &\equiv \overline{u}(-x) \\
w(x) &\equiv u * v(x).
\end{aligned}$$

Therefore, $v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $w \in L^1(\mathbb{R}^n) \cap L(\mathbb{R}^n)$.

First, we have

$$\begin{aligned}
\widehat{v}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} v(x)\,dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \overline{u}(-x)\,dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{iy\cdot\xi} \overline{u}(y)\,dy \\
&= \frac{1}{(2\pi)^{n/2}} \overline{\int_{\mathbb{R}^n} e^{-iy\cdot\xi} u(y)\,dy} \\
&= \overline{\widehat{u}}(\xi).
\end{aligned}$$

Therefore, using Claim 8,

$$\widehat{w}(\xi) = (2\pi)^{n/2}\widehat{u}(\xi)\widehat{v}(\xi) = (2\pi)^{n/2}\widehat{u}(\xi)\overline{\widehat{u}}(\xi) = (2\pi)^{n/2}|\widehat{u}|^2.$$

Next, we use the fact that if f, g are in $L^1(\mathbb{R}^n)$, then $\widehat{f}, \widehat{g}$ are in $L^\infty(\mathbb{R}^n)$, and, moreover,

$$\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)g(\xi)\,dx. \tag{2.9}$$

This fact can be seen by direct substitution, as shown below,

$$\begin{aligned}
\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx &= \int_{\mathbb{R}^n} f(x)\left[\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}g(\xi)\,d\xi\right]\,dx \\
&= \int_{\mathbb{R}^n}\left[\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}f(x)\,dx\right]g(\xi)\,d\xi \\
&= \int_{\mathbb{R}^n} \widehat{f}(\xi)g(\xi)\,d\xi.
\end{aligned}$$

Therefore, letting $f(x) = e^{-\epsilon|x|^2}$ and letting $g(x) = w(x)$ as defined above, substituting f and g into (2.9) and using Claim 7 to calculate the Fourier transform of f, we have

$$\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2}\widehat{w}(\xi)\,d\xi = \int_{\mathbb{R}^n} \frac{1}{(2\epsilon)^{n/2}}e^{-|x|^2/4\epsilon}w(x)\,dx. \tag{2.10}$$

Now we take the limit of both sides above as $\epsilon \to 0^+$. First,

$$\lim_{\epsilon\to 0^+}\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2}\widehat{w}(\xi)\,d\xi = \int_{\mathbb{R}^n}\widehat{w}(\xi)\,d\xi. \tag{2.11}$$

Second, we claim

$$\lim_{\epsilon\to 0^+}\frac{1}{(2\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}w(x)\,dx = (2\pi)^{n/2}w(0). \tag{2.12}$$

We prove this claim as follows. In particular, we will prove that

$$\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}w(x)\,dx \to w(0) \text{ as } \epsilon \to 0^+.$$

First, we note that

$$\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}\,dx = 1. \tag{2.13}$$

This follows directly from the fact that

$$\int_{-\infty}^{\infty} e^{-z^2}\,dz = \sqrt{\pi}.$$

Therefore,

$$\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}w(x)\,dx - w(0) = \frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx.$$

Now, we will show that for all $\gamma > 0$ there exists an $\widetilde{\epsilon} > 0$ such that

$$\left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| < \gamma$$

for $0 < \epsilon < \widetilde{\epsilon}$, thus, proving (2.12). Let $B(0,\delta)$ be the ball of radius δ about 0. (We will choose δ sufficiently small below.) Now break up the integral above into two pieces, as follows,

$$\begin{aligned}\left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| &\le \left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{B(0,\delta)} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| \\ &\quad + \left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n - B(0,\delta)} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| \\ &\equiv I + J.\end{aligned}$$

First, for term I, we have

$$\begin{aligned}\left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{B(0,\delta)} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| &\le |w(x)-w(0)|_{L^\infty(B(0,\delta))}\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon}\,dx \\ &< \frac{\gamma}{2},\end{aligned}$$

for δ sufficiently small, using the fact that $w \in C(\mathbb{R}^n)$ and (2.13). Now for δ fixed small, we look at term J,

$$\begin{aligned}&\left|\frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n - B(0,\delta)} e^{-|x|^2/4\epsilon}[w(x)-w(0)]\,dx\right| \\ &\qquad \le \frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n - B(0,\delta)} e^{-|x|^2/4\epsilon}|w(x)|\,dx \\ &\qquad\quad + \frac{1}{(4\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n - B(0,\delta)} e^{-|x|^2/4\epsilon}|w(0)|\,dx \\ &\qquad \le \left|\frac{1}{(4\pi\epsilon)^{n/2}}e^{-|x|^2/4\epsilon}\right|_{L^\infty(\mathbb{R}^n - B(0,\delta))}\int_{\mathbb{R}^n}|w(x)|\,dx \\ &\qquad\quad + e^{-\delta^2/8\epsilon}|w(0)|\frac{2^{n/2}}{(8\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n - B(0,\delta)} e^{-|x|^2/8\epsilon}\,dx \\ &\qquad \le C\left|\frac{1}{(4\pi\epsilon)^{n/2}}e^{-|\delta|^2/4\epsilon}\right| + Ce^{-\delta^2/8\epsilon}\frac{1}{(8\pi\epsilon)^{n/2}}\int_{\mathbb{R}^n} e^{-|x|^2/8\epsilon}\,dx \\ &\qquad \le C\left|\frac{1}{(4\pi\epsilon)^{n/2}}e^{-|\delta|^2/4\epsilon}\right| + Ce^{-\delta^2/8\epsilon} \\ &\qquad < \frac{\gamma}{2}\end{aligned}$$

for ϵ sufficiently small, using the fact that for a fixed $\delta \neq 0$,

$$\lim_{\epsilon\to 0^+}\frac{1}{(4\pi\epsilon)^{n/2}}e^{-|\delta|^2/4\epsilon} = 0 = \lim_{\epsilon\to 0^+} e^{-\delta^2/8\epsilon}.$$

Therefore, we conclude that

$$I + J < \gamma$$

for ϵ chosen sufficiently small, and, thus, (2.12) is proven.

Now combining (2.11) and (2.12) with (2.10), we conclude that

$$\int_{\mathbb{R}^n} \widehat{w}(\xi)\, d\xi = (2\pi)^{n/2} w(0).$$

Now $\widehat{w}(\xi) = (2\pi)^{n/2}|\widehat{u}|^2$ and $w(0) = u * v(0) = \int_{\mathbb{R}^n} u(x)v(0-x)\, dx = \int_{\mathbb{R}^n} u(x)\overline{u}(x)\, dx = \int_{\mathbb{R}^n} |u|^2\, dx$. Therefore, we conclude that

$$(2\pi)^{n/2} \int_{\mathbb{R}^n} |\widehat{u}|^2\, d\xi = (2\pi)^{n/2} \int_{\mathbb{R}^n} |u|^2\, dx,$$

or

$$||\widehat{u}||_{L^2} = ||u||_{L^2},$$

as claimed. A similar technique can be used to show that

$$||\check{u}||_{L^2} = ||u||_{L^2}.$$

□

Defining the Fourier Transform on $L^2(\mathbb{R}^n)$. For $f \in L^1(\mathbb{R}^n)$, that is, f such that

$$\int_{\mathbb{R}^n} |f(x)|\, dx < +\infty$$

it is clear that the Fourier transform is well-defined, i.e. - the integral converges. If $f \notin L^1(\mathbb{R}^n)$, the integral may not converge. Here we describe how we define the Fourier transform of a function $f \in L^2(\mathbb{R}^n)$ (but which may not be in $L^1(\mathbb{R}^n)$).

Let $f \in L^2(\mathbb{R}^n)$. Approximate f by a sequence of functions $\{f_k\}$ such that $f_k \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and

$$||f_k - f||_{L^2} \to 0 \text{ as } k \to +\infty.$$

By Plancherel's theorem,

$$||\widehat{f_k} - \widehat{f_j}||_{L^2} = ||\widehat{f_k - f_j}||_{L^2} = ||f_k - f_j||_{L^2} \to 0 \text{ as } k, j \to +\infty.$$

Therefore, $\{\widehat{f_k}\}$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$, and, therefore, converges to some $g \in L^2(\mathbb{R}^n)$. We *define* the Fourier transform of f by this function g. That is,

$$\widehat{f} \equiv g.$$

Other Properties of the Fourier Transform. Assume $u, v \in L^2(\mathbb{R}^n)$. Then

(a) $\displaystyle\int_{\mathbb{R}^n} u(x)\overline{v}(x)\, dx = \int_{\mathbb{R}^n} \widehat{u}(\xi)\overline{\widehat{v}}(\xi)\, d\xi.$

(b) $\widehat{\partial_{x_i}^{\alpha_i} u}(\xi) = (i\xi_i)^{\alpha_i}\widehat{u}(\xi).$

(c) $\widehat{u * v}(\xi) = (2\pi)^{n/2}\widehat{u}(\xi)\widehat{v}(\xi).$

(d) $u = \check{\widehat{u}}$

Proof of (a). Let $\alpha \in \mathbb{C}$. By Plancherel's Theorem, we have

$$||u + \alpha v||_{L^2} = ||\widehat{u} + \widehat{\alpha v}||_{L^2}$$
$$\implies \int_{\mathbb{R}^n} |u + \alpha v|^2 \, dx = \int_{\mathbb{R}^n} |\widehat{u} + \widehat{\alpha v}|^2 \, d\xi$$

Now using the fact that

$$|a + b|^2 = (a + b) \cdot (\overline{a + b}) = |a|^2 + \overline{a}b + a\overline{b} + |b|^2,$$

we have

$$\int_{\mathbb{R}^n} |u|^2 + \overline{u}(\alpha v) + u(\overline{\alpha v}) + |\alpha v|^2 \, dx = \int_{\mathbb{R}^n} |\widehat{u}|^2 + \overline{\widehat{u}}(\widehat{\alpha v}) + \widehat{u}(\overline{\widehat{\alpha v}}) + |\widehat{\alpha v}|^2 \, d\xi,$$

which implies

$$\int_{\mathbb{R}^n} \overline{u}(\alpha v) + u(\overline{\alpha v}) \, dx = \int_{\mathbb{R}^n} \overline{\widehat{u}}(\widehat{\alpha v}) + \widehat{u}(\overline{\widehat{\alpha v}}) \, d\xi.$$

Letting $\alpha = 1$, we have

$$\int_{\mathbb{R}^n} \overline{u}v + u\overline{v} \, dx = \int_{\mathbb{R}^n} \overline{\widehat{u}}\widehat{v} + \widehat{u}\overline{\widehat{v}} \, d\xi. \tag{2.14}$$

Letting $\alpha = i$, we have

$$\int_{\mathbb{R}^n} i\overline{u}v - iu\overline{v} \, dx = \int_{\mathbb{R}^n} i\overline{\widehat{u}}\widehat{v} - i\widehat{u}\overline{\widehat{v}} \, d\xi. \tag{2.15}$$

Multiplying (2.15) by i, we have

$$\int_{\mathbb{R}^n} -\overline{u}v + u\overline{v} \, dx = \int_{\mathbb{R}^n} -\overline{\widehat{u}}\widehat{v} + \widehat{u}\overline{\widehat{v}} \, d\xi. \tag{2.16}$$

Adding (2.14) and (2.16), we have

$$2\int_{\mathbb{R}^n} u\overline{v} \, dx = 2\int_{\mathbb{R}^n} \widehat{u}\overline{\widehat{v}} \, d\xi,$$

and, therefore, property (a) is proved. □

Proof of (b). We assume u is smooth and has compact support. We can use an approximation argument to prove the same equality for any u such that $\partial_{x_i}^{\alpha_i} u \in L^2(\mathbb{R}^n)$.

Using the definition of Fourier transform and integrating by parts, we have

$$\begin{aligned}
\widehat{\partial_{x_i} u}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} \partial_{x_i} u(x) \, dx \\
&= -\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} -i\xi_i e^{-ix\cdot\xi} u(x) \, dx \\
&= \frac{i\xi_i}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x) \, dx \\
&= i\xi_i \widehat{u}(\xi).
\end{aligned}$$

Property (b) follows by applying the same argument to higher derivatives.

□

Proof of (c). Property (c) was already proved in Claim 8 in the case when u and v are in $L^1(\mathbb{R}^n)$. One can use an approximation argument to prove the result in the general case.

□

Proof of (d). Fix $z \in \mathbb{R}^n$. Define the function $v_\epsilon(x) \equiv e^{ix\cdot z - \epsilon|x|^2}$. Therefore,

$$\begin{aligned}\widehat{v_\epsilon}(\xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} v_\epsilon(x)\,dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} e^{ix\cdot z} e^{-\epsilon|x|^2}\,dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot(\xi - z)} e^{-\epsilon|x|^2}\,dx \\ &= \widehat{f}(\xi - z)\end{aligned}$$

where

$$f(x) \equiv e^{-\epsilon|x|^2}.$$

By Claim 7, we know that

$$\widehat{f}(\xi) = \frac{1}{(2\epsilon)^{n/2}} e^{-|\xi|^2/4\epsilon}.$$

Therefore, we have

$$\widehat{v_\epsilon}(\xi) = \frac{1}{(2\epsilon)^{n/2}} e^{-|\xi - z|^2/4\epsilon}.$$

Now using (2.9), with $f = u$ and $g = v_\epsilon$, we have

$$\begin{aligned}\int_{\mathbb{R}^n} \widehat{u}(\xi) v_\epsilon(\xi)\,d\xi &= \int_{\mathbb{R}^n} u(x) \widehat{v_\epsilon}(x)\,dx \\ \Longrightarrow \int_{\mathbb{R}^n} \widehat{u}(\xi) e^{i\xi\cdot z - \epsilon|\xi|^2}\,d\xi &= \frac{1}{(2\epsilon)^{n/2}} \int_{\mathbb{R}^n} u(x) e^{-|x-z|^2/4\epsilon}\,dx\end{aligned}$$

Taking the limit as $\epsilon \to 0^+$, we have

$$\lim_{\epsilon\to 0^+} \int_{\mathbb{R}^n} \widehat{u}(\xi) e^{i\xi\cdot z - \epsilon|\xi|^2}\,d\xi = \int_{\mathbb{R}^n} \widehat{u}(\xi) e^{i\xi\cdot z}\,d\xi$$

and

$$\lim_{\epsilon\to 0^+} \frac{1}{(2\epsilon)^{n/2}} \int_{\mathbb{R}^n} u(x) e^{-|x-z|^2/4\epsilon}\,dx = (2\pi)^{n/2} u(z),$$

using (2.12). Therefore, we conclude that

$$(2\pi)^{n/2} u(z) = \int_{\mathbb{R}^n} e^{i\xi\cdot z} \widehat{u}(\xi)\,d\xi,$$

which implies

$$u(z) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i\xi\cdot z} \widehat{u}(\xi)\,d\xi = \check{\hat{u}}(z),$$

as desired.

□

2.4 Solving the Heat Equation in $\mathbb{R}^n$

2.4.1 Using the Fourier Transform to Solve the Initial-Value Problem for the Heat Equation in $\mathbb{R}^n$

Consider the initial-value problem for the heat equation on $\mathbb{R}$,

$$\begin{cases} u_t = ku_{xx} & x \in \mathbb{R}, t > 0 \\ u(x,0) = \phi(x). \end{cases}$$

Applying the Fourier transform to the heat equation, we have

$$\begin{aligned} &\widehat{u_t}(\xi,t) = k\widehat{u_{xx}}(\xi,t) \\ \implies &\widehat{u}_t = k(i\xi)^2\widehat{u} = -k\xi^2\widehat{u} \end{aligned}$$

Solving this ODE and using the initial condition $u(x,0) = \phi(x)$, we have

$$\widehat{u}(\xi,t) = \widehat{\phi}(\xi)e^{-k\xi^2 t}.$$

Therefore,

$$\begin{aligned} u(x,t) &= \frac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} e^{ix\cdot\xi}\widehat{u}(\xi,t)\,d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix\cdot\xi}\widehat{\phi}(\xi)e^{-k\xi^2 t}\,d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix\cdot\xi}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-iy\cdot\xi}\phi(y)\,dy\right] e^{-k\xi^2 t}\,d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(y)\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i(y-x)\cdot\xi}e^{-k\xi^2 t}\,d\xi\right] dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(y)\widehat{f}(y-x)\,dy \end{aligned}$$

where $f(\xi) = e^{-k\xi^2 t}$. By Claim 7, $f(\xi) = e^{-kt\xi^2}$ implies $\widehat{f}(z) = \frac{1}{\sqrt{2kt}}e^{-z^2/4kt}$. Therefore,

$$\widehat{f}(y-x) = \frac{1}{\sqrt{2kt}}e^{-(x-y)^2/4kt},$$

and consequently, the solution of the initial-value problem for the heat equation on $\mathbb{R}$ is given by

$$\boxed{u(x,t) = \frac{1}{\sqrt{4k\pi t}}\int_{-\infty}^{\infty} \phi(y)e^{-(x-y)^2/4kt}\,dy} \tag{2.17}$$

for $t > 0$.

We can use a similar analysis to solve the problem in higher dimensions. Consider the initial-value problem for the heat equation in $\mathbb{R}^n$,

$$\begin{cases} u_t = k\Delta u & x \in \mathbb{R}^n, t > 0 \\ u(x,0) = \phi(x). \end{cases} \tag{2.18}$$

Employing the Fourier transform as in the 1-D case, we arrive at the solution formula

$$\boxed{u(x,t) = \frac{1}{(4k\pi t)^{n/2}} \int_{\mathbb{R}^n} \phi(y) e^{-|x-y|^2/4kt}\, dy} \tag{2.19}$$

for $t > 0$.

Remark. Above, we have shown that *if* there exists a solution u of the heat equation (2.18), then u has the form (2.19). It remains to verify that this solution formula actually satisfies the initial-value problem. In particular, looking at the the formulas (2.17) and (2.19), we see that the functions given are not actually defined at $t = 0$. Therefore, to say that the solution formulas actually satisfy the initial-value problems, we mean to say that $\lim_{t \to 0^+} u(x,t) = \phi(x)$.

Theorem 9. (Ref: Evans, p. 47) *Assume $\phi \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, and define u by (2.19). Then*

1. $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$.
2. $u_t - k\Delta u = 0$ *for all* $x \in \mathbb{R}^n, t > 0$.
3. $\displaystyle\lim_{\substack{(x,t)\to(x_0,0) \\ x_0, x \in \mathbb{R}^n, t>0}} u(x,t) = \phi(x_0)$.

Proof. 1. Since the function $\frac{1}{t^{n/2}} e^{-\frac{|x|^2}{4kt}}$ is infinitely differentiable, with uniformly bounded derivatives of all orders on $\mathbb{R}^n \times [\delta, \infty)$ for all $\delta > 0$, we are justified in passing the derivatives inside the integral inside and see that $u \in C^\infty(\mathbb{R}^n \times (0,\infty))$.

2. By a straightforward calculation, we see that the function $\mathcal{H}(x,t) \equiv \frac{1}{(4k\pi t)^{n/2}} e^{-|x-y|^2/4kt}$ satisfies the heat equation for all $t > 0$. Again using the fact that this function is infinitely differentiable, we can justify passing the derivatives inside the integral and conclude that

$$u_t(x,t) - k\Delta u(x,t) = \int_{\mathbb{R}^n} [(\mathcal{H}_t - k\Delta_x \mathcal{H})(x-y,t)]\phi(y)\, dy = 0.$$

since $\mathcal{H}(x,t)$ solves the heat equation.

3. Fix a point $x_0 \in \mathbb{R}^n$ and $\epsilon > 0$. We need to show there exists a $\delta > 0$ such that

$$|u(x,t) - \phi(x_0)| < \epsilon$$

for $|(x,t) - (x_0,0)| < \delta$. That is, we need to show

$$\left| \frac{1}{(4\pi kt)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/4kt} \phi(y)\, dy - \phi(x_0) \right| < \epsilon$$

for $|(x,t)-(x_0,0)|<\delta$ where δ is chosen sufficiently small. The proof is similar to the proof of (2.12). In particular, using (2.13), we write

$$\begin{aligned}
&\left|\frac{1}{(4\pi kt)^{n/2}}\int_{\mathbb{R}^n} e^{-|x-y|^2/4kt}\phi(y)\,dy-\phi(x_0)\right| \\
&\qquad = \left|\frac{1}{(4\pi kt)^{n/2}}\int_{\mathbb{R}^n} e^{-|x-y|^2/4kt}[\phi(y)-\phi(x_0)]\,dy\right|.
\end{aligned}\tag{2.20}$$

Let $B(x_0,\gamma)$ be the ball of radius γ about x_0. We look at the integral in (2.20) over $B(x_0,\gamma)$. Using the fact that ϕ is continuous, we see that

$$\left|\frac{1}{(4\pi kt)^{n/2}}\int_{B(x_0,\gamma)} e^{-|x-y|^2/4kt}[\phi(y)-\phi(x_0)]\,dy\right| \le |\phi(y)-\phi(x_0)|_{L^\infty(B(x_0,\gamma))}<\frac{\epsilon}{2}$$

for γ chosen sufficiently small. For this choice of γ, we look at the integral in (2.20) over the complement of $B(x_0,\gamma)$. For $y\in\mathbb{R}^n-B(x_0,\gamma)$, $|x_0-x|<\frac{\gamma}{2}$, we have

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

Therefore, on this piece of the integral, we have $|y-x_0|<2|y-x|$. Therefore, this piece of the integral is bounded as follows,

$$\left|\frac{1}{(4\pi kt)^{n/2}}\int_{\mathbb{R}^n-B(x_0,\gamma)} e^{-|x-y|^2/4kt}[\phi(y)-\phi(x_0)]\,dy\right| \le \frac{C}{t^{n/2}}\int_{\mathbb{R}^n-B(x_0,\gamma)} e^{-|x_0-y|^2/16kt}\,dy.$$

We claim the right-hand side above can be made arbitrarily small by taking t arbitrarily small. We prove this in the case of $n=1$. The proof for higher dimensions is similar. By making a change of variables, we have

$$\begin{aligned}
\frac{C}{t^{1/2}}\int_{\mathbb{R}-B(x_0,\gamma)} e^{-(x_0-y)^2/16kt}\,dy &= \frac{C}{t^{1/2}}\int_\gamma^\infty e^{-\widetilde{y}^2/16kt}\,d\widetilde{y} \\
&= C\int_{\gamma/\sqrt{t}}^\infty e^{-z^2}\,dz \to 0 \quad \text{as } t\to 0^+.
\end{aligned}$$

In summary, take γ sufficiently small such that the piece of the integral in (2.20) over $B(x_0,\gamma)$ is bounded by $\epsilon/2$. Then, choose $\delta<\gamma/2$ sufficiently small such that

$$C\int_{\gamma\sqrt{t}}^\infty e^{-z^2}\,dz<\frac{\epsilon}{2}.$$

Part (3) of the theorem follows.

□

2.4.2 The Fundamental Solution

Consider again the solution formula (2.19) for the initial-value problem for the heat equation in $\mathbb{R}^n$. Define the function

$$\boxed{\mathcal{H}(x,t)\equiv\begin{cases}\dfrac{1}{(4\pi kt)^{n/2}}e^{-|x|^2/4kt} & t>0\\[2ex] 0 & t<0\end{cases}}\tag{2.21}$$

As can be verified directly, $\mathcal{H}$ is a solution of the heat equation for $t > 0$. In addition, we can write the solution of (2.18) as

$$u(x,t) = [\mathcal{H}(t) * \phi](x) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy.$$

As $\mathcal{H}$ has these nice properties, we call $\mathcal{H}(x,t)$ the **fundamental solution** of the heat equation.

Let's look at the fundamental solution a little more closely. As mentioned above, $\mathcal{H}$ itself a solution of the heat equation. That is,

$$\mathcal{H}_t - k\Delta\mathcal{H} = 0 \qquad x \in \mathbb{R}^n, t > 0.$$

What kind of initial conditions does $\mathcal{H}$ satisfy? We notice that for $x \neq 0$, $\lim_{t\to 0^+} \mathcal{H}(x,t) = 0$. However, for $x = 0$, $\lim_{t\to 0^+} \mathcal{H}(x,t) = \infty$. In addition, using (2.13), we see that

$$\int_{\mathbb{R}^n} \mathcal{H}(x,t)\,dx = 1$$

for $t > 0$. Therefore, $\lim_{t\to 0^+} \int_{\mathbb{R}^n} \mathcal{H}(x,t)\,dx = 1$. What kind of function satisfies these properties? Well, actually, *no* function satisfies these properties. Intuitively, the idea is that a "function" satisfying these properties represents a point mass located at the origin. This object is known as a *delta function*. We emphasize that the delta function is **not** a function! Instead, it is part of a group of objects called *distributions* which act on functions. We make these ideas more precise in the next section, and then will return to discussing the fundamental solution of the heat equation.

2.5 Distributions

Ref: Strauss, Section 12.1.

2.5.1 Definitions and Examples

We begin by making some definitions. We say a function $\phi : \mathbb{R}^n \to \mathbb{R}$ has **compact support** if $\phi \equiv 0$ outside a closed, bounded set in $\mathbb{R}^n$. We say ϕ is a **test function** if ϕ is an infinitely differentiable function with compact support. Let $\mathcal{D}$ denote the set of all test functions. We say $F : \mathcal{D} \to \mathbb{R}$ is a **distribution** if F is a continuous, linear functional which assigns a real number to every test function $\phi \in \mathcal{D}$. We let (F, ϕ) denote the real number associated with this distribution.

Example 10. Let $g : \mathbb{R} \to \mathbb{R}$ be any bounded function. We define the distribution associated with g as the map $F_g : \mathcal{D} \to \mathbb{R}$ which assigns to a test function ϕ the real number

$$(F_g, \phi) = \int_{-\infty}^{\infty} g(x)\phi(x)\,dx.$$

That is, $F_g : \phi \to \int g(x)\phi(x)\,dx$.

One particular example is the **Heaviside function**, defined as

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0. \end{cases}$$

Then, the distribution F_H associated with the Heaviside function is the map which assigns to a test function ϕ the real number

$$(F_H, \phi) = \int_0^\infty \phi(x)\, dx.$$

That is, $F_H : \phi \to \int_0^\infty \phi(x)\, dx$. $\diamond$

Example 11. The **Delta function** δ_0 (which is not actually a function!) is the distribution $\delta_0 : \mathcal{D} \to \mathbb{R}$ which assigns to a test function ϕ the real number $\phi(0)$. That is,

$$(\delta_0, \phi) = \phi(0).$$

Remarks.

(a) We sometimes write

$$\int_{\mathbb{R}^n} \delta_0(x)\phi(x)\, dx = \phi(0),$$

however, this is rather informal and not accurate because $\delta_0(x)$ is not a function! It should be thought of as purely notational.

(b) We can talk about the delta function centered at a point other than $x = 0$ as follows. For a fixed $y \in \mathbb{R}^n$, we define $\delta_y : \mathcal{D} \to \mathbb{R}$ to be the distribution which assigns to a test function ϕ the real number $\phi(y)$. That is,

$$(\delta_y, \phi) = \phi(y).$$

$\diamond$

2.5.2 Derivatives of Distributions

We now define derivatives of distributions. Let $F : \mathcal{D} \to \mathbb{R}$ be a distribution. We define the **derivative of the distribution** F as the distribution $G : \mathcal{D} \to \mathbb{R}$ such that

$$(G, \phi) = -(F, \phi').$$

for all $\phi \in \mathcal{D}$. We denote the derivative of F by F'. Then F' is the distribution such that

$$(F', \phi) = -(F, \phi')$$

for all $\phi \in \mathcal{D}$.

Example 12. For g a C^1 function, we define the distribution associated with g as $F_g : \mathcal{D} \to \mathbb{R}$ such that

$$(F_g, \phi) = \int_{-\infty}^{\infty} g(x)\phi(x)\,dx.$$

Therefore, integrating by parts, we have

$$\begin{aligned}(F_g, \phi') &= \int_{-\infty}^{\infty} g(x)\phi'(x)\,dx \\ &= -\int_{-\infty}^{\infty} g'(x)\phi(x)\,dx.\end{aligned}$$

By definition, the derivative of F_g, denoted F_g' is the distribution such that $(F_g', \phi) = -(F_g, \phi')$ for all $\phi \in \mathcal{D}$. Therefore, F_g' is the distribution associated with the function g'. That is,

$$(F_g', \phi) = -(F_g, \phi') = \int_{-\infty}^{\infty} g'(x)\phi(x)\,dx.$$

◇

Example 13. Let H be the Heaviside function defined above. Let $F_H : \mathcal{D} \to \mathbb{R}$ be the distribution associated with H, discussed above. Then the derivative of F_H, denoted F_H' must satisfy

$$(F_H', \phi) = -(F_H, \phi')$$

for all $\phi \in \mathcal{D}$. Now

$$\begin{aligned}(F_H, \phi') &= \int_{0}^{\infty} \phi'(x)\,dx \\ &= \lim_{b\to\infty} \int_0^b \phi'(x)\,dx \\ &= \lim_{b\to\infty} \phi(x)|_{x=0}^{x=b} \\ &= -\phi(0).\end{aligned}$$

Therefore,

$$(F_H', \phi) = -(F_H, \phi') = \phi(0).$$

That is, the derivative of the distribution associated with the Heaviside function is the delta function. We will be using this fact below. ◇

2.5.3 Convergence of Distributions

Let $F_n : \mathcal{D} \to \mathbb{R}$ be a sequence of distributions. We say F_n **converges weakly** to F if

$$(F_n, \phi) \to (F, \phi)$$

for all $\phi \in \mathcal{D}$.

2.5.4 The Fundamental Solution of the Heat Equation Revisited

In this section, we will show that the fundamental solution $\mathcal{H}$ of the heat equation (2.21) can be thought of as a solution of the following initial-value problem,

$$\begin{cases} \mathcal{H}_t - k\Delta\mathcal{H} = 0 & x \in \mathbb{R}^n, t > 0 \\ \mathcal{H}(x,0) = \delta_0. \end{cases} \tag{2.22}$$

Now, first of all, we must be careful in what we mean by saying that $\mathcal{H}(x,0) = \delta_0$. Clearly, the function $\mathcal{H}$ defined in (2.21) is not even defined at $t = 0$. And, now we're asking that it equals a distribution? We need to make this more precise. First of all, when we write $\mathcal{H}(x,0) = \delta_0$, we really mean "=" in the *sense of distributions*. In addition, we really mean to say that $\lim_{t\to 0^+} \mathcal{H}(x,t) = \delta_0$ in the sense of distributions. Let's state this more precisely now. Let $F_{\mathcal{H}(t)}$ be the distribution associated with $\mathcal{H}(t)$ defined by

$$(F_{\mathcal{H}(t)}, \phi) = \int_{\mathbb{R}^n} \mathcal{H}(x,t)\phi(x)\,dx.$$

Now to say that $\lim_{t\to 0^+} \mathcal{H}(x,t) = \delta_0$ in the sense of distributions, we mean that

$$\lim_{t\to 0^+} (F_{\mathcal{H}(t)}, \phi) = (\delta_0, \phi) = \phi(0). \tag{2.23}$$

Therefore, to summarize, by saying that $\mathcal{H}$ is a solution of the "initial-value problem" (2.22), we really mean that $\mathcal{H}$ is a solution of the heat equation for $t > 0$ and (2.23) holds.

We now prove that in fact, our fundamental solution defined in (2.21) is a solution of (2.22) in this sense. Showing that $\mathcal{H}$ satisfies the heat equation for $t > 0$ is a straightforward calculation. Therefore, we only focus on showing that the initial condition is satisfied. In particular, we need to prove that (2.23) holds. We proceed as follows.

$$\begin{aligned} \lim_{t\to 0^+} (F_{\mathcal{H}(t)}, \phi) &= \lim_{t\to 0^+} \int_{\mathbb{R}^n} \mathcal{H}(x,t)\phi(x)\,dx \\ &= \lim_{t\to 0^+} \frac{1}{(4\pi kt)^{n/2}} \int_{\mathbb{R}^n} e^{-|x|^2/4kt}\phi(x)\,dx. \end{aligned}$$

But, now in (2.12), we proved that this limit is exactly $\phi(0)$! Therefore, we have proven (2.23). Consequently, we can think of our fundamental solution as a solution of (2.22).

The beauty of this formulation is the following. If $\mathcal{H}$ is a solution of (2.22), then define

$$u(x,t) \equiv [\mathcal{H}(t) * \phi](x) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy.$$

The idea is that u *should* be a solution of (2.18). We give the *formal* argument below.

$$u(x,0) = [\mathcal{H}(0) * \phi](x) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,0)\phi(y)\,dy = \phi(x).$$

In addition,

$$\begin{aligned} u_t - k\Delta_x u &= \int_{\mathbb{R}^n} \mathcal{H}_t(x-y,t)\phi(y)\,dy - \int_{\mathbb{R}^n} k\Delta_x\mathcal{H}(x-y,t)\phi(y)\,dy \\ &= \int_{\mathbb{R}^n} [\mathcal{H}_t(x-y,t) - k\Delta_x\mathcal{H}(x-y,t)]\phi(y)\,dy = 0. \end{aligned}$$

Note: These calculations are formal in the sense that we are ignoring convergence issues, etc. To verify that this formulation actually gives you a solution, you need to deal with these issues. Of course, for "nice" initial data, we have shown that this formulation does give us a solution to the heat equation.

2.6 Properties of the Heat Equation

2.6.1 Invariance Properties

Consider the equation

$$u_t = ku_{xx}, \qquad -\infty < x < \infty.$$

The equation satisfies the following invariance properties,

(a) The *translate* $u(x-y,t)$ of any solution $u(x,t)$ is another solution for any fixed y.

(b) Any *derivative* (u_x, u_t, u_{xx}, etc.) of a solution is again a solution.

(c) A *linear combination* of solutions is again a solution.

(d) An *integral* of a solution is again a solution (assuming proper convergence.)

(e) If $u(x,t)$ is a solution, so is the *dilated* function $u(\sqrt{a}x, at)$ for any $a > 0$. This can be proved by the chain rule. Let $v(x,t) = u(\sqrt{a}x, at)$. Then

$$v_t = au_t$$

and

$$v_x = \sqrt{a}u_x.$$

And, therefore,

$$v_{xx} = \sqrt{a} \cdot \sqrt{a}u_{xx} = au_{xx}.$$

2.6.2 Properties of Solutions

1. *Smoothness of Solutions.* As can be seen from the above theorem, solutions of the heat equation are infinitely differentiable. Therefore, even if there are singularities in the initial data, they are instantly "smoothed out" and the solution $u(x,t) \in C^\infty(\mathbb{R}^n \times (0,\infty))$.

2. *Domain of Dependence.* The value of the solution at the point x, time t depends on the value of the initial data on the whole real line. In other words, there is an infinite domain of dependence for solutions to the heat equation. This is in contrast to hyperbolic equations where solutions are known to have finite domains of dependence.

2.7 Inhomogeneous Heat Equation

Ref: Strauss, Sec. 3.3; Evans, Sec. 2.3.1.c.

In this section, we consider the initial-value problem for the inhomogeneous heat equation on some domain Ω (not necessarily bounded) in $\mathbb{R}^n$,

$$\begin{cases} u_t - k\Delta u = f(x,t) & x \in \Omega, t > 0 \\ u(x,0) = \phi(x). \end{cases} \tag{2.24}$$

We claim that we can use solutions of the homogeneous equation to construct solutions of the inhomogeneous equation.

2.7.1 Motivation

Consider the following ODE:

$$\begin{cases} u_t + au = f(t) \\ u(0) = \phi, \end{cases}$$

where a is a constant. By multiplying by the integrating e^{at}, we see the solution is given by

$$u(t) = e^{-at}\phi + \int_0^t e^{-a(t-s)} f(s)\, ds.$$

In other words, the solution $u(t)$ is the *propagator* e^{-at} applied to the initial data, plus the propagator "convolved" with the nonlinear term. In other words, if we let $S(t)$ be the operator which multiplies functions by e^{-at}, we see that the solution of the homogeneous problem,

$$\begin{cases} u_t + au = 0 \\ u(0) = \phi \end{cases}$$

is given by $S(t)\phi = e^{-at}\phi$. Further, the solution of the inhomogeneous problem is given by

$$u(t) = S(t)\phi + \int_0^t S(t-s) f(s)\, ds.$$

We claim that this same technique will allow us to find a solution of the inhomogeneous heat equation. Being able to construct solutions of the inhomogeneous problem from solutions of the homogeneous problem is known as **Duhamel's principle**. We show below how this idea works.

Suppose we can solve the homogeneous problem,

$$\begin{cases} u_t - k\Delta u = 0 & x \in \Omega \\ u(x,0) = \phi(x). \end{cases} \tag{2.25}$$

That is, assume the solution of (2.25) is given by $u_h(x,t) = S(t)\phi(x)$ for some solution operator $S(t)$. We claim that the solution of the inhomogeneous problem (2.24) is given by

$$u(x,t) = S(t)\phi(x) + \int_0^t S(t-s) f(x,s)\, ds.$$

At least formally, we see that

$$\begin{aligned} u_t - k\Delta u &= [\partial_t - k\Delta]S(t)\phi(x) + [\partial_t - k\Delta]\int_0^t S(t-s)f(x,s)\,ds \\ &= 0 + S(t-t)f(x,t) + \int_0^t [\partial_t - k\Delta]S(t-s)f(x,s)\,ds \\ &= S(0)f(x,t) = f(x,t). \end{aligned}$$

Below, we show how we use this idea to construct solutions of the heat equation on $\mathbb{R}^n$ and on bounded domains $\Omega \subset \mathbb{R}^n$.

2.7.2 Inhomogeneous Heat Equation in $\mathbb{R}^n$

Consider the inhomogeneous problem,

$$\begin{cases} u_t - k\Delta u = f(x,t) & x \in \mathbb{R}^n, t > 0 \\ u(x,0) = \phi(x). \end{cases} \tag{2.26}$$

From earlier, we know the solution of the corresponding homogeneous initial-value problem

$$\begin{cases} u_t - k\Delta u = 0 & x \in \mathbb{R}^n, t > 0 \\ u(x,0) = \phi(x) \end{cases} \tag{2.27}$$

is given by

$$u_h(x,t) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy.$$

That is, we can think of the solution operator $S(t)$ associated with the heat equation on $\mathbb{R}^n$ as defined by

$$S(t)\phi(x) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy.$$

Therefore, we *expect* the solution of the inhomogeneous heat equation to be given by

$$\begin{aligned} u(x,t) &= S(t)\phi(x) + \int_0^t S(t-s)f(x,s)\,ds \\ &= \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy + \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(x-y,t-s)f(y,s)\,dy\,ds. \end{aligned}$$

Now, we have already shown that $u_h(x,t) = \int_{\mathbb{R}^n} \mathcal{H}(x-y,t)\phi(y)\,dy$ satisfies (2.27), (at least for "nice" functions ϕ). Therefore, it suffices to show that

$$u_p(x,t) \equiv \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(x-y,t-s)f(y,s)\,dy\,ds \tag{2.28}$$

satisfies (2.26) with zero initial data. If we can prove this, then $u(x,t) = u_h(x,t) + u_p(x,t)$ will clearly solve (2.26) with initial data ϕ. In the following theorem, we prove that u_p satisfies the inhomogeneous heat equation with zero initial data.

Theorem 14. *Assume $f \in C_1^2(\mathbb{R}^n \times [0,\infty))$ (meaning f is twice continuously differentiable in the spatial variables and once continuously differentiable in the time variable) and has compact support. Define u as in* (2.28)*. Then*

1. $u \in C_1^2(\mathbb{R}^n \times (0,\infty))$.
2. $u_t(x,t) - k\Delta u(x,t) = f(x,t)$ *for all* $x \in \mathbb{R}^n, t > 0$.
3. $\lim_{\substack{(x,t)\to(x_0,0)\\ x\in\mathbb{R}^n, t>0}} u(x,t) = 0$ *for all* $x_0 \in \mathbb{R}^n$.

Proof. 1. Since $\mathcal{H}$ has a singularity at $(0,0)$, we cannot justify passing the derivatives inside the integral. Instead, we make a change of variables as follows. In particular, letting $\widetilde{y} = x - y$ and $\widetilde{s} = t - s$, we have

$$\int_0^t \int_{\mathbb{R}^n} \mathcal{H}(x-y, t-s) f(y,s)\, dy\, ds = \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(\widetilde{y}, \widetilde{s}) f(x-\widetilde{y}, t-\widetilde{s})\, d\widetilde{y}\, d\widetilde{s}.$$

For ease of notation, we drop the $\widetilde{\ }$ notation. Now by assumption $f \in C_1^2(\mathbb{R}^n \times [0,\infty))$ and $\mathcal{H}(y,s)$ is smooth near $s = t > 0$. Therefore, we have

$$\begin{aligned}\partial_t \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) f(x-y,t-s)\, dy\, ds = &\int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) \partial_t f(x-y,t-s)\, dy\, ds \\ &+ \int_{\mathbb{R}^n} \mathcal{H}(y,t) f(x-y,0)\, dy\end{aligned}$$

and

$$\partial_{x_i x_j} \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) f(x-y,t-s)\, dy\, ds = \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) \partial_{x_i x_j} f(x-y,t-s)\, dy\, ds.$$

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

2. Now we need to calculate $u_t - k\Delta u$. Using the same change of variables as above, we have

$$\begin{aligned}&[\partial_t - k\Delta_x] \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) f(x-y,t-s)\, dy\, ds \\ &\quad= \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s)[\partial_t - k\Delta_x] f(x-y,t-s)\, dy\, ds + \int_{\mathbb{R}^n} \mathcal{H}(y,t) f(x-y,0)\, dy \\ &\quad= \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s)[-\partial_s - k\Delta_y] f(x-y,t-s)\, dy\, ds + \int_{\mathbb{R}^n} \mathcal{H}(y,t) f(x-y,0)\, dy.\end{aligned}$$

Now, we would like to integrate by parts to put the derivatives on $\mathcal{H}$ as we know $\mathcal{H}$ is a solution of the heat equation. However, we know $\mathcal{H}$ has a singularity at $s = 0$.

To avoid this, we break up the integral into the intervals $[0,\epsilon]$ and $[\epsilon,t]$. Therefore, we write

$$\begin{aligned}
[\partial_t - k\Delta_x]u &= \int_\epsilon^t \int_{\mathbb{R}^n} \mathcal{H}(y,s)[-\partial_s - k\Delta_y]f(x-y,t-s)\,dy\,ds \\
&\quad + \int_0^\epsilon \int_{\mathbb{R}^n} \mathcal{H}(y,s)[-\partial_s - k\Delta_y]f(x-y,t-s)\,dy\,ds \\
&\quad + \int_{\mathbb{R}^n} \mathcal{H}(y,t)f(x-y,0)\,dy \\
&\equiv I_\epsilon + J_\epsilon + K.
\end{aligned}$$

First, for J_ϵ, we have

$$\begin{aligned}
&\left|\int_0^\epsilon \int_{\mathbb{R}^n} \mathcal{H}(y,s)[-\partial_s - k\Delta y]f(x-y,t-s)\,dy\,ds\right| \\
&\qquad \le (||f_t||_{L^\infty} + k||\Delta f||_{L^\infty}) \int_0^\epsilon \int_{\mathbb{R}^n} \mathcal{H}(y,s)\,dy\,ds \\
&\qquad \le \epsilon C \int_{\mathbb{R}^n} \mathcal{H}(y,t)\,dy \le C\epsilon,
\end{aligned}$$

using (2.13).

For I_ϵ, using the assumption that f has compact support, we integrate by parts as follows,

$$\begin{aligned}
&\int_\epsilon^t \int_{\mathbb{R}^n} \mathcal{H}(y,s)[-\partial_s - k\Delta_y]f(x-y,t-s)\,dy\,ds \\
&\qquad = \int_\epsilon^t \int_{\mathbb{R}^n} [\partial_s - k\Delta_y]\mathcal{H}(y,s)f(x-y,t-s)\,dy\,ds \\
&\qquad\quad - \int_{\mathbb{R}^n} \mathcal{H}(y,s)f(x-y,t-s)\,dy\Big|_{s=\epsilon}^{s=t} \\
&\qquad = 0 + \int_{\mathbb{R}^n} \mathcal{H}(y,\epsilon)f(x-y,t-\epsilon)\,dy - \int_{\mathbb{R}^n} \mathcal{H}(y,t)f(x-y,0)\,dy \\
&\qquad = \int_{\mathbb{R}^n} \mathcal{H}(y,\epsilon)f(x-y,t-\epsilon)\,dy - K.
\end{aligned}$$

Therefore,

$$I_\epsilon + K = \int_{\mathbb{R}^n} \mathcal{H}(y,\epsilon)f(x-y,t-\epsilon)\,dy.$$

Now,

$$\begin{aligned}
u_t - k\Delta u &= \lim_{\epsilon\to 0^+} [I_\epsilon + J_\epsilon + K] \\
&= \lim_{\epsilon\to 0^+} \int_{\mathbb{R}^n} \mathcal{H}(y,\epsilon)f(x-y,t-\epsilon)\,dy \\
&= \lim_{\epsilon\to 0^+} \frac{1}{(4\pi k\epsilon)^{n/2}} \int_{\mathbb{R}^n} e^{-|y|^2/4k\epsilon} f(x-y,t-\epsilon)\,dy \\
&= f(x,t),
\end{aligned}$$

using the same technique we used to prove part (3) of Theorem 9.

3. To prove that the limit as $t \to 0^+$ of our solution $u(x,t)$ is 0, we use the fact that

$$
\begin{aligned}
|u(x,t)|_{L^\infty(\mathbb{R}^n)} &= \left| \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s) f(x-y, t-s)\, dy\, ds \right| \\
&\leq |f|_{L^\infty(\mathbb{R}^n \times [0,t])} \int_0^t \int_{\mathbb{R}^n} \mathcal{H}(y,s)\, dy\, ds \\
&\leq C \int_0^t \frac{1}{(4\pi ks)^{n/2}} \int_{\mathbb{R}^n} e^{-|y|^2/4ks}\, dy\, ds \\
&= Ct,
\end{aligned}
$$

using (2.13). Therefore, as $t \to 0^+$, $u(x,t) \to 0$, as claimed.

□

2.7.3 Inhomogeneous Heat Equation on Bounded Domains

In this section, we consider the initial/boundary value problem for the inhomogeneous heat equation on an interval $\mathcal{I} \subset \mathbb{R}^n$,

$$
\begin{cases}
u_t - ku_{xx} = f(x,t) & x \in \mathcal{I} \subset \mathbb{R}, t > 0 \\
u(x,0) = \phi(x) & x \in \mathcal{I} \\
u \text{ satisfies certain BCs} & t > 0.
\end{cases}
$$

Using Duhamel's principle, we *expect* the solution to be given by

$$
u(x,t) = S(t)\phi(x) + \int_0^t S(t-s) f(x,s)\, ds
$$

where $u_h(x,t) = S(t)\phi(x)$ is the solution of the homogeneous equation and $S(t)$ is the solution operator associated with the homogeneous problem.

As shown earlier, the solution of the homogeneous problem with symmetric boundary conditions,

$$
\begin{cases}
u_t - ku_{xx} = 0 & x \in \mathcal{I}, t > 0 \\
u(x,0) = \phi(x) & x \in \mathcal{I} \\
u \text{ satisfies symmetric BCs} & t > 0
\end{cases}
$$

is given by

$$
u(x,t) = \sum_{n=1}^{\infty} A_n X_n(x) e^{-k\lambda_n t}
$$

where X_n are the eigenfunctions and λ_n the corresponding eigenvalues of the eigenvalue problem,

$$
\begin{cases}
-X'' = \lambda X & x \in \mathcal{I} \\
X \text{ satisfies symmetric BCs,}
\end{cases}
$$

and the coefficients A_n are defined by

$$
A_n = \frac{\langle X_n, \phi \rangle}{\langle X_n, X_n \rangle},
$$

where the inner product is taken over $\mathcal{I}$. Therefore, the solution operator associated with the homogeneous equation is given by

$$S(t)\phi = \sum_{n=1}^{\infty} A_n X_n(x) e^{-k\lambda_n t}$$

where

$$A_n = \frac{\langle X_n, \phi \rangle}{\langle X_n, X_n, \rangle}.$$

Therefore, we expect the solution of the inhomogeneous equation to be given by

$$\begin{aligned} u(x,t) &= S(t)\phi(x) + \int_0^t S(t-s) f(x,s)\, ds \\ &= \sum_{n=1}^{\infty} A_n X_n(x) e^{-\lambda_n t} + \int_0^t \sum_{n=1}^{\infty} B_n(s) X_n(x) e^{-\lambda_n (t-s)}\, ds \end{aligned}$$

where

$$B_n(s) \equiv \frac{\langle X_n, f(s) \rangle}{\langle X_n, X_n \rangle}.$$

In fact, for "nice" functions ϕ and f, this formula gives us a solution of the inhomogeneous initial/boundary-value problem for the heat equation. We omit proof of this fact here. We consider an example.

Example 15. Solve the inhomogeneous initial/boundary value problem for the heat equation on $[0, l]$ with Dirichlet boundary conditions,

$$\begin{cases} u_t - k u_{xx} = f(x,t) & x \in [0,l], t > 0 \\ u(x,0) = \phi(x) & x \in [0,l] \\ u(0,t) = 0 = u(l,t) & t > 0 \end{cases} \tag{2.29}$$

The solution of the homogeneous problem with initial data ϕ,

$$\begin{cases} u_t - k u_{xx} = 0 & x \in [0,l], t > 0 \\ u(x,0) = \phi(x) & x \in [0,l] \\ u(0,t) = 0 = u(l,t) & t > 0 \end{cases}$$

is given by

$$u_h(x,t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l} x\right) e^{-k(n\pi/l)^2 t},$$

where

$$A_n = \frac{2}{l} \int_0^l \sin\left(\frac{n\pi}{l} x\right) \phi(x)\, dx.$$

Therefore, the solution of the inhomogeneous problem with zero initial data is given by

$$u_p(x,t) = \int_0^t \sum_{n=1}^{\infty} B_n(s) \sin\left(\frac{n\pi}{l} x\right) e^{-k(n\pi/l)^2 (t-s)}\, ds$$

where

$$B_n(s) = \frac{2}{l} \int_0^l \sin\left(\frac{n\pi}{l}x\right) f(x,s)\,dx.$$

Consequently, the solution of (2.29) is given by

$$\boxed{u(x,t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{l}x\right) e^{-k(n\pi/l)^2 t} + \int_0^t \sum_{n=1}^{\infty} B_n(s) \sin\left(\frac{n\pi}{l}x\right) e^{-k(n\pi/l)^2(t-s)}\,ds}$$

where A_n and $B_n(s)$ are as defined above.

◇

2.7.4 Inhomogeneous Boundary Data

In this section we consider the case of the heat equation on an interval with inhomogeneous boundary data. We will use the method of shifting the data to reduce this problem to an inhomogeneous equation with homogeneous boundary data. Consider the following example.

Example 16. Consider

$$\begin{cases} u_t - ku_{xx} = 0 & 0 < x < l \\ u(x,0) = \phi(x) & 0 < x < l \\ u(0,t) = g(t) & \\ u(l,t) = h(t). & \end{cases} \tag{2.30}$$

We introduce a new function $\mathcal{U}(x,t)$ such that

$$\mathcal{U}(x,t) = \frac{1}{l}\left[(l-x)g(t) + xh(t)\right].$$

Assume $u(x,t)$ is a solution of (2.30). Then let

$$v(x,t) \equiv u(x,t) - \mathcal{U}(x,t).$$

Therefore,

$$\begin{aligned} v_t - kv_{xx} &= (u_t - ku_{xx}) - (\mathcal{U}_t - k\mathcal{U}_{xx}) \\ &= -\mathcal{U}_t \\ &= -\frac{1}{l}\left[(l-x)g'(t) + xh'(t)\right]. \end{aligned}$$

Further,

$$\begin{aligned} v(x,0) &= u(x,0) - \mathcal{U}(x,0) = \phi(x) - \frac{1}{l}\left[(l-x)g(0) + xh(0)\right] \\ v(0,t) &= u(0,t) - \mathcal{U}(0,t) = 0 \\ v(l,t) &= u(l,t) - \mathcal{U}(l,t) = 0. \end{aligned}$$

Therefore, v is a solution of the inhomogeneous heat equation on the interval $[0,l]$ with homogeneous boundary data,

$$\begin{cases} v_t - kv_{xx} = -\frac{1}{l}\left[(l-x)g'(t) + xh'(t)\right] & 0 < x < l \\ v(x,0) = \phi(x) - \frac{1}{l}\left[(l-x)g(0) + xh(0)\right] & 0 < x < l \\ v(0,t) = 0 = v(l,t). \end{cases}$$

We can solve this problem for v using the technique of the previous section. Then we can solve our original problem (2.30) using the fact that $u(x,t) = v(x,t) + \mathcal{U}(x,t)$. $\diamond$

2.8 Maximum Principle and Uniqueness of Solutions

In this section, we prove what is known as the *maximum principle* for the heat equation. We will then use this principle to prove uniqueness of solutions to the initial-value problem for the heat equation.

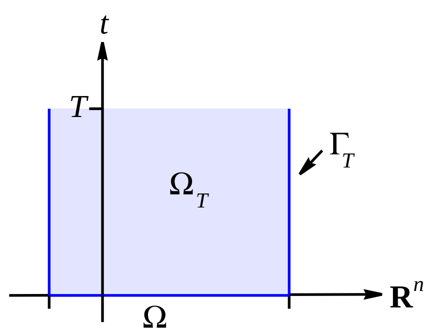
2.8.1 Maximum Principle for the Heat Equation

First, we prove the maximum principle for solutions of the heat equation on bounded domains. Let $\Omega \subset \mathbb{R}^n$ be an open, bounded set. We define the **parabolic cylinder** as

$$\Omega_T \equiv \Omega \times (0,T].$$

We define the **parabolic boundary** of Γ_T as

$$\Gamma_T \equiv \overline{\Omega}_T - \Omega_T.$$



We now state the maximum principle for solutions to the heat equation.

Theorem 17. (Maximum Principle on Bounded Domains) *(Ref: Evans, p. 54.) Let Ω be an open, bounded set in $\mathbb{R}^n$. Let Ω_T and Γ_T be as defined above. Assume u is sufficiently smooth, (specifically, assume $u \in C_1^2(\Omega_T) \cap C(\overline{\Omega}_T)$) and u solves the heat equation in Ω_T. Then,*

1. $\max_{\overline{\Omega}_T} u(x,t) = \max_{\Gamma_T} u(x,t)$. (Weak maximum principle)

2. If Ω is connected and there exists a point $(x_0, t_0) \in \Omega_T$ such that

$$u(x_0, t_0) = \max_{\overline{\Omega}_T} u(x, t),$$

then u is constant in $\overline{\Omega}_{t_0}$. (Strong maximum principle)

Proof. (*Ref: Strauss, p. 42 (weak); Evans, Sec. 2.3.3 (strong)*)

Here, we will only prove the weak maximum principle. The reader is referred to Evans for a proof of the strong maximum principle. We will prove the weak maximum principle in the case $n = 1$ and $\Omega = (0, l)$.

Let

$$M = \max_{\Gamma_T} u(x, t).$$

We must show that $u(x, t) \leq M$ throughout Ω_T. Let $\epsilon > 0$ and define

$$v(x, t) = u(x, t) + \epsilon x^2.$$

We claim that $v(x, t) \leq M + \epsilon l^2$ throughout Ω_T. Assuming this, we can conclude that $u(x, t) \leq M + \epsilon(l^2 - x^2)$ on Ω_T and consequently, $u(x, t) \leq M$ on Ω_T.

Therefore, we only need to show that $v(x, t) \leq M + \epsilon l^2$ throughout Ω_T. By definition, we know $v(x, t) \leq M + \epsilon l^2$ on Γ_T (that is, on the lines $t = 0$, $x = 0$ and $x = l$). We will prove that v cannot attain its maximum in Ω_T, and, therefore, $v(x, t) \leq M + \epsilon l^2$ throughout Ω_T. We prove this as follows.

First, by definition of v, we see v satisfies

$$v_t - k v_{xx} = u_t - k(u + \epsilon x^2)_{xx} = u_t - k u_{xx} - 2\epsilon k = -2\epsilon k < 0. \tag{2.31}$$

Inequality (2.31) is known as the *diffusion inequality*. We will use (2.31) to prove that v cannot achieve its maximum in Ω_T.

Suppose v attains its maximum at a point $(x_0, t_0) \in \Omega_T$ such that $0 < x_0 < l$, $0 < t_0 < T$. This would imply that $v_t(x_0, t_0) = 0$ and $v_{xx}(x_0, t_0) \leq 0$, and, consequently that

$$v_t(x_0, t_0) - k v_{xx}(x_0, t_0) \geq 0,$$

which contradicts the diffusion inequality (2.31). Therefore, v cannot attain its maximum at such a point.

Suppose v attains its maximum at a point (x_0, T) on the top edge of Ω_T. Then $v_x(x_0, T) = 0$ and $v_{xx}(x_0, T) \leq 0$, as before. Furthermore,

$$v_t(x_0, T) = \lim_{\delta \to 0^+} \frac{v(x_0, T) - v(x_0, T - \delta)}{\delta} \geq 0.$$

Again, this contradicts the diffusion inequality. But v must have a maximum somewhere in the closed rectangle $\overline{\Omega_T}$. Therefore, this maximum must occur somewhere on the parabolic boundary Γ_T. Therefore, $v(x, t) \leq M + \epsilon l^2$ throughout Ω_T, and, we can conclude that $u(x, t) \leq M$ throughout Ω_T, as desired. □

We now prove a maximum principle for solutions to the heat equation on all of $\mathbb{R}^n$. Without having a boundary condition, we need to impose some growth assumptions on the behavior of the solutions as $|x| \to +\infty$.

Theorem 18. (Maximum Principle on $\mathbb{R}^n$) *(Ref: Evans, p. 57) Suppose u (sufficiently smooth) solves*

$$\begin{cases} u_t - k\Delta u = 0 & x \in \mathbb{R}^n \times (0,T) \\ u = \phi & x \in \mathbb{R}^n \end{cases}$$

and u satisfies the growth estimate

$$u(x,t) \leq Ae^{a|x|^2} \qquad x \in \mathbb{R}^n, 0 \leq t \leq T$$

for constants $A, a > 0$. Then

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

Proof. (For simplicity, we take $k = 1$, but the proof works for arbitrary $k > 0$.) Assume

$$4aT < 1.$$

Therefore, there exists $\epsilon > 0$ such that

$$4a(T + \epsilon) < 1.$$

Fix $y \in \mathbb{R}^n$ and $\mu > 0$. Let

$$v(x,t) \equiv u(x,t) - \frac{\mu}{(T+\epsilon-t)^{n/2}} e^{\frac{|x-y|^2}{4(T+\epsilon-t)}}.$$

Using the fact that u is a solution of the heat equation, it is straightforward to show that v is a solution of the heat equation,

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

Fix $r > 0$ and let $U \equiv B(y,r), U_T = B(y,r) \times (0,T]$. From the maximum principle, we know

$$\max_{\overline{U}_T} v(x,t) = \max_{\Gamma_T} v(x,t).$$

We will now show that $\max_{\Gamma_T} v(x,t) \leq \phi(x)$. First, look at (x,t) on the base of Γ_T. That is, take $(x,t) = (x,0)$. Then

$$\begin{aligned} v(x,0) &= u(x,0) - \frac{\mu}{(T+\epsilon)^{n/2}} e^{\frac{|x-y|^2}{4(T+\epsilon)}} \\ &\leq u(x,0) = \phi(x). \end{aligned}$$

Now take (x,t) on the sides of Γ_T. That is, choose $x \in \mathbb{R}^n$ such that $|x-y| = r$ and t such that $0 \leq t \leq T$. Then

$$\begin{aligned} v(x,t) &= u(x,t) - \frac{\mu}{(T+\epsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\epsilon-t)}} \\ &\leq Ae^{a|x|^2} - \frac{\mu}{(T+\epsilon-t)^{n/2}} e^{\frac{r^2}{4(T+\epsilon-t)}} \\ &\leq Ae^{a(|y|+r)^2} - \frac{\mu}{(T+\epsilon)^{n/2}} e^{\frac{r^2}{4(T+\epsilon)}}. \end{aligned}$$

By assumption, $4a(T+\epsilon) < 1$. Therefore, $1/4(T+\epsilon) = a + \gamma$ for some $\gamma > 0$. Therefore,

$$v(x,t) \leq Ae^{a(|y|+r)^2} - \mu(4(a+\gamma))^{n/2} e^{(a+\gamma)r^2} \leq \sup_{\mathbb{R}^n} \phi(x)$$

for r sufficiently large.

Therefore, we have shown that $v(x,t) \leq \sup_{\mathbb{R}^n} \phi(x)$ for $(x,t) \in \Gamma_T$, and, thus, by the maximum principle

$$\max_{\overline{U}_T} v(x,t) = \max_{\Gamma_T} v(x,t),$$

we have $v(x,t) \leq \sup_{\mathbb{R}^n} \phi(x)$ for all $(x,t) \in \overline{U}_T$.

We can use this same argument for all $y \in \mathbb{R}^n$, to conclude that

$$v(y,t) \leq \sup_{\mathbb{R}^n} \phi(x)$$

for all $y \in \mathbb{R}^n$, $0 \leq t \leq T$ as long as $4aT < 1$. Then using the definition of v and taking the limit as $\mu \to 0+$, we conclude that

$$u(y,t) \leq \sup_{\mathbb{R}^n} \phi(x)$$

for all $y \in \mathbb{R}^n$, $0 \leq t \leq T$ as long as $4aT < 1$.

If $4aT \geq 1$, we divide the interval $[0,T]$ into subintervals $[0,T_1]$, $[T_1, 2T_1]$, etc., where $T_1 = \frac{1}{8a}$ and perform the same calculations on each of these subintervals. $\square$

2.8.2 Using the Maximum Principle to Prove Uniqueness

We now use the maximum principles proven above to prove uniqueness of solutions to the heat equation.

Theorem 19. (Uniqueness on bounded domains) *Let Ω be an open, bounded set in $\mathbb{R}^n$ and define Ω_T and Γ_T as above. Consider the initial/boundary value problem,*

$$\begin{cases} u_t - k\Delta u = f & x \in \Omega_T \\ u = g & x \in \Gamma_T \end{cases} \tag{2.32}$$

Assume f and g are continuous. Then there exists at most one (smooth) solution of (2.32).

Proof. Suppose there exist two smooth solutions u and v. Let

$$w \equiv u - v.$$

Then w is a solution of

$$\begin{cases} w_t - k\Delta w = 0 & x \in \Omega_T \\ w = 0 & x \in \Gamma_T. \end{cases} \tag{2.33}$$

By the weak maximum principle,

$$\max_{\overline{\Omega}_T} w(x,t) = \max_{\Gamma_T} w(x,t) = 0.$$

Therefore, $w(x,t) \leq 0$ on Ω_T, which implies $u(x,t) \leq v(x,t)$ on Ω_T.

Next, let $\widetilde{w} = v - u$. Therefore, $\widetilde{w}$ is also a solution of (2.33), and, by the weak maximum principle,

$$\max_{\overline{\Omega}_T} \widetilde{w}(x,t) = \max_{\Gamma_T} \widetilde{w}(x,t) = 0.$$

Consequently, $\widetilde{w}(x,t) \leq 0$ on Ω_T, which implies $v(x,t) \leq u(x,t)$ on Ω_T.

Therefore, we have

$$\begin{aligned} u(x,t) &\leq v(x,t) \quad (x,t) \in \overline{\Omega_T} \\ v(x,t) &\leq u(x,t) \quad (x,t) \in \overline{\Omega_T}. \end{aligned}$$

Consequently, we conclude that $u = v$ for in $\overline{\Omega_T}$, as desired. $\square$

Theorem 20. (Uniqueness on $\mathbb{R}^n$) *(Ref: Evans, p. 58.) Consider the initial-value problem*

$$\begin{cases} u_t - k\Delta u = f & x \in \mathbb{R}^n \times (0,T) \\ u(x,0) = g & x \in \mathbb{R}^n. \end{cases} \tag{2.34}$$

Assume f and g are continuous. Then there exists at most one (smooth) solution of (2.34) *satisfying the growth estimate*

$$|u(x,t)| \leq Ae^{a|x|^2}.$$

Proof. Assume u, v are two solutions. Let $w \equiv u - v$ and let $\widetilde{w} = v - u$. Now apply the maximum principle on $\mathbb{R}^n$ to show that $u \equiv v$. $\square$

Remark. There are, in fact, infinitely many solutions of

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

(Ref: John, Chapter 7.) All of those solutions other than $u(x,t) \equiv 0$ grow faster than $e^{|x|^2}$ as $|x| \to +\infty$.

2.8.3 Energy Methods to prove Uniqueness

We now present an alternate technique to prove uniqueness of solutions to the heat equation on bounded domains.

Theorem 21. *Let Ω be an open, bounded set in $\mathbb{R}^n$. Let $T > 0$. Let Ω_T, Γ_T be the parabolic cylinder and parabolic boundary defined earlier. Consider the initial/boundary value problem,*

$$\begin{cases} u_t - k\Delta u = f & (x,t) \in \Omega_T \\ u(x,t) = g & (x,t) \in \Gamma_T. \end{cases} \tag{2.35}$$

There exists at most one (smooth) solution u of (2.35).

Proof. Suppose there exist two solutions u and v. Let $w = u - v$. Then w solves

$$\begin{cases} w_t - k\Delta w = 0 & (x,t) \in \Omega_T \\ w = 0 & (x,t) \in \Gamma_T. \end{cases} \tag{2.36}$$

Let

$$E(t) \equiv \int_\Omega w^2(x,t)\, dx.$$

Now, $E(0) = \int_\Omega w^2(x,0)\, dx = 0$. We claim that $E'(t) \leq 0$ and, therefore, $E(t) = 0$ for $0 \leq t \leq T$. By using the fact that w is a solution of (2.36) and integrating by parts, we see that

$$\begin{aligned} E'(t) &= 2\int_\Omega w w_t\, dx \\ &= 2k\int_\Omega w\Delta w\, dx \\ &= -2k\int_\Omega |\nabla w|^2\, dx + 2k\int_{\partial\Omega} w\frac{\partial w}{\partial \nu}\, dS(x) \\ &= -2k\int_\Omega |\nabla w|^2\, dx \leq 0, \end{aligned}$$

using the fact that $w = 0$ on Γ_T and $k > 0$. Therefore, $E(t) = 0$ for $0 \leq t \leq T$. Using the assumption that w is smooth, this implies $w \equiv 0$ in $\overline{\Omega}_T$, and, therefore, $u \equiv v$ in $\overline{\Omega}_T$.

□